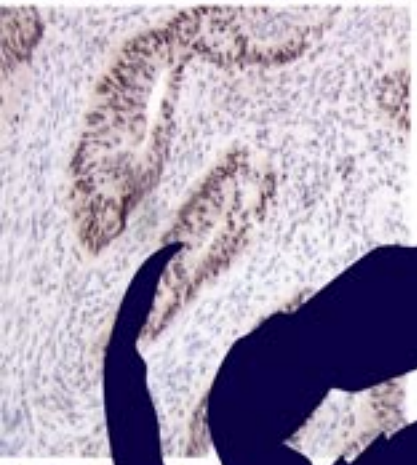
Diet, Lifestyle, Heritable Factors and Colorectal Carcinogenesis:

Associations with Histopathological and Molecular Endpoints



Petra A. Wark

Diet, Lifestyle, Heritable Factors and Colorectal Carcinogenesis: Associations with Histopathological and Molecular Endpoints

Petra A. Wark

Promotoren:

Prof. Dr. Ir. P. van 't Veer, hoogleraar Voeding en Epidemiologie, Wageningen Universiteit
Prof. Dr. Ir. F.J. Kok, hoogleraar Voeding en Gezondheid, Wageningen Universiteit

Co-promotor:

Dr. Ir. M.P. Weijenberg, universitair docent, capaciteitsgroep Epidemiologie, Universiteit Maastricht

Promotiecommissie:

Dr. Ir. R.A. Bausch-Goldbohm, TNO Kwaliteit van Leven, Zeist
Prof. Dr. L.A.L.M. Kiemeney, Radboud Universiteit, Nijmegen
Prof. Dr. J.H.J.M van Krieken, Radboud Universiteit, Nijmegen
Prof. Dr. Ir. I.M.C.M. Rietjens, Wageningen Universiteit

Dit onderzoek is uitgevoerd binnen de onderzoeksschool VLAG.

Diet, Lifestyle, Heritable Factors and Colorectal Carcinogenesis: Associations with Histopathological and Molecular Endpoints

Petra A. Wark

Proefschrift
ter verkrijging van de graad van doctor
op gezag van de rector magnificus
van Wageningen Universiteit,
Prof. dr. M.J. Kropff,
in het openbaar te verdedigen
op dinsdag 16 januari 2007
des namiddags te vier uur in de Aula

Petra A. Wark

Diet, Lifestyle, Heritable Factors and Colorectal Carcinogenesis:
Associations with Histopathological and Molecular Endpoints

Thesis Wageningen University - With references - With summary in Dutch

ISBN: 90-8504-575-4

"... the idea of a sharp distinction
between health and disease is a
medical artefact for which nature,
if consulted, provides no support."

Geoffrey A Rose,
referring to Hamilton M et al, Clin Sci (Lond). 1954; 13(1): 11-35.
In: Rose GA, The strategy of preventive medicine. Oxford: Oxford University Press; 1992:6.

ABSTRACT

Background: Diet, lifestyle and heritable factors have been related to colorectal cancer risk; to date, their relevance to the overall scope of colorectal carcinogenesis, has not been clearly established.

Aim and Methods: To evaluate whether distinguishing colorectal tissue by its histopathological and molecular characteristics sheds further light on the etiology of colorectal cancer. Five research questions addressed associations between diet, lifestyle and heritable factors, and specific tissue characteristics.

Results: First, we observed that consumption of fruits, in particular citrus fruits, was associated with increased rectal glutathione *S*-transferase activity in a cross-sectional study of 94 Dutch individuals. Consumption of cruciferous vegetables was also associated with increased activity, but only among individuals who carried the *GSTM1* genotype.

Second, we observed that intake of vitamin B2 was inversely associated with adenomas with a K-*ras* mutation (n=81) but not with adenomas without a K-*ras* mutation (n=453) in a case-control study conducted in the Netherlands. A positive association with monounsaturated fat was confined to K-*ras* mutation-negative adenomas. We found indications for differential associations with some additional factors, but the epidemiological evidence on risk factors and K-*ras* mutations remains inconsistent.

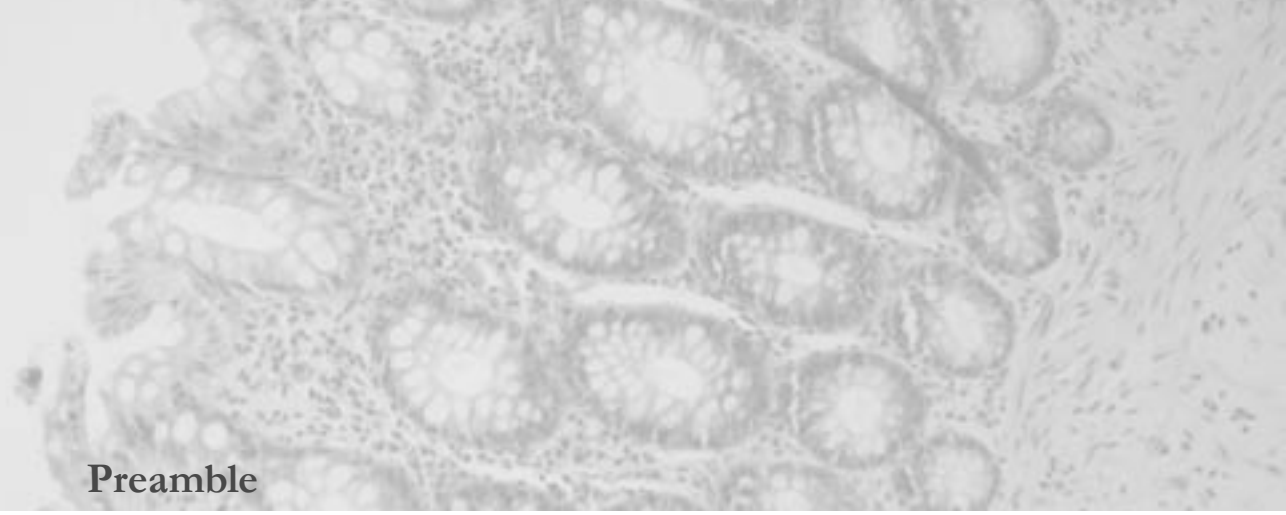
Third, in a cohort study of 26,769 American men, we observed that most risk factors were similarly associated with advanced (≥1cm or with any villous characteristics or carcinoma in situ) and non-advanced colorectal adenomas after 17 years of follow-up. However, smoking had a stronger positive association with advanced adenomas than with non-advanced adenomas, and a high glycemic index was inversely associated with advanced but not with non-advanced adenomas.

Fourth, associations with family history of colorectal cancer were stronger for men with multiple distal adenomas than for men with a single distal adenoma at first diagnosis, in the aforementioned cohort study among US men. Associations between family history, and advanced and non-advanced adenomas, were of similar strength, but a tendency towards a somewhat stronger association with non-advanced adenomas was found.

Fifth, fruit consumption was inversely associated with hMLH1 protein-deficient colon cancer (n=54) but not with hMLH1 protein-proficient colon cancer (n=387) in a cohort study of 120,852 people who were followed-up over 7.3 years, while ignoring information from the initial 2.3 years of follow-up. Clear associations with consumption of vegetables, or nutrients related to fruits and vegetables, could not be detected.

Conclusions: We showed that distinguishing colorectal tumors by their histopathological and molecular characteristics may indeed shed further light on the role of diet, lifestyle and heritable factors in colorectal carcinogenesis. Such an approach may alleviate some of the weaknesses of traditional epidemiology, but also adds another layer of complexity. It is a challenge for the future to develop a framework into which specific associations can be integrated, using risk markers signaling the molecular and biochemical pathways from normal to cancerous tissue.

Contents



Preamble

This chapter intends to provide a comprehensive introduction to how studies using histopathological and molecular markers, such as those described later in this thesis, may improve the understanding of the role of dietary, lifestyle and heritable factors in the causation of colorectal cancer. The chapter consists of four main parts:

- **Part I**: focuses on cancer in general. It gives an overview of the main characteristics of cancer development, and it introduces relevant medical and scientific terminology at a basic level. If you are already familiar with this, you can skip this part and continue with part II, III or even IV.
- **Part II**: focuses on colorectal cancer. It explains what colorectal cancer is and how often it occurs; it mentions the main characteristics of various steps and stages that are involved in its development. Postulated risk factors are also mentioned.
- **Part III**: describes how studies incorporating histopathological and molecular markers may help to improve the understanding of the role of risk factors in (colorectal) carcinogenesis.
- **Part IV**: describes the aims and outline of this thesis, and broad and specific research questions are formulated.

An **index**, containing a listing of the terminology introduced in this chapter and their corresponding page numbers, can be found at the last page of this chapter.

1 Introduction

PART I: WHAT IS CANCER?

To many people, cancer is one of the most feared diseases. They believe that cancer is always painful, that the treatment is worse than the disease and that death is inevitable. Besides, the fact that something grows in the body over which one has no control is fearful to many, and once being diagnosed with the disease, cancer is a sword of Damocles that often continues to hang over the patient for the rest of his life[1].

CANCER is a group of diseases characterized by abnormal growth and unlimited division of CELLS, which are the basic structural and functional units of our body. Groups of cells form masses, TUMORS, which may be benign or malignant. Cancer is a synonym for malignant tumors, and they may eventually spread to other parts of the body.

Cancer is a major public health problem. In many countries, cancer causes more than a quarter of the deaths. Cancer accounted for about 12.5% of the deaths worldwide in 2002[2], when seven million people died of cancer[3]. In the same year, 11 million people were diagnosed with cancer and nearly 25 million people were alive with cancer (within three years of their diagnosis)[3]. In 2020, there could be as many as 15 million new cases per year[2].

But there is good news too. At least a third of all cancers are likely to be preventable[4], as many cancers are due to modifiable risk factors. Cancer prevention strategies include increasing awareness of people of unhealthy behavior (*e.g.* warning people of eating blackened meat), enhancing healthy behavior (*e.g.* promoting physical exercise and consumption of fruits and vegetables), and governmental actions such as banning smoking from public places. However, evidence on which risk factors are involved in the causation of cancer is not entirely consistent, and the story becomes more complex because genetic factors play a role in cancer causation, either directly or by modulating the effect of environmental factors on the development of cancer.

In this part, we explain the basics of cancer development and how dietary, lifestyle and heritable factors may interfere with cancer development.

From Normal Cells to Cancer

Processes involved in the transformation of a normal cell to a cancerous cell are not completely understood. Most knowledge on this transformation process, which is called CARCINOGENESIS, has been gained through careful comparison of cancerous tissue with normal tissue.

Changes that occur during carcinogenesis can be classified according to the level at which they are visible. HISTOPATHOLOGICAL CHANGES may occur in the organization of groups of cells that are similar in structure and that perform one or more common functions: a TISSUE. When we have a closer look at what happens within the tissue, we observe that cancerous cells look and behave differently than normal cells: CYTOPATHOLOGICAL CHANGES have occurred in the cancerous cells. The genetic makeup of cancerous cells also differs from the genetic makeup of normal cells. We summarize the most important changes for each of these levels separately below.

Histopathological Changes: Changes to the Tissue

A common first signal of carcinogenesis is the occurrence of hyperplasia, which refers to the presence of an abnormally high amount of cells in a specific area. When the transformation to cancer goes on, normal cellular architecture and orientation are gradually lost and the tissue loses its normal organization little by little; dysplasia is said to occur. Hyperplasia and dysplasia can occur in many types of cells, but they occur most frequently in epithelial cells, which are the cells that form the skin and line the organs. Cancers that arise from such cells are called carcinomas.

It is very hard to distinguish between a benign tumor with severe dysplasia and carcinoma in situ. Carcinoma in situ is cancer that is confined to the very local area where the cancer first developed, and it has not yet spread to surrounding areas or other tissues. When the tumor grows through the thin layer that separates the epithelial cells from the underlying supporting framework, it is called invasive or infiltrative cancer. Invasive cancers have a poorly defined boundary, their surrounding area is frequently inflamed, and they are often surrounded by a network of blood vessels. The cancer is able to create such blood vessels itself, the process of which is called angiogenesis; by doing so, the tumor ensures its own oxygen and nutrients supplies.

Invasive cancers may eventually grow into blood or lymph vessels, which transport cells that are important in the defense against infections. When a cancerous cell enters the bloodstream or lymph, it may be transported to other parts of the body, where the cancerous cell may settle down and divide. This may result in the creation of new tumors, which are called metastases. How large a cancer has grown and how far it has spread are reflected in the cancer's stage. The stage of a cancer is closely related to the prognosis, the expected clinical course and predicted outcome of the disease.

Cytopathological Changes: Changes to the Cell

Under healthy conditions, most cells divide from time to time to replace cells that die. In cancer, the balance between cell division and cell death is disturbed: the cancer cells stimulate themselves to grow and divide, they do not respond to signals that tell them to stop dividing and they often escape apoptosis, which is programmed cell death[5]. Consequently, fewer cells die than new cells are formed, which is called increased proliferation. Cells that allow the cancer to grow faster or let it survive longer are selected, and these cells proliferate further.

Whereas hyperplastic cells still preserve normal architecture and retain a normal shape, these are gradually lost in dysplasia. Cancer cells also often have an abnormal size as well as a much larger and irregularly shaped nucleus, which is the cell's controlling centre responsible for growth and reproduction. As a result of containing an overload of DNA - the genetic material that carries all the instructions for the composition, growth and development of our body - the nuclei of tumor cells often appear to be darker under the microscope after coloring with certain dyes than those of normal cells.

Cancer cells lack the ability to communicate well with their neighboring cells[6] and cancer cells adhere to their neighbors badly; the latter partly accounts for their ability to metastasize. Cancer

cells also stop maturing early on; they lack the specialized functions of normal adult cells and are said to show lack of DIFFERENTIATION.

MOLECULAR CHANGES: CHANGES TO THE GENETIC MAKEUP OF A CELL

During cancer development, the genetic makeup of a cell undergoes several changes. In this section, we briefly explain the structure and function of the DNA. Thereafter, we describe the different types of genetic alterations and their possible relation to cancer causation.

Background: the DNA

Each cell contains a complete package of someone's genetic blueprint, the GENOME, the vast majority of which is found in the nucleus. The human genome is organized in 46 CHROMOSOMES that are arranged in 23 pairs, with one member of each pair inherited from each parent at the time of conception. Each CHROMOSOME consists of a long string of coiled-up DNA. This string consists of two long chains of NUCLEOTIDES, the basic units of the DNA, which are twisted into a double helix and are joined by so-called hydrogen bonds[7]. Each nucleotide consists of a sugar component, a phosphate molecule and one of four molecular building blocks, BASES, which are GUANINE (G), CYTOSINE (C), THYMINE (T) and ADENINE (A). Related to their comparable chemical structure, adenine and guanine are classified as PURINES, and cytosine and thymine as PYRIMIDINES. An adenine at one chain of the helix always combines with a thymine at the other chain, and a cytosine always pairs with a guanine base[7].

Each person has a unique genetic blueprint as defined by a unique sequence of nucleotides, which defines the hereditary characteristics. The major part of the sequence does not vary between people however, as humans share many basic characteristics. Around one percent of the DNA is CODING DNA that determines which PROTEINS are produced; proteins are required for the structure, function and regulation of the body's cells and tissues, and consist of amino acids. There are 20 different AMINO ACIDS, which are simple organic compounds containing carbon, hydrogen, oxygen, nitrogen and sometimes sulphur. Three adjacent nucleotide bases determine which amino acid might be formed; the next three bases determine the next amino acid, etcetera. GENES, consisting of a sequence of nucleotides, dictate which protein is produced. There are two copies of each gene, called ALLELES, in each cell as chromosomes come in pairs.

Genes have to be turned on before protein production can start. Protein production starts by TRANSCRIPTION, during which a single strand of DNA is copied in a single strand of nucleotides called RNA. After a few intermediate steps, the RNA gives off a message to produce specific proteins, and this process is called TRANSLATION. The function of the majority of the NON-CODING DNA is unknown, but many non-coding sequences are thought to have regulatory roles[8].

Types of Genetic Alterations

During the development of cancer, many genetic alterations occur. These alterations comprise MUTATIONS, which are direct alterations to the nucleotide sequence, and EPIGENETIC MODIFICATIONS, which are changes that may indirectly influence the function of a gene without affecting the

nucleotide sequence. If mutations or epigenetic modifications occur, altered versions of proteins may be produced and their function in the body could be affected.

Mutations can be inherited or can be acquired over the course of life. Inherited mutations are called GERMLINE MUTATIONS, and they already occur in sperm and egg. However, the majority of the mutations are acquired during lifetime and they occur in cells other than sperm and egg. These are called SOMATIC MUTATIONS.

Many different types of mutations exist. POINT MUTATIONS are also called BASE SUBSTITUTIONS because one nucleotide base is converted into another one. Point mutations can be further subdivided into four subtypes according to the functional change: missense, nonsense, silent and frameshift mutations. A NONSENSE MUTATION is a specific type of point mutation that results in an incomplete protein. A MISSENSE MUTATION refers to the replacement of one amino acid by another one, which may lead to a protein with altered function. However, a point mutation is not always harmful; the same amino acid is still produced when a SILENT MUTATION occurs. FRAMESHIFT mutations occur when a nucleotide is deleted or inserted in the coding section of a gene. Frameshift mutations are likely to result in a protein that was not intended to be produced; as a sequence of three adjacent nucleotide bases determines which amino acid is produced, insertion and deletions change the reading frame of the nucleotide sequence and thereby easily the proteins that are produced.

Alternatively, point mutations can be classified according to the bases that are substituted. When a purine base is exchanged for another purine or a pyrimidine base is exchanged for another pyrimidine, *i.e.* A for G, G for A, C for T, or T for C, a TRANSITION MUTATION has occurred. When a purine is exchanged for a pyrimidine, or vice versa, a TRANSVERSION MUTATION has occurred.

Mutations can also occur at the chromosomal level. They comprise TRANSLOCATIONS, where a segment of a chromosome is moved to another chromosome, typically to one of a different pair; LOSS OF HETEROZYGOSITY (LOH) when part of a chromosome is deleted; and gene AMPLIFICATIONS, where there is an increase in the number of copies of a specific DNA fragment.

Epigenetic modifications form a different type of category of genetic alterations. Epigenetic modifications refer to changes in the PHENOTYPE, the observable traits or characteristics of a person, without accompanying changes in the GENOTYPE, the specific genetic makeup of an individual. Many epigenetic events are part of normal cell function, but may be deregulated in carcinogenesis. They include DNA METHYLATION. METHYL GROUPS, small organic group consisting of one carbon which is single-bonded to three hydrogen molecules, may bind to the DNA and block expression of a specific gene thereby acting as a switch to turn genes on and off, to control which proteins are produced and at which moment. Furthermore, such binding affects the structural stability of the DNA. When such regulatory functions are deregulated, such as in HYPO- (too few methyl groups attached to the DNA) or HYPERMETHYLATION (too many methyl groups attached to the DNA), they can contribute to the development of cancer[9], together with mutations.

Fortunately, not all mutations are harmful. Besides, our body does its utmost best to protect itself against DNA damage. For instance, a class of proteins checks whether the DNA was

replicated well during cell division, repair proteins mend mistakes that occur during this replication process and the immune system kills cells that are severely damaged or such cells commit suicide themselves. Nevertheless, if DNA errors are not repaired or cells with severe damage are not killed, cancer could start developing if the errors occur in certain critical genes.

Mutations in Critical Cancer Genes

Alterations to three classes of genes are thought to be important in turning a cell into a cancerous one: genes that produce proteins which stimulate the cell to divide, genes that produce proteins that stop cell division, and stability genes that keep genetic alterations to a minimum[10].

Genes that produce proteins that stimulate the cell to divide into two are called ONCOGENES. Under normal circumstances, signals for cell division are given from time to time. However, when a genetic alteration occurs in a PROTO-ONCOGENE, which then becomes an oncogene, the cell constantly keeps getting signals to divide, and cell division gets out of control.

Genes encoding for proteins that restrain cell growth and division are called TUMOR SUPPRESSOR GENES. If both alleles of a specific tumor suppressor gene are genetically altered, the cell does not give off a signal to stop division. Again, cell division gets out of control.

The class of STABILITY GENES includes DNA REPAIR GENES, which correct errors that occur in the DNA[10], and genes that constantly monitor chromosomes and the DNA. However, mutations may also occur in these genes. If this happens, mutations in other genes may start accumulating[11].

It takes mutations in several oncogenes, tumor suppressor genes or stability genes in a single cell before this cell turns into a malignant one and starts replicating. Because of this, not everyone who inherits a mutation in one such a gene will eventually develop cancer.

Carcinogens, Mutations, Defense Mechanisms and Cancer Development

Somatic mutations may occur spontaneously, typically because errors are made during cell division when the entire genome has to be replicated and distributed to both daughter cells. Insertion of the wrong base, a so-called MISMATCH**,** is the most common error. Mutations may also be induced by cancer-causing agents that are ingested, inhaled, enter the body via the skin, or are produced by the body itself: CARCINOGENS. Carcinogens attack the DNA and may alter its structure. The good news is that damage can be prevented by substances present in the environment assisted by the body's defense mechanisms, even after a cell is hit.

Let us have a look at what happens when a spontaneous mutation occurs or when the body is exposed to a harmful compound (Figure 1-1). A spontaneous mutation may directly result in DNA damage if the damage is not repaired. Harmful compounds generally need to be activated in the body before they can exert their carcinogenic action. The non-activated carcinogen is called the PROCARCINOGEN and the activated carcinogen is called the ULTIMATE CARCINOGEN.

To protect itself from potential damage, our body does its utmost best to dispose of such compounds: special proteins, PHASE I ENZYMES, detect procarcinogens and make them more reactive. Subsequently, PHASE II ENZYMES attack these ultimate carcinogens, inactivate them, enhance their water-solubility and guide them from the body. Some factors present in the environment may modulate the activity of phase I and II enzymes; for instance, tobacco smoke may induce phase I enzymes with potentially deleterious effects[12], whereas substances from

broccoli may prevent carcinogenic damage by increasing the activity of phase II enzymes[13, 14]. ANTIOXIDANTS such as VITAMIN C, E and CAROTENOIDS that are present in our diet may also help to prevent DNA damage. They attack and render REACTIVE OXYGEN SPECIES inert; these highly reactive compounds are constantly formed in the body during normal processes but they may also be derived from environmental factors, such as tobacco smoke[15].

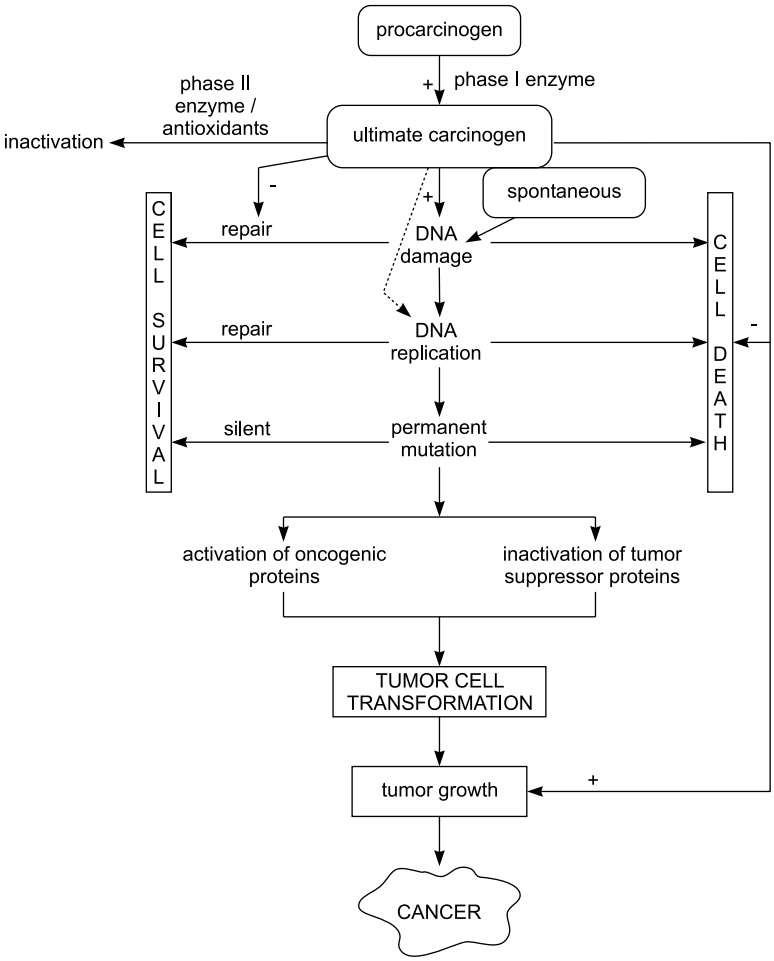


Figure 1-1. Carcinogens, spontaneous mutation and the process of the causation of cancer. After Figures 3-19 and 3-21 in Casarett, 1996[16] and Figure 2.7 in the WCRF report, 1997[17].

In spite of this, ultimate carcinogens may escape inactivation or the carcinogen may cause damage before inactivation can take place. The ultimate carcinogens may bind to the DNA and form complexes, ADDUCTS, or induce other types of DNA damage. If the damage by ultimate carcinogens or the spontaneous mutations are fixable, the cell sends in its DNA repair machinery. If the damage is too extensive, the cell triggers its own death. However, when a mutated cell replicates before it is repaired or fails to die, the mutation is passed on to the two new daughter

cells and becomes permanent. If such an irreversible mutation occurs in a tumor suppressor gene or oncogene, INITATION[18] is said to have taken place. PROMOTION[18] follows initiation, and leads to clonal expansion and tumor development; this process might be reversible if exposure to promoting agents discontinues. As the cells continue to proliferate, additional mutations may lead to PROGRESSION[18] towards cancer. The terminology of initiation, promotion and progression is mainly used for data from the laboratory, but in humans, carcinogens can also act at more than one stage, *e.g.* early and late in carcinogenesis[19].

PART II: COLORECTAL CANCER

Having introduced the basic concepts of cancer development, we next focus on colorectal cancer only. We explain what colorectal cancer is and how often it occurs, and we highlight the main characteristics of the multistep process of colorectal carcinogenesis. Thereafter we discuss whether colorectal cancer is caused by environmental or genetic factors or a combination of both.

What Is Colorectal Cancer and How Often Does It Occur?

COLORECTAL cancer is the medical term for cancer that occurs in the LARGE BOWEL, which forms the last part of the digestive tract. Here, water and useful compounds are removed from digested food passing through, and transferred back into the body. The waste is excreted through the ANUS, where it leaves the body.

Colorectal cancer accounted for about 1 million of new cancer diagnoses and about 529,000 deaths in 2002[20]. It is the third most common cancer in women and the fourth most common cancer in men worldwide, although among both men and women about 9.4% of all new cancers occur in the large bowel[20]. The disease becomes more common with age: around age 40, the risk of colorectal cancer begins to rise, and this risk is highest around age 70[21, 22]. The risk that someone develops colorectal cancer over their lifetime is 5.6% in the Western world[22]. The disease occurs in all regions of the world, but occurs more frequently in the Western world compared with the non-Western world[20].

The large bowel can be subdivided into the colon, the rectosigmoid, and rectum or even further, as illustrated in Figure 1-2. The COLON is the part of the large bowel that starts where the small bowel ends. It is 1.5-1.8 meters long when stretched. The RECTUM forms the final 10-15 cm of the large bowel, and it opens to the outside at the anus. The RECTOSIGMOID is the transitional zone between the colon and the rectum.

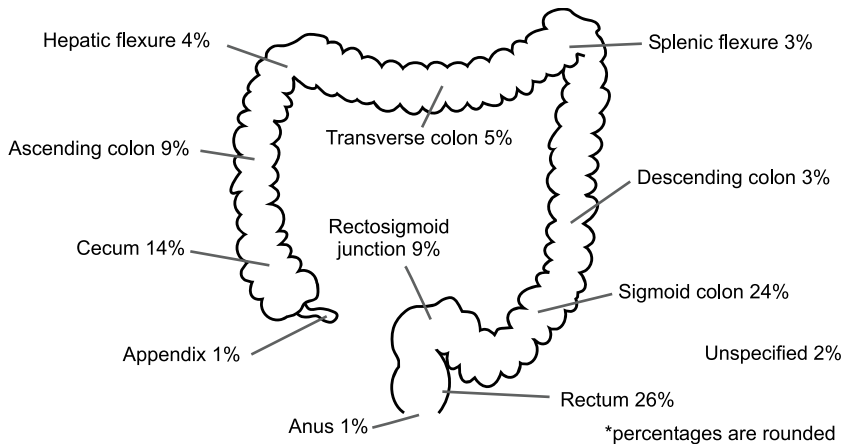


Figure 1-2. Distribution of colorectal cancer over the different areas of the large bowel, based on cancer registrations in the Netherlands Cancer Registry in 2003.

Figure 1-2 also illustrates that colorectal cancer can occur almost everywhere in the large bowel. Although colon cancer occurs more frequently than rectum cancer, the number of patients with rectum cancer is high given the length of this organ. Within the colon, the disease occurs most frequently in the part between the rectum and splenic flexure, the so-called DISTAL COLON or LEFT COLON. In the past, even more cancers occurred in this area; nowadays cancers tend to occur more frequently in the area from the splenic flexure to the cecum - the PROXIMAL or RIGHT COLON - than they did before[23, 24]. Colon cancer occurs as often in men as in women worldwide[17]. However, cancer of the rectum is more common among men[17].

From Healthy Tissue to Colorectal Cancer

Various changes occur during colorectal carcinogenesis. The general histopathological, cytopathological and molecular changes that occur in the development of all cancers, including colorectal carcinogenesis, have already been described. In this section, we introduce specific changes that may occur during colorectal carcinogenesis. Like the corresponding section in part I on cancer in general, this section is subdivided according to the level (tissue, cell or DNA) at which the changes occur.

HISTOPATHOLOGICAL CHANGES: CHANGES TO THE TISSUE

We first describe the specific changes that occur to the tissue during the development of colorectal cancer. We then discuss the implications of these changes in terms of colorectal cancer risk.

Description of Histopathological Changes

Colorectal carcinogenesis starts with hyperplasia of the epithelial cells, and the tissue becomes more and more dysplastic. This process results in the earliest identifiable lesion, the ABERRANT CRYPT FOCUS[25]. The dysplastic tissue may further develop into so-called POLYPS, which are benign tumors. Several types of polyps exist. The ADENOMATOUS POLYP, or ADENOMA, which consists of glandular epithelial tissue that lines the inner layer of the wall of the large bowel, is regarded as the most important type of polyp in colorectal carcinogenesis. Not all adenomas look alike[26]. TUBULAR adenomas, which are the most common subtype[27-30], look like a tube due to branching or infolding of the glandular epithelium. VILLOUS adenomas, which occur least frequently[27-30], have a finger-like appearance. TUBULOVILLOUS adenomas have both finger- and tube-like characteristics. They are more common than villous but less common than tubular adenomas[27-30]. A small adenoma may develop into an intermediate adenoma and subsequently into a large adenoma.

Colorectal cancer can also be classified according to stage or pathological characteristics. Approximately 98% of the colorectal cancers are ADENOCARCINOMAS, which originate from glandular epithelial cells[17]. Traditionally, the system of DUKES has been used for defining the stage of colorectal cancer[31, 32]. In DUKES' A colorectal cancer, the cancer has grown through several layers of the large bowel, but did not yet grow through its muscular wall. DUKES' B colorectal cancers have grown through the wall, but have not yet reached the lymph nodes. When the cancer has spread to at least one lymph node in the nearby area, but not to other body parts,

the cancer is classified as being DUKES' C. DUKES' D is the most advanced stage of colon cancer; the cancer has reached distant organs or tissue, such as the liver or the lungs.

The vast majority of colorectal cancers appear to develop from adenomas[33, 34]. An association between adenomas and colorectal cancer was firstly suggested in 1928[35]. The current body of evidence for this association includes the following:

I. Adenoma tissue can occur within a carcinoma, and carcinoma tissue can occur within an adenoma[36, 37].
II. Adenomas and carcinomas have a similar distribution over the colon[38] and just like carcinomas, nowadays adenomas occur more frequently in the proximal colon than they did before[39].
III. Adenomas occur, on average, earlier in life than carcinomas do[33, 38].
IV. In countries where the incidence of colon cancer is high, adenomas also occur frequently[26, 39].
V. Dietary and lifestyle factors seem to be mostly similarly associated with colorectal adenomas and carcinomas[26].
VI. First-degree relatives of patients with colorectal cancer and those of patients with colorectal adenomas seem to have a similar increased risk of colorectal tumors[40].
VII. Virtually all patients with the familial adenomatous polyposis syndrome, whose colon is scattered with hundreds of adenomas from late childhood or adolescence onwards, develop carcinoma provided the colon is not removed[41].
VIII. Similar changes in the DNA, protein expression and chromosomal constitution are observed in many adenomas and carcinomas. These changes occur to a smaller extent in adenomas than in carcinomas[42-44].
IX. Adenoma patients are at higher risk of getting and dying from colorectal cancer than patients without adenomas[30, 45-47].
X. Adenoma patients whose adenomas were not removed had a 25% risk of developing colorectal cancer at the site of the adenoma after 20 years of follow-up, which is much higher than the risk in the general population[48].
XI. Removal of adenomas reduces the incidence of carcinomas in adenoma patients[49-54].

It has been estimated that colorectal cancer takes at least five years[34] to develop, although most studies estimate that it takes between 10 and 30 years[33, 48, 55-57]. Most of that time is thought to be needed for adenoma formation. The trajectory from small adenomas, large adenomas, carcinoma in situ and eventually to invasive colorectal cancer has been named the ADENOMA-CARCINOMA SEQUENCE.

Carcinomas that do not appear to develop through the adenoma-carcinoma sequence are sometimes called "DE NOVO CANCERS". However, they may have developed from flat adenomas, which are hard to diagnose because they are located close to the epithelial surface[58]. They may also have developed from lesions that do not appear to be adenomas at first glance but that show adenomatous features at the cellular level[59, 60].

Implications of Specific Histopathological Changes

Whereas the lifetime risk of colorectal cancer is 5.6%[22] in Western countries, adenomas are found in about 30-40% of people aged 65 years or older[26, 28, 34, 39, 55]. This suggests that only a subset of adenomas develop into cancer.

Not all adenoma patients are at equal risk of colorectal cancer. Patients with villous adenomas are thought to have a higher risk of colorectal cancer[30, 45, 57, 61] than patients with tubular adenomas, whereas patients having tubulovillous adenomas have a risk that is thought to be intermediate[30, 33, 45, 57, 61]. Large adenomas are thought to be more likely to transfer into a malignant tumor than smaller adenomas[30, 45, 57, 61], and people having multiple adenomas have a higher risk of colorectal cancer than people with only one adenoma[57]. Most adenomas are detected during COLONOSCOPY or SIGMOIDOSCOPY, which are procedures during which the large bowel is inspected using a flexible, lighted tube. Adenomas are generally removed when detected, and patients having adenomas are put under surveillance, as they are likely to develop adenomas again.

Ninety-three percent of the patients with Dukes' A colorectal cancer, 86% of the patients with Dukes' B colorectal cancer, 60% of the patients with Dukes' C colorectal cancer and only 8% of the patients with Dukes' D colorectal cancer are still alive within five years of diagnosis[62]. Thus, the prognosis is better if cancer is detected in an early stage. Colorectal cancer screening aims at doing so; in the US, population screening projects are already established and pilot studies are underway in Europe, *e.g.* in the UK[63] and the Netherlands[64], whereas screening of high-risk patients is already common policy[65].

Cytopathological Changes: Changes to the Cell

The cytopathological chances that occur in colorectal carcinogenesis are equal to those occurring in carcinogenesis in general. During carcinogenesis, cell proliferation increases, the shape and size of the cell and its nucleus change, cell-cell communication becomes impaired, and cancerous cells lack specialized functions. These, and other occurring changes to the cell, have already been described in part I.

Molecular Changes: Changes to the Genetic Makeup of a Cell

During the development of colorectal cancer, several changes in the genetic makeup of colorectal cancer cells occur. Two pathways – the CHROMOSOMAL INSTABILITY PATHWAY and the HYPERMUTABILITY PATHWAY – have been described in colorectal carcinogenesis[66]. These pathways are illustrated in Figure 1-3. The figure shows that different mutations are involved in these pathways, and that additional mutations may occur in the tissue after colorectal cancer development. Whereas cancer arising as a result of the chromosomal instability pathway occurs mainly in the distal colon, cancer arising as the result of the hypermutability pathway mainly arises in the proximal colon[67]. The proximal and distal colon may behave differently because they developed from different parts of the embryo[67]. The chromosomal instability and hypermutability pathway are described in more detail in the next two subsections.

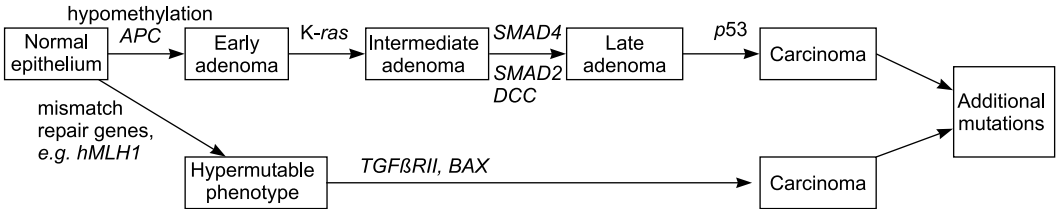


Figure 1-3. Two main pathways involved in the development of colorectal cancer. The boxes indicate the different stages of colorectal carcinogenesis. Arrows connect different stages; the main genes and epigenetic modifications that are thought to play a role in the transition between two stages are mentioned above the arrows. The majority of colorectal cancers are thought to develop via the chromosomal instability pathway, which is the pathway displayed on top. The pathway below is called the hypermutability pathway.

Chromosomal Instability Pathway

The chromosomal instability pathway is largely equal to the model on the key genetic aberrations that are involved in the development of colorectal cancer[42] by Fearon and Vogelstein, which was proposed in 1990[42]. Fearon and Vogelstein's model covered mutations in the Adenomatous Polyposis Coli (*APC*), Kirsten-ras (K-*RAS*), Deleted in Colorectal Carcinoma (*DCC*), and *p53* genes. Additionally, the model covered hypomethylation, which refers to insufficient methylation in large regions of the genome.

Mutations in the tumor suppressor gene *APC* are considered to be early events in the chromosomal instability pathway. The importance of *APC* in the development of the disease is easily visible in patients with the hereditary syndrome Familial Adenomatous Polyposis (FAP). FAP patients are born with mutations in the *APC* gene, and as a result, their colon surface becomes covered with hundreds to thousands of adenomas[41] in late childhood or adolescence. Loss of APC function triggers a chain of other molecular and histological changes[68], including losses and gains of chromosomes and chromosome rearrangements, and *APC* has therefore been proposed to function as a gatekeeper. APC indirectly regulates the function of a number of genes that are critical in the development and maintenance of tissue organization. One of its functions is removal of excess free ß-catenin, a protein that governs events leading to cell proliferation[69], from the cytoplasm of the cell.

Hypomethylation could also be an early event in the chromosome instability pathway. Hypomethylation may lead to further genetic instability[70] and in turn to secondary mutations.

A second gene involved in the chromosome instability pathway is the K-*ras* oncogene, which encodes a protein that conveys growth signals. Fearon and Vogelstein observed that mutations in the K-*ras* gene occurred in approximately 50% of the colorectal carcinomas and adenomas larger than 1 cm, but in fewer than 10% of the adenomas smaller than 1 cm; they hypothesized that mutations in K-*ras* could be responsible for the conversion of small adenomas to larger and more dysplastic ones, which are more likely to progress to cancer[42].

A third gene that is thought to play a role in the chromosomal instability pathway is located at chromosome 18. Fearon and Vogelstein thought that the *DCC* gene was involved, which codes for a protein regulating that cells stick to each other or to non-cellular components[42]. However, it is still not demonstrated that *DCC* itself is relevant. Other genes at chromosome 18 located

closely to *DCC*, such as the tumor suppressor genes *SMAD4 (*Mothers against decapentaplegic homolog-4) and *SMAD2*, may be more important. SMAD2 and SMAD4 proteins form a complex that inhibits growth when they are stimulated by the growth factor Transforming growth factor-beta (TGF-ß)[71]. A role of *SMAD4* is also suggested by its role in the familial Juvenile Polyposis Syndrome, although it only accounts for a few cases[72].

The *p53* tumor suppressor gene is mutated in over 75% of all colorectal carcinomas but mutations are rarely present in adenomas. Therefore, Fearon and Vogelstein hypothesized that *p53* is mainly important in the late stage of colorectal carcinogenesis[42]. *P53* normally activates proteins that prevent DNA replication of cells containing damaged DNA or stimulate the cell to commit suicide when cellular damage occurs[73]. Thus, mutations in *p53* could provide cells with a selective growth advantage that may lead to tumor progression.

Hypermutability Pathway

Approximately 10-20% of all colorectal cancers are thought to arise through the hypermutability pathway. The hypermutability pathway is related to defects in the DNA mismatch repair system[74]. Mismatch repair proteins remove bases that should not have been inserted during DNA replication, and they can also correct insertions or deletions of short stretches of DNA. Defects in mismatch repair occur both in hereditary and sporadic colorectal cancer, but the causes of the defects are different ones. Individuals with the syndrome hereditary non-polyposis colorectal cancer (HNPCC) have germline mutations in one of the mismatch repair genes, *e.g.* in *hMLH1*, whereas sporadic colorectal cancer patients may have inactivated mismatch repair genes due to hypermethylation of their promoter regions, which are the regions that control gene expression, *i.e.* the actual production of the protein that the gene encodes[75-81].

Tumors that are thought to arise via the hypermutability pathway characterize themselves by few mutations in traditional tumor suppressor genes and oncogenes as *APC*, K-*ras* and *p53*. However, mutations in other genes occur, such as in the BCL2-associated X protein gene (*BAX*) and Transforming growth factor-beta type II receptor (*TGF*β*RII*); BAX protein plays a part in regulating programmed cell death[66] and TGFβ is a multifunctional protein that controls, among others, cell proliferation and differentiation. When a mutation in *TGF*β*RII* occurs, the cells are no longer able to respond to TGFβ[66] and control of cell proliferation and differentiation fails.

Is Colorectal Cancer Caused by Genes, the Environment or Both?

A Genetic Disease?

Around five percent of colorectal cancers are caused by inherited genetic mutations, which occur in patients with hereditary colorectal cancer syndromes[41]. People who inherited an altered gene that is strongly related to colorectal cancer have a much higher risk for developing colorectal cancer than people who inherited normal copies. Typically, they also get the disease at a younger age than people with sporadic cancer. Such hereditary syndromes are mostly the result of a single inherited mutation, which is said to have a high penetrance; this means that the vast majority of people who have such a mutation get the disease.

The most common hereditary forms of colorectal cancer are HNPCC and FAP, which have been mentioned before. HNPCC is related to inherited mutations in one of the mismatch repair genes and FAP is caused by inherited mutations in the *APC* gene. Individuals with HNPCC have a lifetime colorectal cancer risk up to 80%[82] and individuals with FAP have a colorectal cancer risk of virtually 100 percent[41].

A study on twins estimated that about 35 percent of all colorectal cancer cases are related to heritable factors[83]. As the majority of colorectal cancers are unrelated to hereditary syndromes, genetic factors must also be involved in sporadic colorectal cancer. Indeed, individuals with a first-degree relative with colorectal cancer have a twofold higher risk of developing colorectal cancer themselves compared with individuals who do not have a first-degree relative with colorectal cancer[84-86]. The risk of colorectal cancer is four times higher for people who have more than one such relative compared with those without one[86]. It is poorly understood which genetic factors are involved. They may include yet unidentified germline mutations of high penetrance, but it is more likely that multiple LOW-PENETRANCE mutations that each confer a low risk of colorectal cancer are involved; an individual carrying a large number of low-penetrance susceptibility alleles may be at relatively high risk of colorectal cancer[87]. Low-penetrance mutations include GENETIC POLYMORPHISMS, which are variations in a gene that appear in at least two percent of a population[88].

A Disease Caused by Environmental Risk Factors?

The rate of colorectal cancer is higher among migrants who moved from a region with a low colorectal cancer incidence to an area in which the disease is more common. The higher rate may be visible within one generation, but even within the migrating generation itself[17, 89]. It points to a role of environmental, modifiable risk factors.

In 1981, Doll and Peto estimated that 35 percent of all cancer deaths may be related to dietary factors[90]. In 1995, Willett estimated that 70-90 percent of colorectal cancers may be avoidable through more health-promoting nutritional and lifestyle practices[91], whereas the aforementioned study on twins suggested that 60 percent of all colorectal cancers is likely to be related to non-shared environmental factors[83].

Although the data are not entirely consistent, a set of modifiable determinants of colorectal cancer has emerged over the years[17, 92-94]:

- ***Red and processed meat*** are believed to increase risk of colorectal cancer[17, 94-97]. Especially cooking meat at high temperatures or over an open fire may increase the risk, as carcinogens are formed during these processes. Examples are heterocyclic amines and polycyclic aromatic hydrocarbons[98]. The possibly harmful effect of processed meat such as sausages, luncheon meats, bacon and hot dogs is thought to be due to the carcinogenic *N*-nitroso compounds that are formed from nitrite preservatives that are added to the meat[98]. It has also been suggested that the high heme iron content of red meat contributes to enhanced cell division and growth[99].

- ***Dietary fat*** is hypothesized to increase risk of colorectal cancer, although evidence is inconsistent[100-102]. A high fat intake increases the amount of primary bile acids in the large bowel. Bacteria may then convert these bile acids to secondary bile acids, which are tumor-promoting agents[103].
- ***Fish*** is hypothesized to decrease risk of colorectal cancer[17, 94, 97]. Omega-3-fatty acids, which occur in fatty fish such as salmon and mackerel, may, among others, slow down cell proliferation[104].
- ***Calcium and vitamin D*** are thought to decrease risk of colorectal cancer[105-107]. Calcium may, among others, bind secondary bile acids and ionized fatty acids in the colon and thereby reduce their proliferative stimuli[108]. Vitamin D gives calcium a hand in reducing cell proliferation and inducing cell differentiation, it helps the body to absorb calcium and also has comparable anticarcinogenic effects itself[105-107].
- ***Fruits and vegetables*** are hypothesized to decrease risk of colorectal cancer, although evidence is inconsistent[17, 94, 109-111]. They are rich in antioxidants that catch harmful species (*e.g.* vitamin C), they contain compounds that increase the activity of phase II detoxification systems (*e.g.* glucosinolates in broccoli), and they contain a lot of fiber[110] (see below).
- ***Fiber*** is hypothesized to decrease risk of colorectal cancer, although studies have shown conflicting results[112, 113]. Fiber-rich food helps to move waste through the colon. It also binds primary bile acids, which inhibits their conversion to secondary bile acids. The latter are thought to promote cancer development[103].
- ***Folate*** is believed to decrease risk of colorectal cancer[114, 115]. Folate is needed to produce intact DNA and RNA, and plays a role in gene expression[115, 116]. Other B-vitamins are also needed for these processes[117]; research on their role in colorectal carcinogenesis is ongoing[118, 119].
- ***Smoking*** is believed to increase risk of colorectal cancer[120, 121]. Tobacco smoke contains at least 50 compounds that are carcinogenic to humans[15]. These carcinogens may, among others, induce DNA strand breaks and thereby cause mutations[15]. Smoking may also increase the activity of phase I enzymes[12]. Much of its damage is done through provision of reactive oxygen species that bind to the DNA.
- ***Alcohol consumption*** is believed to increase risk of colorectal cancer[122-124]. Its breakdown product acetaldehyde is a carcinogen. Alcohol may also increase the carcinogenicity of other carcinogens. In the past, beer has been associated with rectal cancer, possibly via nitrosamines that are currently not present in beer anymore[125]. Alcohol also decreases folate levels[126].
- ***Aspirin and other non-steroidal anti-inflammatory drugs*** are believed to decrease risk of colorectal cancer[127]. They inhibit the enzyme CYCLO-OXYGENASE 2 (COX 2), which plays, among others, a role in inflammation and angiogenesis[128, 129].
- ***High body fatness*** is believed to increase risk of colorectal cancer[92]. This may be related to a high intake of fat or a lack of exercise. The higher risk may also be related to an increased risk of insulin resistance, which exposes the cells to high levels of insulin and possibly high levels of INSULIN GROWTH FACTOR 1 (IGF-1).

