UNIVERSITA' DEGLI STUDI DI MILANO

Facoltà di Medicina e Chirurgia

Dipartimento di Scienze Cliniche e di Comunità

Scuola di Dottorato in **Scienze biomediche cliniche e sperimentali**

Dottorato in **STATISTICA BIOMEDICA – XXV CICLO**

FAMILY HISTORY OF CANCER AND FAMILY HISTORY SCORES FOR ASSESSING THE LEVEL OF DISEASE RISK IN FAMILIES

Dott. **Federica Turati**

Matricola R08568

Tutor: Ch.mo Prof. **Adriano Decarli**

Coordinatore del Dottorato: Ch.mo Prof. **Adriano Decarli**

Anno accademico 2011-2012

Chapter 1 – Family history of cancer

Chapter 2 – Family history scores for assessing the level of disease risk in families

SUMMARY

In the last decades genetic factors are playing an increasingly important role in medical research, given the evidence for the existence of a heritable susceptibility for various diseases, including common cancers, based on reports of families with multiple affected relatives. Epidemiologists have utilized family history, usually of first-degree relatives, as a surrogate for genetic risk, aware that family history reflects the consequences of genetic susceptibilities, shared environment, and common behaviors.

During my PhD I have dealt with two different aspects of family history, i.e., the role of family history of cancer in epidemiological cancer research (Chapter 1) and the use of complex family history score for assessing the level of disease risk in families (Chapter 2).

In particular, I have systematically examined the extent to which a family history of cancer might be a risk factor for cancer within the same cancer site and across multiple cancer sites, analyzing a large and comprehensive dataset based on a network of integrated case-control studies, conducted in Italy and Switzerland since the early 90's. The database included 1468 cases of cancer of the oral cavity and pharynx, 198 of the rhinopharynx, 505 of the esophagus, 230 of the stomach, 2390 of the colorectum, 185 of the liver, 326 of the pancreas, 852 of the larynx, 3034 of the breast, 367 of the endometrium, 1031 of the ovary, 1294 of the prostate, 767 of the renal cell, and a total of 16022 corresponding controls. Unconditional multiple logistic regression models, adjusted for the major possible confounding factors, and a procedure for controlling for multiplicity using a false discovery rate were used. The risk of developing cancer at a particular site was increased, although not always significantly, in subjects with a first-degree relative affected by cancer at the same site, with odds ratios ranging from 1.4 for

pancreatic cancer, to 7.4 for ovarian cancer. Several across sites associations emerged, some of which possibly due to shared environmental exposures or lifestyle practices among family members (e.g., alcohol, smoking, unhealthy diet, infections) or to the inheritance of one or more predisposing gene mutations (high penetrance gene mutations, such as BRCA1/2 in breast and ovarian cancer, and/or low penetrance polymorphisms, as those involved in carcinogens metabolism, such as GST genes in oral cancer) or to a combination of both. The analysis I performed confirmed that several associations were stronger for a younger age at diagnosis in relatives. A detailed discussion of the findings is reported in paragraph 4 of Chapter 1.

In addition to the investigation of the role of family history of cancer in cancer etiology, I have performed a statistical evaluation of the performance of different family history scores to recommend the measure that performs best. Family history scores summarize familial information and are used for estimating the familiar risk, i.e. the level of risk for a particular disease among members of that family. The simplest and most common family history scores are the dichotomous measure indicator, positive in families that have at least one relative with the disease, the number of affected family members, and the proportion of affected relatives, which takes into account the size of the family. The other family history scores proposed in the literature are statistics that describe the deviation of the observed situation from the expected risk for each family. More detailed information on family members (affected and unaffected) as well as incidence rates of the diseases of interest in strata of selected covariates are needed to compute these more complex family history scores. To evaluate family history scores' performance I used two different complementary approaches: a data-derived approach, using data from the Italian HI-WATE study, with the aim of examining the power of

various family history scores in predicting a particular diseases (i.e., colorectal cancer), and a simulation approach to evaluate their accuracy of predicting the true familial risk. From 200 simulations for 48 different settings, Reed's score and FHS2 seem to perform slightly better than the other scores. However, the simple proportion of affected relatives is not so far in terms of predictivity of the true familial risk. The use of this simple score seems therefore justified, at least until stronger evidence is brought for the advantages of using a more complex score.

Chapter 1

Family history of cancer

1.1 Introduction

Overall, an estimated 12.7 million new cancer cases and 7.6 million cancer deaths occurred in 2008 worldwide [1]. By 2020, these figures are estimated to rise to over 16 million new cases, with 10 million deaths. There may be more than 20 million new cases of cancer in 2030 [2].

The role of the environment on carcinogenesis has been actively investigated in many site-specific cancers, and a number of different types of exogenous factors are known causes of cancer. These include some aspects of food and nutrition, tobacco smoking, infectious agents, medication, radiation, and industrial chemicals, and also carcinogenic agents in food and drink [3].

In the last decades genetic factors are playing an increasingly important role in cancer research, given the evidence for the existence of an heritable susceptibility for various common cancers [4], based on reports of families with multiple affected relatives.

Epidemiologists have utilized family history (FH), usually of first-degree relatives (FDR), as a surrogate for genetic risk, aware that FH reflects the consequences of genetic susceptibilities, shared environment, and common behaviors [4, 5].

For many common cancers, including those of the digestive tract, and smoking- and alcohol-related cancers, the risk is significantly increased among subjects with a FDR affected by cancer at the same site [6-15]. However, the magnitude of the associations with FH have varied considerably between studies, with estimated relative risks (RR) ranging from about 2 to 5 for most cancers.

A number of dominant susceptibility genes cause cancer at several sites. These include clustering of adenocarcinomas of the colorectum and endometrium in hereditary non-polyposis colorectal cancer (HNPCC) families [16], the association of soft tissue sarcomas, leukaemia, brain and breast tumours in the Li–Fraumeni syndrome [17] and clustering in BRCA1 and BRCA2 families of cancers of the breast and ovary, with smaller risks for some other sites including colon, prostate and pancreas [18, 19]. It is then justified to ask whether there is a general susceptibility to cancer, causing clustering of discordant or different cancers in families. Some studies have addressed the question. A systematic study of familial correlations between 28 different types of cancer using the Utah population database revealed a number of significant associations [20]. Environmental and behavioral risk factors which are shared within families contribute to some associations, such as the familial clustering for cancers of smoking-related sites. Clustering of cancers of the female genitalia, lip and oral cavity may reflect a common viral aetiology. Some associations are similar to those caused by rare penetrant genes. BRCA1 and BRCA2 presumably contribute to the associations between breast, colon and prostate and between ovary and pancreas, and the increased

risk of soft tissue cancers among relatives of breast cancer probands maybe due in part to the Li–Fraumeni syndrome. Truncating mutations in these genes are probably too rare to account for all of these associations, suggesting the existence of less penetrant variants, or perhaps other genes that affect some of the same pathways. Evidence of the HNPCC complex of cancers was seen in relatives of young uterine cancer probands. In addition to these expected associations between cancer sites, various significant associations such as thyroid cancer and non-Hodgkin's lymphoma with breast cancer and leukaemias with colorectal cancer suggest previously unrecognized hereditary effects. A case-control study nested within a large cohort, the American Cancer Society Cancer Prevention Study-1 [21], found that the associations between FH and cancer mortality were generally stronger within cancer sites than across cancer sites. Within site associations were found for breast cancer, colorectal cancer, stomach cancer, and lung cancer. Across-site associations were observed for a FH of breast cancer as a risk factor for ovarian cancer mortality, stomach cancer as a risk factor for ovarian cancer mortality, and uterine cancer as a risk factor for pancreatic cancer mortality. In the Family-Cancer Database from the nationwide Swedish registries, familial aggregation between parents and offspring was observed for *5* concordant and 14 discordant cancer sites and 10 parental sites at which all cancer was increased in the offspring [22]. The concordant sites between the parent and offspring were colorectum, breast, melanoma, skin (squamous cell carcinoma), and thyroid. The aggregation at discordant sites in the parents and the offspring included stomach-breast, colorectum-sailvary glands, colorectum-breast, colorectum-lymphoma, colorectum-leukemia, liver-breast, pancreas-breast, breast-melanoma, ovary-breast, prostate-breast, prostate-cervix, prostate-multiple myeloma, kidney-melanoma, and nervous tissue-melanoma. More recently, a

study based on the Icelandic population database [23] analyzed familial aggregation of cancer cases both within and between pairs of cancer sites. It found 17 cancer sites involved in 20 significant pairs of sites. The estimated RRs for the 20 pairs are between 1.1 and 1.7 for FDR relatives and between 1.1 and 1.5 for second-degree relatives. The highest RRs in FDR between cancer sites were seen for esophagus–cervix, with a RR of 1.74, pancreas–ovary, with a RR of 1.66, and colon–rectum, with a RR of 1.64. Other cancer pairs that demonstrated significant familial co-clustering were: melanoma-kidney, cervix-lung, stomach-brain, colon-stomach, prostate-colon, prostate-breast.

Moreover, recent interest in the clustering of cancers across sites has stemmed from genome-wide association studies, which have identified more than 100 genetic variants spanning at least 20 cancers [24, 25].

The genetic contribution to diseases of complex origin, such as cancer, often is most salient to families of patients with early onset. Scanty information, however, is available on the variation of risk for FH in relation age at diagnosis of the affected relative. During my three-year PhD experience I have had the opportunity to handle a large amount of data from a series of well-conducted observational studies on cancer, systematically collecting information on cancer in FDR. This allowed me to explore the role of FH of cancer in cancer etiology. In particular, I have systematically examined the extent to which a FH of cancer might be a risk factor for cancer within the same cancer site and across multiple cancer sites, analyzing a large and comprehensive dataset based on a network of integrated case-control studies, conducted in Italy and Switzerland since the early 90's, and I have also investigated the possible modifying effect of age at diagnosis and sex of the affected relative [8, 26, 27]. Moreover, I focused on liver cancer and I quantitatively combined in a systematic meta-analysis all

published data on FH of liver cancer and liver cancer risk from observational studies [8].

1.2 Methods

1.2.1 The network of case-control studies

Recruitment. Between 1991 and 2009, a series of hospital-based case-control studies on several neoplasms were carried out in various areas of northern (the greater Milan area; the provinces of Pordenone, Padua, Udine, Gorizia and Forlì; the urban area of Genoa), central (the provinces of Rome and Latina), and southern (the urban area of Naples) Italy, and in the Canton of Vaud, Switzerland. The studies included a total of 1468 cases of cancer of the oral cavity and pharynx (OP), 198 of the rhinopharynx (RP), 505 of the esophagus (i.e., squamous cell carcinoma of the esophagus, SCCE), 230 of the stomach, 2390 of the colorectum (CR), 185 of the liver, 326 of the pancreas, 852 of the larynx, 3034 of the breast, 367 of the endometrium, 1031 of the ovary, 1294 of the prostate, 767 of the renal cell, and a total of 16022 corresponding controls (Table 1).

All studies included incident cases, identified in the major teaching and general hospitals of the study areas. Controls were subjects admitted to the same network of hospitals as cases for a wide spectrum of acute, non-neoplastic conditions unrelated to known or potential risk factors for the corresponding cancer site. Overall, 8.5% of controls were admitted for traumatic conditions, 24.2% for non-traumatic orthopedic conditions, 29.4% for acute surgical conditions, and 37.9% for miscellaneous other illnesses.

The proportion of refusals of subjects approached was less than 5% in Italian centers, and about 15% in Switzerland. The study protocol was revised and approved by local ethics committees of the hospitals involved according to the regulations at the time of each study conduction, and all participants gave informed consent.

Data collection. Subjects were face-to-face interviewed using similar structured questionnaires, administered by trained interviewers. The questionnaire included information on socio-demographic characteristics, anthropometric measures, lifestyle habits (e.g., tobacco smoking, alcohol drinking), dietary habits, a personal medical history, and, for women, menstrual and reproductive factors, and use of oral contraceptives (OC) and hormone replacement therapy (HRT). Subjects were specifically asked for how many sisters and brothers they had, and whether their parents, siblings, children, grandparents or spouse had ever had any cancer (excluding nonmelanoma skin cancer). For each relative with a history of cancer, the subject was asked to report the vital status at the time of interview, current age or the age at death, cancer site and age at diagnosis. The history of cancer in first-degree only, i.e., parents, siblings and sons/daughters, was considered in the analysis. On account of recall and classification difficulties, some sites were combined (i.e., all CRC, all Hodgkin and non-Hodgkin lymphoma, myeloma and leukemia, as well as cervix and corpus uteri).
In some of the studies in the network of case-control studies a blood sample was collected from each subject.

Statistical analysis. Odds ratios (OR) of 13 different cancers according to FH of selected cancers in FDR and the corresponding 95% confidence intervals (CI) were estimated by unconditional multiple logistic regression models [28]. The models included terms for quinquennia of age, sex (when appropriate), study centre (when appropriate), year of interview, education, alcohol drinking, tobacco smoking, body mass index (BMI), number of brothers and sisters. For female cancers, models included further terms for parity, menopausal status, age at menopause, and OC and HRT use,

and, for breast cancer only, age at first birth. Additional models were used to assess the potential modifying effect of sex and age at diagnosis of the affected FDR. In further analyses, I accounted for the problem of multiple comparisons using a false discovery rate controlling procedure, according to the Benjamini-Hochberg method [29].

Additional stratified analyses, as well as analyses of interaction between FH and selected lifestyle exposure were also performed for liver and laryngeal cancer.

All statistical analyses were performed with SAS 9.1 statistical software (SAS Institute, Cary, NC).

1.2.2 Meta-analysis: a focus on liver cancer

I performed a Medline search in PubMed up to April 2011 using the string "(liver OR hepatocellular) AND (cancer OR neoplasm OR tumor OR carcinoma) AND ("family history"))", limiting the search to the publications written in English language and following the Meta-analysis Of Observational Studies in Epidemiology (MOOSE) guidelines [30]. The PubMed search identified 232 articles. From these, I selected 14 reports giving information on the association between FH of liver cancer and liver cancer risk. Studies considering FH of any cancer or FH of liver diseases were not considered. Some reports were excluded as based on data later updated [31-33], or because reporting data from case-control studies nested in previously identified cohorts [34, 35]. Three additional papers were identified through the review of the reference lists of the publications retrieved [36-38]. Finally, besides our present case-control study, the meta-analysis included other 8 case-control [36-43] and 4 cohort studies [44-47].

Whenever available, I considered multivariate risk estimates, adjusted for the largest number of potential confounding factors; otherwise I computed the crude OR (and the corresponding 95% CI) from the distribution of cases and controls according to FH of liver cancer. In a study reporting the adjusted OR but not the corresponding 95% CI [38], I used the standard error of the crude OR. I pooled the RR estimates of liver cancer for FH of liver cancer from each study using random-effects models, which consider both within- and between-study variation [48]. Heterogeneity among studies was assessed using the $\chi 2$ test (results were defined as heterogeneous for a p-value<0.10) [49].

1.3 Results

1.3.1 Results from the network of case-control studies

Table 2 gives the distribution of cases of cancer at 13 different sites and the corresponding controls according to a positive FH of selected cancers. The corresponding ORs and 95% CIs are reported in Table 3 (in bold, significant ORs at the 0.05 level after accounting for the problem of multiple comparisons). In general, the risk of developing cancer at a particular site was increased, although not always significantly, in subjects with a FDR affected by cancer at the same site, with ORs ranging from 1.4 for pancreatic cancer, to 7.4 for ovarian cancer.

With regard to the across sites associations, without controlling for the problem of multiple comparisons, significant associations or associations of borderline significance emerged between cancer at OP and FH of laryngeal (OR=3.3, significant [s]), skin (OR=3.3, not significant [ns]), or breast cancer (OR=1.5, ns); cancer at RP and FH of colorectal (OR=3.1, s) or skin (OR=4.6, ns) cancer; SCCE and FH of OP (OR=4.1, s) or stomach (OR=1.8, ns) cancer; cancer at CR and FH of stomach (OR=1.2, ns), liver (OR=1.4, ns), skin (OR=2.2, ns), ovarian (OR=2.1, ns), prostate (OR=1.6, ns) cancer, or of all hemolymphopoietic (HLP, OR=1.4, ns) cancers; cancer at the pancreas and FH of stomach (OR=2.4, s) or bone (OR=2.7, ns) cancer; cancer at the larynx and FH of colorectal (OR=1.5, ns), or skin (OR=8.4, s) cancer.

For female cancers, the risk of breast cancer was increased among those with FH of stomach (OR=1.2, ns), colorectal (OR=1.5, s), skin (OR=3.0, s), uterine (OR=1.4, ns), prostate (OR=1.6, s), or of all HLP (OR=1.7, s) cancers; the risk of endometrial cancer among those with FH of oral and pharyngeal cancer (OR=2.4, ns), stomach (OR=1.8,

ns), kidney (OR=4.2, ns), or brain (OR=4.2, s) cancer; the risk of ovarian cancer among those with FH of colorectal (OR=1.6, s), laryngeal (OR=1.8, ns), breast (OR=2.3, s), or of all HLP cancers (OR=1.6, ns).

Prostate cancer risk was elevated in subjects with at least one FRD affected by colorectal (OR=1.5, ns), lung (OR=1.5, s), ovarian (OR=7.4, s), bladder (OR=3.4, s), or kidney (OR=3.4, s) cancer; the risk of renal cell cancer among those with a FDR with laryngeal (OR=2.2, ns), uterine (OR=1.7, ns), ovarian (OR=4.1, ns), or prostate (OR=1.7, ns) cancer.

Individuals with a FH of any type of cancer are generally at a higher risk of developing all the cancers considered. When FH of the cancer under investigation was not considered, the magnitude of these associations decreased.

Significance at the 0.05 level after Benjamini-Hochberg correction was found for the following associations: cancer at OP and FH the corresponding cancer (corrected-pvalue=0.004) and laryngeal cancer (corrected-pvalue=0.010); SCCE and FH of oral and pharyngeal cancer (corrected-pvalue=0.034); cancer at stomach and FH of the corresponding cancer (corrected-pvalue=0.012); cancer at CR and FH of cancer at the same site (corrected-pvalue<0.001); cancer at liver and FH of liver cancer (corrected-pvalue=0.062, borderline); cancer at the larynx and FH of the corresponding cancer (corrected-pvalue= 0.028); cancer at the breast and FH of the corresponding cancer (corrected-pvalue<0.001), colorectal (corrected-pvalue=0.049), skin (corrected-pvalue=0.062), or all HLP cancers (corrected-pvalue=0.036); cancer at the ovary with FH of cancer at the same site (corrected-pvalue<0.001) or at breast (corrected-pvalue<0.001); cancer at prostate and FH of prostatic (corrected-pvalue<0.001) or

bladder (corrected-pvalue=0.034) cancer; cancer at the renal cell and FH of kidney cancer (corrected-pvalue=0.028).

Sex and age at diagnosis of the affected relative

Table 4 presents results for FH according to sex and age at cancer diagnosis of the affected FDR. For each cancer site investigated in this report, we showed RR estimates only for FH of cancer at 1) the same site, 2) sites for which a significant increased risk emerged in the previous analysis, 3) any sites, and 4) any sites excluding the one under investigation.

Some of the associations resulted restricted to (or stronger for) having a male rather than a female affected FDR. This is the case of the associations between oral and pharyngeal cancer and FH of cancer at the same site (OR for a male affected=3.1 [s] - OR for a female affected=0.8), rhinopharyngeal cancer and FH of colorectal cancer (OR=4.0 [s], vs. 0.5), colorectal cancer and FH of colorectal cancer (OR=3.5 [s] vs. 2.2 [s]), liver cancer and FH of cancer at the same site (OR=3.6 [s] vs 1.1 [ns]), pancreatic cancer and FH of stomach cancer (OR=3.5 [s] vs 1.4 [ns]), and breast cancer and FH of skin cancer (OR=10.7 [s] vs 1.0). The associations between prostate cancer and FH of bladder cancer (OR for a female affected=11.3 [s], for a male affected=2.5 [borderline]) and between renal cell cancer and FH of kidney cancer (OR=6.2 [s] vs 2.9 [ns]) were stronger for having a female rather than a male affected relative.

Concerning age at cancer diagnosis in FDR, several associations were stronger for earlier diagnosed FDR. This is particularly evident for SCCE and FH of cancer at OP, stomach cancer and laryngeal cancer and FH of cancer at the same sites, and breast cancer and FH of cancer at the skin. Opposite findings emerged for a few associations.

Additional analyses

Laryngeal cancer. Table 5 gives the ORs for FH of laryngeal cancer in strata of age, tobacco and alcohol consumption. The ORs were somewhat higher in younger probands, current smokers and in heavy drinkers, although the tests of heterogeneity were not significant. The combined effect of tobacco, alcohol and FH of laryngeal cancer is shown in Figure 1. As compared to the lowest risk category, i.e., non-smokers, drinkers of less than 28 drinks per week, without FH, the risk was increased in those with one or more factors in the highest risk category: the OR was 1.4 (95% CI, 0.3–6.2) for nonsmokers and moderate drinkers with FH of laryngeal cancer, 18.2 (95% CI, 13.7–24.2) for current smokers and heavy drinkers without FH and 37.1 (95% CI, 9.9–139.4) for smokers and heavy drinkers who also reported a FDR with laryngeal cancer.

Liver cancer. After further adjustment for hepatitis B virus (HCV) and/or hepatitis C virus chronic infection, the OR of primary liver cancer for FH of liver cancer was 2.38 (95% CI, 1.01-5.58) (data not shown). Figure 2 shows the interaction between FH of liver cancer, and HBV surface antigen (HBsAg) and/or antibodies against HCV (anti-HCV) positivity. Compared to the lowest risk category (i.e. subjects without FH and with no chronic B/C hepatitis), the ORs were 2.94 (95% CI, 0.94-9.21) for subjects not chronically infected by hepatitis viruses and with FH, 38.19 (95% CI, 21.97-66.39) for those with chronic infection with hepatitis viruses and no FH, and 72.48 (95% CI, 21.92-239.73) for those exposed to both risk factors. No significant interaction emerged between these two factors (p=0.61).

1.3.2 Results from the meta-analysis

Besides the case-control study on liver cancer of our research group, from the literature search I identified 8 other case–control and 4 cohort studies, for a total of 3627 liver cancer cases. The main study characteristics are reported in Table 6. Most studies were conducted in South-eastern Asia, where the prevalence of HCV/HBV infection is high; only 3 studies, besides the present one, were carried out in Western countries. Figure 3 shows the forest plot for the association between FH of liver cancer and HCC risk. The pooled RRs for liver cancer were 2.80 (95% CI, 2.19-3.58) for case-control, 2.28 (95% CI, 1.58-3.29) for cohort, and 2.50 (95% CI, 2.06-3.03) for all studies, with no heterogeneity between study type (p=0.17). When studies not allowing for hepatitis were excluded from the analyses, the overall RR was 2.28 (95% CI, 1.85-2.18) (RR estimate based on studies either including a term for hepatitis in the logistic regression model or performed on subjects with hepatitis infection only). Analyses by sex showed a pooled RR of 2.80 (95% CI, 2.14-3.66, p for heterogeneity=0.21) for males from 9 studies (including the present one) [36-39, 41, 44-46], and of 1.55 (95% CI, 0.92-2.64, p for heterogeneity=0.77) for females from 6 studies (including the present one). When only the 6 studies (including the present one) reporting risk estimates for both males and females separately were considered, the pooled RR for males became 2.39 (95% CI, 2.03-2.81) while the one for females remained the same.

1.4 Discussion

During my PhD, I have provided a comprehensive picture on the role of FH of cancer at different sites on the risk of several cancers, using original data from a network of Italian and Switzerland case-control studies, with the same inclusion criteria and questionnaire. The associations between FH of cancer at a particular site and the risk of developing the same disease were substantially confirmed. An original aspect of this study is the systematic investigation of all cancer sites in FDR, and consequently the possibility of obtaining quantitative estimates of risks of 13 selected cancers with reference to familial aggregation of cancer at other sites. For most of the cancers investigated in this report, a somewhat increased risk according to FH of cancer at other sites emerged. Some of these associations may be explained by shared environmental factors among members of the same family, some of others by inherited genetic susceptibilities, some of others by a combination of both. Concerning the possible modifying effect of age at diagnosis, this analysis confirmed that several associations were stronger for a younger age at diagnosis in relatives.

In the following, a brief discussion of the major findings for each cancer site investigated.

Oral and pharyngeal cancer. Our findings of an elevated risk of oral and pharyngeal cancer in subjects with FH of oral and pharyngeal or laryngeal cancers are in broad agreement with other reports [20, 50, 51]. An inherited component of susceptibility to oral and pharyngeal cancer has been suggested by case reports of families with multiple affected members [52-54], by epidemiologic studies indicating familial tendency to oral

and pharyngeal cancer or other cancers of upper aerodigestive tract [55-62], by segregation analysis in FDR [63], by elevated risks associated with genes involved in DNA repair maintenance of genetic stability (e.g., P53) [64, 65], and by elevated risks associated with polymorphic genes involved in the metabolism of tobacco, alcohol and other carcinogens (e.g, ADH, NAT2, GSTM1, GSTT1, CYP) [66-70], also in consideration that carcinogen-matabolizing enzimes are expressed in the oral cavity, suggesting that matabolism of carcinogens could occur at this site.

Familial aggregation of oral and pharyngeal cancers may also be due to shared environmental exposure to the main risk factors, i.e., alcohol and tobacco. This is supported by the stronger association for having an affected male than female DFR, since smoking and alcohol drinking are more common among the former, and by the observed elevated risk associated with FH of laryngeal cancer. However, the risk of oral and pharyngeal cancer was only moderately increased in individuals with a FH of esophageal cancer and not related to a FH of cancer at other sites related to alcohol or tobacco (i.e., liver or lung) [71, 72], suggesting that tobacco and alcohol cannot totally explain this association. The association with FH of skin cancer has not to my knowledge been previously reported in the literature, and needs therefore independent confirmation. Although the point estimate of the OR was around 3.3, this estimate was based on 8 cases and 7 controls only; moreover, significance was lost after controlling for multiple comparisons.

Rhinopharyngeal cancer. A modest, not significant, 20% increased risk of rhinopharyngeal cancer emerged for subjects with FH oral and pharyngeal cancer. A FH of rhinopharyngeal cancer have been reported to significantly increase the risk of

developing the disease in other epidemiological studies on the issue, with ORs ranging from 4 to 20 for individuals with an affected FDR as compared with those with no affected relatives [73-85]. On the contrary Abdulamir et al. [86] found that a positive FH of head and neck cancer do not increase rhinopharyngeal cancer risk. Almost all these studies, however, were conducted in high-risk populations, including those from southern China and south-eastern Asia, which may be substantially different from this Southern European population in terms of tumor histology, and lifestyle behaviors. Moreover, I have considered in the analysis FH of oral and pharyngeal cancers combined, which included rhinopharyngeal cancer, since no study subjects reported a FDR affected by that specific tumor. Inherited cancer syndromes are not presumed to account for a high proportion of cases. Rather, common genetic variation, particularly involved in genes that play a role in the immune response to Epstein-Barr virus infection (i.e., HLA) or the metabolism of environmental carcinogens, including tobacco (e.g., CYP, GSTM, GSTT) are likely to be an important factor in rhinopharyngeal cancer development [87].

The positive association emerged with FH of colorectal cancer should be interpreted with caution, in the absence of confirmatory findings from other studies [83, 85]. Moreover, the corresponding OR lost its significance after multiple comparisons controlling.

SCCE. An elevated but not significant risk of SCCE emerged for subjects reporting at least one FDR affected by esophageal cancer. Several epidemiological studies from high-risk esophageal cancer areas in China and the Caspian Littoral of Iran found that esophageal/gastric cancer aggregates in families, and that a FH of esophageal/gastric

cancer is a risk factor for the disease [88-95]. Results have been less consistent in studies from other areas of the world. No association between FH of esophageal cancer and risk of SCCE was found in three case-control studies from the U.S. [96, 97] and Sweden [98], including about 200 cases each. Conversely, a case-control study from Japan on 167 cases with squamous cell carcinoma of the hypopharynx or cervical esophagus found an OR of 5.1 (95% CI: 0.7-36.1) in subjects with a FH of esophageal cancer, and an OR of 2.6 (95% CI, 1.1-6.3) for a FH of cancers of the esophagus, head and neck or lung [99]. The Swedish Family-Cancer Database, which includes 10.1 million individuals and about 6,000 cases with esophageal cancers, found a standardized incidence ratio of 4.9 (95% CI, 1.8-9.6) for having a parent with SCCE and of 12.6 (95% CI, 1.2-36.2) for having a SCCE in a sibling [100].

This data showed that a FH of oral/pharyngeal and stomach cancer increases the risk of esophageal cancer. It is possible that some misclassification between contiguous parts of the digestive tract has occurred in reporting cancer in the relatives. In Western countries, however, cancers of the oral cavity/pharynx and esophagus show remarkable similarities in etiology and geographic distribution [101]. Moreover, synchronous and metachronous cancers in both the esophagus and the OP have been observed [100, 102, 103]. Familial aggregation of SCCE and oral/pharyngeal cancers may be due to shared environmental exposure to the main risk factors, mainly alcohol and tobacco. However, the risk of SCCE was not increased in individuals with a FH of cancer at other sites related to smoking (lung), alcohol (liver) or both factors (larynx). A joint inherited susceptibility to both cancers, thus, cannot be ruled out. Similarly, in high-risk areas of China, cancers of the esophagus and the gastric cardia are both very frequent, and a link between the two is conceivable [95].

Stomach cancer. Our study confirms that a FH of stomach cancer in FDR increases the risk of stomach cancer, particularly when the relative was diagnosed in earlier age. No differences emerged according to the sex of such relative. The results are in line with most epidemiologic studies investigating the issue, which reported a risk of gastric cancer between 1.5 and 3.5 for subjects with relatives with stomach cancer [4, 20, 23, 97, 98, 104-121]. Both the long term decline in gastric cancer in developing countries and the results of migrant studies suggest that environmental factors predominate in the etiology of gastric cancer. Nonetheless, the observed association may therefore reflect a shared environment, but genetic factors might also contribute to this familial clustering. Helicobacter pylori (HP) infection, one of the most important risk factor for stomach cancer, tends to aggregate among family members [122-124]; thus, the excess risk associated with FH may be at least partly due to concordance in HP infection status. However some studies found that FH of gastric cancer and HP infection were independent risk factors [125, 126], although in one study this was found only for women [126]. Similarity in lifestyle (e.g., smoking) and diet (e.g, processed/salted food) could also partially explain the familial aggregation of stomach cancer. On the other hand, the role of genetic susceptibility has been strengthened by the fact that discovery mutations in the E-cadherin gene have recently been associated with an increased risk of familial or sporadic gastric cancer [127] and by the observation that genetic polymorphisms of Interleukin-1 (IL-1) promote development of the intestinal type of gastric cancer associated with HP infection [128, 129]. Among the studies that investigated the association with FH of other cancers, a cohort study from Utah found an association between a FH of brain/central nervous system, female genital cancer and

stomach cancer [20]. A cohort study from Iceland found an increased risk of stomach cancer for a FH of cervix, endometrium, ovary, colon, esophagus, thyroid and brain cancer [23]. In a case-control study from Sweden the risk of cancer of the cardia was increased among persons who reported having a FDR with breast cancer [98]. In the ‘‘Swedish Family Cancer Database’’ sites associated with stomach cancer were endometrium and urinary bladder in siblings, and male genital cancers other than prostate in fathers [106]. Some studies reported no significant association between a FH of any cancer other than stomach cancer [97, 113]. In our study no increase in risk was found for a FH of cancer at all sites excluding stomach. The number of exposed subjects were however small and the estimates were subject to substantial random variation for any specific cancer site.

Colorectal cancer. These analyses showed that FH of colorectal cancer increased the risk of developing colorectal cancer of about three-fold. The significance of the association persisted even after multiple testing correction. No appreciable difference in risk emerged in this study with the age at diagnosis in relatives. Elevated, but not significant, risk of colorectal cancer emerged for FH of stomach, liver, ovarian, prostate, all HL, and skin cancers. However, for these across sites associations, significance at 0.05 level was lost after Benjamini-Hochberg correction.

Many epidemiological studies estimated the risk of colorectal cancer in individuals with a FH of the disease, with RR estimates ranging from 1.21- to 9.33-fold [10, 130] as compared to individual without FH. Colorectal cancer clustering in families may be explained by common unhealthy behaviors among members of the same family, including a diet rich red and processed meat, smoking and alcohol, inherited genetic

susceptibility and/or their interaction. The two major familial syndromes that underline inherited susceptibility are the familial adenomatous polyposis and HNPCC, also called Lynch syndrome. Familial adenomatous polyposis is a rare autosomal dominant syndrome, caused by mutations in the APC gene (the prevalence of mutations is about 1 in 10,000 individuals), and characterized by the development of multiple colorectal adenomas. One or more of the polyps will almost progress to cancer. HNPCC is an autosomal dominant syndrome, caused in most cases by germinline mutations in two genes in DNA mismatch repair, MSH2 and MLH1. HNPCC is most strongly associated with colon cancer but also to other malignancy, involving the endometrium, urinary tract, stomach, and biliary system. In summary, no more than 10% of colorectal cancers are due to inherited mutations in these single genes. A higher but very uncertain proportion may partially caused by inheritance of variants in metabolic or other genes that alter the way in which lifestyle factors influence the risk. These include selected polymorphisms of genes involved in the metabolism of aromatic and heterocyclic aromatic amines (present in cigarette smoke and red meat cooked by high-temperature cooking techniques), such as NAT2, NAT1, and CYP1A2 [131], folate, such as MTHFR, and nonsteroidal anti-inflammatory drugs (i.e., NSAIDs), such as the UGTs gene. Other genes that have been implicated in the aetilogy or progression of colorectal cancer include other DNA repair genes [132-135], hormone receptors [135], and insulin, insulin-like growth factors.

