Department of Medical Epidemiology and Biostatistics
Karolinska Institutet, Stockholm, Sweden

Population-based family studies: Genetic contribution to cancer development and survival?

Linda Sofie Lindström

Stockholm 2008

Published by Karolinska Institutet.
Printed by Larserics Digital Print AB

Linda.Lindstrom@ki.se
ISBN 978-91-7409-196-0

Till min älskade familj

Ronny
min bästa vän & livs kärlek
lycklig är jag att med dig genom livet få vandra

Erik & Emma
ni mig om livet lär, tydlig nu är livets mening
stoltare & lyckligare kan ingen mamma vara

Nu och för evigt finns ni i det allra innersta av mitt hjärta

ABSTRACT

Cancer affects essentially everyone, directly or indirectly. The aim of this thesis was to study the genetic and environmental factors in cancer development and survival. Our studies were based on a record linkage between several Swedish population-based registries, principally the Multi-Generation Register, which records familial relationships, the Swedish Cancer Registry, and the Cause of Death Registry. In summary, the Swedish Family-Cancer database comprised over 11 million individuals organized into around three million families, including more than one million cancers.

In study I, we developed a generalized linear mixed model, enabling analyses of genetic and environmental effects in two- and three-generational families, considering all relationships in a family at once avoiding pairwise analyses of dependent family members. The two- and three- generational family design resulted in similar genetic and environmental estimates. In the two-generational families, no noteworthy differences were observed correcting for the unequal follow-up time in parents and children. Further, in our second study, the genetic contribution to melanoma was estimated at 18% (95% Confidence Interval [CI]=13% to 22%) in the analysis of all body sites. Contrasting the family-shared environment in sun-covered and sun-exposed body sites, the contribution was higher in covered sites, possibly conveying the benefit from cautious sunbathing on sensitive skin. The estimated childhood-shared environment for both melanoma and squamous cell carcinoma of the skin (SCC) elucidated the impact of sun habits and the avoidance of risk inflicted sunburns during infancy and youth. Moreover, in SCC, the familial shared environment at 18% (95% CI=16% to 19%) is important in defining the susceptibility to the disease. Genetic variability in individuals enhancing sensitivity to accumulated sun exposure will most probably also be involved in the aetiology. Finally, we propose that genetic factors are vital in the common liability to both melanoma and SCC. We estimated that 47% (95% CI=43% to 51%) of the susceptibility was estimated to be attributed to genetic factors.

It has been established that genetic variability influences the susceptibility to cancer; still little is known about the inheritance of cancer survival. In study III and IV, we present the first population-based comprehensive analyses of cancer survival concordance among family members. In study III, we noted a significantly increased risk of poor survival in children with poor parental survival compared to the risk in children with good parental survival in colorectal cancer (Hazard Ratio [HR]=1.44, 95% CI=1.01 to 2.01), lung cancer (HR=1.39, 95% CI=1.00 to 1.94), breast cancer (HR=1.75, 95% CI=1.13 to 2.71), ovarian cancer (HR=2.23, 95% CI=0.78 to 6.34) and prostate cancer (HR=2.07, 95% CI=1.13 to 3.79). All hazard ratio estimates, except for ovarian cancer, were statistically significant with trends of increasing risk of death among offspring by degree of survival outcome among parents. In study IV, lung cancer survival in children was associated with the lung cancer survival in their parents with a decreased hazard ratio for death in children with good parental survival (Hazard Ratio [HR]=0.71, 95% CI=0.51 to 0.99), compared to those with poor parental survival. We also found a strong protective effect (HR=0.14, 95% CI=0.030 to 0.65) for siblings, while no effect was seen on spouse survival. Genetic background of an individual may be more important than lifestyle factors such as smoking in lung cancer survival. The very strong protective effect in siblings compared to parent-child pairs further suggests a possible recessive pathway of inheritance. In light of study III and IV, we propose that genetic background is of importance in foreseeing an individual's cancer specific survival.

In conclusion, genetic factors are vital in the familial aggregation of melanoma in addition to the co-aggregation of melanoma and SCC. The ability to fight cancer disease and survive may also be inherited. In the future, I envision that population-based studies will help in identification of genetic variation influencing both the liability to cancer disease development and subsequent survival.

LIST OF PUBLICATIONS

This thesis is based on the following papers, which will be referred to in the text by their Roman numerals (I-IV):

I. Lindström LS, Pawitan Y, Reilly M, Hemminki K, Lichtenstein P, Czene K.
Estimation of genetic and environmental factors for melanoma onset using population based family data.
Statistics in Medicine 2006;25:3110-23.

II. Lindström LS, Yip B, Lichtenstein P, Pawitan Y, and Czene K.
Etiology of familial aggregation in melanoma and squamous cell carcinoma of the skin
Cancer Epidemiology Biomarkers & Prevention 2007;16(8):1639-43.

III. Lindström LS, Hall P, Hartman M, Wiklund F, Grönberg H, Czene K.
Familial concordance in cancer survival: a Swedish population-based study.
Lancet Oncology 2007;8:1001-06.

IV. Lindström LS, Hall P, Hartman M, Wiklund F and Czene K.
A genetic background to lung cancer survival?
Manuscript submitted.

TABLE OF CONTENTS

ABBREVIATIONS

CI	Confidence Interval
FAP	Familial Adenomatous Polyposis
GLMM	Generalized Linear Mixed Model
HNPCC	Hereditary NonPolyposis Colorectal Cancer
HPV	Human Papilloma Virus
HR	Hazards Ratio
ICD	International Classification of Diseases
PSA	Prostate Specific Antigen
SCC	Squamous Cell Carcinoma of the skin
SNP	Single Nucleotide Polymorphism

INTRODUCTION

In history, there has been increasing support for the idea that hereditary factors are involved in the development of cancer. The earliest reports of aggregation of cancer in families dates back to breast cancer in the wife and family of the French physician Broca and gastric cancer in Napoleon Bonaparte's family.[1] In the 19th century, an early childhood tumour to the eye, retinoblastoma, was recognised to cluster in families. This phenomenon puzzled Alfred Knudson and by 1971 based upon the clinical records of 48 retinoblastoma patients, he concluded that individuals inheriting a mutation in the retinoblastoma gene will only need one additional mutation and are thus far more likely to develop the malignancy earlier and bilaterally in contrast to sporadic cases who require two mutational hits in retinal cells.[2] The gene for retinoblastoma was the first hereditary cancer gene to be discovered, starting the molecular genetic era of modern research.[3,4]

The study of extended families has been important through the history of cancer research, mainly in the study of Mendelian diseases. This approach is still today vital in for instance understanding the aetiology of cancers, in the identification of cancer related genes and in clinical decisions and counselling. Certain limitations are however posed on the use of families in the study of cancer as a common and complex disease. Traditionally, cancer research was based on studies of family pedigrees with clear inheritance of cancer from generation to generation. Very large pedigrees with information on relatives could be retrieved by interviews of family members in several generations. These studies yield good results when the interest is set on high penetrance Mendelian inheritance genes. However, extended families are less likely to give insights to the general and complex disease nature of cancer, in a very limited set of individuals. Also, the recollection of relatives often stops at individuals' second or third cousins along with their disease status. Now, the majority of genes segregating in a high-penetrant Mendelian fashion may have been discovered leaving the high risk genes as more narrow examples of the inheritance of cancer disease. The focus may shift to the discovery of particular genetic dysfunctions involving heterogeneous expression in multiple genes and functional pathways, guided by indication of genetic importance in addition to environmental effects on cancer susceptibility. Consequently, advancement in one field of study will expand our knowledge of other intrinsically related disciplines, see figure 1.

It is only in the last twenty years we have started to identify abnormalities in cancer-predisposing genes. However, these abnormalities still today explain no more than a small proportion of the observed familial aggregation of cancer.[5-7] Cancer research has only entered a fascinating epoch leaving the vital findings in understanding the true nature of cancer still to be unravelled. In the

future, rather than traditional high-risk family studies, larger population-based family studies may be fundamental in determining individual cancer susceptibility and ability to fight cancer disease.

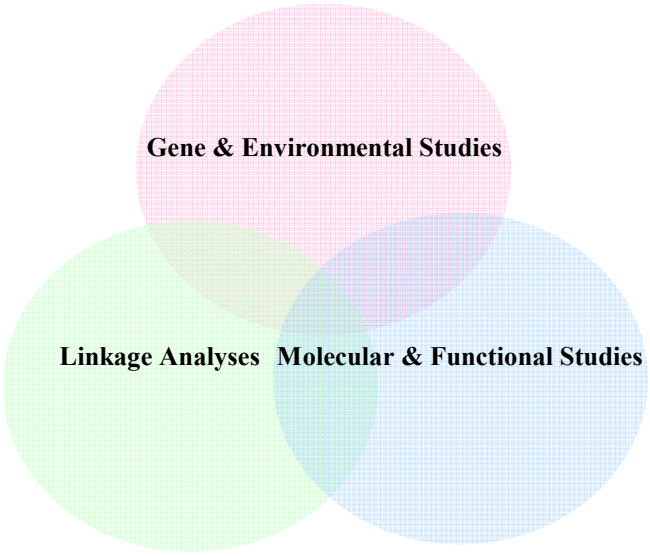


Figure 1 - Studies on genes & environment, molecular & functional traits and linkage analyses are inherently related

CANCER DEVELOPMENT

Neoplasia is originally a Greek word denoting the growth of new tissue, inhabiting cells with growing ability beyond their normal boundaries, see figure 2. Hence, the balance between proliferation and programmed cell death is disturbed, allowing cells with genetic abnormalities to survive. Neoplasias may be both benign (localized) or malignant (invading host tissues). Nonetheless, some benign neoplasias are precancerous lesions and develop over time into invasive cancers. Cancer research still attempts to understand the genetic basis of cancer to explain its progressive nature, but it is now beyond all dispute that cancer is a genetic disease at the cellular level.

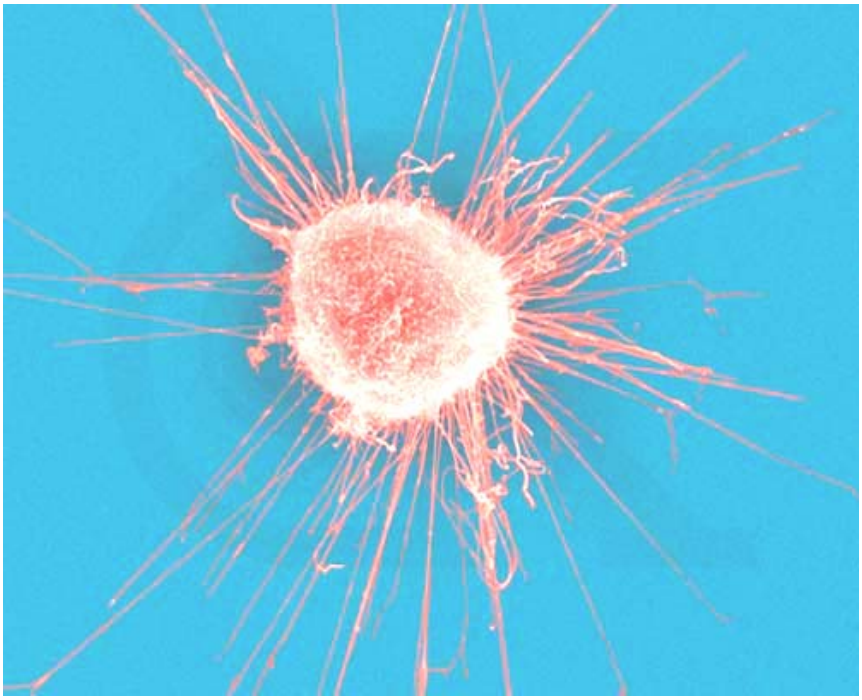


Figure 2 - A breast cancer cell

In Sweden, the most common cancers are prostate and breast cancer contributing to 35% and 29% of all cancers in men and women, respectively, see figure 3.[8] Skin cancer (excluding melanoma) and colorectal cancer are the second and third most common cancers accounting for around 15% of all cancer cases. During the last decade melanoma has increased by 3% annually and the yearly increase of other skin cancers has been even higher. The underlying reason may at least partly be due to changed sun-tanning and vacation habits. However, an increase has also been seen for both prostate (3% yearly) and breast cancer (1.3% yearly) during the last decades, which may be partly explained by enhanced diagnostic methods such as mammography screening and prostate specific antigen testing (PSA) in Sweden. In women lung cancer has increased by 3% yearly the last twenty years and almost 4% in the last decade, whereas the trend is the

opposite in men with a yearly decrease of 1%. It has been suggested that the increasing smoking in women since the 1950s in Sweden explain most of the increase.[9]

Percent of all cancer cases	
18	Prostate
14	Breast
8	Skin
7	Colon
6	Lung
5	Urinary bladder
5	Melanoma
4	Rectal
3	Malignant lymphoma
3	Cervical

Percent of all cancer cases	
35	Prostate
9	Skin
7	Colon
7	Urinary bladder
6	Lung
5	Rectal
4	Melanoma
3	Malignant lymphoma
2	Kidney
2	Leukemia

Percent of all cancer cases	
29	Breast
8	Colon
8	Skin
6	Lung
6	Cervical
5	Melanoma
4	Rectal
3	Ovarian
3	Brain and nerve system
3	Malignant lymphoma

Figure 3 - The ten most common cancers divided by gender

FAMILIAL CANCER

Studies of familial cancer have been the main approach in the assessment of the hereditary effects in cancer. Population-based family studies provide reliable quantitative estimates on familial risks once family relationships and cancers in relatives have been confirmed.[10-15]

Familial aggregation can be described by familial risk and in population-based family studies an individual with an affected parent or sibling develop the specific cancer twice as often as compared to individuals without family history, see table 1.[16] Moreover, individuals with two or more affected first-degree relatives are even more susceptible to cancer. In breast cancer for instance, these individuals are three times more likely to develop the disease.[17]

Table 1 - Familial risk of cancer

Risk in offspring by	Parent concordant cancer			Sibling concordant cancer		
Site or type	**SIR**	**95% CI**		**SIR**	**95% CI**	
Colorectal[a]	1.86	1.73	2.01	2.87	2.40	3.40
Lung	2.09	1.88	2.32	3.13	2.53	3.84
Breast	1.84	1.76	1.93	2.03	1.89	2.17
Ovary	3.15	2.56	3.85	4.25	3.01	5.84
Prostate	2.45	2.30	2.62	4.46	3.85	5.15
Melanoma	2.62	2.23	3.05	2.93	2.38	3.57
Skin	2.52	1.99	3.15	3.63	1.87	6.37

[a]Including only adenocarcinoma

Some cancers co-aggregate in families unravelling hereditary cancer syndromes, while for others the common aetiology is still unknown. For instance, melanoma and squamous cell carcinoma of the skin (SCC) have been shown to co-aggregate within families.[18,19] The reason for this familial aggregation and relatively high familial risk is still unknown but a shared sensitivity to ultraviolet radiation has been suggested to be the main risk factor for both diseases.[20]

Second or multiple cancers in the same individual may also signal an inherited susceptibility. For instance in Li-Fraumeni syndrome, primary breast cancer tumours often develops followed by secondary brain tumours or leukaemia. Further, secondary endometrial cancers are commonly seen in families with HNPCC (hereditary nonpolyposis colorectal cancer) and in familial breast cancer syndrome families, ovarian and pancreatic cancer are known to co-aggregate as second cancers.[21-25] However, second or multiple cancers in the same individual may also signal sensitivity to specific environmental factors, such as ultraviolet radiation in skin cancers or susceptibility to carcinogens in smoking or immunological factors. Thus, secondary cancers offer interesting insights into the risk of cancer. However, the aetiology of secondary cancers may also include effects of radio- and chemotherapy, or an increased probability to find a new tumour due to intensive medical surveillance after the first diagnosis.[23,24,26]

Studies on familial risk of cancer give reliable estimates on familial susceptibility, even though since no more than one order of genetic relationship are exploited such studies are less advantageous in the dissection of the effects of shared genes from the effects of shared environment. Family members share many environmental factors, such as lifestyle, including diet and habits, which may increase or decrease exposures to cancer-causing or protective factors.