- ***Physical activity*** is believed to decrease risk of colorectal cancer, especially of colon cancer[130]. Physical activity may increase bowel movement and enhance disposal of waste, enhance the immune system, decrease obesity and down-regulate hormonal systems that may be related to tumor growth.
- ***Use of birth control pills and postmenopausal hormone use*** are believed to decrease risk of colorectal cancer[93]. Female hormones, ESTROGENS, decrease production of secondary bile acids which are tumor promoters or may lower the level of the growth factor IGF-1[93].

Many other potential modifiable risk factors and protective factors have been suggested, but their role is not yet elucidated or evidence is scarce. Besides, not all evidence is consistent for many of the aforementioned risk factors and many underlying mechanisms remain to be elucidated.

Interplay between Genetic and Environmental Factors?

We have mentioned earlier that, in some cases, a single gene largely determines whether someone gets colorectal cancer. In a few cases, heavy exposure to carcinogens determines whether someone gets colorectal cancer. However, in most cases, the environment and the genetic background of a person determine the risk of colorectal cancer together: the environment affects the activity of the genes and the effect of a certain environmental factor depends on the genes. This interplay is called GENE-ENVIRONMENT INTERACTION. It is generally thought that colorectal cancer is caused by the action of multiple genes, their interactions with each other and the interplay with the environment.

For instance, there is a lot of variation in the way people handle exposure to carcinogens. Let us take red meat consumption as an example. Suppose that a person who can detoxify carcinogens from red meat quickly and another person who can detoxify such carcinogens only slowly consume the same amount of barbecued red meat. The carcinogens present in the meat might be less of a problem for the first person than for the second person, because the first person can get rid of the carcinogens quickly and the carcinogens have less time to cause harm. Thus, the environmental factor (red meat) and the genetic factor (determining detoxification speed) determine the damage together. Concrete examples of genes that determine detoxification activity and capacity are genes that encode the phase II detoxification enzymes GLUTATHIONE *S*-TRANSFERASE M1 (*GSTM1*) and T1 (*GSTT1*), and *N*-ACETYLTRANSFERASE 2 (*NAT2*)[14, 131]. Different versions of these enzymes exist with are associated with different rates of detoxification. Inheritance determines which version a person gets, and subsequently how well this person can deal with certain environmental compounds.

PART III: ELUCIDATING THE ROLE OF RISK FACTORS BY INCORPORATING MOLECULAR AND HISTOPATHOLOGICAL MARKERS

We previously described that many dietary, lifestyle and heritable factors are thought to play a role in colorectal carcinogenesis and that they act in concert; that several changes to the DNA, cells and the tissue go along with carcinogenesis; that colorectal carcinogenesis is a multistage process that gradually presents itself (healthy tissue, adenoma, and carcinoma level), and that colorectal cancer may develop through various pathways (*e.g.* the hypermutability and chromosomal instability pathway).

Studies that associated exposure with colorectal cancer risk have been successful in identifying strong risk factors for colorectal cancer, although the underlying mechanisms have not yet been fully elucidated. Especially the evidence regarding weaker factors is inconsistent; many modifiable factors probably fall into the latter category.

Observational studies now incorporate concepts and techniques from research areas such as genetics and molecular biology. By doing so, researchers do not only hope to improve the quality of the assessment of exposure, but also the understanding of mechanisms underlying associations between risk factors and cancer occurrence; knowledge on such underlying mechanisms enhances the credibility that the risk factor causes cancer. Furthermore, incorporation of concepts and techniques from genetics and molecular biology may help to evaluate whether a risk factor plays a different role in the causation of different cancer subtypes as well as in different stages of the development of cancer.

After summarizing which study designs can be used, we explain why incorporating histopathological and molecular markers within these studies may contribute to improved understanding of the role of dietary, lifestyle and heritable factors in colorectal carcinogenesis in humans.

Background: Overview of Study Designs

Indications for the role of different factors in carcinogenesis can be obtained from *IN VITRO* STUDIES, that study isolated cells or tissues, or from *IN VIVO* STUDIES, in which living organisms are studied. However, studies on humans are the only way to establish whether dietary, lifestyle and heritable factors and pathways are relevant to carcinogenesis in humans. Such studies can be conducted in many ways.

In TRIALS, some people are asked to undertake a certain action (such as taking a supplement) and other people are asked not to undertake that action or to undertake a different action (such as taking a different type of supplement or a fake pill). A random method usually determines who is allocated to which group. Where possible, the researcher usually ensures that the study participant is not aware of the group to which he/she is allocated and that he does not have access to such information himself until the end of the trial. Trials are often regarded as providing the strongest evidence on the role of a certain factor in cancer causation. However, trials generally take long to undertake as people need to be followed for a long time, especially when cancer

occurrence is registered as outcome, and they are usually not conducted until other types of studies suggested that an intervention is promising.

The first identification of potential risk factors come from studies that observe people and their habits in detail, some of which were already conducted before the 20th century[132]. Such initial studies include CASE SERIES that describe exposure to potential risk factors of some selected patients with the disease of interest, and ECOLOGICAL STUDIES that correlate the average exposure to hypothesized risk factors in a certain population with the number of people getting cancer in the same population, *e.g.* they correlate the average meat consumption in various countries with the risk of cancer in these countries. CROSS-SECTIONAL STUDIES link exposure to cancer occurrence in a certain population at a single point in time.

After World War II, more formal approaches to study associations between risk factors and cancer risk came into use, with the formulation of objectives and analysis methods of CASE-CONTROL STUDIES[133-135] and the refinement of the cohort design for large scale studies[136-138]. In a CASE-CONTROL STUDY, both a group of people with a certain disease (the case group) and a group of people without a certain disease (the control group) are selected, after which information on exposure to potential risk factors in the past is collected. The odds of exposure among the cases are then compared with the odds of exposure among the controls. In a COHORT STUDY, exposure to factors of interest is assessed in a group of people without the disease, and these people are followed over time while disease occurrence is recorded. After a number of years, the rate of disease occurrence is associated with the past exposure of interest.

Improving Assessment of Exposure

Traditionally, questionnaires and interviews have been used to assess exposure to potential risk factors in observational studies. However, the assessment of specific dietary risk factors, especially, has been problematic, among others because foods and nutrients are consumed together and are related to other lifestyle habits. Furthermore, the amount of dietary factors that are consumed may not reflect the doses that are relevant to colorectal cancer causation or prevention, because dietary factors have to pass through several absorption barriers and metabolizing circuits before they reach the large bowel. The amount of exposure that eventually reaches the target tissue of interest, *e.g.* the large bowel, is called the INTERNAL DOSE. However, the exposure agents have to reach the DNA of the cells in the large bowel before they can cause damage to the genetic material; that exposure dose is called the MOLECULAR DOSE. Because some environmental agents cause very specific changes to the DNA, such as specific adducts or mutations, these changes can be used as molecular measures of exposure. It is the molecular dose that may lead to mutations in critical cancer genes. If such mutations occur by themselves, they are classified as PRECLINICAL CHANGES, but if they accumulate, they may result in adenomas and eventually carcinomas. Figure 1-4 illustrates this process. The figure also illustrates that factors such as age and genetic susceptibility may influence the associations between external exposure and internal dose, between internal dose and molecular dose, between molecular dose and adenomas, and between adenomas and carcinomas. Taking such internal variation into account in studies may lead to an enhanced understanding of how the external dose is related to the

biological effective dose and thereby to cancer causation, and may result in stronger associations with the disease of interest.

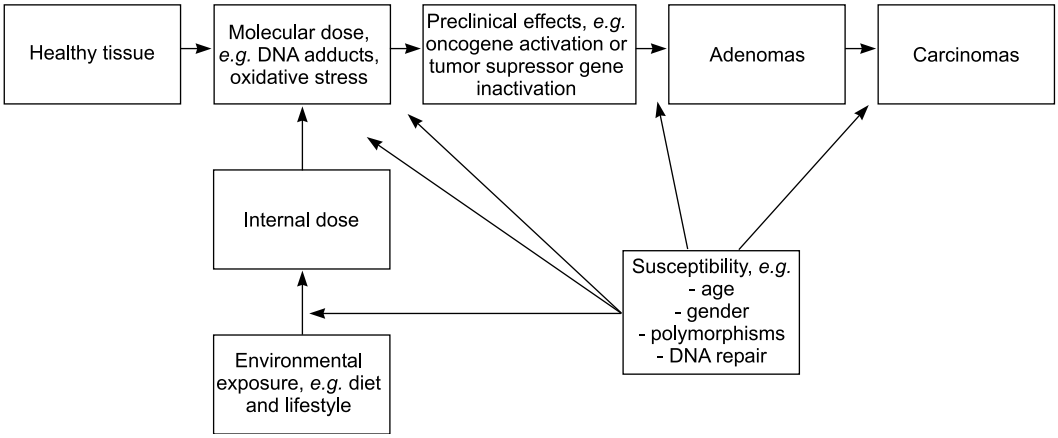


Figure 1-4. Biomarkers may assist in opening the black box behind colorectal cancer development. After Perera, 2000[139].

Elucidating Mechanisms behind Associations of Risk Factors with Cancer Occurrence

Taking molecular or histopathological markers into account may strengthen the evidence regarding the causality of an association between a risk factor and disease occurrence.

First, environmental compounds may cause specific mutations in oncogenes and tumor suppressor genes. Smoking, for instance, has been associated with an excess of G to T transversions in the *p53* gene, which may imply that smoking causes such mutations[140]. However, observing tumors with many of such transversions may also hint at a role of smoking in the causation of these specific tumors.

Second, the relevance of many biological pathways in cancer causation is established or suspected based on research from other fields. If an environmental factor influences such a well-established pathway, this suggests that the factor plays a role in carcinogenesis in humans. Examples of such studies are studies that evaluate whether dietary factors are associated with specific levels of a protein that is relevant in detoxification of carcinogens.

Third, studies incorporating genetic polymorphisms may reveal that some risk factors for cancer are not harmful or beneficial for everyone, and may explain why some associations between risk factors and colorectal cancer differ between study populations even when similar research methodology is used. Some risk factors that are thought to have a weak effect may actually turn out to be strong risk factors but only in a minor subset of people with specific hereditary characteristics.

Fourth, studying intermediate endpoints such as preclinical changes or adenomas focuses on only part of the pathway leading to colorectal cancer; fewer additional factors may affect

the association between the exposure and the endpoint and therefore it is possible that such associations are easier to detect.

Studying Differences between Subtypes of Cancer

We have previously remarked that colorectal cancers vary widely in appearance: some occur in the proximal colon, others occur in the distal colon; some have mutations in a specific cancer gene, others do not have mutations in that gene; some are diagnosed at an early stage, others are diagnosed late; some are well differentiated and others are poorly differentiated, etcetera. These differences may partly be related to a different combination of factors that are involved in their causation, or, in other words, these cancers may have a different ETIOLOGY. If we can distinguish the underlying pathways responsible for these differences, *e.g.* by taking specific molecular or histopathological characteristics into account, the exposure factors that cause cancer may be better understood.

PART IV: AIMS AND OUTLINE OF THIS THESIS

The aim of this thesis is to evaluate whether distinguishing colorectal tissue by its histopathological and molecular characteristics sheds further light on the roles of dietary, lifestyle and heritable factors that are possibly involved in colorectal carcinogenesis. In particular, we focus on the following broad research questions to evaluate this:

- Can habitual consumption levels influence detoxification systems as observed in ***healthy tissue***?
- Are dietary and lifestyle factors that are thought to play a role in colorectal carcinogenesis associated with specific histopathological and molecular characteristics of ***colorectal adenomas***?
- Are dietary and lifestyle factors that are thought to play a role in colorectal carcinogenesis associated with specific histopathological and molecular characteristics of ***colorectal cancer***?

The following set of specific research questions is studied:

- Are parameters reflecting the activity of the rectal glutatione *S*-transferase detoxification system associated with consumption of fruits and vegetables at habitual levels of exposure? If so, do genetic variations in glutathione *S*-transferase genes *GSTM1* and *GSTT1* modify this association? (Chapter 2)
- Do associations between presumed modifiable risk factors for colorectal adenomas depend on K-*ras* mutational status of the adenoma? (Chapter 3)
- Are dietary and lifestyle risk factors for colorectal cancer differently associated with adenomas that are thought to have a high risk of malignancy (advanced adenomas) and adenomas that are thought to have a low risk of becoming malignant? (Chapter 4)?
- Do people with a family history of colorectal cancer have a higher risk of advanced adenomas or do they mostly have a higher likelihood of getting multiple adenomas (Chapter 5)?
- Is consumption of fruits and vegetables differentially associated with colon cancer with and without defects in the mismatch repair system? (Chapter 6)

The biological relevance of the findings is discussed in each corresponding chapter. In Chapter 7, we discuss the used approach and methodology to identify issues that allow improvement of future research.

Acknowledgements

We thank the Netherlands Cancer Registry and Dr. J.A.A.M. van Dijck (Comprehensive Cancer Center East, Nijmegen, the Netherlands) for kindly providing the data presented in Figure 1-2.

References

1. Muzzin LJ, Anderson NJ, Figueredo AT, Gudelis SO. The experience of cancer. *Soc Sci Med.* 1994;38(9):1201-1208.
2. Stewart BW, Kleihues P. *World cancer report.* Lyon: IARC Press; Oxford: Oxford University Press [distributor]; 2003.
3. Parkin DM, Bray F, Ferlay J, Pisani P. Global cancer statistics, 2002. *CA Cancer J Clin.* 2005;55(2):74-108.
4. Danaei G, Vander Hoorn S, Lopez AD, Murray CJ, Ezzati M. Causes of cancer in the world: comparative risk assessment of nine behavioural and environmental risk factors. *Lancet.* 2005;366(9499):1784-1793.
5. Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell.* 2000;100(1):57-70.
6. Sancho E, Batlle E, Clevers H. Signaling pathways in intestinal development and cancer. *Annu Rev Cell Dev Biol.* 2004;20:695-723.
7. Watson JD, Crick FH. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature.* 1953;171(4356):737-738.
8. Shabalina SA, Spiridonov NA. The mammalian transcriptome and the function of non-coding DNA sequences. *Genome Biol.* 2004;5(4):105.
9. Jones PA. DNA methylation errors and cancer. *Cancer Res.* 1996;56(11):2463-2467.
10. Vogelstein B, Kinzler KW. Cancer genes and the pathways they control. *Nat Med.* 2004;10(8):789-799.
11. Peltomäki P. Deficient DNA mismatch repair: a common etiologic factor for colon cancer. *Hum Mol Genet.* 2001;10(7):735-740.
12. Zevin S, Benowitz NL. Drug interactions with tobacco smoking. An update. *Clin Pharmacokinet.* 1999;36(6): 425-438.
13. Hoensch HP, Hartmann F. The intestinal enzymatic biotransformation system: potential role in protection from colon cancer. *Hepatogastroenterology.* 1981;28(4):221-228.
14. Hayes JD, Pulford DJ. The glutathione *S*-transferase supergene family: regulation of GST and the contribution of the isoenzymes to cancer chemoprotection and drug resistance. *Crit Rev Biochem Mol Biol.* 1995;30(6):445-600.
15. Husgafvel-Pursiainen K. Genotoxicity of environmental tobacco smoke: a review. *Mutat Res.* 2004;567(2-3): 427-445.
16. Casarett LJ, Klaassen CD, Amdur MO, Doull J. *Casarett and Doull's toxicology: the basic science of poisons.* 5th ed. New York: McGraw-Hill Health Professions Division; 1996.
17. World Cancer Research Fund, American Institute for Cancer Research. *Food, nutrition and the prevention of cancer: a global perspective.* Washington, D.C.: American Institute for Cancer Research; 1997.
18. Pitot HC, Hikita H, Dragan Y, Sargent L, Haas M. Review article: the stages of gastrointestinal carcinogenesis--application of rodent models to human disease. *Aliment Pharmacol Ther.* 2000;14 Suppl 1:153-160.
19. Peto J. Early- and late-stage carcinogenesis in mouse skin and in man. In: Börzsönyi M, Day NE, Lapis K, Yamasaki H, eds. *Models, mechanisms and etiology of tumour promotion: proceedings of a symposium organized by the Hungarian Cancer Society and the IARC, held in Budapest, 16-18 May 1983.* Vol 56. Lyon: IARC Sci Publ.; 1984:359-371.
20. Ferlay J, Bray F, Pisani P, Parkin D. GLOBOCAN 2002: Cancer Incidence, Mortality and Prevalence Worldwide. *IARC CancerBase No. 5. version 2.0* [Accessed January, 10, 2005.
21. Damhuis RAM, Schouten LJ, Visser O, eds. *Gastrointestinal cancer in the Netherlands 1989-1992.* Utrecht, The Netherlands: Association of Comprehensive Cancer Centres; 1994.
22. Ries LAG, Eisner MP, Kosary CL, et al., eds. *SEER Cancer Statistics Review, 1975-2002.* Bethesda, MD: National Cancer Institute. http://seer.cancer.gov/csr/1975_2002/ based on November 2004 SEER data submission, posted to the SEER web site; 2005.
23. Rabeneck L, Davila JA, El-Serag HB. Is there a true "shift" to the right colon in the incidence of colorectal cancer? *Am J Gastroenterol.* 2003;98(6):1400-1409.
24. White JS, McCallion K, Gardiner KR, et al. Changing patterns of colorectal cancer. *Am J Gastroenterol.* 2004;99(4): 766.
25. Takayama T, Katsuki S, Takahashi Y, et al. Aberrant crypt foci of the colon as precursors of adenoma and cancer. *N Engl J Med.* 1998;339(18):1277-1284.
26. Peipins LA, Sandler RS. Epidemiology of colorectal adenomas. *Epidemiol Rev.* 1994;16(2):273-297.
27. Bond JH. Polyp guideline: diagnosis, treatment, and surveillance for patients with colorectal polyps. Practice Parameters Committee of the American College of Gastroenterology. *Am J Gastroenterol.* 2000;95(11):3053-3063.

28. Lieberman DA, Weiss DG, Bond JH, Ahnen DJ, Garewal H, Chejfec G. Use of colonoscopy to screen asymptomatic adults for colorectal cancer. Veterans Affairs Cooperative Study Group 380. *N Engl J Med.* 2000;343(3):162-168.

29. O'Brien MJ, Winawer SJ, Zauber AG, et al. The National Polyp Study. Patient and polyp characteristics associated with high-grade dysplasia in colorectal adenomas. *Gastroenterology.* 1990;98(2):371-379.

30. Loeve F, Van Ballegooijen M, Boer R, Kuipers EJ, Habbema JD. Colorectal cancer risk in adenoma patients: a nation-wide study. *Int J Cancer.* 2004;111(1):147-151.

31. Dukes C. The classification of cancer of the rectum. *J Pathol Bacteriol.* 1932;35:323-332.

32. Gabriel WB, Dukes C, Bussey HJR. Lymphatic spread in cancer of the rectum. 1935;23:395-413.

33. Chen CD, Yen MF, Wang WM, Wong JM, Chen TH. A case-cohort study for the disease natural history of adenoma-carcinoma and de novo carcinoma and surveillance of colon and rectum after polypectomy: implication for efficacy of colonoscopy. *Br J Cancer.* 2003;88(12):1866-1873.

34. Koretz RL. Malignant polyps: are they sheep in wolves' clothing? *Ann Intern Med.* 1993;118(1):63-68.

35. Lockhart-Mummery H, C. D. The pre-cancerous changes in the rectum and colon. *Surg Gynecol Obst.* 1928;46(5).

36. Rider JA, Kirsner JB, Moeller HC, Palmer WL. Polyps of the colon and rectum; a four-year to nine-year follow-up study of five hundred thirty-seven patients. *J Am Med Assoc.* 1959;170(6):633-638.

37. Vamosi-Nagy I, Koves I. Correlation between colon adenoma and cancer. *Eur J Surg Oncol.* 1993;19(6):619-624.

38. Ponz de Leon M, Antonioli A, Ascari A, Zanghieri G, Sacchetti C. Incidence and familial occurrence of colorectal cancer and polyps in a health-care district of northern Italy. *Cancer.* 1987;60(11):2848-2859.

39. Neugut AI, Jacobson JS, De Vivo I. Epidemiology of colorectal adenomatous polyps. *Cancer Epidemiol Biomarkers Prev.* 1993;2(2):159-176.

40. Ponz de Leon M, Sassatelli R, Sacchetti C, Zanghieri G, Scalmati A, Roncucci L. Familial aggregation of tumors in the three-year experience of a population-based colorectal cancer registry. *Cancer Res.* 1989;49(15):4344-4348.

41. De la Chapelle A. Genetic predisposition to colorectal cancer. *Nat Rev Cancer.* 2004;4(10):769-780.

42. Fearon ER, Vogelstein B. A genetic model for colorectal tumorigenesis. *Cell.* 1990;61(5):759-767.

43. Lin YM, Furukawa Y, Tsunoda T, Yue CT, Yang KC, Nakamura Y. Molecular diagnosis of colorectal tumors by expression profiles of 50 genes expressed differentially in adenomas and carcinomas. *Oncogene.* 2002;21(26): 4120-4128.

44. Notterman DA, Alon U, Sierk AJ, Levine AJ. Transcriptional gene expression profiles of colorectal adenoma, adenocarcinoma, and normal tissue examined by oligonucleotide arrays. *Cancer Res.* 2001;61(7):3124-3130.

45. Simons BD, Morrison AS, Lev R, Verhoek-Oftedahl W. Relationship of polyps to cancer of the large intestine. *J Natl Cancer Inst.* 1992;84(12):962-966.

46. Levi F, Randimbison L, La Vecchia C, Te VC, Franceschi S. Cancer risk following polyps or cancer of the large bowel in Vaud, Switzerland. *Int J Cancer.* 1999;80(4):634-635.

47. Lotfi AM, Spencer RJ, Ilstrup DM, Melton LJ, 3rd. Colorectal polyps and the risk of subsequent carcinoma. *Mayo Clin Proc.* 1986;61(5):337-343.

48. Stryker SJ, Wolff BG, Culp CE, Libbe SD, Ilstrup DM, MacCarty RL. Natural history of untreated colonic polyps. *Gastroenterology.* 1987;93(5):1009-1013.

49. Loeve F, van Ballegooijen M, Snel P, Habbema JD. Colorectal cancer risk after colonoscopic polypectomy: a population-based study and literature search. *Eur J Cancer.* 2005;41(3):416-422.

50. Winawer SJ, Zauber AG, Ho MN, et al. Prevention of colorectal cancer by colonoscopic polypectomy. The National Polyp Study Workgroup. *N Engl J Med.* 1993;329(27):1977-1981.

51. Thiis-Evensen E, Hoff GS, Sauar J, Langmark F, Majak BM, Vatn MH. Population-based surveillance by colonoscopy: effect on the incidence of colorectal cancer. Telemark Polyp Study I. *Scand J Gastroenterol.* 1999;34(4): 414-420.

52. Muller AD, Sonnenberg A. Prevention of colorectal cancer by flexible endoscopy and polypectomy. A case-control study of 32,702 veterans. *Ann Intern Med.* 1995;123(12):904-910.

53. Citarda F, Tomaselli G, Capocaccia R, Barcherini S, Crespi M. Efficacy in standard clinical practice of colonoscopic polypectomy in reducing colorectal cancer incidence. *Gut.* 2001;48(6):812-815.

54. Gilbertsen VA, Nelms JM. The prevention of invasive cancer of the rectum. *Cancer.* 1978;41(3):1137-1139.

55. Loeve F, Boer R, Zauber AG, et al. National Polyp Study data: evidence for regression of adenomas. *Int J Cancer.* 2004;111(4):633-639.

56. Selby JV, Friedman GD, Quesenberry CP, Jr., Weiss NS. A case-control study of screening sigmoidoscopy and mortality from colorectal cancer. *N Engl J Med.* 1992;326(10):653-657.

57. Atkin WS, Morson BC, Cuzick J. Long-term risk of colorectal cancer after excision of rectosigmoid adenomas. *N Engl J Med.* 1992;326(10):658-662.

58. Rembacken BJ, Fujii T, Cairns A, et al. Flat and depressed colonic neoplasms: a prospective study of 1000 colonoscopies in the UK. *Lancet.* 2000;355(9211):1211-1214.

59. Hamilton SR. Origin of colorectal cancers in hyperplastic polyps and serrated adenomas: another truism bites the dust. *J Natl Cancer Inst.* 2001;93(17):1282-1283.

60. Hawkins NJ, Ward RL. Sporadic colorectal cancers with microsatellite instability and their possible origin in hyperplastic polyps and serrated adenomas. *J Natl Cancer Inst.* 2001;93(17):1307-1313.

61. Jørgensen OD, Kronborg O, Fenger C. The Funen Adenoma Follow-up Study. Incidence and death from colorectal carcinoma in an adenoma surveillance program. *Scand J Gastroenterol.* 1993;28(10):869-874.

62. O'Connell JB, Maggard MA, Ko CY. Colon cancer survival rates with the new American Joint Committee on Cancer sixth edition staging. *J Natl Cancer Inst.* 2004;96(19):1420-1425.

63. UK Colorectal Cancer Screening Pilot Group. Results of the first round of a demonstration pilot of screening for colorectal cancer in the United Kingdom. *BMJ.* 2004;329(7458):133.

64. De Visser M, Van Ballegooijen M, Bloemers SM, et al. Report on the Dutch consensus development meeting for implementation and further development of population screening for colorectal cancer based on FOBT. *Cell Oncol.* 2005;27(1):17-29.

65. Syngal S, Bandipalliam P, Boland CR. Surveillance of patients at high risk for colorectal cancer. *Med Clin North Am.* 2005;89(1):61-84, vii-viii.

66. Boland CR. Genetic pathways to colorectal cancer. *Hosp Pract (Off Ed).* 1997;32(11):79-84, 87-96.

67. Iacopetta B. Are there two sides to colorectal cancer? *Int J Cancer.* 2002;101(5):403-408.

68. Fodde R, Kuipers J, Rosenberg C, et al. Mutations in the APC tumour suppressor gene cause chromosomal instability. *Nat Cell Biol.* 2001;3(4):433-438.

69. Morin PJ. beta-catenin signaling and cancer. *BioEssays.* 1999;21(12):1021-1030.

70. Eden A, Gaudet F, Waghmare A, Jaenisch R. Chromosomal instability and tumors promoted by DNA hypomethylation. *Science.* 2003;300(5618):455.

71. Miyaki M, Kuroki T. Role of Smad4 (DPC4) inactivation in human cancer. *Biochem Biophys Res Commun.* 2003;306(4): 799-804.

72. Houlston R, Bevan S, Williams A, et al. Mutations in DPC4 (SMAD4) cause juvenile polyposis syndrome, but only account for a minority of cases. *Hum Mol Genet.* 1998;7(12):1907-1912.

73. Levine AJ, Finlay CA, Hinds PW. P53 is a tumor suppressor gene. *Cell.* 2004;116(2 Suppl):S67-69, 61 p following S69.

74. Atkin NB. Microsatellite instability. *Cytogenet Cell Genet.* 2001;92(3-4):177-181.

75. Herman JG, Umar A, Polyak K, et al. Incidence and functional consequences of hMLH1 promoter hypermethylation in colorectal carcinoma. *Proc Natl Acad Sci U S A.* 1998;95(12):6870-6875.

76. Wheeler JM, Loukola A, Aaltonen LA, Mortensen NJ, Bodmer WF. The role of hypermethylation of the hMLH1 promoter region in HNPCC versus MSI+ sporadic colorectal cancers. *J Med Genet.* 2000;37(8):588-592.

77. Kim HC, Kim CN, Yu CS, Roh SA, Kim JC. Methylation of the hMLH1 and hMSH2 promoter in early-onset sporadic colorectal carcinomas with microsatellite instability. *Int J Colorectal Dis.* 2003;18(3):196-202.

78. Kuismanen SA, Holmberg MT, Salovaara R, De la Chapelle A, Peltomaki P. Genetic and epigenetic modification of MLH1 accounts for a major share of microsatellite-unstable colorectal cancers. *Am J Pathol.* 2000;156(5): 1773-1779.

79. Kamory E, Kolacsek O, Otto S, Csuka O. hMLH1 and hMSH2 somatic inactivation mechanisms in sporadic colorectal cancer patients. *Pathol Oncol Res.* 2003;9(4):236-241.

80. Cunningham JM, Christensen ER, Tester DJ, et al. Hypermethylation of the hMLH1 promoter in colon cancer with microsatellite instability. *Cancer Res.* 1998;58(15):3455-3460.

81. Cunningham JM, Kim CY, Christensen ER, et al. The frequency of hereditary defective mismatch repair in a prospective series of unselected colorectal carcinomas. *Am J Hum Genet.* 2001;69(4):780-790.

82. Aarnio M, Sankila R, Pukkala E, et al. Cancer risk in mutation carriers of DNA-mismatch-repair genes. *Int J Cancer.* 1999;81(2):214-218.

83. 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.

84. Butterworth AS, Higgins JP, Pharoah P. Relative and absolute risk of colorectal cancer for individuals with a family history: a meta-analysis. *Eur J Cancer.* 2006;42(2):216-227.

85. Fuchs CS, Giovannucci EL, Colditz GA, Hunter DJ, Speizer FE, Willett WC. A prospective study of family history and the risk of colorectal cancer. *N Engl J Med.* 1994;331(25):1669-1674.

86. Johns LE, Houlston RS. A systematic review and meta-analysis of familial colorectal cancer risk. *Am J Gastroenterol.* 2001;96(10):2992-3003.

87. Houlston RS, Peto J. The search for low-penetrance cancer susceptibility alleles. *Oncogene.* 2004;23(38):6471-6476.

88. Khoury MJ, Beaty TH, Cohen BH. *Fundamentals of genetic epidemiology.* New York: Oxford University Press; 1993.

89. Parkin DM, Khlat M. Studies of cancer in migrants: rationale and methodology. *Eur J Cancer.* 1996;32A(5):761-771.

90. Doll R, Peto R. The Causes of Cancer - Quantitative Estimates of Avoidable Risks of Cancer in the United-States Today. *Journal of the National Cancer Institute.* 1981;66(6):1191-1308.

91. Willett WC. Diet, nutrition, and avoidable cancer. *Environ Health Perspect.* 1995;103 Suppl 8:165-170.

92. Giovannucci E. Diet, body weight, and colorectal cancer: a summary of the epidemiologic evidence. *J Womens Health (Larchmt).* 2003;12(2):173-182.

93. Giovannucci E. Modifiable risk factors for colon cancer. *Gastroenterol Clin North Am.* 2002;31(4):925-943.

94. Key TJ, Schatzkin A, Willett WC, Allen NE, Spencer EA, Travis RC. Diet, nutrition and the prevention of cancer. *Public Health Nutr.* 2004;7(1A):187-200.

95. Norat T, Lukanova A, Ferrari P, Riboli E. Meat consumption and colorectal cancer risk: dose-response meta-analysis of epidemiological studies. *Int J Cancer.* 2002;98(2):241-256.

96. Sandhu MS, White IR, McPherson K. Systematic review of the prospective cohort studies on meat consumption and colorectal cancer risk: a meta-analytical approach. *Cancer Epidemiol Biomarkers Prev.* 2001;10(5):439-446.

97. Norat T, Bingham S, Ferrari P, et al. Meat, fish, and colorectal cancer risk: the European Prospective Investigation into cancer and nutrition. *J Natl Cancer Inst.* 2005;97(12):906-916.

98. Cross AJ, Sinha R. Meat-related mutagens/carcinogens in the etiology of colorectal cancer. *Environ Mol Mutagen.* 2004;44(1):44-55.

99. Lee DH, Anderson KE, Harnack LJ, Folsom AR, Jacobs DR, Jr. Heme iron, zinc, alcohol consumption, and colon cancer: Iowa Women's Health Study. *J Natl Cancer Inst.* 2004;96(5):403-407.

100. Giovannucci E, Goldin B. The role of fat, fatty acids, and total energy intake in the etiology of human colon cancer. *Am J Clin Nutr.* 1997;66(6 Suppl):1564S-1571S.

101. Zock PL. Dietary fats and cancer. *Curr Opin Lipidol.* 2001;12(1):5-10.

102. Woutersen RA, Appel MJ, van Garderen-Hoetmer A, Wijnands MV. Dietary fat and carcinogenesis. *Mutat Res.* 1999;443(1-2):111-127.

103. Nagengast FM, Grubben MJAL, Van Munster IP. Role of bile acids in colorectal carcinogenesis. *Eur J Cancer.* 1995;31A(7-8):1067-1070.

104. Roynette CE, Calder PC, Dupertuis YM, Pichard C. n-3 polyunsaturated fatty acids and colon cancer prevention. *Clin Nutr.* 2004;23(2):139-151.

105. Cho E, Smith-Warner SA, Spiegelman D, et al. Dairy foods, calcium, and colorectal cancer: a pooled analysis of 10 cohort studies. *J Natl Cancer Inst.* 2004;96(13):1015-1022.

106. Shaukat A, Scouras N, Schunemann HJ. Role of supplemental calcium in the recurrence of colorectal adenomas: a metaanalysis of randomized controlled trials. *Am J Gastroenterol.* 2005;100(2):390-394.

107. Martinez ME, Willett WC. Calcium, vitamin D, and colorectal cancer: a review of the epidemiologic evidence. *Cancer Epidemiol Biomarkers Prev.* 1998;7(2):163-168.

108. Newmark HL, Wargovich MJ, Bruce WR. Colon cancer and dietary fat, phosphate, and calcium: a hypothesis. *J Natl Cancer Inst.* 1984;72(6):1323-1325.

109. Trock B, Lanza E, Greenwald P. Dietary fiber, vegetables, and colon cancer: critical review and meta-analyses of the epidemiologic evidence. *J Natl Cancer Inst.* 1990;82(8):650-661.

110. Steinmetz KA, Potter JD. Vegetables, fruit, and cancer prevention: a review. *J Am Diet Assoc.* 1996;96(10): 1027-1039.

111. Flood A, Schatzkin A. Colorectal cancer: does it matter if you eat your fruits and vegetables? *J Natl Cancer Inst.* 2000;92(21):1706-1707.

112. Ferguson LR. Does a diet rich in dietary fibre really reduce the risk of colon cancer? *Dig Liver Dis.* 2005;37(3): 139-141.

113. Bingham SA, Day NE, Luben R, et al. Dietary fibre in food and protection against colorectal cancer in the European Prospective Investigation into Cancer and Nutrition (EPIC): an observational study. *Lancet.* 2003;361(9368): 1496-1501.

114. Sanjoaquin MA, Allen N, Couto E, Roddam AW, Key TJ. Folate intake and colorectal cancer risk: a meta-analytical approach. *Int J Cancer.* 2005;113(5):825-828.

115. Giovannucci E. Epidemiologic studies of folate and colorectal neoplasia: a review. *J Nutr.* 2002;132(8 Suppl): 2350S-2355S.

116. Kim YI. Folate and cancer prevention: a new medical application of folate beyond hyperhomocysteinemia and neural tube defects. *Nutr Rev.* 1999;57(10):314-321.

117. Mason JB. Biomarkers of nutrient exposure and status in one-carbon (methyl) metabolism. *J Nutr.* 2003;133 Suppl 3: 941S-947S.

118. Van den Donk M, Buijsse B, Van den Berg SW, et al. Dietary Intake of Folate and Riboflavin, MTHFR C677T Genotype, and Colorectal Adenoma Risk: A Dutch Case-Control Study. *Cancer Epidemiol Biomarkers Prev.* 2005;14(6): 1562-1566.

119. Wei EK, Giovannucci E, Selhub J, Fuchs CS, Hankinson SE, Ma J. Plasma vitamin B6 and the risk of colorectal cancer and adenoma in women. *J Natl Cancer Inst.* 2005;97(9):684-692.

120. Giovannucci E. An updated review of the epidemiological evidence that cigarette smoking increases risk of colorectal cancer. *Cancer Epidemiol Biomarkers Prev.* 2001;10(7):725-731.

121. Colditz GA, Yaus KP. Smoking causes colon carcinoma. *Cancer.* 2004;100(2):223-224.

122. Cho E, Smith-Warner SA, Ritz J, et al. Alcohol intake and colorectal cancer: a pooled analysis of 8 cohort studies. *Ann Intern Med.* 2004;140(8):603-613.

123. Giovannucci E. Alcohol, one-carbon metabolism, and colorectal cancer: recent insights from molecular studies. *J Nutr.* 2004;134(9):2475S-2481S.

124. Kune GA, Vitetta L. Alcohol consumption and the etiology of colorectal cancer: a review of the scientific evidence from 1957 to 1991. *Nutr.Cancer.* 1992;18:97-111.

125. Riboli E, Cornee J, Macquart-Moulin G, Kaaks R, Casagrande C, Guyader M. Cancer and polyps of the colorectum and lifetime consumption of beer and other alcoholic beverages. *Am J Epidemiol.* 1991;134(2):157-166.

126. Hillman RS, Steinberg SE. The effects of alcohol on folate metabolism. *Annu Rev Med.* 1982;33:345-354.

127. Chan AT. Aspirin, non-steroidal anti-inflammatory drugs and colorectal neoplasia: future challenges in chemoprevention. *Cancer Causes Control.* 2003;14(5):413-418.

128. Stoehlmacher J, Lenz HJ. Cyclooxygenase-2 inhibitors in colorectal cancer. *Semin Oncol.* 2003;30(3 Suppl 6):10-16.

129. Sanborn R, Blanke CD. Cyclooxygenase-2 inhibition in colorectal cancer: boom or bust? *Semin Oncol.* 2005;32(1): 69-75.

130. Samad AK, Taylor RS, Marshall T, Chapman MA. A meta-analysis of the association of physical activity with reduced risk of colorectal cancer. *Colorectal Dis.* 2005;7(3):204-213.

131. Brockton N, Little J, Sharp L, Cotton SC. N-acetyltransferase polymorphisms and colorectal cancer: a HuGE review. *Am J Epidemiol.* 2000;151(9):846-861.

132. Dos Santos Silva I, International Agency for Research on Cancer. *Cancer epidemiology: principles and methods.* Lyon, France: International Agency for Research on Cancer; 1999.

133. Breslow NE, Day NE. Statistical methods in cancer research. Volume I -The analysis of case-control studies. *IARC Sci Publ.* 1980(32):5-338.

134. Paneth N, Susser E, Susser M. Origins and early development of the case-control study: Part 1, Early evolution. *Soz Praventivmed.* 2002;47(5):282-288.

135. Paneth N, Susser E, Susser M. Origins and early development of the case-control study: Part 2, The case-control study from Lane-Claypon to 1950. *Soz Praventivmed.* 2002;47(6):359-365.

136. Breslow NE, Day NE. Statistical methods in cancer research. Volume II-The design and analysis of cohort studies. *IARC Sci Publ.* 1987(82):1-406.

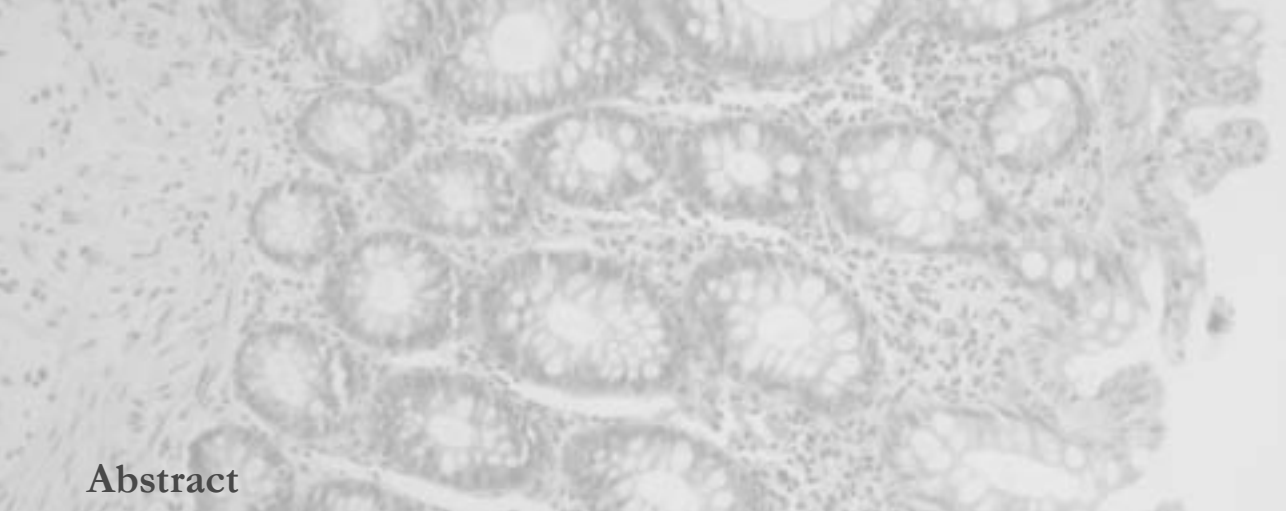
137. Doll R. Cohort studies: history of the method. I. Prospective cohort studies. *Soz Praventivmed.* 2001;46(2):75-86.

138. Doll R. Cohort studies: history of the method. II. Retrospective cohort studies. *Soz Praventivmed.* 2001;46(3): 152-160.

139. Perera FP, Weinstein IB. Molecular epidemiology: recent advances and future directions. *Carcinogenesis.* 2000;21(3): 517-524.

140. Hainaut P, Olivier M, Pfeifer GP. TP53 mutation spectrum in lung cancers and mutagenic signature of components of tobacco smoke: lessons from the IARC TP53 mutation database. *Mutagenesis.* 2001;16(6):551-553; author reply 555-556.

Index



Abstract

Background: The glutathione (GSH)/glutathione *S*-transferase (GST) system is an important detoxification system in the gastrointestinal tract. A high activity of this system may benefit cancer prevention.

Aim: To assess whether habitual consumption of fruits and vegetables, especially citrus fruits and brassica and allium vegetables, is positively associated with parameters reflecting the activity of the GSH/GST enzyme system in human rectal mucosa. GST enzyme activity, GST isoenzyme levels of GST-alpha (A1-1, A1-2 and A2-2), -mu (M1-1) and -pi (P1-1), and GSH levels were measured in rectal biopsies from 94 subjects. Diet, lifestyle, *GSTM1* and *GSTT1* null polymorphisms were assessed.

Results: Mean GST enzyme activity was 237 nmol/min/mg protein (SD=79). Consumption of citrus fruits was positively associated with GST enzyme activity (difference between high and low consumption: 28.9, 95% confidence interval (CI)= 9.3-48.6 nmol/min/mg protein), but was not associated with the other parameters. A positive association with brassica vegetables was found among carriers of the *GSTM1*-plus genotype (difference between high and low consumption: 22.6, 95% CI=0.2-45.0 nmol/min/mg protein), but not among *GSTM1*-null individuals (-25.8 nmol/min/mg protein, 95% CI=-63.3-11.8). This is in line with a positive association between consumption of brassica vegetables and GSTM1-1 isoenzyme level (difference between high and low consumption: 67.5%, 95% CI=6.8-162.7). Consumption of allium vegetables was not associated with GST enzyme activity, but negatively with GSTP1-1 levels (difference between high and low consumption: -23.3%, 95% CI=-35.5–-8.6). Associations were similar among those with the *GSTT1*-plus and *GSTT1*-null genotype.

Conclusion: Variations in habitual consumption of fruits, particularly citrus fruits, and of vegetables, in particular brassica vegetables, among those with the *GSTM1*-plus genotype, may contribute to variations in human rectal GST enzyme activity.

Habitual Consumption of Fruits and Vegetables: Associations with Human Rectal Glutathione *S*-transferase

2

Petra A. Wark
Marina J.A.L. Grubben
Wilbert H.M. Peters
Fokko M. Nagengast
Ellen Kampman
Frans J. Kok
Pieter van 't Veer

Carcinogenesis. 2004;25(11):2135-42

Introduction

Given their role in absorption, digestion and transport, the colon and rectum are constantly challenged by potentially harmful compounds, including mutagens and carcinogens. The large intestine possesses several defense mechanisms to counteract damage of the colorectal mucosa by such reactive compounds. These include the ability to up-regulate detoxification systems[1, 2].

Essential is the glutathione (GSH)/glutathione *S*-transferase (GST) detoxification system, which comprises antioxidant reduced GSH and GSTs (EC 2.5.1.18); a family of phase II enzymes, which, in humans, consists of four main subgroups: alpha(α), mu(μ), pi(π) and theta(θ).

Because this system plays an important role in detoxification of a broad range of carcinogens[3], high GST enzyme activity has been suggested as being beneficial to cancer prevention[4]. Considering the low GST enzyme activity in the colon and rectum compared with tissues in which cancer occurs less frequently, we hypothesized previously that GST enzyme activity might be critically low and related to high rates of carcinogenesis in this organ[5, 6].

Colonic GST enzyme activity and GST protein levels vary considerably between individuals[7, 8], which may be related to differential susceptibility to colorectal cancer. Individuals with homozygous deletions of *GSTM1* or *GSTT1* (null genotype) do not have detectable GSTM1 or GSTT1 enzyme activity, respectively, and were postulated to be at higher risk of colorectal cancer. However, no consistent associations of *GSTM1* and *GSTT1* null polymorphisms with risk of colorectal cancer have been observed[9]. Apart from inherited polymorphisms in GSTs, individuals may differ in GST enzyme activity due to differential exposure to bioactive compounds.

In vivo and *in vitro* studies have shown that a variety of dietary compounds or their metabolites can induce the GSH/GST detoxification system. These include glucosinolate metabolites and dithiolthiones present in brassica vegetables, diallyl sulfides present in allium vegetables, limonoids and flavonoids present in citrus fruits[3, 10-17], and butyrate produced by colonic fermentation of fiber[8].

Evidence for induction of the human rectal GSH/GST detoxification system was found in a crossover study among ten volunteers consuming 300 g/d of cooked Brussels sprouts during seven days. This yielded 30 and 15% increases of GST-alpha and GSTP1-1 protein levels, respectively. However, no effect upon GST enzyme activity was found[11], and taking 3 g/d of broccoli supplements for 14 days also did not influence GST enzyme activity in lymphocytes or colon mucosa[10].

Nonetheless, indications of the up-regulation of GST enzyme activity by brassica and allium vegetables were found in blood plasma, urine and saliva[11-15]. Moreover, duodenal GST-α and GST-π protein levels were higher among subjects consuming vegetables at least four times a week, and antral GSTT1 protein levels were higher among subjects consuming fruits at least four times a week compared with those who consumed these products less frequently. However, no associations between frequency of consumption of fruits and vegetables and GST enzyme activity were found in these tissues[16].

Induction of the GSH/GST detoxification system by fruits and vegetables may partially account for the observed inverse associations between their consumption and risk of colorectal

cancer[12, 18]. We investigated whether habitual consumption of fruits and vegetables, in particular of brassica and allium vegetables and citrus fruits, is positively associated with the following components of the rectal GSH/GST detoxification system: GST enzyme activity, GST-α, GSTM1-1 and GSTP1-1 isoenzyme and GSH levels.

Material and Methods

This cross-sectional study comprises a sub-study nested in a case-control study on dietary factors, genetic susceptibility and somatic mutations in sporadic and hereditary colorectal adenomas[19]. In addition to the data collected in the main study, rectal biopsies were taken from participants who enrolled in one of the participating hospitals.

Study Population

Between December 1995 and February 1998, subjects undergoing colonoscopy or sigmoidoscopy in the outpatient clinic of the Department of Gastroenterology and Hepatology of University Medical Centre Nijmegen, The Netherlands, were recruited by their gastroenterologist. Eligible subjects were Dutch speaking Caucasians, between 18 and 75 years old at day of endoscopy and having no history of colorectal resection, polyposis coli, colorectal cancer or chronic inflammatory bowel disease. Subjects with and without adenomas, either with or without a family background of hereditary non-polyposis colorectal cancer (HNPCC) according to the Amsterdam I criteria[20] were included. All HNPCC family members were first-degree relatives of patients with colorectal or endometrial cancer, but were free of cancer themselves. Subjects who did not belong to a HNPCC family were excluded if diagnosed with colorectal adenoma >3 years prior to recruitment. The medical ethical committees of Wageningen University and University Medical Centre Nijmegen approved the study protocol.

Rectal biopsies were taken from 106 (82%) of the 130 eligible subjects. Twelve subjects did not return dietary and/or lifestyle questionnaires. Thus, the final study population consists of 94 subjects, including 44 members of 26 HNPCC families.

Data Collection

After providing written informed consent, participants underwent endoscopy during which six biopsies from healthy mucosa were taken from rectal mucosa within 10 cm from the anal verge. Biopsy specimens were immediately frozen in liquid nitrogen and stored at -80°C. Additionally, 30 ml of EDTA blood was drawn and stored at -20°C.