The elevated colorectal cancer risk found for FH of ovarian and prostate cancer suggests an involvement of the BRCA1 gene, which has been linked to colon cancer [19, 136]. Moreover, ovarian cancer risk is increased among HNPCC families. Also liver cancer has been moderately associated to BRCA1 mutation [136]; this, together with similar

alcohol and smoking behaviors among family members could at least in part explained the association between colorectal cancer and FH of liver cancer. However, against the hypothesis of un involvement of the BRCA1 gene, no association emerged with FH of breast cancer.

FH of gastric cancer has been found to increase colorectal cancer in another study [137]. This association may be explained by shared unhealthy behaviors among family members, including smoking and alcohol, as well as by HNPCC, which increases the risk of developing both colorectal and stomach cancer. However, the absence of association with FH of other cancer sites associated to HNPCC, e.g., endometrium and urinary tract, do not support such explanation. Moreover, a recent meta-analysis suggested a small elevation in colorectal cancer risk for HP infection, a strong risk factor for stomach cancer.

Beside colorectal cancer, polymorphisms within the MTHFR gene have been recently related to hematopoietic cancers in some [138, 139] but not all studies [140, 141]. This possible relation may partially explain the observed association of colorectal cancer risk with FH of HLP cancers. However, this association is yet to be confirmed by other studies.

Liver cancer. A significant threefold increased HCC risk was found for subjects with a FH of liver cancer, with stronger association for a male affected relative and a younger age at relative diagnosis; further adjustment for serological markers of hepatitis B and/or C viruses (HBV, HCV) infection did not substantially change these results (OR= 2.38, 95% CI, 1.01-5.58, data not shown). The RR estimate in this database is consistent with the overall evidence from published data. The pooled OR from the meta-analysis I

performed was RR of 2.50 (95% CI, 2.06-3.03) from 13 studies [8]. HBV and HCV transmission among family members, together with other shared environmental risk factors (e.g. alcohol drinking), may be responsible for part of the observed familial aggregation of liver cancer. Familial clustering of HCC has been frequently reported in eastern Asia [44, 142], where HBV infection is common [143]. However, FH was found to be related to HCC risk even after adjustment for other risk factors, and in subjects without hepatitis B and C serum markers [39, 40, 45]. Several inherited disorders, including α1-antitrypsin deficiency, hemocromatosis, and porphrya cutanea tarda, have been associated to HCC [144]. Specific genes studies have yielded conflicting results. Directly opposite effects have been reported for the association of several enzymes, including cytocrome P450 2E1 (involved the metabolism of smoking and alcohol), the glutathioneS-transferases and the epoxide hydrolases (involved in the process of aflatoxine B1 detoxification in hepatocytes). No apparent associations were observed with polymorphisms in the gene encoding alcohol-metabolizing enzymes [144].

Pancreatic cancer. The current analysis showed a 40% not significant increased risk of pancreatic cancer among subjects with FH of the same cancer, and an OR of pancreatic cancer of 2.4 (95% CI, 1.2-4.7) for those with FH of stomach cancer. This association emerged for a male affected FDR. Epidemiological literature showed nearly a two-fold increased risk for developing pancreatic cancer in subjects with FH of cancer at the pancreas [145, 146]. To my knowledge, among the few available studies analyzing across sites associations [21, 41, 88, 100, 147, 148], only one cohort of about 1.1 subjects in the Cancer Prevention Study-II supported a role of FH of stomach cancer in pancreatic cancer aetiology, with RRs of 1.22 (95% CI, 1.09-1.36) for one and 1.53

(95% CI, 1.02-2.30) for two or more affected relatives [148]. Although chance could not be excluded as a possible explanation of the association observed, also in consideration that significance was lost after Benjamini-Hochberg p-value correction, HP infections within families may also play a role [149, 150]. HP infection is an established risk factor for stomach cancer [151], being responsible for two-thirds of all gastric cancers [152]; moreover, a recent meta-analysis of six studies involving a total of 2,335 patients found a significant association between HP seropositivity and the development of pancreatic cancer [153]. However, the fact that the association between FH of stomach cancer and pancreatic cancer is limited to having a male affected relative weights against a major role of HP infection in familial clustering. Also tobacco smoking, which is positively related to both pancreatic and stomach cancer, may also explain part of the observed association. However, the absence of any relation between stomach cancer and FH of pancreatic cancer does not support such explanations.

Laryngeal cancer. These analyses confirm that a FH of laryngeal cancer in FDR increases the risk of laryngeal cancer, and provides further evidence that the risk is higher when the affected relative is younger than 60 years. Moreover, further analyses on this database (data not showed) found that the risk is independent from that of tobacco smoking and alcohol drinking. A FH of skin, colorectal or kidney cancer are associated to a somewhat increased laryngeal cancer risk. However, these across sites associations are not longer found after correction for multiple comparisons. A few epidemiologic studies have investigated the risk of laryngeal cancer in subjects who have a FH of cancer. A study based on a Utah population database found a standardized incidence ratio of laryngeal cancer of 8.0 (95% CI, 2.1-17.9) in 246 subjects with FH of

laryngeal cancer [20]. A case-control study conducted in China on 288 laryngeal cancer cases found an OR of 2.3 (95% CI, 1.2-4.5) in subjects with a FH of that malignancy [154]. The International Head and Neck Cancer Epidemiology Consortium (INHANCE) pooled data from case-control studies, including a total of 2,357 laryngeal cancer cases, showed an OR of 2.1 (95% CI, 1.6-2.7) in subjects reporting a FDR with head and neck cancer [155]. A study from Poland on 2839 FDR of 760 laryngeal cancer patients, found that the incidence rates of laryngeal, lung, stomach, and breast cancers (only early-onset breast cancer) were significantly increased in families with laryngeal cancer, while that of colon cancer was significantly decreased, as compared to the general population [156]. Several genetic polymorphisms in genes involved in the metabolism of carcinogens, DNA repair or in several other processes have been associated to laryngeal cancer risk, although results were not always consistent [157, 158]. Given that the differential ability to metabolize carcinogens matters only when exposure occurs, it is also possible that the familial risk reflects both a higher genetic susceptibility to laryngeal cancer together with an aggregation of exposures. However, no increased risk was found for FH of oral cavity and esophageal cancer, i.e, other cancers also caused by alcohol and tobacco. For lung cancer, too, no significant excess emerged. The associations with FH of skin, kidney and colorectal cancer are not confirmed in other studies and should therefore be interpreted with caution.

Breast cancer. These results confirm that breast cancer risk is increased in women with a FH of breast cancer. The risk is somewhat higher when the affected FDR was diagnosed before the age of 60. Elevated risk of breast cancer emerged for FH of colorectal, skin, prostate, and all HLP cancers combined. After adjustment for multiple

testing, in addition to that with FH of the corresponding cancer, the association with FH of colorectal cancer persisted.

The overall OR of 2.4 associated with FH of breast cancer is comparable with most estimates in other populations [6, 7]. In particular, a collaborative re-analysis combining individual data from 52 epidemiological studies found ORs of 1.80 (95% CI, 1.70-1.91), 2.93 (95% CI, 2.37-3.63), and 3.90 (95% CI, 2.03-7.49) for one, two, and three or more affected FDR, respectively, as compared to no affected relative. Moreover, the risk ratios were greater the younger the relative was diagnosed [7].

It has been estimated that 5–10% of all breast cancers can be attributed to highly penetrant germinline mutations [159]. A further proportion is caused by a number of moderate penetrance genes. At present, around 20 low penetrance genes have been identified cumulatively; when all have been discovered, these may contribute a higher proportion of familial breast cancer [160]. These genes appear to interact, although this has yet to be completely elucidated. The most important identified with high-risk genes, conferring 40–85% lifetime risk, included BRCA1, BRCA2 and P53. BRCA1 and BRCA2 are human genes that belong to a class of genes known as tumor suppressors. In normal cells, BRCA1 and BRCA2 help ensure the stability of the cell's genetic material (DNA) and help prevent uncontrolled cell growth. Whilst pathogenic mutations in BRCA1/2 have high penetrance, they occur relatively rarely: 1 in 500-1000 individuals carry a pathogenic mutation in BRCA1, and 1 in 600-800 in BRCA2. Pathogenic mutations in BRCA1/BRCA2 are known to increase the risk of breast cancer by 10- to 20-fold, and to confer an increased lifetime risk of ovarian cancer of 40-60% for BRCA1 and up to 30% for BRCA2. Mutations in the tumor suppressor gene P53 (Li-Fraumeni syndrome) also give a high risk of breast cancer (i.e., an 18- to 60-fold

increased risk as compared to the general population), although the frequency of germline mutations is even lower (<1%). Moderate risk genes (20–40% risk) included PALB2 (which encodes for a protein that was identified as a binding partner of BRCA1), BRIP1 (which encodes for a protein that interacts with BRCA2), ATM (gene of ataxia telangiectasia) and CHEK2 (tumor suppressor gene, with a mutation prevalence of about 0.2-1%). A number of common polymorphisms have now been identified to be associated with a slightly increased or decreased risk of breast cancer; most of these are involved in the metabolism of carcinogen.

Some but not all studies reported an increased risk of breast cancer in women with a FH of ovarian cancer. In this study the estimated OR was 1.7 (not significant). The excess of ovarian cancer in relatives of women with breast cancer has been related to BRCA1 mutations, which have been shown to account for the majority, if not all, of hereditary breast-ovarian cancer families [161]. An increase in the risk of breast cancer of about 50% in women with a FH of colorectal cancer was observed in this and other studies. Cancers of the breast and CRC share some etiological factors, including positive social class correlates, a diet poor in fruit and vegetables, and alcohol drinking and a low level of physical activity [162]. Moreover, an increased risk of colon cancer has been observed in carriers of *BRCA*1 mutations [18]. The same study showed an increased risk of prostate cancer, as well. An elevated risk of prostate cancer in relatives of breast cancer patients has been reported in a few but not all studies. In this study a significant 60% increase in breast cancer risk was observed for FH of prostate cancer. We found a significant association between breast cancer and FH of all HLP cancers, with an OR of 1.5 (95% CI, 1.0-2.7) for FH of leukemia (data not shown). Breast cancer and leukemia have been linked within families that have rare germ-line mutations in either the P53

gene (Li-Fraumeni Syndrome) or ataxia telangiectasia gene (ATM) [17, 163, 164]. However, these mutations are rare and thus responsible for few cancer cases.

We found a significant increased breast cancer risk for FH of skin cancer. Epidemiological studies have provided suggestive evidence of a link between melanoma and breast cancer. Moreover, registry-based and hospital-based studies have shown an increased risk of breast cancer among female melanoma survivors and vice versa [165-169]. Breast cancer and cutaneous melanoma occur at a higher frequency than expected by chance in the same individual. In addition, a high risk of breast cancer in relatives of cutaneous melanoma patients from families with strong aggregations of cancer has been suggested [170]. The commonly reported environmental risk factors for breast cancer and cutaneous melanoma are dissimilar and could not explain the observed clustering; otherwise, a genetic relationship between cutaneous melanoma and breast cancer has been suggested by reports of a higher risk of cutaneous melanoma among BRCA2 mutation carriers and an elevated frequency of breast cancer in CDKN2A 113insArg mutation families [19, 171]. Furthermore, in Polish population, a common missense variant of the CDKN2A gene (A148T) appeared as a low-penetrance breast cancer susceptibility gene [172].

Endometrial cancer. The present analysis found that a FH of uterine cancer in FDR is associated with an increased risk of endometrial cancer, and the risk seemed to be higher for FDR younger than 60 years at diagnosis. A FH of cancer at other sites, including brain, CRC, stomach, and OP, was also associated with a somewhat increased risk of endometrial cancer. After multiple testing corrections, none of these associations however were statistically significant at 0.05 level. Only a few studies have evaluated

the risk of endometrial cancer associated with history of cancer in FDR [173-179]. Most studies [173, 174, 176, 177, 179] found an association with FH of endometrial cancer, but in one study [175] there was no association. Results from this database are therefore compatible with the overall epidemiologic evidence. Cancer of the cervix and corpus uteri differ considerably in terms of aetiology [180, 181]. However, many women were not able to distinguish between cancer of the cervix and of the corpus uteri in relative. When subgroup analyses were performed, the OR of endometrial cancer was 2.97 (95% CI, 0.74-11.16) for FH of endometrial cancer, and 0.84 (95% CI, 0.04-16.53) for FH of cervical cancer. Specific genetic bases of endometrial cancer have been studied. HNPCC is most strongly associated with colon and endometrial cancer [16]. The lifetime risk of endometrial cancer among women with HNPCC is 50%–60%. However, overall, HNPCC accounts for only 2% of all endometrial cancers. Other potential genetic etiologies have been investigated using DNA analysis to examine potential mutations with likely physiologic relationships to development of endometrial malignancies. Genes with mutations that are associated with endometrial cancer with unknown magnitude and clinical significance include the sex hormone binding globulin gene; Fas gene promoter; P53 gene; mitochondrial gene polymorphisms; the methylenetetrahydrofolate reductase gene; the methylguanine DNA methyltransferase gene; the androgen receptor gene; CYP19 and CYP17 gene mutations; ATM, CHEK2, ERBB2 haplotypes; the MDM2 gene mutation and the catechol-O-methyltransferase gene [179]. Two studies [174, 177] found an association between FH of cancer of the CRC (OR=1.9) or the colon (standardized incidence ratio=3.3) and endometrial cancer risk. With regard to clusters of intestinal cancer, in this database one case (aged 45 years) reported four relatives with intestinal cancer, that is, the mother, two brothers and

one sister. One case and one control (both aged 66 years) reported two FDR with intestinal cancer. However, the case also reported a grandfather with intestinal cancer and three relatives with a history of cancer at other sites, whereas the control reported a grandmother with breast cancer. As mentioned before, HNPCC is characterized by an increased risk of colon cancer and other malignancies, including cancers of the endometrium, ovary, stomach, small intestine, and brain. This may at least partially explain some observed elevated risk of endometrial cancer associated to FH of these cancers [16]. The association with brain cancer, moreover, may be also in part attributed to misclassification of metastasis. Although endometrial and breast cancers share some reproductive, hormonal and genetic risk factors, in this and other studies [173-175, 177, 178], a FH of breast cancer was not associated with endometrial cancer risk.

Ovarian cancer. This study confirms that a FH of ovarian cancer in FDR increases the risk of ovarian cancer. A FH of a few other cancer sites, including breast, CRC, larynx and all HLP, was also positively associated with ovarian cancer risk, and the OR was increased for FH of any cancer. After controlling for multiple comparisons, however, the only significant associations (with 0.05 level) were those between ovarian cancer and FH of cancer at ovary and breast.

The finding of an elevated risk of ovarian cancer in subjects with a FH of cancer at the same site are in broad agreement with other reports [20, 21, 177, 182-192], although the point estimate obtained in this analysis is somewhat higher than in other studies (i.e., 7.4). Given the broad confidence interval due to the small number of subjects reporting a FH of ovarian cancer, results from this analysis none the less do not contrast with the pooled OR of 3.1 (95% CI 2.6–3.7) estimated from a meta-analysis of published studies

[15], which became 4.0 when only case–control studies were considered. A similar association with FH of ovarian cancer emerged according the age of the onset of ovarian cancer in the affected FDR, consistently with other reports [15].

The OR of 2.3 in women with a history of breast cancer in FDR is slightly higher than in most studies, ranging between 1.3 and 1.8 [21, 182-185, 187, 189, 190, 193]. The associations between ovarian malignancies and FH of the cancer at the same as well as of breast cancer may be attributable to *BRCA1/2* mutations and to a shared response of ovary and breast to pregnancy hormones [194, 195]. The familial aggregation of ovarian The 4-fold increase in prostate malignancy with colorectal and oral and pharyngeal cancer may be attributed in some degree to HNPCC-related genes, since HNPCC shows colonic malignancies with these cancers, as well as to shared environmental factors, including smoking habits (for both associations) and diet/obesity (for the association with colorectal cancer).

Prostate cancer. In our study, a FH of prostate cancer in FDR was directly associated with prostate cancer, with no meaningful differences according to age of the affected relatives. The risk of prostate cancer was also increased in subjects with a FH of cancers of the ovary, bladder, kidney, lung, and, to a lesser extent, CRC and breast. However, after adjustment for multiple testing, most of the associations across sites lost their significance.

cancer risk in subjects with a FH of the disease is somewhat higher than the risk of 2–3 estimated by some systematic meta-analyses [14, 196-198]. Both common environmental risk factors and genetic influences can contribute to the familial clustering of prostate carcinoma. Genome-wide linkage searches have implicated

several regions in inherited prostate cancer, e.g. HPC1, PCaP, HPCX, CAPB, HPC20, HPC2/ECAC2 [196, 197], but confirmatory studies at these loci have produced discordant results. Other susceptibility genes with an associated increased risk for prostate cancer include BRCA1, with carriers having a 3.3-fold increased risk of prostate cancer compared to the general population [18] and BRCA2 associated with a 2–5% increased risk of disease [199]. With regard to shared environmental factors among family member, obesity, which runs into family, a diet rich in unsaturated fats, and cigarette smoking may be in part responsible to the prostatic cancer clustering. The epidemiologic studies investigating the association of FH of other cancers and risk of prostate cancer have yielded contrasting results: significant associations with a FH of cancers of the ovary [200, 201], stomach [202, 203], CRC [20, 201, 204], skin [205], breast [200-202, 206-208], kidney [200, 202, 209], central nervous system [20, 210], Hodgkin's and non-Hodgkin’s lymphomas [20, 201], and liver [201] have been reported. Other studies found no significant association with FH of any cancer other than prostate cancer [211-213]. An association of prostate cancer risk with FH of ovarian cancer was found in this database. Although the point estimate of the OR was around 7, this was based on 10 cases and 3 controls only, with a lower confidence limit of 1.4. Most of the ovarian cancers reported appeared to be rapidly lethal. This rules against minor ovarian conditions (e.g., ovarian cysts) being misreported as cancer. A relation between these 2 cancers is not implausible. Male carriers of germline mutations of the genes *BRCA1* and *BRCA2*, 2 genes predisposing to breast and ovarian cancer, have been found to be at high risk of prostate cancer. Compared to noncarriers, men carrying a mutation of *BRCA1* had a relative risk of prostate cancer of 3.3 [18] and men carrying a mutation of *BRCA2* of 4.7, which became 7.3 below age 65 years [19]. The

50% not significant elevate prostate cancer risk emerged for FH of breast cancer may be explained by the same mutations of the genes *BRCA1* and *BRCA2*. A further suggestion that in a few prostate cancer cases breast/ovarian cancer predisposition genes may play a role is given by the clusters of these cancers found in relatives: 1 prostate cancer case (aged 61 years) reported ovarian cancer in 3 sisters, and 3 prostate cancer cases (aged 70, 72 and 73 years) reported 3 cases of breast cancer each in female FDR. I also found an association with FH of cancers of the bladder and the kidney; for the former the association was higher for female relatives, and could thus not be explained by misreporting of prostate cancers. It is possible that cases had urinary symptoms before diagnosis and were thus more sensitized to urinary tract cancers in the family than controls. However, a high risk of kidney cancer in FDR of prostate cancer cases has been previously reported. In a cohort of Swedish families with hereditary prostate cancer, the risk of kidney cancer was 2.5 (95% CI, 1.2– 4.8) overall and 4.6 (95% CI, 1.8 –9.5) for women [202]. In a record linkage study from Iceland, the risk of kidney cancer in FDR of prostate cancer cases was 1.7 (95% CI, 0.7–3.8) for males and 2.5 (95% CI, 1.1–5.7) for females [209]. The association between the risk of prostate cancer and FH of lung cancer has not been reported by others in my knowledge. A possible explanation of the observed elevated risk associated to FH of cancer at CRC may be the sharing of a diet rich in red meat among family members, as it was related to both intestinal and prostate cancer.

Renal cell cancer. These analyses confirm that a FH of kidney cancer in FDR increases the risk of renal cell cancer. An elevated, not significant, risk of renal cell cancer emerged for FH of uterine and ovarian cancer.

FH of renal cell cancer was positively associated to the risk of the corresponding cancer even after multiple comparisons correction. Findings of an elevated risk of renal cell cancer in subjects with a FH of kidney cancer are in broad agreement with other reports, although our point estimate is somewhat higher than in other studies [20, 214-220]. Given the broad confidence interval due to the small number of subjects reporting a FH of kidney cancer, our results are nonetheless in line with other reports. A multicenter case-control study in Central Europe, including 1,097 cases of kidney cancer and 1,555 controls, found an OR of kidney cancer of 1.40 (0.71-2.76) for FH of kidney cancer [220]. A population-based case-control study from Canada [214], based on 518 RCCs and using a mailed questionnaire, did not find an association between FH and risk of RCC (OR= 1.1 in both sexes). Conversely, three other population-based case-control studies [215-217] reported significantly increased risks: a stud from Denmark on 368 RCCs found an OR of 4.1 in men and 4.8 in women [215]; an international study conducted in Denmark, Sweden, Germany, Australia, and the United States and including 1,732 RCCs found an OR of 1.6 for one FDR affected, whereas 7 cases and no controls reported two affected relatives [216]; and a study from Los Angeles, United States, based on 550 RCCs, found an OR of 2.5 for an affected FDR [217]. Three linkage studies analyzed the familial risk of kidney cancer. In a systematic population-based assessment using the Utah Population database [20], based on 687 kidney cancers, the familial RR was 2.5. In the nationwide Swedish Family Cancer database [218], including 23,137 kidney cancers, the standardized incidence ratio of kidney cancer was 1.6 in offspring and 4.7 in siblings of kidney cancer cases. Finally, a population-based familial aggregation analysis based on 1,078 renal cell cancer cases in Iceland [219] found relative risks of 2.5 for siblings and 2.2 for parents. Several

autosomal dominant inherited syndromes predisposing to renal cell cancer have been described, the most common being the von Hippel-Lindau syndrome, characterized by excesses of renal cell cancer and other neoplasms, including those of the central nervous system, eye, inner ear, endocrine glands, and pancreas, and caused by germ-line mutations in the von Hippel-Lindau tumor suppressor gene (i.e., VHL gene) [221, 222]. However, these syndromes are rare and probably, as for many other cancers, most of the familial risk in older patients is not due to these highly penetrant genes [4]. Other susceptibility genes may exist, with lower penetrance but much higher frequency in the population, which might account for more cases of renal cell cancer. Identification of these genes is extremely difficult because their low penetrance does not cause striking familial aggregations [4]. A common lifestyle characterized by smoking and overweight/obesity may shared by members of the same family may also explained some of the familial clustering observed. As concerning the association of FH of other cancers with renal cell cancer, in the Utah Population database [20], no significant associations were found between FH of other cancers and risk of kidney cancer, whereas in the Swedish Family Cancer database [218], discordant sites that were associated with kidney cancer in siblings were ovaries, endocrine glands. In our study too, cases reported a FH of ovarian cancer more frequently than controls, although the association was not significant, probably given the small numbers.

Strengths and limits. These case–control studies have a few limitations that should be considered in interpreting these findings. Selection bias should be limited as in the control group were included subjects admitted for a wide spectrum of acute, non-

neoplastic conditions, unrelated to the major risk factors for cancer. Moreover, the almost complete participation has likely reduced selection bias.

Data on FH of cancer was self-reported, and it is possible that cancer patients may be more interested in understanding their family cancer history in greater detail, especially if multiple family members have been affected by a specific cancer [223]. However, an analysis in this population showed a good reliability of data on FH of all cancers provided by hospital controls, with a kappa statistic of 0.70 for all cancers, 0.70 for liver cancer, and 0.80 for any digestive tract cancers [224]. In a recent systematic review from The Agency for Healthcare Research and Quality (AHRQ), self-reported FH about common cancers appeared to be fairly accurate. Specificity across all cancers types ranged from 91 to 99%, while the sensitivity values showed greater variability, with breast cancer having the highest values (around 85-90%) [225]. In the Connecticut Family Health Study, reports from FDRs were more accurate than those from second degree relatives; we therefore considered FDRs only in our analysis [226].

With reference to confounding, RR estimates we adjusted for the main recognized risk factors, including tobacco smoking, alcohol drinking, overweight/obesity, and, for female cancers, reproductive factors.

A possible limitation in these studies is the lack of information on important risk factors for cancer among family members, including alcohol drinking (in particular for laryngeal, oral, and esophageal cancer), smoking habits (in particular for oral, laryngeal cancer, pancreatic cancer), some inherited diseases, diet (in particular for colorectal and stomach cancer), body weight, HP and HBV/HCV infection (for stomach and liver cancer, respectively), which may partly explain the observed familial aggregation across sites.

In the present study 13 cancer sites and 17 exposure variables were covered and some associations were likely to appear by chance. We presented also significant results after multiple comparisons adjustment. However, without adjustment, only one OR was significantly decreased, compared with 25 significantly increased ones (Table 3), suggesting that chance findings could be limited.

1.5 Tables and Figures of Chapter 1

Table 1. Number of cases of selected cancer sites and controls in the network of case-control studies, and corresponding median age. Italy and Switzerland, 1991-2009.

Cancer site	Cases (M/F)	Median age (yrs)	Controls (M/F)	Median age (yrs)	Total (cases/controls)
OP	1190/278	58	2553/1208	58	1468/3761
Rhinopharynx	157/471	52	41/123	52	198/594
SCCE	438/67	60	919/340	60	505/1529
Stomach	143/87	63	286/261	63	230/547
Colorectum	1401/989	62	2586/2357	58	2390/4943
Liver	149/36	66	278/126	65	185/404
Pancreas	174/152	63	348/304	63	326/652
Larynx	770/82	62	1564/406	61	852/1970
Breast	-/3034	55	-/3392	56	3034/3392
Endometrium	-/367	60	-/798	61	367/798
Ovary	-/1031	56	-/2411	57	1031/2411
Prostate	1294/-	66	1451/-	63	1294/1451
Renal cell cancer	494/273	62	988/546	62	767/1534

Abbreviations: M, males; F, females; OP, oral cavity and pharynx; SCCE, squamous cell carcinoma of the esophagus; yrs, years.

Table 2. Distribution of cases of cancer at 13 different sites and corresponding controls according to a positive family history of selected cancers. Italy and Switzerland, 1991-2009.

	Cancer site (cases:controls)												
	OP	RP	SCCE	Stomach	CR	Liver	Pancreas	Larynx	Breast	Endometrium	Ovary	Prostate	Renal cell
FH													
OP[a]	53:47	3:6	21:12	4:14	35:63	1:16	10:16	21:43	32:37	10:11	17:32	25:36	21:33
Esophagus	16:32	1:5	12:14	3:4	17:32	2:2	4:4	6:15	21:20	2:10	6:17	4:15	6:15
Stomach	63:113	4:16	29:39	30:31	132:224	13:23	21:20	48:82	148:138	22:32	62:109	63:75	49:74
CR	47:116	12:17	22:47	10:31	221:166	6:28	19:33	49:80	150:112	28:38	60:89	84:58	39:63
Liver	39:86	6:19	15:28	13:20	91:141	22:18	20:25	34:63	107:107	15:32	44:88	51:50	32:60
Pancreas	18:61	6:6	6:17	2:14	37:62	4:6	10:15	15:37	38.49	6.23	21:37	26:22	11:32
Larynx	33:25	1:3	4:14	0:3	30:49	0:5	9:15	29:27	43:50	1:12	24:42	18:20	9:14
Lung	89:170	11:38	32:66	22:42	143:270	17:39	29:58	64:111	205:194	29:54	81:168	111:88	58:94
Bone	3:16	1:2	3:4	2:3	13:24	3:2	6:7	4:15	20:33	4:7	1:22	12:5	0:6
Skin	8:7	3:3	1:3	1:2	12:11	1:3	0:1	7:3	26:10	2:2	6:9	8:3	7:6
Breast	65:111	3:21	15:42	8:22	107:206	10:25	22:32	31:81	311:145	21:36	104:111	82:64	41:65
Uterus[b]	20:53	3:11	9:20	5:10	60:96	4:16	8:21	8:21	108:84	19:25	31:52	40:35	26:32
Ovary	2:10	1:3	1:3	0:2	15:17	1:1	1:2	3:5	19:13	3:4	27:9	10:2	5:3
Prostate	12:34	7:7	7:19	3:10	39:55	5:7	8:13	12:27	59:42	13:16	17:35	90:28	19:21
Bladder	11:24	1:0	0:4	1:7	26:47	4:3	2:6	5:15	42:37	5:7	10:29	31:10	6:15
Kidney	1:20	3:5	1:3	0:5	19:29	2:3	3:4	11:7	32:25	6:5	8:12	16:6	18:8
Brain	13:37	5:9	8:14	2:4	41:66	3:6	9:9	9:24	41:42	10:7	13:36	26:28	17:21
All HLP cancers[c]	26:65	6:15	6:21	4:13	52:82	2:10	10:16	10:16	92:57	9:16	32:54	32:32	27:31
All sites	441:922	69:181	161:342	96:199	919:1450	84:166	162:250	162:250	1256:1039	164:287	468:829	596:515	321:514
excluding cancer index	388:875	66:175[d]	149:328	66:168	698:1284	62:148	152:235	152:235	945:894	145:262[e]	441:820	788:964	303:506

Abbreviations: OP, oral cavity and pharynx; RP, rhinopharynx, CR, colorectum; HLP, hemolymphopoietic; SCCE, squamous cell carcinoma of the esophagus. [a] Including rhinopharyngeal cancer. [b] Including cancer of the cervix, endometrium and uterus not otherwise specified. [c] Including Hodgkin lymphoma, Non Hodgkin lymphoma or unspecified, leukemia and mieloma. [d] Family history of all cancer sites, excluding cancer of the oral cavity and pharynx. [e] Family history of all cancer sites excluding cancer of the uterus.

Table 3. Odds ratios[a] and 95% confidence intervals of cancer at 13 different sites according to family history of selected cancers. In bold, significant associations at level 0.05 after adjustment for multiple testing according to the Benjamini-Hochberg method. Italy and Switzerland, 1991-2009.