GENES VERSUS ENVIRONMENT

Cancer is caused by environmental and genetic events, see figure 4, which transform a benign cell to a malignant cell through a sequence of molecular changes. In the human cell a portion of the genes regulate cell growth and cell division. Mutations altering the expression and function of these genes in cells may lead to cancer. These mutations can either be inherited, occur spontaneously or be due to environmental exposures (carcinogens, physical mutagens, certain

viruses, or by epigenetic events). Even without the environmental hazardous factors, spontaneous mutations will occur because of limitations of the accuracy of the DNA replication and DNA repair. Still, all these mutations and those caused by environmental factors together occur much more frequently than do cancer in humans.

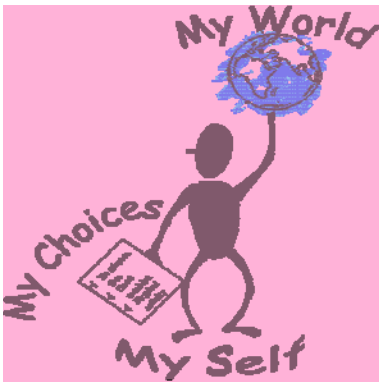


Figure 4 - Myself (genetic factors), my choices (lifestyle and behaviour), and my world (physical environment in which you live)

Some individuals inherit a mutation in a cancer related gene. In their families the number of additional mutations required to develop cancer are fewer, resulting in cancer occurring with a higher frequency and on average at an earlier age.[27] These families are more susceptible to a certain type of cancer and identification of these susceptibility genes have led to a deeper understanding of the carcinogenic process. Today 1-5 % of all cancers are believed to be associated with single-gene, dominant traits.[28] In addition, accumulated data now point at cancer being a complex disease, where the causal pathway may be several genes which each have a minor importance and hence account for a larger part of all cancer cases.[5,7,29] However, family studies have suggested that life-style related factors may explain at least a small part of the familial aggregation observed in some cancers.[12,13]

One strategy to unravel the underlying reasons of familial aggregation of a certain cancer is to disentangle the genetic and environmental factors by use of population-based family data of different degrees of relationships. Quantitative genetic methods are well developed for twin studies because of the innate simplicity comparing the similarity among monozygotic twins and dizygotic twins. Due to the rarity of twinning (1-2%), the sample size is a common problem especially when studying rare diseases such as cancer. Even in a large study disentangling genetic and environmental factors, combining twins from Swedish, Finnish, and Danish registers, significant heritability estimates were only obtained for colorectal, breast and prostate cancer, see table 2.[30]

Table 2 - Effects of genetic and environmental factors, twin study

Site or type	Percentage of variance (95% CI)	
	Genetic	Family-shared environment
Colorectal	35 (10-48)	5 (0-23)
Breast	27 (4-41)	6 (0-22)
Prostate	42(29-50)	0 (0-9)

Families are in contrast to twins more abundantly available and population-based family studies have previously estimated the genetic and environmental contribution to cancer susceptibility, see table 3.[31]

Family studies have inherent methodological concerns related to the dependence among individuals in a family. To enable analysis, individuals have been analysed in a pairwise manner assuming independence between different relative pairs in a family such as sibling-sibling and parent-offspring pairs.

Table 3 - Effects of genetic and environmental factors in cancer, family study

Site or type	Percentage of variance (95% CI)		
	Genetic	Family-shared environment	Childhood-shared environment
Stomach	1 (1-6)	15 (15-16)	13 (12-16)
Colon	13 (12-18)	12 (11-13)	6 (5-7)
Rectum	12 (8-13)	9 (9-9)	3 (3-5)
Lung	8 (5-9)	9 (8-9)	4 (0-4)
Breast	25 (23-27)	9 (9-9)	6 (6-6)
Cervix invasive	22 (14-27)	0 (0-12)	3 (0-8)
Cervix in situ	13 (6-15)	0 (0-3)	13 (12-14)
Testis	25 (15-37)	0 (0-0)	17 (8-24)
Kidney	8 (7-9)	8 (8-9)	6 (6-7)
Urinary bladder	7 (2-11)	12 (9-12)	4 (4-12)
Melanoma	21 (12-23)	2 (1-2)	8 (7-10)
Nervous system	12 (10-18)	5 (4-7)	3 (0-6)
Nervous system: Age >15 yrs	13 (6-20)	5 (2-8)	2 (0-6)
Thyroid	53 (52-53)	1 (1-2)	10 (9-11)
Endocrine glands	28 (27-28)	3 (3-3)	11 (11-11)
Non-Hodgkin lymphoma	10 (8-10)	6 (5-6)	2 (2-4)
Leukemia	1 (0-1)	8 (5-8)	4 (4-5)
Leukemia: Age >15 yrs	9 (9-16)	9 (6-9)	4 (4-7)

ENVIRONMENTAL RISK FACTORS

In the 1960s, cancer was generally viewed as being caused by environmental and lifestyle factors and thus preventable.[32] Today, shared environmental factors in a family have been observed to slightly increase the predisposition to many cancers, see table 2. The understanding on what the 'environment' actually is and on the mechanisms through which the 'environment' is able to cause cancer are still truly fragmentary.[33] Some established environmental carcinogens described below are genuinely vital in influencing cancer susceptibility.

The carcinogenic effect of tobacco is possibly the most important discovery in the history of cancer epidemiology.[33] In Sweden, an estimated 8000 to 10 000 new cancer cases per year are caused by smoking and the majority of lung cancer patients smoke.[9] Smoking also increases the risk of other cancers such as cancer in the pancreas, urinary bladder, kidney, liver, stomach and most probable also cancer to the cervix.[34-36]

In the last twenty years, the discovery that infectious pathogens are potentially carcinogenic greatly enhanced the understanding of the fundamentals of cancer. For instance, a gastric bacteria called *helicobacter pylori* known to cause ulcers, was found to be a major risk factor in the development of stomach cancer.[33,37] In addition, human papilloma virus (HPV), in particular HPV16, 18 and 45, cause cervical cancer, and a common belief is that HPV infection is a necessary cause for cervical cancer development.[38] In leukaemia, many epidemiological studies have associated infection with disease onset, but no specific pathogen has been discovered. [39]

It is widely believed that dietary factors are important in the susceptibility to cancer. However, still no dietary factors have been shown to consistently increase or decrease the risk to develop a certain type of cancer,[33,40] other than drinking alcohol and certain local customs as for instance giving salted fish to infants causing nasopharyngeal cancer and consumption of food with aflatoxin.[33,41]

Figure 5 - Suntan with reason

Ultraviolet radiation in sunlight and sun tanning beds scar the skin when exposed repeatedly and excessively, increasing the risk to develop skin cancer, figure 5.[20] In Nordic countries, the population exposure to ultraviolet radiation has increased dramatically during the twentieth century and the average exposed skin area has doubled.[42] The pattern of ultraviolet exposure seems to be vital. For melanoma, ultraviolet exposure of intermittent nature has been seen to induce malignant changes in melanocytes increasing the risk of invasive disease.[20,43-47] For SCC, on the other hand, the accumulated exposure of ultraviolet radiation during an individual's lifetime is fundamental to disease development and indeed the majority of squamous cell carcinomas appear on sun-exposed areas such as the face, lower-lip, neck, ears, and hands.[20,48]

INHERITED PREDISPOSITION

Essentially, inherited susceptibility to cancer has been seen in rare genetic syndromes resulting in cancer occurring with a higher frequency and on average at an earlier age. More commonly, predisposition to cancer is also seen in families with a family history of one or more common malignancies. It has been proposed that cancer may occur mainly in genetically predisposed individuals, indicating that the greater part of the inherited susceptibility to cancer is due to the combined effects of many genetic variants at a number of different loci.[5-7] Identifying these susceptibility gene variations will have the potential to fundamentally deepen the understanding of the carcinogenic process.

HIGH-PENETRANCE GENES

Studies of extended families with site-specific aggregation of cancer have provided important leads in identifying the genetic susceptibility to cancer. Rare predisposing genes associated with a high risk of cancer with multiple cases of disease in families, were first identified through linkage analysis or positional cloning in the beginning of the era of genetic epidemiology. Key signalling molecules such as p53 (*TP53*) were originally discovered as important mutation sites for viruses[49-51] and somatic mutations in tumours. Later these signalling molecules were identified as germline cancer susceptibility genes.[52]

Many fundamental and unanticipated insights into the mechanisms of carcinogenesis have been provided by high-penetrance alleles. For instance, in the rare autosomal dominant condition familial adenomatous polyposis (FAP), responsible for less than 1% of all colorectal cancers, hundreds to thousands of adenomatous polyps in the colon and rectum usually have developed by the age of 30 years.[53,54] An individual with FAP will if untreated almost unavoidable develop cancer at an average age of 39 years.[53,55] By genetic linkage *APC*, the gene for FAP, was localized on chromosome 5q21-q22.[56] A straightforward correlation was established between the genetic change in *APC* and the corresponding phenotype of the colorectal cancer. Mutations in the early part of the gene often resulted in late onset non-aggressive polyps, while further up (from exon 9 to exon 15) mutations led to more severe disease.[54] In the end of the gene, mutations gave features recognised as Gardner's syndrome and mild polyp disease. In sporadic cancer, the *APC* gene is likewise vital, however, with the difference that the losses of both functioning copies of the gene are acquired and thus the onset of disease is generally later.

Familial Cancer syndrome	**Gene**	**Type of Tumours observed**
Li-Fraumeni Syndrome	*TP53*	Breast cancer, brain tumours and leukemia
Familial Adenomatous Polyposis	*APC*	Colon cancer
Familial Melanoma	*CDKN2A*	Melanoma and pancreatic cancer
Hereditary Nonpolyposis Colon Cancer	*MSH2 MLH1*	Colon cancer and extracolonic cancer
Familial Breast Cancer	*BRCA1 BRCA2*	Breast and ovarian cancer

Families with hereditary nonpolyposis colorectal cancer (HNPCC) also called Lynch syndrome, have an autosomal dominant syndrome accounting for 2-5% of all colorectal cancer cases.[57] Families with HNPCC are characterised by development of extra colonic cancer in the endometrium, ovary, stomach, small bowel and brain among others with early disease onset (median age around 45 years).[58] In HNPCC families, mutations in the DNA mismatch repair genes including genes *MSH2*, *MLH1* and *MSH6*, affect the expansion or contraction of short repeat sequences of DNA known as microsatellite instability (MSI) and cause disease.[54]

In the 1990s two major susceptibility genes, *BRCA1* and *BRCA2*, for breast cancer were identified.[59-61] Later several other genes including *TP53*, *PTEN*, *STK11*, and *CDH1* have also been associated with an increased risk of breast cancer.[62-67] Other identified variants in the DNA repair genes *CHEK2*, *ATM*, *BRIP1* and *PALB2* confer a roughly twofold increase in the risk of breast cancer.[68-74] Susceptibility to melanoma in families with many affected individuals is known to be linked to mutations in the cell cycle regulator *CDKN2A* gene on chromosome 9p21[75] and less frequently to *CDK4*.[76-79] The cell cycle regulator *CDKN2A* controls the cyclin-dependent kinase inhibitor 2a, that inhibit *CDK4* and *CDK6* ability to phosphorylate the retinoblastoma protein.[44] *CDKN2A* has been found to be deleted homozygously or mutated in nearly 75% of melanoma cell lines.[75] The *CDK4* gene controls one of the binding partners of the cyclin-dependent kinase inhibitor 2a[76] and the *CDK4* mutations occur exclusively at the inhibitor binding domain.[44] In addition, a yet unidentified mutation has also been proposed at chromosome 1p22.[80] However, mutations are rare in kindreds with a fewer number of affected individuals, and only 7.8% of the Swedish melanoma families have *CDKN2A* mutations.[80]

IS CANCER A POLYGENIC DISEASE?

Shared alterations in high risk genes merely explain a small part of the familial aggregation of cancer, whereas the genetic variation of an individual including common single nucleotide polymorphisms (SNPs) and a number of low penetrance susceptibility genes may explain a larger part of cancer familiality. Accumulating data suggest that cancer is a truly complex disease since the liability to develop cancer may be caused by many genes or SNPs that each have a minor importance for the individual susceptibility.[5-7,81,82] Predisposition by combinations of weak genetic variants may be of greater significance to public health than the individual risks seen in the inherited cancer syndromes.

In melanoma few low-penetrance risk alleles have been discovered and only variants in the melanocortin receptor gene *MC1R* have been validated.[83-86] However, recent large-scale genome-wide association studies have suggested many loci associated with melanoma susceptibility[87-89] and liability is now believed to be caused by an interaction between the presence of inherited susceptibility genes, sun exposure, and other genes that together moderate the skin's responses to the sun.[44,90,91]

In breast cancer, known susceptibility genes have been estimated to account for around 15%-20% of the observed familial clustering of disease.[5] The remainder of the familial risk (80-85%) may be due to genetic or environmental factors. However, several studies on breast cancer support the belief that genetics predominate.[5,33,92,93] Indeed, in the past two years many genome-wide association studies have investigated breast cancer susceptibility identifying new loci associated

with risk of disease.[94-102] Five SNPs in novel independent loci were strongly associated with breast cancer development and in particular four genes were of functional interest.[95]

Various genes involved in the inflammation pathways were associated with prostate cancer risk, for instance, the macrophage regulating factor MIC1,[103] the Toll-like receptor TLRs helping the innate immunity to recognize pathogen-associated molecular patterns,[104,105] and the pro-inflammatory inhibitor IL1RN.[106] Also, in a more global genome-wide approach, new SNPs were found to be significantly associated with prostate cancer.[107]

CANCER SURVIVAL

Cancer is the second largest cause of death in the Western world. Family history of cancer is a well established risk factor for essentially all common cancers[10,16] but less is known about cancer survival. It may potentially be determined by various factors such as metastatic potential of the tumour, treatment, behavioural and sociodemographic factors. Today, it is not possible to accurately identify patients that do not benefit from available therapy or to differentiate between slow and fast growing tumours at the time of diagnosis. Most patients die from metastases and not the primary tumour, which means that prognosis is more strongly associated with distant spread of the disease.[108] The inefficient multi-step nature of metastasis in cancer progression has puzzled researchers for decades. It has been suggested that the genetic background of a cancer patient may be essential for the metastatic ability of the tumour, since genetic variants have the potential to increase or decrease the probability of a tumour to metastasize.[109-111]

Generally, cancer survival has generally improved during the last decades and for some types of cancers, the survival increase has been substantial. The relative Swedish 5-year cancer survival is presented in figure 6, illustrating the survival for cancer patients compared to the general population of the same age.[8] In the 1970s, the 5- year relative survival for all cancer patients was 36% for both males and females in Sweden. This has almost doubled to around 62% in 2008. For prostate cancer the 5-year relative survival has increased from 37% to 78%. The prostate-specific antigen (PSA) tests may be the true reason behind most of the survival increase, because the proportion of men diagnosed with prostate cancer at ages younger than 70 years has increased along with well differentiated tumours. The decision of whom to treat is problematic and many patients with prostate cancer are today overtreated due to lack of good prognostifying factors. In breast cancer, improved treatment and diagnostics are believed to be the true reasons behind the increased relative survival with a 5-year survival increase from 61% to 86% during the last three decades. A continuous 5-year relative survival increase from 33% to 57% has been observed for colorectal cancer. Several factors are believed to have contributed, for instance earlier diagnosis, increased use of cytostatic agents and improved surgical procedures enabling old individuals to be operated successfully. However, despite intense research, efforts to detect tumours earlier and more aggressive surgical procedures and radiotherapy, lung cancer survival has only slightly improved during the last three decades. Still today only around one in ten patients is alive 5 years after diagnosis.