Clinical information regarding the presence and characteristics of adenomas, HNPCC and indication for endoscopy was abstracted from medical records.

We requested that participants complete a questionnaire on lifestyle and socio-economic factors and a validated semi-quantitative food-frequency questionnaire[21]. The questionnaire assesses consumption of fruits and vegetables with high reproducibility after a year (fruits, Spearman's r=0.61 and 0.77; vegetables, r=0.76 and 0.65 for males and females, respectively), while relative validity versus the means of 12 24-h recalls was moderate (fruits, Spearman's r=0.68 and 0.56;

vegetables, r=0.38 and 0.31 for males and females, respectively)[21]. Nutrient intake was calculated using the 1996 computerized version of the Dutch food composition table[22].

Laboratory Assays

Rectal biopsies were processed into cytosolic fractions[11]. Total GST enzyme activity was assayed by spectrophotometric determination of 1-chloro-2,4-dinitrobenzene (CDNB) conjugation with GSH[23]. Total intracellular reduced glutathione (GSH) level was quantified by high performance liquid chromatography after reaction with monobromobimane using a modification of the method of Fahey and Newton[24,25]. GST-α (GSTA1-1, A1-2 and A2-2), GSTM1-1 and GSTP1-1 isoenzyme levels were determined by western blotting[11]. GST enzyme activity and levels of isoenzymes and GSH were expressed per milligram of intracellular protein, as determined colorimetrically according to the method of Lowry *et al.*[26]. Protein and enzyme activity measurements were done in duplicate. The within- and between-assay coefficients of variations were 10-15% for protein and 5-10% for GST enzyme activity measurements, respectively. As a result of limited tissue availability, not all biopsy samples could be analyzed completely; this resulted in 78 samples analyzed for GST enzyme activity, 94 for GSH, 91 for GST-α and GSTP1-1 and 76 for GSTM1-1 isoenzyme levels.

DNA was extracted from 200 µl frozen whole blood (QIAamp blood kit, Qiagen Inc.), diluted to a concentration of ~20 ng/µl and stored at 4°C until use. *GSTM1* and *GSTT1* genotypes were determined using a general multiplex polymerase chain reaction method followed by electrophoresis[27]. A ß-globin gene fragment of 350 bp was present as a positive control.

Data Analysis

Because all biopsies in which GSTM1-1 protein was not detected were obtained from *GSTM1*-null genotype carriers, we excluded these observations in analyses regarding this isoenzyme. We log-transformed GST-α, GSTM1-1 and GSTP1-1 protein levels to yield approximately normally distributed variables. Nutrient, alcohol and fat intakes were adjusted, separately for men and women, for total energy intake using the residual method[28].

Spearman correlation coefficients were computed to quantify the strength of the association between GST isoenzymes, GSH and GST enzyme activity. Linear mixed models were fitted to associate dietary, lifestyle and medical determinants with GST enzyme activity, GSH and log-transformed GSTM1-1 and GSTP1-1 levels. In these models, family membership was incorporated as random intercept to account for the presence of relatives in the study population, and HNPCC and presence of adenomas at index endoscopy were included as indicator variables to account for the study design. Because no null polymorphism has been identified in any of the genes encoding GST-α, we used Tobit regression[29], which explicitly allows for censored distributions of the dependent variable, to model this isoenzyme.

Two series of models were fitted: a first to investigate the association between lifestyle and medical factors, and a second to address the association between consumption of fruits and vegetables with the aforementioned GSH/GST parameters. In the latter, total energy intake was included to control for potential confounding and to reduce the impact of under- and

over-reporting of intake[28]. To compensate for potential confounding, the following factors that potentially relate to both consumption of fruits and vegetables and to GST enzyme activity were added separately and mutually to the model: current and past smoking, sex and use of NSAIDs or paracetamol as indicator variables, and as continuous variables: age (yr), body mass index (kg/m^2), consumption of coffee, tea, (red) meat, energy-adjusted (saturated) fat, alcohol, red wine in grams per day and vitamin E (equivalents/d). Current smoking was the only factor that changed the regression coefficient of interest with >5 U in most models. To enhance comparability, we adjusted all models for this factor. Additionally, regression equations regarding GST isoenzymes were adjusted for energy-adjusted fat.

Fruits and vegetables were treated as continuous variables, tertiles and quadratic terms. Because Akaike's Information Criteria and likelihood ratio tests - for tertiles and quadratic terms, respectively - showed a similar fit to continuous terms, only models containing the latter are presented.

To evaluate hypothesized effect modification by *GSTM1* and *GSTT1* genotype and explore effect modification by smoking, age, sex, HNPCC and adenoma status, interaction terms were included and tested by Wald tests. Subsequently, to evaluate whether the observed associations may be attributed to fractions related to dietary fiber, vitamin A, vitamin C, ß-carotene and folate, we added these food components to the models. Finally, we checked whether influential outliers were present. This appeared not to be the case.

All tests of statistical significance were two-sided and considered to be significant at the level of 5%. The analyses were conducted using Statistical Analysis Software (SAS version 8.0, SAS Institute, Cary, NC).

Results

In Table 2-1 the characteristics of the entire study population are presented. The characteristics of subjects in whose rectal biopsies the GST enzyme activity, one or more of the isoenzymes and/or GSH could not be assessed, did not differ from those subjects in which these factors were assessed (results not presented).

Figure 2-1 shows the distribution of GST enzyme activity, isoenzymes and GSH according to *GSTM1* genotype. As expected, GSTM1-1 protein could not be detected in any of the biopsies of *GSTM1*-null genotype carriers, whereas detectable levels were present in all *GSTM1*-plus genotype carriers. GST-α protein could not be detected in 51 (56.0%) of the 91 samples.

Among the 76 subjects in whom GST-α, GSTM1-1 and GSTP1-1 levels could be assessed, GSTP1-1 attained highest levels in 71 (93.4%) subjects, while GSTM1-1 levels were highest in the remaining five subjects; GST-α was only present in minor quantities. No differences in GST enzyme activity, GSH and GST isoenzyme levels could be detected according to *GSTM1* and *GSTT1* genotype or their combination (results not presented).

No correlation was detected between GST isoenzymes, GSH and GST enzyme activity. However, after adjusting for HNPCC, adenoma status, smoking behavior, sex, age, *GSTM1* and *GSTT1* genotype, a weak, negative correlation between GSH and GST enzyme activity appeared (r =-0.24, P=0.049).

Table 2-1. Characteristics of the study population*

Characteristic	Men (*n*=45)	Women (*n*=49)
Source of study population		
Members of HNPCC families†, *n* (%)	18 (40.0%)	26 (53.1%)
Non-HNPCC subjects‡, *n* (%)	27 (60.0%)	23 (46.9%)
Demographic and lifestyle factors		
Age (yr), mean ± SD	45.7 ± 13.3	47.9 ± 13.9
Body mass index (kg/m²), mean ± SD	25.8 ± 2.5	24.9 ± 3.6
Current smokers, *n* (%)	14 (31.1%)	16 (33.3%)
Intake of paracetamol (tablets/yr), median (p25, p75)	1.0 (0.5; 4.0)	2.6 (1.0; 6.0)
Intake of NSAIDs (tablets/yr), median (p25, p75)	1.3 (0.5; 2.0)	1.6 (1.0; 2.3)
Dietary factors		
Total energy intake (kJ/d), mean ± SD	10,515 ± 3,368	8,343 ±1,768
Protein (g/d), mean ± SD	90.6 ± 27.2	76.2 ± 14.5
Carbohydrates (g/d), mean ± SD	276.0 ± 94.1	221.1 ± 55.1
Fat (g/d), median (p25, p75)	99.7 (73.1; 113.5)	76.2 (65.5; 89.8)
Fruits (g/d), median (p25, p75)	149.3 (62.8; 253.0)	122.5 (84.7; 236.6)
Vegetables (g/d), median (p25, p75)	110.4 (87.0; 147.0)	118.6 (85.1; 157.6)
Fiber (g/d), median (p25, p75)	27.7 (22.1; 31.1)	21.7 (18.8; 26.9)
Meat (g/d), median (p25, p75)	129.4 (85.4; 154.0)	100.7 (60.7; 123.0)
Alcohol (glasses/wk), median (p25, p75)	7.6 (3.0; 15.7)	3.7 (0.9; 14.2)
Adenomas		
Adenomas at index endoscopy, *n* (%)	16 (35.6%)§	16 (32.7%)§
History of adenomas, *n* (%)	10 (22.2%)	10 (20.4%)
*Self-reported bowel complaints***		
Rectal bleeding, *n* (%)	9 (22.0%)	8 (17.8%)
Bowel release issues, *n* (%)	2 (4.9%)	13 (28.9%)
Diarrhea, *n* (%)	4 (9.8%)	7 (15.6%)
Cramps, *n* (%)	4 (9.8%)	11 (24.4%)
GST genotype		
GSTM1-plus and *GSTT1*-plus, *n* (%)	15 (33.3%)	20 (40.8%)
GSTM1-plus and *GSTT1*-null, *n* (%)	1 (2.2%)	8 (16.3%)
GSTM1-null and *GSTT1*-plus, *n* (%)	25 (55.6%)	21 (42.9%)
GSTM1-null and *GSTT1*-null, *n* (%)	4 (8.9%)	0 (0.0%)

* Because of missing values, not all numbers sum up to 45 and 49, respectively.

† All HNPCC family members were free of cancer themselves. Six (four male, two female) and three (one male, two female) of the HNPCC family members were known to carry a germline mutation in hMLH1 and hMSH2, respectively, whereas 24 (nine male, 15 female) were known to be no germline mutation carriers. One female subject classified as not having HNPCC had potentially late onset HNPCC, although her pedigree did not fulfil the Amsterdam criteria.

‡ Of the non-HNPCC subjects, six (23.1%) of the females and three (13.0%) of the males had a positive first-degree family history of colorectal cancer. For one subject this was unknown.

§ Seven male and four female subjects had rectal adenomas.

** In the year prior to recruitment.

Table 2-2 describes GST enzyme activity according to *GSTM1* and *GSTT1* genotype, lifestyle and medical factors. Smokers had a higher GST enzyme activity compared with never smokers; former smoking subjects tended to have a higher GST enzyme activity compared with those who had never smoked, although the difference was not significant. Strikingly, the presence of adenomas at index endoscopy was associated with a 44.9 nmol/min/mg protein lower GST enzyme activity. Restricting the analysis to subjects having rectal adenomas increased this

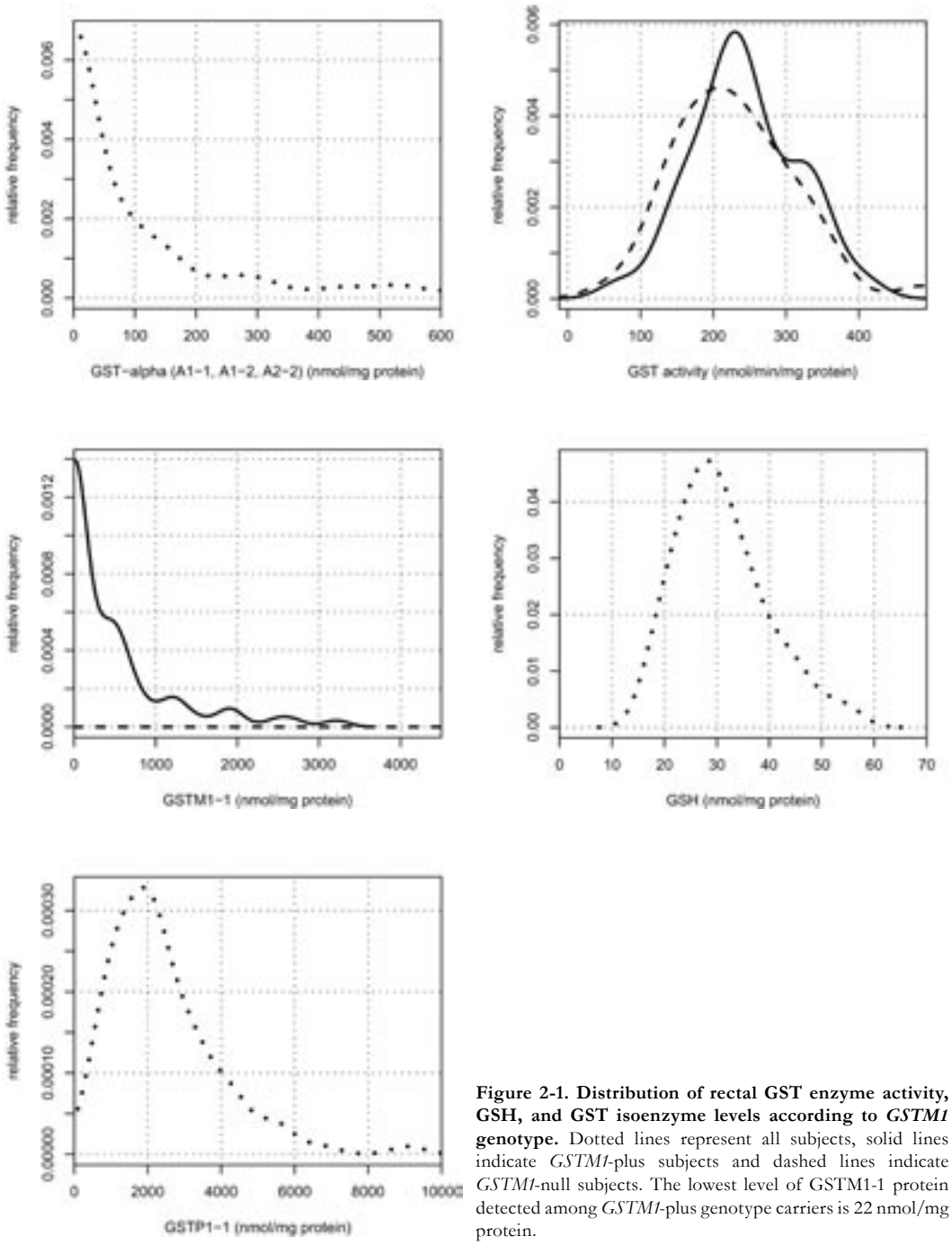


Figure 2-1. Distribution of rectal GST enzyme activity, GSH, and GST isoenzyme levels according to *GSTM1* genotype. Dotted lines represent all subjects, solid lines indicate *GSTM1*-plus subjects and dashed lines indicate *GSTM1*-null subjects. The lowest level of GSTM1-1 protein detected among *GSTM1*-plus genotype carriers is 22 nmol/mg protein.

Table 2-2. Description of rectal GST enzyme activity according to *GSTM1* and *GSTT1* genotype, lifestyle and medical factors

Characteristics	Subgroup	*n**	GST enzyme activity (nmol/min/mg protein) Mean ± SD	Mean difference in GST enzyme activity in nmol/min/mg protein (95% CI)	
				Adjusted for HNPCC and current adenoma	Adjusted for HNPCC, current adenoma and current smoking
Total group		77	237 ± 78	-	-
GSTM1 genotype	*GSTM1*-plus	38 (49.4%)	243 ± 73	9.8 (-28.6-48.2)	21.3 (-16.9-59.4)
	GSTM1-null	39 (50.6%)	231 ± 84	(reference)	(reference)
GSTT1 genotype	*GSTT1*-plus	67 (87.0%)	242 ± 79	30.9 (-30.0-91.9)	28.6 (-31.2-88.4)
	GSTT1-null	10 (13.0%)	202 ± 69	(reference)	(reference)
Gender	Men	36 (46.8%)	239 ± 93	3.6 (-34.6-41.8)	5.6 (-31.5-42.6)
	Women	41 (53.2%)	235 ± 63	(reference)	(reference)
Age	<50 years old	43 (55.8%)	238 ± 83	14.8 (-26.2-55.7)	21.3 (-18.4-61.1)
	≥50 years old	34 (44.2%)	235 ± 73	(reference)	(reference)
Smoking status	Current smokers†	25 (32.5%)	265 ± 61	55.5 (11.0-99.9)	-
	Ex-smokers†‡	25 (32.5%)	232 ± 83	20.6 (-23.1-64.3)	-
	Never smokers‡	27 (35.1%)	215 ± 83	(reference)	-
HNPCC family membership§	Yes	32 (41.6%)	242 ± 88	-0.1 (-45.0-44.8)	-5.4 (-49.9-39.1)
	No	45 (58.4%)	233 ± 71	(reference)	(reference)
Adenoma at index endoscopy	Yes	27 (35.1%)	207 ± 80	-41.8 (-82.9–-0.6)	-44.9 (-84.9–-4.8)
	No	50 (64.9%)	253 ± 73	(reference)	(reference)
History of adenoma prior to index endoscopy	Yes	16 (20.8%)	272 ± 57	44.6 (-0.3-89.5)	32.1 (-14.0-78.1)
	No	61 (79.2%)	228 ± 81	(reference)	(reference)

* Based on *n*=77, because smoking status was missing for one subject.

† The mean difference between current smokers and ex-smokers was 34.9 (95% CI=-8.0; 77.7).

‡ In further data analyses, ex-smokers and never smokers were collapsed in one category.

§ All HNPCC family members were free of cancer themselves.

Table 2-3. Association between consumption of fruits, vegetables and rectal GST enzyme activity

Exposure	Consumption (g/d)		Difference in GST enzyme activity (nmol/min/mg protein) between individuals with a relatively high and those with a relatively low consumption*,†						*P*-value of interaction
	n=77‡,§		Total group (*n*=77‡)		*GSTM1*-plus (*n*=38)		*GSTM1*-null (*n*=39)		
	Median	(p25; p75)	Difference	(95% CI)	Difference	(95% CI)	Difference	(95% CI)	
*Fruits, vegetables, juices and apple sauce***	347.8	(227.1-516.9)	23.4	(-0.5-47.3)	30.3	(-11.9-72.5)	18.2	(-13.9-50.3)	0.60
Fruits	122.9	(62.8-244.3)	25.6	(5.0-46.2)	27.3	(-7.5-62.0)	23.4	(-4.3-51.1)	0.84
Citrus fruits	30.0	(13.0-76.0)	28.9	(9.3-48.6)	24.3	(-2.6-51.2)	34.2	(-1.1-69.6)	0.61
Other fruits	92.9	(42.0-175.9)	20.8	(-0.7-42.2)	21.6	(-16.8-59.9)	20.3	(-6.9-47.6)	0.95
*Vegetables***	105.3	(84.9-149.0)	12.7	(-13.8-39.2)	36.5	(0.6-72.4)	-27.1	(-69.0-14.8)	0.03
Brassica vegetables	19.2	(11.9-31.2)	10.6	(-8.3-29.4)	22.6	(0.2-45.0)	-25.8	(-63.3-11.8)	0.04
Raw brassica vegetables	2.1	(0.6-4.5)	7.4	(-4.2-19.0)	18.7	(-0.02-37.5)	0.1	(-15.4-15.6)	0.11
Cooked brassica vegetables	16.2	(7.4-28.3)	8.0	(-13.8-29.8)	20.0	(-6.2-46.2)	-30.2	(-74.2-13.8)	0.06
Allium vegetables	7.0	(3.5-12.3)	3.9	(-13.9-21.8)	3.9	(-22.3-30.1)	2.6	(-25.6-30.8)	0.94
Green leafy vegetables	13.2	(7.3-25.5)	-10.4	(-36.2-15.4)	-0.1	(-33.7-33.4)	-36.4	(-81.2-8.5)	0.16
Other vegetables	55.2	(44.0-79.1)	8.6	(-11.9-29.1)	24.3	(-6.4-54.9)	-12.0	(-43.8-19.7)	0.10
Juices and apple sauce	61.5	(18.5-146.6)	1.5	(-19.2-22.2)	-3.2	(-30.0-23.6)	14.4	(-25.5-54.3)	0.42
Citrus fruit juices	28.8	(3.7-69.2)	3.0	(-15.3-21.2)	-2.8	(-24.9-19.4)	20.3	(-21.1-61.7)	0.29
Non-citrus fruit juices and apple sauce	30.8	(10.0-69.8)	-2.4	(-23.9-19.2)	-4.8	(-38.7-29.0)	3.9	(-28.1-36.0)	0.66
Vegetable juices	3.0	(0.6-6.9)	2.4	(-6.8-11.7)	1.3	(-12.4-14.9)	4.2	(-10.2-18.6)	0.73

* Adjusted for presence of adenomas at index endoscopy, HNPCC, total energy intake and current smoking.

† Relatively high and relatively low are defined as the 75th and 25th percentile (p75, p25) of the consumption of each food group, respectively.

‡ One subject was excluded, because no information on smoking status was available.

§ The consumption of fruits and vegetables did not differ between the subjects whose GST activity was assessed and all 94 subjects participating in the study.

** Excluding potatoes and legumes.

estimate to 57.7 nmol/min/mg protein (95% confidence interval (CI)=-6.1-121.6), although statistical significance was lost. GST enzyme activity was not associated with age, sex, HNPCC, *GSTM1* or *GSTT1* genotype and history of adenomas. None of the associations was modified by *GSTM1* or *GSTT1* genotype (results not shown).

Those belonging to a family with known HNPCC had 48.8% (95% CI=20.0-67.3) lower GST-α and 98.9% (95% CI=15.1-100.0) lower GSH protein levels compared with individuals who do not belong to a HNPCC family. These associations did not depend on *GSTM1* and *GSTT1* genotype. Current smokers had 103.2% (95% CI=20.6-242.5) higher GST-α and 53.1% (95% CI=9.6-113.7) higher GSTP1-1 protein levels compared with never smokers, while levels of current and ex-smokers did not differ.

GSTP1-1 protein levels were lower for those with a previous diagnosis of colorectal adenomas compared with those without such a diagnosis (-37.8%, 95% CI=-55.4–-13.3), although no differences between those with and without current adenomas could be detected. For other isoenzymes and characteristics mentioned in Table 2-2, no differences were found.

Associations of fruits and vegetables with GST enzyme activity, according to *GSTM1* genotype, are presented in Table 2-3. Consumption of fruits, in particular citrus fruits, was positively associated with GST enzyme activity. Total consumption of vegetables was positively associated with GST enzyme activity among *GSTM1*-plus carriers, but not *GSTM1*-null carriers. This tended to be true for all subgroups except allium, although the modification was only statistically significant for total vegetables and brassica vegetables. No association between juices and GST enzyme activity could be detected. When brassica vegetables and citrus fruits were included together in the model, their parameter estimates remained essentially the same.

Consumption of citrus fruits, which was associated with total GST enzyme activity, was not associated with GST-α (difference between high and low consumption: 9.7%, 95% CI=-10.7-34.7), GSTM1-1 (8.1%, 95%, CI=-26.3-58.3), GSTP1-1 (0.9%, 95% CI=-18.4-24.7) and GSH (-26.1%, 95% CI=-97.1-1,796.3). Consumption of brassica vegetables, which was associated with GST enzyme activity, was also associated with GSTM1-1 protein levels (difference between high and low consumption: 67.5%, 95% CI=6.8-162.7), but not with GST-α (-1.5%, 95% CI=-24.4-28.4), GSTP1-1 (8.3%, 95% CI=-17.3-41.8) and GSH (177.7%, 95% CI=-95.0-15,405.8). Consumption of allium vegetables was not associated with GST enzyme activity, but was associated negatively with GSTP1-1 levels (difference between high and low consumption: -23.3%, 95% CI=-35.5–-8.6). Consumption of green leafy vegetables, which was also not associated with GST enzyme activity, was positively associated with GSTM1-1 levels (133.3%, 95% CI=11.7-387.4), and negatively with GST-α levels (26.5%, 95% CI=-44.2–-3.3).These associations were not modified by *GSTM1* and *GSTT1* genotype, and remained after adding consumption of brassica vegetables to the model.

Other associations of fruits and vegetables with GST isoenzyme and GSH levels were absent.

Dietary fiber, vitamin A, ß-carotene and folate did not affect the parameter estimates, whereas the effect of vitamin C could not be disentangled from the contribution of citrus fruits due to their high correlation (r=0.76).

Discussion

We observed that consumption of fruits, in particular citrus fruits, is positively associated with higher rectal GST enzyme activity. Consumption of vegetables, in particular brassica vegetables, was positively associated with GST enzyme activity, although this was limited to *GSTM1*-plus genotype carriers. Consumption of brassica vegetables was also positively associated with GSTM1-1 protein levels. No indication was found of an association between consumption of allium vegetables and GST enzyme activity, but allium vegetables were negatively associated with GSTP1-1 protein levels.

We requested subjects to report consumption of fruits and vegetables according to the year preceding endoscopy by filling out a food-frequency questionnaire. While the validity of the assessment was moderate, the reproducibility of the assessment was high. Although the assessment, like all assessments, is not perfect and may be subject to bias, this method allows assessing habitual consumption with minimal intrusion upon the participants. However, it is hard to determine whether we have assessed consumption in the time window relevant to GST induction. In rat hepatoma cells, GST induction was first observed 8 h after induction, with highest values after 24-48 h[30]. In humans, rectal GST induction is likely to be a delayed process, as bioactive compounds have to reach the intestinal tissue, either via blood or the gastrointestinal tract. Thereupon, the higher activity might persist for some time for biological reasons or prolonged exposure. Indeed, Szarka and co-workers reported that human rectal GST enzyme activity was stable over a 2-4 week interval in another cross-sectional study[31].

We included subjects undergoing endoscopy for medical reasons instead of healthy subjects. This is unlikely to cause problems as subjects with colorectal cancer and inflammatory bowel diseases were excluded since these conditions may influence metabolism and gene expression; meanwhile controlling for self-reported bowel complaints did not importantly alter the estimates. Nevertheless, given subject selection on HNPCC or adenoma status, we adjusted all parameter estimates for these conditions, although this did not influence them. Likewise, these selection criteria modify none of the examined associations.

Our study design allowed us to evaluate the association between GST enzyme activity and adenoma presence, sex, age, HNPCC and genetic polymorphisms in GSTs. In line with the postulated role of GST induction in cancer prevention and previously published results[31], the presence of adenomas at index endoscopy was associated with a lower GST enzyme activity in normal rectal tissue compared with rectal tissue of patients without adenomas. Cancer-free members of HNPCC families had lower rectal GST-α and GSH protein levels than those who did not belong to a HNPCC family. These differences, which could not be attributed to differences in *GSTM1* and *GSTT1* genotype, were unexpected and have not been reported before; therefore, these findings should be treated with caution. Interestingly, smoking was associated with a higher GST enzyme activity (Table 2-2) and higher GST-α and GSTP1-1 protein levels. Indeed, rodent studies showed that several polycyclic aromatic hydrocarbons present in cigarette smoke induce GST enzyme activity as part of an adaptive response mechanism to chemical stress[3].

The main hypothesis concerned the association between fruits, vegetables and GST enzyme activity, isoenzyme levels and GSH. First, fruits, in particular citrus fruits, were associated

positively with GST enzyme activity. This might be attributed to the coumarin auraptene and to limonoids (highly oxidized triterpenoids), which are present in high concentrations in citrus fruit tissues[17, 32]. Indeed, auraptene increased GST enzyme activity in rat colon tissue[33] and limonoids induced GST enzyme activity in small intestinal mucosa and liver of mice[34]. However, auraptene and limonoids are also present in citrus fruit juices[17], which were not associated with GST enzyme activity in this study. As processed juices comprise the larger part of the intake of juices, compounds that are lost during the processing of citrus fruits to juices may be responsible for these findings. Additionally, bioavailability of phytochemicals might differ between whole fruits and fruit juices. Other compounds present in citrus fruits may also account for the higher GST enzyme activity, such as citrus flavonoids and δ-limonene[35]. The latter enhanced colonic GST enzyme activity in rats[36], albeit at doses that far exceed human exposure.

Second, consumption of vegetables was positively associated with GST enzyme activity among those with the *GSTM1*-plus genotype, but not those with the *GSTM1*-null genotype. This association was most evident for brassica vegetables, which are rich in glucosinolates. Isothiocyanates may be primarily responsible for this GST inducing capacity[37], although glucosinolate content may depend on species, maturity and processing of the plants[38].

So far, only one study examined consumption of brassica vegetables in relation to rectal GST enzyme activity. In this study, ten human volunteers consumed 300 g/d of Brussels sprouts for seven days. No effect upon rectal GST enzyme activity and GSH levels was found, while GSTA and GSTP1-1 levels increased by 30 and 15%, respectively[11]. In our study, consumption of brassica vegetables was not associated with GST-α and GSTP1-1 levels. However, it was associated with GSTM level, which is in line with the observed association between brassica vegetables and GST enzyme activity among *GSTM1*-plus genotype carriers.

Furthermore, the synthetic dithiolethione oltipraz - dithiolethiones are a class of chemical compounds, which also occur naturally in brassica vegetables - increased colonic mucosal GST enzyme activity at single oral doses of 125 and 250, but not at 500 or 1000 mg/m^2, in 24 subjects having a family history of colorectal cancer or a personal history of colorectal polyps or carcinomas[39]. However, this could not be confirmed in two other trials[40, 41]. Other studies have evaluated the effect of brassica vegetables upon GST enzyme activity in other tissues[11-15], but because of organ-specific patterns of expression[7] and differences in bioavailability of bioactive compounds able to induce GST, it is hard to compare their results to those of the present study.

Although compounds of allium vegetables did positively affect the GSH/GST detoxification system in many *in vitro* or *in vivo* studies[3], as well as in a human study on lymphocytes[12], we observed a negative association between allium vegetables and GSTP1-1 levels. However, a previous study in which six healthy volunteers were given 250 g/d of mixed vegetables (among which allium vegetables) for a period of three weeks reported decreased GSTP1 protein levels in five of the six subjects, which is in line with our observation[42].

Assessing overall GST enzyme activity, no correlation between GST isoenzyme levels and GST enzyme activity appeared. This could be partially attributed to our substrate CDNB, which reacts with moderate activity to GSTA1, GSTA2 and GSTP1, and with high activity to GSTM1 and GSTM2[3, 43]. The GST-μ and GST-π classes belong to the most important GST

isoenzymes in the rectum[11]. While this makes CDNB a good general substrate, it is not highly specific to a particular isoenzyme, which might be related to the lack of correlation between GST isoenzymes and GST enzyme activity. Besides, GSTM2 protein levels were not assessed, while GSTM2 contributes to GST enzyme activity. As GSTM2 is expressed in the human colon and is inducible[8], induction of GSTM2 might further explain the lack of correlation. Alternatively, the assessment of GST enzyme activity might be slightly biased, as GSTs also serve as binding and transport proteins, resulting in a temporary loss of enzyme activity[3, 43].

Associations with fruits and vegetables appeared to be different for GST isoenzymes and GST enzyme activity. For citrus fruits, we found an association with GST enzyme activity, but this could not be attributed to any of the assessed isoenzymes. Theoretically, we cannot exclude the possibility that citrus fruits induce GSTM2 exclusively. However, it seems more likely that effects on specific isoenzymes are too small to be detected, whereas effects on GST enzyme activity - which reflects isoenzyme levels of GSTA, GSTM and GSTP1-1 in combination with their specific activity - can be detected. This might also explain why consumption of green leafy vegetables was associated with GSTM1-1 and GST-α levels, while no association with GST enzyme activity was found. As these associations have not been examined before, the actual nature of the relationships remains to be established.

The underlying idea of this study is that high GST enzyme activity is believed to be beneficial to cancer prevention. GST enzyme activity in normal mucosa along the gastrointestinal tract was inversely correlated with tumor incidence[6]; inhibition of rat bowel carcinogenesis was shown to correlate with induction of GST[33]; and compounds that induce GST generally counteract cytogenetic damage[44]. However, it should be noted that certain food components (*e.g.* sulforaphane and indole 3-carbinol present in broccoli[45, 46]) may affect phase I as well as phase II enzymes; the balance between induction and inhibition of different detoxification systems is of utmost relevance[47].

In conclusion, our study shows positive associations of habitual levels of consumption of fruits and vegetables with rectal GST enzyme activity. This is especially true for citrus fruits and brassica vegetables, but positive associations with the latter only existed for *GSTM1*-plus individuals.

Acknowledgements

We are grateful to all patients who kindly participated in this study. We thank Dr. Dorien Voskuil for collecting data, Claudia van den Braak for assessing GST/GSH measurements, Annelies Bunschoten and Jan Harryvan for assessing *GSTM1* and *GSTT1* genotypes, Dr. Marga Ocké for calculation of food and nutrient intake, Ivon Milder for contributing to this study during her MSc training, and Rick Yagodich (UK) for editing and proofreading. This work was supported by grants from the Netherlands Digestive Diseases Foundation (grant number WS 94-54) and Netherlands Organisation for Health Research and Development (grant number 980-10-026).

References

1. Roediger WE, Babidge W. Human colonocyte detoxification. *Gut.* 1997;41(6):731-734.
2. Hoensch HP, Hartmann F. The intestinal enzymatic biotransformation system: potential role in protection from colon cancer. *Hepatogastroenterology.* 1981;28(4):221-228.
3. Hayes JD, Pulford DJ. The glutathione *S*-transferase supergene family: regulation of GST and the contribution of the isoenzymes to cancer chemoprotection and drug resistance. *Crit Rev Biochem Mol Biol.* 1995;30(6):445-600.
4. Talalay P, Fahey JW, Holtzclaw WD, Prestera T, Zhang Y. Chemoprotection against cancer by phase 2 enzyme induction. *Toxicol Lett.* 1995;82-83:173-179.
5. Peters WHM, Kock L, Nagengast FM, Kremers PG. Biotransformation enzymes in human intestine: critical low levels in the colon? *Gut.* 1991;32(4):408-412.
6. Peters WHM, Roelofs HMJ, Hectors MPC, Nagengast FM, Jansen JBMJ. Glutathione and glutathione *S*-transferases in Barrett's epithelium. *Br J Cancer.* 1993;67(6):1413-1417.
7. Coles BF, Chen G, Kadlubar FF, Radominska-Pandya A. Interindividual variation and organ-specific patterns of glutathione *S*-transferase alpha, mu, and pi expression in gastrointestinal tract mucosa of normal individuals. *Arch Biochem Biophys.* 2002;403(2):270-276.
8. Ebert MN, Klinder A, Peters WH, et al. Expression of glutathione *S*-transferases (GSTs) in human colon cells and inducibility of GSTM2 by butyrate. *Carcinogenesis.* 2003;24(10):1637-1644.
9. Cotton SC, Sharp L, Little J, Brockton N. Glutathione *S*-transferase polymorphisms and colorectal cancer: a HuGE review. *Am.J.Epidemiol.* 2000;151(1):7-32.
10. Clapper ML, Szarka CE, Pfeiffer GR, et al. Preclinical and clinical evaluation of broccoli supplements as inducers of glutathione *S*-transferase activity. *Clin Cancer Res.* 1997;3(1):25-30.
11. Nijhoff WA, Grubben MJAL, Nagengast FM, et al. Effects of consumption of Brussels sprouts on intestinal and lymphocytic glutathione *S*-transferases in humans. *Carcinogenesis.* 1995;16(9):2125-2128.
12. Lampe JW, Chen C, Li S, et al. Modulation of human glutathione *S*-transferases by botanically defined vegetable diets. *Cancer Epidemiol Biomarkers Prev.* 2000;9(8):787-793.
13. Bogaards JJP, Verhagen H, Willems MI, Van Poppel G, Van Bladeren PJ. Consumption of Brussels sprouts results in elevated alpha-class glutathione *S*-transferase levels in human blood plasma. *Carcinogenesis.* 1994;15(5):1073-1075.
14. Nijhoff WA, Mulder TPJ, Verhagen H, van Poppel G, Peters WHM. Effects of consumption of brussels sprouts on plasma and urinary glutathione *S*-transferase class-alpha and -pi in humans. *Carcinogenesis.* 1995;16(4):955-957.
15. Steinkellner H, Hietsch G, Sreerama L, et al. Induction of glutathione-*S*-transferases in humans by vegetable diets. *Dietary anticarcinogens and antimutagens: chemical and biological aspects.* Cambridge, UK: Royal Society of Chemistry; 2000: 193-198.
16. Hoensch H, Morgenstern I, Petereit G, et al. Influence of clinical factors, diet, and drugs on the human upper gastrointestinal glutathione system. *Gut.* 2002;50(2):235-240.
17. Hasegawa S, Miyake M. Biochemistry and biological functions of citrus limonoids. *Food Rev Int.* 1996;12(4): 413-435.
18. Steinmetz KA, Potter JD. Vegetables, fruit, and cancer prevention: a review. *J Am Diet Assoc.* 1996;96(10): 1027-1039.
19. Voskuil DW, Kampman E, Grubben MJAL, et al. Meat consumption and meat preparation in relation to colorectal adenomas among sporadic and HNPCC family patients in The Netherlands. *Eur J Cancer.* 2002;38(17):2300-2308.
20. Vasen HF, Mecklin JP, Khan PM, Lynch HT. The International Collaborative Group on Hereditary Non-Polyposis Colorectal Cancer (ICG-HNPCC). *Dis Colon Rectum.* 1991;34(5):424-425.

21. Ocké MC, Bueno de Mesquita HB, Goddijn HE, et al. The Dutch EPIC food frequency questionnaire. I. Description of the questionnaire, and relative validity and reproducibility for food groups. *Int J Epidemiol.* 1997;26 Suppl 1:S37-S48.

22. Stichting Nederlands Voedingsstoffenbestand. *Dutch Food Composition Table. Nevo tabel: Nederlands voedingsstoffenbestand*: The Hague : Voorlichtingsbureau voor de Voeding; 1996.

23. Habig WH, Jakoby WB. Assays for differentiation of glutathione *S*-transferases. *Methods Enzymol.* 1981;77:398-405.

24. Fahey RC, Newton GL. Determination of low-molecular-weight thiols using monobromobimane fluorescent labeling and high-performance liquid chromatography. *Methods Enzymol.* 1987;143:85-96.

25. Nijhoff WA, Groen GM, Peters WHM. Induction of rat hepatic and intestinal glutathione *S*-transferases and glutathione by dietary naturally-occurring anticarcinogens. *Int J Oncol.* 1993;3(6):1131-1139.

26. Lowry OR, Rosebrough NH, Fall AR, Randall RJ. Protein measurement with folin phenol reagent. *J Biol Chem.* 1951;193:265-275.

27. Arand M, Muhlbauer R, Hengstler J, et al. A multiplex polymerase chain reaction protocol for the simultaneous analysis of the glutathione *S*-transferase GSTM1 and GSTT1 polymorphisms. *Anal.Biochem.* 1996;236(1):184-186.

28. Willett W. *Nutritional Epidemiology.* 2nd ed. New York, NY: Oxford University Press; 1998.

29. Long JS. Limited Outcomes: The Tobit Model. *Regression Models for Categorical and Limited Dependent Variables.* Vol 7. Thousand Oaks, CA, USA: Sage Publications; 1997:187-216.

30. 't Hoen PA, Rooseboom M, Bijsterbosch MK, Van Berkel TJ, Vermeulen NP, Commandeur JN. Induction of glutathione-*S*-transferase mRNA levels by chemopreventive selenocysteine Se-conjugates. *Biochem Pharmacol.* 2002;63(10):1843-1849.

31. Szarka CE, Pfeiffer GR, Hum ST, et al. Glutathione *S*-transferase activity and glutathione *S*-transferase mu expression in subjects with risk for colorectal cancer. *Cancer Res.* 1995;55(13):2789-2793.

32. Ogawa K, Kawasaki A, Yoshida T, et al. Evaluation of auraptene content in citrus fruits and their products. *J Agric Food Chem.* 2000;48(5):1763-1769.

33. Tanaka T, Kawabata K, Kakumoto M, et al. Citrus auraptene exerts dose-dependent chemopreventive activity in rat large bowel tumorigenesis: the inhibition correlates with suppression of cell proliferation and lipid peroxidation and with induction of phase II drug-metabolizing enzymes. *Cancer Res.* 1998;58(12):2550-2556.

34. Lam LK, Hasegawa S. Inhibition of benzo[a]pyrene-induced forestomach neoplasia in mice by citrus limonoids. *Nutr Cancer.* 1989;12(1):43-47.

35. Hakim IA, Hartz V, Graver E, Whitacre R, Alberts D. Development of a questionnaire and a database for assessing dietary d-limonene intake. *Public Health Nutr.* 2002;5(6A):939-945.

36. Van Lieshout EM, Posner GH, Woodard BT, Peters WHM. Effects of the sulforaphane analog compound 30, indole-3-carbinol, D-limonene or relafen on glutathione *S*-transferases and glutathione peroxidase of the rat digestive tract. *Biochim Biophys Acta.* 1998;1379(3):325-336.

37. Fahey JW, Zhang Y, Talalay P. Broccoli sprouts: an exceptionally rich source of inducers of enzymes that protect against chemical carcinogens. *Proc Natl Acad Sci U S A.* 1997;94(19):10367-10372.

38. Mithen R, Dekker M, Verkerk R, Rabot S, Johnson IT. The nutritional significance, biosynthesis and bioavailability of glucosinolates in human foods. *J Sci Food Agric.* 2000;80:967-984.

39. O'Dwyer PJ, Szarka CE, Yao KS, et al. Modulation of gene expression in subjects at risk for colorectal cancer by the chemopreventive dithiolethione oltipraz. *J Clin Invest.* 1996;98(5):1210-1217.

40. Benson III AB, Olofunmilayo OI, Ratain MJ, et al. Chronic daily low dose of 4-methyl-5-(2-pyrazinyl)-1,2-dithiole-3-thione (Oltipraz) in patients with previously resected colon polyps and first degree female relatives of breast cancer patients. *Clin Cancer Res.* 2000;6(10):3870-3877.

41. Szarka CE, Yao KS, Pfeiffer GR, et al. Chronic dosing of oltipraz in people at increased risk for colorectal cancer. *Cancer Detect Prev.* 2001;25(4):352-361.

42. Persson I, He L, Fang C, Normen L, Rylander R. Influence of vegetables on the expression of GSTP1 in humans - a pilot intervention study (Sweden). *Cancer Causes Control.* 2000;11(4):359-361.

43. Eaton DL, Bammler TK. Concise review of the glutathione *S*-transferases and their significance to toxicology. *Toxicol. Sci.* 1999;49(2):156-164.

44. Grubben MJAL, Nagengast FM, Katan MB, Peters WHM. The glutathione biotransformation system and colorectal cancer risk in humans. *Scand J Gastroenterol Suppl.* 2001(234):68-76.

45. Lampe JW, King IB, Li S, et al. Brassica vegetables increase and apiaceous vegetables decrease cytochrome P450 1A2 activity in humans: changes in caffeine metabolite ratios in response to controlled vegetable diets. *Carcinogenesis.* 2000;21(6):1157-1162.

46. Kall MA, Vang O, Clausen J. Effects of dietary broccoli on human drug metabolising activity. *Cancer Lett.* 1997;114(1-2):169-170.

47. Paolini M, Mesirca R, Gialluca N, Bauer C, Biagi GL, Cantelli-Forti G. On cancer chemoprevention: complications and limitations of some proposed strategies. *Carcinogenesis.* 1995;16(4):971-973.

Abstract

Background: K-*ras* mutation-positive (K-*ras*$^+$) and -negative (K-*ras*$^-$) colorectal adenomas may differ clinically and pathologically.

Aim: As environmental compounds may cause mutations in the growth-related K-*ras* oncogene or affect clonal selection depending on mutational status, we evaluated whether the etiology of K-*ras*$^+$ and K-*ras*$^-$ adenomas differs.

Results: K-*ras* mutations in codons 12 and 13 were assessed in colorectal adenoma tissue (K-*ras*$^+$: n=81, K-*ras*$^-$: n=453). Dietary and lifestyle data were collected through questionnaires that were also administered to 709 polyp-free controls. Multiple logistic regression analyses showed that intake of vitamin B2 and monounsaturated fat were differently associated with risk of K-*ras*$^+$ and K-*ras*$^-$ adenomas; vitamin B2 was inversely associated with K-*ras*$^-$ (highest versus lowest tertile: odds ratio (OR)=0.70, 95% confidence interval (CI)=0.50-0.97, P_{trend}=0.020), but not with K-*ras*$^+$ adenomas, and a positive association with monounsaturated fat was confined to K-*ras*$^-$ adenomas (OR=1.57, 95% CI=1.06-2.34, P_{trend}=0.029). Besides, potential, not statistically significant, differences in risk arose because red meat was distinctly positively associated with K-*ras*$^+$ adenomas (OR=1.70, 95% CI=0.94-3.09, P_{trend}=0.061); total dietary and polyunsaturated fat tended to be inversely associated with risk of K-*ras*$^+$ but not of K-*ras*$^-$ adenomas; inverse associations with dairy products, calcium, protein and tea were confined to K-*ras*$^-$ adenomas, and smoking was more markedly positively associated with K-*ras*$^-$ adenomas. No differences in risk of K-*ras*$^+$ and K-*ras*$^-$ adenomas could be detected for other factors.

Conclusion: Dietary and lifestyle factors may influence risk of K-*ras*$^+$ and K-*ras*$^-$ adenomas differently. However, epidemiological literature on diet, lifestyle and colorectal K-*ras* mutations is inconsistent.

Diet, Lifestyle and Risk of K-*ras* Mutation-Positive and -Negative Colorectal Adenomas

3

Petra A. Wark
Wieke van der Kuil
Janneke Ploemacher
Goos N.P. van Muijen
Chris J.J. Mulder
Matty P. Weijenberg
Frans J. Kok
Ellen Kampman

Int J Cancer. 2006;119(2):398-405

Introduction

Over the course of time, a set of modifiable risk factors for colorectal cancer and their presumed precursors colorectal adenomas has emerged. These risk factors include excess body weight, physical inactivity, alcohol consumption, smoking, consumption of red and processed meat, low intake of folate, calcium and non-steroidal anti-inflammatory drugs (NSAIDs)[1, 2]. However, findings are not entirely consistent.

Mutations in the Kirsten-*ras* (K-*ras*) gene are the most common abnormality of oncogenes in human tumors[3], including colorectal adenomas[4] and carcinomas[5, 6]. The K-*ras* gene encodes the p21ras protein: a guanine nucleotide binding-protein that plays a role in transmitting growth stimulatory signals from membrane-bound tyrosine kinases through a group of downstream regulators to the nucleus. When a point mutation occurs in the K-*ras* gene, the ras protein is locked in its active state leading to a prolonged mitotic signal and thereupon to cell growth and tissue proliferation[7].

K-*ras* mutations occur more frequently in adenomas with a villous histology[8]. Especially glycine-to-valine mutations in codon 12 of the gene may predispose to more aggressive clinical behavior in colorectal cancer[9]. It has also been suggested that K-*ras* mutational status may distinguish two groups of familial non-hereditary non-polyposis colorectal cancer patients whose cancers currently have unknown genetic origins[10]. The aforementioned pathological and clinical differences may imply that the etiology of K-*ras* mutation-positive and mutation-negative adenomas may also differ, potentially also according to specific type of mutation.

Indeed, environmental carcinogens, such as *N*-nitroso compounds that are endogenously formed after red meat consumption or derived from cigarette smoke or processed meat[11, 12], and polycyclic aromatic hydrocarbons present in cigarette smoke and formed during grilling of meat over a direct flame[11], may induce (specific) K-*ras* mutations[13-15]. Environmental factors may also select for K-*ras* mutation-negative lesions as suggested for calcium[16] and heterocyclic amines[17] that are formed during meat cooking at high temperature[11, 12], for instance by reducing the mutation rate; the latter was shown for *n*-3 polyunsaturated fatty acids in rats[18]. Moreover, environmental factors may affect growth of tumors of K-*ras* mutation-positive and -negative tumors differently. The NSAID sulindac, for example, selectively modified amplification of cells with a K-*ras* mutation in codon 12 in rats[19].

In spite of this, the potential environmental influences on the occurrence and persistence of K-*ras* mutations are largely unknown[3], especially in humans. To date, only one cross-sectional study has evaluated whether the distributions of risk factors differ between patients with K-*ras* mutation-positive and -negative colorectal adenomas: K-*ras* mutation-positive adenoma patients were more likely to have a lower intake of total folate and tended to have higher level of vigorous aerobic exercise than did K-*ras* mutation-negative adenoma patients[20]. We conducted a case-control study in the Netherlands to further evaluate whether associations between presumed dietary and lifestyle risk factors and risk of colorectal adenomas depend on K-*ras* mutational status.