	Cancer site in the proband												
	OP	RP	SCCE	Stomach	CR	Liver	Pancreas	Larynx	Breast[b]	Edometrium[b]	Ovary[b]	Prostate	Renal cell
FH													
OP[c]	**2.6 (1.6-4.2)**	1.2 (0.3-5.7)	**4.1 (1.7-9.8)**	0.6 (0.2-1.9)	1.2 (0.8-1.8)	0.2 (0.0-1.2)	1.2 (0.5-2.9)	0.8 (0.5-1.5)	0.9 (0.6-1.5)	2.4 (0.9-6.1)	1.0 (0.5-1.8)	0.6 (0.4-1.1)	1.4 (0.8-2.4)
Esophagus	1.5 (0.7-3.2)	0.7 (0.1-6.2)	1.7 (0.6-4.5)	1.7 (0.4-7.9)	1.0 (0.6-1.9)	3.7 (0.4-32.0)	1.7 (0.4-7.5)	0.9 (0.3-2.5)	1.1 (0.6-2.0)	0.4 (0.1-2.0)	0.8 (0.3-2.2)	0.3 (0.1-1.0)	0.9 (0.3-2.5)
Stomach	1.4 (0.9-2.0)	0.5 (0.2-1.5)	1.8 (1.0-3.3)	**2.7 (1.6-4.6)**	1.2 (1.0-1.6)	1.8 (0.8-4.1)	2.4 (1.2-4.7)	1.2 (0.8-1.9)	1.2 (1.0-1.6)	1.8 (1.0-3.4)	1.2 (0.8-1.7)	0.9 (0.6-1.3)	1.3 (0.9-2.0)
Colorectum	1.0 (0.6-1.4)	3.1 (1.4-7.0)	1.1 (0.6-2.1)	0.8 (0.4-1.6)	**2.8 (2.3-3.5)**	0.5 (0.2-1.3)	0.9 (0.5-1.7)	1.5 (1.0-2.3)	**1.5 (1.1-1.9)**	1.6 (0.9-2.9)	1.6 (1.1-2.4)	1.5 (1.0-2.2)	1.3 (0.8-2.0)
Liver	1.1 (0.7-1.8)	0.8 (0.3-2.4)	0.9 (0.4-1.9)	1.8 (0.8-3.7)	1.4 (1.0-1.8)	**3.0 (1.4-6.5)**	1.5 (0.7-3.0)	1.3 (0.8-2.1)	1.2 (0.9-1.6)	1.2 (0.6-2.4)	1.3 (0.9-2.0)	1.2 (0.8-1.8)	1.1 (0.6-1.7)
Pancreas	0.9 (0.5-0.6)	1.8 (0.5-6.4)	0.8 (0.3-2.4)	0.3 (0.1-1.5)	1.4 (0.9-2.1)	2.3 (0.5-10.1)	1.4 (0.5-3.3)	1.3 (0.7-2.6)	0.8 (0.5-1.2)	0.7 (0.3-1.9)	1.6 (0.9-2.8)	1.3 (0.7-2.5)	0.7 (0.3-1.4)
Larynx	**3.3 (1.7-6.3)**	1.0 (0.1-14.5)	0.8 (0.2-2.9)	NE	1.3 (0.8-2.1)	NE	1.1 (0.4-2.8)	**2.8 (1.5-5.1)**	1.0 (0.6-1.5)	0.4 (0.0-2.9)	1.8 (1.0-3.2)	1.1 (0.6-2.3)	2.2 (0.9-5.3)
Lung	1.2 (0.9-1.7)	0.8 (0.4-1.7)	0.9 (0.6-1.6)	1.3 (0.7-2.3)	1.1 (0.9-1.4)	0.9 (0.5-1.8)	0.9 (0.5-1.5)	1.4 (0.9-2.0)	1.2 (0.9-1.4)	1.1 (0.7-1.8)	1.1 (0.8-1.5)	1.5 (1.1-2.0)	1.2 (0.9-1.8)
Bone	0.7 (0.2-2.6)	1.4 (0.1-16.2)	2.9 (0.5-16.0)	2.0 (0.3-12.8)	1.4 (0.7-2.8)	4.8 (0.6-36.8)	2.7 (0.8-9.0)	0.5 (0.1-1.7)	0.7 (0.4-1.2)	0.6 (0.1-2.3)	0.2 (0.0-1.4)	1.5 (0.5-4.7)	NE
Skin	3.3 (1.0-10.7)	4.6 (0.9-24.1)	1.1 (0.1-11.9)	1.5 (0.1-17.2)	2.2 (0.9-5.2)	0.4 (0.0-7.0)	NE	8.4 (1.7-41.8)	**3.0 (1.4-6.4)**	3.1 (0.4-22.9)	1.3 (0.4-4.2)	2.2 (0.5-9.1)	2.3 (0.7-7.2)
Breast	1.5 (1.0-2.2)	0.4 (0.1-1.3)	1.1 (0.5-2.2)	0.9 (0.4-2.2)	1.1 (0.8-1.4)	1.2 (0.5-2.8)	1.4 (0.7-2.6)	0.8 (0.5-1.2)	**2.6 (2.1-3.2)**	1.3 (0.7-2.4)	**2.3 (1.7-3.2)**	1.5 (0.9-2.0)	1.3 (0.8-2.0)
Uterus[d]	0.9 (0.5-1.7)	0.6 (0.1-2.4)	1.0 (0.4-2.6)	1.4 (0.5-4.3)	1.3 (0.9-1.9)	0.5 (0.1-1.7)	1.0 (0.4-2.3)	1.2 (0.6-2.3)	1.4 (1.0-1.9)	2.0 (1.0-4.0)	1.4 (0.8-2.2)	1.2 (0.7-2.0)	1.7 (1.0-3.0)

...CONTINUED

	Cancer site in the proband												
	OP	RP	Esophagus	Stomach	CR	Liver	Pancreas	Larynx	Breast[b]	Edometrium[b]	Ovary[b]	Prostate	Renal cell
FH													
Ovary	0.7 (0.1-3.4)	0.7 (0.1-7.6)	1.4 (0.1-15.8)	NE	2.1 (1.0-4.2)	9.7 (0.3-367.7)	1.7 (0.1-22.3)	2.6 (0.5-13.9)	1.7 (0.8-3.4)	3.3 (0.6-17.0)	**7.4 (3.3-16.6)**	7.4 (1.4-38.2)	4.1 (0.9-18.2)
Prostate	0.9 (0.4-2.1)	2.4 (0.7-7.6)	0.9 (0.3-2.8)	0.7 (0.2-2.7)	1.6 (1.0-2.4)	1.7 (0.5-6.3)	0.7 (0.3-2.0)	1.2 (0.5-2.6)	1.6 (1.1-2.4)	1.4 (0.6-3.3)	1.1 (0.6-2.0)	**3.9 (2.4-6.2)**	1.7 (0.9-3.2)
Bladder	1.0 (0.4-2.5)	NE	NE	0.3 (0.0-2.8)	1.2 (0.7-2.0)	2.6 (0.5-13.2)	0.9 (0.2-5.2)	0.8 (0.2-2.5)	1.2 (0.8-1.9)	3.0 (0.8-10.7)	0.9 (0.4-1.9)	**3.4 (1.6-7.3)**	0.9 (0.3-2.4)
Kidney	0.1 (0.0-0.6)	1.3 (0.3-6.3)	0.5 (0.0-5.4)	NE	1.6 (0.9-3.0)	0.8 (0.1-5.5)	2.1 (0.4-10.5)	2.9 (0.9-9.1)	1.4 (0.8-2.4)	4.2 (1.0-16.5)	1.7 (0.6-4.6)	3.4 (1.2-9.4)	**4.2 (1.8-9.8)**
Brain	1.0 (0.5-2.0)	1.8 (0.6-6.1)	1.5 (0.5-4.3)	1.4 (0.3-8.1)	1.4 (0.9-2.1)	1.0 (0.2-4.9)	2.2 (0.8-6.1)	0.8 (0.3-1.8)	1.2 (0.7-1.8)	4.2 (1.3-13.3)	0.8 (0.4-1.6)	1.1 (0.6-2.1)	1.4 (0.7-2.7)
All HLP cancers[e]	1.0 (0.6-1.8)	1.2 (0.4-3.3)	0.7 (0.2-2.0)	0.8 (0.2-2.4)	1.4 (1.0-2.0)	0.5 (0.1-2.9)	0.8 (0.3-2.1)	1.1 (0.6-2.2)	**1.7 (1.2-2.4)**	1.2 (0.5-2.8)	1.6 (1.0-2.5)	1.1 (0.6-1.9)	1.4 (0.8-2.5)
All sites	1.2 (1.0-1.5)	1.2 (0.8-1.7)	1.2 (0.9-1.5)	1.3 (1.0-1.8)	1.6 (1.4-1.8)	1.5 (1.0-2.2)	1.4 (1.0-1.9)	1.3 (1.1-1.6)	1.6 (1.4-1.7)	1.5 (1.2-2.0)	1.6 (1.4-1.9)	1.6 (1.3-1.8)	1.4 (1.2-1.7)
excluding cancer index	1.1 (0.9-1.3)	1.2[f] (0.8-1.7)	1.1 (0.8-1.5)	0.9 (0.7-1.3)	1.2 (1.1-1.4)	1.1 (0.7-1.6)	1.3 (1.0-1.8)	1.2 (1.0-1.4)	1.2 (1.1-1.4)	1.4[g] (1.0-1.9)	1.5 (1.2-1.7)	1.3 (1.1-1.5)	1.3 (1.1-1.6)

Abbreviations: OP, oral cavity and pharynx; RP, rhinopharynx, CR, colorectum; HLP, hemolymphopoietic; SCCE, squamous cell carcinoma of the esophagus.

[a] Adjusted for age, sex (when appropriate), study centre (when appropriate), year of interview, education, body mass index, alcohol drinking, tobacco smoking, and number of brothers and sisters. Reference category: no family history of the selected cancer. [b] Further adjusted for menopausal status, age at menopause, oral contraceptive and hormone replacement therapy use, parity. Odds ratios for breast cancer were further adjusted for age at first birth. [c] Including rhinopharyngeal cancer. [d] Including cancer of the cervix, endometrium and uterus not otherwise specified. [e] Including Hodgkin lymphoma, Non Hodgkin lymphoma or unspecified, leukemia and myeloma. [f] Odds ratio for family history of all cancer sites, excluding cancer of the oral cavity and pharynx. [g] Odds ratio for family history of all cancer sites, excluding cancer of the uterus.

Table 4. Odds ratios[a] (ORs) and 95% confidence intervals (CIs) of cancer at 13 different sites according to family history of selected cancers by sex and age at diagnosis[b] of the affected first-degree relative. Italy and Switzerland, 1991-2009.

	Male		Female		<60 years		≥60 years	
FH, cancer site in relatives	ca:co	OR (95% CI)	ca:co	OR (95% CI)	ca:co	OR (95% CI)	ca:co	OR (95% CI)
Cancer site in the proband: Oral cavity/pharynx								
Oral cavity/pharynx	50:37	3.1 (1.8-5.1)	3:10	0.8 (0.2-3.8)	25:16	3.1 (1.4-6.6)	20:23	2.1 (1.0-4.5)
Larynx	29:24	3.1 (1.6-6.0)	5:2	7.0 (0.7-66.7)	15:11	3.7 (1.4-9.7)	15:13	3.1 (1.2-7.7)
Any site	308:599	1.3 (1.1-1.59)	204:453	1.2 (1.0-1.5)	203:371	1.4 (1.1-1.8)	227:497	1.2 (1.0-1.5)
excluding oral cavity/pharynx	258:562	1.1 (0.9-1.4)	201:443	1.2 (0.9-1.5)	178:355	1.3 (1.0-1.6)	207:474	1.1 (0.9-1.4)
Cancer site in the proband: Rhinopharynx								
Oral cavity/pharynx	2:5	1.3 (0.2-7.8)	1:1	1.4 (0.1-17.2)	2:2	2.7 (0.3-26.3)	1:3	0.9 (0.1-10.7)
Colorectum	7:8	4.0 (1.3-12.2)	5:9	0.5 (0.1-1.7)	10:14	3.1 (1.3-7.6)	12:17	3.1 (1.4-7.0)
Any site	51:116	1.4 (0.9-2.2)	28:96	0.9 (0.5-1.4)	32:92	1.1 (0.7-1.8)	42:91	1.5 (1.0-2.4)
excluding oral cavity/pharynx	49:111	1.4 (0.9-2.2)	27:95	0.9 (0.5-1.4)	30:90	1.0 (0.6-1.7)	41:88	1.5 (0.9-2.4)
Cancer site in the proband: SCCE								
Esophagus	10:12	1.6 (0.5-4.9)	2:2	1.9 (0.2-16.9)	6:7	1.3 (0.4-5.2)	6:7	2.1 (0.5-9.2)
Oral cavity/pharynx	17:11	3.8 (1.5-9.7)	5:1	7.6 (0.5-109.9)	11:5	4.1 (1.3-12.9)	8:7	2.8 (0.7-11.2)
Any site	124:229	1.3 (0.9-1.7)	72:158	1.2 (0.8-1.7)	75:155	1.2 (0.8-1.8)	93:209	1.1 (0.8-1.5)
excluding esophagus	114:217	1.2 (0.9-1.7)	70:156	1.2 (0.8-1.7)	69:148	1.2 (0.8-1.8)	87:202	1.0 (0.7-1.5)
Cancer site in the proband: Stomach								
Stomach	17:18	2.7 (1.3-5.5)	13:13	2.6 (1.2-5.9)	8:5	4.4 (1.4-14.2)	9:19	1.3 (0.6-2.9)
Any site	62:136	1.3 (0.9-1.9)	48:101	1.3 (0.9-2.0)	36:74	1.4 (0.9-2.3)	36:91	1.1 (0.7-1.7)
excluding stomach	45:118	0.9 (0.6-1.4)	35:88	1.0 (0.6-1.5)	28:69	1.3 (0.8-2.0)	27:72	0.9 (0.6-1.5)
Cancer site in the proband: Colorectum								
Colorectum	129:79	3.5 (2.6-4.7)	101:93	2.2 (1.7-3.0)	197:119	3.5 (2.7-4.4)	206:134	3.2 (2.6-4.1)
Any site	604:945	1.6 (1.4-1.8)	470:720	1.6 (1.4-1.8)	432:726	1.5 (1.3-1.8)	536:828	1.6 (1.4-1.8)
excluding colorectum	475:866	1.3 (1.1-1.4)	369:627	1.3 (1.1-1.5)	335:656	1.2 (1.0-1.4)	379:722	1.2 (1.0-1.3)

...CONTINUED

	Male		Female		<60 years		≥60 years	
FH, cancer site in relatives	ca:co	OR (95% CI)	ca:co	OR (95% CI)	ca:co	OR (95% CI)	ca:co	OR (95% CI)
Cancer site in the proband: Liver								
Liver	15:10	3.6 (1.4-9.1)	8:8	1.1 (0.5-2.3)	9:9	3.6 (1.2-11.4)	7:8	1.7 (0.5-5.6)
Any site	60:113	1.5 (0.9-2.4)	41:89	1.4 (0.8-2.4)	33:78	1.4 (0.8-2.5)	45:105	1.2 (0.7-2.0)
excluding liver	45:103	1.1 (0.7-1.8)	33:81	1.1 (0.6-1.9)	24:69	1.0 (0.5-1.8)	38:97	1.0 (0.6-1.7)
Cancer site in the proband: Pancreas								
Pancreas	6:11	1.1 (0.4-3.5)	4:4	2.0 (0.4-8.8)	3:3	1.7 (0.3-9.6)	6:11	1.3 (0.4-3.9)
Stomach	16:11	3.5 (1.5-8.1)	6:10	1.4 (0.4-4.3)	7:7	2.3 (0.7-7.5)	11:10	2.3 (0.9-5.9)
Any site	113:165	1.5 (1.1-2.1)	77:128	1.3 (0.9-1.9)	79:113	1.6 (1.1-2.3)	84:149	1.2 (0.9-1.8)
excluding pancreas	107:154	1.5 (1.1-2.1)	73:124	1.2 (0.8-1.8)	76:110	1.6 (1.1-2.3)	78:138	1.2 (0.8-1.7)
Cancer site in the proband: Larynx								
Larynx	24:23	2.7 (1.4-5.3)	5:4	3.2 (0.7-14.0)	15:13	3.6 (1.5-8.8)	10:16	1.3 (0.5-3.2)
Skin	3:3	4.8 (0.7-31.9)	4:0	∞	3:2	3.2 (0.4-26.5)	5:1	23.5 (2.2-248.3)
Any site	217:402	1.4 (1.1-1.8)	164:301	1.4 (1.1-1.8)	142:256	1.4 (1.1-1.8)	185:367	1.2 (1.0-1.6)
excluding larynx	193:379	1.3 (1.0-1.6)	159:297	1.3 (1.0-1.7)	127:243	1.2 (0.9-1.6)	175:351	1.2 (0.9-1.5)
Cancer site in the proband: Breast [d]								
Breast	1:3[c]	-	310:143	2.6 (2.1-3.2)	213:93	2.7 (2.1-3.4)	92:52	2.2 (1.5-3.1)
Colorectum	73:54	1.5 (1.0-2.1)	82:61	1.5 (1.1-2.2)	109:79	1.6 (1.2-2.1)	120:81	1.7 (1.2-2.2)
Skin	18:2	10.7 (2.4-46.9)	8:8	1.0 (0.4-2.8)	15:2	10.2 (2.3-45.8)	11:8	1.3 (0.5-3.2)
Prostate	59:42	1.6 (1.1-2.4)	-	-	8:5	2.2 (0.7-6.8)	51:35	1.6 (1.0-2.5)
All HLP [e]	57:38	1.7 (1.1-2.6)	36:20	1.8 (1.1-3.2)	55:30	2.0 (1.3-3.2)	34:26	1.4 (0.8-2.4)
Any site	709:663	1.3 (1.2-1.5)	754:528	1.9 (1.6-2.1)	654:536	1.6 (1.4-1.8)	723:590	1.5 (1.4-1.8)
excluding breast	709:663	1.3 (1.2-1.5)	444:385	1.4 (1.2-1.6)	441:443	1.2 (1.0-1.3)	631:538	1.3 (1.2-1.5)
Cancer site in the proband: Endometrium [d]								
Uterus	-	-	19:25	2.0 (1.0-4.0)	18:22	2.1 (1.0-4.2)	4:8	1.6 (0.4-6.5)
Brain	5:5	3.5 (0.9-14.7)	5:2	5.8 (0.8-42.9)	4:3	2.9 (0.5-15.7)	4:4	5.0 (1.0-24.9)
Any site	96:190	1.4 (1.0-2.0)	92:136	1.8 (1.2-2.5)	78:126	2.1 (1.4-3.0)	86:151	1.6 (1.1-2.3)
excluding uterus	96:190	1.4 (1.0-2.0)	73:111	1.6 (1.1-2.3)	65:109	1.9 (1.3-2.8)	82:143	1.5 (1.0-2.1)

...CONTINUED

	Male		Female		<60 years		≥60 years	
FH, cancer site in relatives	ca:co	OR (95% CI)	ca:co	OR (95% CI)	ca:co	OR (95% CI)	ca:co	OR (95% CI)
Cancer site in the proband: Ovary [d]								
Ovary	-	-	27:9	7.4 (3.3-16.6)	17:6	7.3 (2.7-19.6)	10:3	7.5 (1.9-29.9)
Colorectum	29:44	1.6 (0.9-2.7)	33:48	1.6 (1.0-2.7)	54:57	2.1 (1.4-3.2)	50:65	1.8 (1.2-2.7)
Breast	1:6[c]	-	103:106	2.4 (1.8-3.2)	72:67	2.6 (1.8-3.8)	31:47	1.6 (1.0-2.7)
Any site	275:550	1.4 (1.2-1.7)	282:392	2.1 (1.2-1.8)	276:411	1.9 (1.6-2.4)	233:467	1.4 (1.2-1.8)
excluding ovary	275:550	1.4 (1.2-1.7)	255:383	1.8 (1.5-2.2)	259:405	1.8 (1.4-2.2)	223:464	1.3 (1.1-1.6)
Cancer site in the proband: Prostate								
Prostate	90:28	3.9 (2.4-6.2)	-	-	11:4	4.8 (1.3-17.0)	75:23	3.9 (2.3-6.6)
Lung	88:72	1.4 (1.0-2.1)	27:17	1.6 (0.8-3.2)	3:2	1.7 (0.2-13.2)	70:50	1.4 (0.9-2.1)
Ovary	-	-	10:2	7.4 (1.4-38.2)	7:2	6.0 (1.1-33.4)	2:0	∞
Bladder	9:22	2.5 (1.0-5.8)	12:1	11.3 (1.4-90.8)	7:2	2.4 (0.5-12.1)	21:8	3.3 (1.4-7.9)
Kidney	11:4	3.2 (1.0-10.7)	5:2	3.9 (0.6-25.9)	5:3	2.7 (0.5-13.0)	11:3	4.0 (1.0-15.5)
Any site	400:337	1.6 (1.3-2.0)	301:256	1.5 (1.2-2.8)	269:242	1.6 (1.3-2.0)	389:297	1.7 (1.4-2.0)
excluding prostate	310:309	1.2 (1.0-1.5)	301:256	1.5 (1.2-2.8)	258:238	1.4 (1.1-1.8)	314:274	1.3 (1.1-1.6)
Cancer site in the proband: Renal cell								
Kidney	8:5	2.9 (0.9-9.4)	10:3	6.2 (1.7-22.9)	9:4	4.3 (1.3-14.5)	8:2	7.6 (1.6-36.7)
Any site	206:324	1.5 (1.2-1.8)	175:259	1.5 (1.2-1.9)	145:234	1.5 (1.2-1.9)	182:284	1.5 (1.2-1.9)
excluding kidney	198:319	1.4 (1.1-1.7)	165:256	1.4 (1.1-1.8)	136:230	1.4 (1.1-1.8)	174:282	1.4 (1.1-1.8)

Abbreviations: Ca, cases; Co, controls; OR, odds ratio; CI, confidence interval; HLP, hemolymphopoietic; SCCE, squamous cell carcinoma of the esophagus.

[a] Adjusted for age, sex (when appropriate), study centre (when appropriate), year of interview, education, body mass index, alcohol drinking, tobacco smoking, and number of brothers and sisters. Reference category: no family history of the selected cancer. [b] The sum of cases and controls across strata may not add to the total because of some missing values of age at cancer diagnosis in relatives. [c] First-degree male relatives with history of male breast cancer. [d] Further adjusted for menopausal status, age at menopause, oral contraceptive and hormone replacement therapy use, parity. Odds ratios for breast cancer were further adjusted for age at first birth. [e] Including Hodgkin lymphoma, Non Hodgkin lymphoma or unspecified, leukemia and myeloma

Table 5. Odds ratios (OR) of laryngeal cancer and corresponding 95% confidence intervals (CI) according to family history of laryngeal cancer in first degree relatives in strata of selected covariates.

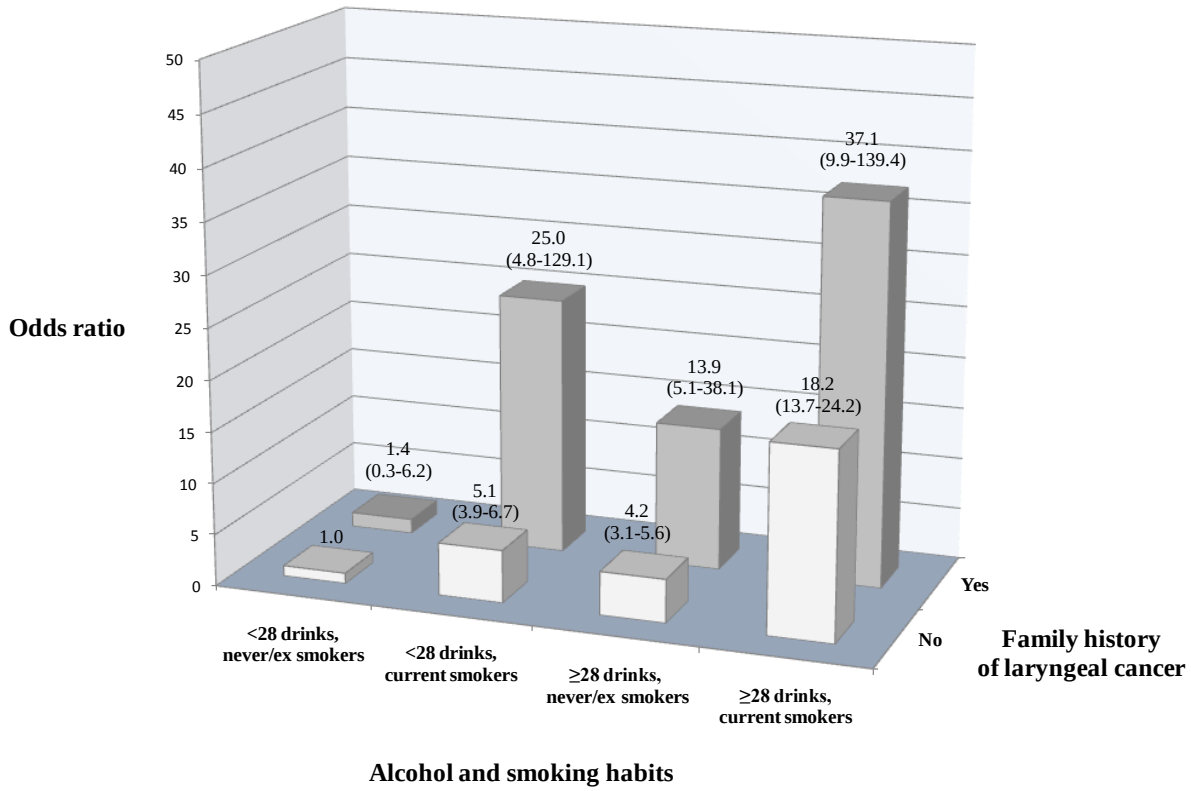
	Family history of laryngeal cancer			
	Cases:Controls		OR[a] (95% CI)	OR[b] (95% CI)
	No	Yes		
Age (years)				
< 60	341:820	14:13	2.8 (1.3-6.7)	3.5 (1.4-8.8)
≥ 60	482:1123	15:14	2.3 (1.1-4.9)	2.3 (1.0-5.5)
Smoking status[c]				
Never/ex smokers	304:1441	12:22	2.8 (1.3-5.9)	2.5 (1.1-5.7)
Current smokers	519:497	17:5	3.3 (1.2-9.0)	3.5 (1.2-10.0)
Alcohol (drinks/week)[c]				
< 28	303:1355	8:17	2.2 (0.9-5.3)	2.3 (0.9-6.4)
≥ 28	514:583	20:10	2.8 (1.3-6.1)	3.0 (1.3-7.0)

[1] Estimates from logistic regression models, adjusted for age, sex and center. Reference category: no family history.

[2] Further adjusted for education, tobacco smoking, alcohol drinking and number of brothers and sisters.

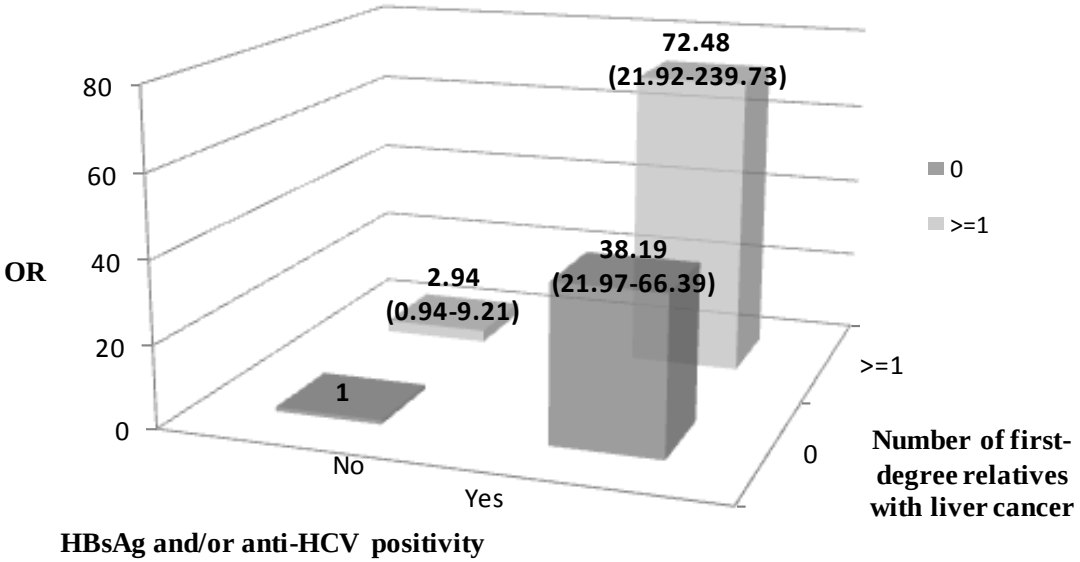
[3] The sum does not add up to the total because of some missing values.

Figure 1. Odds ratios (OR)[a] and 95% confidence intervals (CI) for laryngeal cancer according to alcohol drinking, tobacco smoking and family history of laryngeal cancer in first-degree relatives.



[a] Adjusted for age, sex, study center, education and number of brothers and sisters.

Figure 2. Odds ratios (OR)[a] and 95% confidence intervals (CI) for liver cancer according to HBsAg and/or anti-HCV positivity and family history of liver cancer in first-degree relatives.



Abbreviations: HBsAg, hepatitis B surface antigen; anti-HCV, antibodies against hepatitis C virus.

[a] Adjusted for age, sex, centre, education, alcohol drinking and smoking habits.

Table 6. Main characteristics of the studies on liver cancer and family history of liver cancer included in the meta-analysis

Study	Country	Gender	No. of Cases	No. and type of controls / size of Cohort	Years of study/ duration of follow-up	Adjusting factors and matching variables	Notes
Case-control studies							
Tsukuma et al, 1990 [37]	Japan	Both	229	266 hb	1983-87	Age, sex, alcohol, smoking, history of blood transfusion, HBsAg	
Chen et al, 1991 [38]	Taiwan	Male	200	200 pb	1985-87	Age, sex, ethnic group, residence, alcohol, smoking, HBsAg/HBeAg	95% CI for the adjusted OR was calculated on the distribution of cases and controls according to FH of liver cancer
Tanaka et al, 1992 [36]	Japan	Both	204	410 pb	1985-89	Age, sex	

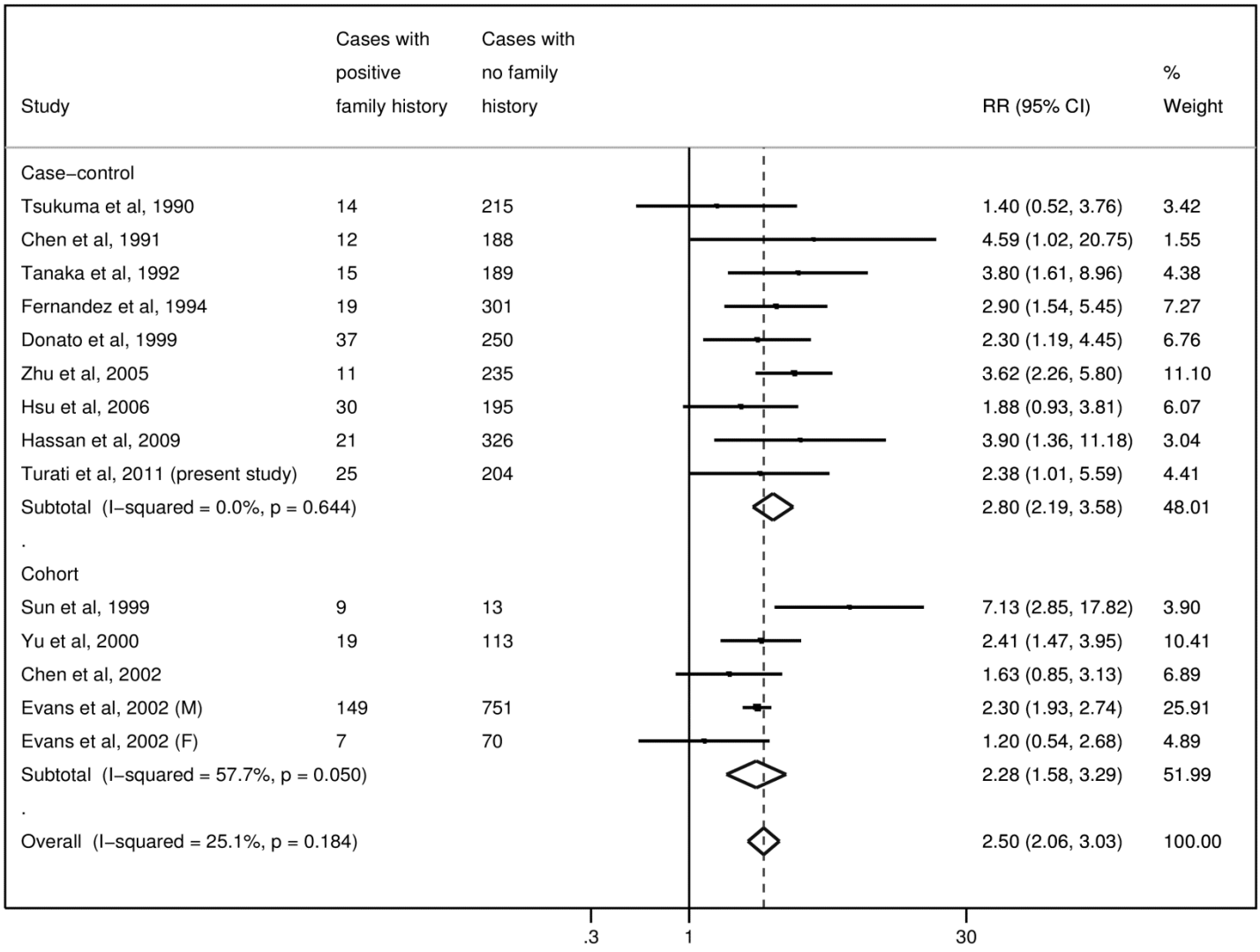
Fernandez et al, 1994 [41]	Italy	Both	320	1408 hb	1983-92	Age, sex, residence, education, alcohol, smoking, history of cirrhosis and hepatitis	
Donato et al, 1999 [40]	Italy	Both	287	450 hb	1996-98	Age, sex, date of admission to hospital, residence, education, alcohol, HCV and HBV infection, FH of all cancer excluding liver	
Zhu et al, 2005 [42]	China	Both	246	549 hb	2001-03	Age, sex	OR and 95% CI calculated on the distribution of cases and controls according to FH of liver cancer
Hsu et al, 2006 [43]	Taiwan	Both	225	225 hb	1999-2001	Age, sex, alcohol, TCR-γ STR genotype 16	Cases were HBV- or HCV-related cirrhotic patients with HCC; controls were HBV- or HCV-related cirrhotic patients without HCC
Hassan et al, 2009 [39]	USA	Both	347	1075 pb	2000-08	Age, sex, ethnic group, education, alcohol, smoking, diabetes, anti-HCV, HBsAg, anti-HBc	

Turati et al, 2011 (present study)	Italy	Both	229	431 hb	1999-2002	age, sex, centre, education, alcohol, smoking, HBsAg and/or anti-HCV	
Cohort studies							
Sun et al, 1999 [46]	China	Male	22	145 pr	1988-98 10 years	Age, aflatoxin, anti-HCV	The cohort consisted of men with chronic HBV hepatitis.
Yu et al, 2000 [44]	Taiwan	Male	132	4808 pr	1988-92 8.9 years (average)	Age, education, alcohol, smoking, number of siblings	The cohort consisted of men HBV carriers. The paper reports also data from a case-control familial study.

Chen et al, 2002 [45]	Taiwan	Both	94	4843 pr	1991-98 7 years (average)	Age, sex, HBsAg, HCV, AFP, AST, ALT	The cohort consisted of subjects selected among members of a screened population who were positive for at least one of 6 investigated risk factors of HCC (i.e., HBsAg, anti-HCV, AFP≥20 ng/mL, AST≥40 IU/L, ALT≥45 IU/L and family history of HCC). I selected ORs for incidence rather than mortality.
Evans et al, 2002 [45]	China	Both	1092	83,794 pr 434,718 py	1992-2000 8 years	Males: age, occupation, alcohol, tea, well water, history of acute hepatitis, HBsAg Females: age, alcohol, smoking, history of acute hepatitis, HBsAg	The paper reports results from a cohort of males and one of females; results were stratified by sex. The outcome was death for HCC

Abbreviations: hb, hospital-based controls; pb, population-based controls; pr, persons at risk; py, person-years; HBsAg, hepatitis B surface antigen; HBeAg, hepatitis B e antigen; CI, confidence interval; OR, odds ratio; HCV, hepatitis C virus; HBV, hepatitis B virus; HCC, hepatocellular carcinoma; anti-HCV, antibodies against hepatitis C virus; anti-HBc, hepatitis B core antibody; AFP, alpha-fetoprotein; AST, aspartate transaminase; ALT, alanine transminase.

Figure 3. Forest plot for the association between family history of liver cancer and liver cancer risk.



Abbreviations: RR, relative risk; CI, confidence intervals; M, males; F, females.