5-year relative survival (percent)	
90	Skin
88	Melanoma
86	Breast (female)
78	Prostate
68	Cervical
57	Colon
57	Rectal
12	Lung

5-year relative survival (percent)	
89	Skin
84	Melanoma
80	Skin
78	Prostate
71	Urinary bladder
56	Colon
56	Rectal
10	Lung

5-year relative survival (percent)	
91	Melanoma
91	Skin
86	Breast
58	Colon
58	Rectal
45	Ovarian
15	Lung
68	Cervical

Figure 6 - 5-year relative survival for the most common cancers

THE METASTASIS PROCESS IN CANCER

In cancer, the metastasis process is fundamental given that 60% to 70% of cancer patients may have initiated the metastatic process by the time of diagnosis.[113] Even patients with no signs of tumour dissemination at primary diagnosis are at risk. Roughly one-third of sentinel lymph node negative women at primary breast cancer diagnosis are affected by metastatic disease.[114] Approximately 15% to 25% of patients with small primary and node negative tumours will develop secondary tumours.[115] Hence, better understanding of the factors leading to tumour dissemination is essential, especially since the first stages of metastasis can be early events.[116]

Despite the relatively high percentage of cancer patients developing distant metastasis, the metastatic cascade is very inefficient.[111] A cancer cell must go through a complete cascade before it can be successful in colonising at a secondary site, see figure 7. If all steps are not completed, the result will be failure to colonize and proliferate. The cells need to separate from the primary tumour, invade surrounding tissue and basement membranes, proceed to the blood vessel or lymphatic system and survive in the circulation. Finally the cancer cells may arrest in a distant target organ and more often than not continue to extravasate into the surrounding tissue. To survive in the new microenvironment and to be able to proliferate, the cancer cells continuously have to escape apoptotic death and immunological response. Although the mentioned metastatic cascade may explain the metastatic inefficiency, it is likely that we are aware of some but not all of these key regulatory points. For instance, cells have been seen to extravaste from capillary

beds in the bloodstream with high efficiency and in the secondary sites reside dormant for long periods of time,[117,118] sometimes for years.[119]

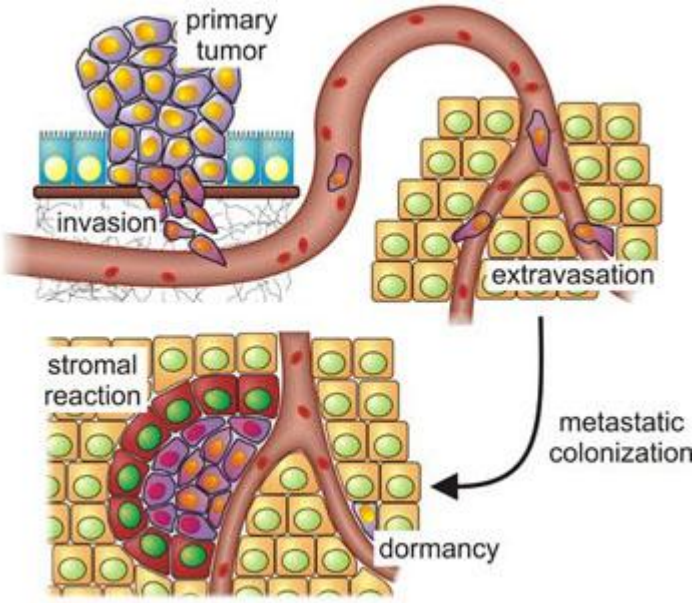


Figure 7 - The metastatic process in cancer

Cancer has been argued to generally be a local disease that may spread and metastasise over time. On the other hand, some argue that cancer is rather a systemic disease with formation of distant micrometastases very early in cancer development leading to high probability of disease spread before diagnosis. These different views on cancer development and metastatic process have immense implications for treated patients.[120]

Today, the common perception of the carcinogenic process combines historical contradictory views and cancer is generally considered to be a heterogeneous disease including cancers that are local in their nature continuously and cancers that are systemic very early in the cancer development. It is known that metastasising tumour cells primary use direct haematogenous routes and indeed, the higher the probability for distant metastasis to have formed the less will local therapies influence the patient's survival. Studies support the current view that cancer is a heterogeneous disease and in breast cancer, for instance, mammographic screening has been seen to reduce breast cancer mortality, with a relative risk of death of 0.85 in screened compared to unscreened populations.[121] Hence, earlier diagnosis can in some patients prevent the development of distant metastases since the length of a patient's survival is heavily dependent on if the disease has spread or not. In addition, this point to the fact that the probability of metastasis formation is influenced by the time of diagnosis and some tumours may develop the ability to spread to distant sites as a function of time, leaving strict definitions of either having the propensity or not to spread less likely. Local control of a tumour has been associated with better overall survival in many studies and indeed a significantly increased overall survival has been seen in women with breast cancer with a high risk of recurrence when adding radiation therapy after mastectomy.[122-126]

GENETIC BACKGROUND & METASTATIC DISEASE

The genetic background of an individual has been suggested to be essential for metastasis potential and cancer survival. It is hypothesised that allelic variants may modify the likelihood of tumour metastasis through vital secondary events, such as deletions, amplifications, and epigenetic modulations in the metastatic cascade.[110,111]

The microenvironment of the tissue where tumour cells escape is known to be important in metastasis formation,[127] because tumour cells are highly dependent on for instance normal stroma for their signalling events.[128] Studies have proposed that a majority of tumour cells are capable to spread,[118,129] which suggests that success in proliferating at secondary sites may determine whether cancer cells proliferate into a secondary tumour or undergo apoptosis. Furthermore, allelic variation may affect escape from the immune surveillance, because small variations in the ability of an individual to rise an effective cytolytic defence, together with the tumour cell's ability to downregulate specific antigens,[130] may also be important in the metastatic potential.

Currently, little is known about the involvement of genetic variants in cancer survival. Nevertheless, some genetic variations modulating disease progression and thus survival have been proposed in the literature. A locus altering the aggressiveness of prostate cancer has been mapped and confirmed by linkage analysis, and recently also *CAV-1* involvement situated in this region was presented.[131-134] In lung cancer, fifteen SNPs in the DNA repair pathway and five gene signatures significantly influenced lung cancer survival,[135] and the *EGFR* gene polymorphic simple sequence repeat length,[136] the Y/X polymorphism of the innate-immunity gene *MBL2* with haplotypes,[137] and glutathione-related genes have been associated with improved lung cancer survival.[138] Moreover, several investigators have explored the possible relationship between breast cancer survival and genetic polymorphisms in growth factor receptors as well as genes involved in angiogenesis, DNA repair, cell cycle checkpoints, and in extracellular and carcinogen metabolism.[139-141] Recently, metastasis suppressor genes inhibiting metastases without blocking tumour formation were discovered,[109,142-145] and a large population-based study assessed the association of genetic polymorphisms in DNA repair, hormone metabolism, carcinogen metabolism and other genes with breast cancer survival.[139] A gene involved in the DNA double-strand break repair, LIG4, had the largest effect on survival. Recently, genomic alterations in chromosome 16 have been associated with survival as well.[146] Finally, in colorectal cancer, genetic variation in genes involved in response to inflammation, DNA repair, and cell cycle checkpoints have been associated with survival.[147]

AIMS OF THE STUDIES

The aim of this thesis was to study familial aggregation of cancer with interest in genetic and environmental factors affecting disease development and survival.

CANCER DEVELOPMENT

Study I:

Extend the generalised linear mixed model (GLMM), enabling the analyses of genetic and environmental effects in two- and three-generational families, using melanoma as a model cancer.

Study II:

Enable analyses of co-aggregation of two diseases in a population-based setting by extending the GLMM model from study I, with special interest in disease aetiology of melanoma and squamous cell carcinoma of the skin.

CANCER SURVIVAL

Study III:

Test the novel hypothesis that cancer survival may be inherited by applying analyses of survival correlation in parent-child pairs.

Study IV:

Investigate the importance of genetic and environmental factors in lung cancer survival by means of analysis of pairs with different degrees of familial relationships.

MATERIALS & METHODS

POPULATION-BASED FAMILY DATA

National population-based registers with information on family members and possible cancer diagnosis are unique tools in epidemiologic cancer research. For instance, the Icelandic registers comprise good documentation of more than 687 500 individuals tracing the population back to the founding days including practically complete records of cancer cases from 1955 and today approximately 95% of cases in the register are histologically verified.[148] Also, the Utah population database, assembled originally from records from the Utah Family History Library called ”Family Group Sheet”, constitutes high-quality family information with more than 6 million included individuals linked to the Utah Cancer Registry recording cancer cases since 1966.[149]

In Sweden excellent resources for register-based research are found in the Multi-Generation Register and the Cancer Registry. All studies described in this thesis were based on Swedish population-based register data. Record linkage of personal information is possible since all residents have unique national registration numbers. Our studies are based on a record linkage between several population-based registers: the Multi-Generation Register, the Swedish Cancer Registry, the Cause of Death Registry, and the Migration Registry. Finally, additional linkages were made to the Censuses of 1960, 1970, 1980 and 1990 that holds information on individual socioeconomic status, figure 8.

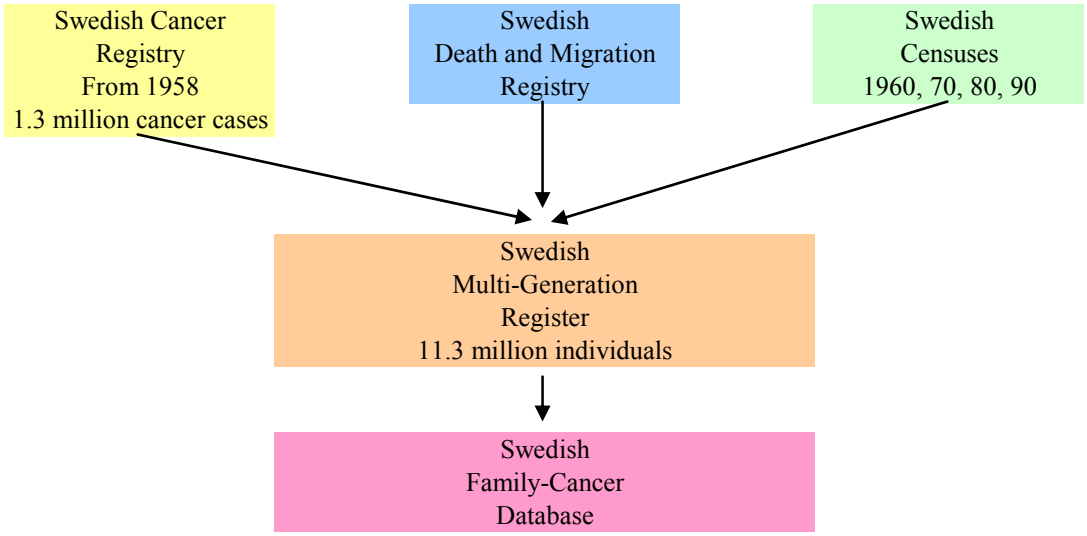


Figure 8 - Linkage in the creation of the Swedish Family-Cancer Database

The Multi-Generation Register includes individuals born in Sweden since January 1932 with their biological parents.[150] The registry was initiated in 1961 from written records in church parishes and country registration offices. A prerequisite to be included in the Multi-Generational Register, however, is that the individual is registered in Sweden at some time after 1961. In the Multi-Generation Register children exist only once while parents are present as many times as number of children. An individual can be in the database both as offspring and parent, and parents are those that admit to parenthood at birth, thus not only married individuals.

Incident cancers in Sweden have been recorded since 1958 in the Swedish Cancer Registry. All tumours in an individual are recorded separately in the Cancer Registry together with a unique personal number using a four digit diagnostic code according to the seventh revision of the International Classification of Diseases (ICD-7), together with information on date and county of diagnosis and histopathological type.[151] The records are based on compulsory reports from all treating physicians in both the private and public health care system. In the 1970s, the completeness of cancer registration (with cytological or histological verification) was assessed to be around 95% and has been regarded to be close to 100% since the 1990s.[152] In our studies, deaths caused by cancer were collected from the Cause of Death Registry, which has a reported accuracy of 96% from 1961 onwards and socioeconomic status was assembled from the Censuses. In summary, our database comprised over 11 million individuals organized into around three million families, including more than one million tumours.

The Swedish population-based registers are ideal for the conduct of large epidemiological studies. But as with all registers, there are some limitations by reasons of incompleteness and missingness.[153] In particular, about half of the individuals dying before the 1990s have missing parental information. This means that, for instance, early onset or fatal cancer cases are more likely to have missing parent information compared to non-fatal cases. Analyses only including cases with parental information may as a consequence potentially lead to biased estimates as it would selectively remove the fatal cases. Comfortingly, Leu et al concluded that missing familial information results in little or no bias in familial risk estimates when the mortality for familial and nonfamilial cases do not differ immensely.[154] Further, in the Swedish Cancer Registry, cancers before the start of the registry in 1958 were not recorded causing the family history of cancer to be misclassified in some individuals. However, left truncation of the cancer registry, has been seen to result in small downward non-differential bias, especially at moderate familial risk of disease.[154]

STATISTICAL ANALYSIS

STUDY I & STUDY II

SINGLE-TRAIT MODEL

In our analysis, cancer was defined as a binary trait with the assumption of an underlying normal distribution of liability defining the liability as the sum of the genetic and environmental effects in an individual.