Material and Methods

Study Population

Cases and controls were recruited among those undergoing endoscopy of the large bowel in the outpatient clinics of ten hospitals in the Netherlands between June 1997 and June 2000. Medical review boards of all participating hospitals and Wageningen University approved the study protocol. Potential participants were informed about the study through printed material handed to them, at the time of the endoscopy, by staff, or sent by mail within three months thereof. The overall response rate was 54% and varied from 35 to 91% in the various outpatient clinics. To be eligible, participants had to be between 18 and 75 years old at the time of endoscopy, be a Caucasian and be able to speak Dutch. They also had to be free of hereditary colorectal cancer syndromes, chronic inflammatory bowel diseases and should not have a history of colorectal cancer or (partial) bowel resection. All participants provided written informed consent. Cases were defined as participants who had at least one histologically confirmed colorectal adenoma within three years of recruitment (n=658). Controls were never diagnosed with any type of polyps (n=709). Ninety-two percent of the cases and 85% of the controls underwent a full endoscopy (*i.e.* full colonoscopy or sigmoidoscopy combined with X-ray).

Exposure Assessment

Participants were requested to fill out a questionnaire on lifestyle and socio-economic factors, a validated semi-quantitative food-frequency questionnaire (FFQ)[21, 22] and a questionnaire on meat consumption and meat preparation methods[23] covering the year preceding study inclusion or bowel complaints.

The FFQ assessed the habitual consumption of 178 foods and covered at least 90% of the population mean intake of the food groups and nutrients of interest in the Netherlands[21]. For most items, consumption could be indicated per day, month, or year. For some foods that are consumed less regularly, consumption frequency could be reported by ticking always/most, often, sometimes or rarely/never. Quantity of consumption could be indicated in pre-specified units for most foods, photographs of two to four different-sized portions were shown to estimate this for 21 foods, and standard portion sizes were assumed for a few foods. Daily consumption in grams per day was calculated by multiplying consumption frequency per day by portion size. Energy and nutrient intake were calculated using the 1996 computerized version of the Dutch food composition table[24]. The questionnaire seems adequate for ranking according to intake of most food groups and nutrients ($r \geq 0.40$ *vs.* dietary records), although the relative validity for vegetables (r=0.38 for men and 0.31 for women), fish (r=0.32 for men and 0.37 for women), β-carotene (r=0.23 for both sexes), vitamin C for men (r=0.37) and vitamin E for women (r= 0.35) was low. The 6- and 12-month reproducibility of the assessment was high for all food groups and nutrients ($r \geq 0.45$)[21, 22].

The additional questionnaire on meat, described in detail elsewhere[23], assessed consumption frequency and portion sizes of 16 types of meat, gravy, and meat preparation methods (*e.g.* height of heat source, addition of water, use of lid). Besides, color photographs of fried beef patties, pork

chops, steak and bacon, which originated from a Swedish questionnaire[25], were included to assess habitual color of meat surface as proxy for preparation at 225, 200, 175 and 150ºC.

Adenoma Characteristics and K-*ras* Mutation Analysis

We reviewed relevant medical records to obtain information on the indication for endoscopy, polyp history, completeness of colonoscopy and adenoma location, size and number. Besides, the national pathology number of the tumor tissue was recorded to enable retrieval. We obtained tumor tissue from 615 of the 658 eligible adenoma patients (93%) for K-*ras* mutation analysis. One experienced gastrointestinal pathologist judged the histopathology of all obtained specimens. One sample turned out to be a carcinoma and six to be hyperplastic polyps, 15 samples did not contain any polyp tissue, tissue was sparse in 44 and the quality of one sample was poor; these 67 samples were excluded.

DNA was extracted from the formalin-fixed paraffin-embedded tissue (10-12 sections, 10 μm thick) using the Puregene™ DNA isolation kit (Gentra Systems, Minneapolis, MN). Microdissection was performed by a hematoxylin and eosin-stained 4 μm section, and only those areas containing at least 60% tumor cells were used. The tissue was incubated overnight at 55ºC in 500 μl lysis solution containing 0.5 mg/ml Proteinase K (Roche Diagnostics, Mannheim, Germany), followed by 72 hours at 37ºC. Proteins were removed by precipitation according to the manufacturer's protocol. DNA was precipitated with 500 μl 100% isopropanol at 4ºC for 30 minutes. The pellet was washed with 500 μl 70% ethanol, air-dried, and subsequently the DNA was rehydrated in 30 μl DNA hydration solution.

Codon 12 and 13 of the K-*ras* gene were examined for mutations by direct-sequencing in most samples. Polymerase Chain Reactions (PCRs) were performed in 35 cycles of 30 seconds at 94ºC, 45 seconds at 55ºC and one minute at 72ºC, followed by a final extension step at 72ºC for five minutes. The PCR reaction mixture consisted of 100 ng of purified DNA, 20 mmol/l $(NH_4)_2SO_4$, 75 mmol/l Tris-HCl (pH 9.0), 0.01% Tween, 200 μmol/l dNTPs, 2.5 mmol/l $MgCl_2$ and 0.4 μmol/l of the following primers: ACTCATGAAAATGGTCAGAG (3'-primer) and GTACTGGTGGAGTATTTGATAG (5'-primer). PCR products were checked on an ethidium bromide-stained 2% agarose gel and purified with the QIAquick PCR Purification Kit (Qiagen, Valencia, CA) according to the manufacturer's protocol. Subsequently, the PCR products were sequenced using ABI PRISM® BigDye™ Terminators v3.0 (Applied Biosystems, Foster City, CA) and cycle sequencing with AmpliTaq® DNA polymerase, FS (Applied Biosystems). Sequencing was performed in both directions using the same primers. Samples were analyzed on an ABI PRISM® 3700 DNA Analyzer (Applied Biosystems). Forty-two samples were analyzed using Mutant-Allele-Specific Amplification (MASA)-PCR as described previously[26], 12 of which were mutation-positive. The direct sequencing method and MASA-PCR methods allowed detection of the presence of 5% and 10-15% mutant DNA, respectively. Additionally, we performed sequence analyses on a number of samples (n=8), each with a mutation in codon 12 or 13 according to MASA. In all samples, the same mutation was found.

K-*ras* mutation analysis was successful in 534 of the 548 available samples, yielding a study population of 81 cases with K-*ras* mutation-positive adenomas, 453 cases with K-*ras* mutation-negative adenomas and 709 controls.

Statistical Analysis

In the analyses, we evaluated factors that are thought to be risk factors for colorectal adenomas and carcinomas[1, 2], hypothesized to be related to K-*ras* mutational status[3, 13, 20, 26-29] and factors that were previously associated with colorectal adenomas in our study population[30, 31].

The distributions of demographic and tumor characteristics and presumed dietary and lifestyle risk factors were examined visually and numerically for cases according to K-*ras* mutational status and controls. All nutrients bar alcohol were adjusted for total energy intake using the residual method[32]. Continuous exposure variables were divided into tertiles based on the distribution among controls. The distributions of risk factors among K-*ras* mutation-positive and -negative cases were compared to evaluate etiologic heterogeneity using multinomial logistic regression (case-case analyses). Simultaneously, odds ratios (OR) and 95% confidence intervals (CI) were estimated for K-*ras* mutation-positive and -negative adenomas separately versus controls to evaluate whether risk factors were positively or inversely associated with the disease (case-control analyses).

Separate models, which always included age (continuous), gender and total energy intake (kJ/d, continuous), were fitted for each factor under investigation. Either one of the other examined factors, family history of colorectal cancer, indication for endoscopy (bowel-complaint related yes/no), education level or hormone replacement therapy (set to 0 for men) was added in turn to evaluate potential confounding. Red meat changed the parameter estimate for monounsaturated fat by more than 10% and was therefore added to the model for this factor. None of the other factors changed specific parameter estimates by more than 10%, and they were therefore not included in the final models.

An ordinal score value based on the median value within each tertile among controls was used to test for trend across tertiles. We also evaluated whether the distribution of risk factors over tertiles or exposure categories differed for patients having codon 12 transitions and codon 12 transversion mutations by Fisher's exact test, whereas we visually examined the distribution of codon 13 mutations.

As sensitivity analyses, we evaluated whether the parameter estimates changed importantly when we excluded participants who indicated to have ever changed their diet, who did not undergo a full endoscopy, cases who were previously diagnosed with an adenoma, had multiple adenomas, advanced adenomas (≥1cm or (tubulo)villous structure) or no adenomas in the distal colon and samples that were analyzed by MASA-PCR in turn. Furthermore, we evaluated whether the associations were driven by potential differential consumption patterns of men and women by using gender-specific tertiles. We also applied quartiles, instead of tertiles, in the analyses.

All reported *p*-values are two-sided and regarded as statistically significant if less than 0.05. We used Statistical Analysis Software (SAS version 9.1, SAS Institute, Cary, NC) for all analyses.

Results

Table 3-1 shows the characteristics of K-*ras* mutation-positive (K-*ras*$^+$) and -negative (K-*ras*$^-$) cases, cases in whose tissue K-*ras* mutational status could not be assessed and controls. The entire case group was slightly older and consisted of relatively more men and smokers. A smaller proportion of cases than controls underwent endoscopy for bowel complaints. The case group had a higher intake of total energy, meat, fruits, vegetables and alcohol compared with the control group. The characteristics of cases in whose tissue K-*ras* mutational status could not be assessed were largely comparable to the characteristics of K-*ras*$^-$ cases, although they resembled K-*ras*$^+$ cases more with respect to gender and consumption of fruits. Fifteen percent of the cases in which we could assess K-*ras* mutational status harbored mutations in the gene. There were fewer smokers among K-*ras*$^+$ cases than among K-*ras*$^-$ cases. On average, K-*ras*$^+$ cases had a higher total energy intake and consumed more fruit compared with K-*ras*$^-$ cases. A smaller proportion of K-*ras*$^+$ cases had adenomas before, but a larger proportion was diagnosed with multiple adenomas compared with K-*ras*$^-$ cases. The majority of K-*ras*$^+$ adenomas had villous or tubulovillous histology and were large, whereas the majority of K-*ras*$^-$ adenomas were tubular and small. Relatively more K-*ras*$^-$ adenomas occurred in the distal colon.

Among the 81 specimens with K-*ras* mutations, 66 carried mutations in codon 12 exclusively (62 monoallelic, four biallelic), 14 a monoallelic mutation in codon 13 exclusively, and one carried both a biallelic codon 12 mutation and a monoallelic codon 13 mutation. Transversion from K-*ras* codon 12 GGT to GTT (Gly→Val) was the most common mutation (n=29), followed by transition to GAT (→Ala, n=23). Other mutations in codon 12 were to TGT (→Cys, n=6), AGT (→Ser, n=6), CGT (→Arg, n=5), and GCT (→Ala, n=3). In codon 13, only transition mutations from GGC to GAC (Gly→Asp) occurred.

Table 3-2 presents case-case comparisons according to potential modifiable risk factors. Cases with K-*ras*$^+$ adenomas were more likely to have a lower intake of monounsaturated fat and a higher intake of vitamin B2 than cases with K-*ras*$^-$ adenomas. The distributions of total dietary and polyunsaturated fat, calcium, and red meat, and less strikingly, of protein, dairy products, tea and (duration of) smoking also appeared to differ between K-*ras*$^+$ and K-*ras*$^-$ cases, although not statistically significantly. Dairy products, calcium and protein were strongly correlated with intake of vitamin B2 (Spearman's $r \geq 0.60$). The inverse association with dairy products was still visible when we added vitamin B2 or calcium to the model, albeit the case-case comparison hardly showed differences anymore (highest *vs.* lowest tertile: OR=0.74 and P_{trend}=0.58; OR=1.09 and P_{trend}=0.82, respectively). Nonetheless, when we added dairy products, calcium or protein to the model for vitamin B2, the case-case comparison still indicated differences between K-*ras*$^+$ and K-*ras*$^-$ adenomas (OR=2.37 and P_{trend}=0.075; OR=1.67 and P_{trend}=0.20; OR=1.78 and P_{trend}=0.083, respectively). This was also the case when we added dairy products (OR=1.70 and P_{trend}=0.23) or protein (OR=1.64 and P_{trend}=0.18) to the model on calcium. However, addition of vitamin B2 decreased the strength of the case-case association considerably (OR=1.20 and P_{trend}=0.73). Similarly, the potential differential association with protein disappeared when we added vitamin B2 to the model (OR=1.13 and P_{trend}=0.71). We did not find indications for differential associations according to meat preparation methods (estimates not shown). No

Table 3-1. Description of the study population

Characteristic	K-*ras* mutation-positive adenomas (K-*ras*$^+$) *n*=81
Demographic and lifestyle factors	
Age (yr), mean ± SD	60.3 ± 9.6
Female, *n* (%)	39 (48.2%)
BMI, kg/m^2, mean ± SD	26.1 ± 4.2
Low education level, *n* (%)‡	30 (44.1%)
Ever smoking, *n* (%)	49 (61.3%)
Duration of smoking among smokers (yr), median (p25; p75)	25 (20; 41.5)
High physical activity, *n* (%)§	24 (32.9%)
Medical history, n (%)	
Family history of colorectal cancer	18 (22.2%)
> 12 times NSAIDs/yr	23 (28.4%)
Bowel complaints or defecation problems as indication of endoscopy	36 (46.2%)
Dietary factors, median (p25; p75)	
Total energy intake (kJ/d)	8,897 (7,224; 9,717)
Protein (g/d)	77.5 (69.0; 94.4)
Carbohydrates (g/d)	227.9 (183.6; 269.9)
Fat (g/d)	82.5 (63.8; 94.7)
Fruits (g/d)**	152.4 (86.5; 246.8)
Vegetables (g/d)††	118.9 (95.7; 161.4)
Dietary fiber (g/d)	23.2 (20.0; 28.5)
Meat (g/d)	108.8 (69.7; 134.3)
Alcohol (g/d)	11.6 (1.7; 21.8)
Adenoma, n (%)	
Adenoma history	12 (15.0%)
Histopathology‡‡	
Tubular	33 (40.7%)
Tubulovillous	22 (27.2%)
Villous	26 (32.1%)
Location§§	
Rectum	14 (25.9%)
Distal colon	31 (57.4%)
Proximal colon	9 (16.7%)
Large (≥1 cm)	48 (64.0%)
Multiple adenomas	46 (59.0%)

* Because of missing values, not all numbers sum up to the column totals.

† No tissue obtained from 43, 67 samples did not contain (sufficient) polyp tissue, K-*ras* assessment unsuccessful for 14; see methods section.

‡ Primary school or lower vocational training only.

§ Individuals classified in the highest tertile of a continuous activity score (based on distribution among controls).

** Excluding juices and apple sauce.

†† Excluding potatoes.

‡‡ As assessed by the study pathologist for tissue that was available and obtained from the medical charts when the tumor block was not present or could not be judged, which was the case for most adenomas in which K-*ras* mutation status could not be determined. Histopathology of the examined polyp was not evaluated for 4 K-*ras*$^-$ adenomas and 17 adenomas in which K-*ras* mutation status could not be assessed.

§§ Recorded from medical charts, in which the location of the retrieved adenoma was not always possible to retrieve, in particular when multiple polyps were diagnosed.

Cases* K-*ras* mutation-negative adenomas (K-*ras*$^-$) *n*=453		Tissue unavailable or assessment unsuccessful† *n*=124		Controls* *n*=709	
58.7 ± 10.1		58.8 ± 10.1		51.5 ± 13.6	
203	(44.8%)	61	(49.2%)	437	(61.6%)
26.1 ± 3.8		26.7 ± 4.1		25.5 ± 4.1	
142	(35.0%)	45	(38.1%)	215	(33.0%)
310	(68.7%)	79	(64.2%)	389	(55.2%)
30	(20; 40)	30	(20; 39)	22.5	(13; 32)
124	(29.1%)	32	(27.8%)	228	(33.5%)
102	(22.5%)	30	(24.2%)	136	(19.2%)
117	(25.8%)	30	(24.2%)	207	(29.2%)
241	(54.2%)	49	(39.8%)	525	(77.0%)
8,336	(7,035; 10,177)	8,224	(6,984; 9,748)	8,112	(6,758; 9,935)
76.2	(64.3; 89.9)	78.4	(65.6; 91.0)	74.6	(62.2; 90.0)
216.1	(177.3; 265.0)	216.2	(179.5; 267.7)	223.1	(180.3; 267.7)
79.4	(62.4; 98.4)	79.8	(63.2; 94.7)	75.1	(58.4; 97.2)
125.9	(71.1; 248.9)	169.9	(74.0; 248.9)	124.7	(72.5; 243.9)
115.7	(90.4; 140.8)	122.0	(95.7; 150.4)	109.3	(83.7; 139.2)
23.7	(19.2; 28.3)	23.8	(19.4; 27.9)	22.6	(18.8; 27.2)
107.2	(70.1; 144.0)	106.8	(70.6; 136.7)	101.7	(63.1; 133.0)
8.7	(1.0; 24.1)	8.1	(0.3; 15.4)	4.2	(0.3; 15.4)
100	(22.1%)	37	(30.1%)	0	(0%)
369	(82.2%)	84	(78.5%)	-	
56	(12.5%)	18	(16.8%)	-	
24	(5.4%)	5	(4.7%)	-	
40	(13.8%)	16	(22.5%)	-	
205	(70.7%)	44	(62.0%)	-	
45	(15.5%)	11	(15.5%)	-	
160	(37.0%)	39	(34.2%)	-	
185	(42.0%)	53	(42.7%)	-	

Table 3-2. Case-case comparisons: odds ratios of K-*ras* mutation-positive (K-*ras*$^+$) versus -negative adenomas (K-*ras*$^-$) according to tertiles or categories of potential modifiable risk factors*,†

Characteristic	Cut-off points for classification of exposure		Number of cases, low/intermediate/high	
	Low	High	K-*ras*$^+$	K-*ras*$^-$
Foods				
Dairy products	≤238.2 g/d	>474.6 g/d	25/27/29	173/148/132
Total meat	≤74.6 g/d	>121.7 g/d	22/28/31	125/147/181
Red meat‡	≤38.2 g/d	>70.5 g/d	22/21/38	138/156/159
Processed meat	≤18.5 g/d	>39.9 g/d	27/31/23	144/145/164
Poultry	≤5.1 g/d	>13.7 g/d	26/30/25	128/165/160
Fish	≤4.28 g/d	>14.5 g/d	31/24/26	138/170/145
Fruits	≤105.5 g/d	>230.8 g/d	21/31/29	149/144/160
Vegetables	≤91.9 g/d	>128.5 g/d	18/29/33	119/162/167
Macronutrients				
Total energy	≤7,173.8 kJ/d	>9,227.9 kJ/d	20/28/33	127/162/164
Carbohydrates	≤212.3 g/d	>241.6 g/d	24/35/22	161/131/161
Total dietary fat	≤71.2 g/d	>84.5 g/d	29/19/33	108/138/207
Polyunsaturated fat	≤13.1 g/d	>16.5 g/d	32/22/27	119/150/184
Monounsaturated fat	≤26.9 g/d	>32.3 g/d	30/18/33	110/138/205
Saturated fat	≤29.6 g/d	>35.7 g/d	20/32/29	108/143/202
Protein	≤71.0 g/d	>81.6 g/d	21/23/37	133/141/179
Animal protein	≤44.0 g/d	>54.1 g/d	21/22/38	130/152/171
Micronutrients				
Dietary fiber	≤20.8 g/d	>24.8 g/d	23/23/35	128/138/187
Vitamin B2	≤1.35 mg/d	>1.75 mg/d	18/27/36	126/193/134
Vitamin B6	≤1.44 mg/d	>1.70 mg/d	22/26/33	109/147/197
Folate	≤170.8 µg/d	>202.2 µg/d	17/27/37	106/146/201
Vitamin B12	≤3.62 mg/d	>4.83 mg/d	17/29/35	127/136/190
Calcium	≤916.0 mg/d	>1,208.4 mg/d	18/31/32	145/171/137
Lifestyle factors				
Alcohol	<1 glass/wk	≥10 glasses/wk	20/26/34	131/153/170
Coffee	<3 cups/d	>5 cups/d	21/36/24	157/140/156
Tea	<1 cup/d	≥3 cups/d	29/25/27	197/138/118
Smoking**	Never smokers	Current smokers	31/32/17	141/177/133
Duration of smoking	Never	≥22.5 yr	31/20/30	141/115/197
Physical activity††	Low	High	25/24/24	178/124/124
Body mass index	≤23.5 kg/m^2	>26.7 kg/m^2	23/25/33	124/158/163
Regular NSAID use	<12 /yr	≥12 /yr	58/-/23	336/-/117

* Adjusted for sex, age, total energy. All nutrients, except alcohol, were adjusted for energy by the residual method.

† The case-control comparisons for factors for which the *P*-value for trend was < 0.25 and showed an odds ratio ≥1.4 or equivalently, ≤0.7 are presented in Table 3-3.

‡ Does not include organ meat.

§ Also adjusted for red meat. Without this factor, the estimates were a: 0.46 (0.24-0.90), b: 0.53 (0.26-1.10) and c: 0.13.

** Intermediate category consists of former smokers.

†† As assessed by a continuous score that was subdivided into tertiles.

Intermediate *vs.* low exposure OR (95% CI)	**High *vs.* low exposure OR (95% CI)**	P_{trend} †	**Characteristic** (repeated)
			Foods
1.24 (0.69-2.25)	1.52 (0.83-2.81)	0.17	Dairy products
1.15 (0.62-2.13)	1.10 (0.56-2.15)	0.78	Total meat
0.92 (0.48-1.77)	1.69 (0.92-3.11)	0.063	Red meat‡
1.16 (0.65-2.06)	0.78 (0.40-1.53)	0.41	Processed meat
0.93 (0.52-1.67)	0.81 (0.44-1.49)	0.51	Poultry
0.62 (0.35-1.11)	0.76 (0.42-1.35)	0.38	Fish
1.45 (0.79-2.68)	1.21 (0.65-2.26)	0.77	Fruits
1.17 (0.62-2.22)	1.27 (0.68-2.39)	0.48	Vegetables
			Macronutrients
1.20 (0.64-2.25)	1.57 (0.81-3.05)	0.28	Total energy
1.84 (1.03-3.31)	0.97 (0.48-1.97)	0.99	Carbohydrates
0.48 (0.24-0.94)	0.50 (0.23-1.06)	0.11	Total dietary fat
0.53 (0.29-0.99)	0.52 (0.27-0.98)	0.062	Polyunsaturated fat
0.40 (0.20-0.80) §a	0.42 (0.20-0.90) §b	0.044§c	Monounsaturated fat
1.17 (0.62-2.20)	0.73 (0.35-1.49)	0.29	Saturated fat
1.12 (0.58-2.15)	1.57 (0.80-3.11)	0.16	Protein
0.91 (0.48-1.76)	1.44 (0.78-2.65)	0.19	Animal protein
			Micronutrients
0.93 (0.50-1.74)	1.06 (0.58-1.92)	0.81	Dietary fiber
0.97 (0.51-1.87)	1.88 (0.99-3.55)	0.031	Vitamin B2
0.89 (0.47-1.69)	0.88 (0.45-1.73)	0.46	Vitamin B6
1.16 (0.60-2.26)	1.20 (0.63-2.32)	0.60	Folate
1.64 (0.85-3.16)	1.43 (0.74-2.76)	0.43	Vitamin B12
1.39 (0.74-2.62)	1.80 (0.96-3.38)	0.072	Calcium
			Lifestyle factors
1.12 (0.60-2.10)	1.38 (0.74-2.58)	0.29	Alcohol
1.95 (1.08-3.52)	1.26 (0.66-2.42)	0.46	Coffee
1.18 (0.66-2.13)	1.46 (0.81-2.63)	0.21	Tea
0.83 (0.48-1.45)	0.62 (0.32-1.20)	0.16	Smoking**
0.86 (0.46-1.61)	0.70 (0.40-1.23)	0.21	Duration of smoking
1.38 (0.75-2.54)	1.38 (0.75-2.57)	0.18	Physical activity††
0.87 (0.47-1.63)	1.10 (0.61-1.97)	0.66	Body mass index
-	1.11 (0.65-1.88)	-	Regular NSAID use

indications for differential associations with K-*ras*$^+$ and K-*ras*$^-$ adenomas could be detected for the other examined factors, including specific subgroups of fruits and vegetables (estimates not shown) and coffee consumption (median consumption: 4 cups/d, 25th percentile=3 cups/d, 75th percentile=6 cups/d, equal for K-*ras*$^+$ and K-*ras*$^-$ cases). When we compared regular consumption of coffee (≥2 glasses/d) with low/intermediate consumption (<2 glasses/d), also no difference in distribution among K-*ras*$^+$ and K-*ras*$^-$ adenoma patients could be detected (OR=1.41, 95% CI= 0.53-3.73).

Differences in risk of K-*ras*$^+$ and K-*ras*$^-$ adenomas were only statistically significant for intake of monounsaturated fat and vitamin B2. The case-control analyses presented in Table 3-3 illustrate why these differences in risk arose; the table also presents case-control analyses for other factors

Table 3-3. Case-control comparisons: odds ratios and 95% confidence intervals for K-*ras* mutation-positive adenomas (K-*ras*$^+$) and -negative adenomas (K-*ras*$^-$) according to selected modifiable risk factors*

Characteristic	*N*			Case-control comparisons	
	K-*ras*$^+$	K-*ras*$^-$	Controls	K-*ras*$^+$ *vs.* controls	K-*ras*$^-$ *vs.* controls
Foods					
Dairy products, excluding cheese and butter (g/d)					
≤238.2	25	173	237	1 (reference)	1 (reference)
>238.2-≤474.6	27	148	235	0.89 (0.50-1.60)	0.72 (0.53-0.97)
>474.6	29	132	237	0.94 (0.51-1.71)	0.61 (0.45-0.85)
				P_{trend}=0.85	P_{trend}=0.028
Red meat (g/d)†					
≤38.2	22	138	237	1 (reference)	1 (reference)
>38.2-≤70.5	21	156	235	1.05 (0.55-2.00)	1.14 (0.83-1.56)
>70.5	38	159	237	1.70 (0.94-3.09)	1.00 (0.73-1.39)
				P_{trend}=0.061	P_{trend}=0.98
Macronutrients					
Total dietary fat (g/d, energy-adjusted)					
≤71.2	29	108	237	1 (reference)	1 (reference)
>71.2 -≤84.5	19	138	235	0.55 (0.28-1.06)	1.15 (0.82-1.61)
>84.5	33	207	237	0.62 (0.30-1.32)	1.26 (0.84-1.88)
				P_{trend}=0.30	P_{trend}=0.29
Polyunsaturated fat (g/d, energy-adjusted)					
≤13.1	32	119	236	1 (reference)	1 (reference)
>13.1-≤16.5	22	150	236	0.62 (0.34-1.13)	1.16 (0.84-1.60)
>16.5	27	184	237	0.58 (0.31-1.10)	1.13 (0.80-1.59)
				P_{trend}=0.11	P_{trend}=0.56
Monounsaturated fat (g/d, energy-adjusted)‡					
≤26.8	30	110	237	1 (reference)	1 (reference)
>26.8-≤32.3	18	138	236	0.51 (0.25-1.00) ‡a	1.27 (0.90-1.78) ‡d
>32.3	33	205	236	0.67 (0.32-1.40) ‡b	1.57 (1.06-2.34) ‡e
				P_{trend}=0.37‡c	P_{trend}=0.029‡f
Protein (g/d, energy-adjusted)					
≤71.0	21	133	236	1 (reference)	1 (reference)
>71.0-≤81.6	23	141	236	0.85 (0.44-1.62)	0.76 (0.55-1.05)
>81.6	37	179	237	1.10 (0.56-2.15)	0.70 (0.49-1.00)
				P_{trend}=0.70	P_{trend}=0.054

Footnotes: see next page

Table 3-3. Continued

Characteristic	*N*			Case-control comparisons	
	K-*ras*$^+$	K-*ras*$^-$	Controls	K-*ras*$^+$ *vs.* controls	K-*ras*$^-$ *vs.* controls
Micronutrients					
Calcium (mg/d, energy-adjusted)					
≤ 916.0	18	145	237	1 (reference)	1 (reference)
>916.0-≤1,208.4	31	171	235	1.37 (0.73-2.57)	0.99 (0.73-1.34)
>1,208.4	32	137	237	1.27 (0.68-2.36)	0.71 (0.51-0.97)
				P_{trend} =0.50	P_{trend} =0.028
Vitamin B2 (mg/d, energy-adjusted)					
≤1.35	18	126	237	1 (reference)	1 (reference)
>1.35-≤1.75	27	193	235	1.10 (0.58-2.11)	1.13 (0.83-1.55)
>1.75	36	134	237	1.31 (0.71-2.44)	0.70 (0.50-0.97)
				P_{trend} =0.33	P_{trend} =0.020
Lifestyle factors					
Tea (cups/d)					
<1	29	197	290	1 (reference)	1 (reference)
1-<3	25	138	208	1.02 (0.57-1.82)	0.86 (0.64-1.16)
≥3	27	118	211	1.08 (0.60-1.94)	0.74 (0.54-1.02)
				P_{trend} =0.79	P_{trend} =0.065
Smoking					
Never	31	141	316	1 (reference)	1 (reference)
Former	32	177	216	1.20 (0.70-2.08)	1.45 (1.08-1.96)
Current	17	133	173	1.27 (0.66-2.42)	2.04 (1.47-2.83)
				P_{trend} =0.44	P_{trend} =<0.0001
Duration of smoking					
Never	31	141	316	1 (reference)	1 (reference)
0-<22.5 yr	20	115	203	1.30 (0.70-2.42)	1.52 (1.10-2.11)
≥ 22.5 yr	30	197	190	1.25 (0.71-2.18)	1.79 (1.33-2.42)
				P_{trend}=0.45	P_{trend}=0.0002

* Only factors for which the *P*-value for trend was < 0.25 and showed a odds ratio ≥1.4 or ≤0.7 in the case-case analyses (see Table 3-2) are presented in this table. Odds ratios are adjusted for sex, age and total energy.

† Does not include organ meat.

‡ Also adjusted for red meat. Without this factor, the estimates were a: 0.58 (0.30-1.12), b: 0.81 (0.40-1.66), c: 0.68, d: 1.26 (0.90-1.76), e: 1.53 (1.04-2.25) and f: 0.034.

that were possibly differently associated with risk of K-*ras*$^+$ and K-*ras*$^-$ adenomas. Consumption of red meat was positively associated with K-*ras*$^+$ adenomas, but not with K-*ras*$^-$ adenomas. The intake of total dietary, polyunsaturated and monounsaturated fat tended to be inversely associated with risk of K-*ras*$^+$ adenomas. However, especially intake of monounsaturated fat was positively associated with risk of K-*ras*$^-$ adenomas. Intake of vitamin B2 was somewhat positively, but not statistically significantly, associated with K-*ras*$^+$ adenomas. However, vitamin B2 was inversely associated with K-*ras*$^-$ adenomas. Similar patterns were observed for dairy products, calcium and protein. No association between tea consumption and K-*ras*$^+$ adenomas could be detected, but tea was inversely, although not statistically significantly, associated with risk of K-*ras*$^-$ adenomas.

Smoking, especially its duration, was relatively strongly positively associated with risk of K-*ras*$^-$ adenomas, but only weakly so with K-*ras*$^+$ adenomas.

The aforementioned findings were supported by the sensitivity analyses. The analyses in which gender-specific tertiles were used and with restriction to those who underwent full endoscopy suggested that high physical activity tended to decrease the risk of K-*ras*$^-$ adenomas, but somewhat increase the risk of K-*ras*$^+$ adenomas, albeit this was not statistically significant.

Strata according to specific mutations were very small. However, remarkably, none or a very low number of patients who developed an adenoma harboring a codon 13 mutation were classified in the lowest tertile of consumption or intake of fruit and vegetables, folate and less strikingly, of β-carotene and vitamin C (results not shown). It was also remarkable that a higher proportion of patients with codon 12 transversion mutations were classified in the highest tertile of fruits, green leafy vegetables and folate compared with patients with codon 12 transition mutations ($P<0.05$; results not shown).

Discussion

We observed that intake of vitamin B2 and monounsaturated fat were differently associated with risk of K-*ras* mutation-positive and -negative adenomas. Similar, not statistically significant, differences in risk were observed for red meat, dairy products, total dietary and polyunsaturated fat, protein, calcium, tea and (duration of) smoking. These differences in risk mainly arose because inverse associations with vitamin B2, dairy products, calcium and possibly protein and tea were confined to K-*ras* mutation-negative adenomas. Intake of total dietary, monounsaturated and polyunsaturated fat tended to be inversely associated with risk of K-*ras* mutation-positive adenomas, but not, or somewhat positively, with risk of K-*ras* mutation-negative adenomas. Smoking and its duration were more clearly positively associated with K-*ras* mutation-negative than with mutation-positive adenomas. A positive association with red meat was distinctly observed for K-*ras* mutation-positive adenomas. Additionally, we found hints that the distribution of specific K-*ras* mutations may vary according to intake of folate and consumption of fruits and vegetables.

In the present analyses, we mainly focused on the comparisons of the distributions of risk factors among K-*ras* mutation-positive and -negative cases. Such comparisons are unlikely to be affected by information bias; reporting of diet and lifestyle is implausible to be related to mutational status. We cannot completely rule out selection bias that occurs when factors determining tumor block availability or successful K-*ras* assessment are also associated with exposure factors of interest and K-*ras* mutational status[33, 34]. However, its impact is likely to be minor; the distribution of risk factors and adenoma characteristics of the tumors in which mutational status could not be assessed appeared to be largely comparable to the weighted average of the distributions among K-*ras* mutation-positive and -negative adenoma cases. Risk of selection and information bias remains a weak point regarding the case-control comparisons, although it is reassuring that excluding participants who indicated to have changed their diet did not affect the results.

In our study, 15% of the adenomas harbored a K-*ras* mutation. This percentage is in the same range of magnitude as the 18% observed in a cross-sectional study on K-*ras* mutations

in colorectal adenomas[20]. Reported mutation rates are generally higher, *e.g.* a compilation of 12 reports on 646 adenomas, which did not include the cross-sectional study, reported a rate of 37%[4]. The latter estimate may be higher because it was partly based upon adenomas diagnosed in cancer patients, whereas the cross-sectional study and ours were based upon a cancer-free endoscopy population. Alternatively, we may have underestimated the mutation rate because we studied one adenoma per patient and could not always study the entire adenoma, whereas mutational patterns may differ between[35] and within adenomas of the same patient[36]. Besides, we did not evaluate the presence of K-*ras* codon 61 mutations, but these are rare[37]. We expect that bias due to the use of two different methods of K-*ras* assessment is small, because exclusion of samples analyzed with MASA-PCR did not change the results. More substantial errors could have arisen through the multiple comparisons made and the relatively small sample size of K-*ras* mutation-positive cases. Consequently, some detected differences may be due to chance, and we may not have had sufficient power to detect existing associations.

Only one study evaluated associations between presumed risk factors and K-*ras* mutation-positive and -negative adenomas before[20]. In that cross-sectional study, K-*ras* mutation-positive adenoma cases (n=120) were less likely to have a high folate intake, and possibly less likely to participate in vigorous aerobic exercise than K-*ras* mutation-negative adenoma cases (n=558)[20]. Differences in study population or design, and the relatively small sample size of K-*ras* mutation-positive samples in both studies may partly explain why we observed associations with different factors. Levels of intake may not have been high enough to allow detection of an association with calcium in the cross-sectional study[20] and with folate in our study; indeed, the cross-sectional study observed an association with total but not with dietary folate[20]. Even so, it is unlikely that the aforementioned factors account for all discrepancies, as epidemiological studies on environmental risk factors and risk of K-*ras* mutation-positive and -negative colorectal cancers also yielded inconsistent results[26-29, 38-43].

The inverse association with calcium, for instance, was restricted to K-*ras* mutation-negative adenomas in our study, which is in keeping with a study in which calcium-supplemented rats did not develop K-*ras* mutation-negative tumors[16]. However, it was restricted to mutation-positive colorectal cancers in another study[42]. Two colorectal cancer studies[26, 29] could not detect any such differential association, although in one of them codon 12 mutations were more likely to be observed than codon 13 mutations when calcium intake was high[26]. Part of these differences may be attributed to differences in exposure to compounds associated with the intake of calcium or dairy products, such as vitamin B2 and organochlorines[44]. Indeed, the association with vitamin B2, a co-factor in folate metabolism[30], tended to be stronger than the association with calcium in our study, and organochlorines, that none of the aforementioned studies assessed, were associated with K-*ras* mutations in other tumors[44].

Inconsistent findings were also observed regarding other factors, including consumption of red meat[26, 28, 29, 38, 43], and intake of fat[26, 29, 40, 42] and folate[29, 39, 42]. Other factors have been studied less frequently[27-29, 41, 42, 45], and coffee and tea, whose many constituents exert strong effects upon a wide variety of physiological, cellular and molecular systems[46], have not been studied in relation to K-*ras* mutations in adenomas before. Patients with K-*ras* mutation-positive exocrine pancreatic

cancer were more likely to be in the upper category of coffee consumption than patients with K-*ras* mutation-negative exocrine pancreatic cancer[46], but we could not confirm this regarding colorectal adenomas. Whereas in *vivo* and in *vitro* studies indicate that specific exposures are associated with specific K-*ras* mutations[11-15, 17, 18, 44], the general picture regarding associations between exposure and specific K-*ras* mutations obtained from observational studies is also not clear[20, 26, 29, 38-40, 45].

Besides small sample sizes, the inconsistent findings of epidemiological studies on K-*ras* mutations in colorectal tumors may be partly due to the use of questionnaires to assess exposure. Such an assessment may not be sensitive enough to detect associations; assessment of specific agents hypothesized to affect K-*ras* mutations, including *N*-nitroso compounds[11, 12, 14], polycyclic aromatic hydrocarbons[13], heterocyclic amines[17], *n*-3 polyunsaturated fatty acids[18] and organochlorine compounds[44] in body fluids or tissue samples may be more appropriate. Besides, associations may be diluted because K-*ras* mutations by themselves may not be the best to distinguish unique etiological pathways[47], and some compounds could also be related to mutations in oncogenes, tumor suppressor genes or DNA repair genes in K-*ras* mutation-negative adenomas. On top of this, K-*ras* mutations occur more frequently in strong dysplastic areas[48], which is also reflected in the heterogeneous K-*ras* mutation pattern within an adenoma[36] and the lack of concordance between K-*ras* mutational status in multiple adenomas of the same patient[35]. Thus, associations of diet and lifestyle with K-*ras* mutational status may be hard to confirm in epidemiological studies.

In conclusion, our findings add to the body of evidence that risk factors may play a different role in the etiology of K-*ras* mutation-positive and mutation-negative colorectal adenomas. Nevertheless, findings on dietary and lifestyle factors and K-*ras* mutations in colorectal tumors remain highly inconsistent in the literature.

Acknowledgements

We are grateful to the people who kindly participated in this study. We would like to thank Maria van Vugt, Elly Monster, Drs. Edine Tiemersma, Dorien Voskuil, Brenda Diergaarde, Maureen van den Donk and Mariken Tijhuis for their roles in the conduct of the study; Hanneke Braam for the laboratory analyses; Dr. M.C. Ocké (National Institute of Public Health and the Environment, Bilthoven, the Netherlands) for calculation of food and nutrient intake and Prof. Dr. J.H.J.M. van Krieken for pathological evaluation of the tumor samples. The endoscopy staff and gastroenterologists of the following hospitals assisted with recruitment of participants and the departments of pathology of these hospitals provided adenoma tissue: Slingeland Hospital, Doetinchem; Gelderse Vallei Hospital, Ede; Radboud University Nijmegen Medical Centre, Nijmegen; Anthonius Hospital, Nieuwegein; Meander Medical Center, Amersfoort; Rijnstate Hospital, Arnhem; Rivierenland Hospital, Tiel; Slotervaart Hospital, Amsterdam; Jeroen Bosch Hospital, Den Bosch; Canisius Wilhelmina Hospital, Nijmegen. The Netherlands Organisation for Health Research and Development supported this work (grant number 980-10-026).

References

1. Giovannucci E. Modifiable risk factors for colon cancer. *Gastroenterol Clin North Am.* 2002;31(4):925-943.
2. World Cancer Research Fund, American Institute for Cancer Research. *Food, nutrition and the prevention of cancer : a global perspective.* Washington, D.C.: American Institute for Cancer Research; 1997.
3. Porta M, Ayude D, Alguacil J, Jariod M. Exploring environmental causes of altered ras effects: fragmentation plus integration? *Mol Carcinog.* 2003;36(2):45-52.
4. Zauber P, Sabbath-Solitare M, Marotta SP, Bishop DT. Molecular changes in the Ki-ras and APC genes in primary colorectal carcinoma and synchronous metastases compared with the findings in accompanying adenomas. *Mol Pathol.* 2003;56(3):137-140.
5. Andreyev HJ, Norman AR, Cunningham D, Oates JR, Clarke PA. Kirsten ras mutations in patients with colorectal cancer: the multicenter "RASCAL" study. *J Natl Cancer Inst.* 1998;90(9):675-684.
6. Brink M, De Goeij AFPM, Weijenberg MP, et al. K-ras oncogene mutations in sporadic colorectal cancer in The Netherlands Cohort Study. *Carcinogenesis.* 2003;24(4):703-710.
7. Egan SE, Weinberg RA. The pathway to signal achievement. *Nature.* 1993;365(6449):781-783.
8. Maltzman T, Knoll K, Martinez ME, et al. Ki-ras proto-oncogene mutations in sporadic colorectal adenomas: relationship to histologic and clinical characteristics. *Gastroenterology.* 2001;121(2):302-309.
9. Andreyev HJ, Norman AR, Cunningham D, et al. Kirsten ras mutations in patients with colorectal cancer: the 'RASCAL II' study. *Br J Cancer.* 2001;85(5):692-696.
10. Johnson V, Lipton L, Cummings C, et al. Analysis of somatic molecular changes, clinicopathological features, family history and germline mutations in colorectal cancer families: evidence for efficient diagnosis of HNPCC and for distinct groups of non-HNPCC families. *J Med Genet.* 2005.
11. Cross AJ, Sinha R. Meat-related mutagens/carcinogens in the etiology of colorectal cancer. *Environ Mol Mutagen.* 2004;44(1):44-55.
12. Bingham SA, Pignatelli B, Pollock JR, et al. Does increased endogenous formation of N-nitroso compounds in the human colon explain the association between red meat and colon cancer? *Carcinogenesis.* 1996;17(3):515-523.
13. Tachino N, Hayashi R, Liew C, Bailey G, Dashwood R. Evidence for ras gene mutation in 2-amino-3-methylimidazo[4,5-f]quinoline-induced colonic aberrant crypts in the rat. *Mol Carcinog.* 1995;12(4):187-192.
14. Ziegel R, Shallop A, Jones R, Tretyakova N. K-ras gene sequence effects on the formation of 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK)-DNA adducts. *Chem Res Toxicol.* 2003;16(4):541-550.
15. Bos JL. ras oncogenes in human cancer: a review. *Cancer Res.* 1989;49(17):4682-4689.
16. Llor X, Jacoby RF, Teng BB, Davidson NO, Sitrin MD, Brasitus TA. K-ras mutations in 1,2-dimethylhydrazine-induced colonic tumors: effects of supplemental dietary calcium and vitamin D deficiency. *Cancer Res.* 1991;51(16):4305-4309.

17. Kakiuchi H, Ushijima T, Ochiai M, et al. Rare frequency of activation of the Ki-ras gene in rat colon tumors induced by heterocyclic amines: possible alternative mechanisms of human colon carcinogenesis. *Mol Carcinog.* 1993;8(1): 44-48.

18. Davidson LA, Lupton JR, Jiang YH, Chapkin RS. Carcinogen and dietary lipid regulate ras expression and localization in rat colon without affecting farnesylation kinetics. *Carcinogenesis.* 1999;20(5):785-791.

19. Singh J, Reddy BS. Molecular markers in chemoprevention of colon cancer. Inhibition of expression of ras-p21 and p53 by sulindac during azoxymethane-induced colon carcinogenesis. *Ann N Y Acad Sci.* 1995;768:205-209.

20. Martinez ME, Maltzman T, Marshall JR, et al. Risk factors for Ki-ras protooncogene mutation in sporadic colorectal adenomas. *Cancer Res.* 1999;59(20):5181-5185.

21. Ocké MC, Bueno de Mesquita HB, Goddijn HE, et al. The Dutch EPIC food frequency questionnaire. I. Description of the questionnaire, and relative validity and reproducibility for food groups. *Int J Epidemiol.* 1997;26 Suppl 1:S37-S48.

22. Ocké MC, Bueno-de-Mesquita HB, Pols MA, Smit HA, Van Staveren WA, Kromhout D. The Dutch EPIC food frequency questionnaire. II. Relative validity and reproducibility for nutrients. *Int J Epidemiol.* 1997;26 Suppl 1: S49-58.

23. Voskuil DW, Kampman E, Grubben MJAL, et al. Meat consumption and meat preparation in relation to colorectal adenomas among sporadic and HNPCC family patients in The Netherlands. *Eur J Cancer.* 2002;38(17):2300-2308.

24. Stichting Nederlands Voedingsstoffenbestand. *Dutch Food Composition Table 1986-1987; Nederlands voedingsstoffenbestand 1986-1987*: The Hague : Voorlichtingsbureau voor de Voeding; 1996.

25. Augustsson K, Skog K, Jagerstad M, Steineck G. Assessment of the human exposure to heterocyclic amines. *Carcinogenesis.* 1997;18(10):1931-1935.

26. Kampman E, Voskuil DW, Van Kraats AA, et al. Animal products and K-ras codon 12 and 13 mutations in colon carcinomas. *Carcinogenesis.* 2000;21(2):307-309.

27. Slattery ML, Anderson K, Curtin K, et al. Lifestyle factors and Ki-ras mutations in colon cancer tumors. *Mutat Res.* 2001;483(1-2):73-81.

28. Slattery ML, Curtin K, Ma K, Schaffer D, Potter J, Samowitz W. GSTM-1 and NAT2 and genetic alterations in colon tumors. *Cancer Causes Control.* 2002;13(6):527-534.

29. Slattery ML, Curtin K, Anderson K, et al. Associations between dietary intake and Ki-ras mutations in colon tumors: a population-based study. *Cancer Res.* 2000;60(24):6935-6941.

30. Van den Donk M, Buijsse B, Van den Berg SW, et al. Dietary Intake of Folate and Riboflavin, MTHFR C677T Genotype, and Colorectal Adenoma Risk: A Dutch Case-Control Study. *Cancer Epidemiol Biomarkers Prev.* 2005;14(6): 1562-1566.

31. Diergaarde B, Tiemersma EW, Braam H, et al. Dietary factors and truncating APC mutations in sporadic colorectal adenomas. *Int J Cancer.* 2005;113(1):126-132.

32. Willett W. *Nutritional Epidemiology.* 2nd ed. New York, NY: Oxford University Press; 1998.

33. Hoppin JA, Tolbert PE, Taylor JA, Schroeder JC, Holly EA. Potential for selection bias with tumor tissue retrieval in molecular epidemiology studies. *Ann Epidemiol.* 2002;12(1):1-6.

34. Schroeder JC, Weinberg CR. Use of missing-data methods to correct bias and improve precision in case-control studies in which cases are subtyped but subtype information is incomplete. *Am J Epidemiol.* 2001;154(10):954-962.

35. Rashid A, Zahurak M, Goodman SN, Hamilton SR. Genetic epidemiology of mutated K-ras proto-oncogene, altered suppressor genes, and microsatellite instability in colorectal adenomas. *Gut.* 1999;44(6):826-833.

36. Saraga E, Bautista D, Dorta G, et al. Genetic heterogeneity in sporadic colorectal adenomas. *J Pathol.* 1997;181(3): 281-286.

37. Toyooka S, Tsukuda K, Ouchida M, et al. Detection of codon 61 point mutations of the K-ras gene in lung and colorectal cancers by enriched PCR. *Oncol Rep.* 2003;10(5):1455-1459.

38. Brink M, Weijenberg MP, De Goeij AFPM, et al. Meat consumption and K-ras mutations in sporadic colon and rectal cancer in The Netherlands Cohort Study. *Br J Cancer.* 2005;92(7):1310-1320.

39. Brink M, Weijenberg MP, De Goeij AFPM, et al. Dietary folate intake and k-ras mutations in sporadic colon and rectal cancer in The Netherlands Cohort Study. *Int J Cancer.* 2005;114(5):824-830.

40. Brink M, Weijenberg MP, De Goeij AF, et al. Fat and K-ras mutations in sporadic colorectal cancer in The Netherlands Cohort Study. *Carcinogenesis.* 2004;25(9):1619-1628.