Chapter 2

A new approach: family history scores

2.1 Introduction on family history scores

Estimates of disease RR in families have important utilities in investigations of disease etiology. They are used to examine whether the disease of interest clusters in certain families and whether its etiology has a familial component. They are also used to adjust for familial aggregations when evaluating the effects of other non-familial etiologic factors in epidemiologic studies. Furthermore, familial RR estimates are used to examine effect modification of an etiologic factor according to levels of disease relative risk in families. Finally, a valid assessment of familial RR may have important clinical utility in triaging persons for more involved genetic screening and informing family members about potential risks.

FH scores are used for estimating the familiar risk (FR), that is the level of risk for a particular disease among members of that family. A strong association between FH and the risk of a disease is not a proof of the genetic basis for the disease, since both genetic and environmental factors influenced this risk level. However, these measures can be useful as proxies of a genetic transmission of the disease when there are no genetic tests.

FH scores are created from reports about the disease status of the relatives of a family and can be regarded as an estimate of FR, or a measure of FR that is prone to measurement error.

A good score should have some *desirable properties*:

- *a score should consider the risk profile of a family*, taking into account the risk profile of each relative in terms of relevant covariates (age, sex, smoking status and so on). For example, because early onset is often characteristic of familial cases of disease, such families should have a higher score.
- *a score should be robust to family size and time at risk*, not changing systematically with the number of relatives. Families with the same proportion of relatives affected should have a similar score;
- *a score should consider relationship of relatives,* close blood relatives being given more weight.
- *a score should not be inflated by a single individual; s*mall families are particularly vulnerable to the influence of a single affected member.

Some definitions before to go through the different formulations of the FH scores:

- Disease status of the relative: It was determined by the reported lifetime disease status or cause of death. For the relative *i* (*i*=1,…, n_j , where n_j is the family size of family *j*) in family *j* (*j*=1,…,*n*, where *n* is the number of subjects included in the study), the disease status is referred to o_{ij}, where o_{ij} =1 if affected and 0 if not affected. Thus, o_j represents the number of affected relatives in family *j*.

- Time at risk: For each relative it was defined as age at diagnosis if diseased, current age if not diseased and alive, or age at death if unaffected and deceased. Each individual contributed only with his/her disease-free life experience.

- The expected risk of disease: The expected risk of disease for relative *i* in the family *j* is referred to e_{ij} (thus, e_j represents the total expected number of affected relative in family *j*). Usually, the expected risk for each relative comes from an external source according to strata of important risk factors like age and gender. In particular the expected risk for each family member is obtained by multiplying age-, sex-, and time-specific incident rates for the disease in the area under study (taken from, e.g., regional registries) by age-, sex-, and birth cohort-specific person-years at risk. Person-years at risk, as it was mentioned before, were accumulated until age at interview or age at death for people without the disease or age at diagnosis for people with a history of the disease. In our knowledge, only a FH score, which will be called "FHS2" from this point on, calculates the expected risk of disease from internal data, using a logistic regression model fitted in a dataset in which all the probands' relatives are combined.

2.2 Review of the family history scores

Table 7 shows results from a review of the literature on FH measures, which let me identify 13 different FH scores, including the binary indicator and the number of affected relatives [227-236]. They have been used in heterogeneous settings and for different study designs and topics. Besides the "binary indicator" and the "number of affected relatives", largely used in the medical literature, most of the more complex FH scores have been proposed for common diseases, in particular for coronary heart diseases.

The most common FH score in use is the dichotomous measure, defined to be positive in families that have at least one relative with the disease. Another summary that carries a little more information is the number of affected family members. These crude summaries have two critical deficiencies in view of their use as familial relative risk estimates. First, they do not account for family size, structure, or ages of family members. Larger families and families with older members are naturally more likely to have members who had developed chronic diseases such as cancer. Second, the summaries do not take chance into account: families with identical familial relative risk levels, sizes, structures, and ages can yield different numbers of affected members by chance alone. The proportion of affected relatives is still a simple summary measure with the advantage to taking into account the size of the family (but not its structure in term of age, and sex of its members).

The others FH scores proposed in the literature are statistics that describe the deviation from the expected risk for each family. Briefly, the bigger is the difference between the observed and the expected situation in a family, the bigger is the FH score of that family (i.e., greater is the estimate of the FR), in absolute value. Some FH scores are

aggregated scores (Slack and Evans's score, Fain and Goldar's score, Lynch's score, Reed's score, Kee's score): they use the total expected *(e_j)* and the total observed *(o_j)* number of affected relatives in a family (*j).* Non-aggregated scores (Chakraborty's score, Schwartz's score, Groenelved and Hitzeroh's score, FHS1 and FHS2), instead, link the observed status (o_{ij}) to the expected risk of disease (e_{ij}) for each proband's relative; they are therefore sensitive to which relative has the disease. Non-aggregated scores are more sensitive to the risk profile of the affected relatives, comparing the observed status to the expected risk of each relative separately. In general, an affected relative at low risk of disease contributes a large positive value to the non-aggregate score; conversely, a high-risk affected relative gives only a small contribution. If unaffected, the negative contribution of high risk relatives is greater than that of low-risk relatives. The main deficiency of a non-aggregated score is that it tends to be unstable in small families, i.e. such scores enable a low risk (e.g. young) affected relative to inflate the score when the number of relatives is small.

The expected risk of disease for each family member (or the expected number of affected relatives in a family) is usually estimated by using a set of external reference rates for the disease, on strata of age, sex, and possibly other risk factor of the disease. These scores, therefore, take into account the structure of the family in terms of age, sex, and other factors when possible.

When no affected relative is observed in a family, Slack and Evans's score, comparable to the standardised mortality ratio (SMR), is zero and Lynch's score is equal to -1; Chakraborty's score is always positive defined; FHS1 differs from Schwartz's score in dividing by n_j, thus adjusting for family size.

FHS2 merits a separate discussion: proposed by Silberberg et al [235], together with FHS1, this score does not use external data to estimate the expected risks of disease in

relatives. It estimates them from a logistic regression model, in which all subjects' relatives were combined in a dataset and their disease status (1 if the relative had had the disease under investigation, 0 otherwise) is fitted as a function of their sex, age and other risk factor for the disease. Expected probability of disease estimated by this model represents the expected risk of disease that is then compared to the observed relatives' status. The residuals from this model are then averaged over each family to form FHS2.

2.3 Family history score comparison: Methods

I made a comparison of several FH scores using two different complementary approaches. Firstly, I've used a data-derived approach, applying each identified FH score on the Italian HI-WATE study on colorectal cancer, with the aim of examining the power of different FH scores in predicting colorectal and liver cancer, respectively. In that observational study, cases and controls were asked to report selected information (e.g., age and sex) of both affected and unaffected relatives; this allowed me the calculation of the values of the various FH scores for each study participant. The comparison of the FH scores was also performed using a simulation approach. In this situation the gold standard, i.e., the true disease risk in the family (i.e., the FR), is a known quantity. Thus, the performance of FH scores may be evaluated relative to the true FR. In brief, the research question was: how well does a FH score predict whether a family is at high or low risk? A simple situation in which the level of diseases risk is dichotomous and families may be classified as either high or low risk was assumed.

2.3.1 Data-derived approach

2.3.1.1 The Italian HI-WATE study

HI-WATE data derive from a case-control study of colorectal cancer conducted between 2007 and 2010 in the greater Milan area and in the provinces of Pordenone and Udine, Northern Italy, and in the province of Barcelona, Spain. In the current analysis, only Italian data were analyzed.

Cases were 474 subjects aged 20-85 years with a diagnosis of incident, histologically confirmed colon or rectal cancer and with no previous diagnosis of cancer. They were

identified in major general and teaching hospitals in the study areas (Ospedale Niguarda, Istituto Nazionale dei Tumori, Istituto Europeo di Oncologia, Policlinico di Monza, Ospedale Fatebene Fratelli, Milan; Centro di Riferimento Oncologico, Ospedale S. Maria degli Angeli, Pordenone; S. Maria della Misericordia, Udine). Identification of new cases was achieved through periodical visits to the hospital departments where cases were diagnosed or treated. The local coordinator kept a record of all cases identified, of those who refused participation, and indicated the reasons for refusal of interview or blood sample collection.

Controls were 561 subjects with no previous diagnosis of cancer, randomly chosen among subjects admitted as in-patients or out-patients in the same hospital as cases for acute, non chronic diseases, unrelated to alcohol, tobacco or dietary habits nor to known or potential risk factors for colorectal cancer. Diseases of controls mainly included traumas; minor surgical interventions (i.e., appendectomy, hernia, etc.) and genitourinary, skin, subcutaneous tissue, musculoskeletal, peripheral arterial or venous, ear, eye and mastoid disorders, with a proportion of controls within a specific diagnostic group not over 33% of the overall group. An interviewer contacted the individuals to schedule an appointment, asked for the consent and did the interview and blood collection.

A questionnaire similar to that used in the network of case-control studies was administered to both cases and controls. The FH section was however more detailed than that in the questionnaire of network of case-control studies, and therefore allowed the construction of the various FH scores. Besides the number of brothers, sisters, sons, and daughter, for each FDR (parents, siblings, and children), whether affected or unaffected by cancer, the vital status, current age/age at death, history of cancer (excluding non-melanoma skin cancer), site of cancer, and age at diagnosis were

recorded. A specific section of the questionnaire included detailed information on water consumption and water-related habits in order to evaluate the long-term exposure to various disinfection by-products through ingestion, inhalation and dermal absorption. However, this information was not used in the current analyses.

The study protocols have been submitted to approval by the local ethics committees, and, during the study, national and international directives were followed (deontological code, Helsinki declaration). A signed informed consent to participate were obtained from all study subjects prior participation. An identification code were assigned to all cases and controls at the study entry and databases including information from interview and biological samples contained only the code of the subject. The code assigned to the subject and the personal information was kept on a separate file in a different computer.

Table 8 shows the main characteristic of cases and controls.

2.3.1.2 Family history score calculation

The detailed information of FH of cancer that was collected allowed for the construction of different FH scores. Colorectal cancer risk among families was therefore measured using the following 13 FH scores:

(1) Binary indicator

(2) Number of affected relatives

(3) Proportion of affected relatives

(4) Slack and Evans's score

(5) Chakraborty's score

(6) Fain and Goldar's score

(7) Lynch's score

(8) Reed's score

(9) Schwartz's score

(10) Groenelved and Hitzeroth's score

(11) Kee's score

(12) FHS1

(13) FHS2;

whose formulations and references are reported in Table 5.

For each score, two different methods for the calculation of the expected risk of colorectal cancer in relatives were used: (a) a method which needs internal data only, and (b) a method requiring the incidence rates of the disease under investigation according to age, sex, and possibly other factors involved in the disease aetiology, in the area under study.

In the following, a description of these methods.

(a) Use of internal data only

Methods: This approach for e_{ij} calculation was proposed by Silberberg et al. (13) in 1999, when introducing FHS2. This method does not require external data on the expected risk of disease in the population since e_{ij} are estimated directly from the dataset. This is a one of its advantage, but it could also be a disadvantage since there is no validation versus external data. Briefly, in order to estimate the expected risk of the disease for each relative, his/her disease status (affected, if had the disease; not affected otherwise) was fitted as a function of his/her sex, age, and possibly other potential risk factors for the disease (if available), using a logistic regression model. In this model, age represents the disease-free life experience of each relative, and is defined as age at diagnosis for relatives affected by the disease, current age for those free of the disease

and alive, and age at death for those unaffected by the disease and deceased. For each relative in the dataset, the probability predicted from the model represents his/her the expected risk of disease (for relative *i* in the family *j* is referred to as e_{ij}).

Application: In the Italian HI-WATE study the only information on potential risk factors for colorectal cancer recorded in relatives were age and sex. I therefore combined all the FDR (n=6538) of the study subjects' in a unique dataset; then their disease status (1 if the relative had a history of colorectal cancer, 0 otherwise) was fitted as a function of their age (in continuous) and sex, using a logistic regression model. Age was defined as age at diagnosis for FDR with a history of colorectal cancer (regardless of their vital status at the time of data collection), current age for those without a history of colorectal cancer and alive, and age at death for those unaffected by colorectal cancer and deceased.

(b) Use of cancer registry information

Methods: The most commonly method for calculating e_{ij} presupposes the availability of the incidence rates for the disease under investigation according to age group, sex, time period, and possibly other potential risk factors for the disease (e.g, race and smoking).

For the *i*th family's *j*th member, the expected risk e_{ij} is given by the cumulative risk of the disease under observation:

$$e_{ij} = 1 - \exp(-\sum_{k} \lambda_k * t_{ijk})$$

where λ_k is the external reference rate for the *k*th stratum (age, -sex-, and other factor-defined stratum) and t_{ijk} is the length of time that *i*th family's *j*th member spent under

observation in the *k*yh stratum. Ages and sex of the family members are therefore accounted for in the computation of the expected risk.

Application: From the age- and sex-specific incidence rates of colorectal cancer from cancer registries of Varese 1998 (for study subjects from the great Milan area) and Friuli Venezia Giulia 1998-2002 (for study subjects from Pordenone and Udine), I therefore obtained the expected risk of colorectal cancer for each FDR in my dataset using the formula above.

From the expected risks of colorectal cancer for each FDR, obtained from the two methods presented above (i.e., use of internal data only - use of external data) the value of the 13 FH scores was calculated for each family in the HI-WATE study. Except for the binary indicator, which contemplates only two possible alternatives (Yes/No), study subjects were divided into 3 different categories according to their value on each FH score. Subjects reporting no affected relative were included in the same category labeled "low", since the interest was in the ORs comparing subjects with a positive FH (divided in two or more groups on the basis of FH scores) to those with no FH (i.e., no affected relatives in the family). Subjects with at least one relative affected by colorectal cancer were divided into two categories: the category labeled "intermediate", included subjects with the FH score value lower than the median FH score calculated among controls with a positive score, and the category labeled "high", included subjects with the FH score value greater than it.

2.3.1.3 Statistical analysis

Multiple logistic regression models were used to assess the association between a FH of colorectal cancer and colorectal cancer risk according to different methods to define FH of colorectal cancer, and to examine the goodness-of-fit and the accuracy of these

different methods in predicting the development of colorectal cancer. A series of multiple logistic regression models in which the outcome was the proband's disease status (case/control) and the main explanatory variable was each of the FH score, together with the same set of covariates (age, sex, center and education, BMI, physical activity, tobacco smoking, alcohol drinking, and number of siblings) were fitted. The comparison between the different scores was made using the area under the curve (AUC) of the receiver operating characteristic (ROC) curve, ranging between 0.5 and 1, with higher values indicating better discrimination of cases and controls. The comparison between the AUCs of different FH scores with the AUC of the simple proportion was performed using a nonparametric approach for correlated ROC curves (i.e., model fitted on the same dataset) [237].

2.3.2 Simulation approach

Assuming that FR is dichotomous, data of 1000 families with high risk and 1000 families with low risk were simulated from a hypothetical population. Different scenarios were set: the proportion (π) of families with high risk in the population was chosen to be 0.0001, 0.001, 0.05, 0.10, 0.20, or 0.40 and RR of disease comparing high-risk and low-risk individuals to be 1.2, 2.0, 3.0, 4.0. Then, FH data for each individual were generated: the number of family members was generated by a truncated Poisson distribution with mean 8.5 (average family size in the network of Italian and Switzerland case-control studies), truncated so that each individual had at least two relatives (the parents). Second, age of the family members (<45, 56-55, 56-65, >65) was generated according to a multinomial distribution with probabilities 0.1321, 0.2383, 0.3352 and 0.2944 (these probabilities were derived from the age distribution of controls in the network of Italian and Switzerland case-control studies). For simplicity,

the gender of the subjects was ignored. The expected risks of disease for each family member (e_{ij}) were chosen to be 0.001, 0.01, 0.10 or 0.20 or 0.04, 0.08, 0.12, 0.16 for the four age strata, respectively. Then, assuming that the disease is not correlated with family size (i.e., the disease is not associated with the parental choice to for having additional children) and age stratum, the disease variable for each family member (o_{ij}) was generated according to a Bernoulli distribution. The probabilities of disease depended on age stratum and FR, and were computed as follows:

P(D | FR=low) = $\frac{e_{ij}}{1+\pi(RR-1)}$, where D is the disease.

This comes from the following equation:

$$e_{ij} = P(D \mid FR = low) * P(FR = low) + P(D / FR = high) * P(FR = high) =$$
$$= P(D / FR = low) * (1-\pi) + P(D / FR = high) * \pi$$

Rewriting P(D | FR=high)= P(M | FR=low)*RR

$$e_{ij} = P(D / FR = low) * (1-\pi) + P(D / FR = low) * RR * \pi =$$
$$= P(D / FR = low) * (1-\pi+\pi * RR)$$

And then:

$P(D / FR = low) = \frac{e_{ij}}{1+\pi(RR-1)}$; consequently: P(D | FR=high) = RR* P(D | FR=low)

Simulated data were analyzed using a logistic regression model for FR (dependent variable) on the 13 FH scores described earlier one by one (in continuous). For each setting, 200 simulations were conducted, and the mean and standard deviation of the AUC for each FH score were computed from the results.

2.4 Family history score evaluation: Results

2.4.1 Data derived-approach

(a) Expected risk of colorectal cancer in relatives

Use of internal data only: Among the 6538 cases and controls' FDR, 114 had had a history of colorectal cancer. The main results for the logistic regression model fitted on the relatives' dataset (outcome: relatives' colorectal cancer status; covariates: age (in continuous) and sex (Females/Males)) are reported in Table 9. The logistic regression estimated an OR of colorectal cancer of 1.02 (95% CI, 1.01-1.03) for every 1 year increase in age and an OR of 0.77 (95% CI, 0.53-1.13) for females, as compared to males (i.e., β_{age}=0.021, β_{sex} = -0.258).

The predicted probabilities of colorectal cancer estimated from the regression coefficients are relatively small for each subject, ranging from a minimum of 0.0042 to a maximum of 0.0401, with a median value of 0.0173 and a standard deviation of 0.0067.

Use of external data: When the colorectal cancer incidence rates from registries in Varese and Friuli Venezia Giulia were used, the median value of the expected risk of colorectal cancer for FDR was 0.0167 with a standard deviation of 0.0307 (minimum<0.00001 and maximum=0.1863).

Figure 4 shows the distribution of the expected risk of colorectal cancer in FDR using both methods. When external data were used, the expected risk of disease has a skewed (skewed=1.506) distribution with a long tail to the right of 0, which arises from the small expected values due to the low prevalence of the disease in the population. The distribution is also kurtotic (kurtosis=2.366), with "heavier tails" than for a normal distribution. The distribution of the expected risk of disease when internal data where

used is less asymmetrical (skewed=0.324) with slightly “lighter tails" than for a normal distribution.

(2) FH scores

Summary descriptive measures for the different FH scores (only the 11/13 continuous FH scores) by the two methods for expected risk calculation are presented in Table 10. Some of the FH scores reached values that exceed 100, other had maximum values in the order of some tens, others did not reach 10, and some others don’t exceed 1. For each score, the larger ranges were found when using the external data method, reflecting the distribution of the expected risk of colorectal cancer presented before (see Figure 4). Because of their formulations, some FH scores are always positive (i.e., proportion, Slack and Evans’ score, Chakraborty’s score), some others assume also negative values (i.e., Fain and Goldar’s score, Lynch’s score, Reed’s score, Schwartz’s score, Groenveld and Hitzeroth’s score, Kee’s score, FHS1 and FHS2). The form distribution is similar for each FH score, with the great part of the subjects (substantially those with no FH) characterized by the lowest scores and a small part of them (substantially those with a positive FH) characterized by relatively high values as compared to the others.

Table 11 reports the joint distribution of the case-control study participants according to the FH scores categories and the observed number of FDR with colorectal cancer for both methods for expected risk calculation. For each score, the “low” category was constructed so that it includes subjects with no FH. With both methods (internal / external data), Slack and Evans’ score, Fain and Goldar’s score, Lynch’s score and Reed’s score classified study subjects in the same manner. When using the method of internal data only, the 10 study subjects reporting two or more relatives affected by colorectal cancer were classified in the “high” category for each FH score, with the only

exception of the Kee's score, which classified 8 of these subjects in the "high" category and 2 in the "intermediate" one. When using the method requiring incidence rates of colorectal cancer, 6 scores (i.e., Chakraborty's, Schwartz's, Groenveld and Hitzeroth's, and Kee's score, FHS1 and FHS2) classified 9/10 subjects with two or more affected relatives in their "high" category and 1/10 of those subjects in their "intermediate" one. Chakraborty's, Schwartz's, and Groenveld and Hitzeroth's scores, FHS1 and FHS2 (and not Kee's score) classified these 10 subjects in the same manner (the subject in the "intermediate" category and those in the "high" category were the same).

(1) FH scores' comparison

Table 12 presents selected characteristics of the subjects' families according to different FH scores, when using both methods for expected risk calculation. Briefly, the average family size of subjects with no FH was 6.3 (sd=2.5), the average age of the FDR was 59.7 (sd=9.1) and the average expected number of colorectal cancer cases in the family was 0.110 (sd=0.046) when using the method of internal data and 0.168 when using that of external data. The corresponding values for subjects with a positive FH of colorectal cancer were 6.7 (sd=2.9), 60.8 (sd=8.0) and 0.117 (sd=0.052) when using the method of internal data and 0.168 when using that of external data, respectively. Families with 2+ affected relatives were somewhat bigger and older than those with 1 affected relative.

When we considered the two categories of the FH scores in which subjects with a positive FH were divided (i.e., the "intermediate" and the "high" categories), we found that, for most scores, families in the high categories were smaller, characterized by a lower average age and by a lower average number of expected colorectal cancer cases than those in the intermediate ones. These results were similar for both methods for the expected risk calculation in FDR.

In Table 13, the effect of using different methods of defining a positive FH on the odds of colorectal cancer is compared, as well as the goodness-of-fit and accuracy of these methods in predicting the development of such cancer. The analysis was performed for FH scores obtained with both the method of internal data and that of external data. Slack and Evans's, Fain and Goldar's, Lynch's, and Reed's scores gave the same information in terms of odds of colorectal cancer and goodness-of-fit, since they classified subjects in the same manner. Using the binary indicator (Yes *vs* No) resulted in an OR of 1.55 (95% CI, 1.02-2.37) for subjects with one or more FDR with colorectal cancer as compared to those with no affected relatives. Other FH scores divided subjects with at least one affected relative in an "intermediate" category and a "high" category. For most of them, in both scenarios (internal data and external data), the magnitude of the ORs increased with subsequent categories of the scores, although the differences in ORs were not so noticeable. However, statistical significance at α=0.05 was not reached for most of the associations. On the contrary some FH scores found no increased risk (or a slight not significant risk) for the intermediate category, which included subjects with a positive FH but characterized by a relatively low value of the score, and a significant increased risk of colorectal cancer for the high category, in which subjects with a positive FH and characterized by a relatively high value of the score were included, as compared to no FH. In particular, in the internal data setting, FHS1 found a slightly 30% not significant increased risk of colorectal cancer for subjects in the intermediate category and a 80% significant increased risk for those in the high category. Similar results were observed for the Groenveld and Hitzeroth's score in the external data setting. FHS2 fund ORs of 1.17 (95% CI, 0.62-2.12) for the intermediated category and 1.92 (95% CI, 1.09-3.38) for the high category as compared to no FH of colorectal cancer. Only the Chakraborty's score in the internal data setting

and the Kee's score and the Slack and Evans's/Fain and Goldar's/Lynch's/Reed's scores found an higher OR for the intermediate category than for the high one.

With regard to the goodness-of-fit of the model, similar AUC were found for each FH score in both settings; no significant differences in AUC between each FH scores and the proportion emerged.

2.4.2 Simulation

Results for the simulations performed on various scenarios of π and RR, with expected risks of disease equal to 0.001, 0.01, 0.10 and 0.20 (Panel A) or 0.04, 0.08, 0.12, and 0.16 (Panel B) according to the 4 age groups are reported in Table 14. Table 15 shows FH scores' final ranking according to mean AUC. Based on mean AUC observation, Reed's score appears the most efficient FH measure in all the 48 simulation settings, with mean AUC ranging from 0.551 to 0.944, while the binary indicator is the poorest all the scenarios, with mean AUC ranging from 0.534 to 0.739.

FHS2 ranks second in 20 of the 48 simulations, third in 27/48 simulations, and fourth in one simulation only. However, this score, as it was mentioned before, is different from the others in its original formulation, since it calculates the expected risk of disease in relatives using information (i.e., disease status, the disease-free life experience, sex, and other factors) collected among these relatives, instead of using incidence rate of the disease in the area under investigation. In my simulation, in order to make FH scores comparable, I've used the same expected risk of disease in age strata for each score; thus results from the simulation do not take into account the differences between FHS2 and the other FH scores in terms of the calculation of the expected risk of disease.

Schwartz's score ranks third, with in general better performances with higher π when compared with other scores.

Slacnck and Evans' and Lynch scores have very similar performances in all the simulations, and rank fourth, followed by FHS1, Groenelved and Hitzeroth's score, proportion of affected relatives, and Fain and Goldar's score. Besides the binary indicator, Chakraborty's score (twelfth in the final ranking), Kee's score (eleventh in the final ranking) and the observed number of affected relatives (tenth in the final ranking) give the worst prediction of the level of risk of the families.

Figure 5 and Figure 6 show how the performance of the different FH scores varies according to RR (Figure 5) and prevalence of high risk family in the population (Figure 6), with expected risks of disease equal to 0.001, 0.01, 0.10 and 0.20 (Panel A) or 0.04, 0.08, 0.12, 0.16 (Panel B) according to the 4 subsequent age groups. The following observation can be made from the results displayed:

(a) except for the binary indicator, the performance of all the other FH scores decreases with the increase of the prevalence of families with high risk. This is particularly evident when the RR of FH is high (RR=3.0 or 4.0).

(b) In general, the performance of all the FH score increases with the increase of the RR, regardless of the prevalence of families with high risk. However, the expected risks of diseases for the 4 age groups were set to 0.002, 0.01, 0.1, and 0.2 and π=0.0001, the mean AUC of most of the FH scores slightly decreases when RR increases from 3.0 to 4.0 (the mean AUC from all the FH scores at RR=3.0 is 0.898, while the mean AUC at RR=3.0 is 0.896).

(c) The RR for FH contributes to the FH score performance with a higher extent than the prevalence of families with high risk.

If we focus our attention on the proportion, which is the FH summary measure which performs better among the simplest ones (i.e., among those FH measures which do not require the estimate of the expected risk of disease for each relative; i.e., the binary indicator and the number of affected relatives), we found relatively small differences in mean AUC with the better scores (Table 16). In particular, the difference in mean AUC between the proportion and the Reed's score ranged from 0.003 (all the simulations with RR=1.2 when the expected risk of diseases were set to 0.04, 0.08, 0.12, and 0,16) to 0.030 (simulation with π=0.4 and RR=4 when the expected risk of diseases were set to 0.001, 0.01, 0.10 and 0.20), with a mean difference of 0.014. Differences between the proportion and the other scores are lower when the expected risks of disease were 0.04, 0.08, 0.12, and 0,16 for the 4 subsequent age strata than when they were set to 0.001, 0.01, 0.10 and 0.20. The gap in mean AUC between the proportion and the other scores increases with the increase of the RR.

2.5 Discussion

The first preliminary approach in order to compare several different FH scores for measuring FR and recommend the measure that performs best, was to exploit real data collected from a well-conducted Italian case-control study on colorectal cancer. The values of 13 different FH scores were obtained for each study's subjects on the basis of FH information on FDR. Two methods for the calculation of the expected risk of colorectal cancer in FDR were used: a method based on internal data only, as proposed by Silberberg et al [235], and a method using information on incidence rates of colorectal cancer in the area where enrolled subjects come from. Although ranges were different, the majority of the continuous FH scores ranked subjects very similarly. For most of the scores, for both methods for expected risk calculation, the magnitude of the ORs increased with subsequent categories of the scores, although significance at α=0.05 was not reached for most of the associations. On the other hand, FHS1 (when the internal data method for expected risk calculation was used), and the Groenveld and Hitzeroth's score and FHS2 (when the external data method was used) seemed to be able to identify subjects at "intermediate" and at "high" risk of HCC among those with a positive FH. They have been shown to be able of identify subjects with positive FH of colorectal cancer who may have no increased risk of colorectal cancer, since coming from larger and older families, in which the expected risk of disease is higher. However, these FH scores did not provide a higher degree of predictive accuracy of colorectal cancer incidence than simply predicting that risk by designating FH as either positive or negative or using the proportion of affected relatives in a family. In fact, differences in the goodness-of-fit are small and not meaningful, and therefore seem to justify the use

of simplest FH scores, i.e., the binary indicator or the proportion of affected relatives, for the construction of which limited information on FDR is needed.

Some considerations must be done. Colorectal cancer is a rare disease and the association with FH of the same disease is probably modest; in our dataset a small number of subjects reported therefore a positive FH of colorectal cancer, and only the 10% of these (i.e., 6 cases and 4 controls) had 2 or more affected relatives. Consequently, the "intermediate" and the "high" categories of the FH scores included a small number of cases and controls, and this resulted in OR estimates characterized by wide CIs. We calculated FH scores on the basis of information about FDR only. Presumably, the possible effects of family structure and age on familial risk of colorectal cancer would be greater among extended families (i.e., considering also second-degree families) than among families defined on the basis of a first-degree relationship. Moreover, the number of colorectal cancer cases would be greater among extended families. The value of using FH scores to assess the impact of a positive FH on colorectal cancer risk should increase with a greater number of relatives and with the complexity of the pedigrees. Another point is about the calculation of the expected risks of disease in relatives that was performed in two ways: (a) directly from the data by a logistic model, without any validation versus external data, and (b) using age- and sex-specific incidence rates of colorectal cancer from Varese and Friuli registers. A better estimation of the expected risks of colorectal cancer in relatives would be obtained if (a) information on other colorectal cancer risk factor were collected in relatives or (b) incidence rate of the disease in strata of other colorectal cancer risk factors were available. In light of these considerations, caution in interpreting or generalizing these results obtained by the Italian HI-WATE study is needed. Our consideration are only indicative for situation similar to our own, in which the outcome is a disease with

relatively low incidence rates and characterized by a positive modest association with FH of the same disease.

Given the problems in comparing FH scores in a setting characterized by a rare outcome weakly associated with the FR, this issue has been also addressed through a more systematic approach. Different FH scores were compared using simulations of various settings in terms of prevalence of high risk families and RR of disease comparing high-risk and low-risk individuals, in order to evaluate their accuracy of predicting the true FR. The advantage of this approach, as compared to the data-derived approach discussed before, is that scores may be evaluated on the basis of a known quantity, which is chosen at the beginning and that FH scores should estimate. It has been found that the performance of FH scores decreases with the increase of the prevalence of families with high risk, particularly with relatively high RRs. Moreover, in general, the performance increases with the increase of the RR of FH, regardless of the prevalence of families with high risk. More importantly, differences in predicting the true FR between FH scores were minor, although Reed's score and FHS2 performed slightly better than most of the other scores, followed by the Schwartz's score and the Slanck and Evans's score. The binary indicator was the worst predictor of the true FR, while the simple proportion of affected relatives, although ranking 8th in the final ranking, was characterized by small differences in mean AUC with the Reed's score, and with the other scores with a better predictivity. These gaps in mean AUC between the proportion and the other scores increased with the increase of the RR and with the increase of the prevalence of high risk families, although to a lesser extent. Consequently, the use of the simple proportion seems justified, at least until stronger evidence is brought for the advantages of using a more complex score.

2.6 Tables and Figures of Chapter 2

Table 7. Family history score.

Score [year] (reference)	**Score formulation**
Binary indicator	Positive when $o_j \geq 1$
Number of affected relatives	o_j
Proportion of affected relatives [1994] [227]	$p_j = \frac{o_j}{n_j}$
Slack and Evans's score [1996] [228]	$\frac{o_j}{e_j}$
Chakraborty's score [1984] [230]	$\sum_{i=1}^{n_j} \left(o_{ij} - e_{ij}\right)^2 / \left(e_{ij}(1-e_{ij})\right)$
Fain and Goldar's score [1986] [231]	$\frac{o_j - e_j}{\sqrt{e_j(1-e_j)}}$
Lynch's score [1986] [229]	$\frac{o_j - e_j}{e_j}$
Reed's score [1986] [236]	$\frac{o_j - e_j}{\sqrt{e_j}}$
Schwartz's score [1988] [234]	$\sum_{i=1}^{n_j} \left(o_{ij} - e_{ij}\right) / \sqrt{e_{ij}}$
Groenelved and Hitzeroth's score [1991] [232]	$\frac{1}{n_j} \sum_{i=1}^{n_j} \left(o_{ij} - e_{ij}\right) \ln\left[\left(1-e_{ij}\right)/e_{ij}\right]$
Kee's score [1991] [233]	$\frac{o_j - e_j}{e_j \sqrt{e_j}}$
Silberberg's scores [1999] [235]	$FHS\_1 = \frac{1}{n_j} \sum_{i=1}^{n_j} \left(o_{ij} - e_{ij}\right) / \sqrt{e_{ij}}$ $FHS\_2 = \frac{1}{n_j} \sum_{i=1}^{n_j} \left(o_{ij} - \hat{e}_{ij}\right)$

Table 8. Distribution of 474 cases of colorectal cancer and 561 controls according to center, sex, age, and other selected variables (Italian HI-WATE study).