We utilized a mixed effects model, assuming that the mean outcome can be explained by a set of fixed effects and a set of random effects. The probability to get cancer (p_{ij}) for each member (j) in a family (i) was assumed to be Bernoulli distributed and a transformation to the standard normal scale with the probit link was used for modelling reasons.

The random effects ($\mathbf{z}'_{ij}\mathbf{b}_i$) capture all the effects that explain the aggregation of cancer in families, while the fixed effects ($\mathbf{x}'_{ij}\boldsymbol{\beta}$) in our case describe the prevalence levels of the disease within in the different generations.

The model is given by

$$\Phi^{-1}(p_{ij}) = \mathbf{x}'_{ij}\boldsymbol{\beta} + \mathbf{z}'_{ij}\mathbf{b}_i ,$$

where $\Phi()$ is the normal distribution function.

Family members share genes and common environment including all lifestyle factors. We define several types of environment shared by different members in our families. The family-shared environment is shared by all members in a family, while the adult-shared environment is shared only by the spouses in a family whereas the children in a family share childhood environment.

In **study I,** we modelled melanoma and specifically wanted to separate the adult-shared and childhood-shared environment. The shared factors were modelled as a sum of genetic (g_{ij}), adult-shared environment (a_{ij}) and childhood-shared (c_{ij}).

$$\Phi^{-1}(p_{ij}) = \mathbf{x}'_{ij}\boldsymbol{\beta} + \mathbf{g}_{ij} + \mathbf{a}_{ij} + \mathbf{c}_{ij},$$

In **study II**, to shed light on the aetiology of melanoma and SCC we wanted to analyse the relative environmental and genetic burden in covered sites and exposed sites. Thus, we modelled the shared factors as a sum of genetic (g_{ij}), family-shared environmental (f_{ij}), and childhood-shared environmental (c_{ij}) effects, so that

$$\Phi^{-1}(p_{ij}) = \mathbf{x}'_{ij}\boldsymbol{\beta} + \mathbf{g}_{ij} + \mathbf{f}_{ij} + \mathbf{c}_{ij},$$

The definition of the random effects, the variance components for each model and the covariance for the individuals in each family are seen below.

	Variance (Random effects)	**Covariance (spouse-spouse)**	**Covariance (sibling-sibling)**	**Covariance (parent-child)**
Genetic, childhood-shared and adult-shared environment model	$\sigma_g^2 + \sigma_c^2 + \sigma_a^2$	σ_a^2	$\frac{1}{2}\sigma_g^2 + \sigma_c^2$	$\frac{1}{2}\sigma_g^2$
Genetic, family-shared and childhood-shared environment model	$\sigma_g^2 + \sigma_f^2 + \sigma_c^2$	σ_f^2	$\frac{1}{2}\sigma_g^2 + \sigma_f^2 + \sigma_c^2$	$\frac{1}{2}\sigma_g^2 + \sigma_f^2$

In both studies, different family structures are assumed to be independent, but the random effects within the families are correlated according to the usual assumptions in quantitative genetic analysis, where the correlations between relatives for environmental and genetic effects are set to fixed values according to their degree of genetic and environmental relationship. Hence, all first-degree relatives (parent-offspring and full siblings) were assumed to be correlated by 0.5 for the genetic factors while the grandparents and grandchildren in the three-generational model were correlated by 0.25 for the genetic effects. A higher correlation among relatives that are more closely related to each other genetically indicates the importance of genetic effects. Moreover, all members in a family shared the family-shared environment, while only spouses are fully correlated for the shared adult environmental factors, and siblings share childhood environment. Childhood-shared environmental effects result in greater environmental resemblance among siblings than among parents and children, inhabiting also recessive genetic effects. Non-shared environmental effects are seen in the within-family differences.

Within family differences in liability to disease are seen in non-shared environmental effects and is implied by the randomness of the probit model. The variance of these non-shared environmental effects was set to 1 for identifiability.

In quantitative genetics, the estimated variances of liability attributable to various factors are usually presented. The heritability, in particular, is defined as the proportion of susceptibility due to genetic factors given by

$$h_S^2 = \frac{\sigma^2{}_g}{\sigma^2{}_g + \sigma^2{}_e + 1},$$

where e stands for the defined environmental effects in the model, s stands for single trait and finally 1 in the denominator is derived from the non-shared environment as explained above. To reduce the amount of computation and without any impact on the results, we first summarized the data according to all possible configurations in a family with regard to the event. The probability of each family pattern was then computed using a fast Monte Carlo integration method. The total likelihood was retrieved by summing the individual log-likelihood contributions from all families, assuming that the family patterns were independent.

Optimization of the total log-likelihood was done by employing optim(), a derivative-free simplex algorithm available in the statistical package R. Although the exact location of the maximum of the likelihood could not be found because of the Monte Carlo approximation used in the computation of the likelihood, we confirmed that the estimate was within the statistical uncertainty in the data by smoothing the log likelihood around each estimated parameter value.

FAMILIAL CO-AGGREGATION MODEL

In study II the occurrence of melanoma and SCC in an individual or in a family were studied with our newly developed co-aggregation model. The co-aggregation of two diseases results in the binary outcome of disease with the possibility of each individual being affected with melanoma and SCC. Thus, for a family including mother, father, and two children, the vector has a length of 8, consisting of four family members with binary outcomes for each disease.

Similar to the single-trait model, in the co-aggregation model the random effects were assumed to be independent. Analysing the variance, the covariance matrix between the two diseases was defined as a matrix combining the effects of melanoma (m) and SCC (s) including two variance components (σ^2_m and σ^2_s) and one covariance component (σ_{ms}). The covariance component is presented below,

$$\sigma_{ms} = \begin{pmatrix} M_{ms} & M_m F_s & M_m C1_s & M_m C2_s \\ F_m M_s & F_{ms} & F_m C1_s & F_m C2_s \\ C1_m M_s & C1_m F_s & C1_{ms} & C1_m C2_s \\ C2_m M_s & C2_m F_s & C2_{ms} & C2_m C2_s \end{pmatrix}$$

where M stands for mother, F stands for father, and C1 and C2 stand for sibling 1 and sibling 2, respectively.

The genetic, family-shared environmental, childhood-shared environmental and the unshared environmental factors were modelled in the familial co-aggregation model inhabiting two malignancies. The unshared effect (u_i) describes exposures that affect the co-aggregation of two diseases within an individual and are not shared with the rest of the family. The proportion of variance for genetic and environmental effects can be defined employing the quantitative genetic approach in a similar manner as for the single-trait model as given by,

$$h^2_{Co} = \frac{\sigma_{msg}}{\sigma_{msg} + \sigma_{msf} + \sigma_{msc} + \sigma^2_u}, \; f^2_{Co} = \frac{\sigma_{msf}}{\sigma_{msg} + \sigma_{msf} + \sigma_{msc} + \sigma^2_u}, \; c^2_{Co} = \frac{\sigma_{msc}}{\sigma_{msg} + \sigma_{msf} + \sigma_{msc} + \sigma^2_u}$$

where Co stands for co-aggregation.

STUDY III & STUDY IV

In study III, we analysed the cancer-specific survival in children with a parent diagnosed with the same cancer. All parents with a first primary cancer diagnosed between 1961, and 2001, who had a child diagnosed with cancer between 1991 and 2001, were included in the study.

In study IV, all parent-child, sibling and spouse pairs concordantly diagnosed with a first primary invasive lung cancer were included. Cause-specific lung cancer death within 5 years was the outcome of interest. Children and siblings in our analyses were diagnosed between January 1991 and December 2001, while for parents and spouses the follow-up was unrestricted (January 1961 until December 2001).

Depending on the modelled survival in proband the survival in proband's relative (child, sibling, spouse) was analysed. First, by use of the Kaplan-Meier method following individuals for five years after diagnosis, and secondly with a multivariate proportional hazard model adjusting for possible confounders on survival such as calendar year of diagnosis, age at diagnosis, socioeconomic factors, county of diagnosis, tumour histology and gender.

In both studies, we restricted our offspring analysis to 1991 and onwards, because complete data for parents of children who died from 1991 are available in the Multi-Generation Register, whereas before this date the data are incomplete. The start of follow-up was the date of cancer diagnosis and follow-up continued until death, emigration, or the end of follow-up. Since the Cause of Death Registry has a high reported accuracy of 96% from 1961 onwards, we limited our follow-up back to 1961.

In study III, the survival in children in relation to parental survival was analysed using the Kaplan-Meier method. The parents were grouped into either dead within ten years (from the same cancer as their child) or alive ten years after diagnosis. To ensure that all parents had the possibility of a 10-year survival, only parents diagnosed between 1961 and 1991 were included. In study IV, the proband survival was defined from the multivariate proportional hazards model described below.

In both studies the survival in the probands (parent, sibling, spouse) was modelled with a multivariate proportional hazards (Cox) model adjusting for the calendar year of diagnosis and age at diagnosis. The residuals from this model were used to describe proband survival compared to the cumulative baseline hazard, adjusting for calendar year of diagnosis and age at diagnosis, resulting in residual values below, above, and around zero. Subsequently, by defining groups according to quartiles of survival the probands were categorized to better than expected survival group as the best quartile of survival, the expected survival group as the middle two quartiles of survival, and the worse than expected survival group as the worst quartile of survival. For simplicity, we refer to these categories as good, expected and poor.

Finally, the proportional hazard assumption for the main exposure variable was assessed using Schoenfeld's test statistics;[155] no significant deviation was noted for the family pairs studied. An arbitrary level of 5% statistical significance was used.

FINDINGS & INTERPRETATION

STUDY I

Population-based family studies have inherent problems in contrasting gene and environmental factors due to the dependence among individuals in a family. Until now, family members been analysed in a pairwise manner assuming independence between different relative pairs in a family, such as sibling-sibling and parent-offspring pairs.

In study I, we extended the previously developed generalized linear mixed model enabling analyses of genetic and environmental effects in two- and three-generational families considering all relationships in a family at once avoiding pairwise analyses of dependent family members.

More specifically, we aimed at comparing the relative importance of genetic and environmental factors in the development of melanoma. Melanoma was selected as a model cancer since it is a common sex-unspecific cancer with reasonably early onset of disease. From our population-based Swedish database we defined two-generational families as families consisting of a mother and a father together with their oldest two children (or their only child). We also linked our nuclear family data into three-generational families, constituting a grandmother and a grandfather, their two oldest children with spouses, and the two oldest grandchildren, see figure 9.

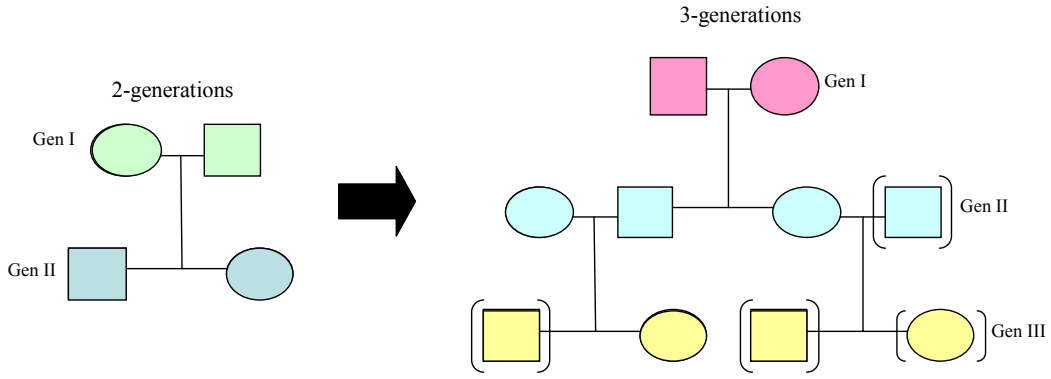


Figure 9 - Two- and three-generational families where individuals within brackets are optional in the family structures accommodated by our analysis

To reduce the computational burden, we restricted our three-generational families by only including the two oldest siblings in generation II, excluding families with less than two siblings in generation II, in order to obtain a good estimation of the childhood-shared effect. We included the two oldest siblings in generation II since the oldest sibs are more likely to have their own children contributing to more person–time of observation in the analysis. Our model can easily be modified to inhabit different structured pedigrees.

After the restriction, 2.6 million two-generational families and one million three-generational families could be included in our analyses. In the three-generational families, individuals in generations II and III belong both to the family with their own parents and to the family with their spouses' parents, while the grandchildren in generation III are biological members of both their mothers' and fathers' family with grandparents.

In the two-generational families, the median age of children and parents was 34 and 58 years, respectively. The median attained age up to the end of follow-up in families with three generations of relatives was 20 years for children, 47 years for parents and 69 years for grandparents. Thus, a proportion of the children belonging to the three-generational families were too young to be at risk to develop melanoma, so in estimating the environmental and genetic contribution we included a baseline risk for each generation in the two- and three-generational families.

In previous studies, family members have often been assumed to either share environment in the whole family (familial environment) or the environment has been assumed to be shared by the spouses (adult) or siblings (childhood) separately. In our case, defining a familial environment across three generations assuming common environment in grandchildren and grandparents did not seem reasonable. Instead, a separate childhood-shared and adult-shared environment was defined to be shared in the families. Thus in summary, we assumed that all first-degree relatives (parent-offspring and full siblings) shared 50% of their genes while grandparents and grandchildren shared 25% of their genes, as generally the case when thinking in average inheritance terms. Spouses are assumed to share adult environment and the siblings in addition to genes are also assumed to share childhood environment in our model. Childhood-shared environmental effects result in greater environmental resemblance among siblings than among parents and children and a higher concordance of melanoma among closely related family members would indicate an importance of genetic effects.

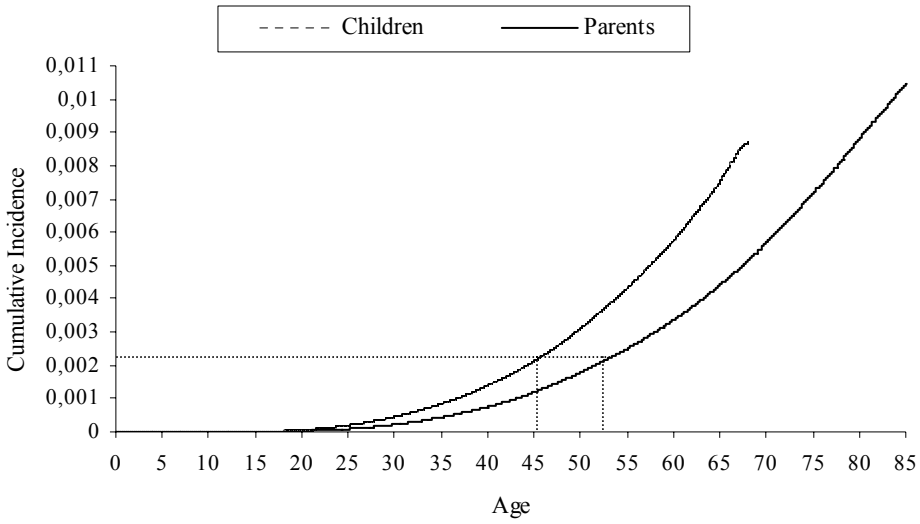
The estimated contributions of the genetic and adult-shared environment in the two- and three-generational families were 29% and 6% in both family structures, table 4. The childhood-shared environmental effect differed slightly and was estimated at 8% and 4% in the two- and three-generational families, respectively, see table 4.