41. Moreno V, Guino E, Bosch FX, et al. Diet and K-ras mutations in colorectal cancer. *IARC Sci Publ.* 2002;156: 501-502.

42. Bautista D, Obrador A, Moreno V, et al. Ki-ras mutation modifies the protective effect of dietary monounsaturated fat and calcium on sporadic colorectal cancer. *Cancer Epidemiol Biomarkers Prev.* 1997;6(1):57-61.

43. O'Brien H, Matthew JA, Gee JM, et al. K-ras mutations, rectal crypt cells proliferation, and meat consumption in patients with left-sided colorectal carcinoma. *Eur J Cancer Prev.* 2000;9(1):41-47.

44. Howsam M, Grimalt JO, Guino E, et al. Organochlorine exposure and colorectal cancer risk. *Environ Health Perspect*. 2004;112(15):1460-1466.
45. Diergaarde B, Vrieling A, Van Kraats AA, Van Muijen GNP, Kok FJ, Kampman E. Cigarette smoking and genetic alterations in sporadic colon carcinomas. *Carcinogenesis*. 2003;24(3):565-571.
46. Porta M, Malats N, Alguacil J, et al. Coffee, pancreatic cancer, and K-ras mutations: updating the research agenda. *J Epidemiol Community Health*. 2000;54(9):656-659.
47. Hermsen M, Postma C, Baak J, et al. Colorectal adenoma to carcinoma progression follows multiple pathways of chromosomal instability. *Gastroenterology*. 2002;123(4):1109-1119.
48. Ohnishi T, Tomita N, Monden T, et al. A detailed analysis of the role of K-ras gene mutation in the progression of colorectal adenoma. *Br J Cancer*. 1997;75(3):341-347.

Abstract

Background: Adenomas ≥1 cm or with a villous histology are postulated to have the highest risk of progression to colorectal cancer. Risk factors for advanced and non-advanced adenomas might differ from each other as these might reflect different stages in the adenoma-carcinoma sequence.

Methods: We examined whether previously identified potential risk factors for single colorectal adenomas and cancer are differentially associated with advanced (≥1 cm, report of a villous structure or carcinoma in situ) and non-advanced colorectal adenomas using data on 26,769 U.S. men participating in the Health Professionals Follow-Up Study who underwent an endoscopy. During 17 years of follow-up (1986 through 2002) 4,075 histologically confirmed cases with a diagnosis of colorectal adenoma were ascertained, 1,246 of which were classified as advanced and 1,661 as non-advanced. Medical and lifestyle information was collected biennially, and dietary information was collected every four years using questionnaires. Multinomial logistic regression was used to compare the distribution of risk factors among advanced adenoma patients, non-advanced adenoma patients and men not known with adenomas.

Results: Most risk factors were similarly associated with advanced and non-advanced colorectal adenomas. Smoking was more strongly associated with risk of advanced adenomas (current *vs.* never: odds ratio (OR)=1.59, 95% confidence interval (CI)=1.56-2.01, P_{trend}≤0.0001) than with risk of non-advanced adenomas (OR=1.14, 95% CI=0.91-1.43, P_{trend}≤0.0079). A high glycemic index was inversely associated with risk of advanced adenomas (highest versus lowest quintile: OR=0.78, 95% CI=0.64-0.94, P_{trend}=0.019) but not with non-advanced adenomas (OR=1.04, 95% CI=0.88-1.23, P_{trend}=0.94). Not statistically significant indications for different associations with advanced and non-advanced adenomas were also found for physical activity, use of multivitamins and intake of starch. Height, aspirin and possibly fish may also be differently associated with these outcomes, but only regarding specific adenoma subsites.

Conclusion: Associations between smoking, glycemic index and adenomas may depend on the stage of adenoma development.

Diet, Lifestyle and Advanced and Non-Advanced Adenomas: A Prospective Study among Men

4

Petra A. Wark
Kana Wu
Pieter van 't Veer
Charles S. Fuchs
Walter C. Willett
Edward L. Giovannucci

Introduction

The development of colorectal cancer extends over many years[1] and is thought to occur mainly via the adenoma-carcinoma sequence. Adenomas tend to recur in polypectomized adenoma patients[2-7], and these patients are at increased risk of colorectal cancer[8-11]. However, only a subset of colorectal adenomas eventually progress to cancer[2, 12]. Patients having adenomas with a villous component or large (≥ 1 cm) adenomas[2-7, 9-11] have a greater risk of colorectal cancer than patients with other types of adenomas.

Over the years, a set of modifiable risk factors for colorectal cancer has been identified, although results are not entirely consistent[13]. These risk factors may influence adenoma formation, but they may also enhance the transition of adenomas to more malignant lesions. Smoking[14], consumption of alcohol[15, 16] or fruits[17], intake of carbohydrates[17], use of non-steroidal anti-inflammatory drugs (NSAIDs)[18] and high body fatness[19] have been associated with adenoma growth or cell proliferation. The increased cell proliferation that has been observed across the adenoma-carcinoma sequence[20] might also be associated with calcium, omega-3-fatty acids from fish, red meat, vegetables, vitamin/mineral supplements, β-carotene, starch and sucrose (ref 21, 22, and references therein), although a trial on adenoma growth did not observe an effect of calcium and antioxidants[23]. A stronger association between serum levels of Insulin-Growth Factor (IGF)-I and IGF-I/insulin-like growth factor binding protein 3 (IGFBP3) ratio and advanced adenomas than between these levels and non-advanced adenomas[24] also suggests that the effect of risk factors may vary across adenoma stage, as the Insulin-Growth Factor (IGF) axis has been associated with factors like obesity and physical activity[25], several nutritional parameters[26, 27] and height as a marker of pre-adult insulin and IGF-I bioactivity[28, 29]. Various modifiable risk factors were indeed differently associated with risk of advanced (or large) and non-advanced (or small) adenomas[30-33], although this was not consistent across studies.

In this study, we further evaluate whether associations between presumed modifiable risk factors differ for advanced and non-advanced colorectal adenomas.

Material and Methods

Study Population

The current study comprises a subset of men of the Health Professionals Follow-Up Study (HPFS): an ongoing prospective study among 51,529 male US health professionals who responded to a mailed questionnaire in 1986 when they were between 40 and 75 years old. The Human Subjects Committee of the Harvard School of Public Health approved the HPFS.

At enrollment in 1986, and every two years thereafter, the cohort members were requested to fill out follow-up questionnaires to update information on various risk factors and to identify newly professionally diagnosed cases of various diseases. In 1986, 1990, 1994, 1998 and 2002, participants also completed a semi-quantitative food-frequency questionnaire. Questionnaires were mailed up to four times to non-responders. At the time of the 2002 questionnaire, 37,433 men were alive and still participating.

Only men who completed the 1986 questionnaire and who underwent colonoscopy or sigmoidoscopy between 1986 and 2002 were eligible for the present analyses. Participants who

reported to have had cancer (except non-melanoma skin cancer), ulcerative colitis or Crohn's disease before 1986 were excluded. We also excluded cohort members who reported implausible caloric intakes, *i.e.* <800 kcal/d or >4,200 kcal/d) as well as those who left 70 more of the items blank at the food-frequency questionnaire and 49 people with missing information for BMI or physical activity. This resulted in a base population of 26,769 men.

Case Ascertainment

For each man who reported to have had an adenoma on a follow-up questionnaire for the first time, we asked for permission to request and review relevant medical records. A study investigator, who was blinded to exposure status, reviewed endoscopy and pathology reports. Only when the self-reported diagnosis was confirmed by a histopathological report, case status was assigned. The self-report was reliable; a review of the medical records obtained from a random sample of 200 patients who reported a negative endoscopic result confirmed the absence of adenomas in all cases.

Adenomas in the cecum, ascending colon, hepatic flexure, transverse colon or splenic flexure were classified as being in the proximal colon. Adenomas in the descending or sigmoid colon were classified as being in the distal colon, and adenomas in the rectum or at the rectosigmoid junction were classified as rectal.

For diagnoses up to 1990, the number of adenoma, the size of the largest adenoma as assessed endoscopically and pathologically, and location of the most proximally located adenoma were recorded. For later diagnoses, number and size of the largest adenoma according to endoscopy and pathology reports were recorded for distal, proximal and rectal adenomas separately. In all years, only the most severe histological subtype of adenoma (in ascending order: tubular, tubulovillous, villous, carcinoma in situ) was registered.

Between 1986 and 2002, 4,075 participants were diagnosed with adenomas. We used adenoma size according to the endoscopy report, but if such information was lacking, data from the pathology report was used.

Men having at least one adenoma ≥1 cm, with a (tubulo)villous structure or carcinoma in situ were classified as having advanced adenomas (n=1,246); men having only adenomas that were smaller than 1 cm and with no mention of a villous structure were classified as having non-advanced adenomas (n=1,661). Adenomas of 1,168 men could not be classified due to missing information on size combined with either a tubular structure or no available information on histology (see also Table 4-1).

Exposure Assessment

The semi-quantitative food-frequency questionnaire covered more than 90% of the major nutrient intake of participants, and inquired after vitamin and mineral supplements[34]. Foods for which specific hypotheses of interest existed were also included. The questionnaire included about 130 items, each with a specified commonly used unit or portion size, and an open-ended section for unlisted foods. Participants were asked to indicate how often they consumed a specific food on average during the past year; they could tick one of the nine multiple choice responses,

which ranged from never or less than once per month to six or more times per day. Individual nutrient values were computed according to the US Department of Agriculture publications[35], supplemented with other data. The mean correlation coefficients between intakes determined by two one-week diet records and the dietary questionnaire were 0.65 for nutrients and 0.63 for specific foods, after adjusting for week-to-week variation in the diet records[34, 36]. Intraclass correlation coefficients for nutrient intakes assessed by questionnaires one year a part ranged from 0.47 for vitamin E without supplements to 0.80 for vitamin C with supplements[34].

We used values from published estimates[37], or from direct testing of food items at the Nutrition Center of the University of Toronto (D. Jenkins) to obtain glycemic index (GI) for individual foods and mixed meals, which reflects the postprandial blood glucose response compared with a reference food. We computed dietary glycemic load (GL) by multiplying the carbohydrate content of each food by its GI, after which we added the values from all foods. Each unit of dietary GL represents the glycemic equivalent of 1 g carbohydrate from white bread. As an indicator for overall quality of carbohydrate intake for each cohort member, we calculated the overall dietary GI by dividing GL by the amount of carbohydrate.

We also assessed (1) alcohol consumption calculated from daily servings of beer, white wine, red wine and liquor, and whether participants were former alcohol consumers[38]; (2) physical activity based on questions that measured the average weekly time spent at specified activities during the past year[39], expressed as the intensity of the activity in metabolic equivalent task hours (MET-h), *i.e.* the ratio of the metabolic rate during activity to the resting metabolic rate[40] per week; (3) body mass index, calculated as weight in kilograms divided by the square of height in meters[41]; (4) smoking status, the average number of cigarettes smoked per day and details on age or year when participants started or quit smoking, (5) family history of colorectal cancer, which was restricted to parents in 1986 and to parents and siblings in later years; and (6) use of aspirin that exceeded twice a week.

Statistical Analysis

We selected postulated modifiable risk factors for colorectal neoplasm, hypothesized risk factors that were previously associated with colorectal neoplasm in our study population before and factors that were associated with adenoma growth or cell proliferation in human studies before (listed in Table 4-3). Nutrients, except alcohol, were adjusted for total energy intake using the residual method[42]. To best represent long-term exposure and reduce within-person variation[43], we calculated the average of food and nutrient variables from all questionnaires up to diagnosis, most recent endoscopy or 1998, whichever occurred first, and we used these values in our analyses. Height was assessed at baseline, and the other non-dietary exposures were updated every two years. Participants were grouped into quartiles, tertiles or categories according to the exposure of interest.

Distributions of risk factors among patients having an advanced adenoma and patients having a non-advanced adenoma were compared with each other to evaluate etiological heterogeneity (case-case analyses). If the case-case analyses suggested potential differences, *i.e.* either by showing a *P*-value ≤ 0.10 or odds ratio (OR)≥ 1.4 or ≤ 0.7, we compared both categories of patients

with the people who did not report any adenomas to obtain information on the direction of the associations using a multinomial logistic regression model[44].

The main models included total energy intake, age, history of endoscopy prior to study entry, routine screening versus other indications for any endoscopy, aspirin use, use of multivitamins, smoking, consumption of red meat, alcohol, intake of folate, calcium, BMI, physical activity, family history of colorectal cancer, and the risk factor of interest. The case-case analyses were also adjusted for adenoma location (distal colon, proximal colon or rectal adenoma). As none of the other studied factors influenced any of the risk estimates by more than 10%, we did not add additional factors to the models. Tests for trend across categories of exposure were conducted by treating the median values of the exposure categories among the men who did not report adenomas as ordinal variables in the regression models.

Whether the association between exposure and advanced versus non-advanced adenomas depended on the location of the adenoma was explored in the group of men with adenoma at only one location. We did so by fitting a model containing two indicator variables for location, the continuous exposure variable that was also used in the test for trend and two cross-product terms of one of the indicator variables for location and the continuous exposure variable. The *P*-value for interaction was obtained by comparing a model with and without the two cross-product terms using a likelihood ratio test. Subsequently, we fitted three separate multinomial logistic regression models according to adenoma location for those exposures for which the case-case *P*-value of interaction was ≤ 0.10; these models included the men who did not report an adenoma.

As sensitivity analyses, we evaluated whether the associations depended on the presence of synchronous adenomas, year of most recent endoscopy and age at most recent endoscopy in the full dataset, using similar methods.

All reported *P*-values are two-sided. *P*-values < 0.05 were considered statistically significant. The analyses were performed using SAS version 9.1 (SAS Institute, Inc., Cary, NC).

Results

Histopathological Characteristics and Size of Adenomas

Table 4-1 presents the characteristics of the adenomas in the study population. Whereas tubular adenomas were more likely to be small, patients with tubulo(villous) adenomas and carcinoma in situ were more likely to have a large adenoma. Patients with villous adenomas or carcinoma in situ were 1.28 times more likely to have a large adenoma than were patients with tubulovillous adenoma, but this difference was not statistically significant (95% confidence interval (CI)=0.86-1.92).

We cannot be sure that the histopathological characteristics and sizes in Table 4-1 correspond to the same adenoma because only the worst characteristic of all adenomas was recorded, *i.e.* men with a small villous adenoma could also have a large tubular adenoma and they would end up in the category "large and villous" in Table 4-1. Therefore we repeated these analyses in men who had only one adenoma at index endoscopy. Doing so resulted in stronger associations between size and histopathological characteristics because men with multiple adenomas were more likely to have advanced adenomas (odds ratio (OR)=1.96, 95% CI=1.67-2.30). In the group of men

Table 4-1. Histopathology and size of the adenomas in the Health Professionals' Follow-Up Study*

Most severe histopathology of all diagnosed adenoma	All *n* (%)	Size *n* (%)			Large *vs.* small adenomas Odds ratio (95% confidence interval)
		Large (≥1cm)	Small (<1cm)	Unknown	
Total	4,075 (100%)	985 (24.2%)	1,846 (45.3%)	1,244 (30.5%†)	-
Tubular	1,731 (42.5%†)	360 (36.6%†)	**1,153 (62.5%†)**	*218 (17.5%†)*	1 (reference)
Tubulovillous	537 (13.2%†)	334 (33.9%†)	144 (7.8%†)	59 (4.7%†)	7.47 (5.93-9.40) ‡a
Villous or carcinoma in situ§	180 (4.4%†)	122 (12.4%†)	41 (2.2%†)	17 (1.4%†)	9.53 (6.55-13.8) ‡b
Unknown	1,627 (39.9%†)	169 (17.2%†)	**508 (27.5%†)**	*950 (76.4%†)*	-

* Darker grey shading indicates advanced adenomas (*n*=1,246), bold fond indicates non-advanced adenomas (*n*=1,661) and italic fond indicates adenomas (*n*=1,168) that were not further classified.

† Percentages of adenomas with specific histopathological characteristics according to size are presented.

‡ Adjusted for age at diagnosis; if we had also adjusted for year of endoscopy, history of large bowel endoscopy before study entry and whether any endoscopy for routine screening had been conducted, the estimates would be a:7.51 (5.95-9.48) and b:9.42 (6.45-13.8).

§ 148 were villous adenomas and 32 were carcinoma in situ.

Table 4-2. Description of the study population

Characteristic	Men with advanced adenomas *n*=1,246	Men with non-advanced adenomas *n*=1,661	Men with adenomas that were not classified *n*=1,168	Men not known with adenomas *n*=22,694
*Most proximally located adenoma, %**				
Rectum	13.0	13.7	14.3	-
Distal colon	46.4	42.1	43.2	-
Proximal colon	40.6	44.2	42.6	-
Age at study entry, yr	62.5 ± 9.0	60.0 ± 8.7	59.5 ± 8.9	60.2 ± 9.5
Age at most recent endoscopy/ diagnosis†	65.3 ± 8.8	63.6 ± 8.6	63.3 ± 8.5	65.2 ± 9.3
Year of most recent endoscopy or diagnosis, %‡				
1986-1992	30.7	30.1	12.5	12.3
1993-1998	38.9	36.3	51.9	24.8
1999-2002	30.4	33.6	35.6	62.9
Large bowel endoscopy before study entry, %‡	13.4	18.0	14.2	16.8
Ever a large bowel endoscopy for routine screening, %‡	81.9	83.5	82.0	79.4
Multiple adenomas, %‡	40.4	28.0	8.1	-
Family history of colorectal cancer, %‡	22.3	24.6	21.1	13.9

* This figure reflects both the presence of adenomas at each location and whether sigmoidoscopy or colonoscopy was conducted, so the percentage of people with proximal adenomas should not be interpreted as reflecting the prevalence of adenomas.

† This figure represents the age-adjusted date of most recent endoscopy for the men who did not report adenomas and the date of diagnosis for adenoma patients. The values were very similar before adjustment for age.

‡ Adjusted for age at most recent endoscopy.

Table 4-3. Age-adjusted means or percentages of modifiable factors in the study population, and determinants of advanced versus non-advanced adenomas*

Exposure	Unit of exposure	Age-adjusted means or percentages*			
		Men with advanced adenomas *n*=1,246	Men with non-advanced adenomas *n*=1,661	Men with non-classified adenomas *n*=1,168	Men not known with adenomas *n*=22,694
Lifestyle factors					
BMI	kg/m²	26.0	25.8	26.0	25.7
Height§	in.	70.1	70.1	70.1	70.1
Physical activity	MET-h/wk	28.4	29.4	29.2	31.7
Smoking status	% current	7.7	5.5	7.1	5.1
Alcohol consumption	g/d	12.4	12.3	12.4	10.6
Aspirin use§	% ≥2 tablets/wk	41.3	40.8	41.7	42.5
Multivitamin use	%	46.6	51.3	52.7	55.3
Macronutrients					
Carbohydrates	g/d	240.1	242.4	243.7	247.5
Starch	g/d	77.0	78.9	79.8	81.5
Total fat	g/d	69.9	68.8	68.5	68.0
Sucrose	g/d	43.7	43.6	43.9	44.3
Glycemic index and load					
Glycemic index	/d	52.9	53.0	52.9	53.1
Glycemic load	/d	127.0	128.6	128.9	131.5
Foods					
Red meat	servings/d	0.59	0.54	0.56	0.55
Processed meat	servings/d	0.34	0.30	0.30	0.30
Poultry	servings/d	0.38	0.38	0.38	0.40
Fish§	servings/d	0.38	0.41	0.39	0.39
High-fat dairy	servings/d	0.85	0.80	0.81	0.79
Low-fat dairy	servings/d	1.05	1.09	1.15	1.14
Fruits	servings/d	1.65	1.69	1.67	1.76
Vegetables	servings/d	3.62	3.68	3.69	3.80
Cruciferous	servings/d	0.48	0.48	0.49	0.50
Whole grain	servings/d	1.41	1.56	1.53	1.54
Refined grain	servings/d	1.25	1.23	1.25	1.25
Micronutrients					
Fiber	g/d	21.7	22.1	22.1	22.6
Folate	µg/d	509	527	525	549
Methionine	g/d	2.13	2.13	2.12	2.13
Vitamin D	IU/d	425	452	449	455
Calcium	mg/d	898	912	933	941
β-carotene	µg/d	5,066	5,288	5,320	5,603
Marine fatty acids**	g/d	0.32	0.34	0.33	0.33

* Means for continuous variables, percentages of categorized variables. (Cumulative) updated values up to year of diagnosis for cases or most recent endoscopy for men without adenomas are presented.

† Adjusted for: total energy intake (quintiles), age (in 5-yr age groups), history of endoscopy prior to study entry (yes/no), routine screening versus other indications for any endoscopy, aspirin use (currently, past or never ≥2 times/wk), use of multivitamins (current, former, never), smoking (never, quit ≤10 yr ago, quit>10 yr ago, current, missing) consumption of red meat (quintiles), alcohol (0, 0-<10, 10-<20, 20-<30, ≥30 g/d); intake of folate (quintiles), calcium (quintiles), BMI (quintiles), physical activity (quintiles), family history of colorectal cancer (yes/no), proximal adenomas (yes/no), adenomas in distal colon, adenomas in rectal colon (yes/no).

Highest category of exposure		Lowest category of exposure		Advanced *vs.* non-advanced adenomas†‡	
Value	*N* advanced/ non-advanced	Value	*N* advanced/ non-advanced	Odds ratio (95% confidence interval)	P_{trend}
≥27.9	268/343	<23.2	206/308	1.08 (0.84-1.40)	0.30
>72	214/299	≤68	356/419	0.91 (0.71-1.15)§	0.54§
≥48.3	205/294	<9.87	321/360	0.84 (0.66-1.08)	0.30
Current	96/96	Never	481/694	1.33 (0.97-1.84)	0.0032
≥30	152/196	0	181/219	0.83 (0.61-1.14)	0.29
Currently ≥2	514/677	Never ≥2	540/727	0.98 (0.83-1.17)§	0.85§
Current	586/842	Never	379/439	0.82 (0.65-1.02)	0.074
≥277	198/273	<216	322/386	1.04 (0.76-1.43)	0.73
≥96.1	186/279	<65.4	356/394	0.78 (0.60-1.01)	0.97
≥78.3	315/350	<58.2	217/300	0.96 (0.72-1.29)	0.70
≥55.0	239/311	<32.6	273/362	1.06 (0.82-1.39)	0.27
≥55.3	226/350	<51.0	280/338	0.70 (0.54-0.90)	0.0030
≥149.1	182/284	<112.7	332/392	0.81 (0.60-1.10)	0.28
≥0.83	285/312	<0.24	236/313	1.35 (0.92-1.97)	0.96
≥0.50	299/352	<0.07	198/287	1.06 (0.79-1.42)	0.43
≥0.56	216/309	<0.19	289/377	1.02 (0.79-1.31)	0.80
≥0.57	257/374	<0.17	247/275	0.89 (0.69-1.16)§	0.24§
≥1.14	284/356	<0.26	245/342	1.05 (0.81-1.35)	0.91
≥1.79	235/292	<0.34	299/364	1.00 (0.72-1.39)	0.96
≥2.50	228/293	<0.81	285/388	1.15 (0.87-1.51)	0.49
≥5.06	222/334	<2.25	285/358	0.90 (0.68-1.19)	0.98
≥0.71	245/309	<0.20	216/267	1.06 (0.81-1.38)	0.54
≥2.32	231/348	<0.57	288/309	0.78 (0.59-1.02)	0.069
≥1.78	254/325	<0.58	261/334	0.94 (0.72-1.22)	0.33
≥27.1	226/271	<17.3	319/383	1.18 (0.89-1.58)	0.48
≥722	207/303	<348	325/376	0.96 (0.71-1.28)	1.00
≥2.41	258/336	<1.83	241/336	1.04 (0.80-1.33)	0.85
≥642	225/328	<239	295/353	0.82 (0.58-1.17)	0.19
≥1,179	218/306	<655	304/389	1.01 (0.78-1.30)	0.91
≥7,477	203/277	<3,115	303/398	1.06 (0.81-1.39)	0.97
≥0.46	235/370	<0.14	278/340	0.88 (0.68-1.14)	0.20

‡ All intermediate categories were part of the model, but only the comparison of the extreme categories is presented.

§ Odds ratios and *P*-values for trend should not be directly interpreted, because potential interaction according to adenoma location was detected. See text and Table 4-4.

** Dietary docosahexaenoic acid+eicosapentaenoic acid.

Table 4-4. Associations of selected exposure factors and risk of advanced and non-advanced adenomas versus men not known with adenomas

Exposure	Location	Category	Men with advanced adenomas	Men with non-advanced adenomas
			N	
Smoking	All	Never smokers	481	694
		Former smokers, quit > 10 yr ago	499	684
		Former smokers, quit ≤ 10 yr ago	135	139
		Current smokers	96	96
Physical activity	All	< 9.87 MET-h/wk	321	360
		9.87-<18.9 MET-h/wk	262	253
		18.9-<30.1 MET-h/wk	234	340
		30.1-<48.3 MET-h/wk	224	314
		≥48.3 MET-h/wk	205	294
Multivitamins	All	Never	379	439
		Former user	281	380
		Current user	586	842
Whole grain	All	<0.57 servings/d	288	309
		0.57-<1.03 servings/d	282	361
		1.03-<1.50 servings/d	219	313
		≥2.32 servings/d	231	348
Starch	All	<65.4 g/d	356	394
		65.4-<75.9 g/d	241	371
		75.9-<84.9 g/d	248	323
		84.9-<96.1 g/d	215	294
		≥96.1 g/d	186	279
Glycemic index	All	<51.0 /d	280	338
		51.0-<52.6 /d	241	332
		52.6-<53.8 /d	251	347
		53.8-<55.3 /d	248	294
		≥55.3 /d	226	350
Aspirin†	Rectum	Never ≥2 tablets/wk	68	122
		≥2 tablets/wk in the past	23	42
		Currently ≥2 tablets/wk	71	64
	Distal colon	Never ≥2 tablets/wk	252	289
		≥2 tablets/wk in the past	84	97
		Currently ≥2 tablets/wk	196	281
	Proximal colon	Never ≥2 tablets/wk	98	228
		≥2 tablets/wk in the past	39	78
		Currently ≥2 tablets/wk	116	240

Continued: page 90

* Odds ratios (OR) and 95% confidence intervals (95% CI). Adjusted for the same factors as listed in Table 4-3.

† Only men with adenomas at one specific location were included in these comparisons. Adjusted for the same factors as listed in Table 4-3, but tertiles were used instead of quintiles.

Men not known with adenomas	Advanced *vs.* non-advanced adenomas*		Advanced adenomas *vs.* no adenomas*		Non-advanced adenomas *vs.* no adenomas*	
	OR (95% CI)	P_{trend}	OR (95% CI)	P_{trend}	OR (95% CI)	P_{trend}
10,504	1 (reference)	0.0032	1 (reference)	<.0001	1 (reference)	0.0079
8,610	1.00 (0.84-1.20)		1.23 (1.08-1.41)		1.22 (1.09-1.37)	
1,565	1.32 (1.00-1.75)		1.69 (1.37-2.07)		1.25 (1.03-1.51)	
1,150	1.33 (0.97-1.84)		1.59 (1.26-2.01)		1.14 (0.91-1.43)	
4,429	1 (reference)	0.30	1 (reference)	0.0002	1 (reference)	0.0017
4,514	0.87 (0.69-1.10)		0.84 (0.71-0.99)		0.96 (0.82-1.12)	
4,523	0.82 (0.64-1.03)		0.76 (0.64-0.91)		0.92 (0.78-1.08)	
4,574	0.88 (0.69-1.12)		0.75 (0.62-0.90)		0.85 (0.72-0.99)	
4,654	0.84 (0.66-1.08)		0.68 (0.57-0.83)		0.79 (0.67-0.93)	
5,170	1 (reference)	0.074	1 (reference)	0.0004	1 (reference)	0.16
4,963	0.88 (0.71-1.10)		0.83 (0.70-0.97)		0.94 (0.81-1.08)	
12,561	0.82 (0.65-1.02)		0.73 (0.62-0.86)		0.88 (0.77-1.03)	
4,241	1 (reference)	0.069	1 (reference)	0.79	1 (reference)	0.0012
4,738	0.91 (0.72-1.15)		0.97 (0.82-1.16)		1.13 (0.96-1.33)	
4,440	0.78 (0.60-1.01)		0.88 (0.73-1.06)		1.13 (0.95-1.34)	
4,566	0.78 (0.59-1.02)		1.02 (0.83-1.25)		1.36 (1.13-1.62)	
4,346	1 (reference)	0.97	1 (reference)	0.0001	1 (reference)	0.0001
4,515	0.70 (0.56-0.88)		0.69 (0.58-0.82)		0.93 (0.80-1.08)	
4,526	0.86 (0.68-1.09)		0.75 (0.63-0.89)		0.83 (0.71-0.97)	
4,603	0.83 (0.65-1.06)		0.66 (0.55-0.79)		0.74 (0.63-0.87)	
4,703	0.78 (0.60-1.01)		0.57 (0.47-0.70)		0.69 (0.58-0.82)	
4,476	1 (reference)	0.0030	1 (reference)	0.019	1 (reference)	0.94
4,526	0.86 (0.67-1.09)		0.90 (0.76-1.08)		1.02 (0.87-1.19)	
4,532	0.80 (0.63-1.02)		0.92 (0.77-1.11)		1.05 (0.90-1.23)	
4,601	0.97 (0.76-1.25)		0.90 (0.75-1.08)		0.89 (0.76-1.06)	
4,559	0.70 (0.54-0.90)		0.78 (0.64-0.94)		1.04 (0.88-1.23)	
9,024	1 (reference)	0.0034	1 (reference)	0.87	1 (reference)	<0.0001
4,027	1.04 (0.57-1.89)		0.82 (0.51-1.33)		0.79 (0.55-1.13)	
9,643	2.02 (1.27-3.20)		1.03 (0.73-1.45)		0.51 (0.37-0.70)	
9,024	1 (reference)	0.069	1 (reference)	0.013	1 (reference)	0.93
4,027	1.01 (0.72-1.42)		0.83 (0.64-1.07)		0.82 (0.65-1.04)	
9,643	0.79 (0.61-1.02)		0.79 (0.65-0.95)		0.99 (0.84-1.18)	
9,024	1 (reference)	0.67	1 (reference)	0.66	1 (reference)	0.92
4,027	1.13 (0.71-1.78)		0.87 (0.59-1.27)		0.77 (0.59-1.00)	
9,643	1.07 (0.77-1.50)		1.06 (0.80-1.40)		0.99 (0.82-1.19)	

Table 4-4. Continued

Exposure	Location	Category	Men with advanced adenomas	Men with non-advanced adenomas *N*
Height†	Rectum	≤68.0 in.	39	63
		68.1-≤70.0 in.	39	70
		70.1-≤72.0 in.	53	53
		>72.1 in.	31	42
	Distal colon	≤68.0 in.	175	161
		68.1-≤70.0 in.	142	200
		70.1-≤72.0 in.	137	174
		>72.1 in.	78	132
	Proximal colon	≤68.0 in.	72	143
		68.1-≤70.0 in.	68	175
		70.1-≤72.0 in.	60	140
		>72.1 in.	53	88
Fish †	Rectum	≤0.24 servings/d	55	70
		0.25-≤0.43 servings/d	56	79
		>0.43 servings/d	51	79
	Distal colon	≤ 0.24 servings/d	194	211
		0.25-≤0.43 servings/d	191	227
		>0.43 servings/d	147	229
	Proximal colon	≤ 0.24 servings/d	75	158
		0.25-≤0.43 servings/d	76	190
		>0.43 servings/d	102	198

Footnotes: see page 88

with only one adenoma, tubulovillous adenoma were 10.2 times (95% CI=7.5-13.9) and villous adenoma/carcinoma in situ were 13.9 times (95% CI=8.36-23.0) more likely to be large than were tubular adenomas. In the same group of men, adenomas were more likely to occur in the rectum (OR=2.78, 95%=1.18-6.61) and distal colon (OR=2.16, 95% CI=1.06-4.40) than in the proximal colon.

The endoscopy and pathology reports of the majority of adenomas that we did not classify further did not include any information on size and histopathology. Relatively more adenomas with an unknown histopathology were small, and relatively more adenomas whose size was unknown were tubular. Information as to whether a person had synchronous adenomas was unavailable for 73% of the men whose adenomas we did not classify as advanced or non-advanced.

Endoscopy-Related Characteristics and Family History of Colorectal Cancer

The majority of men who did not report an adenoma underwent an endoscopy in the more recent years of follow-up (Table 4-2). Men with non-advanced adenomas and not classified adenomas were on average two years younger than men with advanced adenomas at study entry and at diagnosis. Although being of similar age as people who developed non-advanced or non-classified adenomas at study entry, men who did not report an adenoma during the follow-up

Men not known with adenomas	Advanced *vs.* non-advanced adenomas*		Advanced adenomas *vs.* no adenomas*		Non-advanced adenomas *vs.* no adenomas*	
	OR (95% CI)	P_{trend}	OR (95% CI)	P_{trend}	OR (95% CI)	P_{trend}
6,018	1 (reference)	0.089	1 (reference)	0.21	1 (reference)	0.25
6,670	0.96 (0.55-1.69)		0.90 (0.58-1.41)		0.94 (0.66-1.32)	
6,120	1.78 (1.02-3.11)		1.30 (0.86-1.98)		0.73 (0.51-1.06)	
3,886	1.36 (0.73-2.52)		1.21 (0.75-1.95)		0.89 (0.60-1.33)	
6,018	1 (reference)	0.0032	1 (reference)	0.0018	1 (reference)	0.38
6,670	0.67 (0.49-0.90)		0.72 (0.57-0.90)		1.07 (0.87-1.33)	
6,120	0.75 (0.55-1.02)		0.74 (0.59-0.93)		0.99 (0.79-1.23)	
3,886	0.58 (0.40-0.82)		0.68 (0.51-0.89)		1.17 (0.92-1.49)	
6,018	1 (reference)	0.33	1 (reference)	0.47	1 (reference)	0.50
6,670	0.81 (0.54-1.21)		0.89 (0.64-1.24)		1.10 (0.88-1.38)	
6,120	0.93 (0.61-1.41)		0.87 (0.62-1.24)		0.95 (0.74-1.20)	
3,886	1.34 (0.86-2.11)		1.26 (0.88-1.82)		0.94 (0.71-1.23)	
7,292	1 (reference)	0.39	1 (reference)	0.83	1 (reference)	0.28
7,950	0.92 (0.56-1.51)		0.95 (0.65-1.40)		1.04 (0.75-1.45)	
7,452	0.79 (0.47-1.35)		0.95 (0.63-1.43)		1.20 (0.85-1.69)	
7,292	1 (reference)	0.062	1 (reference)	0.15	1 (reference)	0.25
7,950	0.96 (0.73-1.27)		0.95 (0.77-1.16)		0.99 (0.81-1.20)	
7,452	0.76 (0.56-1.03)		0.85 (0.67-1.07)		1.11 (0.91-1.36)	
7,292	1 (reference)	0.55	1 (reference)	0.021	1 (reference)	0.021
7,950	0.84 (0.57-1.23)		0.95 (0.68-1.31)		1.14 (0.91-1.41)	
7,452	1.06 (0.72-1.57)		1.39 (1.01-1.91)		1.31 (1.04-1.63)	

underwent, on average, their most recent endoscopy at the age at which men were diagnosed with advanced adenomas.

Especially men with non-advanced adenomas and, to a smaller extent, men who did not report to have had adenomas were more likely to have undergone an endoscopy of the large bowel before study entry. Men with adenomas, but in particular non-advanced adenomas, were more likely to have undergone endoscopy for routine screening and to have a positive family history of colorectal cancer.

Characteristics of the Men with Adenomas that Were not Further Classified

The distributions of modifiable risk factors (Table 4-3) for men with adenomas that we did not classify seemed to be a mixture of the distributions among men with advanced and non-advanced adenomas.

Characteristics of the Men without Adenomas

As the main focus of the current analyses is to evaluate whether the distributions of risk factors differ between advanced and non-advanced adenoma patients, we do not include an extensive comparison of the adenoma patients versus the men who did not report to have an adenoma here. However, mean values of exposure variables within this group are included in Table 4-3, and we

will present the comparisons versus adenoma-free men for those factors for which we detected potential different associations with advanced and non-advanced adenomas in Table 4-4.

Men with Advanced Adenomas versus Men with Non-Advanced Adenomas

Before comparing men with advanced and non-advanced adenomas, we explored whether the associations of advanced versus non-advanced adenomas depended on adenoma location in men who had adenomas at only one location. Most associations did not depend on location ($P_{interaction} \geq 0.10$), but the differences in distribution between men with advanced and non-advanced adenomas regarding height ($P_{interaction}=0.052$), aspirin ($P_{interaction}=0.0028$) and possibly fish ($P_{interaction}=0.067$) varied according to location.

Table 4-3 shows that the men with advanced adenomas and non-advanced adenomas were exposed to most factors we studied to a similar extent. However, men with advanced adenomas were more likely to smoke, but they were less likely to have a high consumption of whole grains and to consume foods with high glycemic index. Men with advanced adenomas also tended to be less physically active and to take multivitamin supplements and to have a somewhat higher intake of starch, but these differences were not statistically significant.

Men with Advanced and Non-Advanced Adenomas versus Men without Reported Adenomas

Table 4-4 illustrates why the aforementioned differences arose. Smoking was associated with risk of advanced and non-advanced adenomas but more prominently so with risk of advanced adenomas. Physical activity and intake of starch were inversely associated with risk of advanced adenomas but to a smaller extent with risk of non-advanced adenomas. The same applied to use of multivitamins, but the difference between the associations was not statistically significant. The difference between the associations regarding consumption of whole grains was also not statistically significant, although it appeared to be not associated with risk of advanced adenomas but seemed to be positively associated with risk of non-advanced adenomas. Men who consumed foods with a high glycemic index were less likely to be diagnosed with advanced adenomas but they were as likely to be diagnosed with non-advanced adenomas as men who consumed foods with a low glycemic index.

The dependency of the case-case association on location regarding aspirin was due to the fact that use of aspirin was solely associated with reduced risk of rectal non-advanced adenomas, but on the contrary, aspirin was solely associated with reduced risk of advanced adenomas in the distal colon. The interaction between height and location could be attributed to a different association within the distal colon: height was not associated with non-advanced adenomas, irrespective of site; however, height was inversely associated with distal advanced adenomas, whereas height tended to be not or somewhat positively associated with advanced adenomas at other locations. Although the interaction between fish consumption and adenoma location was not statistically significant, fish consumption tended to be inversely associated with risk of advanced adenomas in the distal colon but fish consumption was not associated with risk of non-advanced adenomas

at that location. Overall, fish tended to be positively associated with non-advanced adenomas and adenomas in the proximal colon.

Sensitivity Analyses

The association between the examined factors and advanced versus non-advanced adenomas did generally not depend on the presence of synchronous adenomas, as indicated by *P*-values for interaction ≥0.10. Only the association with smoking might depend on adenoma multiplicity; current smokers with one adenoma were 1.63 (95% CI=1.15-2.33) times more likely to have an advanced adenoma than men who never smoked, but current smokers diagnosed with synchronous adenomas were not more likely to have an advanced adenoma (OR=0.84, 95% CI=0.52-1.43). However, the difference between these associations was not statistically significant ($P_{interaction}$=0.064).

The associations of risk factors with advanced and non-advanced adenomas did also not depend on the year of most recent endoscopy, except the association with calcium intake ($P_{interaction}$=0.038). However, as calcium intake differed between men with advanced and non-advanced adenomas only among men who underwent endoscopy between 1986 and 1992, we do not present these data here.

Discussion

We observed that most dietary and lifestyle factors were not differently associated with advanced and non-advanced colorectal adenomas. However, smoking was a stronger risk factor for advanced adenomas than for non-advanced adenomas and a high glycemic index was solely inversely associated with risk of non-advanced adenomas. Indications for different associations were also found for physical activity, use of multivitamins and intake of starch, height, aspirin and fish. The differences in height, use of aspirin and possibly fish consumption between men with advanced and non-advanced adenomas depended on adenoma location. Before we compare our findings to the literature and interpret them, we have to consider the methodological drawbacks and strengths of our study.

This study is the largest study on advanced and non-advanced adenomas conducted to date, in which participants were followed-up over time while exposure information was repeatedly assessed in a standardized way. To make such a large-scale study feasible, adenoma data that were collected for clinical practice were used. However, endoscopy and pathology reports often, but not always, mention how many adenomas a person has and the size of the largest adenoma, but individual adenoma characteristics are often not given. Not all adenomas are removed or sent for histopathological evaluation, because the clinician may judge the adenoma as not being clinically relevant; indeed, many of the adenomas for which we did not have sufficient data available for further categorization seemed to be tubular or small adenomas. However, the distributions of most modifiable factors among men with such non-classified adenomas mimicked the weighted average of the distributions among men with advanced and non-advanced adenomas, which supported our decision not to include them in either category. It is likely that the presence or absence of data on histopathology and size of the adenoma does not depend on exposure level

of the participant. Thus, we may have underestimated the strenghts of the associations but it is unlikely that the incomplete case ascertainment has resulted in systematic errors. Indeed, had we decided to treat the non-classified adenomas as being non-advanced or even as being advanced, the estimates would have pointed in the same directions, which suggest that the impact of such a bias is small.

Like in any adenoma study, some adenomas will have been missed at endoscopy, especially so for small adenomas[45]. Thus, some of the men without an adenoma diagnosis may actually belong to the group of non-advanced adenomas, and a proportion of the group classified as having non-advanced adenomas may, in fact, have advanced adenomas, also because not everybody underwent a full colonoscopy. In spite of this, it is unlikely that the rigorousness of the endoscopic examination depends on modifiable factors. The resulting non-differential misclassification might have diluted the risk estimates, but the impact is likely to be reduced by the large number of truly adenoma-free men in our study population.

A further complexity related to the use of available clinical data is that it is generally impossible to tell where in the colon or rectum a particular advanced adenoma was found in the presence of adenomas at multiple locations. As a non-advanced distal adenoma diagnosed by sigmoidoscopy does not exclude the presence of an advanced proximal adenoma[46] while risk factors may vary by subsite[47], we evaluated whether the case-case analyses depended on adenoma location in the group of people of whom only adenomas at one location were registered. We observed some differences according to location, which suggests that future studies should pay close attention to adenoma location when comparing groups according to size and histopathology. This study was the first that acknowledges that differences between adenomas stages may reflect different distributions of correlated characteristics such as adenoma multiplicity and adenoma location.

We classified adenomas as advanced or non-advanced as a proxy method for separating them according to their stage of development and the chance of becoming malignant. Adenomas with a villous component, with severe dysplasia or large (≥ 1 cm) adenomas[2-7, 9-11] are thought to be most likely to develop further into colorectal cancer. We did not incorporate dysplasia in our definition of advanced adenomas, however, because the consistency of the classification into high-grade and low-grade dysplasia has been shown to be poor when different community pathologists were involved like in our study[48, 49]. Nonetheless, the majority of adenomas showing high-grade dysplasia are probably captured in our definition of advanced adenomas, as high-grade dysplasia is strongly correlated with size and histopathological characteristics[50-52]. If endoscopy had been planned later in time, however, some of the non-advanced adenomas may have progressed to advanced adenomas and some adenomas may have regressed in size[53, 54]. The men with advanced adenomas were two years older at diagnosis than the men with non-advanced adenomas, which could suggest that the observed differences between advanced and non-advanced adenomas might simply reflect that the men with non-advanced adenomas were examined at younger age. We cannot completely rule out that explanation, despite adjusting for age at most recent endoscopy in all analyses, but it is reassuring that in general, age did not modify the association between the examined factors and advanced versus non-advanced adenomas.

We acknowledge that we examined a large number of potential risk factors, which may have lead to chance findings. It is therefore important to consider the results of this study in the context of existing literature.

To our knowledge, only two studies explicitly studied whether associations differ between advanced and non-advanced adenomas[31] before. One was a relatively small-sized study that pooled data from four case-control studies where advanced adenomas were defined as those with severe dysplasia[31], carcinoma in situ or intramucosal carcinoma. In line with our study, the pooled case-control study could not detect differences in the distributions of most risk factors among patients with advanced (n=119) and non-advanced colorectal adenomas (n=441)[31]. Among men but not women, increased physical activity was more strongly associated with reduced risk for advanced than with reduced risk for non-advanced adenomas[31], for which we also found indications. The other study included 70 advanced and 132 non-advanced adenoma patients, and observed stronger associations of serum levels of IGF-I and IGF-I/IGFBP-3 with advanced than with non-advanced adenomas[24], which may correspond with our observation that glycemic index, height (but only regarding the distal colon) and to a lesser extent physical activity tended to be inversely associated with advanced but not with non-advanced adenomas, although this was not the case for BMI.

Our study also bears similarities to a small case-control study that compared risk factors for large (n=208) and small (n=154) adenomas. In that study, patients with large adenomas were more likely to have a high intake of animal fats compared with patients with small adenomas, but they were less likely to have a high intake of yoghurt[30]. No differences in the distribution of other dietary factors[30] as well as of BMI and physical activity[32] could be detected. The prevalence of smoking did also not differ between the two patient groups but alcohol consumption was higher among in the group with large adenomas than in the group with small adenomas[33], which is not in line with the findings of our study. Analyses among participants in a trial that monitored growth of small adenomas left in situ suggested that high body fatness[19] was strongly associated with increased adenoma growth of ≥1mm after three years, which could possibly also apply to alcohol consumption[16]. Fruits and carbohydrates were weakly inversely associated with growth[17], whereas a similar study that left adenomas in situ for two years suggested that intake of dietary fiber and possibly non-fiber carbohydrate and cruciferous vegetables were associated with adenoma growth[55]. Thus, the findings of the two growth studies are not in concordance with each other, nor with our study and the study on large and small adenomas[30, 32, 33]. It cannot be completely excluded that these different findings are caused by the different endpoints and differences in exposure assessment, but chance may also be an explanation especially because previous studies were small. However, our results regarding fish and whole grains should be interpreted with caution, because they pointed in an unexpected direction.

Adenoma advancement most likely results from the interaction of multiple parallel processes that lead to elevated cell proliferation, evasion of apoptosis or selection for specific sort of lesions, whilst mutations in critical cancer genes accumulate[56]. The dynamics of these processes may depend on colorectal mucosal and luminal conditions, which could be influenced by dietary or lifestyle factors, and the balance between them will determine the association of advanced versus

non-advanced adenomas. If risk factors would only interfere with the causation or accumulation of mutations, then it is not surprising that associations are more clearly visible in the group of advanced adenomas. This idea is consistent with most of our data, and it seems biologically plausible for smoking, which is well known to be mutagenic[57]. The stronger associations of use of multivitamins and, to a smaller extent, consumption of red meat with advanced adenomas than with non-advanced adenomas could also reflect mutagenic effects, although the differences between these associations were not statistically significant. Nonetheless, it is unlikely that such an explanation accounts for the different associations of glycemic index, physical activity, starch and height (but the latter only in the distal colon) with advanced and non-advanced adenomas, as it is hard to imagine how these factors would interfere with the causation of mutations. The IGF growth factor axis may offer an explanation for these findings, as the IGF system may especially be involved in the later stages of adenoma development. The observed effects of sulindac on suppression of growth of adenomas via inhibition of the enzyme cyclooxygenase (COX) in patients with familial adenomatous polyposis[18] and the higher levels of COX expression in the distal versus the proximal colon as well as in advanced versus non-advanced adenomas[58] support our finding that aspirin tended to be more strongly associated with reduced risk of advanced than with risk of non-advanced adenomas in the distal colon. However, it is hard to explain why aspirin was associated with reduced risk of rectal non-advanced adenomas but not with rectal advanced adenomas.

In conclusion, our study suggests that associations of modifiable factors with advanced and non-advanced adenomas are mostly similar. Most of the factors for which we observed different effects seem to be associated with the IGF or COX systems, which may explain the stronger effects on advanced adenomas, although other mechanisms cannot be cancelled out. Hypothetically, the stronger associations of smoking and other mutagenic factors could simply reflect accumulation of mutations, but again, other mechanisms cannot be excluded. Other studies will need to confirm our findings and they may shed further light on the mechanisms involved.

Acknowledgements

The HPFS is supported by NCI Research Grant CA 55075. Petra Wark's visit to the Harvard School of Public Health was made possible with support from the Dutch Cancer Society, and she was further supported by the Netherlands Organisation for Health Research and Development.