	Cases		Controls		Total	
	N	%	N	%	N	%
Centre						
Pordenone/Udine	240	(50.6)	247	(44.0)	487	(47.1)
Milano	234	(49.4)	314	(56.0)	548	(52.9)
Age (yrs)						
<60	116	(24.5)	149	(26.6)	265	(25.6)
60-64	85	(17.9)	97	(17.3)	182	(17.6)
65-69	110	(23.2)	126	(22.5)	236	(22.8)
70-74	89	(18.8)	109	(19.4)	198	(19.1)
≥75	74	(15.6)	80	(14.3)	154	(14.9)
Sex						
Male	310	(65.4)	366	(65.2)	676	(65.3)
Female	164	(34.6)	195	(34.8)	359	(34.7)
Education (years)						
<7	161	(34.0)	225	(40.1)	386	(37.3)
7-11	154	(32.5)	164	(29.2)	318	(30.7)
≥12	159	(33.5)	172	(30.7)	331	(32.0)
Body mass index (kg/m^2) [a]						
<20	19	(4.0)	26	(4.6)	45	(4.3)
20-25	174	(36.7)	179	(31.9)	353	(34.1)
25-30	188	(39.7)	249	(44.4)	437	(42.2)
>30	92	(19.4)	103	(18.4)	195	(18.8)
Physical activity [a]						
Hig	91	(19.2)	132	(23.5)	223	(21.5)
Medium	119	(25.1)	115	(20.5)	234	(22.6)
Low	263	(55.5)	307	(54.7)	570	(55.1)
Alcohol (drinks/week) [a]						
≤7	150	(31.6)	225	(40.1)	375	(36.2)
8-21	161	(34.0)	198	(35.3)	359	(34.7)
>22	162	(34.2)	138	(24.6)	300	(29.0)
Smoking habits [a,b]						
Never	191	(40.3)	235	(41.9)	426	(41.2)
Ex	192	(40.5)	216	(38.5)	408	(39.4)
Current,<15 cig/day	40	(8.4)	36	(6.4)	76	(7.3)
Current, ≥15 cig/day	51	(10.8)	73	(13.0)	124	(12.0)
Number of siblings						
0	53	(11.2)	56	(10.0)	109	(10.5)
1	119	(25.1)	145	(25.8)	264	(25.5)
2	90	(19.0)	109	(19.4)	199	(19.2)
3	76	(16.0)	81	(14.4)	157	(15.2)
≥4	136	(28.7)	170	(30.3)	306	(29.6)

[a] The sum does not add up to the total because of some missing values. [b] Cigar equals 3 cigarettes, 1 g of pipe tobacco equals 1 cigarette.

Table 9. Results for the logistic regression model fitting the relative's colorectal cancer status as a function of their age (in continuous) and sex.

Analysis of Maximum Likelihood Estimates				
Parameter	DF	Estimate	SE	Wald Chi-square
Intercept	1	-4.973	0.434	131.081
sex	1	-0.258	0.192	1.811
age	1	0.021	0.005	15.195

	Odds Ratio Estimates		
Effect	Point Estimate	95% Wald Confidence Limits	
sex	0.773	0.531	1.125
age	1.021	1.010	1.032

Figure 4. Distribution of expected risk of colorectal cancer in first-degree relatives using internal data only (in pink) and external data (i.e., cancer incidence rates from cancer registries; in blue).

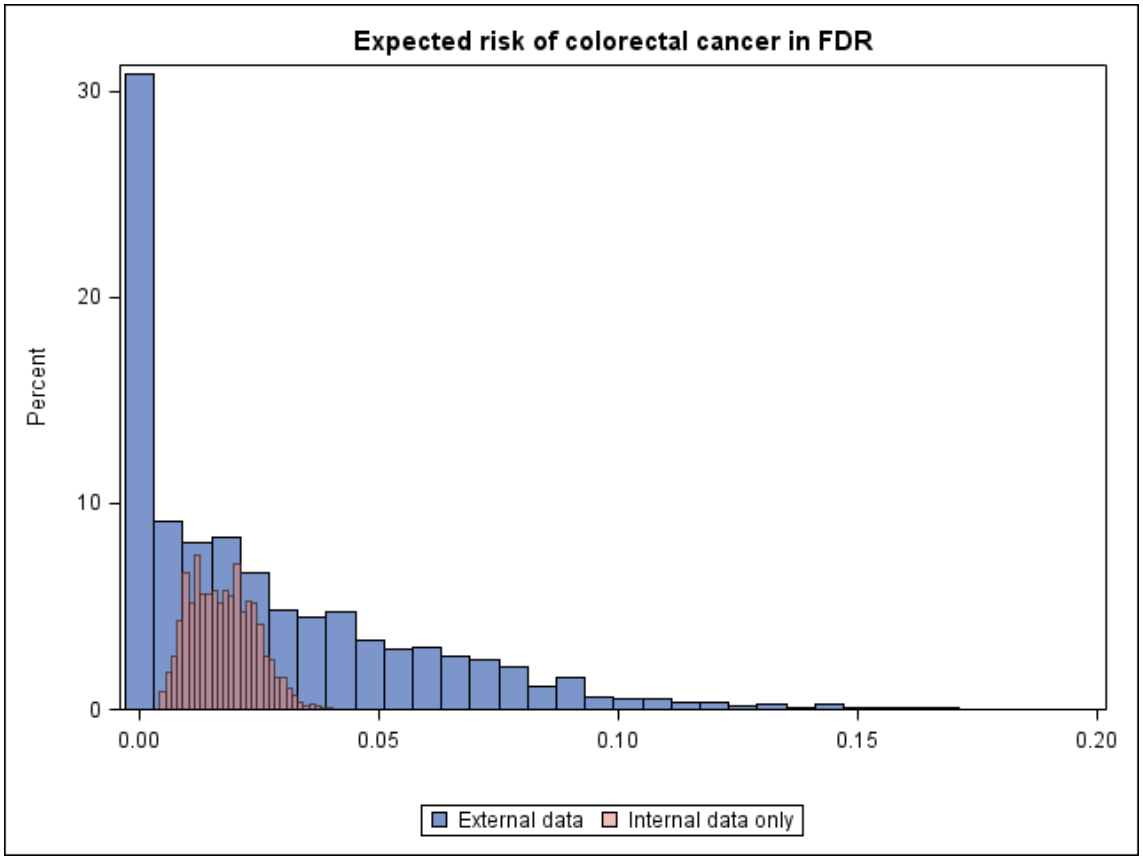


Table 10. Summary measures describing the distribution of the different family history scores by the two methods for expected risk in FDR calculation. Only the distribution of the 11/13 continuous FH scores was presented.

	Internal data only					External data				
FH score	mean	median	sd	min	max	mean	median	sd	min	max
Proportion	0.02	0.00	0.07	0.00	0.50	0.02	0.00	0.07	0.00	0.50
Slack and Evans	1.11	0.00	3.70	0.00	30.69	0.89	0.00	3.54	0.00	59.63
Chakraborty	6.14	0.11	20.38	0.03	231.52	9.68E+13	0.16901	3.11E+15	0.0082	1.00E+17
Fain and Goldar	0.02	-0.32	1.15	-0.63	7.08	-0.12	-0.40	1.07	-1.83	7.66
Lynch	0.11	-1.00	3.70	-1.00	29.69	-0.11	-1.00	3.54	-1.00	58.63
Reed	0.02	-0.31	1.09	-0.53	6.79	-0.10	-0.37	0.99	-0.88	7.59
Schwartz	-0.01	-0.71	2.62	-1.97	24.74	306125.71	-0.68	9838974.58	-2.79	316227765.51
Groenelved and Hitzeroth	0.01	-0.07	0.26	-0.10	1.98	-0.01	-0.07	0.35	-0.22	7.77
Kee	0.54	-3.08	13.75	-6.09	164.51	0.20	-2.49	18.48	-11.10	452.73
FHS1	0.01	-0.13	0.49	-0.17	3.76	61225.15	-0.13	1967794.92	-0.33	63245553.10
FHS2	0.00	-0.02	0.07	-0.03	0.48	-0.01	-0.02	0.07	-0.11	0.49

Table 11. Joint distribution of the Italian HI-WATE case-control study participants according to the family history (FH) scores categories and the observed number of first-degree relatives (FDR) with colorectal cancer, for both methods for expected risk calculation.

	INTERNAL DATA N. of affected FDR			**EXTERNAL DATA N. of affected FDR**		
FH score	**0**	**1**	**2+**	**0**	**1**	**2+**
Proportion						
Low	930	0	0	930	0	0
Intermediate	0	54	0	0	54	0
High	0	39	10	0	39	10
Slack and Evans, Fain and Goldar, Lynch, Reed						
Low	930	0	0	930	0	0
Intermediate	0	48	0	0	52	0
High	0	45	10	0	41	10
Chakraborty[a]						
Low	930	0	0	930	0	0
Intermediate	0	51	0	0	47	1
High	0	42	10	0	46	9
Schwartz[a]						
Low	930	0	0	930	0	0
Intermediate	0	47	0	0	47	1
High	0	46	10	0	46	9
Groenelved and Hitzeroth[b]						
Low	930	0	0	930	0	0
Intermediate	0	48	0	0	44	1
High	0	45	10	0	49	9
Kee						
Low	930	0	0	930	0	0
Intermediate	0	47	2	0	52	1
High	0	46	8	0	41	9
FHS1[b]						
Low	930	0	0	930	0	0
Intermediate	0	46	0	0	44	1
High	0	47	10	0	49	9
FHS2						
Low	930	0	0	930	0	0
Intermediate	0	48	0	0	43	1
High	0	45	10	0	50	9

[a] The classification of the study subjects by the Chakraborty's score is different from that by the Schwartz's score.

[b] The classification of the study subjects by the Groenelved and Hitzeroth's score is different from that by the FHS1.

Table 12. Characteristics of the study subjects' families according to different family history scores.

	INTERNAL DATA						
	N. of families	**Average family size**	**SD**	**Average age of the family**	**SD**	**Average expected risk of disease in a family**	**SD**
No family history	930	6.3	2.5	59.7	9.1	0.110	0.046
Binary indicator							
Yes	103	6.7	2.9	60.8	8.0	0.117	0.052
N. of affected relatives							
1	93	6.6	2.8	60.6	8.2	0.116	0.053
2+	10	7.3	3.1	63.3	5.9	0.128	0.047
Proportion							
Intermediate	54	8.4	2.4	59.9	6.7	0.147	0.048
High	49	4.8	2.1	61.8	9.1	0.084	0.033
Slanck and Evan, Fain and Goldar, Lynch, Reed							
Intermediate	48	8.6	2.4	61.1	5.8	0.154	0.046
High	55	4.9	2.0	60.6	9.6	0.085	0.031
Chakraborty							
Intermediate	51	6.2	2.8	61.6	8.9	0.113	0.056
High	52	7.1	2.8	60.0	7.0	0.121	0.048
Swartz							
Intermediate	47	7.1	3.2	61.4	8.3	0.129	0.061
High	56	6.3	2.5	60.3	7.7	0.107	0.042
Groenelved and Hitzeroth							
Intermediate	48	8.6	2.4	59.9	6.9	0.152	0.048
High	55	4.9	2.0	61.6	8.8	0.086	0.032
Kee							
Intermediate	49	8.8	2.5	61.2	5.5	0.157	0.046
High	54	4.7	1.3	60.5	9.7	0.081	0.020
FHS1							
Intermediate	46	8.7	2.5	59.4	7.0	0.152	0.050
High	57	5.1	2.0	62.0	8.6	0.089	0.033
FHS2							
Intermediate	48	8.6	2.4	60.9	6.4	0.154	0.047
High	55	4.9	2.0	60.8	9.2	0.085	0.031

	EXTERNAL DATA						
	N. of families	**Average family size**	**SD**	**Average age of the family**	**SD**	**Average expected risk of disease in a family**	**SD**
No family history	930	6.3	2.5	59.7	9.1	0.168	0.101
Binary indicator							
Yes	103	6.7	2.9	60.8	8.0	0.171	0.099
N. of affected relatives							
1	93	6.6	2.8	60.6	8.2	0.172	0.103
2+	10	7.3	3.1	63.3	5.9	0.165	0.050
Proportion							
Intermediate	54	8.4	2.4	59.9	6.7	0.218	0.107
High	49	4.8	2.1	61.8	9.1	0.120	0.054
Slanck and Evan, Fain and Goldar, Lynch, Reed							
Intermediate	52	7.9	2.9	62.2	7.4	0.232	0.100
High	51	5.4	2.2	59.4	8.4	0.109	0.044
Chakraborty							
Intermediate	48	6.3	3.0	61.8	8.1	0.173	0.112
High	55	7.0	2.7	60.0	7.9	0.170	0.086
Swartz							
Intermediate	48	6.6	3.0	61.6	7.8	0.182	0.114
High	55	6.7	2.7	60.1	8.1	0.162	0.083
Groenelved and Hitzeroth							
Intermediate	45	8.2	3.0	60.1	6.6	0.215	0.118
High	58	5.4	2.0	61.4	8.9	0.137	0.063
Kee							
Intermediate	53	7.8	2.9	62.5	7.5	0.232	0.099
High	50	5.5	2.3	59.1	8.2	0.107	0.041
FHS1							
Intermediate	45	8.0	3.1	60.1	6.6	0.212	0.120
High	58	5.6	2.2	61.3	9.0	0.140	0.063
FHS2							
Intermediate	44	8.8	2.8	61.4	5.8	0.229	0.116
High	59	5.0	1.6	60.4	9.3	0.129	0.053

Table 13. Summary of estimates from logistic regression analysis by different family history (FH) measures, using both methods for the expected risk calculation.

FH score	INTERNAL DATA OR[a] (95% CI)	INTERNAL DATA AUC (p[b])	EXTERNAL DATA OR[a] (95% CI)	EXTERNAL DATA AUC (p[b])
Binary indicator				
Yes	1.55 (1.02-2.37)	0.629 *(0.771)*	1.55 (1.02-2.37)	0.629 (*0.711)*
Number of affected relatives				
1	1.54 (0.99-2.41)	0.629	1.54 (0.99-2.41)	0.629
2+	1.63 (0.45-6.00)	*(0.980)*	1.63 (0.45-6.00)	*(0.980)*
Proportion				
Intermediate	1.45 (0.81-2.59)	0.629	1.45 (0.81-2.59)	0.629
High	1.67 (0.91-3.07)	-	1.67 (0.91-3.07)	-
Slack and Evans, Fain and Goldar, Lynch, and Reed				
Intermediate	1.38 (0.75-2.55)	0.628	1.58 (0.88-2.85)	0.629
High	1.72 (0.96-3.06)	*(0.661)*	1.52 (0.84-2.74)	*(0.821)*
Chakraborty				
Intermediate	1.68 (0.93-3.02)	0.630	1.47 (0.81-2.68)	0.628
High	1.44 (0.80-2.57)	*(0.294)*	1.63 (0.92-2.88)	*(0.723)*
Schwartz				
Intermediate	1.44 (0.79-2.63)	0.628	1.52 (0.84-2.76)	0.629
High	1.66 (0.94-2.93)	*(0.748)*	1.58 (0.89-2.80)	*(0.958)*
Groenelved and Hitzeroth				
Intermediate	1.41 (0.77-2.60)	0.629	1.33 (0.72-2.48)	0.628
High	1.69 (0.95-3.00)	*(0.733)*	1.76 (1.00-3.08)	*(0.463)*
Kee				
Intermediate	1.44 (0.78-2.64)	*0.629*	1.64 (0.92-2.95)	0.630
High	1.67 (0.93-2.98)	*(0.867)*	1.46 (0.81-2.65)	*(0.517)*
FHS-1				
Intermediate	1.31 (0.70-2.43)	0.628	1.35 (0.73-2.51)	0.627
High	1.79 (1.01-3.16)	*(0.781)*	1.74 (0.99-3.05)	*(0.304)*
FHS-2				
Intermediate	1.39 (0.76-2.57)	0.629	1.17 (0.62-2.21)	0.629
High	1.71 (0.96-3.04)	*(0.724)*	1.92 (1.09-3.38)	*(0.837)*

[a] Estimated from unconditional multiple logistic regression model adjusted for age, sex, center and education, body mass index, physical activity, tobacco smoking, alcohol drinking, and number of siblings. Reference category: no family history of colorectal cancer.

[b] p value for nonparametric tests comparing AUCs for models including continuous FH scores with AUC for model including the proportion indicator.

Table 14. AUC for 200 simulations with expected risk of diseases equal to 0.001, 0.01, 0.10 and 0.20 (Panel A) or 0.04, 0.08, 0.12, 0.16 (Panel B) for the four age strata (i.e., <45, 45-55, 56-65, >56), respectively, according to different prevalences of families at high risk in the population (i.e., 1‰, 1 ‰, 5%, 10%, 20% or 40%) and to different relative risks (RR) for high vs low risk families (i.e., 1.2, 2.0, 3.0 or 4.0). The four highest values of mean AUC were underline in orange; the four lowest values in blue.

PANEL A

	Prevalence=0.0001							
	RR=1.2		RR=2.0		RR=3.0		RR=4.0	
FH SCORE	**mean (AUC)**	**std (AUC)**	**mean (AUC)**	**std (AUC)**	**mean (AUC)**	**std (AUC)**	**mean (AUC)**	**std (AUC)**
Binary indicator	0.534	0.008	0.625	0.009	0.604	0.010	0.705	0.012
Number of affected	0.545	0.006	0.690	0.008	0.895	0.009	0.890	0.008
FHS1	0.549	0.006	0.707	0.008	0.934	0.007	0.914	0.007
FHS2	0.550	0.006	0.716	0.008	0.950	0.006	0.931	0.006
Proportion	0.547	0.007	0.702	0.008	0.936	0.006	0.913	0.007
Slanck and Evans	0.547	0.006	0.706	0.008	0.950	0.006	0.927	0.007
Groenelved and Hitzeroth	0.548	0.006	0.706	0.008	0.922	0.007	0.915	0.007
Fain and Goldar	0.546	0.011	0.697	0.011	0.869	0.019	0.907	0.009
Kee	0.540	0.016	0.690	0.009	0.938	0.007	0.910	0.008
Lynch	0.547	0.006	0.706	0.008	0.950	0.006	0.927	0.007
Reed	0.552	0.005	0.721	0.008	0.953	0.005	0.937	0.006
Chakraborty	0.537	0.008	0.659	0.009	0.843	0.011	0.845	0.010
Schwartz	0.551	0.005	0.713	0.008	0.930	0.007	0.914	0.007

	Prevalence=0.001							
	RR=1.2		RR=2.0		RR=3.0		RR=4.0	
FH SCORE	**mean (AUC)**	**std (AUC)**	**mean (AUC)**	**std (AUC)**	**mean (AUC)**	**std (AUC)**	**mean (AUC)**	**std (AUC)**
Binary indicator	0.534	0.008	0.625	0.009	0.680	0.012	0.705	0.011
Number of affected	0.545	0.006	0.690	0.008	0.814	0.010	0.889	0.008
FHS1	0.549	0.006	0.707	0.008	0.838	0.009	0.914	0.007
FHS2	0.550	0.006	0.715	0.008	0.853	0.009	0.930	0.006
Proportion	0.547	0.007	0.702	0.008	0.834	0.009	0.913	0.007
Slanck and Evans	0.547	0.006	0.706	0.008	0.845	0.009	0.927	0.007
Groenelved and Hitzeroth	0.548	0.006	0.706	0.008	0.838	0.009	0.915	0.007
Fain and Goldar	0.546	0.011	0.697	0.011	0.828	0.012	0.907	0.009
Kee	0.540	0.016	0.690	0.009	0.825	0.009	0.909	0.008
Lynch	0.547	0.006	0.706	0.008	0.845	0.009	0.927	0.007
Reed	0.552	0.005	0.721	0.008	0.859	0.008	0.936	0.006
Chakraborty	0.537	0.008	0.659	0.009	0.770	0.010	0.845	0.010
Schwartz	0.551	0.005	0.713	0.008	0.842	0.009	0.914	0.007

FH SCORE	Prevalence=0.05							
	RR=1.2		RR=2.0		RR=3.0		RR=4.0	
	mean (AUC)	std (AUC)	mean (AUC)	std (AUC)	mean (AUC)	std (AUC)	mean (AUC)	std (AUC)
Binary indicator	0.534	0.007	0.625	0.009	0.687	0.012	0.719	0.011
Number of affected	0.545	0.006	0.685	0.008	0.802	0.010	0.872	0.008
FHS1	0.548	0.006	0.701	0.008	0.825	0.009	0.897	0.008
FHS2	0.550	0.006	0.709	0.008	0.838	0.009	0.911	0.007
Proportion	0.547	0.007	0.696	0.008	0.820	0.009	0.893	0.008
Slanck and Evans	0.547	0.006	0.699	0.008	0.828	0.009	0.904	0.007
Groenelved and Hitzeroth	0.548	0.006	0.700	0.008	0.824	0.009	0.897	0.007
Fain and Goldar	0.545	0.011	0.691	0.011	0.815	0.012	0.889	0.009
Kee	0.540	0.016	0.683	0.009	0.808	0.010	0.885	0.008
Lynch	0.547	0.006	0.699	0.008	0.828	0.009	0.904	0.007
Reed	0.552	0.005	0.715	0.007	0.845	0.008	0.917	0.006
Chakraborty	0.537	0.008	0.655	0.009	0.759	0.010	0.828	0.010
Schwartz	0.551	0.005	0.708	0.007	0.831	0.009	0.900	0.007

FH SCORE	Prevalence=0.1							
	RR=1.2		RR=2.0		RR=3.0		RR=4.0	
	mean (AUC)	std (AUC)	mean (AUC)	std (AUC)	mean (AUC)	std (AUC)	mean (AUC)	std (AUC)
Binary indicator	0.534	0.007	0.626	0.009	0.690	0.010	0.729	0.012
Number of affected	0.545	0.006	0.681	0.008	0.789	0.008	0.856	0.009
FHS1	0.548	0.006	0.697	0.008	0.812	0.008	0.880	0.008
FHS2	0.550	0.006	0.704	0.008	0.823	0.008	0.893	0.007
Proportion	0.546	0.006	0.690	0.008	0.805	0.008	0.874	0.008
Slanck and Evans	0.547	0.006	0.694	0.008	0.812	0.008	0.882	0.008
Groenelved and Hitzeroth	0.548	0.006	0.696	0.008	0.811	0.008	0.880	0.008
Fain and Goldar	0.545	0.011	0.686	0.011	0.801	0.010	0.872	0.010
Kee	0.540	0.016	0.677	0.009	0.791	0.009	0.862	0.009
Lynch	0.547	0.006	0.694	0.008	0.812	0.008	0.882	0.008
Reed	0.552	0.005	0.709	0.007	0.830	0.008	0.899	0.007
Chakraborty	0.536	0.008	0.651	0.009	0.748	0.009	0.812	0.010
Schwartz	0.550	0.005	0.703	0.007	0.819	0.008	0.886	0.008

FH SCORE	Prevalence=0.2							
	RR=1.2		RR=2.0		RR=3.0		RR=4.0	
	mean (AUC)	std (AUC)	mean (AUC)	std (AUC)	mean (AUC)	std (AUC)	mean (AUC)	std (AUC)
Binary indicator	0.534	0.007	0.626	0.009	0.694	0.010	0.737	0.011
Number of affected	0.544	0.006	0.673	0.007	0.769	0.008	0.827	0.008
FHS1	0.548	0.006	0.688	0.008	0.790	0.009	0.850	0.008
FHS2	0.549	0.006	0.694	0.008	0.800	0.009	0.861	0.008
Proportion	0.546	0.006	0.681	0.007	0.781	0.008	0.841	0.007
Slanck and Evans	0.546	0.006	0.683	0.008	0.786	0.008	0.846	0.008
Groenelved and Hitzeroth	0.547	0.006	0.686	0.008	0.789	0.009	0.849	0.008
Fain and Goldar	0.545	0.011	0.678	0.011	0.779	0.011	0.840	0.010
Kee	0.539	0.015	0.667	0.009	0.763	0.009	0.822	0.009
Lynch	0.546	0.006	0.683	0.008	0.786	0.008	0.846	0.008
Reed	0.551	0.005	0.700	0.007	0.807	0.008	0.868	0.007
Chakraborty	0.536	0.008	0.644	0.009	0.729	0.010	0.784	0.009
Schwartz	0.550	0.005	0.695	0.007	0.799	0.008	0.859	0.007

FH SCORE	Prevalence=0.4							
	RR=1.2		RR=2.0		RR=3.0		RR=4.0	
	mean (AUC)	std (AUC)	mean (AUC)	std (AUC)	mean (AUC)	std (AUC)	mean (AUC)	std (AUC)
Binary indicator	0.534	0.007	0.624	0.008	0.691	0.010	0.733	0.009
Number of affected	0.543	0.006	0.658	0.007	0.737	0.008	0.782	0.008
FHS1	0.547	0.006	0.672	0.008	0.757	0.009	0.804	0.009
FHS2	0.548	0.006	0.677	0.008	0.763	0.009	0.811	0.009
Proportion	0.545	0.006	0.664	0.007	0.745	0.008	0.790	0.008
Slanck and Evans	0.545	0.006	0.666	0.007	0.747	0.008	0.792	0.008
Groenelved and Hitzeroth	0.546	0.006	0.671	0.008	0.755	0.009	0.803	0.009
Fain and Goldar	0.544	0.011	0.661	0.011	0.743	0.011	0.789	0.011
Kee	0.537	0.017	0.649	0.009	0.722	0.009	0.764	0.009
Lynch	0.545	0.006	0.666	0.007	0.747	0.008	0.792	0.008
Reed	0.550	0.005	0.683	0.007	0.771	0.008	0.820	0.008
Chakraborty	0.535	0.008	0.632	0.009	0.700	0.010	0.741	0.009
Schwartz	0.549	0.005	0.679	0.007	0.766	0.008	0.814	0.008

PANEL B

FH SCORE	Prevalence=0.0001							
	RR=1.2		RR=2.0		RR=3.0		RR=4.0	
	mean (AUC)	std (AUC)	mean (AUC)	std (AUC)	mean (AUC)	std (AUC)	mean (AUC)	std (AUC)
Binary indicator	0.535	0.007	0.621	0.010	0.667	0.012	0.685	0.013
Number of affected	0.550	0.006	0.708	0.009	0.834	0.009	0.909	0.008
FHS1	0.553	0.006	0.723	0.009	0.859	0.009	0.935	0.006
FHS2	0.554	0.006	0.728	0.009	0.865	0.008	0.941	0.005
Proportion	0.553	0.006	0.723	0.009	0.860	0.008	0.937	0.006
Slanck and Evans	0.553	0.006	0.725	0.009	0.863	0.008	0.940	0.006
Groenelved and Hitzeroth	0.553	0.006	0.723	0.009	0.859	0.009	0.935	0.006
Fain and Goldar	0.550	0.014	0.705	0.014	0.832	0.013	0.908	0.010
Kee	0.550	0.007	0.714	0.010	0.851	0.009	0.930	0.006
Lynch	0.553	0.006	0.725	0.009	0.863	0.008	0.940	0.006
Reed	0.556	0.005	0.733	0.009	0.870	0.008	0.944	0.005
Chakraborty	0.545	0.008	0.688	0.010	0.809	0.010	0.885	0.009
Schwartz	0.556	0.005	0.729	0.009	0.861	0.008	0.933	0.006

FH SCORE	Prevalence=0.001							
	RR=1.2		RR=2.0		RR=3.0		RR=4.0	
	mean (AUC)	std (AUC)	mean (AUC)	std (AUC)	mean (AUC)	std (AUC)	mean (AUC)	std (AUC)
Binary indicator	0.535	0.007	0.621	0.010	0.667	0.012	0.685	0.013
Number of affected	0.550	0.006	0.708	0.009	0.834	0.009	0.909	0.008
FHS1	0.553	0.006	0.723	0.009	0.858	0.009	0.935	0.006
FHS2	0.554	0.006	0.727	0.009	0.865	0.008	0.941	0.006
Proportion	0.553	0.006	0.723	0.009	0.860	0.008	0.937	0.006
Slanck and Evans	0.553	0.006	0.725	0.009	0.863	0.008	0.940	0.006
Groenelved and Hitzeroth	0.553	0.006	0.723	0.009	0.858	0.009	0.935	0.006
Fain and Goldar	0.550	0.014	0.705	0.014	0.832	0.013	0.908	0.010
Kee	0.550	0.008	0.714	0.010	0.850	0.009	0.930	0.007
Lynch	0.553	0.006	0.725	0.009	0.863	0.008	0.940	0.006
Reed	0.556	0.005	0.733	0.009	0.870	0.008	0.944	0.006
Chakraborty	0.545	0.008	0.688	0.010	0.809	0.010	0.885	0.009
Schwartz	0.556	0.005	0.729	0.009	0.861	0.008	0.933	0.006

FH SCORE	Prevalence=0.05							
	RR=1.2		RR=2.0		RR=3.0		RR=4.0	
	mean (AUC)	std (AUC)	mean (AUC)	std (AUC)	mean (AUC)	std (AUC)	mean (AUC)	std (AUC)
Binary indicator	0.535	0.007	0.623	0.010	0.676	0.012	0.703	0.013
Number of affected	0.550	0.006	0.703	0.009	0.822	0.009	0.893	0.008
FHS1	0.553	0.006	0.717	0.009	0.844	0.009	0.917	0.007
FHS2	0.554	0.006	0.721	0.009	0.850	0.009	0.922	0.006
Proportion	0.553	0.006	0.717	0.009	0.845	0.009	0.917	0.006
Slanck and Evans	0.553	0.006	0.718	0.009	0.847	0.009	0.920	0.007
Groenelved and Hitzeroth	0.553	0.006	0.717	0.009	0.844	0.009	0.916	0.007
Fain and Goldar	0.550	0.014	0.700	0.014	0.820	0.013	0.892	0.011
Kee	0.549	0.007	0.707	0.010	0.835	0.009	0.909	0.007
Lynch	0.553	0.006	0.718	0.009	0.847	0.009	0.920	0.007
Reed	0.556	0.005	0.728	0.009	0.856	0.008	0.927	0.006
Chakraborty	0.545	0.008	0.683	0.010	0.797	0.010	0.868	0.010
Schwartz	0.555	0.005	0.723	0.009	0.849	0.009	0.918	0.007

FH SCORE	Prevalence=0.1							
	RR=1.2		RR=2.0		RR=3.0		RR=4.0	
	mean (AUC)	std (AUC)	mean (AUC)	std (AUC)	mean (AUC)	std (AUC)	mean (AUC)	std (AUC)
Binary indicator	0.535	0.007	0.625	0.010	0.683	0.012	0.717	0.013
Number of affected	0.550	0.006	0.698	0.009	0.811	0.009	0.877	0.008
FHS1	0.553	0.006	0.712	0.009	0.831	0.009	0.898	0.008
FHS2	0.554	0.006	0.716	0.009	0.837	0.009	0.904	0.007
Proportion	0.553	0.006	0.711	0.008	0.831	0.009	0.899	0.007
Slanck and Evans	0.553	0.006	0.712	0.009	0.833	0.009	0.901	0.008
Groenelved and Hitzeroth	0.553	0.006	0.711	0.009	0.831	0.009	0.898	0.008
Fain and Goldar	0.549	0.015	0.695	0.014	0.808	0.013	0.875	0.011
Kee	0.549	0.008	0.701	0.009	0.820	0.010	0.889	0.009
Lynch	0.553	0.006	0.712	0.009	0.833	0.009	0.901	0.008
Reed	0.556	0.005	0.722	0.008	0.844	0.009	0.910	0.007
Chakraborty	0.544	0.008	0.679	0.010	0.785	0.010	0.851	0.010
Schwartz	0.555	0.005	0.718	0.008	0.837	0.009	0.903	0.007

FH SCORE	Prevalence=0.2							
	RR=1.2		RR=2.0		RR=3.0		RR=4.0	
	mean (AUC)	std (AUC)	mean (AUC)	std (AUC)	mean (AUC)	std (AUC)	mean (AUC)	std (AUC)
Binary indicator	0.535	0.007	0.626	0.010	0.691	0.012	0.732	0.012
Number of affected	0.549	0.006	0.689	0.009	0.790	0.009	0.849	0.008
FHS1	0.552	0.006	0.701	0.009	0.808	0.009	0.868	0.009
FHS2	0.553	0.006	0.705	0.009	0.813	0.009	0.873	0.008
Proportion	0.552	0.006	0.701	0.008	0.807	0.009	0.867	0.008
Slanck and Evans	0.552	0.006	0.702	0.009	0.808	0.009	0.868	0.008
Groenelved and Hitzeroth	0.552	0.007	0.701	0.009	0.808	0.009	0.868	0.009
Fain and Goldar	0.549	0.015	0.686	0.014	0.787	0.014	0.847	0.012
Kee	0.548	0.008	0.690	0.010	0.794	0.010	0.853	0.009
Lynch	0.552	0.006	0.702	0.009	0.808	0.009	0.868	0.008
Reed	0.555	0.005	0.712	0.008	0.821	0.009	0.881	0.008
Chakraborty	0.544	0.008	0.670	0.010	0.765	0.011	0.823	0.010
Schwartz	0.554	0.005	0.709	0.008	0.817	0.009	0.877	0.008