Table 4 - Two- and three-generational families and genetic and environmental factors in melanoma

Family structure	**Percentage of variance (95% CI)**		
	Genetic	Adult-shared environmental	Childhood-shared environmental
Two generations	29 (27-31)	6 (3-9)	8 (2-13)
Three generations	29 (26-32)	6 (3-9)	4 (0-9)

In our data, we have unequal observation time on children compared to their parents. In addition, the incidence of melanoma has changed with a steep increase during the last twenty years. In order to correct for this potential survival length bias, we made a cumulative incidence restriction in our two-generational model within each family, resulting in equal cumulative hazard in the two

generations. For example, in a family with children with an average age of 45 years, the parents were followed for around 53 years to achieve to same cumulative incidence in the family, see the age cumulative incidence graph below.



No noteworthy differences in the estimated genetic and environmental components were obtained from the restriction, particularly when considering the confidence intervals of the estimates, table 5.

Table 5 - Cumulative incidence restriction in two-generational families

Family structure	**Percentage of variance (95 % CI)**		
	Genetic	Adult-shared environmental	Childhood-shared environmental
Two generations	24 (17-30)	13 (6-20)	10 (5-15)

Our newly developed models enable analyses on complete family structures. Although we consider our developed GLMM model to be statistically superior to standard structural equation modelling,[156] still these two models resulted in very similar estimates with respect to melanoma. Further, the two- and three-generational family design resulted in very similar genetic and shared-environmental estimates. However, because of inherent difficulties in the three-generational family approach, such as for instance truncation problems of our database leading to young average age in grandchildren, we favour the two-generational design. Lastly, since the correction for differential length of follow-up in our families resulted in estimates of similar magnitude, the potential survival length bias owing to unequal follow-up in families did not distort our final genetic and environmental estimates. In view of that and by reasons of reduction in number of disease concordances within the families, we decided not to restrict the survival time within the families in the subsequent familial studies.

STUDY II

AETIOLOGY OF MELANOMA & SCC

In both melanoma and SCC, genetic and environmental susceptibility to ultraviolet exposure is mainly unexplored. The pattern of sunlight to sun-covered sites has been suggested to be of a more intermittent nature than the sunlight pattern of sun-exposed sites.

The estimates on genetic contribution in sun-covered and sun-exposed sites in melanoma, see table 6, indicate that the pattern of sun exposure does not interact with the genetic susceptibility to melanoma, since the genetic effect was stable over different patterns of environmental exposure.

Table 6 - Effects of genetic and environmental factors in melanoma and SCC

Site or type	Percentage of variance (95% CI)		
	Genetic	Family-shared environmental	Childhood-shared environmental
Melanoma			
All sites	18 (13-22)	6 (4-7)	8 (3-12)
Exposed sites	12 (6-18)	6 (3-8)	8 (1-15)
Covered sites	13 (5-21)	9 (6-11)	9 (1-17)
SCC			
All sites	8 (4-12)	18 (16-19)	7 (3-11)
Exposed sites	8 (4-13)	17 (16-19)	7 (2-12)

Family-shared environment in sun-covered sites seem more important than in the sun-exposed sites for melanoma. Hence, familial habits of sunbathing especially on sensitive winter skin unaccustomed to ultraviolet radiation may increase the susceptibility to melanoma. Higher familial risks at sun-covered compared with sun-exposed sites support our finding.[157] Generally, familial risks can be attributed to both genetic and environmental factors, but in light of our results we can now suggest that the higher familial risk in sun-covered compared to sun-exposed sites may mainly be explained by familial environment. The difference in genetic effects for sun-exposed, sun-covered, and all-sites may be due to random variation or be a true difference in the aetiology of the sites since the sun-covered and sun-exposed analyses exclude discordant sites in families. Also, childhood-shared environment estimated at 8-9% suggests that sun habits during infancy and youth are important in the susceptibility to melanoma.

Only a small proportion of the estimated genetic effect in melanoma can today be explained by identified mutations. In families with many affected individuals however, susceptibility has been linked to mutations in the cell cycle regulator *CDKN2A* (p16) gene and the *CDK4* gene.[76-79] Nevertheless, genetic susceptibility to melanoma may also include low-penetrance genes and SNPs, such as the melanocortin-1 receptor as well as variability in various DNA repair mechanisms.[83-86] Generally today the susceptibility is believed to be caused by an interaction

between the presence of inherited susceptibility genes, and other genes that together moderate the skin's responses to the sun.[44,90,91]

Previously, a reasonably high familial risk of 2.7 in SCC has revealed a familial aggregation.[18] In light of our results estimating familial shared environment at 18%, we propose that familial suntanning habits are of vital importance in determining the liability to SCC, consequently familial habits such as lengthy sunlight exposure, seems to be a cornerstone in defining the familial risk. However, genetic factors and childhood-shared environment are also clearly involved in the susceptibility to SCC, but to a seemingly smaller extent. Nonetheless, we also conclude that inherited factors including skin type and pigmentation or other so far unknown mechanisms may increase the sensitivity to accumulated sun exposure. In addition, sunburn at a young age has been found to increase the risk of SCC supporting our findings that childhood-shared environment is involved in SCC susceptibility.[158]

The estimation of genetic and environmental effects in melanoma including all body sites resulted in very similar estimates compared to the estimates in study I fitting a model with adult-shared environment effects instead of family-shared environment for model comparison reasons. However, it is our belief that the present model (including genetic, family-shared environment and childhood-shared environment) is favourable in a biological perspective in addition to being statistically superior.

CO-AGGREGATION OF MELANOMA & SCC

Melanoma and SCC co-aggregate in families, still the underlying familial susceptibility is unknown. A familial shared sensitivity to ultraviolet radiation has been proposed to be the reason, but high relative risks in families with co-aggregation of both diseases imply possible involvement of genetic factors.[18,19]

Table 7 - Genetic and environmental contribution to the familial co-aggregation of melanoma and SCC

Co-aggregation	**Percentage of variance (95% CI)**		
	Genetic	Family-shared environmental	Childhood-shared environmental
Melanoma and SCC	47 (43-51)	36 (33-39)	8 (4-13)

We disentangled the familial co-aggregation and apportioned it to genetic and environmental factors, table 7. Interestingly, our results indicate that inherited factors are vital in the common aetiology of melanoma and SCC. Today, the suggested genetic susceptibility is unexplained and only in a few rare syndromes such as Werner's syndrome and xeroderma pigmentosum, genetic variation has been seen to lead to co-aggregation of melanoma and SCC. In addition, familial sun-tanning habits also appear to be important in the common susceptibility to melanoma and SCC.

We conclude that genetic factors are vital in familial aggregation of melanoma. However, only a small portion of the genetic effect can today be explained by known genetic alterations. Nonetheless, inherited susceptibility genes, in addition to other genes that together moderate the skin's response to the sun will most probably determine the susceptibility to melanoma. The family-shared environmental contribution in sun-covered sites of the body was higher compared to sun-exposed sites, probably conveying the benefit from cautious sunbathing on sensitive skin. Further, suntanning habits in childhood seem to be significant in both melanoma and SCC. In SCC, both family-shared environment and genetic variability influence the susceptibility to disease. Lastly, genetic components are involved in the familial co-aggregation of melanoma and SCC.

STUDY III & IV

Today, it is not possible to accurately identify patients that do not benefit from available therapy. However, it has been suggested that the genetic background of a cancer patient is essential for the metastatic ability of the tumour.[110,111] As a first nation-wide Swedish population-based epidemiological study, we analysed the familial correlation in cancer survival.

In study III, we investigated whether cancer specific survival was concordant among parents and children, identifying all pairs of parents and children diagnosed with colorectal, lung, female breast, ovarian, and prostate cancer along with the number of cancer-specific deaths.

Table 8 - Hazard ratio of poor, expected or good survival in children depending on parental survival

		Follow-up (parent, child) 10, 5 years	
Site or type	Parental[a] survival	Adjusted[b] HR (95% CI)	Trend test p-values[c]
Colorectal	Good	1.0 ref.	0.045
	Expected	1.20 (0.87-1.66)	
	Poor	1.44 (1.01-2.01)	
Lung	Good	1.0 ref.	0.047
	Expected	1.21 (0.90-1.62)	
	Poor	1.39 (1.00-1.94)	
Breast	Good	1.0 ref.	0.010
	Expected	1.29 (0.86-1.94)	
	Poor	1.75 (1.13-2.71)	
Ovary	Good	1.0 ref.	0.13
	Expected	1.62 (0.60-4.37)	
	Poor	2.23 (0.78-6.34)	
Prostate[d]	Good	1.0 ref.	0.026
	Expected	2.02 (1.16-3.51)	
	Poor	2.07 (1.13-3.79)	

[a]Parental survival was defined by a separate proportional hazard model adjusted for parental age and period of diagnosis

[b]Adjusted for age of diagnosis, socioeconomic status, and area of diagnosis in children

[c]Trend test with one degree of freedom

[d]Children are followed between January 1961 and December 2001

We noted a significantly increased risk of poor survival in children with poor parental survival compared with the risk in children with good parental survival in all cancer sites assessed, except of ovarian cancer, see table 8. A significant trend of increasing risk of death in children by worsening parental survival was also seen for colorectal, lung, female breast and prostate cancer.

Both genetic and environmental factors could explain the cancer-specific survival concordance between family members in a family, because individuals who share environment share behavioural and other exposures that may affect cancer survival. Therefore we conducted **study IV** aiming at disentangling the importance of genetic and environmental factors in lung cancer. We selected lung cancer because it is a sex unspecific cancer and so far adjuvant therapeutic survival gains have been small compared to other common cancers making interpretations on survival less complicated.

In study IV, lung cancer survival in an individual was seen to be dependent on the lung cancer survival in his/her parents or siblings, see table 9. However, in spouses, no effect on spouse survival was seen.

Table 9 - Risk of lung cancer-specific death in proband's relative depending on proband survival

		Risk of lung cancer related death in proband's relative	
Pairs of relatives	Survival in proband[a]	Adjusted[b] HR (95% CI)	Trend test p values[c]
Parent[d]-Child[e]	Good	0.71 (0.51-0.99)	0.04
	Expected	0.86 (0.65-1.13)	
	Poor	1.0 ref.	
Sibling-Sibling[e]	Good	0.14 (0.030-0.65)	0.05
	Expected	1.26 (0.48-3.27)	
	Poor	1.0 ref.	
Spouse-Spouse[d]	Good	0.85 (0.64-1.13)	0.26
	Expected	0.90 (0.70-1.15)	
	Poor	1.0 ref.	

[a]Multivariate proportional hazard (Cox) model adjusted for calendar year of diagnosis and age at diagnosis

[b]Multivariate proportional hazard (Cox) model adjusted for age, year and place of diagnosis, socioeconomic status, gender and histology

[c]One degree of freedom

[d]Parents (Spouses) diagnosed between January 1961 and December 2001

[e]Children (siblings) diagnosed between January 1991 and December 2001

A potential limitation to our study was the absence of information about smoking. However, we believe that our inability to adjust for smoking habits will have only very small effects on our results for several reasons. Firstly, in literature the impact of smoking on lung cancer survival seems to be dependent on a number of factors such as sex, histological type and years since smoking cessation,[159-162] and the overall effect of smoking on lung cancer survival has been seen to at most, in certain histologies and in women, increase the risk with around 30% of dying.[159,160] Secondly, although smoking is a genuinely established risk factor for lung cancer development, with utmost risk in individuals who begin to smoke at young age and continue through life,[33] familial cases of lung cancer can not be attributed to shared smoking habits.[163] Lastly, previous reports do not support higher correlation of smoking habits between siblings or parent-offspring as compared to spouses.[164-169] Based on our strong protective findings and the relatively limited effect of therapy on survival, we believe that an individual's genetic makeup is more important

than lifestyle factors such as smoking in explaining the similarities in lung cancer survival. In addition, we should not rule out the possibility of a recessive pathway of inheritance as indicated by the sibling estimates.

In both study III and IV, the absence of information on prognostic factors such as stage of disease is a limitation. However, adjusting for such factors is problematic because if familial correlation in cancer survival is a genuine biological effect, this correlation would be seen in prognostic factors called mediators, and adjusting for them would weaken or eliminate the association. Nonetheless, such information would give us an opportunity to study whether the cancer prognosis of a parent is an independent survival prognosticator in the newly diagnosed individual.

Little is currently known on the involvement of genetic variants in cancer survival, nevertheless, some genetic variations modulating progression and thus survival such as the *CAV-1* locus in prostate cancer, [131-134] several genes such as the *EGFR*[136] and *MBL2*[137] and fifteen SNPs in the DNA repair pathway[135] have been discovered in lung cancer. For breast cancer genetic polymorphisms in growth factor receptors in addition to genes involved in angiogenesis, DNA repair, cell cycle checkpoints, and in extracellular and carcinogen metabolism influence survival.[139-141] Also, in colorectal cancer, genetic variation in genes involved in response to inflammation, DNA repair, and cell cycle checkpoints have been associated with survival.[147,170,171]

In study III and IV, we tested the novel hypothesis that cancer survival is inherited. Encouraged by our results, we now suggest that genetic background influences the individual's ability to fight cancer and survive. We see many advantages in our study design, for instance an almost complete ascertainment of cancers along with a complete follow-up of cancer patients in addition to robust estimates only slightly differing on the inclusion of potential confounding factors. Additionally, genetic background appears to be more important than lifestyle factors in lung cancer survival. Summing up, currently little is known about genetic variants influencing cancer survival, nevertheless, some have been suggested in literature and the list is growing by the day.

CONCLUSIONS & FUTURE PERSPECTIVES

DEVELOPMENT OF CANCER

In a very large population-based family study of two- and three- generational families, we conclude that it is possible to use our GLMM model in estimating the contribution of genetic and environmental effects within a feasible time. We believe that our GLMM model is superior to the standard Mx methodology. The resulting estimates from the two- and three-generational family analyses were very similar and complicating factors in the use of three-generational family approach, such as increased analysis time and truncation of our database leading to young average age in grandchildren, leaves us to prefer the two-generational design. Even so, when the inclusion of additional covariates to the model such as age is possible, we will favour the three-generational design since it enriches the analyses contributing with more genetic and environmental relations among the family members comparing to the two-generational setting. The correction for differential length of follow-up in our families resulted in estimates of similar magnitude. We conclude that the potential survival length bias by reasons of unequal follow-up in our families does not seem to distort our final genetic and environmental estimates. However, since the restriction on observation time in the parental generation reduces the concordances of disease within the families and the power of the study, we chose not to restrict our families in the following studies. In the future, the next step would be to further develop our model enabling analyses of multiple covariates, for instance age and period explicitly in families in a population-based setting.