References

1. Chen CD, Yen MF, Wang WM, Wong JM, Chen TH. A case-cohort study for the disease natural history of adenoma-carcinoma and de novo carcinoma and surveillance of colon and rectum after polypectomy: implication for efficacy of colonoscopy. *Br J Cancer.* 2003;88(12):1866-1873.
2. Martinez ME, Sampliner R, Marshall JR, Bhattacharyya AK, Reid ME, Alberts DS. Adenoma characteristics as risk factors for recurrence of advanced adenomas. *Gastroenterology.* 2001;120(5):1077-1083.
3. Winawer SJ, Zauber AG, O'Brien MJ, et al. Randomized comparison of surveillance intervals after colonoscopic removal of newly diagnosed adenomatous polyps. The National Polyp Study Workgroup. *N Engl J Med.* 1993;328(13): 901-906.
4. Külling D, Christ AD, Karaaslan N, Fried M, Bauerfeind P. Is histological investigation of polyps always necessary? *Endoscopy.* 2001;33(5):428-432.
5. Van Stolk RU, Beck GJ, Baron JA, Haile R, Summers R. Adenoma characteristics at first colonoscopy as predictors of adenoma recurrence and characteristics at follow-up. The Polyp Prevention Study Group. *Gastroenterology.* 1998;115(1):13-18.
6. Noshirwani KC, Van Stolk RU, Rybicki LA, Beck GJ. Adenoma size and number are predictive of adenoma recurrence: implications for surveillance colonoscopy. *Gastrointest Endosc.* 2000;51(4 Pt 1):433-437.
7. Jorgensen OD, Kronborg O, Fenger C. The Funen Adenoma Follow-Up Study. Characteristics of patients and initial adenomas in relation to severe dysplasia. *Scand J Gastroenterol.* 1993;28(3):239-243.
8. Loeve F, Van Ballegooijen M, Boer R, Kuipers EJ, Habbema JD. Colorectal cancer risk in adenoma patients: a nation-wide study. *Int J Cancer.* 2004;111(1):147-151.
9. Atkin WS, Morson BC, Cuzick J. Long-term risk of colorectal cancer after excision of rectosigmoid adenomas. *N Engl J Med.* 1992;326(10):658-662.
10. Simons BD, Morrison AS, Lev R, Verhoek-Oftedahl W. Relationship of polyps to cancer of the large intestine. *J Natl Cancer Inst.* 1992;84(12):962-966.
11. Jorgensen OD, Kronborg O, Fenger C. The Funen Adenoma Follow-up Study. Incidence and death from colorectal carcinoma in an adenoma surveillance program. *Scand J Gastroenterol.* 1993;28(10):869-874.
12. Neugut AI, Jacobson JS, De Vivo I. Epidemiology of colorectal adenomatous polyps. *Cancer Epidemiol Biomarkers Prev.* 1993;2(2):159-176.
13. Giovannucci E. Modifiable risk factors for colon cancer. *Gastroenterol Clin North Am.* 2002;31(4):925-943.
14. Giovannucci E. An updated review of the epidemiological evidence that cigarette smoking increases risk of colorectal cancer. *Cancer Epidemiol Biomarkers Prev.* 2001;10(7):725-731.
15. Pöschl G, Seitz HK. Alcohol and cancer. *Alcohol Alcohol.* 2004;39(3):155-165.
16. Almendingen K, Hofstad B, Vatn MH. Does intake of alcohol increase the risk of presence and growth of colorectal adenomas followed-up in situ for three years? *Scand J Gastroenterol.* 2002;37(1):80-87.
17. Almendingen K, Hofstad B, Vatn MH. Dietary habits and growth and recurrence of colorectal adenomas: results from a three-year endoscopic follow-up study. *Nutr Cancer.* 2004;49(2):131-138.
18. Giardiello FM, Hamilton SR, Krush AJ, et al. Treatment of colonic and rectal adenomas with sulindac in familial adenomatous polyposis. *N Engl J Med.* 1993;328(18):1313-1316.
19. Almendingen K, Hofstad B, Vatn MH. Does high body fatness increase the risk of presence and growth of colorectal adenomas followed up in situ for 3 years? *Am J Gastroenterol.* 2001;96(7):2238-2246.
20. Risio M. Cell proliferation in colorectal tumor progression: an immunohistochemical approach to intermediate biomarkers. *J Cell Biochem Suppl.* 1992;16G:79-87.
21. Bostick RM, Fosdick L, Grandits GA, et al. Colorectal epithelial cell proliferative kinetics and risk factors for colon cancer in sporadic adenoma patients. *Cancer Epidemiol Biomarkers Prev.* 1997;6(12):1011-1019.
22. Caderni G, Palli D, Lancioni L, et al. Dietary determinants of colorectal proliferation in the normal mucosa of subjects with previous colon adenomas. *Cancer Epidemiol.Biomarkers.Prev.* 1999;8(3):219-225.
23. Hofstad B, Almendingen K, Vatn M, et al. Growth and recurrence of colorectal polyps: a double-blind 3-year intervention with calcium and antioxidants. *Digestion.* 1998;59(2):148-156.
24. Schoen RE, Weissfeld JL, Kuller LH, et al. Insulin-like growth factor-I and insulin are associated with the presence and advancement of adenomatous polyps. *Gastroenterology.* 2005;129(2):464-475.

25. Giovannucci E. Insulin, insulin-like growth factors and colon cancer: a review of the evidence. *J Nutr.* 2001;131(11 Suppl):3109S-3120S.

26. Giovannucci E, Pollak M, Liu Y, et al. Nutritional predictors of insulin-like growth factor I and their relationships to cancer in men. *Cancer Epidemiol Biomarkers Prev.* 2003;12(2):84-89.

27. Voskuil DW, Vrieling A, Van 't Veer LJ, Kampman E, Rookus MA. The insulin-like growth factor system in cancer prevention: potential of dietary intervention strategies. *Cancer Epidemiol Biomarkers Prev.* 2005;14(1):195-203.

28. Juul A, Bang P, Hertel NT, et al. Serum insulin-like growth factor-I in 1030 healthy children, adolescents, and adults: relation to age, sex, stage of puberty, testicular size, and body mass index. *J Clin Endocrinol Metab.* 1994;78(3): 744-752.

29. Caprio S. Insulin: the other anabolic hormone of puberty. *Acta Paediatr Suppl.* 1999;88(433):84-87.

30. Senesse P, Boutron-Ruault MC, Faivre J, Chatelain N, Belghiti C, Meance S. Foods as risk factors for colorectal adenomas: a case-control study in Burgundy (France). *Nutr Cancer.* 2002;44(1):7-15.

31. Terry MB, Neugut AI, Bostick RM, et al. Risk factors for advanced colorectal adenomas: a pooled analysis. *Cancer Epidemiol Biomarkers Prev.* 2002;11(7):622-629.

32. Boutron-Ruault MC, Senesse P, Meance S, Belghiti C, Faivre J. Energy intake, body mass index, physical activity, and the colorectal adenoma-carcinoma sequence. *Nutr Cancer.* 2001;39(1):50-57.

33. Boutron MC, Faivre J, Dop MC, Quipourt V, Senesse P. Tobacco, alcohol, and colorectal tumors: a multistep process. *Am J Epidemiol.* 1995;141(11):1038-1046.

34. Rimm EB, Giovannucci EL, Stampfer MJ, Colditz GA, Litin LB, Willett WC. Reproducibility and validity of an expanded self-administered semiquantitative food frequency questionnaire among male health professionals. *Am J Epidemiol.* 1992;135(10):1114-1126; discussion 1127-1136.

35. U. S. Department of Agriculture. *Composition of foods-raw, processed, and prepared, 1963-1992*. Washington (DC): Department of Agriculture, Government Printing Office; 1993.

36. Feskanich D, Rimm EB, Giovannucci EL, et al. Reproducibility and validity of food intake measurements from a semiquantitative food frequency questionnaire. *J Am Diet Assoc.* 1993;93(7):790-796.

37. Foster-Powell K, Miller JB. International tables of glycemic index. *Am J Clin Nutr.* 1995;62(4):871S-890S.

38. Giovannucci E, Colditz G, Stampfer MJ, et al. The assessment of alcohol consumption by a simple self-administered questionnaire. *Am J Epidemiol.* 1991;133(8):810-817.

39. Chasan-Taber S, Rimm EB, Stampfer MJ, et al. Reproducibility and validity of a self-administered physical activity questionnaire for male health professionals. *Epidemiology.* 1996;7(1):81-86.

40. Ainsworth BE, Haskell WL, Leon AS, et al. Compendium of physical activities: classification of energy costs of human physical activities. *Med Sci Sports Exerc.* 1993;25(1):71-80.

41. Rimm EB, Stampfer MJ, Colditz GA, Chute CG, Litin LB, Willett WC. Validity of self-reported waist and hip circumferences in men and women. *Epidemiology.* 1990;1(6):466-473.

42. Willett W. *Nutritional Epidemiology.* 2nd ed. New York, NY: Oxford University Press; 1998.

43. Hu FB, Manson JE, Liu S, et al. Prospective study of adult onset diabetes mellitus (type 2) and risk of colorectal cancer in women. *J Natl Cancer Inst.* 1999;91(6):542-547.

44. Begg CB, Gray R. Calculation of Polychotomous Logistic-Regression Parameters Using Individualized Regressions. *Biometrika.* 1984;71(1):11-18.

45. Rex DK, Cutler CS, Lemmel GT, et al. Colonoscopic miss rates of adenomas determined by back-to-back colonoscopies. *Gastroenterology.* 1997;112(1):24-28.

46. Levin TR, Palitz A, Grossman S, et al. Predicting advanced proximal colonic neoplasia with screening sigmoidoscopy. *Jama.* 1999;281(17):1611-1617.

47. Wei EK, Giovannucci E, Wu K, et al. Comparison of risk factors for colon and rectal cancer. *Int J Cancer.* 2004;108(3):433-442.

48. Jensen P, Krogsgaard MR, Christiansen J, Braendstrup O, Johansen A, Olsen J. Observer variability in the assessment of type and dysplasia of colorectal adenomas, analyzed using kappa statistics. *Dis Colon Rectum.* 1995;38(2):195-198.

49. Terry MB, Neugut AI, Bostick RM, Potter JD, Haile RW, Fenoglio-Preiser CM. Reliability in the classification of advanced colorectal adenomas. *Cancer Epidemiol Biomarkers Prev.* 2002;11(7):660-663.

50. O'Brien MJ, Winawer SJ, Zauber AG, et al. The National Polyp Study. Patient and polyp characteristics associated with high-grade dysplasia in colorectal adenomas. *Gastroenterology.* 1990;98(2):371-379.

51. Gschwantler M, Kriwanek S, Langner E, et al. High-grade dysplasia and invasive carcinoma in colorectal adenomas: a multivariate analysis of the impact of adenoma and patient characteristics. *Eur J Gastroenterol Hepatol.* 2002;14(2): 183-188.

52. Jørgensen OD, Kronborg O, Fenger C. The Funen Adenoma Follow-Up Study. Characteristics of patients and initial adenomas in relation to severe dysplasia. *Scand J Gastroenterol.* 1993;28(3):239-243.

53. Loeve F, Boer R, Zauber AG, et al. National Polyp Study data: evidence for regression of adenomas. *Int J Cancer.* 2004;111(4):633-639.
54. Hofstad B, Vatn MH, Andersen SN, et al. Growth of colorectal polyps: redetection and evaluation of unresected polyps for a period of three years. *Gut.* 1996;39(3):449-456.
55. Hoff G, Moen IE, Trygg K, et al. Colorectal adenomas and food. A prospective study of change in volume and total mass of adenomas in man. *Scand J Gastroenterol.* 1988;23(10):1253-1258.
56. Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell.* 2000;100(1):57-70.
57. Husgafvel-Pursiainen K. Genotoxicity of environmental tobacco smoke: a review. *Mutat Res.* 2004;567(2-3): 427-445.
58. Einspahr JG, Krouse RS, Yochim JM, et al. Association between Cyclooxygenase expression and colorectal adenoma characteristics. *Cancer Res.* 2003;63(14):3891-3893.

Abstract

Background: A family history of colorectal cancer may increase colorectal cancer risk by influencing adenoma growth or enhancing the formation of new lesions.

Aim and methods: Data of men from the prospective Health Professionals Follow-Up Study who underwent an endoscopy between 1986 and 2002 were used to evaluate whether a family history of colorectal cancer is associated with adenoma multiplicity or advanced adenoma stage (≥1cm or any mention of villous characteristics or carcinoma in situ). 22.9% of the 4,037 adenoma patients and 13.9% of the 22,498 adenoma-free men had a first-degree relative with colorectal cancer. 1,229 men were classified as having advanced and 1,652 as having non-advanced adenomas. 528 men had multiple and 1,768 single adenomas in the distal colon and rectum.

Results: After adjustment for presumed risk factors, a family history of colorectal cancer was similarly associated with advanced and non-advanced adenomas (advanced versus non-advanced, odds ratio (OR)=0.91, 95% confidence interval (CI)=0.76-1.08; advanced versus adenoma-free: OR=1.84, 95% CI=1.60-2.12; non-advanced versus adenoma-free: OR=2.03, 95% CI=1.80-2.29). However, a family history of colorectal cancer was stronger associated with multiple distal adenomas (multiple versus single, OR=1.39, 95% CI=1.11-1.74; multiple distal versus no distal adenomas: OR=2.24, 95% CI=1.83-2.73; single distal versus no distal adenomas: OR=1.61, 95% CI=1.42-1.82). The number of adenomas was also associated positively with a family history of colorectal cancer.

Conclusion: Adenoma advancement was not associated with family history in this study, but adenoma multiplicity was, which may suggest that heritable factors are more important in earlier than in the later stages of adenoma formation at the population level.

Family History of Colorectal Cancer: A Determinant of Advanced Adenoma Stage or Adenoma Multiplicity?

5

Petra A. Wark
Kana Wu
Pieter van 't Veer
Charles S. Fuchs
Edward L. Giovannucci

Introduction

First-degree relatives of individuals with colorectal cancer are approximately twice as likely to get colorectal cancer or their precursor lesions adenomatous polyps as people with unaffected relatives[1-3]. The risk of colorectal cancer is 4.3 times higher for people who have more than one such relative compared with those without one[1]. A study on twins suggested that about 35 percent of all colorectal cancer cases can be attributed to heritable factors[4]. Given that approximately 5 percent of all cancers are due to hereditary syndromes[5], genetic factors must also be involved in sporadic colorectal carcinogenesis.

There are various ways by which a family history of colorectal cancer may affect colorectal cancer risk. Two possibilities are that a family history of colorectal cancer conveys a genetic susceptibility that enhances the formation of new lesions or the transition from adenomas to carcinomas. Shared behavioural risk factors have also been proposed to underlie part of the association between a family history of colorectal cancer and risk of colorectal neoplasm.

Indications for effects of heritable factors on adenoma growth were found in a study in which small polyps were left in situ for three years before removal[6]: net adenoma growth was observed in nine out of 14 (64%) patients with a positive family history of colorectal cancer but only in 22 out of 73 (30%) of the patients without such a family history (odds ratio (OR)=3.9, 95% confidence interval (CI)=1.2-13.4). A stronger association with large adenomas (OR=2.1, $P<0.01$) than with small adenomas (OR=1.2, $P>0.10$) was indeed suggested in a case-control study on 362 patients[7]. Although the same impression can be obtained from smaller studies[8-11], a pooled analyses on 518 adenoma patients observed a stronger association between a family history of colorectal cancer and adenomas with no, mild, or moderate dysplasia (OR=1.5, 95% CI=1.1-2.1) than with adenomas with severe dysplasia, carcinoma in situ, or intramucosal carcinoma (OR=1.0, 95% CI=0.6-1.9). Nonetheless, the difference between these associations was not statistically significant[12].

The influence of genetic factors on adenoma multiplicity is clearly visible in patients with inherited polyposis syndromes[5]. Responsible genes such as Adenomatous Polyposis Coli (*APC*) or human MutY homologue (*MYH*) could have a more subtle influence in sporadic carcinogenesis[5]. Indeed, the expression of several genes in macroscopically normal rectosigmoid mucosa was altered in individuals with a sporadic family history of colon cancer compared with individuals without such history[13], and one study noted that the prevalence of multiple adenomas was higher in the presence of a positive family history[7].

We examined associations between family history of colorectal cancer, adenoma multiplicity and adenoma advancement in a large prospective cohort study, which allowed for adjustment for known behavioral risk factors for colorectal neoplasia. Evidence regarding such associations at the population level may yield clues to the search of genetic factors involved in the development of sporadic colorectal adenomas and cancer.

Material and Methods

Study Population

The current study comprises a subset of men of the Health Professionals Follow-Up Study (HPFS): an ongoing prospective study among 51,529 male US health professionals who responded to a mailed questionnaire in 1986 when they were between 40 and 75 years old. The Human Subjects Committee of the Harvard School of Public Health approved the HPFS.

At enrollment in 1986, and every two years thereafter, the cohort members were requested to fill out follow-up questionnaires to update information on various risk factors and to identify newly professionally diagnosed cases of various diseases. In 1986, 1990, 1994, 1998 and 2002, participants also completed a semi-quantitative food-frequency questionnaire. Questionnaires were mailed up to four times to non-responders. At the time of the 2002 questionnaire 37,433 men were alive and still participating.

Only men who completed the 1986 questionnaire and who underwent colonoscopy or sigmoidoscopy between 1986 and 2002 were eligible for the present analyses. Participants who reported to have had colorectal polyps, cancer (except non-melanoma skin cancer), ulcerative colitis or Crohn's disease before 1986 were excluded. We also excluded cohort members who reported implausible caloric intakes, *i.e.* smaller than 800 kcal/day or larger than 4,200 kcal/day) as well as those who left 70 more of the items blank on the food-frequency questionnaire. This resulted in a base population of 26,818 men.

Case Ascertainment

For each man who reported to have had an adenoma on a follow-up questionnaire for the first time, we asked for permission to request and review relevant medical records. A study investigator, who was blinded to exposure status, reviewed endoscopy and pathology reports. Only when the self-reported diagnosis was confirmed by a histopathological report, case status was assigned. The self-report of a negative endoscopy was reliable; a review of the medical records obtained from a random sample of 200 patients who reported a negative endoscopic result confirmed the absence of adenomas in all cases.

Adenomas in the cecum, ascending colon, hepatic flexure, transverse colon or splenic flexure were classified as being in the proximal colon. Adenomas in the descending or sigmoid colon were classified as being in the distal colon, and adenomas in the rectum or at the rectosigmoid junction were classified as rectal.

For diagnoses up to 1990, the number of adenoma, the size of the largest adenoma as assessed endoscopically and pathologically, and location of the most proximally located adenoma were recorded. For later diagnoses, number and size of the largest adenoma according to endoscopy and pathology reports were recorded for distal, proximal and rectal adenomas separately. In all years, only the most severe histological subtype of adenoma (in ascending order: tubular, tubulovillous, villous, carcinoma in situ) was registered.

Between 1986 and 2002, 4,102 of the participants were diagnosed with adenomas. For classification purposes, we used the adenoma size according to the endoscopy report, but data

from the pathology report were used if such information was lacking. Men having at least one adenoma ≥1 cm, with a (tubulo)villous structure or carcinoma in situ were classified as having advanced adenomas (n=1,254); men having only adenomas that were smaller than 1 cm and with no mention of a villous structure were classified as having non-advanced adenomas (n=1,675). No information on size and histopathology was present in the pathology and endoscopy reports of 955 adenomas, and the size of 218 tubular adenomas was unknown; these adenomas were not classified as being either advanced or non-advanced in the main analyses. Because not all men underwent colonoscopy, the classification according to adenoma multiplicity was based on distal and rectal adenoma only, under the assumption of a single adenoma where the number of adenomas was not mentioned in the endoscopy and pathology reports (single distal adenomas: n=1,794, multiple distal adenomas: n=542, no distal adenomas at all: n=23,644). The adenoma location was unknown for 838, and these adenoma patients were excluded from the analyses on multiplicity.

Assessment of Family History of Colorectal Cancer

The 1986 questionnaire included questions on the diagnosis of colorectal cancer and corresponding age (before age 50, age 50 to 59, age 60 to 69, age 70+, age unknown) in the father and mother separately. In 1990 and 1992 the men were asked whether any of the siblings had colorectal cancer, and this was also asked for the father and the mother. The information on the presence or absence of colorectal cancer in the family was updated in 1996 when a question on age of first diagnosis was included; the question referred to any parent, sibling and an additional sibling separately and the same categories were used as in 1986. Later questionnaires did not contain a question on family history of colorectal cancer. A report of colorectal cancer in first-degree relatives appears to be reliable[14, 15]. Nonetheless, we excluded 235 men who reported to have a positive family history of colorectal cancer at only one of the questionnaires but who did not report so on at least two subsequent questionnaires.

The remaining 26,583 men were classified as to whether at least one of their first-degree relatives was known with colorectal cancer and according to age of diagnosis of the youngest affected first-degree relative. For this classification, we used information up to 1996 because a family history of colorectal cancer can be regarded as a surrogate or indicator of inherited genetic susceptibility rather than a time-varying factor.

Assessment of Other Exposure Factors

The questionnaires, which have been described in detail elsewhere[16], requested information on age, race, height, weight, physical activity, use of aspirin, smoking history and habits, alcohol consumption and whether the men underwent either colonoscopy or sigmoidoscopy in the past two years. The 2004 questionnaire requested specifically whether the men ever had a colonoscopy, and if so in which time period. The 1996 questionnaire included questions on the numbers of biological brothers and sisters (0, 1, 2, 3, 4, ≥5). The semi-quantitative food frequency questionnaires, which included about 130 items and an open-ended question for

unlisted foods, covered more than 90% of the major nutrient intake of participants and inquired after vitamin and mineral supplements[16].

Derived nutrients, except alcohol, were adjusted for total energy intake using the residual method[17]. The average intake up to the date of diagnosis for cases and the date of last endoscopy for adenoma-free men was calculated to best represent long-term exposure and reduce within-person variation[18]. Subsequently, men were grouped into quantiles or categories according to the exposure factors of interest. Forty-eight people did not complete questions on BMI or physical activity and were excluded, which resulted in a final study population of 26,535 men (including 4,037 adenoma patients).

Statistical Analysis

To plot the prevalence of adenomas against age on a logarithmic scale, we determined whether an adenoma was present at the first endoscopy that was registered after study entry, and calculated the prevalence across strata of age. Linear regression, weighted by the inverse of the variance of the estimated proportion, was used to obtain the slope of the resulting curves. In a sub-analysis we excluded the men who reported having undergone an endoscopy before entering the study and those who underwent their first endoscopy for reasons other than routine screening.

Multinomial logistic regression was used to compare the distributions of a positive family history of colorectal cancer among patients having advanced adenoma and patients having non-advanced adenoma to evaluate etiological heterogeneity (case-case analyses). Within the same multinomial logistic regression model[19], we also compared both categories of patients with the people who did not report adenomas during the follow-up period. As a sensitivity analysis, we checked whether the associations remained similar when studying the group with single adenoma only. Whether the association between family history and advanced versus non-advanced adenomas depended on the location of the adenoma was also explored in the group of men with adenoma at only one location. We did so by fitting a model containing two indicator variables for location, the family history indicator variable and two cross-product terms of one of the indicator variables for location and family history. The *P*-value for interaction was obtained by comparing a model with and without the two cross-product terms using a likelihood ratio test.

Multinomial logistic regression was also used to compare the distributions of a positive family history of colorectal cancer among patients with multiple distal and single distal adenomas, and men who did not report any adenomas. Data from all adenoma patients were used when studying the association between the number of adenomas in the distal colon or rectum, but data from patients with at least one proximal adenoma were used when studying the association between a positive family history of colorectal cancer and the number of proximal adenomas because the entire colon was not always inspected.

The main models included total energy intake, age, history of endoscopy prior to study entry, routine screening versus other indications for any endoscopy, aspirin use, use of multivitamins, smoking, consumption of red meat, alcohol, intake of folate, calcium, BMI, and physical activity, in addition to a family history of colorectal cancer. As these presumed risk factors did not noteworthily affect the risk estimates corresponding to a family history of colorectal cancer,

we did not check whether inclusion of additional dietary and lifestyle factors affected the risk estimates. We also adjusted the models referring to affected siblings for the number of siblings (0, 1, 2, 3, 4, ≥5).

Additional sensitivity analyses evaluated whether the case-case associations depended on indication of endoscopy (screening versus complaints), race and age (continuous term) by including cross-products term in the model and evaluating them using a likelihood ratio test.

All reported *P*-values are two-sided. *P*-values<0.05 were considered statistically significant. The analyses were performed using SAS version 9.1 (SAS Institute, Inc., Cary, NC).

Results

Description of the Study Population

In 1986, 9.3 percent of the men reported a positive family history of colorectal cancer. This figure increased to 11.7% in 1990, 13.3% in 1992 and 15.3% in 1996, which most likely reflect ageing of the family members of the men.

One relative was affected of 92.1% of the 4,062 men with a positive family history of colorectal cancer (father: *n*=1,521; mother: *n*=1,275; sibling: *n*=638; at least one of the parents, but unknown who: *n*=306), whereas 7.7% had two types of affected relatives (father+mother: *n*=127; father+sibling: *n*=55; mother+sibling: *n*=59; at least one of the parents+sibling: *n*=71) and the mother, father and at least one sibling of 10 men (0.3%) were affected. Multiple siblings of 83 men were diagnosed with colorectal cancer.

Table 5-1 illustrates that dietary and lifestyle characteristics were comparable for men with and without a family history of colorectal cancer. Men with a family history of colorectal cancer however, were more likely to have undergone endoscopy before study entry, but they were less likely to have been examined for routine screening. A similar pattern was observed for the characteristics at study entry (not shown).

Description of Adenoma Characteristics

Of the study participants, 15.3% had at least one adenoma during the follow-up. A total of 1,721 (71.0%) patients had only adenomas that were classified as tubular, 527 (21.7%) had at least one tubulovillous adenoma but no villous adenoma or carcinoma in situ, and 145 (6.0%) had at least one villous adenoma but no carcinoma in situ and 32 (1.3%) men had carcinoma in situ. The histopathology was unknown for the remaining 1,612 adenoma patients. A total of 976 (34.8%) adenomas were 10 mm in diameter or larger and 1,832 were smaller than 10 mm. Size was unknown for 1,229 adenoma patients, and 912 of them also belonged to the group without information on histopathology. A total of 644 adenoma patients had adenomas in the rectum, 1,787 men had adenomas in the distal colon and 1,368 men had adenomas in the proximal colon. No information on location was available for 826 adenoma patients. A total of 528 men had multiple adenomas in the distal colon or rectum combined and 348 men were registered with multiple adenomas in the proximal colon.

Men with (tubulo)villous adenomas were more likely to have large adenomas than men with tubular adenomas (OR for tubulovillous *vs.* tubular=7.62, 95% CI=6.03-9.63; OR for

Table 5-1. Age-adjusted characteristics of the men of the Health Professionals Follow-Up Study according to the presence or absence of a positive family history of colorectal cancer*

Characteristic	Family history of colorectal cancer *n*=4,062	No family history of colorectal cancer *n*=22,473
Age at most recent endoscopy (mean ± sd, yr)	65.7 ± 9.2	64.9 ± 9.2
Race (%)		
Southern European	22.9	22.4
Northern European	71.1	72.7
Other	6.0	4.9
History of smoking (%)	52.7	53.0
Body mass index (mean, kg/m^2)	25.7	25.7
Physical activity (mean, met-h/w)†	31.4	31.3
Total energy intake (mean, kcal/d)	1,962	1,953
Alcohol intake (mean, g/d)	10.9	10.9
Mean dietary intake		
Carbohydrates (g/d)	247	247
Protein (g/d)	90.9	91.1
Fat (g/d)	68.2	68.2
Methionine (g/d)	2.13	2.13
Folate (μg/d)‡	545	544
Calcium (mg/d)‡	927	938
Vitamin D (IU/d)‡	451	453
Dietary fiber (g/d)	22.6	22.5
Red meat (servings/d)	0.55	0.55
Use of multivitamins (%)	53.8	54.6
Regular use of aspirin (%)§	40.8	42.5
History of endoscopy before study entry (%)	19.0	16.1
Any screening endoscopy (%)	74.4	80.9

* Updated variables are used for time-varying exposures (see method section). Mean values are presented. Missing values are excluded.

† Met-h: metabolic equivalent task hours: the ratio of the metabolic rate during activity to the resting metabolic rate[20].

‡ Includes usage of supplements.

§ Usage of ≥ 2 times per week.

villous *vs.* tubular=9.59, 95% CI=6.56-14.0). Men with adenomas in the distal colon were 1.63 (95% CI=1.34-2.01) times more likely and men with adenomas in the rectum were 1.97 (95% CI=1.67-2.33) times more likely to have advanced adenomas than men with proximal adenomas.

Men with multiple adenomas in the distal colon were 1.50 (95% CI=1.22-1.84) times more likely to be diagnosed with advanced adenomas than patients with single adenomas which was also the case when patients who had only distal adenomas were studied (OR=1.59, 95% CI=1.25-2.02).

Family History and Prevalence of Adenomas

Adenomas occurred more frequently among men who underwent a first endoscopy at older age than among men who were younger at first endoscopy (first panel of Figure 5-1). Adenomas were more common in men with a positive family history of colorectal cancer than in men without such a history in all age categories, but the higher prevalence in men with a positive

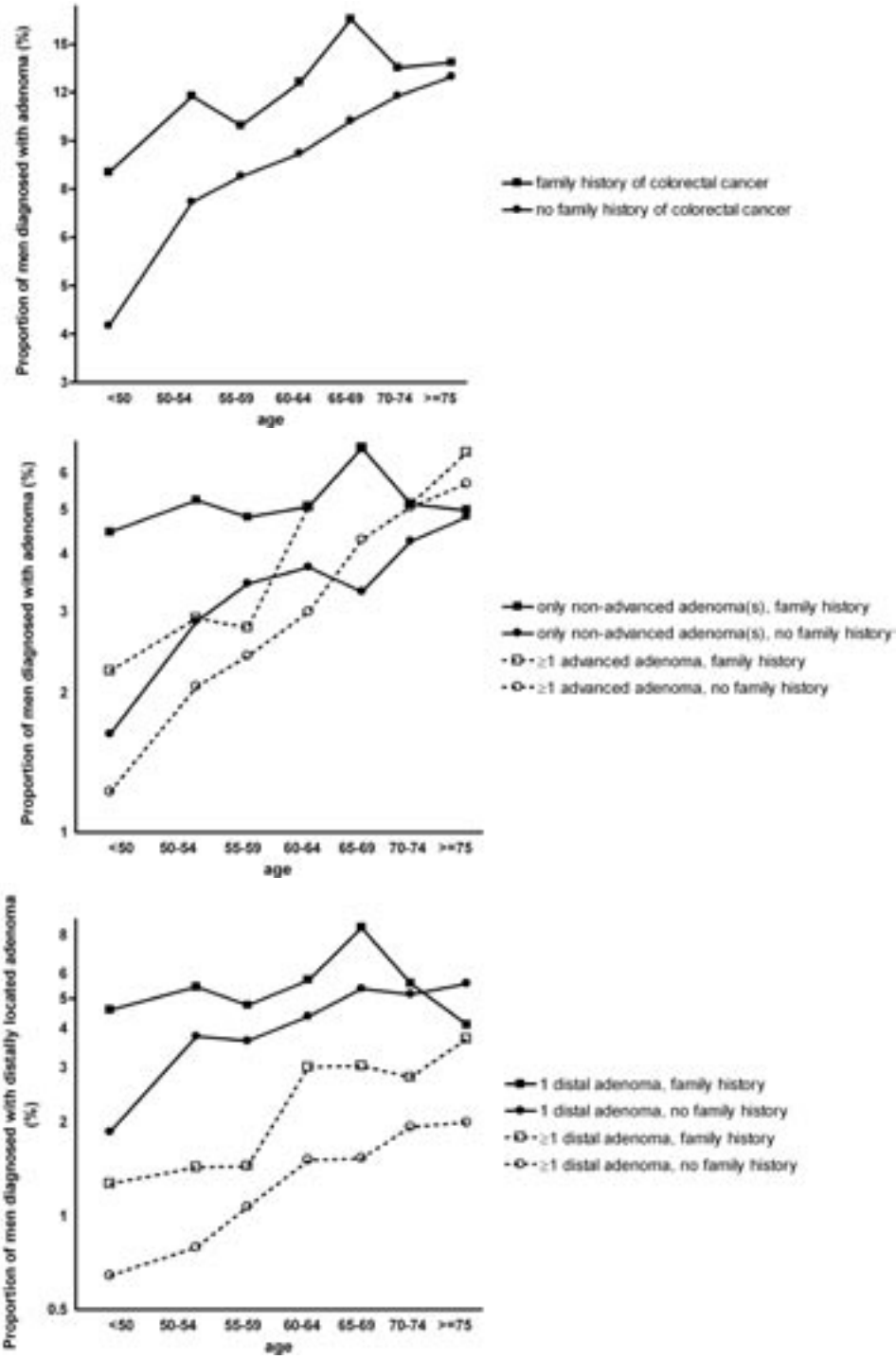


Figure 5-1. Prevalence of adenomas at first endoscopy after study entry across strata of age and family history of colorectal cancer.

family history was notably visible in the younger age categories. Apparent straight lines could be drawn independently in this plot for men with and without a family history of colorectal cancer, which suggests that adenoma prevalence increases as a power function of age. The slope of the weighted regression line was 1.0 for men with a positive family history of colorectal cancer and 2.3 for men without, which suggests that one of the rate-limiting steps in adenoma development has already occurred in men with a positive family history of colorectal cancer[21]. Therefore, the association between family history and adenomas appears stronger in the young age categories. A similar plot was obtained when we restricted the analyses to men who underwent their first endoscopy because of routine screening (not shown).

The aforementioned findings were supported by analyses in which we took variation in modifiable risk factors into account. Men with a positive family history of colorectal cancer were 1.91 times more likely to get colorectal adenomas than men without an affected first-degree relative (95% CI=1.75-2.07); this association was similar among men who underwent all registered endoscopies for routine screening (OR=1.89, 95% CI=1.71-2.07) and men who underwent at least one endoscopy for another reason (OR=1.98, 95% CI=1.65-2.38; $P_{interaction}$=0.51). The associations were similar for men with Southern European (OR=2.01, 95% CI=1.70-2.38), Northern European (OR=1.83, 95% CI=1.66-2.02), or other ethnic background (OR=2.07, 95% CI=1.44-2.96; $P_{interaction}$=0.56). As also visible in the figure, the strength of the association was stronger for men who were diagnosed at young age than for men who were diagnosed at older age (for those aged ≤55 years old OR=2.91, 95% CI=2.37-3.57; for those aged >55 years OR=1.75, 95% CI=1.60-1.92; $P_{interaction}$=0.0001).

The overall risk was somewhat higher for individuals with an affected sibling (OR=1.93, 95% CI=1.63-2.28) as for individuals with an affected parent (OR=1.75, 95% CI=1.60-1.92). Men with multiple affected relatives were more likely to be diagnosed than men with only one affected relative (OR=2.59, 95% CI=2.01-3.33; and OR=1.92, 95% CI=1.76-2.10, respectively).

Family History of Colorectal Cancer and Advanced and Non-Advanced Adenomas

The second panel of Figure 5-1 shows that non-advanced adenomas were more common than advanced adenomas up to age 65, after which advanced adenomas were more frequently found. The proportion of men with non-advanced adenomas also increases with age, but less strikingly so for non-advanced adenomas among men with a family history of colorectal cancer. Figure 5-1 suggests a steeper slope of prevalence with age for non-advanced adenomas among men without a family history of colorectal cancer than among men with a positive family history of colorectal cancer, which suggests a hereditary component in an early stage of carcinogenesis.

After adjustment for presumed risk factors, a family history of colorectal cancer was similarly associated with advanced and non-advanced adenomas (advanced versus non-advanced, OR=0.91, 95% CI=0.76-1.08; advanced versus adenoma-free: OR=1.84, 95% CI=1.60-2.12; non-advanced versus adenoma-free: OR=2.03, 95% CI=1.80-2.29) (Table 5-2). We could not detect differences in the strengths of the association when the whole study population was studied with respect to all family-related characteristics, although a stronger association with non-advanced

adenomas than with advanced adenomas was suggested among men having only adenomas in the distal colon or rectum. Indeed, the case-case analysis indicated interaction with location that was borderline statistically significant ($P_{interaction}$=0.068). A similar indication was found in the subgroup of men with only one adenoma (advanced versus non-advanced: OR= 0.80, 95% CI=0.63-1.01; advanced versus adenoma free: OR=1.54, 95% CI=1.26-1.87; non-advanced versus adenoma-free: OR=1.92, 95% CI=1.66-2.21). The difference between the associations with advanced and non-advanced adenomas did not seem to depend on race ($P_{interaction}$=0.29), having had screening endoscopy or not (P for interaction=0.46), or age ($P_{interaction}$=0.48). Among the men with adenomas, the distribution of adenoma size was comparable among those with and without a family history of colorectal cancer (median=6mm, 25th percentile=4mm, 75th percentile=10mm), although the non-parametric two-sample Wilcoxon test suggested that adenoma size was somewhat greater among those without a family history of colorectal cancer (P= 0.042).

Family History of Colorectal Cancer and Adenoma Multiplicity

The third panel of Figure 5-1 shows the distribution of multiple adenomas in the distal colon and rectum combined according to age groups and presence or absence of a family history of colorectal cancer. Across all age groups and both among men with and without a positive family history of colorectal cancer, single distal adenomas were more prevalent than multiple distal adenomas. The slope representing the prevalence of single distal adenomas among men with a family history of colorectal cancer was almost horizontal (suggesting only one rate limiting step), whereas the slope for men without a family history of colorectal cancer was slightly positive, but this may be largely due to the younger age groups. A similar pattern was observed regarding multiple distal adenomas, although the slopes were positive and the parallelism clearer.

The prevalence of multiple distal, but also single distal adenomas was higher among men with a positive family history of colorectal cancer, which is also reflected in the odds ratios presented in Table 5-3. The association with family history was stronger for multiple distal than for single distal adenomas. The difference in associations between family history and multiple distal adenomas, and family history and single distal adenomas, increased in relation to the number of affected relatives. The trend test also suggested that the association became stronger when family members were diagnosed at younger ages. The difference between the associations with multiple and single distal adenomas did not depend on race ($P_{interaction}$=0.72), having had screening endoscopy or not ($P_{interaction}$=0.68), and age ($P_{interaction}$=0.20; see Figure 5-1).

The association between family history and adenoma multiplicity was also clearly visible when we studied the number of adenomas. Of the 1,768 men with one distal adenoma 365 (20.6%) had a positive family history; this also applied to 99 of the 398 (24.9%) men with two distal adenomas, 24 of the 92 (26.1%) men with three distal adenomas, 8 of the 22 (36.4%) men with four distal adenomas, 4 of the 9 (44.4%) men with five distal adenomas, 3 of the 4 (75%) men with six distal adenomas, the only case with seven distal adenomas, the only case with eight distal adenomas, but not the only case with nine distal adenomas. A similar pattern was visible when the number of proximal adenomas was studied. Of the 957 men with one adenoma in the proximal colon

Table 5-2. Risk of advanced and non-advanced colorectal adenomas according to location or strength and type of family history of colorectal cancer (CRC)*

Category	Number of men†			Odds ratio (95% CI)		
	Advanced adenomas *n*=1,229	Non-advanced adenomas *n*=1,652	No reported adenomas *n*=22,498	Advanced *vs.* non-advanced adenomas	Advanced adenomas *vs.* no adenomas	Non-advanced adenomas *vs.* no adenomas
Risk among all men	277/952‡	397/1,255‡	3,137/19,361‡	0.91 (0.76-1.08)	1.84 (1.60-2.12)	2.03 (1.80-2.29)
Risk of having only adenoma in the distal colon§	166/497‡	110/417‡	-	0.72 (0.59-1.03)	1.70 (1.37-2.12)	2.19 (1.82-2.62)
Risk of having only adenomas in the rectum§	41/186‡	23/138‡	-	0.74 (0.42-1.29)	1.07 (0.69-1.67)	1.45 (1.03-2.06)
Risk of having only proximal adenomas§	138/406‡	73/175‡	-	1.19 (0.85-1.67)	2.53 (1.92-3.34)	2.12 (1.74-2.59)
Age at diagnosis of the youngest first-degree relative with CRC						
No family history	1,255	952	19,361	1 (reference)	1 (reference)	1 (reference)
Before age 50	156	97	1105	0.88 (0.50-1.54)	1.71 (1.08-2.69)	1.95 (1.33-2.85)
Age 50-59	112	82	904	1.06 (0.73-1.54)	2.16 (1.61-2.90)	2.04 (1.56-2.65)
Age 60-69	66	53	508	0.97 (0.72-1.30)	1.96 (1.54-2.48)	2.03 (1.65-2.49)
Age ≥70	31	21	251	0.80 (0.61-1.05)	1.83 (1.47-2.28)	2.27 (1.90-2.72)
P_{trend}				0.67	<0.0001	<0.0001
*Affected family member**						
Parent						
No	999	1,322	19,841	1 (reference)	1 (reference)	1 (reference)
Yes	230	330	2657	0.93 (0.77-1.12)	1.72 (1.47-2.00)	1.85 (1.62-2.10)
Sibling						
No	1,158	1,561	21,880	1 (reference)	1 (reference)	1 (reference)
Yes	71	91	618	0.98 (0.71-1.37)	1.96 (1.50-2.55)	1.99 (1.57-2.52)
Number of first-degree relatives with CRC						
None	952	1,255	19,361	1 (reference)	1 (reference)	1 (reference)
1	237	332	2,654	0.92 (0.76-1.11)	1.86 (1.60-2.16)	2.02 (1.77-2.30)
≥2	27	44	228	0.77 (0.47-1.26)	2.43 (1.62-3.67)	3.16 (2.26-4.40)
P_{trend}				0.22	<.0001	<.0001

* Models adjusted for age (in 5-yr age groups), history of endoscopy prior to study entry (yes/no), routine screening versus other indications for any endoscopy, aspirin use (currently, past or never ≥2 times/wk), use of multivitamins (current, former, never), smoking (never, quit ≤10 yr ago, quit>10 yr ago, current, missing) consumption of red meat (quintiles), alcohol (0, 0-<10, 10-<20, 20-<30, ≥30 g/d); intake of folate (quintiles), calcium (quintiles), BMI (quintiles), physical activity (quintiles), total energy intake (quintiles).

† Due to missing values, not all totals add up to 1,229, 1,652 or 22,498.

‡ Men with a family history of CRC / without family history of CRC.

§ Restricted to men with adenomas at one location only. Similar resultss were observed for men who had at least one adenoma at the studied location.

** Categories are not mutually exclusive. Estimates are mutually adjusted.

Table 5-3. Risk of multiple distal versus single distal colorectal adenomas according strength and type of family history of colorectal cancer (CRC)*

Category	Number of men†			Odds ratio (95% CI)		
	Multiple distal	Single distal	No distal adenomas	Multiple *vs.* single adenomas	Multiple distal *vs.* no distal adenomas	Single distal *vs.* no distal adenomas
	n=528	*n*=1,768	*n*=23,413			
Risk among all men	140/388‡	365/1,403‡	3,382/20,031‡	1.39 (1.11-1.74)	2.24 (1.83-2.73)	1.61 (1.42-1.82)
Age at diagnosis of the youngest first-degree relative with CRC						
No family history	388	1,403	20,031	1 (reference)	1 (reference)	1 (reference)
Before age 50	44	144	1,210	1.11 (0.77-1.58)	1.97 (1.44-2.72)	1.78 (1.48-2.14)
Age 50-59	43	100	970	1.57 (1.08-2.29)	2.48 (1.80-3.44)	1.58 (1.27-1.96)
Age 60-69	23	65	550	1.31 (0.80-2.13)	2.26 (1.47-3.49)	1.73 (1.33-2.26)
Age ≥70	11	29	275	1.38 (0.68-2.79)	2.13 (1.15-3.94)	1.55 (1.05-2.28)
P_{trend}				0.029	<0.0001	<0.0001
Affected family member§						
Parent						
No	411	1,464	20,552	1 (reference)	1 (reference)	1 (reference)
Yes	117	304	2,861	1.37 (1.08-1.75)	2.05 (1.66-2.54)	1.50 (1.31-1.71)
Sibling						
No	492	1,686	22,735	1 (reference)	1 (reference)	1 (reference)
Yes	36	82	678	1.31 (0.86-2.00)	2.12 (1.47-3.06)	1.62 (1.27-2.07)
Number of first-degree relatives with CRC						
None	388	1,403	20,031	1 (reference)	1 (reference)	1 (reference)
1	118	308	2,860	1.38 (1.09-1.76)	2.23 (1.81-2.76)	1.61 (1.42-1.84)
≥2	15	37	254	1.47 (0.79-2.72)	3.22 (1.88-5.52)	2.20 (1.54-3.13)
P_{trend}				0.0071	<0.0001	<0.0001

* Adjusted for the same factors as listed in Table 5-2.

† Due to missing values, not all totals add up to 528, 1,768 or 23,413.

‡ Men with a family history of CRC / without family history of CRC.

§ Categories are not mutually exclusive. Estimates are mutually adjusted.

230 (24.0%) had a positive family history; this applied to 65 of the 212 (30.7%) men with two adenomas in the proximal colon, 25 of the 84 (29.8%) men with three adenomas in the proximal colon, 8 of the 27 (29.6%) men with four adenomas in the proximal colon, 5 of the 12 (41.7%) men with five adenomas in the proximal colon, 1 of the 3 (33%) men with six adenomas in the proximal colon, 3 of the 5 (60%) men with seven adenomas in the proximal colon, but to none of the 2 men with 8 and none of the 3 men with 9 adenomas in the proximal colon.

Discussion

To our knowledge, this study is the first that systematically examined whether a family history of colorectal cancer influences adenoma multiplicity in a large number of asymptomatic and symptomatic men. Previous studies have reported on the prevalence of advanced adenomas in men with and without a positive family history of colorectal cancer, but these studies were small[8, 12] and only two of them studied this issue systematically[7, 12].

The most striking observation of this study is that a family history of colorectal cancer is more strongly associated with multiple (distal) adenomas than with single adenomas. A positive family history of colorectal cancer was not differently associated with risk of advanced and non-advanced adenomas, but a tendency towards a stronger influence on non-advanced adenomas was found.

By adjusting the associations for presumed modifiable risk factors for colorectal cancer, we largely excluded the explanation that shared environmental factors underlie the observed associations, which is in line with a study that compared incidence rates of colorectal cancer between siblings and spouses, and between parents and their offspring[22].

Not all participants underwent colonoscopy, which may have led to some bias in the analyses on advanced and non-advanced adenomas. However, since most polypectomies take place during colonoscopy as recommended[23], the impact of bias due to incomplete bowel examinations may be limited with regard to the case-case analyses. Nonetheless, some misclassification of case status has inevitably occurred. As for any adenoma study, some adenomas will have been missed at endoscopy, which are more likely to be small[24]. The miss rate will probably not affect the case-case analyses importantly because endoscopists are likely to examine the colon thoroughly once an adenoma has been diagnosed. Provided endoscopists do not conduct the examinations more thoroughly when a patient with a family history of colorectal cancer presents, the effect of missed adenomas is likely to be largely diluted by the much larger number of true adenoma-free men in the comparisons versus adenoma-free men.

Drawbacks of our study are that different physicians performed the endoscopies and that the classification of size and histopathological characteristics is based on judgement of different community pathologists. It can be argued that the classification of advanced adenomas should incorporate dysplasia as an important determinant of colorectal cancer risk[25], but we decided not to do so because the consistency of the classification into high-grade and low-grade dysplasia has been shown to be poor when different community pathologists are involved[26, 27]. The fair agreement of classification of histopathological types[26-28] between pathologists in combination with the strong association between size, histopathological characteristics and dysplasia[29-31]

further supports our classification. We did not categorize 29% of the adenomas as advanced or non-advanced adenomas however. The non-categorized adenomas were more often found to be tubular and thus more likely to be non-advanced; as one would expect as physicians may be less inclined to report observations deemed clinically unimportant than those deemed clinically important. Reassuringly, our conclusions regarding the case-case analyses are robust given that the OR for advanced versus non-advanced adenomas was 0.96 when we treated the non-classified adenomas as if they were non-advanced adenomas and 0.90 when we treated them as advanced adenomas, which were both not statistically significant.