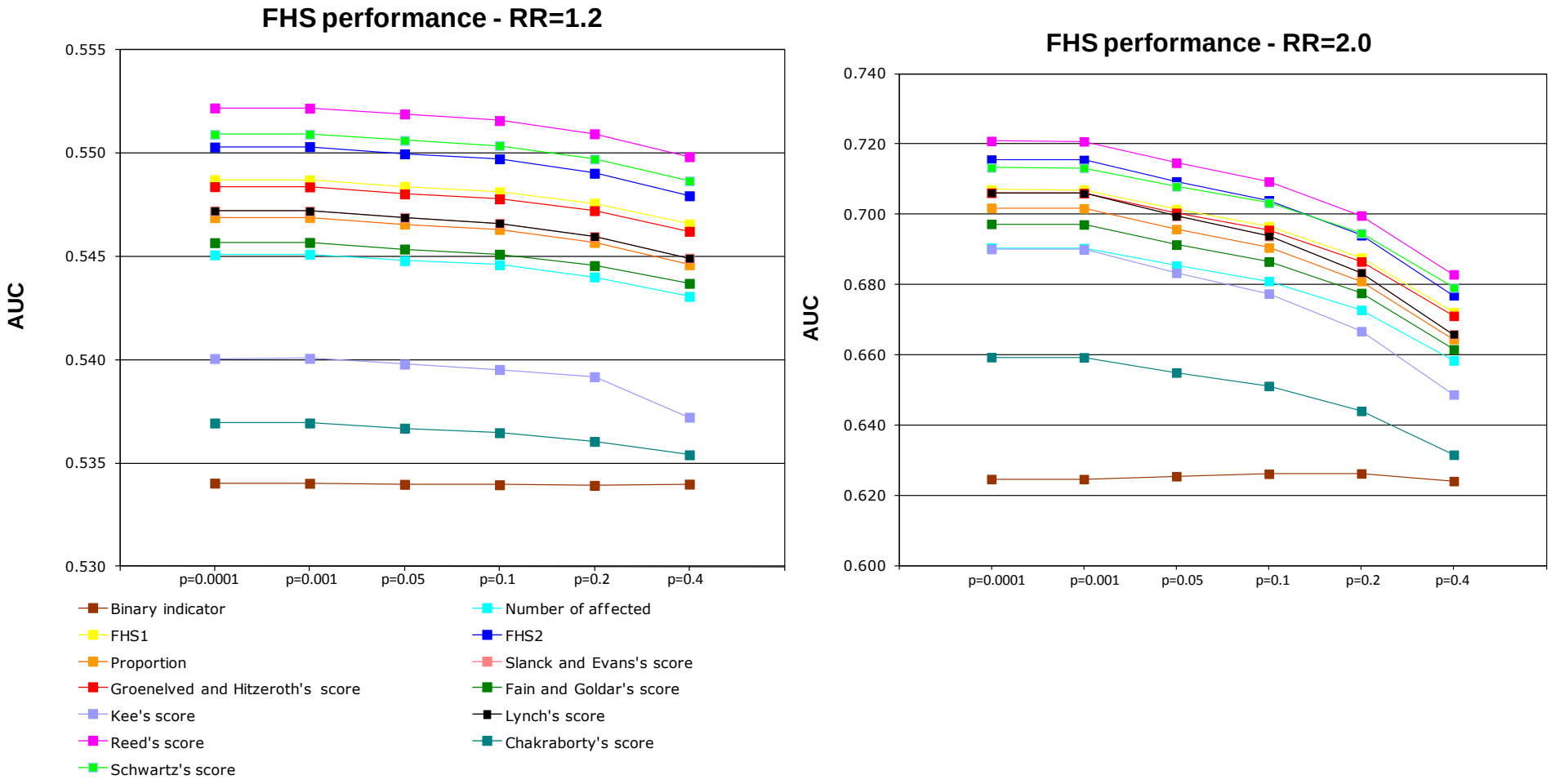
FH SCORE	Prevalence=0.4							
	RR=1.2		RR=2.0		RR=3.0		RR=4.0	
	mean (AUC)	std (AUC)	mean (AUC)	std (AUC)	mean (AUC)	std (AUC)	mean (AUC)	std (AUC)
Binary indicator	0.535	0.008	0.627	0.010	0.695	0.010	0.739	0.011
Number of affected	0.548	0.006	0.674	0.008	0.758	0.009	0.806	0.008
FHS1	0.551	0.007	0.684	0.009	0.771	0.010	0.820	0.009
FHS2	0.552	0.006	0.687	0.009	0.775	0.009	0.824	0.009
Proportion	0.551	0.006	0.683	0.008	0.769	0.009	0.817	0.009
Slanck and Evans	0.551	0.006	0.683	0.008	0.770	0.009	0.817	0.009
Groenelved and Hitzeroth	0.551	0.007	0.684	0.009	0.771	0.009	0.820	0.009
Fain and Goldar	0.548	0.015	0.669	0.015	0.752	0.014	0.799	0.013
Kee	0.547	0.008	0.671	0.010	0.752	0.010	0.798	0.010
Lynch	0.551	0.006	0.683	0.008	0.770	0.009	0.817	0.009
Reed	0.554	0.005	0.695	0.008	0.786	0.009	0.836	0.009
Chakraborty	0.543	0.008	0.656	0.009	0.733	0.010	0.778	0.010
Schwartz	0.553	0.005	0.692	0.008	0.783	0.009	0.833	0.009

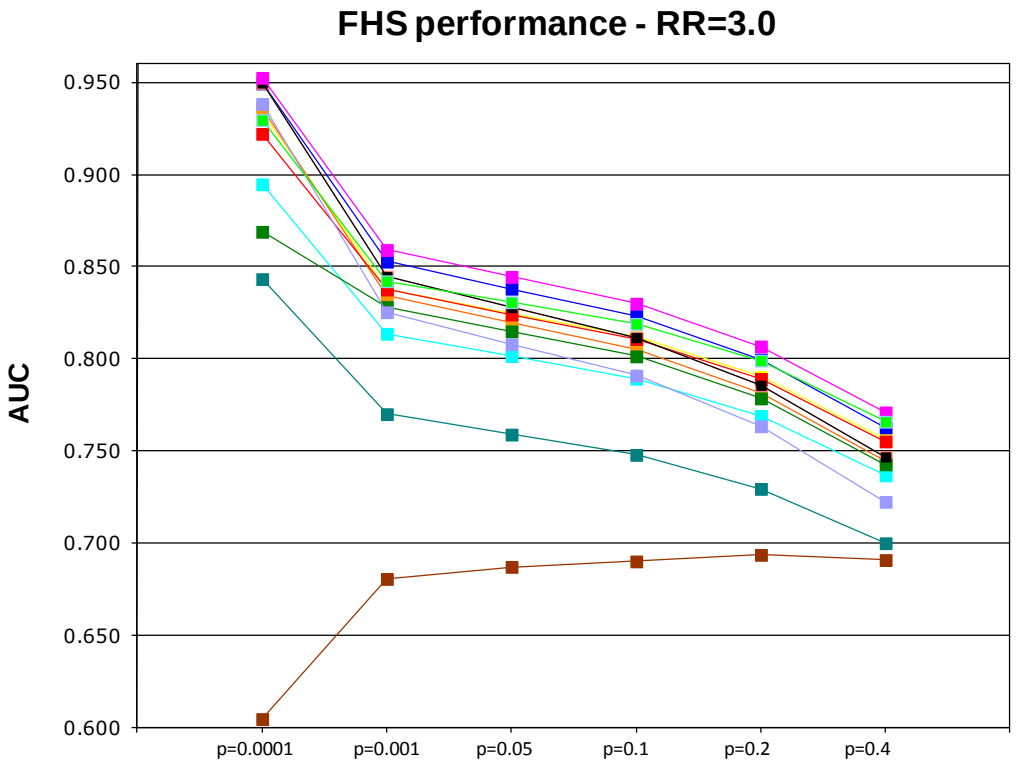
Table 14. Family history scores' final ranking according to mean AUC.

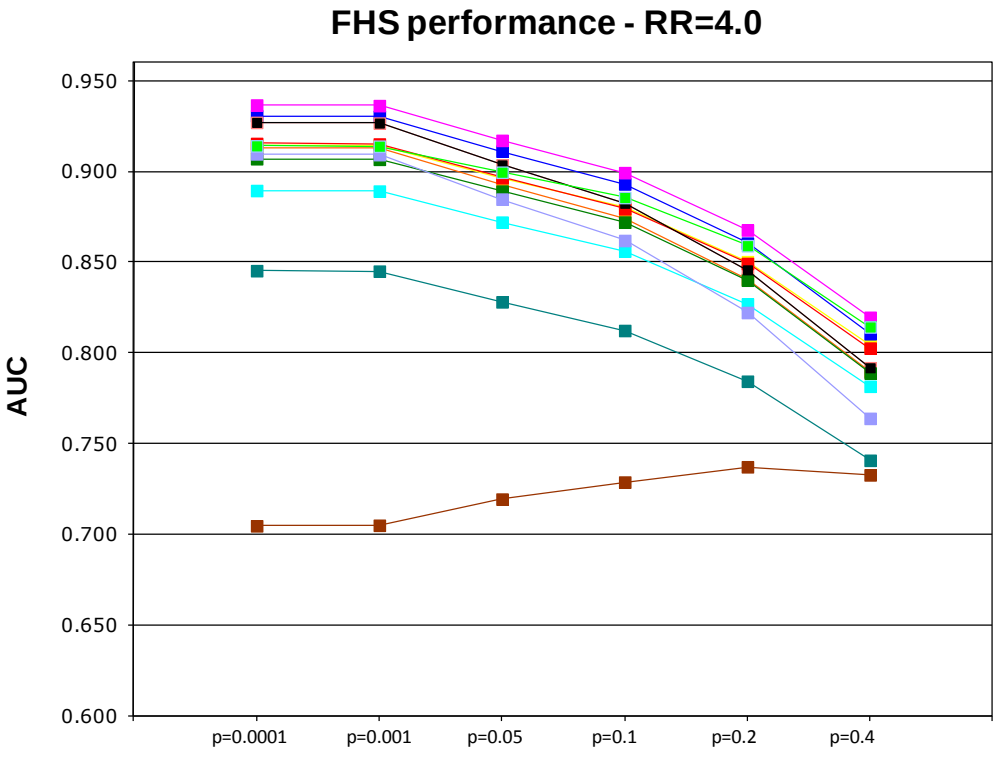
Final Ranking
1 - Reed's score
2 - FHS2
3 - Schwartz's score
4 - Slanck and Evans's score
5 - Lynch's score
6 - FHS1
7 - Groenelved and Hitzeroth's score
8 - Proportion
9 - Fain and Goldar's score
10 - Number of affected relatives
11 - Kee's score
12 - Chakraborty's score
13 - Binary indicator

Figure 5. Mean AUC for different FH scores in relation with prevalence of family at high risk, by different relative risk of FH, with expected risks of disease equal to 0.001, 0.01, 0.10 and 0.20 (Panel A) or 0.04, 0.08, 0.12, 0.16 (Panel B) according to subsequent age groups.

PANEL A

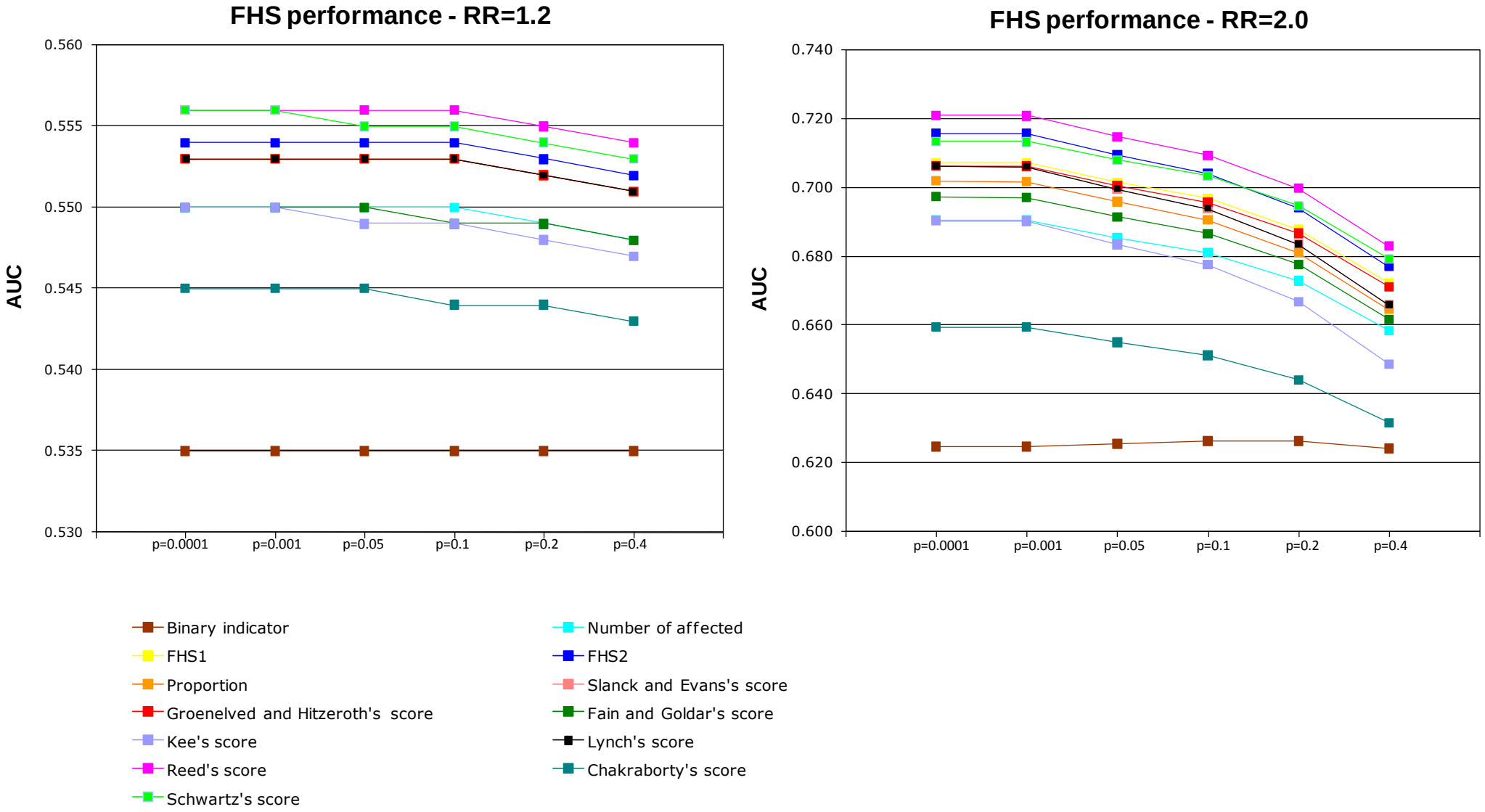


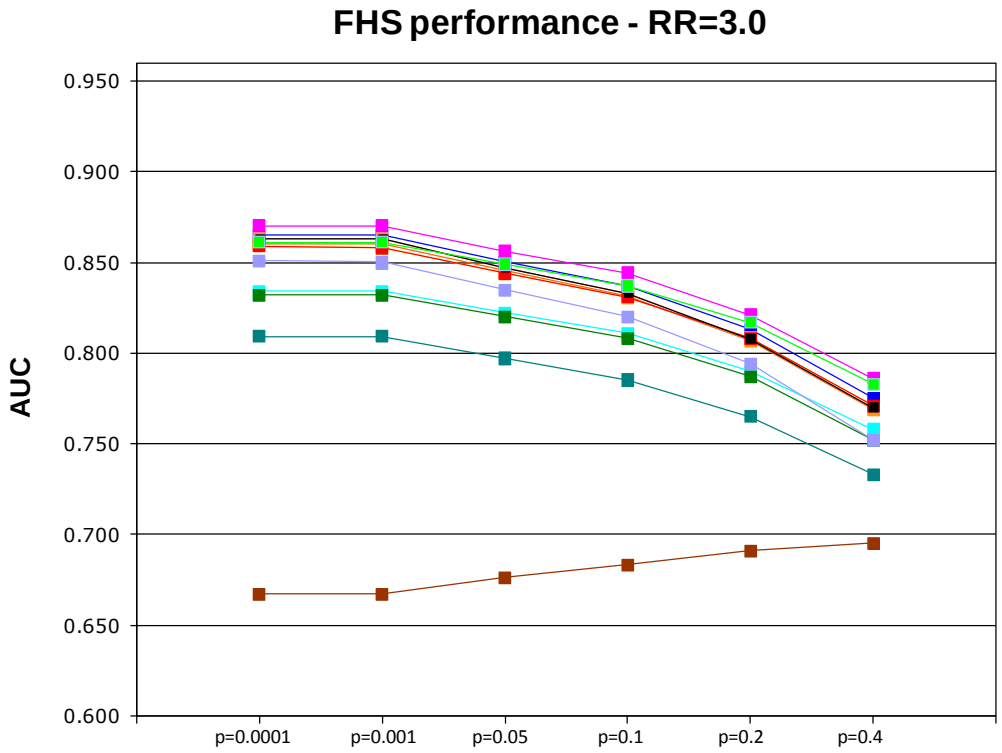


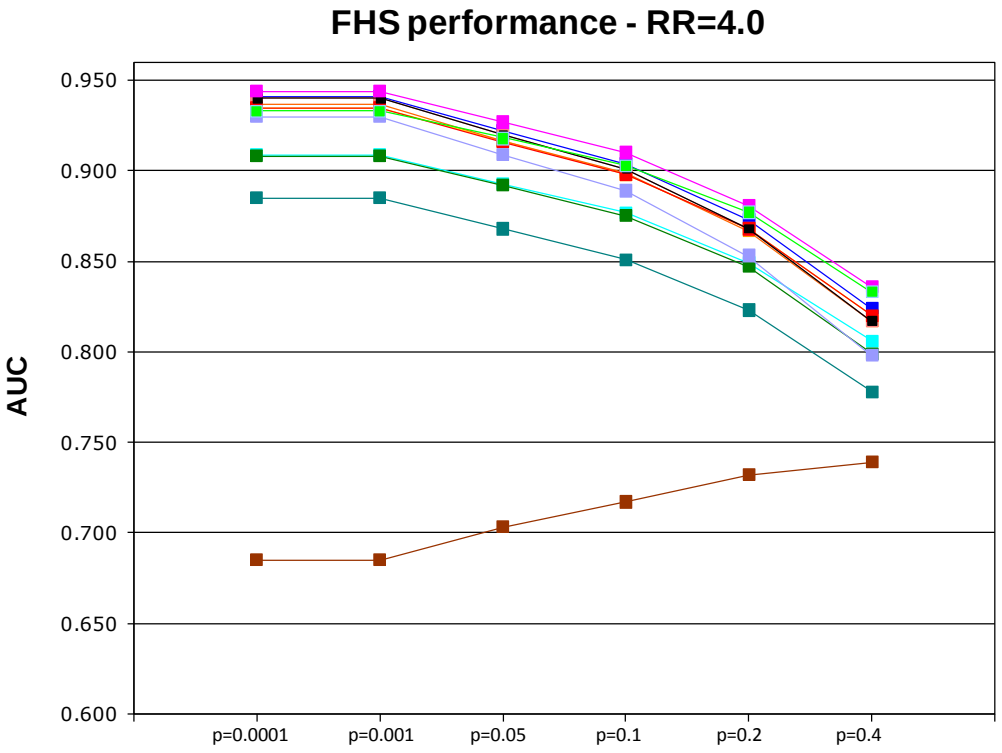


- Binary indicator
- Number of affected
- FHS1
- FHS2
- Proportion
- Slanck and Evans's score
- Groenelved and Hitzeroth's score
- Fain and Goldar's score
- Kee's score
- Lynch's score
- Reed's score
- Chakraborty's score
- Schwartz's score

PANEL B


FHS performance - RR=1.2
AUC
0.560
0.555
0.550
0.545
0.540
0.535
0.530
p=0.0001
p=0.001
p=0.05
p=0.1
p=0.2
p=0.4
FHS performance - RR=2.0
AUC
0.740
0.720
0.700
0.680
0.660
0.640
0.620
0.600
p=0.0001
p=0.001
p=0.05
p=0.1
p=0.2
p=0.4
Binary indicator
Number of affected
FHS1
FHS2
Proportion
Slanck and Evans's score
Groenelved and Hitzeroth's score
Fain and Goldar's score
Kee's score
Lynch's score
Reed's score
Chakraborty's score
Schwartz's score


FHS performance - RR=3.0
AUC
0.950
0.900
0.850
0.800
0.750
0.700
0.650
0.600
p=0.0001
p=0.001
p=0.05
p=0.1
p=0.2
p=0.4


FHS performance - RR=4.0
AUC
0.950
0.900
0.850
0.800
0.750
0.700
0.650
0.600
p=0.0001
p=0.001
p=0.05
p=0.1
p=0.2
p=0.4

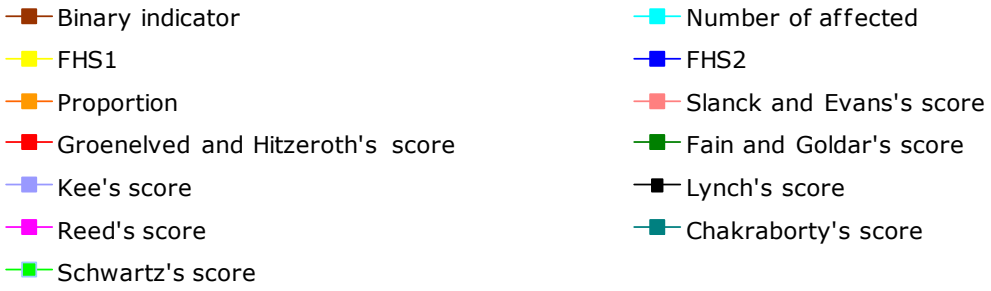
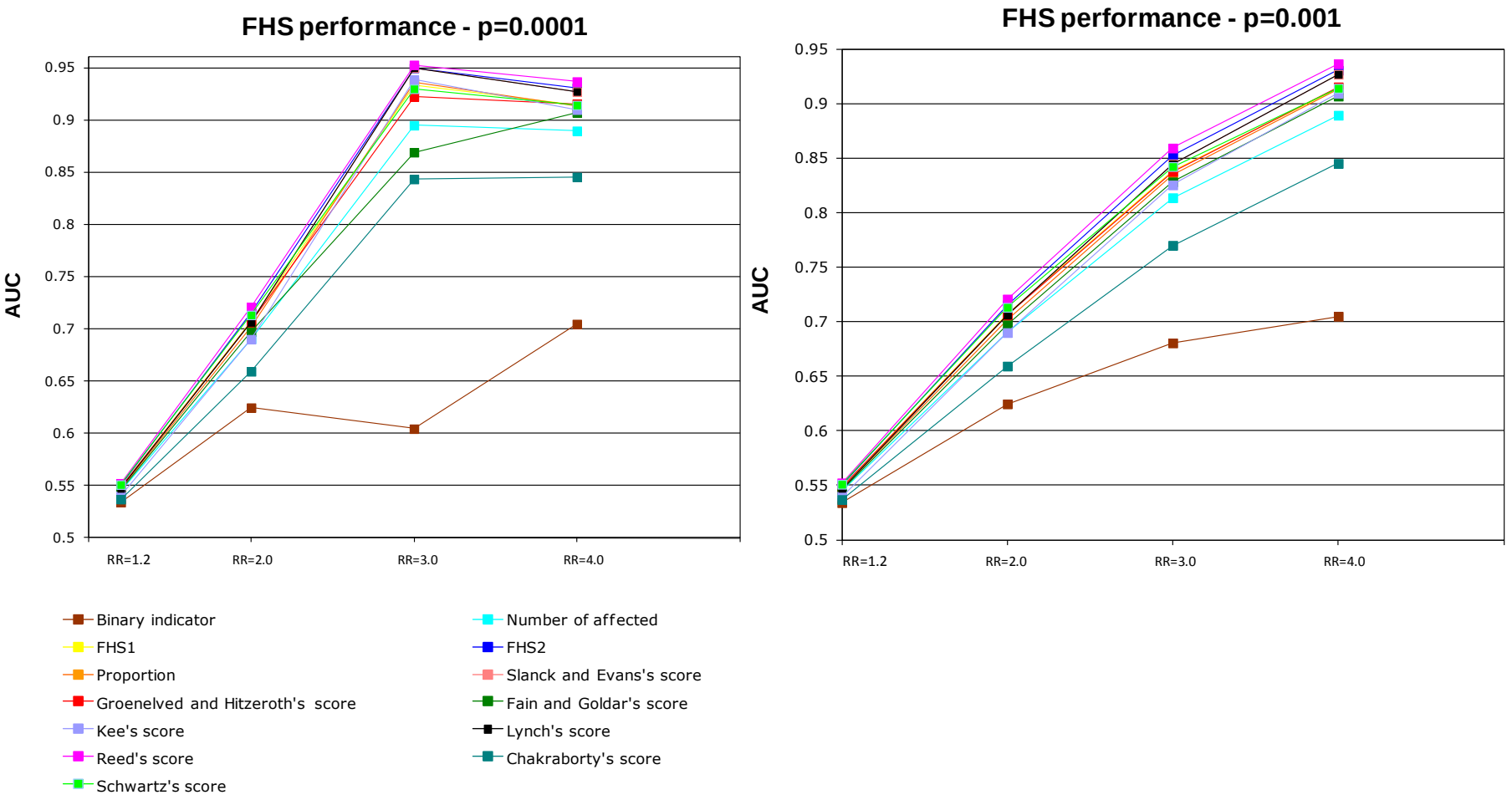
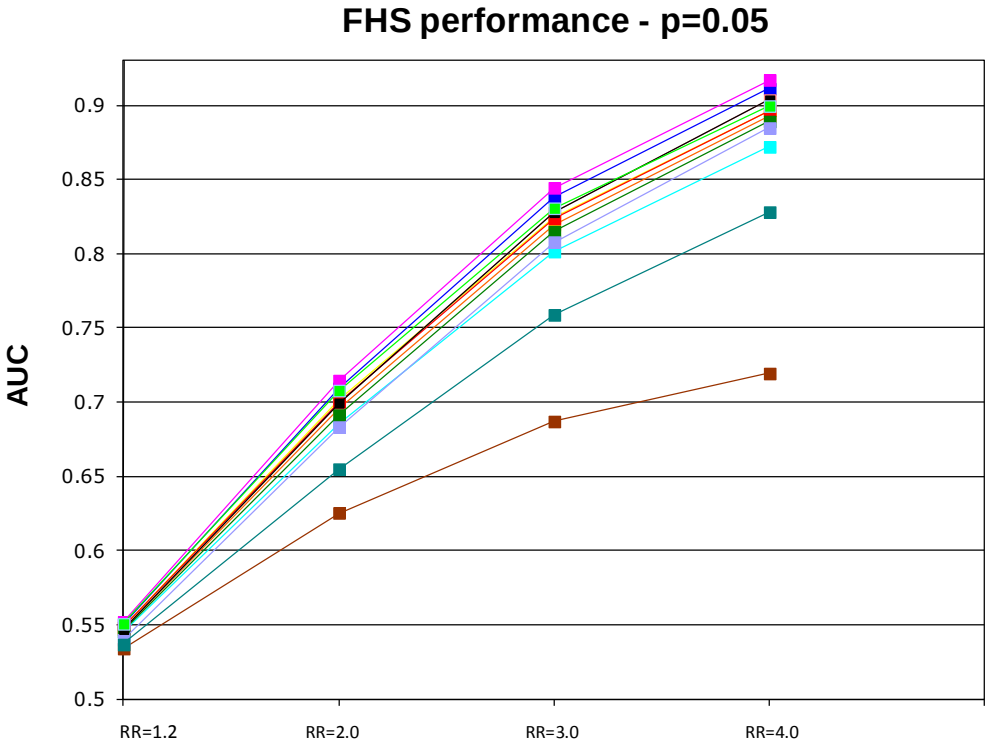
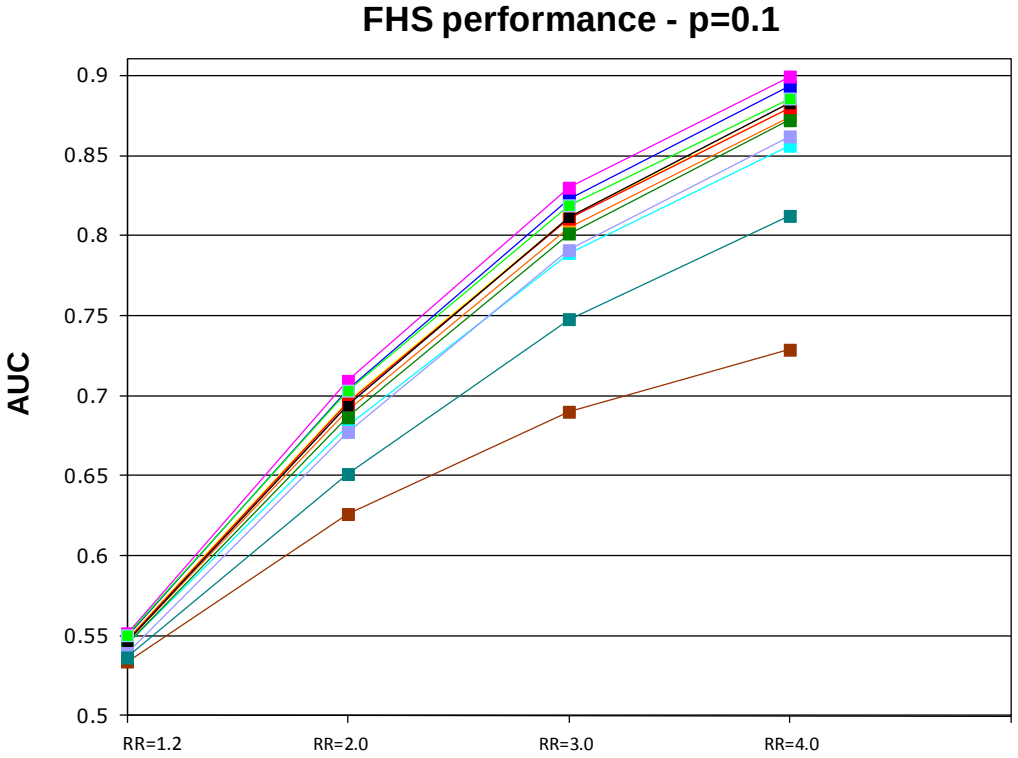

Binary indicator
Number of affected
FHS1
FHS2
Proportion
Slanck and Evans's score
Groenelved and Hitzeroth's score
Fain and Goldar's score
Kee's score
Lynch's score
Reed's score
Chakraborty's score
Schwartz's score

Figure 6. Mean AUC for different FH scores in relation with RR, by different prevalences of family at high risk, with expected risks of disease equal to 0.001, 0.01, 0.10 and 0.20 (Panel A) or 0.04, 0.08, 0.12, 0.16 (Panel B) according to subsequent age groups.