Genetic factors are vital in familial aggregation of melanoma. Still, no more than a small proportion of the estimated genetic effect in melanoma can today be explained by identified mutations. Melanoma susceptibility is likely to be determined by the presence of hereditary susceptibility genes, sun exposure, and other genes that moderate the skin's responses to the sun. In the future, large molecular studies will have the potential to unravel the nature of the genetic susceptibility to melanoma. Additionally, comparing the liability to disease in sun-covered and sun-exposed body sites, the familial environmental contribution was higher in sun-covered sites, conveying the benefit from cautious sunbathing on sensitive skin. Sound childhood environment is important to avoid risk inflicted sunburns that increase the susceptibility to both melanoma and SCC. In SCC family-shared environment appeared to be vital with the highest contribution of family-shared environmental effects ever seen in cancer. Genetic variability in individuals enhancing the sensitivity to accumulated sun exposure is also likely to be involved in the aetiology of SCC. Lastly, genetic components are involved in the familial co-aggregation of melanoma and SCC. The reason behind the genetic susceptibility to both melanoma and SCC is unknown today and the co-aggregation has only been seen in a few rare syndromes.

CANCER SURVIVAL

The novel hypothesis that prognosis is inherited was tested in study III and IV. In light of our studies, we propose that genetic factors are important in determining cancer survival. Our large population-based study had an almost complete ascertainment of cancers along with a complete follow-up of cancer patients. We conclude that our model estimates were robust showing only small differences adjusting for factors that would possibly confound our estimates. Further, we conclude that the cancer-specific survival of an individual may be predicted from the parental survival of a cancer in the same site. Individuals in a family share environmental factors that influence their cancer survival. However, genetic background seems more important than lifestyle factors in lung cancer survival. Little is currently known on the involvement of genetic variants in cancer survival, nevertheless, some genetic variations modulating progression and thus survival have been suggested and the list is growing by the day.

We believe that information about the outcome of cancer among affected first degree relatives may help to foresee the cancer survival of a newly diagnosed individual. However, this novel observation only become relevant for clinical management provided that it is independent to established survival predictors. Further, our results advocate a need for future large powered molecular studies illuminating the genetic determinants of inherited survival in common cancers. Finally, molecular studies unravelling the possible correlation of tumour characteristics among first-degree relatives with cancer may allow a deeper understanding of the biologic mechanisms behind cancer survival.

CONCLUSIONS IN BRIEF

- Our newly developed GLMM model enables us to estimate the genetic and environmental contribution in a two- and three- generational design.

- In the future, if further model developments enable the inclusion of additional variables, such as age and period, the three-generational approach may be preferable since it enriches the analyses with more genetic and environmental relations among the family members.

- Genetic factors seem essential in the familial aggregation of melanoma. The genetic variation enhancing the liability to melanoma may be a combination of inherited susceptibility genes and other genes that moderate the skin's responses to the sun.

- Contrasting the family-shared environment in sun-covered and sun-exposed body sites, the contribution was higher in covered sites, possibly conveying the benefit from cautious sunbathing on sensitive skin.

- In SCC, the family-shared environment appears to be vital with the highest contribution of family-shared environmental effects ever seen in cancer.

- Genetic variability enhancing sensitivity to accumulated sun exposure is probably involved in the aetiology of SCC.

- It is likely that genetic variability is involved in the familial co-aggregation of melanoma and SCC.

- Genetic factors are likely to be important in determining cancer survival. The cancer-specific survival of an individual may be predicted from the parental survival of a cancer in the same site. In addition, for lung cancer we conclude that genetic factors seem more important than lifestyle factors in lung cancer.

- Our model estimates were robust showing only small differences adjusting for factors that would possibly confound our estimates.

- We believe that information about the outcome of cancer among affected first degree relatives may help to foresee the cancer survival of a newly diagnosed individual.

- Finally, our results advocate a need for future large powered molecular studies illuminating the genetic determinants of inherited survival in common cancers.

SVENSK SAMMANFATTNING

Cancer berör i stort sett alla, direkt eller indirekt. Syftet med denna avhandling var att studera det genetiska och miljömässiga bidraget bakom risken att få cancer och att dö i cancerrelaterad sjukdom. Våra studier är baserade på en samkörning av flera svenska populationsbaserade register, däribland Multigenerationsregistret, Svenska Cancerregistret och Dödsorsaksregistret. Sammanfattningsvis baserades studierna på 11 miljoner individer i tre miljoner familjer och på mer än en miljon cancerfall.

I studie I utvecklade vi en statistisk modell för att möjliggöra analys av genetiska och miljömässiga faktorer i två- och tregenerationsfamiljer. Våra familjeanalyser resulterade i lika estimat. I studie II beräknades det genetiska bidraget för melanom till 18%. Vi analyserade även solexponerade och solskyddade kroppsdelar i en separat analys. Familjemiljön var viktigare i de täckta kroppsdelarna jämfört med de solutsatta vilket kan tyda på att känslig hud som är ovan vid sol bör solas varsamt, annars ökar risken för melanom. Barndomsmiljön visade sig vara viktig både för att undvika melanom och skivepitelhudcancer. Man bör särskilt tänka på att undvika att bränna sig som barn och ungdom. Vidare estimerades familjemiljön i skivepitelhudcancer till 18%. De nedärvda anlagen ökar sannolikt också risken att få sjukdomen, kanske genom ökad känslighet för ultraviolett strålning. Slutligen tyder våra estimat på att genetiska faktorer även kan vara avgörande i familjer där individer är känsliga för både melanom och skivepitelhudcancer. I familjer med förekomst av både melanom och skivepitelhudcancer beror nästan hälften av känsligheten på genetiska faktorer.

Numera är det klarlagt att genetiska skillnader hos individer påverkar risken att drabbas av cancer. Lite är dock känt om genetiska faktorers inverkan på canceröverlevnad. Vi presenterar den första populationsbaserade studien av kopplad canceröverlevnad hos familjemedlemmar. I studie III såg vi en signifikant ökat risk för dålig överlevnad hos vuxna barn till föräldrar med dålig prognos vid en jämförelse av vuxna barn med föräldrar med bra överlevnad. För kolorektal cancer var risken 44% högre, 40% för lungcancer, 75% för bröstcancer och mer än dubbelt så hög risk att dö i prostatacancer om ens förälder hade dålig överlevnad i samma typ av cancer. I studie IV analyserade vi föräldrar och vuxna barn, syskon och makar med lungcancer. Vi såg att vuxna barn till föräldrar med bra överlevnad hade signifikant bättre prognos jämfört individer med föräldrar med dålig överlevnad. Vi såg även att syskon med lungcancer hade signifikant bättre överlevnad om syskonet hade bra prognos. Ingen skillnad kunde dock ses hos makar. Nedärvda anlag verkar därför sammanfattningsvis viktigare än livsstilsfaktorer som rökning för överlevnad i lungcancer. Vår samlade slutsats för studie III och IV är att genetiska faktorer är mycket väsentliga för en individs canceröverlevnad.

Sammanfattningsvis tyder våra resultat på att genetisk variabilitet är mycket viktig i en individs känslighet för melanom och även i familjer som är drabbade av både melanom och skivepitelhudcancer. Våra studier stödjer även antagandet att den kroppsliga förmågan att kämpa mot cancer och överleva är delvis nedärvd. I framtiden ser jag populationsbaserade studier som en integrerad del i att identifiera genetiska faktorer som påverkar känsligheten för cancer och möjligheten att överleva då man har drabbats.

ACKNOWLEDGMENTS

My time as a doctoral student has given me the opportunity to meet wonderful people. It has been a truly interesting scientific and inner journey. I want to begin to thank all at MEB for the friendly atmosphere.

I would also particularly like to express my gratitude to the following people:

Kamila Czene, my supervisor and true scientific idol. Thank you for sharing with me your ingenious logical thinking, patience and humour during the last four years. You truly inspire me!

Yudi Pawitan, my co-supervisor. Thank you for introducing me to the field of biostatistics with a smile. It has been a pleasure working with you.

Paul Lichtestein, my co-supervisor. Thank you for working with me during my first two papers even if your main research interest was elsewhere than cancer. It has been a pleasure working with you.

Per Hall, your enthusiasm helps to create the stimulating and friendly atmosphere up on the fifth floor. I feel truly happy to have had the opportunity to be a part of your group.

All my co-authors - thank you for many stimulating scientific discussions, help and support.

Anna Svensson, simply for being who you are. For your support & kindness and for being a very good friend. I am convinced that you will succeed in everything you devote yourself to in the future, simply trust yourself!

Anthony Gunnell, I really miss you as a very good friend and scientist. I remember our chats on life and science and I sincerely hope that we will meet up again in the future. Give Kris and Lisa a big hug from me!

Alexandra Ekman, I truly miss you! You have been a good friend and I hope that we will stay in touch.

Eva Carlström, for being the considerate and truly friendly person you are. I wish you all the best!

Many thanks to Gunilla Sonnebring for so kindly proofreading my thesis!

Mattias Hammarström, for your refreshing honesty, humour and the always so stimulating discussions at lunch or just over the corridor. It has meant a lot to me! Agneta Lönn, for being so considerate and always ready to listen and help whenever you feel there is a need! Milka Krestelica, for your unique caring personality and positive attitude towards life. Ann-Sofie Andersson, for always being very kind and friendly. Edoardo Colzani, for introducing a new Italian charming atmosphere to the corridor! Maria-Pia Hergens and Christina Persson, for so kindly supporting me during my last stressful months of writing!
It has been a true pleasure to get to know all of you! I hope I will see you all in the future.

Elinor Fondell and Sanna Tiikkaja for nice friendship and lunches in Solna Centrum in the early years!

Veronika Remenarova, for being the sympathetic and enthusiastic person you truly are!

Denny Rönngren, Ben Yip, Ulrika Eriksson, Mikael Hartman, Michaela Prochazka, Camilla Björk, Jurgita Narusyte, Pia Fernberg, Anna Bennet, Monica Leu, Lennart Martinsson and Maria Sandberg for great chats over a cup of coffee or lunch.

Gunilla and Ralf Björklund & Maria and Bo Wergedahl. Thank you for your friendship, delicious dinners and company. I am truly happy to know you and your children! Erik and Emma are lucky to have friends in Karl, Jacob, Cecilia and Carl.

Gabriella Stålberg, no matter how long since we spoke it is like we met only the day before. You are a true friend, supporting and encouraging me when needed. Thank you for all the kind words and good spice of advice during the last years of friendship!

Josefine Forssell, "Jossan" my university best of friends. I am so grateful for the opportunity to get to know a person like you. You inspired me! I truly miss you and hope I will have a little more time in the future to visit you up north!

Jenny, Lelle, Martin & Gustav Selander, you are all so very dear to me and my family. Jenny, in you I may confide all aspects of life. Meeting you is always a true pleasure and to me you are the best friend that I could ever have wished for.

Ingegerd & Sören, min nya norrländska familj. Jag är så glad att ni blivit en så viktig del av vårt liv, trots att ni bor nästan 100 mil ifrån oss. Er kärlek & omtanke betyder väldigt mycket för oss alla!

Mina kära syskonbarn: Björn, Angelica, Jonathan, Filip och Hugo. Det är alltid härligt att träffas!

Min kära storasyster Nettan. Du är så ärlig och hjärtlig! Speciellt vill jag tacka dig för din kärlek & omtanke när jag var liten. Den betydde mer för mig än du någonsin kan ana.

Min älskade bror Lars, du är mig så otroligt kär. Tack för ditt stöd och din tilltro till mig. Vi har delat både glädje och sorg. Jag finns alltid här för dig!

Min kära mamma. Tack för din kärlek & närvaro i livets alla stunder. Genom din omsorg & tilltro har jag växt upp till att våga lita på mig själv och göra livet till mitt eget. Nu önskar jag inget hellre än att få hjälpa dig!

Ronny, Erik & Emma. Min älskade familj, ni är mer än jag någonsin vågade hoppas på. Tack för att ni finns!

REFERENCES

1. Hodgson S FW. Inherited susceptibility to cancer. Cambridge: Cambridge University Press, 1998.
2. Knudson AG, Jr. Mutation and cancer: statistical study of retinoblastoma. *Proc Natl Acad Sci U S A* 1971;**68**(4)**:**820-3.
3. Friend SH, Bernards R, Rogelj S, et al. A human DNA segment with properties of the gene that predisposes to retinoblastoma and osteosarcoma. *Nature* 1986;**323**(6089)**:**643-6.
4. Knudson AG, Jr., Meadows AT, Nichols WW, Hill R. Chromosomal deletion and retinoblastoma. *N Engl J Med* 1976;**295**(20)**:**1120-3.
5. Balmain A, Gray J, Ponder B. The genetics and genomics of cancer. *Nat Genet* 2003;**33 Suppl:**238-44.
6. Demant P. The genetic factors in cancer development and their implications for cancer prevention and detection. *Radiat Res* 2005;**164**(4 Pt 2)**:**462-6.
7. Houlston RS, Peto J. The search for low-penetrance cancer susceptibility alleles. *Oncogene* 2004;**23**(38)**:**6471-6.
8. Cancerfonden. Cancerfondsrapporten 2008. Stockholm: Infodata Direct AB, 2008.
9. Cancer i Siffor - populärvetenskapliga fakta om cancer. Sverige. Stockholm: Wassberg Skotte AB, 2005.
10. 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**(21)**:**1600-8.
11. Carstensen B, Soll-Johanning H, Villadsen E, Sondergaard JO, Lynge E. Familial aggregation of colorectal cancer in the general population. *Int J Cancer* 1996;**68**(4)**:**428-35.
12. Hemminki K, Vaittinen P. National database of familial cancer in Sweden. *Genet Epidemiol* 1998;**15**(3)**:**225-36.
13. Hemminki K, Vaittinen P, Kyyronen P. Age-specific familial risks in common cancers of the offspring. *Int J Cancer* 1998;**78**(2)**:**172-5.
14. Easton DF, Matthews FE, Ford D, Swerdlow AJ, Peto J. Cancer mortality in relatives of women with ovarian cancer: the OPCS Study. Office of Population Censuses and Surveys. *Int J Cancer* 1996;**65**(3)**:**284-94.
15. Peto J, Easton DF, Matthews FE, Ford D, Swerdlow AJ. Cancer mortality in relatives of women with breast cancer: the OPCS Study. Office of Population Censuses and Surveys. *Int J Cancer* 1996;**65**(3)**:**275-83.
16. Hemminki K, Li X, Czene K. Familial risk of cancer: data for clinical counseling and cancer genetics. *Int J Cancer* 2004;**108**(1)**:**109-14.
17. Easton DF. Familial risks of breast cancer. *Breast Cancer Res* 2002;**4**(5)**:**179-81.
18. Hemminki K, Zhang H, Czene K. Familial invasive and in situ squamous cell carcinoma of the skin. *Br J Cancer* 2003;**88**(9)**:**1375-80.
19. Hemminki K, Zhang H, Czene K. Familial and attributable risks in cutaneous melanoma: effects of proband and age. *J Invest Dermatol* 2003;**120**(2)**:**217-23.
20. English DR, Armstrong BK, Kricker A, Fleming C. Sunlight and cancer. *Cancer Causes Control* 1997;**8**(3)**:**271-83.
21. Broeks A, Urbanus JH, Floore AN, et al. ATM-heterozygous germline mutations contribute to breast cancer-susceptibility. *Am J Hum Genet* 2000;**66**(2)**:**494-500.
22. Chen Y, Thompson W, Semenciw R, Mao Y. Epidemiology of contralateral breast cancer. *Cancer Epidemiol Biomarkers Prev* 1999;**8**(10)**:**855-61.
23. Meadows AT. Second tumours. *Eur J Cancer* 2001;**37**(16)**:**2074-9; discussion 2079-81.
24. Hemminki K BP. Multiple primary cancers as clues to environmental and heritable causes of cancer and mechanisms of carcinogenesis. *IARC Sci Publ.* 2004;**157:**289-97.