In line with previous studies, we observed that the younger a person was diagnosed with any type of adenoma, the stronger was the observed association with family history of colorectal cancer. A stronger association was also found among men with young affected family members; hence, the etiology of adenomas and cancer occurring early in life appears to be more strongly determined by heritable factors than for those occurring later. In particular adenomas at younger age may have occurred among men who belong to a family with heritable colorectal cancer syndromes, but we are confident that our findings apply to sporadic carcinogenesis as only a few patients were diagnosed with a large number of adenomas and because less than 3% was diagnosed before age 50, whilst the mean age at last endoscopy was 65 years old.

Patients with hereditary colorectal cancer syndromes develop the disease at relatively young age, but colorectal cancer patients with a positive family history also tend to get the disease 10 years earlier than patients without such a history[3]. Advanced and non-advanced adenomas also occurred at a younger age among men with a family history of colorectal cancer in our study. This could point towards family members being more susceptible to the occurrence of mutations. The most striking observation of this study was that patients with a positive family history of adenomas were more likely to have multiple distal adenomas at diagnosis. This corresponds with findings from a small study in which the number of aberrant crypt foci was higher in patients with a positive family history of sporadic colorectal cancer than in patients without[32], and a case-control study also suggested that people with affected relatives were more likely to have multiple adenomas[7]. Perhaps the same susceptibility genes[33] that modify the severity of the familial adenomatous polyposis coli (FAP) syndrome or other polyposis syndromes may also determine the association between a positive family history and adenoma multiplicity in sporadic adenoma patients.

If the genetic component involved in sporadic colorectal carcinogenesis mostly drives the occurrence of key mutations rather than enhancing the growth signals in prevalent adenomas in the gut, it is not surprising that the association between family history of colorectal cancer and multiplicity is stronger than the one with advanced adenoma stage. Indeed, in a cross-sectional study a family history of colorectal cancer was associated with an approximately two-fold higher recurrence rate within three years, whilst no difference in distribution of family history could be detected between patients with and without adenomas at study entry[34] These observations suggest that genetic factors may play a more important role in the earlier stages of the adenoma-carcinoma sequence than in later stages.

However, it cannot be precluded that heritable factors stimulate growth of minuscule adenomas rather than enhancing the occurrence of new lesions. It seems less likely that genetic factors stimulate adenoma growth importantly because the ratio of advanced versus non-advanced adenomas was similar for people with and without a family history of colorectal cancer across all examined strata in this study; similar observations were observed in a pooled analysis that compared adenomas according to degree of dysplasia[12]. On the other hand, a study in which small adenomas were left in situ for up to three years, patients with a first-degree family member with sporadic colorectal cancer were more likely to have adenomas that showed net growth than patients without such a family history[6], and relatively small studies[8-11] as well as a larger case-control study[7] suggested that larger adenomas were more common among those with relatives with colorectal cancer. In the graphs depicting adenoma prevalence, a less steep curve was visible among men with a family history of colorectal cancer diagnosed before age 55. This supports the idea that there is a hereditary subgroup of adenomas that develops faster than the majority. In hereditary non-polyposis colorectal cancer (HNPCC) patients, adenomas are diagnosed less frequently while colorectal cancer is common and occurs at an early age[35]. This suggests that HNPCC adenomas pass quicker through the adenoma-carcinoma sequence than do other adenomas, and this may be the case for other, yet unidentified adenoma subgroups. Likewise, a group of carcinomas may exist with similar potential properties, which may or may not originate from aggressive adenomas. Nonetheless, in the general population, a family history of colorectal cancer seems to play a more important role in the early stages of adenoma development than in later stages. Considering that some adenomas might never turn into a more advanced lesion because they regress in size[36], we cannot conclude that all adenomas have the potential to develop into a carcinoma provided necessary growth conditions will be fulfilled. The selection process that determines the transition of non-advanced to advanced adenomas, however, does not seem to be influenced by heritable factors strongly related to a family history of colorectal cancer.

In conclusion, our data suggest that adenoma multiplicity seems to have a hereditary basis in sporadic colorectal carcinogenesis, but we could not confirm a role of family history in adenoma advancement.

Acknowledgements

The HPFS is supported by NCI Research Grant CA 55075. Petra Wark's visit to the Harvard School of Public Health was made possible with support from the Dutch Cancer Society, and she was further supported by the Netherlands Organisation for Health Research and Development.

References

1. Johns LE, Houlston RS. A systematic review and meta-analysis of familial colorectal cancer risk. *Am J Gastroenterol.* 2001;96(10):2992-3003.

2. Butterworth AS, Higgins JP, Pharoah P. Relative and absolute risk of colorectal cancer for individuals with a family history: a meta-analysis. *Eur J Cancer.* 2006;42(2):216-227.

3. Fuchs CS, Giovannucci EL, Colditz GA, Hunter DJ, Speizer FE, Willett WC. A prospective study of family history and the risk of colorectal cancer. *N Engl J Med.* 1994;331(25):1669-1674.

4. 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.

5. De la Chapelle A. Genetic predisposition to colorectal cancer. *Nat Rev Cancer.* 2004;4(10):769-780.

6. Almendingen K, Hofstad B, Vatn MH. Does a family history of cancer increase the risk of occurrence, growth, and recurrence of colorectal adenomas? *Gut.* 2003;52(5):747-751.

7. Boutron MC, Faivre J, Quipourt V, Senesse P, Michiels C. Family history of colorectal tumours and implications for the adenoma-carcinoma sequence: a case control study. *Gut.* 1995;37(6):830-834.

8. Bazzoli F, Fossi S, Sottili S, et al. The risk of adenomatous polyps in asymptomatic first-degree relatives of persons with colon cancer. *Gastroenterology.* 1995;109(3):783-788.

9. Tung SY, Wu CS. Risk factors for colorectal adenomas among immediate family members of patients with colorectal cancer in Taiwan: a case-control study. *Am J Gastroenterol.* 2000;95(12):3624-3628.

10. Pariente A, Milan C, Lafon J, Faivre J. Colonoscopic screening in first-degree relatives of patients with 'sporadic' colorectal cancer: a case-control study. The Association Nationale des Gastroenterologues des Hopitaux and Registre Bourguignon des Cancers Digestifs (INSERM CRI 9505). *Gastroenterology.* 1998;115(1):7-12.

11. Guillem JG, Forde KA, Treat MR, Neugut AI, O'Toole KM, Diamond BE. Colonoscopic screening for neoplasms in asymptomatic first-degree relatives of colon cancer patients. A controlled, prospective study. *Dis Colon Rectum.* 1992;35(6):523-529.

12. Terry MB, Neugut AI, Bostick RM, et al. Risk factors for advanced colorectal adenomas: a pooled analysis. *Cancer Epidemiol Biomarkers Prev.* 2002;11(7):622-629.

13. Hao CY, Moore DH, Wong P, Bennington JL, Lee NM, Chen LC. Alteration of gene expression in macroscopically normal colonic mucosa from individuals with a family history of sporadic colon cancer. *Clin Cancer Res.* 2005;11(4):1400-1407.

14. Kerber RA, Slattery ML. Comparison of self-reported and database-linked family history of cancer data in a case-control study. *Am J Epidemiol.* 1997;146(3):244-248.

15. Murff HJ, Spigel DR, Syngal S. Does this patient have a family history of cancer? An evidence-based analysis of the accuracy of family cancer history. *Jama.* 2004;292(12):1480-1489.

16. Rimm EB, Giovannucci EL, Stampfer MJ, Colditz GA, Litin LB, Willett WC. Reproducibility and validity of an expanded self-administered semiquantitative food frequency questionnaire among male health professionals. *Am J Epidemiol.* 1992;135(10):1114-1126; discussion 1127-1136.

17. Willett W. *Nutritional Epidemiology.* 2nd ed. New York, NY: Oxford University Press; 1998.

18. Hu FB, Manson JE, Liu S, et al. Prospective study of adult onset diabetes mellitus (type 2) and risk of colorectal cancer in women. *J Natl Cancer Inst.* 1999;91(6):542-547.

19. Begg CB, Gray R. Calculation of Polychotomous Logistic-Regression Parameters Using Individualized Regressions. *Biometrika.* 1984;71(1):11-18.

20. Ainsworth BE, Haskell WL, Leon AS, et al. Compendium of physical activities: classification of energy costs of human physical activities. *Med Sci Sports Exerc.* 1993;25(1):71-80.

21. Armitage P, Doll R. The age distribution of cancer and a multi-stage theory of carcinogenesis. 1954. *Int J Epidemiol.* 2004;33(6):1174-1179.

22. Hemminki K, Chen B. Familial risk for colorectal cancers are mainly due to heritable causes. *Cancer Epidemiol Biomarkers Prev.* 2004;13(7):1253-1256.

23. Bond JH. Polyp guideline: diagnosis, treatment, and surveillance for patients with colorectal polyps. Practice Parameters Committee of the American College of Gastroenterology. *Am J Gastroenterol.* 2000;95(11):3053-3063.

24. Rex DK, Cutler CS, Lemmel GT, et al. Colonoscopic miss rates of adenomas determined by back-to-back colonoscopies. *Gastroenterology.* 1997;112(1):24-28.

25. Winawer SJ, Zauber AG, Fletcher RH, et al. Guidelines for colonoscopy surveillance after polypectomy: a consensus update by the US Multi-Society Task Force on Colorectal Cancer and the American Cancer Society. *CA Cancer J Clin.* 2006;56(3):143-159; quiz 184-145.

26. Jensen P, Krogsgaard MR, Christiansen J, Braendstrup O, Johansen A, Olsen J. Observer variability in the assessment of type and dysplasia of colorectal adenomas, analyzed using kappa statistics. *Dis Colon Rectum.* 1995;38(2):195-198.

27. Terry MB, Neugut AI, Bostick RM, Potter JD, Haile RW, Fenoglio-Preiser CM. Reliability in the classification of advanced colorectal adenomas. *Cancer Epidemiol Biomarkers Prev.* 2002;11(7):660-663.

28. Costantini M, Sciallero S, Giannini A, et al. Interobserver agreement in the histologic diagnosis of colorectal polyps. the experience of the multicenter adenoma colorectal study (SMAC). *J Clin Epidemiol.* 2003;56(3):209-214.

29. O'Brien MJ, Winawer SJ, Zauber AG, et al. The National Polyp Study. Patient and polyp characteristics associated with high-grade dysplasia in colorectal adenomas. *Gastroenterology.* 1990;98(2):371-379.

30. Gschwantler M, Kriwanek S, Langner E, et al. High-grade dysplasia and invasive carcinoma in colorectal adenomas: a multivariate analysis of the impact of adenoma and patient characteristics. *Eur J Gastroenterol Hepatol.* 2002;14(2): 183-188.

31. Jørgensen OD, Kronborg O, Fenger C. The Funen Adenoma Follow-Up Study. Characteristics of patients and initial adenomas in relation to severe dysplasia. *Scand J Gastroenterol.* 1993;28(3):239-243.

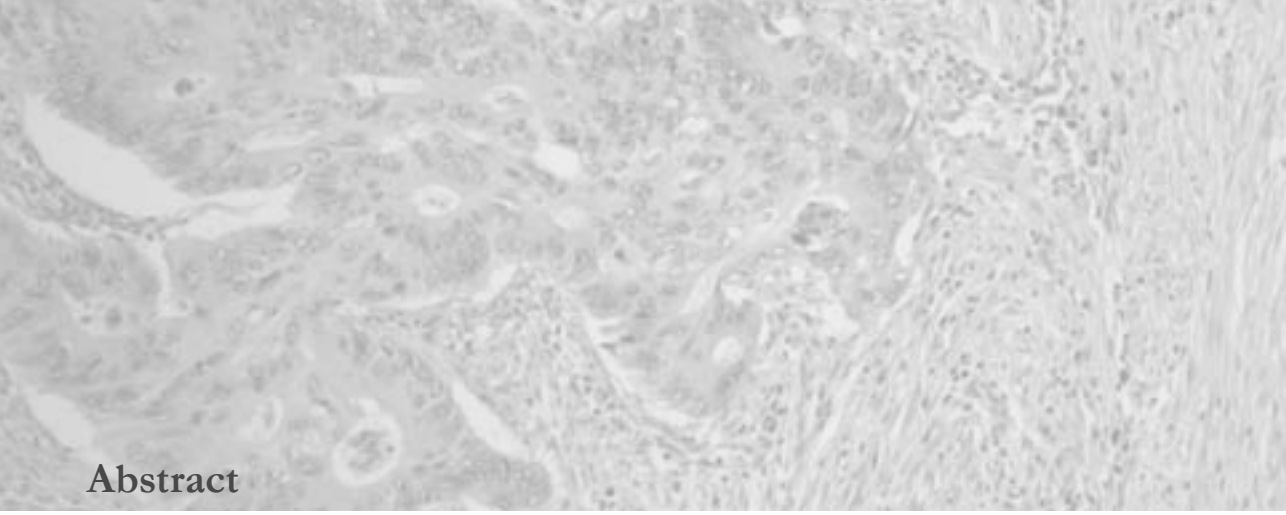
32. Stevens RG, Swede H, Heinen CD, et al. Aberrant crypt foci in patients with a positive family history of sporadic colorectal cancer. *Cancer Lett.* 2006.

33. Houlston R, Crabtree M, Phillips R, Crabtree M, Tomlinson I. Explaining differences in the severity of familial adenomatous polyposis and the search for modifier genes. *Gut.* 2001;48(1):1-5.

34. Fossi S, Bazzoli F, Ricciardiello L, et al. Incidence and recurrence rates of colorectal adenomas in first-degree asymptomatic relatives of patients with colon cancer. *Am J Gastroenterol.* 2001;96(5):1601-1604.

35. Ahlquist DA. Aggressive polyps in hereditary nonpolyposis colorectal cancer: targets for screening. *Gastroenterology.* 1995;108(5):1590-1592.

36. Loeve F, Boer R, Zauber AG, et al. National Polyp Study data: evidence for regression of adenomas. *Int J Cancer.* 2004;111(4):633-639.



Abstract

Background: Clinical and pathologic differences exist between colon carcinomas deficient and -proficient in the mismatch repair protein hMLH1. Animal and *in vitro* studies suggest that fruits, vegetables, folate, and antioxidants are associated with colonic expression of mismatch repair genes.

Methods: Associations between consumption of fruits and vegetables and hMLH1 protein-deficient and -proficient colon cancer were evaluated in the Netherlands Cohort Study on diet and cancer using a case-cohort approach. A self-administered food-frequency questionnaire was completed, in 1986, by 120,852 individuals ages 55 to 69 years. Using immunohistochemistry, hMLH1 protein expression was assessed in colon cancer tissue obtained from 441 patients who were identified over 7.3 years of follow-up excluding the initial 2.3 years. Incidence rate ratios (RR) were estimated for hMLH1 protein-deficient and -proficient colon cancer.

Results: hMLH1 protein expression was absent in 54 tumors (12.2%) and present in 387 tumors. Fruit consumption was associated with hMLH1 protein-deficient colon cancer (highest versus lowest tertile: RR=0.46, 95% confidence interval (CI)=0.23-0.90, P_{trend}=0.029) but not with hMLH1 protein-proficient tumors (highest versus lowest tertile: RR=1.03, 95% CI=0.78-1.35, P_{trend}=0.81). Total consumption of vegetables was not associated with either type of tumor (hMLH1 protein-deficient: RR=0.86, 95% CI=0.45-1.65, P_{trend}=0.67; hMLH1 protein-proficient: RR=0.94, 95% CI=0.72-1.23, P_{trend}=0.72). No associations were observed for folate, fiber, antioxidants, or subgroups of vegetables.

Conclusion: These analyses indicate that an inverse association between consumption of fruits and colon cancer may be confined to the subgroup of tumors with a deficient mismatch repair system.

Fruits, Vegetables, and hMLH1 Protein-Deficient and -Proficient Colon Cancer: The Netherlands Cohort Study

6

Petra A. Wark
Matty P. Weijenberg
Pieter van 't Veer
Gerda van Wijhe
Margreet Lüchtenborg
Goos N.P. van Muijen
Anton F.P.M. de Goeij
R. Alexandra Goldbohm
Piet A. van den Brandt

Cancer Epidemiol Biomarkers Prev. 2005;14(7):1619-25

Introduction

Between 80% and 90% of colorectal carcinomas are thought to arise via the chromosome instability pathway, which is associated with mutations in adenomatous polyposis coli (*APC*), Kirsten-ras (*K-ras*), and *p53* genes. About 10% to 20% of colon carcinomas appear to be arising via the hypermutability pathway involving deficiency of DNA mismatch repair (MMR) enzymes. This pathway is associated with microsatellite instability (MSI) and mutations in genes such as insulin-like growth factor II receptor (*IGFIIR*) and Bcl-2-associated X protein (*BAX*; ref. 1). In sporadic colorectal cancer, *hMLH1* gene inactivation by promoter hypermethylation is thought to be the main mechanism behind MMR deficiency[2-8].

Carcinomas showing MMR deficiency have clinical and pathologic features that distinguish them from carcinomas with MMR proficiency. More often located in the proximal colon and showing a mucinous or undifferentiated histology and prominent lymphocytic tumor infiltration, MMR-deficient carcinomas have a relatively favorable prognosis and respond differently to chemotherapeutic agents compared with MMR-proficient carcinomas[9-13]. Given these distinct tumor features, different sets of dietary and lifestyle factors may be involved in the etiology of MMR-deficient and -proficient colorectal carcinomas.

Fruits and vegetables are candidates for such potential differential effects. Folate, which is present in, among others, green leafy vegetables and citrus fruits, can affect *S*-adenosylmethionine levels, which regulate DNA methylation[14]. Antioxidants could have an influence because MMR enzymes are involved in repair of oxidative DNA damage, and MSI was reduced in MMR-deficient colon cells growing in the presence of ascorbate[15]. Moreover, in a case-control study, fruit consumption was inversely associated with MSI-high colon carcinomas with hypermethylated *hMLH1* but was not associated with MSI-high colon carcinomas without hypermethylated *hMLH1*[16]. In another case-control study, consumption of vegetables was inversely associated with colon carcinomas without MSI[17]. However, an inverse association between consumption of vegetables and colon carcinomas with MSI was also observed, although not statistically significant[17].

In the Netherlands Cohort Study on diet and cancer, we examined whether consumption of fruits and vegetables is differentially associated with hMLH1 protein-deficient and -proficient colon cancer.

Material and Methods

Cohort

The Netherlands Cohort Study on diet and cancer is a population-based prospective study that was initiated in September 1986. At that time, 58,279 men and 62,573 women ages 55 to 69 years completed a mailed, self-administered questionnaire on dietary habits and other potential risk factors for cancer. After baseline exposure assessment, a subcohort of 3,500 people (1,688 men and 1,812 women) was randomly selected and their vital statuses were followed up biennially to estimate accumulated person-years[18].

Follow-Up for Cancer

Cancer follow-up consists of annual linkage of the entire study database to the Netherlands Cancer Registry and the nationwide registry of cytopathology and histopathology (PALGA). The completeness of the follow-up procedure was estimated to be nearly 100%[19, 20]. No subcohort members were lost to follow-up.

The initial 2.3 years of follow-up were excluded because of incomplete coverage of PALGA and to exclude misclassification of exposure related to potential subclinical disease. Subcohort members with prevalent cancer other than non-melanoma skin cancer were also excluded, resulting in a subcohort size of 3,264 individuals.

Between 1989 and 1994, 929 incident cases with histologically confirmed colorectal carcinomas were identified, 819 of which (88%) could be identified in the PALGA database; data from PALGA were needed to identify in which pathologic laboratory tumor tissue was stored. Tumor characteristics (Dukes' stage, sublocalization, and differentiation grade) were obtained from the database of the Netherlands Cancer Registry. We classified cancers as proximal colon cancer (*International Classification of Diseases for Oncology, first edition* codes 153.0, 153.1, 153.4, 153.5, and 153.6), distal colon cancer (codes 153.2, 153.3, and 153.7), rectosigmoid (code 154.0), and rectal cancer (code 154.1).

Tumor Tissue Samples

After approval by the Ethical Review Boards of Maastricht University, PALGA, and the Netherlands Cancer Registry, we were able to retrieve formalin-fixed, paraffin-embedded colorectal tumor tissue from 776 of the 819 patients (95%) from 54 pathology laboratories throughout the country. Thirty-nine tumor blocks could not be used: 20 contained only normal mucosa, 10 turned out to be adenomas, three were non-colorectal malignancies, and tissue was sparse in six. The remaining 737 (95%) specimens were distributed as follows: proximal colon (n=240), distal colon (n=224), colon cancer not otherwise specified (n=12), rectosigmoid (n=85), and rectum (n=176).

hMLH1 Immunohistochemical Analyses

Immunohistochemical analyses were done on 4-μm sections of formalin-fixed, paraffin-embedded colorectal cancer tissue and adjacent normal tissue using monoclonal antibody against hMLH1 (clone G168-15; dilution 1:100; BD PharMingen International/Becton Dickinson, San Diego, CA) as previously described[16]. Absence of hMLH1 protein expression was scored upon the absence of any nuclear staining in tumor cells in combination with a signal in positive internal control tissue (normal epithelial cells, stromal cells, muscle cells, and lymphocytes). Expression of hMLH1 was considered present in case of a nuclear signal in tumor cells and positive internal control tissue. Two investigators (P.A.W. and G. van W.) reviewed the immunostaining independently. Discrepancies were re-examined and discussed with a pathologist until consensus was reached.

hMLH1 expression status was determined successfully in 725 (98%) of the 737 samples: 234 occurred in the proximal colon, 222 in the distal colon, 12 were colon cancers not otherwise specified, 84 occurred in the rectosigmoid, and 173 in the rectum.

Restriction to Colon Cancer

As colon and rectum cancer are considered two distinct disease entities[21, 22], have a different etiology[23], and the rectosigmoid can be considered a clinically applied term rather than an anatomically defined transitional zone between the colon and rectum, it would be ideal to stratify analyses to site (colon, rectum, and rectosigmoid). Because hMLH1 protein expression was detected in cancer tissue specimens of all 84 patients with rectosigmoid cancer and in 169 of the 173 rectal cancer specimens, such analyses were not feasible. Therefore, we restricted the present analyses to the 468 colon cancer patients.

Exposure Assessment

At baseline, a 150-item semi-quantitative food-frequency questionnaire was given to assess habitual consumption of foods and beverages in the year preceding the start of the study[24, 25]. The assessment of fruits and vegetables was described in detail previously[26] and covered all types of fruits and vegetables regularly eaten in 1986 with the exception of chicory, red cabbage, cucumber, and broccoli; broccoli was hardly sold in the Netherlands at that time. The Dutch Food Composition Table was used to calculate the mean daily intake of vitamin C and dietary fiber[27]. Intake of α-carotene, β-carotene, lutein/zeaxanthin, β-cryptoxanthin, and lycopene was calculated as described by Goldbohm *et al.*[28]. Dietary intake of folate was calculated using data from a validated liquid chromatography trienzyme method[29] used to analyze the 125 most important foods contributing to folate intake in the Netherlands[30].

The food-frequency questionnaire was validated against a 9-day dietary record, yielding correlation coefficients of 0.60 for fruits, 0.38 for vegetables, 0.74 for dietary fiber, and 0.58 for vitamin C[24]. Consumption of fruits was, on average, underestimated (mean daily consumption based on dietary records: 207 g, on questionnaires: 189 g) and consumption of vegetables was overestimated by the questionnaire compared with the dietary record (records: 160 g/d, questionnaires: 189 g/d; ref. 24). The reproducibility of the questionnaire was determined from five annually repeated questionnaire administrations in five independent random samples of 400 cohort members, resulting in a test-retest correlation averaged over all nutrients of 0.66. The average decline in correlation amounted to 0.07 after five years[25].

Study Population

Twenty-seven colon cancer cases and 215 subcohort members who left ≥60 items blank on the food-frequency questionnaires and also reported eating <35 items at least once a month, or who left one or more item blocks (groups of items, *e.g.* beverages) empty were excluded[24]. Because questions on vegetables appeared early in the questionnaire, some participants were more prone to make errors on this item block. When more than three errors were made in the various vegetable questions, the participant was excluded from the analyses on vegetables. This

resulted in a final study population of 441 colon cancer cases and 3,048 subcohort members for consumption of fruits and nutrients, and 422 colon cancer cases and 2,884 subcohort members for consumption of vegetables.

Data Analysis

We examined the distributions of fruits, vegetables, several of their subgroups, related nutrients, and supplements (listed in Table 6-2). Because 51% of the study population indicated to drink fruit juices other than fresh orange juice less than once per month, these were not examined separately.

Cases and subcohort members were divided into tertiles based on the distribution among subcohort members. First, distributions among patients with hMLH1 protein-deficient and -proficient colon cancer were compared using logistic regression to evaluate etiologic heterogeneity (case-case analyses). Thereafter, Cox proportional hazard models were used to compute incidence rate ratios and 95% confidence intervals for hMLH1 protein-deficient and -proficient colon cancer separately. In the models, the total person-years at risk were estimated from the subcohort[31] and the robust Huber-White sandwich estimator was used to estimate SEs, accounting for the additional variance introduced by sampling from the cohort. The proportional hazard assumption was evaluated by visual examination of plots of scaled Schoenfeld residuals versus time in combination with a formal test[32, 33] and was met for all presented models.

The standard model included gender, age, family history of colorectal cancer, and total energy intake (as continuous variable). Body mass index, physical activity; consumption of fresh (red) meat, meat products, fish, and alcohol; intake of total fat, calcium, and methionine; use of hormone replacement therapy (set to zero for men and non-users); and smoking (in separate analyses: current/ever/never, duration of smoking, frequency of smoking, and pack-years) were considered as potential confounders. None of these factors changed the rate ratios with >10% and were therefore not added to the standard model. An ordinal score value based on the median value within each tertile in the subcohort was used to test for trend across tertiles.

Three types of sensitivity analyses were conducted. First, subanalyses on proximal colon cancers were conducted to evaluate whether differential associations of fruits and vegetables with hMLH1 protein-deficient and -proficient colon cancer could be attributed to differences in tumor location. Second, we checked whether associations remained similar when restricting to individuals without a positive family history of colorectal cancer in whom *hMLH1* germline mutations are unlikely to occur. Third, we evaluated whether the associations were driven by the higher consumption of fruits and vegetables by women by using gender-specific tertiles. All reported *P*-values are two-sided and considered statistically significant if <0.05. The analyses were conducted using Intercooled STATA for Windows 8.0 (Stata Corp., College Station, TX).

Results

Table 6-1 describes baseline characteristics of the study population. Cases were older than subcohort members and they, especially those with hMLH1 protein-proficient colon cancers ($hMLH1^+$), more often had a positive family history of colorectal cancer. In comparison with

Table 6-1. Baseline demographic, lifestyle, dietary and tumor characteristics of patients with hMLH1 protein-deficient (hMLH1$^-$) and -proficient (hMLH1$^+$) colon cancer and subcohort members

Characteristic	Colon cancer cases		Subcohort
	hMLH1$^-$	hMLH1$^+$	
	n=54	*n* =387	*n* =3,048
Demographic and lifestyle factors			
Age (yr), mean ± SD	62.8 ± 4.5	63.0 ± 4.0	61.4 ± 4.2
Gender, *n* (% men)	22 (40.7%)	216 (55.8%)	1,475 (48.4%)
Family history of colorectal cancer, *n* (%)	5 (9.3%)	49 (12.7%)	170 (5.6%)
Body mass index (kg/m^2), mean ± SD*	25.6 ± 3.5	25.6 ± 3.2	25.1 ± 3.1
Physical activity, > 90 min/d, *n* (%)†	20 (37.0%)	92 (24.2%)	806 (26.8%)
Cigaret smoking status			
Never, *n* (%)	18 (33.0%)	144 (37.2%)	1,128 (37.0%)
Ex-smoker, *n* (%)	20 (37.0%)	174 (45.0%)	1,074 (35.2%)
Smoker, *n* (%)	16 (29.6%)	69 (17.8%)	846 (27.8%)
Pack-years among (ex) smokers, median (p25;p75)‡	26.7 (10.9; 32.4)	20.0 (7.8; 35.3)	18.2 (8.5; 32.0)
Alcohol (g/d), median (p25;p75)§	1.8 (0.00; 15.0)	4.0 (0.00; 18.0)	4.2 (0.25; 14.5)
*Dietary characteristics***			
Total energy intake (kJ/d), mean ± SD	7,505 ± 1,718	8,056 ± 2,063	8,029 ± 2,168
Protein (g/d), mean ± SD	66.4 ± 13.8	70.9 ± 16.4	70.6 ± 17.1
Carbohydrates (g/d), median (p25; p75)	176.8 (150.1; 206.9)	191.7 (160.5; 239.5)	193.6 (158.2; 236.2)
Fat (g/d), median (p25; p75)	79.7 ± 24.5	84.0 ± 26.3	83.9 ± 27.8
Fresh meat (g/d), median (p25;p75)	93.2 (77.5; 121.6)	99.3 (75.7; 121.8)	97.8 (74.5; 123.8)
Fish and shellfish (g/d), median (p25;p75)	6.2 (0.0; 13.2)	6.9 (0.0; 19.9)	8.5 (0.0; 19.9)
Calcium (mg/d), median (p25;p75)	865.6 (657.0; 1,036)	908.8 (694.0; 1,107)	888.6 (700.0; 1,106)
Methionine (mg/d), mean ± SD	1,520 ± 337	1,607 ± 383	1,599 ± 406
Tumor-related characteristics			
Proximally located, *n* (%)††	48 (88.9)	171 (45.4)	-
Dukes' stages C or D, *n* (%)‡‡	14 (26.9)	158 (44.3)	-
Poorly or undifferentiated tumor, *n* (%)§§	18 (40.9)	60 (17.5)	-
Age at diagnosis, mean ± SD	68.6 ± 4.8	68.5 ± 4.1	-

* For body mass index, figures are based on *n*=54 for hMLH1$^-$, *n*=374 for hMLH1$^+$ and *n*=2,948 subcohort members.

† For physical activity, figures are based on *n*=54 for hMLH1$^-$, *n*=381 for hMLH1$^+$ and *n*=3,007 subcohort members.

‡ For pack-years, figures are based on *n*=33 for hMLH1$^-$, *n*=227 for hMLH1$^+$ and *n*=1,779 subcohort members.

§ For alcohol, figures are based on *n*=54 for hMLH1$^-$, *n*=383 for hMLH1$^+$ and *n*=2,957 subcohort members.

** Other than fruits and vegetables and related nutrients.

†† For ten hMLH1+ colon cancers, the exact location is unknown.

‡‡ Dukes' stage is known for 52 hMLH1$^-$ and 357 hMLH1$^-$ colon cancers.

§§ Differentiation grade is known for 44 hMLH1$^-$ and 342 hMLH1$^+$ colon cancers.

Table 6-2. Consumption of fruits, vegetables, related micronutrients and supplements among patients with hMLH1 protein-deficient (hMLH1$^-$) and -proficient (hMLH1$^+$) colon cancer and subcohort members

	Colon cancer cases		Subcohort
	hMLH1$^-$	hMLH1$^+$	
Fruits, median (p25; p75)	*n*=54	*n*=387	*n*=3,048
Total fruit consumption (g/d)[*,†]	120.3 (69.6; 203.0)	159.4 (89.3; 237.4)	156.8 (95.1; 235.1)
Citrus fruits and fresh citrus juices (g/d)	59.5 (11.4; 91.5)	64.2 (21.4; 107.7)	59.7 (20.7; 110.5)
Apples and pears (g/d)[‡]	51.3 (17.1; 97.2)	80.1 (25.6; 115.8)	80.1 (25.0; 119.7)
Other fruits (g/d)	14.5 (6.0; 26.2)	14.3 (6.3; 29.6)	15.6 (6.5; 29.4)
Vegetables, median (p25; p75)	*n*=54	*n*=368	*n*=2,884
Total vegetable consumption (g/d)[§]	181.0 (133.8; 237.6)	173.2 (130.0; 227.8)	177.9 (136.8; 227.9)
Cruciferous vegetables (g/d)[**]	29.0 (19.9; 38.4)	27.2 (17.7; 41.6)	28.3 (18.3; 42.0)
Green leafy vegetables, raw (g/d)	7.1 (4.4; 10.7)	7.1 (3.6; 13.6)	7.1 (3.6; 14.1)
Green leafy vegetables, cooked (g/d)	19.7 (10.2; 32.2)	18.2 (9.1; 26.7)	19.0 (10.2; 29.1)
Allium vegetables (g/d)	30.9 (13.5; 41.6)	21.9 (11.0; 43.0)	24.0 (11.0; 40.9)
Carrots (g/d)	7.5 (2.7; 16.7)	8.1 (3.5; 14.1)	8.6 (3.7; 15.5)
Tomatoes (g/d)	18.8 (9.4; 37.6)	18.8 (4.7; 32.9)	18.8 (9.4; 32.9)
Legumes (g/d)[††]	24.0 (13.8; 37.3)	28.5 (17.0; 42.9)	28.2 (17.6; 41.6)
Other vegetables (g/d)[**]	20.2 (12.9; 37.0)	22.0 (12.7; 31.2)	21.2 (13.3; 31.3)
Micronutrients[‡‡], median (p25; p75)	*n*=54	*n*=387	*n*=3,048
α-carotene (μg/d)	587.4 (295.3; 921.9)	567.4 (348.7; 911.6)	565.1 (333.8; 897.6)
β-carotene (μg/d)	2,645 (1,880; 3,313)	2,706 (1,981; 3,565)	2,653 (1,944; 3,578)
Vitamin C (mg/d)	94.1 (66.2; 116.6)	96.1 (76.5; 127.7)	97.0 (72.4; 127.1)
Lutein/zeaxanthin (μg/d)	2,323 (1,756; 2,863)	2,237 (1,738; 2,955)	2,341 (1,770; 3,001)
β-cryptoxanthin (μg/d)	122.7 (30.6; 255.8)	132.1 (49.5; 270.5)	128.8 (47.7; 265.9)
Lycopene (μg/d)	865.5 (383.4; 1,399)	798.0 (382.5; 1,336)	808.8 (406.2; 1,332)
Folate (μg/d)	190.7 (158.5; 234.1)	197.3 (162.0; 247.3)	200.7 (165.6; 243.9)
Total dietary fiber (g/d)	25.4 (21.7; 28.7)	26.1 (21.2; 32.5)	26.1 (21.4; 31.5)
Users of supplements, n (%)	*n*=54	*n*=387	*n*=3,048
Multivitamins or minerals	0 (0.0%)	15 (3.9%)	145 (4.8%)
Vitamin C	3 (5.6%)	21 (5.4%)	187 (6.1%)
Garlic supplement	7 (13.0%)	34 (8.8%)	271 (8.9%)

[*] Includes fruits noted in an open-ended question on frequently consumed items not listed in the questionnaire.

[†] Processed citrus fruit juices are not included in total fruits.

[‡] Includes apple sauce.

[§] Excludes potatoes but includes vegetables noted in an open-ended question on frequently consumed items not listed in the questionnaire.

[**] Sauerkraut is not included in the group of cruciferous, but of other vegetables, because a lot of potential anticancer agents are destroyed during its processing.

[††] Also includes dried pulses.

[‡‡] Related to fruits and vegetables.

the subcohort, the percentage of men was lower among cases with hMLH1 protein-deficient colon cancer (hMLH1$^-$) and higher among hMLH1$^+$ cases. There were more ex-smokers among hMLH1$^+$ cases, but smokers among hMLH1$^-$ cases smoked more pack-years than did smokers among subcohort members. Total energy intake, intake of macronutrients and methionine, and alcohol consumption were relatively low among hMLH1$^-$ cases, whereas the corresponding distributions among hMLH1$^+$ cases and subcohort members were similar. In particular, hMLH1$^-$ cases tended to consume more meat and less fish than did the subcohort. hMLH1$^-$ tumors were

Table 6-3. Odds ratios, incidence rate ratios and 95% confidence intervals (95% CI) for hMLH1 protein-deficient (hMLH1$^-$) and -proficient (hMLH1$^+$) colon cancer according to consumption categories of fruits, vegetables, folate and vitamin C

Subgroup	Consumption level	Colon cancer cases (*n*)		Person years of observation	Case-case analyses*	Case-cohort analyses†	
		hMLH1$^-$	hMLH1$^+$	in subcohort	hMLH1$^-$ *vs.* hMLH1$^+$	hMLH1$^-$	hMLH1$^+$
All fruits	0-<116.0 g/d	26	130	4,920	1 (reference)	1 (reference)	1 (reference)
	116.0-<204.6 g/d	15	125	4,898	0.55 (0.27-1.11)	0.55 (0.29-1.05)	0.97 (0.74-1.26)
	≥204.6 g/d	13	132	4,942	0.48 (0.23-0.98)	0.46 (0.23-0.90)	1.03 (0.78-1.35)
					P_{trend}=0.046	P_{trend}=0.029	P_{trend}=0.81
Citrus and fresh citrus juices	0-<31.8 g/d	21	121	4,922	1 (reference)	1 (reference)	1 (reference)
	31.8-<89.8 g/d	18	141	4,884	0.67 (0.34-1.32)	0.81 (0.43-1.52)	1.16 (0.89-1.51)
	≥89.8 g/d	15	125	4,954	0.66 (0.32-1.36)	0.67 (0.34-1.32)	1.10 (0.83-1.45)
					P_{trend}=0.25	P_{trend} =0.26	P_{trend}=0.60
Apples/pears	0-<44.5 g/d	21	117	4,477	1 (reference)	1 (reference)	1 (reference)
	44.5-<114.3/d	20	137	5,281	0.80 (0.41-1.55)	0.79 (0.42-1.48)	1.02 (0.78-1.33)
	≥114.3 g/d	13	133	5,002	0.53 (0.25-1.11)	0.55 (0.27-1.09)	1.04 (0.79-1.37)
					P_{trend} =0.094	P_{trend}=0.085	P_{trend}=0.76
Other	0-<8.9 g/d	20	142	4,903	1 (reference)	1 (reference)	1 (reference)
	8.9-<23.5 g/d	19	119	4,903	1.22 (0.61-2.43)	0.96 (0.51-1.81)	0.83 (0.64-1.08)
	≥23.5 g/d	15	126	4,955	0.90 (0.43-1.86)	0.78 (0.39-1.56)	0.89 (0.68-1.16)
					P_{trend}=0.68	P_{trend}=0.46	P_{trend}=0.53
All vegetables	0-<150.6 g/d	21	137	4,630	1 (reference)	1 (reference)	1 (reference)
	150.6-<209.5 g/d	16	108	4,659	1.07 (0.52-2.17)	0.80 (0.41-1.55)	0.81 (0.62-1.07)
	≥209.5 g/d	17	123	4,687	1.00 (0.49-2.01)	0.86 (0.45-1.65)	0.94 (0.72-1.23)
					P_{trend}=0.99	P_{trend}=0.67	P_{trend}=0.72
Micronutrients							
Folate	<177.2 μg/d	23	136	4,917	1 (reference)	1 (reference)	1 (reference)
	177.2-<225.6 μg/d	15	114	4,943	0.91 (0.44-1.87)	0.75 (0.38-1.52)	0.83 (0.62-1.10)
	≥225.6 μg/d	16	137	4,901	0.93 (0.43-2.05)	0.92 (0.43-2.00)	1.04 (0.78-1.39)
					P_{trend}=0.85	P_{trend}=0.86	P_{trend}=0.63
Vitamin C	<80.7 mg/d	18	120	4,919	1 (reference)	1 (reference)	1 (reference)
	80.7-<115.7 mg/d	22	142	4,910	1.08 (0.55-2.13)	1.25 (0.65-2.38)	1.18 (0.90-1.54)
	≥115.7 mg/d	14	125	4,932	0.83 (0.38-1.79)	0.81 (0.39-1.68)	1.08 (0.81-1.43)
					P_{trend}=0.65	P_{trend}=0.51	P_{trend}=0.67

* Odds ratios are presented for case-case analyses.

† Incidence rate ratios are presented for case-control analyses.

more often proximally located, showed a less advanced stage, and were more frequently poorly or not differentiated compared with hMLH1$^+$ tumors.

Table 6-2 presents consumption of fruits and vegetables, related micronutrients, and supplements in the study population. Fruit consumption was lower among hMLH1$^-$ colon cancer cases than among hMLH1$^+$ cases and subcohort members. The same applies to vitamin C, β-cryptoxanthin, folate, and dietary fiber, albeit less strikingly. Consumption of vegetables and intake of α-carotene and lycopene was somewhat higher among hMLH1$^-$ cases and somewhat lower among hMLH1$^+$ cases when compared with the subcohort. The same was true for most subgroups of fruits and vegetables.

In Table 6-3, odds ratios and 95% confidence intervals are presented for hMLH1$^-$ and hMLH1$^+$ colon cancer versus each other, as well incidence rate ratios and corresponding 95% confidence intervals for hMLH1$^-$ and hMLH1$^+$ colon cancer calculated by using the subcohort. Consumption of fruits was differentially associated with hMLH1$^-$ and hMLH1$^+$ colon cancer; their consumption was inversely associated with hMLH1$^-$ colon cancer but not with hMLH1$^+$ colon cancer.

The same pattern was observed in the subgroups citrus fruits and apples/pears, although statistical significance was not reached. Consumption of other types of fruits as well as vegetables was not associated with hMLH1$^-$ and hMLH1$^+$ colon cancer. No associations were found with α-carotene, β-carotene, vitamin C, lutein/zeaxanthin, β-cryptoxanthin, lycopene, folate, and dietary fiber (results only shown for folate and vitamin C), nor did addition of these nutrients to the models on fruits and vegetables alter the rate ratios by >10%.

Sensitivity analyses are presented in Table 6-4. The estimates pointed in the same direction as the main estimates, although the strength of the association between fruits and hMLH1 protein-deficient colon cancer decreased when we applied gender-specific tertiles.

Table 6-4. Sensitivity analyses: odds ratios and incidence rate ratios for hMLH1 protein-deficient (hMLH1$^-$) and -proficient (hMLH1$^+$) colon cancer comparing the third tertile of consumption of fruits or vegetables with the first

Comparison	Case-case analyses*	Case-cohort analyses†	
	hMLH1$^-$ *vs* . hMLH1$^+$	hMLH1$^-$	hMLH1$^+$
Total fruit consumption			
Estimate in Table 6-3	0.48	0.46	1.03
Gender-specific tertiles	0.57	0.54	1.03
Cases with proximal tumors only	0.44	0.51	1.20
Negative family history of colorectal cancer	0.60	0.55	1.00
Total vegetables			
Estimate in Table 6-3	1.00	0.86	0.94
Gender-specific tertiles	0.93	0.82	0.96
Cases with proximal tumors only	0.70	0.70	1.07
Negative family history of colorectal cancer	1.24	1.04	0.91

* Odds ratios are presented for case-case analyses.

† Incidence rate ratios are presented for case-cohort analyses.

Discussion

We observed that fruit consumption was inversely associated with hMLH1 protein-deficient colon cancer but not with hMLH1 protein-proficient colon cancer. This observation is in line

with inverse associations between consumption of citrus fruits and between consumption of apples/pears and hMLH1 protein-deficient colon cancer, although these associations were not statistically significant. No association between other types of fruits and hMLH1 protein-deficient colon cancer was found, and the association of fruit consumption with hMLH1 protein-deficient cancer could not be attributed to antioxidants related to fruits and vegetables, folate, and dietary fiber. Consumption of vegetables was not associated with either hMLH1 protein-deficient or -proficient colon cancer.

Our results on fruits are in line with two Dutch case-control studies. In the first, fruit consumption was associated with MSI-high colon cancer with hypermethylated *hMLH1* but not with MSI-high colon cancer without hypermethylated *hMLH1*[16]. In the other, fruit consumption was associated with a lower risk of colorectal cancer among hereditary non-polyposis colorectal cancer (HNPCC) patients[34]. Conversely, in an American case-control study, no associations between fruit consumption and colon cancer with or without MSI were observed[17]. With regard to consumption of vegetables, in all studies conducted thus far including the current study, case-case analyses showed no differential association for MMR-proficient and -deficient colon cancer[16, 17]. No evidence was found for an association with vegetables for hMLH1 protein-deficient and -proficient colon cancer in our case-cohort analyses and the study on HNPCC-related tumors[34], whereas consumption of vegetables was inversely associated with sporadic colon cancer with and without MSI in two other studies[16, 17]. Intake of antioxidants related to fruits and vegetables, dietary fiber, and folate was not associated with either type of tumor in this study. In contrast, intake of folate, dietary fiber, and β-carotene was inversely associated with colon cancer with and without MSI in an American case-control study[17], although only the associations of folate and fiber with MSI-positive colon cancer were statistically significant.

Hypermethylation status of the promoter region of the *hMLH1* gene determines hMLH1 protein expression in the majority (80-90%) of the sporadic colon cancers[2-8], whereas germline mutations in the *hMLH1* gene are responsible for lack of expression in 50% of the HNPCC-related colon cancers[35]. Our classification of hMLH1 protein-deficient and -proficient colon cancers is likely to be primarily determined by hypermethylation status of the promoter region of the *hMLH1* gene. The youngest patient in our study population was diagnosed at age 57; the proportion of HNPCC was only around 2% in other unselected series[8, 36], and the parameter estimates pointed in the same direction when we restricted the analyses to individuals without a family history of colorectal cancer.

Because dietary methyl donors such as folate are thought to affect methylation status of cancer genes[37, 38], the association between fruits and hMLH1 protein-deficient colon cancer may arise because fruits affect hypermethylation of the promoter region of the *hMLH1* gene. We did not observe an association between intake of folate and the rate of hMLH1 protein-deficient colon cancer. Thus, folate, and its potential role in methylation of *hMLH1*, by itself seems insufficient to explain the association between fruit consumption and hMLH1 protein-deficient colon cancer. As fruit consumption was also inversely associated with HNPCC-related tumors[34], our findings may suggest that fruits play a role in pathways involved in hMLH1 protein-deficient colon carcinogenesis that are determined by other factors than hypermethylation of the *hMLH1*

promoter region. However, another study found that an inverse association between fruits and MSI-high colorectal cancer was restricted to tumors with hypermethylated *hMLH1*, whereas no associations were observed with the methyl donors folate and alcohol[16]. Perhaps components in fruits other than folate influence hypermethylation. Pathways involved are unlikely to depend solely on fruit-related antioxidants and dietary fiber, because they were not associated with hMLH1 protein-deficient colon cancer and hardly influenced the association between fruits and hMLH1 protein-deficient colon cancer. Hence, the mechanism behind this association remains to be elucidated.

Our classification of hMLH1 protein-deficient and -proficient tumors seems reliable. First, the percentage of hMLH1 protein-deficient colon cancers in our study corresponds with the literature[10, 11, 16]. Second, we found a high agreement with classification based on the MSI marker BAT-26[39] as expected[40-42]. Third, the sensitivity (94%) and specificity (100%) of hMLH1 immunohistochemistry was high in an unselected series of colorectal cancer with MSI testing as golden standard[8]. However, a minor percentage of our tumors may show MMR deficiency because of mutations in other MMR genes than *hMLH1* and some hMLH1-proficient tumors may have non-functional protein, resulting in a bias towards the null.

hMLH1 protein-deficient colon cancers occur mostly in the proximal colon. We observed that the association remained when we restricted the analyses to proximal tumors. Thus, distinction of hMLH1 protein-deficient and -proficient colon cancer, in addition to tumor location, provides greater information than tumor location by itself.

The strengths of this study include its large size, prospective nature with 7.3 years of follow-up, the fact that it is based on an unselected series of colon cancer, a very low chance of selective dropout because no subcohort members were lost to follow-up, and high completeness of follow-up for cancer incidence.

As in other epidemiologic studies, it is impossible to completely rule out bias that occurs when factors determining tumor block availability or successful hMLH1 immunohistochemistry are associated with fruit or vegetable consumption and hMLH1 expression status[43, 44]. It is unlikely that such a bias has affected our results considerably. We did not observe differences in distribution of Dukes' stages and differentiation grade between tumors of patients in which hMLH1 expression status could be determined and tumors in which we could not assess this status. Moreover, our study population has, on average, the same age and contains a similar percentage of individuals with a positive family history as the complete study population after 7.3 years of follow-up.