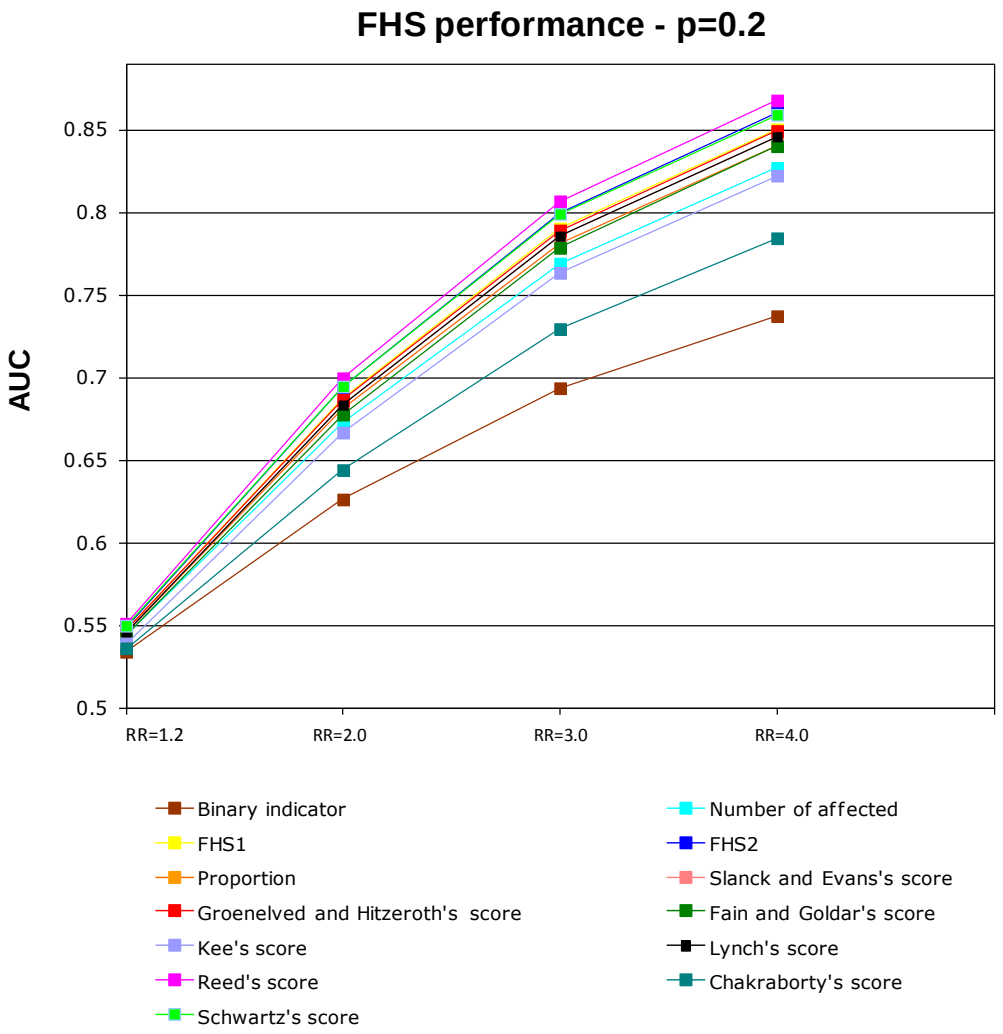
PANEL A

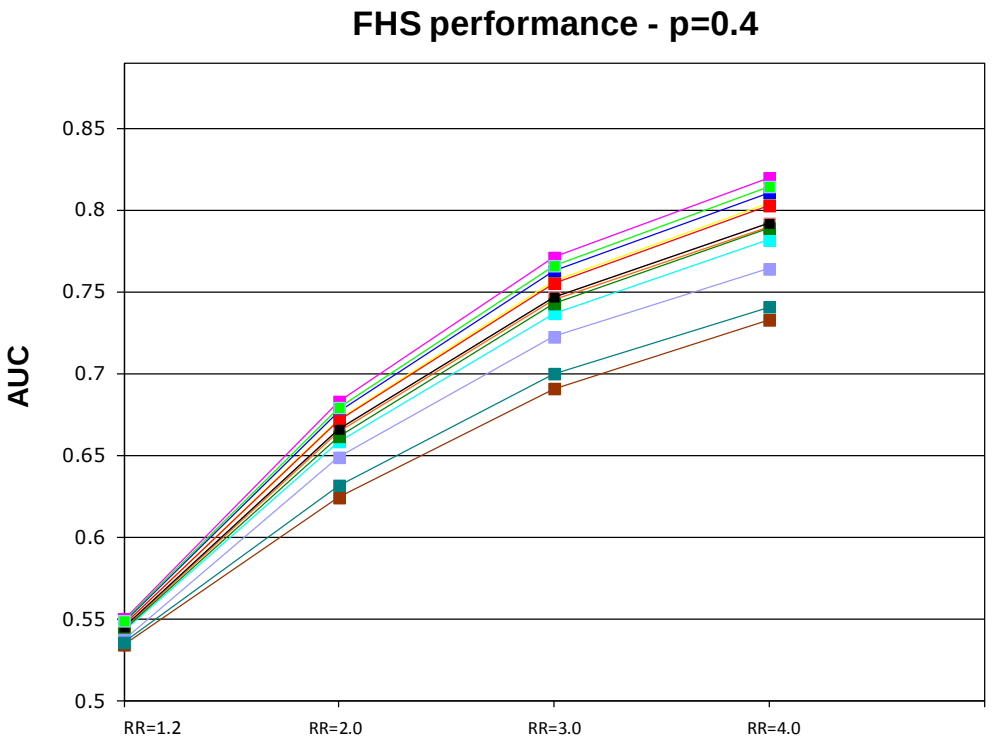



FHS performance - p=0.05
AUC
0.9
0.85
0.8
0.75
0.7
0.65
0.6
0.55
0.5
RR=1.2
RR=2.0
RR=3.0
RR=4.0

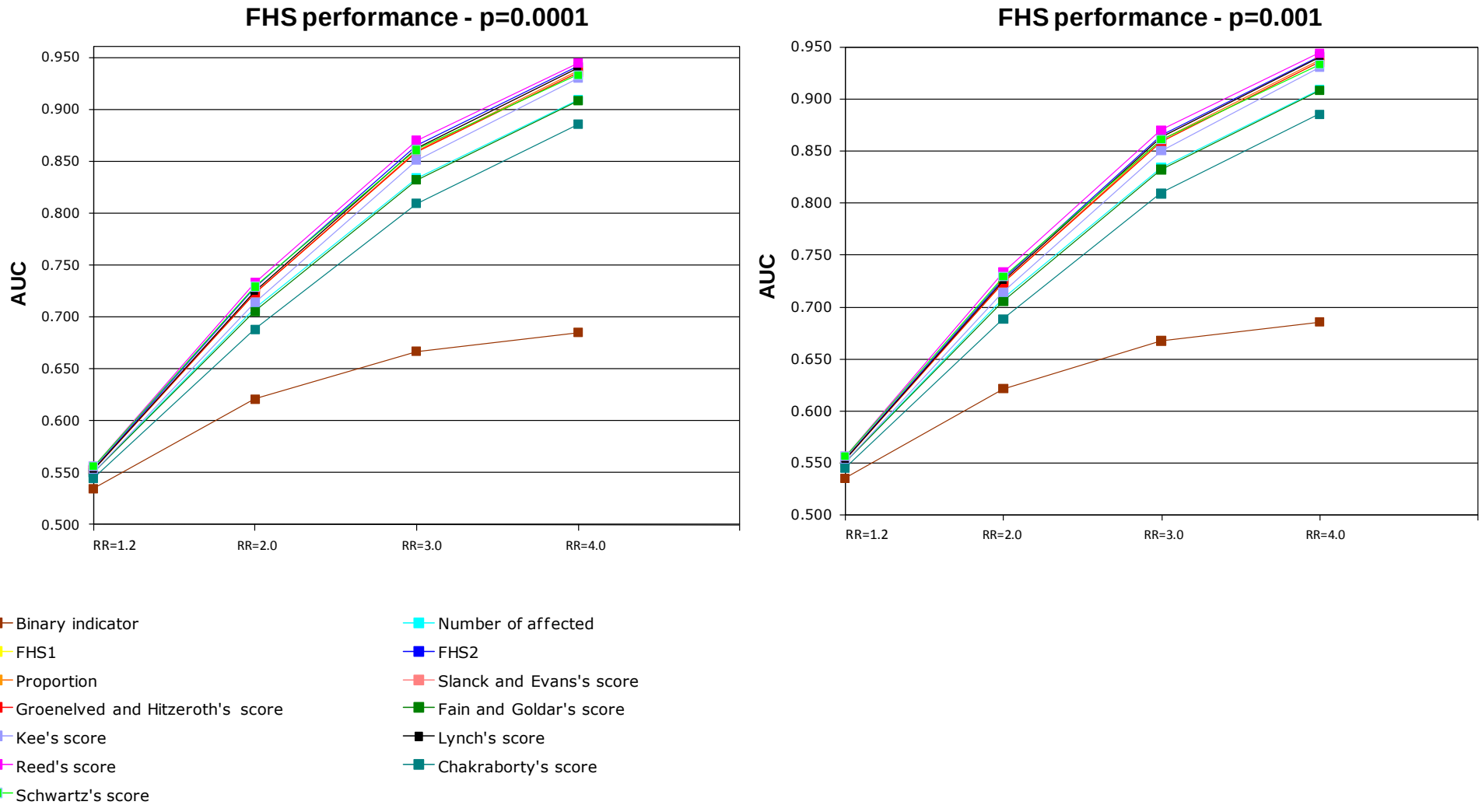

FHS performance - p=0.1
AUC
0.9
0.85
0.8
0.75
0.7
0.65
0.6
0.55
0.5
RR=1.2
RR=2.0
RR=3.0
RR=4.0

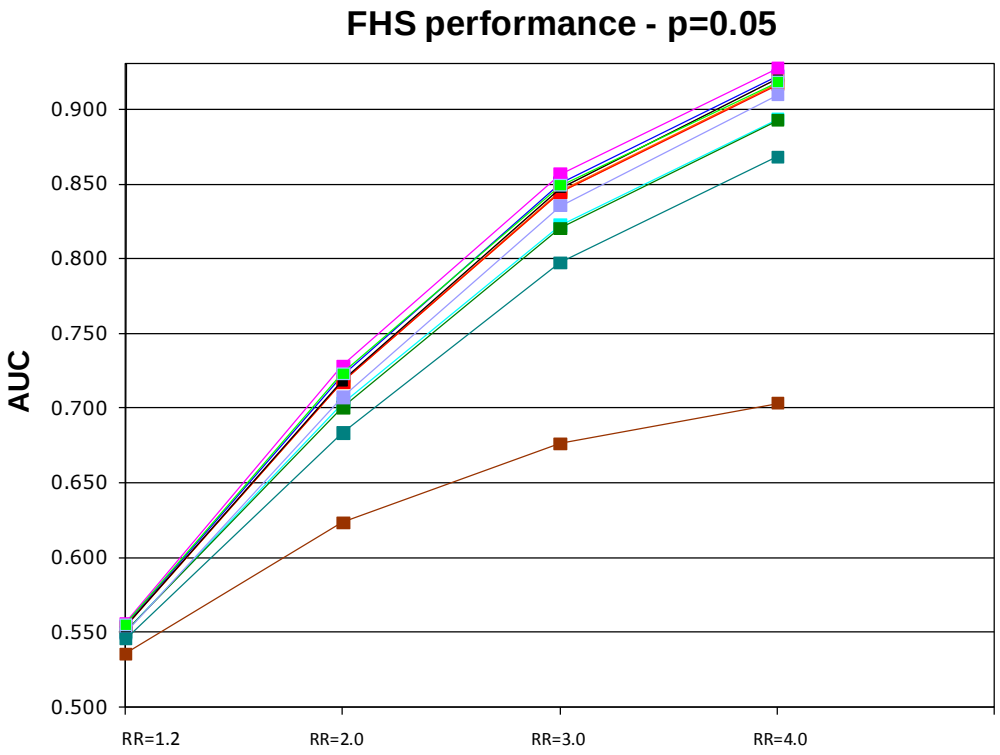

Binary indicator
Number of affected
FHS1
FHS2
Proportion
Slanck and Evans's score
Groenelved and Hitzeroth's score
Fain and Goldar's score
Kee's score
Lynch's score
Reed's score
Chakraborty's score
Schwartz's score

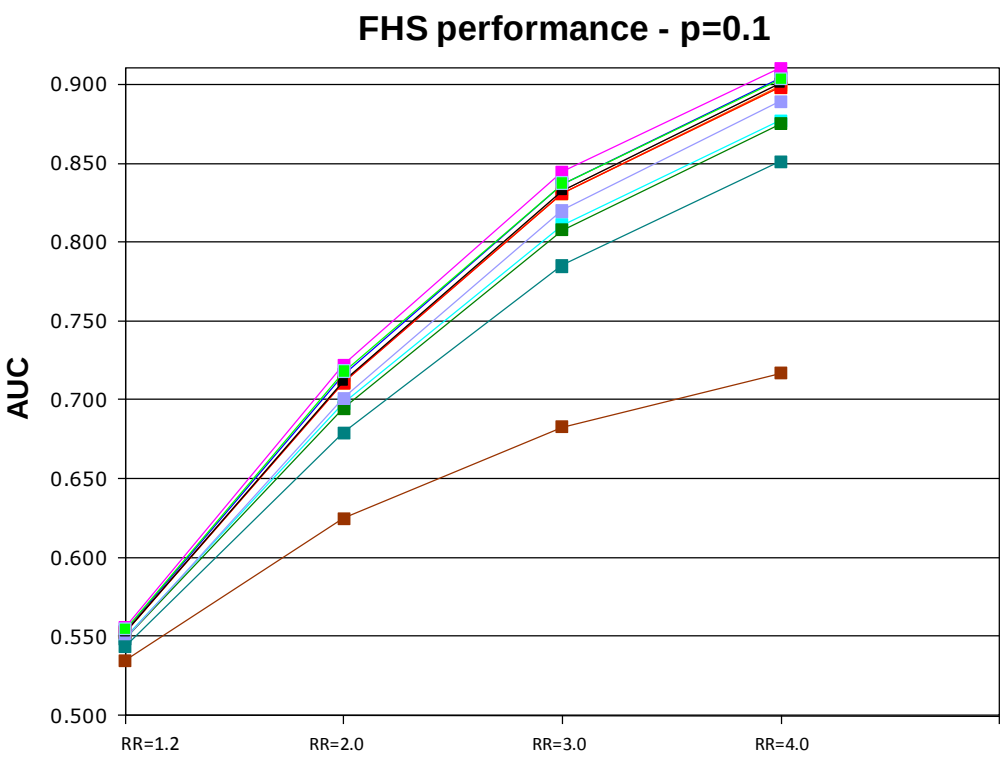

FHS performance - p=0.2
AUC
0.85
0.8
0.75
0.7
0.65
0.6
0.55
0.5
RR=1.2
RR=2.0
RR=3.0
RR=4.0
Binary indicator
Number of affected
FHS1
FHS2
Proportion
Slanck and Evans's score
Groenelved and Hitzeroth's score
Fain and Goldar's score
Kee's score
Lynch's score
Reed's score
Chakraborty's score
Schwartz's score

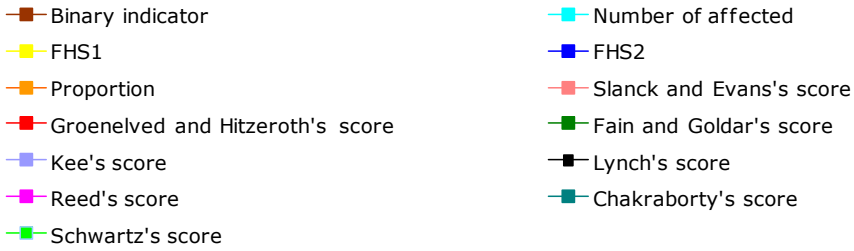

FHS performance - p=0.4
AUC
0.85
0.8
0.75
0.7
0.65
0.6
0.55
0.5
RR=1.2
RR=2.0
RR=3.0
RR=4.0

PANEL B


FHS performance - p=0.0001
AUC
0.950
0.900
0.850
0.800
0.750
0.700
0.650
0.600
0.550
0.500
RR=1.2
RR=2.0
RR=3.0
RR=4.0
FHS performance - p=0.001
AUC
0.950
0.900
0.850
0.800
0.750
0.700
0.650
0.600
0.550
0.500
RR=1.2
RR=2.0
RR=3.0
RR=4.0
Binary indicator
Number of affected
FHS1
FHS2
Proportion
Slanck and Evans's score
Groenelved and Hitzeroth's score
Fain and Goldar's score
Kee's score
Lynch's score
Reed's score
Chakraborty's score
Schwartz's score


FHS performance - p=0.05
AUC
0.900
0.850
0.800
0.750
0.700
0.650
0.600
0.550
0.500
RR=1.2
RR=2.0
RR=3.0
RR=4.0


FHS performance - p=0.1
AUC
0.900
0.850
0.800
0.750
0.700
0.650
0.600
0.550
0.500
RR=1.2
RR=2.0
RR=3.0
RR=4.0


Binary indicator
Number of affected
FHS1
FHS2
Proportion
Slanck and Evans's score
Groenelved and Hitzeroth's score
Fain and Goldar's score
Kee's score
Lynch's score
Reed's score
Chakraborty's score
Schwartz's score

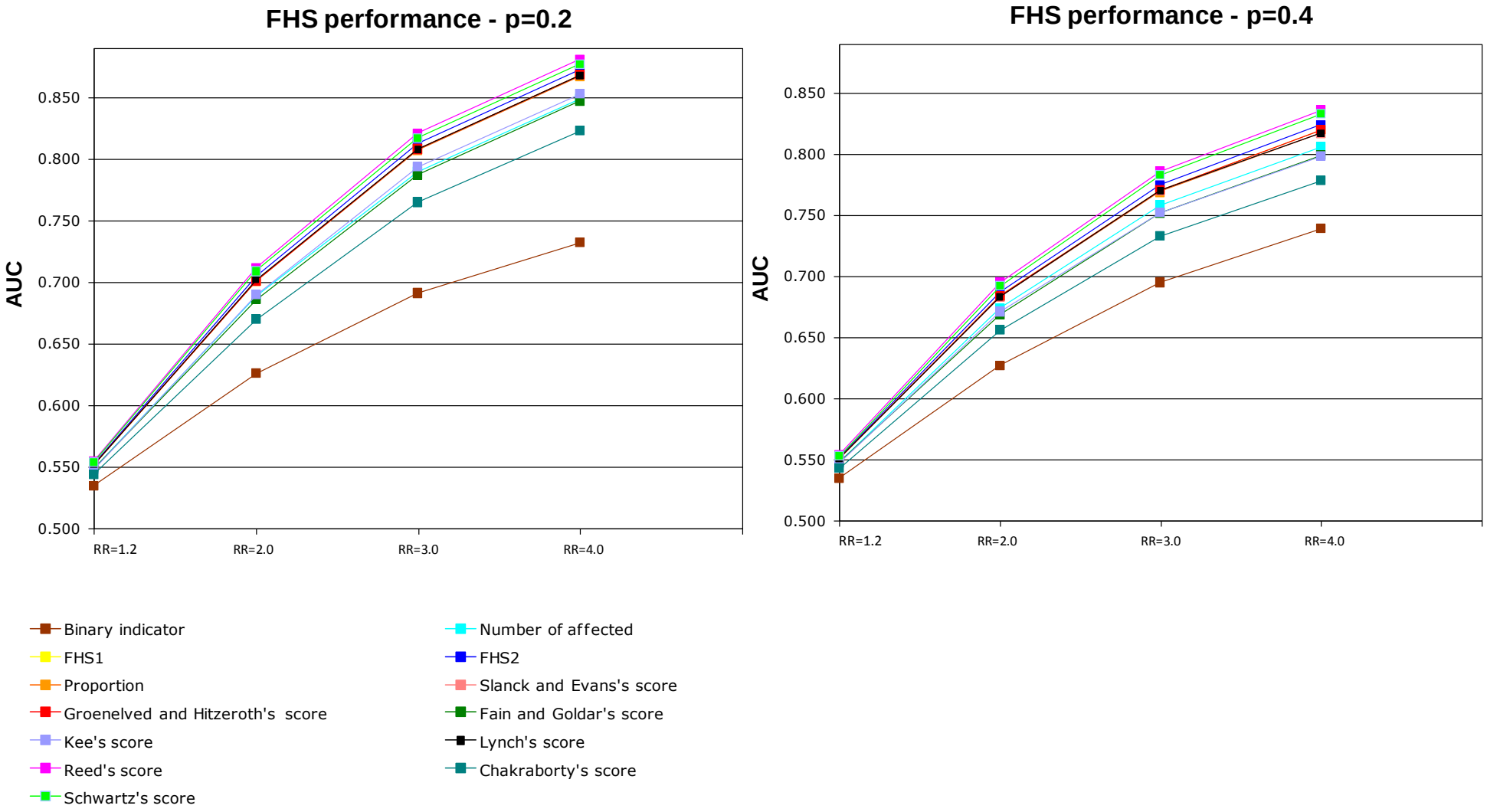

FHS performance - p=0.2
AUC
0.850
0.800
0.750
0.700
0.650
0.600
0.550
0.500
RR=1.2
RR=2.0
RR=3.0
RR=4.0
FHS performance - p=0.4
AUC
0.850
0.800
0.750
0.700
0.650
0.600
0.550
0.500
RR=1.2
RR=2.0
RR=3.0
RR=4.0
Binary indicator
Number of affected
FHS1
FHS2
Proportion
Slanck and Evans's score
Groenelved and Hitzeroth's score
Fain and Goldar's score
Kee's score
Lynch's score
Reed's score
Chakraborty's score
Schwartz's score

Table 16. Differences in mean AUC between the proportion of affected relatives and those FH scores with a better predictivity.

	SIMULATION																							
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
Reed	0.005	0.019	0.016	0.023	0.005	0.019	0.025	0.023	0.005	0.019	0.025	0.024	0.005	0.019	0.025	0.025	0.005	0.019	0.025	0.027	0.005	0.018	0.026	0.030
	0.003	0.01	0.01	0.007	0.003	0.01	0.01	0.007	0.003	0.011	0.011	0.01	0.003	0.011	0.013	0.011	0.003	0.011	0.014	0.014	0.003	0.012	0.017	0.019
FHS2	0.003	0.014	0.013	0.017	0.003	0.014	0.019	0.018	0.003	0.014	0.019	0.018	0.003	0.013	0.018	0.019	0.003	0.013	0.018	0.020	0.003	0.013	0.018	0.021
	0.001	0.005	0.005	0.004	0.001	0.004	0.005	0.004	0.001	0.004	0.005	0.005	0.001	0.005	0.006	0.005	0.001	0.004	0.006	0.006	0.001	0.004	0.006	0.007
Schwartz	0.004	0.011	-0.007	0.001	0.004	0.012	0.008	0.001	0.004	0.012	0.011	0.007	0.004	0.013	0.014	0.012	0.004	0.014	0.018	0.018	0.004	0.015	0.021	0.024
	0.003	0.006	0.001	-0.004	0.003	0.006	0.001	-0.004	0.002	0.006	0.004	0.001	0.002	0.007	0.006	0.004	0.002	0.008	0.01	0.01	0.002	0.009	0.014	0.016
SE	0.000	0.004	0.014	0.014	0.000	0.004	0.010	0.014	0.000	0.004	0.008	0.011	0.000	0.003	0.007	0.008	0.000	0.002	0.004	0.005	0.000	0.001	0.002	0.002
	0	0.002	0.003	0.003	0	0.002	0.003	0.003	0	0.001	0.002	0.003	0	0.001	0.002	0.002	0	0.001	0.001	0.001	0	0	0.001	0
Lynch	0.000	0.004	0.014	0.014	0.000	0.004	0.010	0.014	0.000	0.004	0.008	0.011	0.000	0.003	0.007	0.008	0.000	0.002	0.004	0.005	0.000	0.001	0.002	0.002
	0	0.002	0.003	0.003	0	0.002	0.003	0.003	0	0.001	0.002	0.003	0	0.001	0.002	0.002	0	0.001	0.001	0.001	0	0	0.001	0
FHS1	0.002	0.005	-0.003	0.001	0.002	0.005	0.004	0.001	0.002	0.006	0.005	0.004	0.002	0.006	0.007	0.006	0.002	0.007	0.009	0.010	0.002	0.008	0.012	0.014
	0	0	-0.001	-0.002	0	0	-0.002	-0.002	0	0	-0.001	0	0	0.001	0	-0.001	0	0	0.001	0.001	0	0.001	0.002	0.003
GH	0.001	0.004	-0.014	0.002	0.001	0.004	0.003	0.002	0.001	0.005	0.005	0.004	0.001	0.005	0.006	0.005	0.002	0.006	0.008	0.009	0.002	0.007	0.010	0.013
	0	0	-0.001	-0.002	0	0	-0.002	-0.002	0	0	-0.001	-0.001	0	0	0	-0.001	0	0	0.001	0.001	0	0.001	0.002	0.003

Abbreviations: FH, FH; GH, Groenelved and Hitzeroth; SE, Slanck and Evans.

Legend of Table 16: For each family history score, in the first row are displayed results for 24 different settings (according to different RRs and π) with expected risks of disease equal to 0.001, 0.01, 0.1 an 0.4 for the 4 subsequent age groups; in the second row are displayed results for 24 different settings (according to different RRs and π) with expected risks of disease equal to 0.04, 0.08, 0.12, 0.16. Simulation 1: π=0.0001 - RR=1.2; Simulation 2: π=0.0001 - RR=2.0; Simulation 3: π=0.0001 - RR=3.0; Simulation 4: π=0.0001 - RR=4.0; Simulation 5: π=0.001 - RR=1.2; Simulation 6: π=0.001 - RR=2.0; Simulation 7: π=0.001 - RR=3.0; Simulation 8: π=0.001 - RR=4.0; Simulation 9: π=0.05 - RR=1.2; Simulation 10: π=0.05 - RR=2.0; Simulation 11: π=0.05 - RR=3.0; Simulation 12: π=0.05 - RR=4.0; Simulation 13: π=0.1 - RR=1.2; Simulation 14: π=0.1 - RR=2.0; Simulation 15: π=0. 1 - RR=3.0; Simulation 16: π=0.1 - RR=4.0; Simulation 17: π=0.2 - RR=1.2; Simulation 18: π=0.2 - RR=2.0; Simulation 19: π=0.2 - RR=3.0; Simulation 20: π=0.2 - RR=4.0; Simulation 21: π=0.4 - RR=1.2; Simulation 22: π=0.4 - RR=2.0; Simulation 23: π=0.4 - RR=3.0; Simulation 24: π=0.4 - RR=4.0.

References

1. Ferlay J, Shin HR, Bray F et al. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. Int J Cancer 2010; 127: 2893-2917.

2. Bray F, Jemal A, Grey N et al. Global cancer transitions according to the Human Development Index (2008-2030): a population-based study. Lancet Oncol 2012; 13: 790-801.

3. World Cancer Research Fund and American Institute for Cancer Research. Food, Nutrition, Physical Activity and the Prevention of Cancer: a Global Perspective. Washington, DC: AICR 2007.

4. Peto J, Houlston RS. Genetics and the common cancers. Eur J Cancer 2001; 37 Suppl 8: S88-96.

5. Yoon PW, Scheuner MT, Peterson-Oehlke KL et al. Can family history be used as a tool for public health and preventive medicine? Genet Med 2002; 4: 304-310.

6. Pharoah PD, Day NE, Duffy S et al. Family history and the risk of breast cancer: a systematic review and meta-analysis. Int J Cancer 1997; 71: 800-809.

7. 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.

8. Turati F, Edefonti V, Talamini R et al. Family history of liver cancer and hepatocellular carcinoma. Hepatology 2012; 55: 1416-1425.

9. Baglietto L, Jenkins MA, Severi G et al. Measures of familial aggregation depend on definition of family history: meta-analysis for colorectal cancer. J Clin Epidemiol 2006; 59: 114-124.

10. Butterworth AS, Higgins JP, Pharoah P. Relative and absolute risk of colorectal cancer for individuals with a family history: a meta-analysis. Eur J Cancer 2006; 42: 216-227.

11. Clague J, Lin J, Cassidy A et al. Family history and risk of renal cell carcinoma: results from a case-control study and systematic meta-analysis. Cancer Epidemiol Biomarkers Prev 2009; 18: 801-807.

12. Lissowska J, Foretova L, Dabek J et al. Family history and lung cancer risk: international multicentre case-control study in Eastern and Central Europe and meta-analyses. Cancer Causes Control 2010; 21: 1091-1104.

13. Matakidou A, Eisen T, Houlston RS. Systematic review of the relationship between family history and lung cancer risk. Br J Cancer 2005; 93: 825-833.

14. Kicinski M, Vangronsveld J, Nawrot TS. An epidemiological reappraisal of the familial aggregation of prostate cancer: a meta-analysis. PLoS One 2011; 6: e27130.

15. Stratton JF, Pharoah P, Smith SK et al. A systematic review and meta-analysis of family history and risk of ovarian cancer. Br J Obstet Gynaecol 1998; 105: 493-499.

16. Kohlmann W, Gruber SB. Lynch Syndrome. 1993.

17. Varley JM, Evans DG, Birch JM. Li-Fraumeni syndrome--a molecular and clinical review. Br J Cancer 1997; 76: 1-14.

18. Ford D, Easton DF, Bishop DT et al. Risks of cancer in BRCA1-mutation carriers. Breast Cancer Linkage Consortium. Lancet 1994; 343: 692-695.

19. Cancer risks in BRCA2 mutation carriers. The Breast Cancer Linkage Consortium. J Natl Cancer Inst 1999; 91: 1310-1316.

20. Goldgar DE, Easton DF, Cannon-Albright LA, Skolnick MH. Systematic population-based assessment of cancer risk in first-degree relatives of cancer probands. J Natl Cancer Inst 1994; 86: 1600-1608.

21. Poole CA, Byers T, Calle EE et al. Influence of a family history of cancer within and across multiple sites on patterns of cancer mortality risk for women. Am J Epidemiol 1999; 149: 454-462.

22. Hemminki K, Vaittinen P. Familial cancer in Sweden. Int J Oncol 1997; 11: 273-280.

23. Amundadottir LT, Thorvaldsson S, Gudbjartsson DF et al. Cancer as a complex phenotype: pattern of cancer distribution within and beyond the nuclear family. PLoS Med 2004; 1: e65.

24. Chung CC, Magalhaes WC, Gonzalez-Bosquet J, Chanock SJ. Genome-wide association studies in cancer--current and future directions. Carcinogenesis 2010; 31: 111-120.

25. Fletcher O, Houlston RS. Architecture of inherited susceptibility to common cancer. Nat Rev Cancer 2010; 10: 353-361.

26. Garavello W, Turati F, Bosetti C et al. Family history of cancer and the risk of laryngeal cancer: a case-control study from Italy and Switzerland. Int J Cancer 2012; 130: 665-670.

27. Turati F, Edefonti V, Bravi F et al. Nutrient-based dietary patterns, family history, and colorectal cancer. Eur J Cancer Prev 2011; 20: 456-461.

28. Breslow NE, Day NE. Statistical methods in cancer research. Vol. II. The design and analysis of cohort studies. IARC Sci Publ No. 82. Lyon, France: IARC 1987.

29. Benjamini Y, Drai D, Elmer G et al. Controlling the false discovery rate in behavior genetics research. Behav Brain Res 2001; 125: 279-284.

30. Stroup DF, Berlin JA, Morton SC et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. JAMA 2000; 283: 2008-2012.

31. Hassan MM, Kaseb A, Li D et al. Association between hypothyroidism and hepatocellular carcinoma: a case-control study in the United States. Hepatology 2009; 49: 1563-1570.

32. London WT, Evans AA, McGlynn K et al. Viral, host and environmental risk factors for hepatocellular carcinoma: a prospective study in Haimen City, China. Intervirology 1995; 38: 155-161.

33. Huang YS, Chern HD, Wu JC et al. Polymorphism of the N-acetyltransferase 2 gene, red meat intake, and the susceptibility of hepatocellular carcinoma. Am J Gastroenterol 2003; 98: 1417-1422.

34. Yu MW, Yang SY, Chiu YH et al. A p53 genetic polymorphism as a modulator of hepatocellular carcinoma risk in relation to chronic liver disease, familial tendency, and cigarette smoking in hepatitis B carriers. Hepatology 1999; 29: 697-702.

35. Liu TT, Fang Y, Xiong H et al. A case-control study of the relationship between hepatitis B virus DNA level and risk of hepatocellular carcinoma in Qidong, China. World J Gastroenterol 2008; 14: 3059-3063.

36. Tanaka K, Hirohata T, Takeshita S et al. Hepatitis B virus, cigarette smoking and alcohol consumption in the development of hepatocellular carcinoma: a case-control study in Fukuoka, Japan. Int J Cancer 1992; 51: 509-514.

37. Tsukuma H, Hiyama T, Oshima A et al. A case-control study of hepatocellular carcinoma in Osaka, Japan. Int J Cancer 1990; 45: 231-236.

38. Chen CJ, Liang KY, Chang AS et al. Effects of hepatitis B virus, alcohol drinking, cigarette smoking and familial tendency on hepatocellular carcinoma. Hepatology 1991; 13: 398-406.

39. Hassan MM, Spitz MR, Thomas MB et al. The association of family history of liver cancer with hepatocellular carcinoma: a case-control study in the United States. J Hepatol 2009; 50: 334-341.

40. Donato F, Gelatti U, Chiesa R et al. A case-control study on family history of liver cancer as a risk factor for hepatocellular carcinoma in North Italy. Brescia HCC Study. Cancer Causes Control 1999; 10: 417-421.

41. Fernandez E, La Vecchia C, D'Avanzo B et al. Family history and the risk of liver, gallbladder, and pancreatic cancer. Cancer Epidemiol Biomarkers Prev 1994; 3: 209-212.

42. Zhu ZZ, Cong WM, Liu SF et al. A p53 polymorphism modifies the risk of hepatocellular carcinoma among non-carriers but not carriers of chronic hepatitis B virus infection. Cancer Lett 2005; 229: 77-83.

43. Hsu LM, Huang YS, Yang SY et al. Polymorphism of T-cell receptor gamma short tandem repeats as a susceptibility risk factor of hepatocellular carcinoma. Anticancer Res 2006; 26: 3787-3791.

44. Yu MW, Chang HC, Liaw YF et al. Familial risk of hepatocellular carcinoma among chronic hepatitis B carriers and their relatives. J Natl Cancer Inst 2000; 92: 1159-1164.

45. Evans AA, Chen G, Ross EA et al. Eight-year follow-up of the 90,000-person Haimen City cohort: I. Hepatocellular carcinoma mortality, risk factors, and gender differences. Cancer Epidemiol Biomarkers Prev 2002; 11: 369-376.

46. Sun Z, Lu P, Gail MH et al. Increased risk of hepatocellular carcinoma in male hepatitis B surface antigen carriers with chronic hepatitis who have detectable urinary aflatoxin metabolite M1. Hepatology 1999; 30: 379-383.

47. Chen TH, Chen CJ, Yen MF et al. Ultrasound screening and risk factors for death from hepatocellular carcinoma in a high risk group in Taiwan. Int J Cancer 2002; 98: 257-261.

48. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials 1986; 7: 177-188.

49. Greenland S. Quantitative methods in the review of epidemiologic literature. Epidemiol Rev 1987; 9: 1-30.

50. Goldstein AM, Blot WJ, Greenberg RS et al. Familial risk in oral and pharyngeal cancer. Eur J Cancer B Oral Oncol 1994; 30B: 319-322.

51. Brown LM, Gridley G, Diehl SR et al. Family cancer history and susceptibility to oral carcinoma in Puerto Rico. Cancer 2001; 92: 2102-2108.

52. Ankathil R, Mathew A, Joseph F, Nair MK. Is oral cancer susceptibility inherited? Report of five oral cancer families. Eur J Cancer B Oral Oncol 1996; 32B: 63-67.

53. Tashiro H, Abe K, Tanioka H. Familial occurrence of cancer of the mouth. J Oral Maxillofac Surg 1986; 44: 322-323.

54. Hara H, Ozeki S, Shiratsuchi Y et al. Familial occurrence of oral cancer: report of cases. J Oral Maxillofac Surg 1988; 46: 1098-1102.

55. La Vecchia C, Tavani A, Franceschi S et al. Epidemiology and prevention of oral cancer. Oral Oncol 1997; 33: 302-312.

56. Bondy ML, Spitz MR, Halabi S et al. Association between family history of cancer and mutagen sensitivity in upper aerodigestive tract cancer patients. Cancer Epidemiol Biomarkers Prev 1993; 2: 103-106.

57. Copper MP, Jovanovic A, Nauta JJ et al. Role of genetic factors in the etiology of squamous cell carcinoma of the head and neck. Arch Otolaryngol Head Neck Surg 1995; 121: 157-160.