25. Lynch HT, Smyrk T. Hereditary nonpolyposis colorectal cancer (Lynch syndrome). An updated review. *Cancer* 1996;**78**(6)**:**1149-67.
26. Kaiser HE, Nasir A, Groger AM, Link CJ, Jr. The etiology of second primary neoplasms. *In Vivo* 1998;**12**(1)**:**89-93.
27. Fearon ER. Human cancer syndromes: clues to the origin and nature of cancer. *Science* 1997;**278**(5340)**:**1043-50.
28. Lynch HT, Fusaro RM, Lynch J. Hereditary cancer in adults. *Cancer Detect Prev* 1995;**19**(3)**:**219-33.
29. Pharoah PD, Antoniou AC, Easton DF, Ponder BA. Polygenes, risk prediction, and targeted prevention of breast cancer. *N Engl J Med* 2008;**358**(26)**:**2796-803.
30. Lichtenstein P, Holm NV, Verkasalo PK, et al. Environmental and heritable factors in the causation of cancer--analyses of cohorts of twins from Sweden, Denmark, and Finland. *N Engl J Med* 2000;**343**(2)**:**78-85.
31. Czene K, Lichtenstein P, Hemminki K. Environmental and heritable causes of cancer among 9.6 million individuals in the Swedish Family-Cancer Database. *Int J Cancer* 2002;**99**(2)**:**260-6.
32. World Health Organisation. *Prevention of Cancer* (WHO Tech. Rep. 276) (WHO, Geneva, 1964).
33. Peto J. Cancer epidemiology in the last century and the next decade. *Nature* 2001;**411**(6835)**:**390-5.
34. Liu BQ, Peto R, Chen ZM, et al. Emerging tobacco hazards in China: 1. Retrospective proportional mortality study of one million deaths. *Bmj* 1998;**317**(7170)**:**1411-22.
35. Doll R. Cancers weakly related to smoking. *Br Med Bull* 1996;**52**(1)**:**35-49.
36. Doll R, Peto R. The causes of cancer: quantitative estimates of avoidable risks of cancer in the United States today. *J Natl Cancer Inst* 1981;**66**(6)**:**1191-308.
37. IARC. Schistosomes, Liver Flukes and Heliobacter pylori (IARC Monographs on the Evaluation of Carcinogenic Risks to Humans, 61) 1994, Lyon.
38. Walboomers JM, Jacobs MV, Manos MM, et al. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. *J Pathol* 1999;**189**(1)**:**12-9.
39. Kinlen LJ. Infection and childhood leukemia. *Cancer Causes Control* 1998;**9**(3)**:**237-9.
40. Working group on Diet and Cancer of the Committee on Medical Aspects of Food and Nutrition Policy. *Nutritional Aspects of the Development of Cancer* (Dept Health Rep. Health Social Subjects 48). 1998, London.
41. IARC. Some Naturally Occuring Substances: Food Items and Constituents, Heterocyclic Aromatic Amines and Mycotoxins *(IARC Monographs on the Evaluation of Carcinogenic Risks to Humans, 56)* 1993, Lyon.
42. Kojo K, Jansen CT, Nybom P, et al. Population exposure to ultraviolet radiation in Finland 1920-1995: Exposure trends and a time-series analysis of exposure and cutaneous melanoma incidence. *Environ Res* 2006;**101**(1)**:**123-31.
43. Gilchrest BA, Eller MS, Geller AC, Yaar M. The pathogenesis of melanoma induced by ultraviolet radiation. *N Engl J Med* 1999;**340**(17)**:**1341-8.
44. Bishop JN, Harland M, Randerson-Moor J, Bishop DT. Management of familial melanoma. *Lancet Oncol* 2007;**8**(1)**:**46-54.
45. Gandini S, Sera F, Cattaruzza MS, et al. Meta-analysis of risk factors for cutaneous melanoma: II. Sun exposure. *Eur J Cancer* 2005;**41**(1)**:**45-60.
46. Elwood JM. Melanoma and sun exposure: contrasts between intermittent and chronic exposure. *World J Surg* 1992;**16**(2)**:**157-65.
47. Elwood JM. Melanoma and ultraviolet radiation. *Clin Dermatol* 1992;**10**(1)**:**41-50.
48. Armstrong BK, Kricker A. The epidemiology of UV induced skin cancer. *J Photochem Photobiol B* 2001;**63**(1-3)**:**8-18.
49. Linzer DI, Levine AJ. Characterization of a 54K dalton cellular SV40 tumor antigen present in SV40-transformed cells and uninfected embryonal carcinoma cells. *Cell* 1979;**17**(1)**:**43-52.
50. Lane DP, Crawford LV. T antigen is bound to a host protein in SV40-transformed cells. *Nature* 1979;**278**(5701)**:**261-3.

51. Baker SJ, Fearon ER, Nigro JM, et al. Chromosome 17 deletions and p53 gene mutations in colorectal carcinomas. *Science* 1989;**244**(4901)**:**217-21.
52. Malkin D, Li FP, Strong LC, et al. Germ line p53 mutations in a familial syndrome of breast cancer, sarcomas, and other neoplasms. *Science* 1990;**250**(4985)**:**1233-8.
53. Rowley PT. Inherited susceptibility to colorectal cancer. *Annu Rev Med* 2005;**56:**539-54.
54. de la Chapelle A. Genetic predisposition to colorectal cancer. *Nat Rev Cancer* 2004;**4**(10)**:**769-80.
55. Garber JE, Offit K. Hereditary cancer predisposition syndromes. *J Clin Oncol* 2005;**23**(2)**:**276-92.
56. Nishisho I, Nakamura Y, Miyoshi Y, et al. Mutations of chromosome 5q21 genes in FAP and colorectal cancer patients. *Science* 1991;**253**(5020)**:**665-9.
57. Lynch HT, de la Chapelle A. Hereditary colorectal cancer. *N Engl J Med* 2003;**348**(10)**:**919-32.
58. Lynch HT, de la Chapelle A. Genetic susceptibility to non-polyposis colorectal cancer. *J Med Genet* 1999;**36**(11)**:**801-18.
59. Hall JM, Lee MK, Newman B, et al. Linkage of early-onset familial breast cancer to chromosome 17q21. *Science* 1990;**250**(4988)**:**1684-9.
60. Miki Y, Swensen J, Shattuck-Eidens D, et al. A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science* 1994;**266**(5182)**:**66-71.
61. Wooster R, Bignell G, Lancaster J, et al. Identification of the breast cancer susceptibility gene BRCA2. *Nature* 1995;**378**(6559)**:**789-92.
62. Aylon Y, Oren M. Living with p53, dying of p53. *Cell* 2007;**130**(4)**:**597-600.
63. Li J, Yen C, Liaw D, et al. PTEN, a putative protein tyrosine phosphatase gene mutated in human brain, breast, and prostate cancer. *Science* 1997;**275**(5308)**:**1943-7.
64. Hemminki A, Markie D, Tomlinson I, et al. A serine/threonine kinase gene defective in Peutz-Jeghers syndrome. *Nature* 1998;**391**(6663)**:**184-7.
65. Hearle N, Schumacher V, Menko FH, et al. Frequency and spectrum of cancers in the Peutz-Jeghers syndrome. *Clin Cancer Res* 2006;**12**(10)**:**3209-15.
66. Becker KF, Atkinson MJ, Reich U, et al. E-cadherin gene mutations provide clues to diffuse type gastric carcinomas. *Cancer Res* 1994;**54**(14)**:**3845-52.
67. Masciari S, Larsson N, Senz J, et al. Germline E-cadherin mutations in familial lobular breast cancer. *J Med Genet* 2007;**44**(11)**:**726-31.
68. Savitsky K, Bar-Shira A, Gilad S, et al. A single ataxia telangiectasia gene with a product similar to PI-3 kinase. *Science* 1995;**268**(5218)**:**1749-53.
69. Thompson D, Duedal S, Kirner J, et al. Cancer risks and mortality in heterozygous ATM mutation carriers. *J Natl Cancer Inst* 2005;**97**(11)**:**813-22.
70. The CHEK2 Breast Cancer Case-Control Consortium. CHEK2*1100delC and susceptibility to breast cancer: a collaborative analysis involving 10,860 breast cancer cases and 9,065 controls from 10 studies. *Am J Hum Genet* 2004;**74**(6)**:**1175-82.
71. Weischer M, Bojesen SE, Tybjaerg-Hansen A, Axelsson CK, Nordestgaard BG. Increased risk of breast cancer associated with CHEK2*1100delC. *J Clin Oncol* 2007;**25**(1)**:**57-63.
72. Wang W. Emergence of a DNA-damage response network consisting of Fanconi anaemia and BRCA proteins. *Nat Rev Genet* 2007;**8**(10)**:**735-48.
73. Rahman N, Seal S, Thompson D, et al. PALB2, which encodes a BRCA2-interacting protein, is a breast cancer susceptibility gene. *Nat Genet* 2007;**39**(2)**:**165-7.
74. Seal S, Thompson D, Renwick A, et al. Truncating mutations in the Fanconi anemia J gene BRIP1 are low-penetrance breast cancer susceptibility alleles. *Nat Genet* 2006;**38**(11)**:**1239-41.
75. Kamb A, Shattuck-Eidens D, Eeles R, et al. Analysis of the p16 gene (CDKN2) as a candidate for the chromosome 9p melanoma susceptibility locus. *Nat Genet* 1994;**8**(1)**:**23-6.
76. Zuo L, Weger J, Yang Q, et al. Germline mutations in the p16INK4a binding domain of CDK4 in familial melanoma. *Nat Genet* 1996;**12**(1)**:**97-9.
77. Greene MH. The genetics of hereditary melanoma and nevi. 1998 update. *Cancer* 1999;**86**(11 Suppl)**:**2464-77.

78. Zhu G, Duffy DL, Eldridge A, et al. A major quantitative-trait locus for mole density is linked to the familial melanoma gene CDKN2A: a maximum-likelihood combined linkage and association analysis in twins and their sibs. *Am J Hum Genet* 1999;**65**(2)**:**483-92.
79. Harland M, Meloni R, Gruis N, et al. Germline mutations of the CDKN2 gene in UK melanoma families. *Hum Mol Genet* 1997;**6**(12)**:**2061-7.
80. Gillanders E, Juo SH, Holland EA, et al. Localization of a novel melanoma susceptibility locus to 1p22. *Am J Hum Genet* 2003;**73**(2)**:**301-13.
81. Manolio TA, Brooks LD, Collins FS. A HapMap harvest of insights into the genetics of common disease. *J Clin Invest* 2008;**118**(5)**:**1590-605.
82. McCarthy MI, Abecasis GR, Cardon LR, et al. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nat Rev Genet* 2008;**9**(5)**:**356-69.
83. Palmer JS, Duffy DL, Box NF, et al. Melanocortin-1 receptor polymorphisms and risk of melanoma: is the association explained solely by pigmentation phenotype? *Am J Hum Genet* 2000;**66**(1)**:**176-86.
84. Valverde P, Healy E, Sikkink S, et al. The Asp84Glu variant of the melanocortin 1 receptor (MC1R) is associated with melanoma. *Hum Mol Genet* 1996;**5**(10)**:**1663-6.
85. Suzuki I, Cone RD, Im S, Nordlund J, Abdel-Malek ZA. Binding of melanotropic hormones to the melanocortin receptor MC1R on human melanocytes stimulates proliferation and melanogenesis. *Endocrinology* 1996;**137**(5)**:**1627-33.
86. Duffy DL, Box NF, Chen W, et al. Interactive effects of MC1R and OCA2 on melanoma risk phenotypes. *Hum Mol Genet* 2004;**13**(4)**:**447-61.
87. Brown KM, Macgregor S, Montgomery GW, et al. Common sequence variants on 20q11.22 confer melanoma susceptibility. *Nat Genet* 2008;**40**(7)**:**838-40.
88. Gudbjartsson DF, Sulem P, Stacey SN, et al. ASIP and TYR pigmentation variants associate with cutaneous melanoma and basal cell carcinoma. *Nat Genet* 2008;**40**(7)**:**886-91.
89. Sulem P, Gudbjartsson DF, Stacey SN, et al. Two newly identified genetic determinants of pigmentation in Europeans. *Nat Genet* 2008;**40**(7)**:**835-7.
90. Gandini S, Sera F, Cattaruzza MS, et al. Meta-analysis of risk factors for cutaneous melanoma: I. Common and atypical naevi. *Eur J Cancer* 2005;**41**(1)**:**28-44.
91. Gandini S, Sera F, Cattaruzza MS, et al. Meta-analysis of risk factors for cutaneous melanoma: III. Family history, actinic damage and phenotypic factors. *Eur J Cancer* 2005;**41**(14)**:**2040-59.
92. Cui J, Antoniou AC, Dite GS, et al. After BRCA1 and BRCA2-what next? Multifactorial segregation analyses of three-generation, population-based Australian families affected by female breast cancer. *Am J Hum Genet* 2001;**68**(2)**:**420-31.
93. Antoniou AC, Pharoah PD, McMullan G, et al. A comprehensive model for familial breast cancer incorporating BRCA1, BRCA2 and other genes. *Br J Cancer* 2002;**86**(1)**:**76-83.
94. Antoniou AC, Spurdle AB, Sinilnikova OM, et al. Common breast cancer-predisposition alleles are associated with breast cancer risk in BRCA1 and BRCA2 mutation carriers. *Am J Hum Genet* 2008;**82**(4)**:**937-48.
95. Easton DF, Pooley KA, Dunning AM, et al. Genome-wide association study identifies novel breast cancer susceptibility loci. *Nature* 2007;**447**(7148)**:**1087-93.
96. Ellis NA, Kirchhoff T, Mitra N, et al. Localization of breast cancer susceptibility loci by genome-wide SNP linkage disequilibrium mapping. *Genet Epidemiol* 2006;**30**(1)**:**48-61.
97. Gold B, Kirchhoff T, Stefanov S, et al. Genome-wide association study provides evidence for a breast cancer risk locus at 6q22.33. *Proc Natl Acad Sci U S A* 2008;**105**(11)**:**4340-5.
98. Huijts PE, Vreeswijk MP, Kroeze-Jansema KH, et al. Clinical correlates of low-risk variants in FGFR2, TNRC9, MAP3K1, LSP1 and 8q24 in a Dutch cohort of incident breast cancer cases. *Breast Cancer Res* 2007;**9**(6)**:**R78.
99. Hunter DJ, Kraft P, Jacobs KB, et al. A genome-wide association study identifies alleles in FGFR2 associated with risk of sporadic postmenopausal breast cancer. *Nat Genet* 2007;**39**(7)**:**870-4.

100. Murabito JM, Rosenberg CL, Finger D, et al. A genome-wide association study of breast and prostate cancer in the NHLBI's Framingham Heart Study. *BMC Med Genet* 2007;**8 Suppl 1:**S6.
101. Stacey SN, Manolescu A, Sulem P, et al. Common variants on chromosomes 2q35 and 16q12 confer susceptibility to estrogen receptor-positive breast cancer. *Nat Genet* 2007;**39**(7)**:**865-9.
102. Stacey SN, Manolescu A, Sulem P, et al. Common variants on chromosome 5p12 confer susceptibility to estrogen receptor-positive breast cancer. *Nat Genet* 2008;**40**(6)**:**703-6.
103. Lindmark F, Zheng SL, Wiklund F, et al. H6D polymorphism in macrophage-inhibitory cytokine-1 gene associated with prostate cancer. *J Natl Cancer Inst* 2004;**96**(16)**:**1248-54.
104. Sun J, Wiklund F, Zheng SL, et al. Sequence variants in Toll-like receptor gene cluster (TLR6-TLR1-TLR10) and prostate cancer risk. *J Natl Cancer Inst* 2005;**97**(7)**:**525-32.
105. Zheng SL, Augustsson-Balter K, Chang B, et al. Sequence variants of toll-like receptor 4 are associated with prostate cancer risk: results from the CAncer Prostate in Sweden Study. *Cancer Res* 2004;**64**(8)**:**2918-22.
106. Lindmark F, Zheng SL, Wiklund F, et al. Interleukin-1 receptor antagonist haplotype associated with prostate cancer risk. *Br J Cancer* 2005;**93**(4)**:**493-7.
107. Zheng SL, Liu W, Wiklund F, et al. A comprehensive association study for genes in inflammation pathway provides support for their roles in prostate cancer risk in the CAPS study. *Prostate* 2006;**66**(14)**:**1556-64.
108. Liotta L. Principles of molecular cell biology of cancer: cancer metastasis. In: De Vita SHV RS, ed. Cancer: Principles and Practice of Oncology. Fourth ed. ed. Philadelphia: Lippincott Co, 1993: 134-149.
109. Eccles SA, Welch DR. Metastasis: recent discoveries and novel treatment strategies. *Lancet* 2007;**369**(9574)**:**1742-57.
110. Hunter K, Welch DR, Liu ET. Genetic background is an important determinant of metastatic potential. *Nat Genet* 2003;**34**(1)**:**23-4; author reply 25.
111. Hunter KW. Allelic diversity in the host genetic background may be an important determinant in tumor metastatic dissemination. *Cancer Lett* 2003;**200**(2)**:**97-105.
112. Kvale R, Auvinen A, Adami HO, et al. Interpreting trends in prostate cancer incidence and mortality in the five Nordic countries. *J Natl Cancer Inst* 2007;**99**(24)**:**1881-7.
113. Eccles SA, Box G, Court W, Sandle J, Dean CJ. Preclinical models for the evaluation of targeted therapies of metastatic disease. *Cell Biophys* 1994;**24-25:**279-91.
114. Heimann R, Lan F, McBride R, Hellman S. Separating favorable from unfavorable prognostic markers in breast cancer: the role of E-cadherin. *Cancer Res* 2000;**60**(2)**:**298-304.
115. Heimann R, Hellman S. Clinical progression of breast cancer malignant behavior: what to expect and when to expect it. *J Clin Oncol* 2000;**18**(3)**:**591-9.
116. Schmidt-Kittler O, Ragg T, Daskalakis A, et al. From latent disseminated cells to overt metastasis: genetic analysis of systemic breast cancer progression. *Proc Natl Acad Sci U S A* 2003;**100**(13)**:**7737-42.
117. Holmgren L. Antiangiogenis restricted tumor dormancy. *Cancer Metastasis Rev* 1996;**15**(2)**:**241-5.
118. Naumov GN, MacDonald IC, Weinmeister PM, et al. Persistence of solitary mammary carcinoma cells in a secondary site: a possible contributor to dormancy. *Cancer Res* 2002;**62**(7)**:**2162-8.
119. Riethmuller G, Klein CA. Early cancer cell dissemination and late metastatic relapse: clinical reflections and biological approaches to the dormancy problem in patients. *Semin Cancer Biol* 2001;**11**(4)**:**307-11.
120. Punglia RS, Morrow M, Winer EP, Harris JR. Local therapy and survival in breast cancer. *N Engl J Med* 2007;**356**(23)**:**2399-405.
121. Humphrey LL, Helfand M, Chan BK, Woolf SH. Breast cancer screening: a summary of the evidence for the U.S. Preventive Services Task Force. *Ann Intern Med* 2002;**137**(5 Part 1)**:**347-60.
122. Ragaz J, Olivotto IA, Spinelli JJ, et al. Locoregional radiation therapy in patients with high-risk breast cancer receiving adjuvant chemotherapy: 20-year results of the British Columbia randomized trial. *J Natl Cancer Inst* 2005;**97**(2)**:**116-26.

123. Overgaard M, Hansen PS, Overgaard J, et al. Postoperative radiotherapy in high-risk premenopausal women with breast cancer who receive adjuvant chemotherapy. Danish Breast Cancer Cooperative Group 82b Trial. *N Engl J Med* 1997;**337**(14)**:**949-55.
124. Overgaard M, Jensen MB, Overgaard J, et al. Postoperative radiotherapy in high-risk postmenopausal breast-cancer patients given adjuvant tamoxifen: Danish Breast Cancer Cooperative Group DBCG 82c randomised trial. *Lancet* 1999;**353**(9165)**:**1641-8.
125. Ragaz J, Jackson SM, Le N, et al. Adjuvant radiotherapy and chemotherapy in node-positive premenopausal women with breast cancer. *N Engl J Med* 1997;**337**(14)**:**956-62.
126. Nielsen HM, Overgaard M, Grau C, Jensen AR, Overgaard J. Loco-regional recurrence after mastectomy in high-risk breast cancer--risk and prognosis. An analysis of patients from the DBCG 82 b&c randomization trials. *Radiother Oncol* 2006;**79**(2)**:**147-55.
127. Fidler IJ. Critical determinants of metastasis. *Semin Cancer Biol* 2002;**12**(2)**:**89-96.
128. Alessandro R, Kohn EC. Signal transduction targets in invasion. *Clin Exp Metastasis* 2002;**19**(3)**:**265-73.
129. Luzzi KJ, MacDonald IC, Schmidt EE, et al. Multistep nature of metastatic inefficiency: dormancy of solitary cells after successful extravasation and limited survival of early micrometastases. *Am J Pathol* 1998;**153**(3)**:**865-73.
130. Schirrmacher V, Fogel M, Russmann E, Bosslet K, Altevogt P, Beck L. Antigenic variation in cancer metastasis: immune escape versus immune control. *Cancer Metastasis Rev* 1982;**1**(3)**:**241-74.
131. Haeusler J, Hoegel J, Bachmann N, et al. Association of a CAV-1 haplotype to familial aggressive prostate cancer. *Prostate* 2005;**65**(2)**:**171-7.
132. Witte JS, Goddard KA, Conti DV, et al. Genomewide scan for prostate cancer-aggressiveness loci. *Am J Hum Genet* 2000;**67**(1)**:**92-9.
133. Takahashi S, Shan AL, Ritland SR, et al. Frequent loss of heterozygosity at 7q31.1 in primary prostate cancer is associated with tumor aggressiveness and progression. *Cancer Res* 1995;**55**(18)**:**4114-9.
134. Paiss T, Worner S, Kurtz F, et al. Linkage of aggressive prostate cancer to chromosome 7q31-33 in German prostate cancer families. *Eur J Hum Genet* 2003;**11**(1)**:**17-22.
135. Matakidou A, el Galta R, Webb EL, et al. Genetic variation in the DNA repair genes is predictive of outcome in lung cancer. *Hum Mol Genet* 2007;**16**(19)**:**2333-40.
136. Dubey S, Stephenson P, Levy DE, et al. EGFR dinucleotide repeat polymorphism as a prognostic indicator in non-small cell lung cancer. *J Thorac Oncol* 2006;**1**(5)**:**406-12.
137. Pine SR, Mechanic LE, Ambs S, et al. Lung cancer survival and functional polymorphisms in MBL2, an innate-immunity gene. *J Natl Cancer Inst* 2007;**99**(18)**:**1401-9.
138. Yang P, Yokomizo A, Tazelaar HD, et al. Genetic determinants of lung cancer short-term survival: the role of glutathione-related genes. *Lung Cancer* 2002;**35**(3)**:**221-9.
139. Goode EL, Dunning AM, Kuschel B, et al. Effect of germ-line genetic variation on breast cancer survival in a population-based study. *Cancer Res* 2002;**62**(11)**:**3052-7.
140. Decock J, Long JR, Laxton RC, et al. Association of matrix metalloproteinase-8 gene variation with breast cancer prognosis. *Cancer Res* 2007;**67**(21)**:**10214-21.
141. Hughes S, Agbaje O, Bowen RL, et al. Matrix metalloproteinase single-nucleotide polymorphisms and haplotypes predict breast cancer progression. *Clin Cancer Res* 2007;**13**(22 Pt 1)**:**6673-80.
142. Rinker-Schaeffer CW, O'Keefe JP, Welch DR, Theodorescu D. Metastasis suppressor proteins: discovery, molecular mechanisms, and clinical application. *Clin Cancer Res* 2006;**12**(13)**:**3882-9.
143. Berger JC, Vander Griend DJ, Robinson VL, Hickson JA, Rinker-Schaeffer CW. Metastasis suppressor genes: from gene identification to protein function and regulation. *Cancer Biol Ther* 2005;**4**(8)**:**805-12.
144. Nash KT, Welch DR. The KISS1 metastasis suppressor: mechanistic insights and clinical utility. *Front Biosci* 2006;**11:**647-59.
145. Shevde LA, Welch DR. Metastasis suppressor pathways--an evolving paradigm. *Cancer Lett* 2003;**198**(1)**:**1-20.

146. Nordgard SH, Johansen FE, Alnaes GI, et al. Genome-wide analysis identifies 16q deletion associated with survival, molecular subtypes, mRNA expression, and germline haplotypes in breast cancer patients. *Genes Chromosomes Cancer* 2008;**47**(8)**:**680-96.
147. Wilkening S, Tavelin B, Canzian F, et al. Interleukin promoter polymorphisms and prognosis in colorectal cancer. *Carcinogenesis* 2008;**29**(6)**:**1202-6.
148. Gulcher J, Stefansson K. Population genomics: laying the groundwork for genetic disease modeling and targeting. *Clin Chem Lab Med* 1998;**36**(8)**:**523-7.
149. Eliason MJ, Larson AA, Florell SR, et al. Population-based prevalence of CDKN2A mutations in Utah melanoma families. *J Invest Dermatol* 2006;**126**(3)**:**660-6.
150. Statistics Sweden. Multi-Generation Register 2005, a description of contents and quality: SCB-tryck, 2006. Serie: BE96 Bakgrundsfakta. Befolknings- och Välfärdsstatistik) 2006:6. Available at: http://www.scb.se.
151. Swedish Cancer Registry. Information available at: http://www.socialstyrelsen.se/en/Statistics/statsbysubject/Cancer+Registry.htm
152. Center for Epidemiology. Cancer incidence in Sweden 2000. Stockholm:The National Board of Health and Welfare, 2002.
153. Paltiel O, Schmit T, Adler B, et al. The incidence of lymphoma in first-degree relatives of patients with Hodgkin disease and non-Hodgkin lymphoma: results and limitations of a registry-linked study. *Cancer* 2000;**88**(10)**:**2357-66.
154. Leu M, Czene K, Reilly M. The impact of truncation and missing family links in population-based registers on familial risk estimates. *Am J Epidemiol* 2007;**166**(12)**:**1461-7.
155. Schoenfeld D. Partial residuals for the proportional hazards regression model. *Biometrika* 1982;**69:**239-41.
156. Neale M. Mx statistical modelling. Richmond, Medical College of Virginia, 1994.
157. Hemminki K, Zhang H, Czene K. Incidence trends and familial risks in invasive and in situ cutaneous melanoma by sun-exposed body sites. *Int J Cancer* 2003;**104**(6)**:**764-71.
158. Grodstein F, Speizer FE, Hunter DJ. A prospective study of incident squamous cell carcinoma of the skin in the nurses' health study. *J Natl Cancer Inst* 1995;**87**(14)**:**1061-6.
159. Zhou W, Heist RS, Liu G, et al. Smoking cessation before diagnosis and survival in early stage non-small cell lung cancer patients. *Lung Cancer* 2006;**53**(3)**:**375-80.
160. Ebbert JO, Williams BA, Sun Z, et al. Duration of smoking abstinence as a predictor for non-small-cell lung cancer survival in women. *Lung Cancer* 2005;**47**(2)**:**165-72.
161. Tammemagi CM, Neslund-Dudas C, Simoff M, Kvale P. Smoking and lung cancer survival: the role of comorbidity and treatment. *Chest* 2004;**125**(1)**:**27-37.
162. Yoshino I, Maehara Y. Impact of smoking status on the biological behavior of lung cancer. *Surg Today* 2007;**37**(9)**:**725-34.
163. Lorenzo Bermejo J, Hemminki K. Familial lung cancer and aggregation of smoking habits: a simulation of the effect of shared environmental factors on the familial risk of cancer. *Cancer Epidemiol Biomarkers Prev* 2005;**14**(7)**:**1738-40.
164. Bricker JB, Peterson AV, Jr., Leroux BG, Andersen MR, Rajan KB, Sarason IG. Prospective prediction of children's smoking transitions: role of parents' and older siblings' smoking. *Addiction* 2006;**101**(1)**:**128-36.
165. Griesbach D, Amos A, Currie C. Adolescent smoking and family structure in Europe. *Soc Sci Med* 2003;**56**(1)**:**41-52.
166. West P, Sweeting H, Ecob R. Family and friends' influences on the uptake of regular smoking from mid-adolescence to early adulthood. *Addiction* 1999;**94**(9)**:**1397-411.
167. Sutton GC. Assortative marriage for smoking habits. *Ann Hum Biol* 1980;**7**(5)**:**449-56.
168. Price RA, Vandenberg SG. Spouse similarity in American and Swedish couples. *Behav Genet* 1980;**10**(1)**:**59-71.
169. Clark AE, Etile F. Don't give up on me baby: spousal correlation in smoking behaviour. *J Health Econ* 2006;**25**(5)**:**958-78.

170. Moreno V, Gemignani F, Landi S, et al. Polymorphisms in genes of nucleotide and base excision repair: risk and prognosis of colorectal cancer. *Clin Cancer Res* 2006;**12**(7 Pt 1)**:**2101-8.
171. Kim JG, Chae YS, Sohn SK, et al. Vascular endothelial growth factor gene polymorphisms associated with prognosis for patients with colorectal cancer. *Clin Cancer Res* 2008;**14**(1)**:**62-6.