58. Foulkes WD, Brunet JS, Kowalski LP et al. Family history of cancer is a risk factor for squamous cell carcinoma of the head and neck in Brazil: a case-control study. Int J Cancer 1995; 63: 769-773.

59. Foulkes WD, Brunet JS, Sieh W et al. Familial risks of squamous cell carcinoma of the head and neck: retrospective case-control study. BMJ 1996; 313: 716-721.

60. Yu GP, Zhang ZF, Hsu TC et al. Family history of cancer, mutagen sensitivity, and increased risk of head and neck cancer. Cancer Lett 1999; 146: 93-101.

61. Jefferies S, Eeles R, Goldgar D et al. The role of genetic factors in predisposition to squamous cell cancer of the head and neck. Br J Cancer 1999; 79: 865-867.

62. Mork J, Moller B, Glattre E. Familial risk in head and neck squamous cell carcinoma diagnosed before the age of 45: a population-based study. Oral Oncol 1999; 35: 360-367.

63. de Andrade M, Amos CI, Foulkes WD. Segregation analysis of squamous cell carcinoma of the head and neck: evidence for a major gene determining risk. Ann Hum Genet 1998; 62: 505-510.

64. Sturgis EM, Castillo EJ, Li L et al. Polymorphisms of DNA repair gene XRCC1 in squamous cell carcinoma of the head and neck. Carcinogenesis 1999; 20: 2125-2129.

65. Berwick M, Vineis P. Markers of DNA repair and susceptibility to cancer in humans: an epidemiologic review. J Natl Cancer Inst 2000; 92: 874-897.

66. Harty LC, Caporaso NE, Hayes RB et al. Alcohol dehydrogenase 3 genotype and risk of oral cavity and pharyngeal cancers. J Natl Cancer Inst 1997; 89: 1698-1705.

67. Park JY, Schantz SP, Stern JC et al. Association between glutathione S-transferase pi genetic polymorphisms and oral cancer risk. Pharmacogenetics 1999; 9: 497-504.

68. Jourenkova-Mironova N, Mitrunen K, Bouchardy C et al. High-activity microsomal epoxide hydrolase genotypes and the risk of oral, pharynx, and larynx cancers. Cancer Res 2000; 60: 534-536.

69. Olshan AF, Weissler MC, Watson MA, Bell DA. GSTM1, GSTT1, GSTP1, CYP1A1, and NAT1 polymorphisms, tobacco use, and the risk of head and neck cancer. Cancer Epidemiol Biomarkers Prev 2000; 9: 185-191.

70. Lewis SJ, Smith GD. Alcohol, ALDH2, and esophageal cancer: a meta-analysis which illustrates the potentials and limitations of a Mendelian randomization approach. Cancer Epidemiol Biomarkers Prev 2005; 14: 1967-1971.

71. Franceschi S, Talamini R, Barra S et al. Smoking and drinking in relation to cancers of the oral cavity, pharynx, larynx, and esophagus in northern Italy. Cancer Res 1990; 50: 6502-6507.

72. Pelucchi C, Gallus S, Garavello W et al. Cancer risk associated with alcohol and tobacco use: focus on upper aero-digestive tract and liver. Alcohol Res Health 2006; 29: 193-198.

73. Hsu WL, Yu KJ, Chien YC et al. Familial tendency and risk of nasopharyngeal carcinoma in taiwan: effects of covariates on risk. Am J Epidemiol 2011; 173: 292-299.

74. Chen DL, Huang TB. A case-control study of risk factors of nasopharyngeal carcinoma. Cancer Lett 1997; 117: 17-22.

75. Yu MC, Ho JH, Lai SH, Henderson BE. Cantonese-style salted fish as a cause of nasopharyngeal carcinoma: report of a case-control study in Hong Kong. Cancer Res 1986; 46: 956-961.

76. Chen CJ, Liang KY, Chang YS et al. Multiple risk factors of nasopharyngeal carcinoma: Epstein-Barr virus, malarial infection, cigarette smoking and familial tendency. Anticancer Res 1990; 10: 547-553.

77. Sriamporn S, Vatanasapt V, Pisani P et al. Environmental risk factors for nasopharyngeal carcinoma: a case-control study in northeastern Thailand. Cancer Epidemiol Biomarkers Prev 1992; 1: 345-348.

78. Yu MC, Garabrant DH, Huang TB, Henderson BE. Occupational and other non-dietary risk factors for nasopharyngeal carcinoma in Guangzhou, China. Int J Cancer 1990; 45: 1033-1039.

79. Yuan JM, Wang XL, Xiang YB et al. Non-dietary risk factors for nasopharyngeal carcinoma in Shanghai, China. Int J Cancer 2000; 85: 364-369.

80. Henderson BE, Louie E, SooHoo Jing J et al. Risk factors associated with nasopharyngeal carcinoma. N Engl J Med 1976; 295: 1101-1106.

81. Yan L, Xi Z, Drettner B. Epidemiological studies of nasopharyngeal cancer in the Guangzhou area, China. Preliminary report. Acta Otolaryngol 1989; 107: 424-427.

82. Ung A, Chen CJ, Levine PH et al. Familial and sporadic cases of nasopharyngeal carcinoma in Taiwan. Anticancer Res 1999; 19: 661-665.

83. Friborg J, Wohlfahrt J, Koch A et al. Cancer susceptibility in nasopharyngeal carcinoma families--a population-based cohort study. Cancer Res 2005; 65: 8567-8572.

84. Ji X, Zhang W, Xie C et al. Nasopharyngeal carcinoma risk by histologic type in central China: impact of smoking, alcohol and family history. Int J Cancer 2011; 129: 724-732.

85. Ren ZF, Liu WS, Qin HD et al. Effect of family history of cancers and environmental factors on risk of nasopharyngeal carcinoma in Guangdong, China. Cancer Epidemiol 2010; 34: 419-424.

86. Abdulamir AS, Hafidh RR, Abdulmuhaimen N et al. The distinctive profile of risk factors of nasopharyngeal carcinoma in comparison with other head and neck cancer types. BMC Public Health 2008; 8: 400.

87. Chang E, Adami H-O. Nasopharyngeal carcinoma. In Adami H-O, Hunter D, Trichpoulos D (eds): Cancer Epidemiology. 2nd edition, Edition New York: Oxford Universiy Press 2008; 175-195.

88. Ghadirian P. Familial history of esophageal cancer. Cancer 1985; 56: 2112-2116.

89. Li JY, Ershow AG, Chen ZJ et al. A case-control study of cancer of the esophagus and gastric cardia in Linxian. Int J Cancer 1989; 43: 755-761.

90. Hu N, Dawsey SM, Wu M et al. Familial aggregation of oesophageal cancer in Yangcheng County, Shanxi Province, China. Int J Epidemiol 1992; 21: 877-882.

91. Gao Y, Hu N, Han X et al. Family history of cancer and risk for esophageal and gastric cancer in Shanxi, China. BMC Cancer 2009; 9: 269.

92. Yu Y, Taylor PR, Li JY et al. Retrospective cohort study of risk-factors for esophageal cancer in Linxian, People's Republic of China. Cancer Causes Control 1993; 4: 195-202.

93. Guo W, Blot WJ, Li JY et al. A nested case-control study of oesophageal and stomach cancers in the Linxian nutrition intervention trial. Int J Epidemiol 1994; 23: 444-450.

94. Chang-Claude J, Becher H, Blettner M et al. Familial aggregation of oesophageal cancer in a high incidence area in China. Int J Epidemiol 1997; 26: 1159-1165.

95. Tran GD, Sun XD, Abnet CC et al. Prospective study of risk factors for esophageal and gastric cancers in the Linxian general population trial cohort in China. Int J Cancer 2005; 113: 456-463.

96. Brown LM, Blot WJ, Schuman SH et al. Environmental factors and high risk of esophageal cancer among men in coastal South Carolina. J Natl Cancer Inst 1988; 80: 1620-1625.

97. Dhillon PK, Farrow DC, Vaughan TL et al. Family history of cancer and risk of esophageal and gastric cancers in the United States. Int J Cancer 2001; 93: 148-152.

98. Lagergren J, Ye W, Lindgren A, Nyren O. Heredity and risk of cancer of the esophagus and gastric cardia. Cancer Epidemiol Biomarkers Prev 2000; 9: 757-760.

99. Morita M, Kuwano H, Nakashima T et al. Family aggregation of carcinoma of the hypopharynx and cervical esophagus: special reference to multiplicity of cancer in upper aerodigestive tract. Int J Cancer 1998; 76: 468-471.

100. Hemminki K, Jiang Y. Familial and second esophageal cancers: a nation-wide epidemiologic study from Sweden. Int J Cancer 2002; 98: 106-109.

101. Negri E, La Vecchia C, Levi F et al. Comparative descriptive epidemiology of oral and oesophageal cancers in Europe. Eur J Cancer Prev 1996; 5: 267-279.

102. Day GL, Blot WJ. Second primary tumors in patients with oral cancer. Cancer 1992; 70: 14-19.

103. Shimizu Y, Tsukagoshi H, Fujita M et al. Head and neck cancer arising after endoscopic mucosal resection for squamous cell carcinoma of the esophagus. Endoscopy 2003; 35: 322-326.

104. Bakir T, Can G, Erkul S, Siviloglu C. Stomach cancer history in the siblings of patients with gastric carcinoma. Eur J Cancer Prev 2000; 9: 401-408.

105. Chen MJ, Wu DC, Ko YC, Chiou YY. Personal history and family history as a predictor of gastric cardiac adenocarcinoma risk: a case-control study in Taiwan. Am J Gastroenterol 2004; 99: 1250-1257.

106. Hemminki K, Sundquist J, Ji J. Familial risk for gastric carcinoma: an updated study from Sweden. Br J Cancer 2007; 96: 1272-1277.

107. Ikeguchi M, Fukuda K, Oka S et al. Clinicopathological findings in patients with gastric adenocarcinoma with familial aggregation. Dig Surg 2001; 18: 439-443.

108. Inoue M, Ito LS, Tajima K et al. Height, weight, menstrual and reproductive factors and risk of gastric cancer among Japanese postmenopausal women: analysis by subsite and histologic subtype. Int J Cancer 2002; 97: 833-838.

109. Inoue M, Tajima K, Yamamura Y et al. Family history and subsite of gastric cancer: data from a case-referent study in Japan. Int J Cancer 1998; 76: 801-805.

110. Kerber RA, O'Brien E. A cohort study of cancer risk in relation to family histories of cancer in the Utah population database. Cancer 2005; 103: 1906-1915.

111. Kikuchi S, Nakajima T, Nishi T et al. Association between family history and gastric carcinoma among young adults. Jpn J Cancer Res 1996; 87: 332-336.

112. Kondo T, Toyoshima H, Tsuzuki Y et al. Aggregation of stomach cancer history in parents and offspring in comparison with other sites. Int J Epidemiol 2003; 32: 579-583.

113. La Vecchia C, Negri E, Franceschi S, Gentile A. Family history and the risk of stomach and colorectal cancer. Cancer 1992; 70: 50-55.

114. Lissowska J, Groves FD, Sobin LH et al. Family history and risk of stomach cancer in Warsaw, Poland. Eur J Cancer Prev 1999; 8: 223-227.

115. Munoz N, Plummer M, Vivas J et al. A case-control study of gastric cancer in Venezuela. Int J Cancer 2001; 93: 417-423.

116. Nagase H, Ogino K, Yoshida I et al. Family history-related risk of gastric cancer in Japan: a hospital-based case-control study. Jpn J Cancer Res 1996; 87: 1025-1028.

117. Palli D, Galli M, Caporaso NE et al. Family history and risk of stomach cancer in Italy. Cancer Epidemiol Biomarkers Prev 1994; 3: 15-18.

118. Toyoshima H, Hayashi S, Hashimoto S et al. Familial aggregation and covariation of diseases in a Japanese rural community: comparison of stomach cancer with other diseases. Ann Epidemiol 1997; 7: 446-451.

119. Yatsuya H, Toyoshima H, Mizoue T et al. Family history and the risk of stomach cancer death in Japan: differences by age and gender. Int J Cancer 2002; 97: 688-694.

120. Yu GP, Hsieh CC. Risk factors for stomach cancer: a population-based case-control study in Shanghai. Cancer Causes Control 1991; 2: 169-174.

121. Zanghieri G, Di Gregorio C, Sacchetti C et al. Familial occurrence of gastric cancer in the 2-year experience of a population-based registry. Cancer 1990; 66: 2047-2051.

122. Bamford KB, Bickley J, Collins JS et al. Helicobacter pylori: comparison of DNA fingerprints provides evidence for intrafamilial infection. Gut 1993; 34: 1348-1350.

123. Brenner H, Bode G, Boeing H. Helicobacter pylori infection among offspring of patients with stomach cancer. Gastroenterology 2000; 118: 31-35.

124. Dominici P, Bellentani S, Di Biase AR et al. Familial clustering of Helicobacter pylori infection: population based study. BMJ 1999; 319: 537-540.

125. Brenner H, Arndt V, Sturmer T et al. Individual and joint contribution of family history and Helicobacter pylori infection to the risk of gastric carcinoma. Cancer 2000; 88: 274-279.

126. Yatsuya H, Toyoshima H, Tamakoshi A et al. Individual and joint impact of family history and Helicobacter pylori infection on the risk of stomach cancer: a nested case-control study. Br J Cancer 2004; 91: 929-934.

127. Gayther SA, Gorringe KL, Ramus SJ et al. Identification of germ-line E-cadherin mutations in gastric cancer families of European origin. Cancer Res 1998; 58: 4086-4089.

128. El-Omar EM, Oien K, Murray LS et al. Increased prevalence of precancerous changes in relatives of gastric cancer patients: critical role of H. pylori. Gastroenterology 2000; 118: 22-30.

129. El-Omar EM, Carrington M, Chow WH et al. Interleukin-1 polymorphisms associated with increased risk of gastric cancer. Nature 2000; 404: 398-402.

130. Johns LE, Houlston RS. A systematic review and meta-analysis of familial colorectal cancer risk. Am J Gastroenterol 2001; 96: 2992-3003.

131. Lilla C, Verla-Tebit E, Risch A et al. Effect of NAT1 and NAT2 genetic polymorphisms on colorectal cancer risk associated with exposure to tobacco smoke and meat consumption. Cancer Epidemiol Biomarkers Prev 2006; 15: 99-107.

132. Goode EL, Ulrich CM, Potter JD. Polymorphisms in DNA repair genes and associations with cancer risk. Cancer Epidemiol Biomarkers Prev 2002; 11: 1513-1530.

133. Bigler J, Ulrich CM, Kawashima T et al. DNA repair polymorphisms and risk of colorectal adenomatous or hyperplastic polyps. Cancer Epidemiol Biomarkers Prev 2005; 14: 2501-2508.

134. Jin MJ, Chen K, Song L et al. The association of the DNA repair gene XRCC3 Thr241Met polymorphism with susceptibility to colorectal cancer in a Chinese population. Cancer Genet Cytogenet 2005; 163: 38-43.

135. Slattery ML, Sweeney C, Murtaugh M et al. Associations between ERalpha, ERbeta, and AR genotypes and colon and rectal cancer. Cancer Epidemiol Biomarkers Prev 2005; 14: 2936-2942.

136. Thompson D, Easton DF. Cancer Incidence in BRCA1 mutation carriers. J Natl Cancer Inst 2002; 94: 1358-1365.

137. Hata K, Shinozaki M, Toyoshima O et al. Impact of family history of gastric cancer on colorectal neoplasias in young japanese. Colorectal Dis 2012.

138. Lightfoot TJ, Skibola CF, Willett EV et al. Risk of non-Hodgkin lymphoma associated with polymorphisms in folate-metabolizing genes. Cancer Epidemiol Biomarkers Prev 2005; 14: 2999-3003.

139. Lordelo GS, Miranda-Vilela AL, Akimoto AK et al. Association between methylene tetrahydrofolate reductase and glutathione S-transferase M1 gene polymorphisms and chronic myeloid leukemia in a Brazilian population. Genet Mol Res 2012; 11: 1013-1026.

140. Azhar MR, Rahimi Z, Vaisi-Raygani A et al. Lack of association between MTHFR C677T and A1298C polymorphisms and risk of childhood acute lymphoblastic leukemia in the Kurdish population from Western Iran. Genet Test Mol Biomarkers 2012; 16: 198-202.

141. Kasperzyk JL, Chang ET, Birmann BM et al. Nutrients and genetic variation involved in one-carbon metabolism and Hodgkin lymphoma risk: a population-based case-control study. Am J Epidemiol 2011; 174: 816-827.

142. Chen CH, Huang GT, Lee HS et al. Clinical impact of screening first-degree relatives of patients with hepatocellular carcinoma. J Clin Gastroenterol 1998; 27: 236-239.

143. Mueller NE, Birmann BM, Parsonnet J et al. Infectious agents. In Schottenfeld D, Fraumeni JJ (eds): Cancer Epidemiology and Prevention. Third edition, Edition New York: Oxford Universiy Press 2006; 507-548.

144. Stuver S, Trichopoulos D. Cancer of the liver and biliary tract. In Adami H-O, Hunter D, Trichpoulos D (eds): Cancer Epidemiology. 2nd edition, Edition New York: Oxford Universiy Press 2008; 308-332.

145. Permuth-Wey J, Egan KM. Family history is a significant risk factor for pancreatic cancer: results from a systematic review and meta-analysis. Fam Cancer 2009; 8: 109-117.

146. Jacobs EJ, Chanock SJ, Fuchs CS et al. Family history of cancer and risk of pancreatic cancer: a pooled analysis from the Pancreatic Cancer Cohort Consortium (PanScan). Int J Cancer 2010; 127: 1421-1428.

147. Silverman DT, Schiffman M, Everhart J et al. Diabetes mellitus, other medical conditions and familial history of cancer as risk factors for pancreatic cancer. Br J Cancer 1999; 80: 1830-1837.

148. Jacobs EJ, Rodriguez C, Newton CC et al. Family history of various cancers and pancreatic cancer mortality in a large cohort. Cancer Causes Control 2009; 20: 1261-1269.

149. Raymond J, Thiberg JM, Chevalier C et al. Genetic and transmission analysis of Helicobacter pylori strains within a family. Emerg Infect Dis 2004; 10: 1816-1821.

150. Miyazaki M, Kato M, Takata T, Une H. Intrafamilial transmission of Helicobacter pylori: the association between a parent and an offspring with respect to the presence of anti-CagA antibody. J Infect Chemother 2002; 8: 70-75.

151. Gastric cancer and Helicobacter pylori: a combined analysis of 12 case control studies nested within prospective cohorts. Gut 2001; 49: 347-353.

152. Parkin DM. The global health burden of infection-associated cancers in the year 2002. Int J Cancer 2006; 118: 3030-3044.

153. Trikudanathan G, Philip A, Dasanu CA, Baker WL. Association between Helicobacter pylori infection and pancreatic cancer. A cumulative meta-analysis. JOP 2011; 12: 26-31.

154. Li XY, Guo X, Feng S et al. Relationship between a family history of malignancy and the incidence of laryngeal carcinoma in the Liaoning province of China. Clin Otolaryngol 2009; 34: 127-131.

155. Negri E, Boffetta P, Berthiller J et al. Family history of cancer: pooled analysis in the International Head and Neck Cancer Epidemiology Consortium. Int J Cancer 2009; 124: 394-401.

156. Jaworowska E, Serrano-Fernandez P, Tarnowska C et al. Familial association of laryngeal, lung, stomach and early-onset breast cancer. Breast Cancer Res Treat 2008; 112: 359-361.

157. Canova C, Hashibe M, Simonato L et al. Genetic associations of 115 polymorphisms with cancers of the upper aerodigestive tract across 10 European countries: the ARCAGE project. Cancer Res 2009; 69: 2956-2965.

158. Zhuo WL, Wang Y, Zhuo XL et al. Polymorphisms of CYP1A1 and GSTM1 and laryngeal cancer risk: evidence-based meta-analyses. J Cancer Res Clin Oncol 2009; 135: 1081-1090.

159. Bennett IC, Gattas M, Teh BT. The genetic basis of breast cancer and its clinical implications. Aust N Z J Surg 1999; 69: 95-105.

160. Lalloo F, Evans DG. Familial breast cancer. Clin Genet 2012; 82: 105-114.

161. Ford D, Easton DF. The genetics of breast and ovarian cancer. Br J Cancer 1995; 72: 805-812.

162. Willett W. The search for the causes of breast and colon cancer. Nature 1989; 338: 389-394.

163. Rauscher GH, Sandler DP, Poole C et al. Family history of cancer and incidence of acute leukemia in adults. Am J Epidemiol 2002; 156: 517-526.

164. Malkin D. p53 and the Li-Fraumeni syndrome. Cancer Genet Cytogenet 1993; 66: 83-92.

165. Goggins W, Gao W, Tsao H. Association between female breast cancer and cutaneous melanoma. Int J Cancer 2004; 111: 792-794.

166. Levi F, Te VC, Randimbison L, La Vecchia C. Cancer risk in women with previous breast cancer. Ann Oncol 2003; 14: 71-73.

167. Volk N, Pompe-Kirn V. Second primary cancers in breast cancer patients in Slovenia. Cancer Causes Control 1997; 8: 764-770.

168. Rubino C, de Vathaire F, Diallo I et al. Increased risk of second cancers following breast cancer: role of the initial treatment. Breast Cancer Res Treat 2000; 61: 183-195.

169. Wassberg C, Thorn M, Yuen J et al. Cancer risk in patients with earlier diagnosis of cutaneous melanoma in situ. Int J Cancer 1999; 83: 314-317.

170. Ghiorzo P, Ciotti P, Mantelli M et al. Characterization of ligurian melanoma families and risk of occurrence of other neoplasia. Int J Cancer 1999; 83: 441-448.

171. Borg A, Sandberg T, Nilsson K et al. High frequency of multiple melanomas and breast and pancreas carcinomas in CDKN2A mutation-positive melanoma families. J Natl Cancer Inst 2000; 92: 1260-1266.

172. Debniak T, Gorski B, Huzarski T et al. A common variant of CDKN2A (p16) predisposes to breast cancer. J Med Genet 2005; 42: 763-765.

173. Parazzini F, La Vecchia C, Moroni S et al. Family history and the risk of endometrial cancer. Int J Cancer 1994; 59: 460-462.

174. Gruber SB, Thompson WD. A population-based study of endometrial cancer and familial risk in younger women. Cancer and Steroid Hormone Study Group. Cancer Epidemiol Biomarkers Prev 1996; 5: 411-417.

175. Olson JE, Sellers TA, Anderson KE, Folsom AR. Does a family history of cancer increase the risk for postmenopausal endometrial carcinoma? A prospective cohort study and a nested case-control family study of older women. Cancer 1999; 85: 2444-2449.

176. Parslov M, Lidegaard O, Klintorp S et al. Risk factors among young women with endometrial cancer: a Danish case-control study. Am J Obstet Gynecol 2000; 182: 23-29.

177. Dong C, Hemminki K. Modification of cancer risks in offspring by sibling and parental cancers from 2,112,616 nuclear families. Int J Cancer 2001; 92: 144-150.

178. Kazerouni N, Schairer C, Friedman HB et al. Family history of breast cancer as a determinant of the risk of developing endometrial cancer: a nationwide cohort study. J Med Genet 2002; 39: 826-832.

179. Seger HM, Soisson AP, Dodson MK et al. Familial clustering of endometrial cancer in a well-defined population. Gynecol Oncol 2011; 122: 75-78.

180. Parazzini F, La Vecchia C, Bocciolone L, Franceschi S. The epidemiology of endometrial cancer. Gynecol Oncol 1991; 41: 1-16.

181. Brinton LA. Epidemiology of cervical cancer--overview. IARC Sci Publ 1992; 3-23.

182. Koch M, Gaedke H, Jenkins H. Family history of ovarian cancer patients: a case-control study. Int J Epidemiol 1989; 18: 782-785.

183. Parazzini F, Negri E, La Vecchia C et al. Family history of reproductive cancers and ovarian cancer risk: an Italian case-control study. Am J Epidemiol 1992; 135: 35-40.

184. Houlston RS, Bourne TH, Collins WP et al. Risk of ovarian cancer and genetic relationship to other cancers in families. Hum Hered 1993; 43: 111-115.

185. Kerber RA, Slattery ML. The impact of family history on ovarian cancer risk. The Utah Population Database. Arch Intern Med 1995; 155: 905-912.

186. Auranen A, Pukkala E, Makinen J et al. Cancer incidence in the first-degree relatives of ovarian cancer patients. Br J Cancer 1996; 74: 280-284.

187. Easton DF, Matthews FE, Ford D et al. Cancer mortality in relatives of women with ovarian cancer: the OPCS Study. Office of Population Censuses and Surveys. Int J Cancer 1996; 65: 284-294.

188. Rader JS, Neuman RJ, Brady J et al. Cancer among first-degree relatives of probands with invasive and borderline ovarian cancer. Obstet Gynecol 1998; 92: 589-595.

189. Stratton JF, Thompson D, Bobrow L et al. The genetic epidemiology of early-onset epithelial ovarian cancer: a population-based study. Am J Hum Genet 1999; 65: 1725-1732.

190. Ziogas A, Gildea M, Cohen P et al. Cancer risk estimates for family members of a population-based family registry for breast and ovarian cancer. Cancer Epidemiol Biomarkers Prev 2000; 9: 103-111.

191. Hemminki K, Sundquist J, Brandt A. Incidence and mortality in epithelial ovarian cancer by family history of any cancer. Cancer 2011; 117: 3972-3980.

192. Lorenzo Bermejo J, Rawal R, Hemminki K. Familial association of specific histologic types of ovarian malignancy with other malignancies. Cancer 2004; 100: 1507-1514.

193. Anderson H, Bladstrom A, Olsson H, Moller TR. Familial breast and ovarian cancer: a Swedish population-based register study. Am J Epidemiol 2000; 152: 1154-1163.

194. Stratton JF, Gayther SA, Russell P et al. Contribution of BRCA1 mutations to ovarian cancer. N Engl J Med 1997; 336: 1125-1130.

195. Zaino R, Whitney C, Brady MF et al. Simultaneously detected endometrial and ovarian carcinomas--a prospective clinicopathologic study of 74 cases: a gynecologic oncology group study. Gynecol Oncol 2001; 83: 355-362.

196. Johns LE, Houlston RS. A systematic review and meta-analysis of familial prostate cancer risk. BJU Int 2003; 91: 789-794.

197. Bruner DW, Moore D, Parlanti A et al. Relative risk of prostate cancer for men with affected relatives: systematic review and meta-analysis. Int J Cancer 2003; 107: 797-803.

198. Zeegers MP, Jellema A, Ostrer H. Empiric risk of prostate carcinoma for relatives of patients with prostate carcinoma: a meta-analysis. Cancer 2003; 97: 1894-1903.

199. Edwards SM, Kote-Jarai Z, Meitz J et al. Two percent of men with early-onset prostate cancer harbor germline mutations in the BRCA2 gene. Am J Hum Genet 2003; 72: 1-12.

200. Brandt A, Sundquist J, Hemminki K. Risk for incident and fatal prostate cancer in men with a family history of any incident and fatal cancer. Ann Oncol 2012; 23: 251-256.

201. Hemminki K, Chen B. Familial association of prostate cancer with other cancers in the Swedish Family-Cancer Database. Prostate 2005; 65: 188-194.

202. Gronberg H, Bergh A, Damber JE, Emanuelsson M. Cancer risk in families with hereditary prostate carcinoma. Cancer 2000; 89: 1315-1321.

203. Matikaine MP, Pukkala E, Schleutker J et al. Relatives of prostate cancer patients have an increased risk of prostate and stomach cancers: a population-based, cancer registry study in Finland. Cancer Causes Control 2001; 12: 223-230.

204. Damber L, Gronberg H, Damber JE. Familial prostate cancer and possible associated malignancies: nation-wide register cohort study in Sweden. Int J Cancer 1998; 78: 293-297.

205. Hemminki K, Dong C. Familial prostate cancer from the family-cancer database. Eur J Cancer 2000; 36: 229-234.

206. Valeri A, Fournier G, Morin V et al. Early onset and familial predisposition to prostate cancer significantly enhance the probability for breast cancer in first degree relatives. Int J Cancer 2000; 86: 883-887.

207. Mori M, Masumori N, Fukuta F et al. Weight gain and family history of prostate or breast cancers as risk factors for prostate cancer: results of a case-control study in Japan. Asian Pac J Cancer Prev 2011; 12: 743-747.

208. Chen YC, Page JH, Chen R, Giovannucci E. Family history of prostate and breast cancer and the risk of prostate cancer in the PSA era. Prostate 2008; 68: 1582-1591.

209. Eldon BJ, Jonsson E, Tomasson J et al. Familial risk of prostate cancer in Iceland. BJU Int 2003; 92: 915-919.

210. Isaacs SD, Kiemeney LA, Baffoe-Bonnie A et al. Risk of cancer in relatives of prostate cancer probands. J Natl Cancer Inst 1995; 87: 991-996.

211. Bratt O, Kristoffersson U, Lundgren R, Olsson H. The risk of malignant tumours in first-degree relatives of men with early onset prostate cancer: a population-based cohort study. Eur J Cancer 1997; 33: 2237-2240.

212. Verhage BA, Aben KK, Witjes JA et al. Site-specific familial aggregation of prostate cancer. Int J Cancer 2004; 109: 611-617.

213. Ahn J, Moslehi R, Weinstein SJ et al. Family history of prostate cancer and prostate cancer risk in the Alpha-Tocopherol, Beta-Carotene Cancer Prevention (ATBC) Study. Int J Cancer 2008; 123: 1154-1159.

214. Kreiger N, Marrett LD, Dodds L et al. Risk factors for renal cell carcinoma: results of a population-based case-control study. Cancer Causes Control 1993; 4: 101-110.

215. Mellemgaard A, Engholm G, McLaughlin JK, Olsen JH. Risk factors for renal cell carcinoma in Denmark. I. Role of socioeconomic status, tobacco use, beverages, and family history. Cancer Causes Control 1994; 5: 105-113.

216. Schlehofer B, Pommer W, Mellemgaard A et al. International renal-cell-cancer study. VI. the role of medical and family history. Int J Cancer 1996; 66: 723-726.

217. Gago-Dominguez M, Yuan JM, Castelao JE et al. Family history and risk of renal cell carcinoma. Cancer Epidemiol Biomarkers Prev 2001; 10: 1001-1004.

218. Czene K, Hemminki K. Kidney cancer in the Swedish Family Cancer Database: familial risks and second primary malignancies. Kidney Int 2002; 61: 1806-1813.

219. Gudbjartsson T, Jonasdottir TJ, Thoroddsen A et al. A population-based familial aggregation analysis indicates genetic contribution in a majority of renal cell carcinomas. Int J Cancer 2002; 100: 476-479.

220. Hung RJ, Moore L, Boffetta P et al. Family history and the risk of kidney cancer: a multicenter case-control study in Central Europe. Cancer Epidemiol Biomarkers Prev 2007; 16: 1287-1290.

221. Iliopoulos O, Eng C. Genetic and clinical aspects of familial renal neoplasms. Semin Oncol 2000; 27: 138-149.

222. Pavlovich CP, Schmidt LS. Searching for the hereditary causes of renal-cell carcinoma. Nat Rev Cancer 2004; 4: 381-393.

223. Freedman RA, Garber JE. Family cancer history: healthy skepticism required. J Natl Cancer Inst 2011; 103: 776-777.

224. Bravi F, Bosetti C, Negri E et al. Family history of cancer provided by hospital controls was satisfactorily reliable. J Clin Epidemiol 2007; 60: 171-175.

225. Qureshi N, Wilson B, Santaguida P et al. Collection and Use of Cancer Family History in Primary Care. Evidence Report/Technology Assessment No. 159 (prepared by the McMaster University Evidence-based Practice Center, under Contract No. 290-02-0020). AHRQ Publication No. 08-E001. Rockville, MD: Agency for Healthcare Research and Quality. October 2007.

226. Mai PL, Garceau AO, Graubard BI et al. Confirmation of family cancer history reported in a population-based survey. J Natl Cancer Inst 2011; 103: 788-797.

227. van Esch A, Steyerberg EW, Berger MY et al. Family history and recurrence of febrile seizures. Arch Dis Child 1994; 70: 395-399.

228. Slack J, Evans KA. The increased risk of death from ischaemic heart disease in first degree relatives of 121 men and 96 women with ischaemic heart disease. J Med Genet 1966; 3: 239-257.

229. Lynch HT, Kimberling WJ, Biscone KA et al. Familial heterogeneity of colon cancer risk. Cancer 1986; 57: 2089-2096.

230. Chakraborty R, Weiss KM, Majumder PP et al. A method to detect excess risk of disease in structured data: cancer in relatives of retinoblastoma patients. Genet Epidemiol 1984; 1: 229-244.

231. Fain PR, Goldgar DE. A nonparametric test of heterogeneity of family risk. Genet Epidemiol Suppl 1986; 1: 61-66.

232. Groeneveld H, Hitzeroth H. Quantification of familial predisposition for disease conditions. S Afr Statist 1991; 45-60.

233. Kee F, Collins BJ, Patterson CC. Prognosis in familial non-polyposis colorectal cancer. Gut 1991; 32: 513-516.

234. Schwartz AG, Boehnke M, Moll PP. Family risk index as a measure of familial heterogeneity of cancer risk. A population-based study in metropolitan Detroit. Am J Epidemiol 1988; 128: 524-535.

235. Silberberg J, Fryer J, Wlodarczyk J et al. Comparison of family history measures used to identify high risk of coronary heart disease. Genet Epidemiol 1999; 16: 344-355.

236. Reed T, Wagener DK, Donahue RP, Kuller LH. Family history of cancer related to cholesterol level in young adults. Genet Epidemiol 1986; 3: 63-71.

237. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 1988; 44: 837-845.