Because diet was assessed before the occurrence of disease and the initial 2.3 years of follow-up were excluded, changes in diet due to (subclinical) disease occurrence cannot have affected the assessment. It is no problem that we assessed diet only once, because the small decline in test-retest correlation coefficients over time (0.07 over five years) in the reproducibility study of our food-frequency questionnaire[25] suggests that the exposure is relatively stable over time. However, assessment of diet brings along systematic and random errors. The correlation coefficients between fruit and vegetable consumption assessed by our food-frequency questionnaire and by 9-day dietary records are comparable with those of questionnaires used in other prospective

cohort studies[45, 46], but the moderate ability of the questionnaire to rank according to vegetable consumption (r=0.38 between assessment according to questionnaire and dietary records; ref. 24) results in an attenuation of the associations. Given the substantial correlation coefficient between assessment of fruits by questionnaires and dietary records (r=0.60; ref. 24), it is less likely that misclassification has affected our final conclusion for fruits.

After 7.3 years of follow-up in the Netherlands Cohort Study on diet and cancer, we observed an inverse association between fruits and colon cancer, confined to the subgroup of hMLH1 protein-deficient colon cancer cases. Our findings may provide clues for elucidating pathways involved in carcinogenesis of MMR-deficient colon cancer.

Acknowledgements

We thank the participants of our study; Dr. M. Brink for collecting tumor specimen; H. Braam and G. Roemen for laboratory assistance; Prof. Dr. J.H.J.M. van Krieken for reviewing hMLH1 scoring; Drs. A. Volovics and A. Kester for statistical advice; Drs. L. Schouten, A. de Bruïne, and L.E. Voorrips, S. van de Crommert, H. Brants, J. Nelissen, C. de Zwart, M. Moll, W. van Dijk, M. Jansen, and A. Pisters, for assistance; H. van Montfort, T. van Moergastel, L. van den Bosch, and R. Schmeitz for programming assistance; Prof. J.W. Arends for his participation in the initiation of this study; the regional cancer registries (IKA, IKL, IKMN, IKN, IKO, IKR, IKST, IKW, IKZ); the Dutch nationwide network and registry of histopathology and cytopathology (PALGA); and the departments of pathology of the following hospitals for providing the tissue blocks: Academisch Ziekenhuis Nijmegen Sint Radboud, Academisch Ziekenhuis Groningen, Rijnland Ziekenhuis, Antoni van Leeuwenhoek Ziekenhuis, Academisch Ziekenhuis Rotterdam, Stichting Laboratorium Pathologie Oost Nederland, Pathologisch Instituut Utrecht, Ziekenhuis Rijnstate Arnhem, Laboratorium Volksgezondheid Leeuwarden, Ziekenhuis Bethesda, Stichting Samenwerkende Ziekenhuizen Oost Groningen, Martini Ziekenhuis Groningen, Stichting Samenwerkende Delftse Ziekenhuizen, Leyenburg Ziekenhuis, Academisch Ziekenhuis Vrije Universiteit, Academisch Medisch Centrum, Sint Franciscus Ziekenhuis, Dr. Daniel den Hoed Kliniek, Academisch Ziekenhuis Maastricht, Stichting Laboratorium Goudse Ziekenhuizen, Canisius-Wilhelmina Ziekenhuis, Slotervaart Ziekenhuis, Maaslandziekenhuis, Atrium Heerlen, Atrium Kerkrade and Brunssum, Microbiologie St Medische Stedendriehoek, IJsselmeer Ziekenhuizen, Ziekenhuis Centrum Apeldoorn, Isala Klinieken, Elkerliek ziekenhuis, Groot Ziekengasthuis, Ziekenhuis Gooi-Noord, Medisch Centrum Alkmaar, Regionaal Pathologisch en Cytologisch Laboratorium voor Eemland en Noord-West Veluwe, Diakonesse Ziekenhuis, Sint Antonius Ziekenhuis, Onze Lieve Vrouwe Gasthuis, St. Lucas Andreas Ziekenhuis, Pathologisch Anatomisch Laboratorium SPALK, Ziekenhuis De Heel, Diakonessenhuis, Rode Kruis Ziekenhuis, Ziekenhuis Bronovo, Laurentius Ziekenhuis Roermond, Pathologisch Anatomisch Laboratorium Dordrecht, Zuiderziekenhuis, Sint Clara Ziekenhuis, Medisch Centrum Haaglanden, Stichting Streeklaboratorium Zeeland, Sint Elisabeth Ziekenhuis, Catharina-ziekenhuis, Sint Maartensgasthuis, and Spaarne Ziekenhuis. This study was supported by the Netherlands Organisation for Health Research and Development, and the Dutch Cancer Society.

References

1. Atkin NB. Microsatellite instability. *Cytogenet Cell Genet.* 2001;92(3-4):177-181.
2. Herman JG, Umar A, Polyak K, et al. Incidence and functional consequences of hMLH1 promoter hypermethylation in colorectal carcinoma. *Proc Natl Acad Sci U S A.* 1998;95(12):6870-6875.
3. Wheeler JM, Loukola A, Aaltonen LA, Mortensen NJ, Bodmer WF. The role of hypermethylation of the hMLH1 promoter region in HNPCC versus MSI+ sporadic colorectal cancers. *J Med Genet.* 2000;37(8):588-592.
4. Kim HC, Kim CN, Yu CS, Roh SA, Kim JC. Methylation of the hMLH1 and hMSH2 promoter in early-onset sporadic colorectal carcinomas with microsatellite instability. *Int J Colorectal Dis.* 2003;18(3):196-202.
5. Kuismanen SA, Holmberg MT, Salovaara R, De la Chapelle A, Peltomaki P. Genetic and epigenetic modification of MLH1 accounts for a major share of microsatellite-unstable colorectal cancers. *Am J Pathol.* 2000;156(5): 1773-1779.
6. Kamory E, Kolacsek O, Otto S, Csuka O. hMLH1 and hMSH2 somatic inactivation mechanisms in sporadic colorectal cancer patients. *Pathol Oncol Res.* 2003;9(4):236-241.

7. Cunningham JM, Christensen ER, Tester DJ, et al. Hypermethylation of the hMLH1 promoter in colon cancer with microsatellite instability. *Cancer Res.* 1998;58(15):3455-3460.

8. Cunningham JM, Kim CY, Christensen ER, et al. The frequency of hereditary defective mismatch repair in a prospective series of unselected colorectal carcinomas. *Am J Hum Genet.* 2001;69(4):780-790.

9. Chao A, Gilliland F, Willman C, et al. Patient and tumor characteristics of colon cancers with microsatellite instability: a population-based study. *Cancer Epidemiol Biomarkers Prev.* 2000;9(6):539-544.

10. Ward R, Meagher A, Tomlinson I, et al. Microsatellite instability and the clinicopathological features of sporadic colorectal cancer. *Gut.* 2001;48(6):821-829.

11. Miyakura Y, Sugano K, Konishi F, et al. Extensive methylation of hMLH1 promoter region predominates in proximal colon cancer with microsatellite instability. *Gastroenterology.* 2001;121(6):1300-1309.

12. Thibodeau SN, Bren G, Schaid D. Microsatellite instability in cancer of the proximal colon. *Science.* 1993;260(5109): 816-819.

13. Lothe RA, Peltomaki P, Meling GI, et al. Genomic instability in colorectal cancer: relationship to clinicopathological variables and family history. *Cancer Res.* 1993;53(24):5849-5852.

14. Davis CD, Uthus EO. DNA methylation, cancer susceptibility, and nutrient interactions. *Exp Biol Med (Maywood).* 2004;229(10):988-995.

15. Glaab WE, Hill RB, Skopek TR. Suppression of spontaneous and hydrogen peroxide-induced mutagenesis by the antioxidant ascorbate in mismatch repair-deficient human colon cancer cells. *Carcinogenesis.* 2001;22(10):1709-1713.

16. Diergaarde B, Braam H, Van Muijen GNP, Ligtenberg MJ, Kok FJ, Kampman E. Dietary factors and microsatellite instability in sporadic colon carcinomas. *Cancer Epidemiol Biomarkers Prev.* 2003;12(11 Pt 1):1130-1136.

17. Slattery ML, Anderson K, Curtin K, Ma KN, Schaffer D, Samowitz W. Dietary intake and microsatellite instability in colon tumors. *Int J Cancer.* 2001;93(4):601-607.

18. Van den Brandt PA, Goldbohm RA, Van 't Veer P, Volovics A, Hermus RJ, Sturmans F. A large-scale prospective cohort study on diet and cancer in The Netherlands. *J Clin Epidemiol.* 1990;43(3):285-295.

19. Van den Brandt PA, Schouten LJ, Goldbohm RA, Dorant E, Hunen PM. Development of a record linkage protocol for use in the Dutch Cancer Registry for Epidemiological Research. *Int J Epidemiol.* 1990;19(3):553-558.

20. Schouten LJ, Straatman H, Kiemeney LA, Gimbrère CH, Verbeek AL. The capture-recapture method for estimation of cancer registry completeness: a useful tool? *Int J Epidemiol.* 1994;23(6):1111-1116.

21. Iacopetta B. Are there two sides to colorectal cancer? *Int J Cancer.* 2002;101(5):403-408.

22. Kapiteijn E, Liefers GJ, Los LC, et al. Mechanisms of oncogenesis in colon versus rectal cancer. *J Pathol.* 2001;195(2): 171-178.

23. Wei EK, Giovannucci E, Wu K, et al. Comparison of risk factors for colon and rectal cancer. *Int J Cancer.* 2004;108(3):433-442.

24. Goldbohm RA, Van den Brandt PA, Brants HA, et al. Validation of a dietary questionnaire used in a large-scale prospective cohort study on diet and cancer. *Eur J Clin Nutr.* 1994;48(4):253-265.

25. Goldbohm RA, Van 't Veer P, Van den Brandt PA, et al. Reproducibility of a food frequency questionnaire and stability of dietary habits determined from five annually repeated measurements. *Eur J Clin Nutr.* 1995;49(6): 420-429.

26. Zeegers MP, Goldbohm RA, Van den Brandt PA. Consumption of vegetables and fruits and urothelial cancer incidence: a prospective study. *Cancer Epidemiol Biomarkers Prev.* 2001;10(11):1121-1128.

27. Stichting Nederlands Voedingsstoffenbestand. *Dutch Food Composition Table 1986-1987; Nederlands voedingsstoffenbestand 1986-1987*: The Hague : Voorlichtingsbureau voor de Voeding; 1996.

28. Goldbohm RA, Brants HA, Hulshof KF, Van den Brandt PA. The contribution of various foods to intake of vitamin A and carotenoids in The Netherlands. *Int J Vitam Nutr Res.* 1998;68(6):378-383.

29. Konings EJ. A validated liquid chromatographic method for determining folates in vegetables, milk powder, liver, and flour. *J AOAC Int.* 1999;82(1):119-127.

30. Konings EJ, Roomans HH, Dorant E, Goldbohm RA, Saris WH, Van den Brandt PA. Folate intake of the Dutch population according to newly established liquid chromatography data for foods. *Am J Clin Nutr.* 2001;73(4): 765-776.

31. Volovics A, Van den Brandt PA. Methods for the analyses of case-cohort studies. *Biomed J.* 1997;2:195-214.

32. Schoenfeld DA. Partial residuals for the proportional hazards regression model. *Biometrika.* 1982;69:239-241.

33. Grambsch PM, Therneau TM. Proportional Hazards Tests and Diagnostics Based on Weighted Residuals. *Biometrika.* 1994;81(3):515-526.

34. Diergaarde B, Braam H, Vasen HF, et al. Dietary factors and HNPCC-associated colorectal tumors. Paper presented at: Third Annual AACR International Conference Frontiers in Cancer Prevention Research, 2004; Oct 16-20; Seattle, WA, USA.

35. Peltomäki P. Deficient DNA mismatch repair: a common etiologic factor for colon cancer. *Hum Mol Genet.* 2001;10(7):735-740.

36. Aaltonen LA, Salovaara R, Kristo P, et al. Incidence of hereditary nonpolyposis colorectal cancer and the feasibility of molecular screening for the disease. *N Engl J Med.* 1998;338(21):1481-1487.

37. Friso S, Choi SW. Gene-nutrient interactions and DNA methylation. *J Nutr.* 2002;132(8 Suppl):2382S-2387S.

38. Van Engeland M, Weijenberg MP, Roemen GM, et al. Effects of dietary folate and alcohol intake on promoter methylation in sporadic colorectal cancer: the Netherlands cohort study on diet and cancer. *Cancer Res.* 2003;63(12): 3133-3137.

39. Lüchtenborg M, Weijenberg MP, Wark PA, et al. Mutations in APC, CTNNB1 and K-ras genes and expression of hMLH1 in sporadic colorectal carcinomas from the Netherlands Cohort Study. *BMC Cancer.* 2005;5:160.

40. Cawkwell L, Gray S, Murgatroyd H, et al. Choice of management strategy for colorectal cancer based on a diagnostic immunohistochemical test for defective mismatch repair. *Gut.* 1999;45(3):409-415.

41. Lindor NM, Burgart LJ, Leontovich O, et al. Immunohistochemistry versus microsatellite instability testing in phenotyping colorectal tumors. *J Clin Oncol.* 2002;20(4):1043-1048.

42. Ward R, Meldrum C, Williams R, et al. Impact of microsatellite testing and mismatch repair protein expression on the clinical interpretation of genetic testing in hereditary non-polyposis colorectal cancer. *J Cancer Res Clin Oncol.* 2002;128(8):403-411.

43. Hoppin JA, Tolbert PE, Taylor JA, Schroeder JC, Holly EA. Potential for selection bias with tumor tissue retrieval in molecular epidemiology studies. *Ann Epidemiol.* 2002;12(1):1-6.

44. Schroeder JC, Weinberg CR. Use of missing-data methods to correct bias and improve precision in case-control studies in which cases are subtyped but subtype information is incomplete. *Am J Epidemiol.* 2001;154(10):954-962.

45. Hankin JH, Wilkens LR, Kolonel LN, Yoshizawa CN. Validation of a quantitative diet history method in Hawaii. *Am J Epidemiol.* 1991;133(6):616-628.

46. Ocké MC, Bueno de Mesquita HB, Goddijn HE, et al. The Dutch EPIC food frequency questionnaire. I. Description of the questionnaire, and relative validity and reproducibility for food groups. *Int J Epidemiol.* 1997;26 Suppl 1:S37-S48.

Preamble

This chapter aims to address the underlying research question of this thesis: "does distinguishing colorectal tissue by its histopathological and molecular characteristics shed further light on the role of dietary, lifestyle and heritable factors that are possibly involved in colorectal carcinogenesis?"

We will first describe the main findings (**Section 7.1**) of the studies that we selected, as specified in chapter 1. Thereafter, we will discuss strengths and weaknesses of the chosen approaches in the following three sections:

- **Section 7.2**: Study design: population, effect measures, power and external validity
- **Section 7.3**: Assessment of dietary, lifestyle and heritable factors
- **Section 7.4**: Assessment and advantages of histopathological and molecular endpoints

These sections are further subdivided as follows:

- status quo: describes the difficulties that we encountered and the state of the art.
- future research: provides suggestions on how future research may better deal with these encountered difficulties.

The strenghts and weaknesses analysis will culminate in the concluding remarks (**Section 7.5**), that address the main research question of this thesis, and recommendations for further research.

7

General Discussion

7.1 MAIN FINDINGS

In this thesis, we investigated whether associating dietary, lifestyle and heritable factors with specific histopathological and molecular endpoints, rather than disease occurrence per se, sheds further light on the roles of these factors in colorectal carcinogenesis. We selected five specific research questions on diet, lifestyle or heritable factors as examples of different studies that distinguish colorectal tissue by its histopathological and molecular characteristics. Table 7-1 briefly summarizes the main findings and also gives an overview of the study designs that were used. The table highlights that we assessed endpoints at three levels, which represent different stages in colorectal carcinogenesis: normal tissue, adenomas and carcinomas. The studies were either cross-sectional, case-control or cohort studies; and were conducted in the Netherlands or the US. The studied exposures included dietary and lifestyle factors that are currently thought to influence colorectal cancer risk. Let us have a look at each of the specific research questions separately:

1. Are parameters reflecting the activity of the rectal glutatione S-transferase detoxification system associated with consumption of fruits and vegetables at habitual levels of exposure? If so, do genetic variations in glutathione S-transferase genes GSTM1 and GSTT1 modify this association?

Yes, consumption of (citrus) fruits was positively associated with rectal glutathione *S*-transferase activity in our study on habitual levels of consumption, but this association did not depend on *GSTM1* and *GSTT1* genotype. Consumption of cruciferous vegetables was postively associated with glutathione *S*-transferase activity and GSTM1-1 protein levels among *GSTM1*-plus genotype carriers. No effect modification by *GSTT1* was observed. Hence, habitual levels of consumption of fruits and vegetables may contribute to variations in human rectal GST enzyme activity.

2. Do associations between presumed modifiable risk factors for colorectal adenomas depend on K-ras mutational status of the adenoma?

Associations between some modifiable risk factors and adenomas depended on K-*ras* mutational status in our study. We observed that the intake of vitamin B2 was inversely associated with K-*ras* mutation-negative adenomas, but not with K-*ras* mutation-positive adenomas. A positive association with monounsaturated fat was confined to K-*ras* mutation-negative adenomas. Additional differential associations were also suggested. However, the literature on diet, lifestyle and K-*ras* mutations is inconsistent. Systematic errors inherent to the occurrence of the mutation, confounding by related tumor characteristics and the non-specific assessment of the causal exposures may be responsible for these discrepancies.

3. Are dietary and lifestyle risk factors for colorectal cancer differently associated with adenomas that are thought to have a high risk of malignancy (advanced adenomas) and adenomas that are thought to have a low risk of becoming malignant?

Most dietary and lifestyle factors were similarly associated with advanced and non-advanced adenomas. Smoking was more strongly associated with advanced adenomas than with non-advanced adenomas, and a high glycemic index was inversely associated with advanced but not with non-advanced adenomas. Thus, associations of smoking and glycemic index with adenomas may depend on the stage of adenoma development.

4. Do people with a family history of colorectal cancer have a higher risk of advanced adenomas or do they mostly have a higher likelihood of getting multiple adenomas?

Associations with family history of colorectal cancer were stronger for men with multiple distal adenomas than for men with single distal adenomas at first diagnosis. Associations between family history and advanced and non-advanced adenomas were of similar strength, but a tendency towards a somewhat stronger association with non-advanced adenomas was found. Our findings suggest that heritable factors may be more important in earlier than in later stages of colorectal carcinogenesis at the population level.

5. Is consumption of fruits and vegetables differentially associated with colon cancer with and without defects in the mismatch repair system?

Yes, we observed that the inverse association with fruits and colon carcinomas was confined to hMLH1-deficient tumors. Vegetables were not associated with either type of tumor in our study. These findings indicate that the etiology of colon cancers with and without a defective mismatch repair system may differ; fruits may only influence the etiology of colon cancer with a defective mismatch repair system.

The relevance of our findings with respect to colorectal carcinogenesis has already been discussed in the previous chapters, and we will not repeat this here. Instead, we will reflect upon the methodology of the approaches used, and focus mainly on issues that we encountered in all five conducted studies to be able to address the main research question.

Table 7-1. Main findings corresponding to the five specific research questions

Chapter	Design, country	Tissue	Exposure	Heterogeneity of association between endpoints*	Association with first endpoint†	Association with second endpoint†
2	Cross-sectional, the Netherlands	Normal rectal tissue			*GST enzyme activity*	*GST protein level and GSH*
			Citrus fruits	N/A	+	0
			Brassica vegetables	N/A	+, restricted to *GSTM1* genotype carriers	+ with GSTM1-1 levels only
			Allium vegetables	N/A	0	- with GSTP1-1 levels only
			Other fruits and vegetables	N/A	0	Mixed, but generally 0
3	Case-control, the Netherlands	Adenomas			*K-ras mutation-positive adenomas*	*K-ras mutation-negative adenomas*
			Vitamin B2	Yes	+?	-
			Monounsaturated fat	Yes	0?	+
			Red meat	Likely	+?	0
			Total dietary and polyunsaturated fat	Possibly	-?	0
			Smoking	Likely	0 or +	++
			Dairy products, calcium, protein, tea	Possibly	0?	-
			Other studied factors	No		
4	Cohort, US	Adenomas			*Advanced adenomas*	*Non-advanced adenomas*
			Smoking	Yes	++	+
			Glycemic index	Yes	-	0
			Physical activity, use of multivitamins and intake of starch	Possibly	--?	-
			Height, aspirin and possibly fish	Possibly‡		
			Other studied factors	No		
5	Cohort, US	Adenomas			*Advanced adenomas*	*Non-advanced adenomas*
			Family history of colorectal cancer	No	+	+
					Multiple synchronous adenomas	*No synchronous adenomas*
			Family history of colorectal cancer	Yes	++	+
6	Cohort, the Netherlands	Carcinomas			*hMLH1 protein-proficient colon cancer*	*hMLH1 protein-deficient colon cancer*
			Fruits	Yes	--	0
			Vegetables	No	0	0

* N/A: Not applicable

† --: stronger inverse association; -: inverse association; 0: null association; +: positive association; ++ stronger positive association; ?: indications for a direction of the association only.

‡ For a specific location only; the direction of the associations differed across locations in the colon and rectum

7.2 STUDY DESIGN: POPULATION, EFFECT MEASURES, POWER AND EXTERNAL VALIDITY

7.2.1 Status Quo

As noted earlier, various types of studies are presented in this thesis: a cross-sectional study, a case-control study and two prospective cohort studies. The cross-sectional (Chapter 2), the case-control study (Chapter 3) and one of the cohort studies (Chapter 6) were used to address one specific research question each, whereas the study population of the other cohort study was used to address two specific research questions (Chapters 4 and 5). In this section we will reflect upon the selection of the study population, its relation to the estimated effect measures, the power and study size, and the external validity of the study.

Comparability of the Studied Population with the Source Population of Interest

Independent of the type of study, an appropriate study base should be selected, which represents the source population of people who will be identified as having the endpoint/disease of interest had the endpoint/disease occurred[1]; this is a prerequisite for a valid comparison of the endpoint in exposed and non-exposed people.

For the cross-sectional study on rectal glutathione *S*-transferase (GST) enzyme activity and protein levels described in Chapter 2, participants were recruited from people undergoing endoscopy; they were recruited within the framework of a case-control study, for which they also had to meet the inclusion criteria. As a result, the study population included an oversampling of patients with adenomas and people belonging to families known with the hereditary non-polyposis coli (HNPCC) syndrome. We are not aware of specific medical conditions related to the referral for endoscopy that could affect the association between exposure factors, and GST enzyme activity and protein levels, apart from the conditions already covered by the exclusion criteria. Moreover, adjusting the association for bowel complaints, adenomas and HNPCC did not alter the conclusions, so within the study population the observed findings are probably valid.

In the case-control study described in Chapter 3, associations between presumed modifiable risk factors and adenomas with and without a K-*ras* mutation were of interest. We select cases with adenomas, but a control group was also selected to be able to evaluate whether the exposure factors were positively or inversely associated with adenomas. Cases should form a representative sample of a specifically defined subset of all adenoma patients, whereas controls should constitute a random sample of the source population from which colorectal adenomas emerge[2]. In the Netherlands, where this study was conducted and where population-screening for colorectal cancer has not yet been introduced, people who undergo endoscopy of the large bowel have a medical indication resulting from complaints or concerns about their gut health, sometimes because of a positive (sporadic) family history of colorectal cancer. As adenomas themselves seldom cause complaints, the only subset of the source population that will be identified as

having adenomas consists of people undergoing endoscopy of the large bowel. Therefore, we selected controls among people who underwent endoscopy but were never diagnosed with colorectal adenomas. A population-based control group was regarded as being inappropriate: with such a control group, characteristics of people who underwent endoscopy would have been associated with adenomas due to the design. Selection of controls from among people referred to the same clinics as cases, but for reasons other than endoscopy, was also considered unsuitable; as adenomas are present in about 30-40% of people aged 65 years or older[3-7], such a selection would have had the same result as a substantial degree of misclassification[8] and would have resulted in attenuation of the odds ratios at best. Inviting people for endoscopy without medical indications may also have ethical constraints. Due to our selection method for controls, however, a substantial proportion had bowel complaints or defecation problems. Some of the medical conditions of the controls that underlie their complaints might share etiological factors with adenomas. As a result, we might have underestimated the odds ratio if an exposure factor of interest is overrepresented within the control group. Alternatively, we might have overestimated the odds ratio if the control group contains a lower percentage of people with the exposure factor of interest than the source population. Nonetheless, the presumed variation in medical conditions of the control group guides us towards believing that bias due to associations resulting from shared etiology with adenomas is actually limited in this case-control study.

We compared people who had ever had an adenoma with people who were never diagnosed with adenomas but underwent at least one endoscopy. People who already had an adenoma are more likely to undergo another endoscopy; they are therefore more likely to be included in the study population than men who did not undergo another endoscopy. This may have resulted in selection bias. As determinants of adenoma recurrence may differ from determinants of adenoma incidence, and heritable factors may play a more important role in adenoma occurrence in people with previous adenomas, we excluded people with previous adenomas in a sub-analysis; this did not change the conclusions.

The former reasoning can be extended to the case-case analyses, in which one group of cases is treated as the reference group. Both case groups should consist of a random sample of the adenomas they represent. As we sampled in a clinical setting, it might not be realistic to obtain a fully representative sample of all adenomas, because not all lesions are detected, polypectomized or sent to the pathological laboratory. This applies especially to small lesions, which, if removed, may yield too little tissue to carry out the desired laboratory assessments. In the study described in Chapter 3, these small lesions are the ones more likely to lack K-*ras* mutations, and this group might therefore have been underrepresented in the sample. Such an underrepresentation only results in selection bias if factors determining tumor block availability, or successful assessment of the mutation of interest, are associated with the studied exposure factors as well as with the mutational status[9, 10]. Comparing characteristics of the adenoma patients in whom mutational status was successfully assessed with characteristics of those in whom mutational status was unknown might give some indication of selection bias, but such a comparison can impossibly cover the entire selection procedure. However, we were somewhat reassured because we did not

find any indications for associations between exposure factors and availability of the assessment of K-*ras* mutational status in this study.

In the studies described in Chapters 4 and 5, US Health Professionals were followed up over time. Following similar argumentation as for the case-control study on adenomas, only the subset of the original cohort that underwent endoscopy was selected for the analyses on adenomas. Men who were diagnosed with adenomas before study entry were not eligible, which increased the likelihood that determinants of adenoma occurrence, rather than recurrence, were studied.

As in the case-control study on adenomas, people undergoing multiple endoscopies might have been more likely to be diagnosed with adenomas during the follow-up period; this group not only includes people with medical indications for endoscopy or with a positive family history of colorectal cancer, but also more health-conscious people who adhere to the US national screening guidelines on colorectal cancer. If these people would have been more or less likely to have been exposed to the risk factor under investigation, the increased likelihood that such people might have undergone multiple endoscopies may have resulted in some degree of selection bias. The impact of such a bias is considered smaller than in the case-control study on adenomas, given that a fixed study population was selected (*i.e.* only men of the Health Professionals Follow-Up Study were eligible) rather than a more dynamic study population as was chosen for the case-control study. Comparable issues with regard to potential bias in the case-case analyses could have occurred as described in the previous section. Although tissue retrieval was not necessary for the classification in this chapter, data from endoscopy and pathology reports were used. However, those reports did not contain data required for the classification of advanced and non-advanced adenomas for a large number of people. Sensitivity analyses were conducted to evaluate the impact of such a potential selection bias, which was probably not major in these two studies; also because it is unlikely that the physician's decision as to whether or not to send the adenoma to the pathology laboratory depends on exposure factors.

Colorectal cancer will usually be diagnosed after the patient complains of a condition. Therefore, we did not have to restrict the selection of study participants to those who underwent endoscopy in the study described in Chapter 6. Bias due to unsuccessful hMLH1 immunohistochemistry or restrictions in tumor block availability might also have occurred in this study, analogous to the reasoning within the case-control study regarding the assessment of K-*ras* mutational status.

Estimated Measures of Association

In a case-control setting, the control group should ideally provide an estimate of the relative size of the denominator of the incidence rates for people who were exposed versus people who were not exposed. Because the actual date of occurrence of adenomas may be well before the date of clinical diagnosis, we can only estimate adenoma prevalence in this study. Adenomas occur frequently in the general population[4], and their prevalence depends on the duration of their presence, as determined by incidence and regression[5]. The determinants of this dynamic process cannot be considered equal for exposed and non-exposed individuals[11,12]. Thus, although, prevalence can be considered a summary measure of incidence, duration and regression of adenomas, we are left with estimation of the prevalence odds ratio (POR)[12] for both the case-

control and the case-case comparisons. One of the proposed mechanisms by which exposure factors may influence cancer is indeed enhancing the growth of adenomas, which was one of the reasons leading to the comparison of advanced and non-advanced adenomas presented in Chapters 4 and 5. The estimated PORs, however, remain measures of the strength of association and the POR directly relates to the prevalence based on the same population[12]. The former reasoning applies to all studies on adenomas, even if they are registered prospectively. Thus, also in the studies embedded in the Health Professional Follow-Up Study (Chapters 4 and 5), PORs were estimated.

In studies on colorectal cancer, like in Chapter 6, the date of diagnosis can be regarded as being directly related to incidence, because the disease manifests itself with complaints, leading to its diagnosis. Whether the case-case analyses can be interpreted as an estimation of the (differential) incidence rate ratio depends on whether one regards the two case groups as being independent of each other: this cannot be entirely true in the study in Chapter 6 because loss of hMLH1 expression is essentially a measure of its prevalence in diagnosed tumors, and it may have occurred sooner or later in the adenoma-carcinoma sequence depending on exposure.

Statistical Power and Precision

The sample size of the study in combination with the precision of the measurements determines the magnitude of the difference in exposure that can be detected between the subgroups of interest. Especially when one wants to study differential associations between exposure and various histopathological or molecular endpoints, or incorporate genetic susceptibility like we did, a relatively large number of participants are needed given the large number of distinct subgroups as defined by the combination of exposure and disease categories. Given that some of the studied (sub)groups were small, in particular in Chapters 2, 3 and 6, it is possible that we failed to detect postulated associations between exposure and disease occurrence that truly exist. The magnitude of the estimated association measures guided our interpretation of data on smaller (sub)groups, but we acknowledge that these point estimates are statistically unstable as illustrated by a generally wide confidence interval. Furthermore, small samples - but also highly selected samples as in the cross-sectional study on GSTs presented in Chapter 2 - increase the likelihood of detecting spurious associations because a susceptible subgroup has been selected by chance or through various operating biases[13]. Thus, the findings presented in Chapters 4 and 5, which were based on a large number of adenoma patients of the Health Professionals Follow-Up Study, would likely be more robust when compared to the findings from the other studies in this thesis, if we were to only consider the sample size.

Because we studied associations with many risk factors and multiple endpoints, it is likely that some of the findings described in this thesis have occurred by chance; if one accepts a significance level of 5%, one in every 20 independent analyses is expected to yield an association by chance alone. To avoid too much emphasis on results that might be due to chance fluctuation, we focused mainly on the case-case analyses and the magnitude of the accompanying odds ratios when we interpreted the data. However, sometimes we presented two separate comparisons of two case groups with the reference category even if the case-case analysis suggested that the strength and direction of the association versus the reference category did not depend on

the case group studied; it might have been better to present an overall association within the study population in the absence of such heterogeneity between the two associations rather than illustrating why the case-case association arose.

Generalizability

In the studies described in Chapters 2 to 5 we selected subgroups of the population that underwent endoscopy. Whereas such selection enhances the internal validity of the study, it may be that risk factors are differently distributed among people who undergo endoscopy than among those who do not. This is caused by the fact that people with concerns about their gut health, and people who adhere to the American screenings guidelines for endoscopy, are likely to differ from the general population. This stresses the need for replication of studies in different populations.

The studies presented in Chapters 4 and 5 are based on a cohort of health professionals who are more likely to be more health conscious than the general population. The participants may therefore have had a more favorable distribution of exposure to risk factors, and we may have failed to detect some associations within this study because participants may simply not have been exposed to common levels of the risk factor of interest. Participants in the population-based cohort described in Chapter 6 may also be more health conscious, as well as participants in the cross-sectional study (Chapter 2) and case-control study (Chapter 3). This is due to the fact that health conscious people and/or people with a family history of colorectal cancer are more likely to participate in studies, whereas people who are severely ill may be less inclined to take part in an epidemiological study.

When comparing findings across studies[14], differences between the studied ranges of exposure should also be kept in mind. For example, calcium intake is higher in the Netherlands than in the US, which may partly explain why we observed an association between calcium and K-*ras* mutation-negative colorectal adenomas (Chapter 3) whereas a similar US study did not[15].

A general concern with studies conducted in a single population is the expected homogeneity of dietary and lifestyle patterns. In each study, the effects of individual compounds are examined against the background of average risk associated with other nutrients or foods[16] in a population and it remains to be established whether an association found in one study population can be extended to another study population if complex interactions between foods and nutrients exist.

7.2.2 Future Research

Studies within the Context of Population-Screening

Once population-screening is introduced in a country, studies can be initiated that recruit participants within the framework of these screening programs. These participants can be followed over time. The fact that these people are more likely to be health-conscious, and/or have a family history of colorectal cancer, will not bias associations between exposure and risk of disease when making such a selection, although it could possibly affect generalization of the findings to other populations.

Provided that the time periods between screenings are not overly lengthy, adenoma incidence can be modeled based on the newly diagnosed adenomas having originated in the period between screenings[17], and estimates of adenoma incidence might be obtained. Determinants of adenoma recurrence may also be examined in such a study, which is currently also being studied in trials. Explaining inter-individual variation in the degree of recurrence is also of interest[18] as an indicator of individual susceptibility to adenomas.

Statistical Power and Precision

We raised the issue of relatively low statistical power for subgroup analyses in most studies we presented. An obvious suggestion for improvement is increasing the sample size, but more efficient study designs can also be considered to achieve this. Another way to increase the power is by devising more precise measurements of exposure and endpoints, which will be discussed in Sections 7.3 and 7.4 respectively.

When analyses on histopathological and molecular markers are of interest, a more feasible approach may be to select patients in such a way that they are more likely to show variability in the endpoint; this has additional advantages as laboratory analyses are costly and work-intensive. For example, in the chapter on K-*ras*, we could have opted for selecting participants with specific adenomas, *e.g.* large and villous adenomas that are more likely to harbor such mutations, whereas there is still a reasonable chance that they are mutation-negative. Such a design also enhances the probability that other characteristics associated with K-*ras* mutational status do not account for the observed associations, as will be further discussed in Section 7.4.

Generalizability

Associations between exposure, enzyme activity and protein levels should ideally be confirmed in prospective studies that are population-based or trials that randomly assign people to the exposure of interest. For the hypotheses regarding GST enzyme activity (Chapter 2), trials have already been conducted but they studied higher doses than the usual dietary intakes in which we were interested. Because we were interested in studying habitual patterns, our findings could also be confirmed in a prospective setting where high and low extremes from the exposure distribution (and/or genetic subgroups) are invited to undergo endoscopy to obtain rectal tissue.

7.3 ASSESSMENT OF DIET, LIFESTYLE AND HERITABLE FACTORS

Epidemiological studies have been successful in identifying a range of presumed modifiable risk factors for colorectal cancer. However, their precise contribution is still under investigation, the role of weaker risk factors has yet to be established and the identification of heritable factors has only just begun. This thesis aimed at evaluating whether studying associations with histopathological and molecular endpoints sheds further lights on the role of potential risk factors in colorectal carcinogenesis. Before we can evaluate this, we must reflect upon the state of the art of the assessment of diet, lifestyle and heritable factors, as problems inherent in their assessment influence results of studies between these factors and histopathological and molecular endpoints.

7.3.1 Status Quo

Assessment of Diet, Lifestyle and Heritable Factors in Relation to the Study Objectives

The guiding principle in exposure assessment should be determined by the specific aim of the study. We focused on foods, nutrients and main lifestyle characteristics, which are relevant to the long-term formulation of dietary and lifestyle recommendations. The methods of exposure assessment used were not always suitable to address the role of specific hypothesized (anti)carcinogens. However, in Chapter 2, a variety of specific compounds, such as glucosinolate metabolites derived from brassica vegetables and limonoids present in citrus fruits, guided us to the research hypotheses on the level of a food group; instead of assessing glucosinolate metabolites and limonoids directly, we studied the consumption of brassica vegetables and citrus fruits. Thus, this study should be regarded as hypothesis-generating regarding the role of glucosinolates, limonoids or other components of brassica vegetables or citrus fruits, but it can be regarded as hypothesis-testing regarding the role of brassica vegetables and citrus fruits.

We have not explicitly accounted for the timing of exposure assessment in our studies, although the most etiologically relevant timeframe may differ for the various histopathological and molecular endpoints under study. As tumor development takes years, it is also possible that the exposure related to the development of the disease took place a long time before the tumor manifested itself, which has indeed been observed for the association between smoking and colorectal cancer[19].

Internal Validity and Accuracy of Exposure Assessment

After deciding on the specific exposure and its timeframe of interest, we have to make sure that the assessment reliably reflects what it intends to measure.

In Chapter 2, two observers rather than one assessed *GSTM1* and *GSTT1* genotype using the same method. The validity of the assessment was supported by the observation that GSTM1-1 protein was absent in all biopsies of people with the *GSTM1*-null genotype, whereas detectable

levels were present in all *GSTM1*-plus genotype carriers; indeed, the results of the two observers were comparable in samples that were assessed by both. The validity of the final assessment was further enhanced because samples were interspersed with water controls to check for cross-contamination and a β-globin gene fragment was present as a positive control.

Dietary intake is difficult to assess, because some food items are consumed irregularly and a single food is mostly consumed in combination with other foods. Besides, some people tend to overestimate consumption or habits, whereas others underestimate them[20]. Our primary focus was to describe associations between exposure and characteristics of colorectal tissue, and therefore, adequate ranking of study participants according to dietary and lifestyle behavior was considered more important than assessing absolute intake and exact behavioral characteristics. The questionnaires that were administered in the studies we conducted were able to rank people modestly or well according to most factors[21-25], although ranking was more problematic for vegetables than for most other studied factors (see discussions in separate chapters).

Further complications in assessment of diet are related to the fact that diet consists of a complex array of factors that are often correlated: a single food generally provides a source of multiple nutrients and other components. Besides, dietary habits are often correlated with non-dietary behaviors such as physical exercise and cigarette smoking[26], and dietary patterns tend to cluster according to age, gender, socioeconomic status, and ethnicity. Correlation between dietary and lifestyle factors can be partly taken into account by adjusting for these factors in the analysis. However, associations with two highly correlated exposures cannot always be disentangled, as was the case for vitamin C and citrus fruits in Chapter 2, nor can complex interactive effects of many dietary exposures[27] always been taken into account. Studying dietary patterns[16, 27, 28] instead of single components of the diet could have tackled the correlation issue and possible biological interaction between dietary exposure better but such an approach was not suitable given our study hypotheses.

Systematic misclassification of exposure is likely to be less when it is assessed before the occurrence of the disease (as in Chapters 4, 5 and 6); these assessments are less likely to be affected by disease status and reporting of current exposure is easier than recalling pre-diagnostic or pre-symptomatic exposure status. In the studies described in Chapters 2 and 3, the questionnaire were administered after the participants were diagnosed or underwent endoscopy, which may have affected the reporting or recall of exposure; changes in dietary behavior after undergoing endoscopy may also underlie such different reporting. It is unlikely that reporting methods differ across strata of people with different rectal GST enzyme activity, and thus, in the study of Chapter 2, errors in exposure assessment have most likely resulted in attenuation of the parameter estimates[29]. Misclassification of exposure may be a more taxing issue in Chapter 3, in which patients with colorectal adenomas were compared with adenoma-free men who also underwent endoscopy. A general concern in case-control studies is that cases and controls may recall or report exposure differently, which may result in bias towards or away from the null depending on the size of the systematic error in the comparison groups. For example, people who underwent endoscopy because of bowel complaints may have changed their diet, and are more likely to be members of the control group; therefore, we conducted subanalyses restricted to

people who did not report to have changed their diet to optimize the chance that the assessment reflects the usual exposure during the reference period in Chapter 3. Reassuringly, this exclusion did not change the parameter estimates importantly, which suggests that the influence of errors in exposure assessment due to such changes was small in this study.

Similar reasoning applies to the reporting of a positive family history of colorectal cancer: data up to the last assessment were used rather than data only before diagnosis. However, a family history of colorectal cancer is generally reported accurately and corresponds well with medical-verified data on the relatives[30]. In each case, it is improbable that different groups of cases recall diet, lifestyle or a family history of colorectal cancer differently, which implies that the reported case-case analyses are not subjected to such a bias; these case-case analyses were of main interest in this thesis.

Reproducibility of the Assessment

The chosen assessment methods should be reproducible. The reproducibility of assessment of diet and lifestyle by our questionnaires was generally high[22-24, 31, 32] within individuals. To further reduce random errors in continuous variables, we averaged exposure across different questionnaires in the studies described in Chapters 4 and 5[33], which was not possible in other studies, where exposure was only assessed once. Reassessment of genotype in a subset by the same observers in Chapter 2 yielded similar findings. In general, the reporting of family history of colorectal cancer was consistent between questionnaires, although 235 people reported a positive family history of colorectal cancer in only one of the questionnaires but did not report it in at least two subsequent questionnaires; they were excluded.

However, no assessment is perfect and random errors are hard to avoid. Random errors in exposure assessment result in attenuation of the parameter estimates[29], and if serious, they may reduce the ability of a study to detect associations. This is a more important issue when the power of the study is not very high, such as in Chapters 2 and 3, but also in Chapter 6, in which one of the case groups consisted of only 54 people, despite a large study population.

7.3.2 Future Research

Assessment of Diet, Lifestyle and Heritable Factors in Relation to the Study Objectives

To obtain information on the relevant time-frame of exposure with respect to cancer causation, longitudinal cohort studies with repeated measurements could be used to study the association between diet and colorectal cancer risk for different time intervals between exposure and outcome. Birth cohorts or even conception cohorts on diet, lifestyle and colorectal cancer may be especially useful; such studies cover exposure over the entire lifespan. They may also provide better estimates of disease susceptibility, as exposures early in life may influence fetal growth, cell divisions and organ functioning[34], which may influence the occurrence of cancer on the long term. Studies such as the Norwegian Mother and Child Cohort Study (MoBa)[35] and the Avon Longitudinal Study of Parents and Children (ALSPAC)[36, 37], which include ≥100,000 children,

≥100,000 mothers and ≥75,000 fathers will especially yield valuable information on associations between exposure (including heritable factors), and cancer incidence and mortality in the long term, if study participants can indeed be followed up for the full duration of their lives.

International multicenter studies such as the European Prospective Investigation into Cancer and Nutrition (EPIC)[38] and the Pooling Project of Prospective Studies of Diet and Cancer[39] can evaluate whether diet and cancer associations are consistent across cohort studies comprised of different populations with different dietary habits. It is also possible to examine associations across a wider range of intakes in these studies than in individual studies.

Improvement of Exposure Assessment Methods

Observational studies are hampered by measurement error, which may, at best, result in a reduction in observed effect size, and thus reduce the power of the studies. As part of the inconsistent results between epidemiological studies might be attributed to measurement error, future studies should make the improvement of assessment methods a priority[40]. This applies in particular to studies on diet, which is difficult to assess precisely at the individual level. For research into etiological mechanisms, as well as the purpose of working towards the formulation of concrete measures of prevention, valid and reproducible estimates on consumption of individual foods as well as on nutrient intake are required.

Repeated measurements of exposure will provide more precise estimations of intake, which reduce the random error and thereby increases the power to detect differences in exposure between people. More precise estimates of exposure can also be obtained by combining questionnaire data with information obtained from additional measurements that have independent sources of error: calibration. By doing so, estimates may be adjusted for possible systematic over- or underestimation and corrected for attenuation bias in relative risk estimates[41].

In studies in which exposure assessment has been performed with an instrument of which a likely range of bias is known Monte Carlo sensitivity analysis and Bayesian bias analysis[42] can provide insight into the effect of the bias on the observed effect risk estimates. This allows evaluation of the validity of the assessed exposure-disease relationship.

Furthermore, new methods of exposure assessment should be developed to increase the validity of the exposure assessment, which can be either done by improvement of recording methods of consumption or by development of new biomarkers of intake.

New ways of recording of consumption may follow from new information technologies. In this computer era, multiple stage questionnaires are feasible. Imagine, for instance, a questionnaire in which people first tick foods they consume. Thereafter, a personalized questionnaire can be administered that only requests consumption in detail for those foods that are consumed, or relevant lifestyle habits. On the long term, estimation of consumption may be falicilitated by futuristic fridges that record which food has to be ordered[43], but nowadays supermarkets already register what people buy on loyalty cards. Tracking of consumer purchases through loyalty cards may provide more precise estimation of foods that may be consumed; especially if such a recording method were to be combined with an additional questionnaire that checks actual

consumption of purchased foods, it should be possible to devise more precise and valid tools for estimation of consumption and intake.

Usage of biomarkers rather than self-reported measures of exposure eliminates the issue of recall bias and under- or over-reporting of exposure. Existing biomarkers do not reflect long-time exposure and no such biomarkers for foods are available; they often reflect total exposure to a substance from all sources rather than being specific to one source, and it is not always clear whether they measure the exposure, the biological effect, or some stage of the disease process itself[44]. New biomarkers of exposure may be developed following new developments in genetics and information technology. Gene or protein expression profiles can already be determined by using the microarray technology[45], and it may be that some of these profiles reflect exposure to specific nutrients or even foods. For instance, a person who drinks a lot of alcoholic beverages could have a profile characterized by a high level of expression of proteins that are needed for the detoxification of alcohol. The resulting alcohol-specific gene or protein expression profile can be used as an indicator of exposure to alcohol, provided that the assessment yields a valid and reproducible estimate. New biomarkers may also be obtained by studying genetic aberration profiles in tissue. Certain specific mutations in cancer genes are thought to reflect specific exposures, such as cigaret smoke which is correlated with G:C to T:A transversions[46, 47]. The sensitivity and specificity of using one genetic aberration as assessment of exposure is probably low because exposure factors may cause mutations in multiple genes, and because some of those mutations may be caused by several exposure factors or may have occurred spontaneously. Searching for a combination of genetic aberrations that reflect the exposure may yield more sensitive and specific biomarkers. For instance, studying the point mutational spectra of several genes[48, 49], which do not have to be cancer genes, in normal tissue may identify combinations of mutations that are related to a specific exposure. Another direction in the field of genetic biomarkers is making use of Mendelian randomisation[50-52]. Instead of assessing exposure directly, genetic polymorphisms in genes that are known to mimic the effects of a certain exposure might be studied. For instance, as a polymorphism in the methylenetetrahydrofolate reductase (*MTHFR*) is associated with metabolic effects equivalent to those seen with lower folate intake, it has been proposed that a study of the association between the TT genotype of the gene and risk of colorectal cancer can be regarded as the equivalent to a randomized trial of the effect on disease risk of alteration of the availability of folate[51]. Such associations between genotype and disease occurrence are not confounded, given the random process of genotype determination at conception, but to be a valid alternative to direct assessment of exposure it requires the assumption that the gene variant does not influence behavior[53] and that linkage disequilibrium and alternative pathways through which the gene affects the risk of disease are absent.

7.4 ASSESSMENT AND ADVANTAGES OF HISTOPATHOLOGICAL AND MOLECULAR ENDPOINTS

7.4.1 Status Quo

The relevance of the study design for the availability of tissue specimen has already been discussed in section 7.2. In this present section, we will specifically focus on the handling and storage of tissue samples once they are obtained, and we will discuss the value of studies on normal tissue, adenomas and carcinomas.

Handling and Storage of Tissue

In Chapter 2, rectal biopsies were processed directly after the endoscopy and GST enzyme activity and protein levels were assessed in cytosolic fractions. In the studies described in Chapter 3 on K-*ras*, and Chapter 6 on hMLH1 protein expression, paraffin-tissue was obtained; this was removed as part of the clinical trajectory. This material was used for the assessment of K-*ras* mutational status (adenomas) and for hMLH1 immunohistochemistry (carcinomas).

Because DNA derived from archival tissue is often highly fragmented, misclassification will occur when large DNA segments are required for mutation assessment. However, fragmentation of DNA is unlikely to have been a major problem in the study on the small K-*ras* gene, leaving the degree of misclassification relatively small (Chapter 3). The quality of the hMLH1 immunohistochemistry (Chapter 6) was satisfying, but we noticed that the strength of the immunohistochemical signal depended on factors like the time since storage of the tissue and the conditions of fixation as has been observed before[54]. Whereas it is unlikely that issues related to storage and handling have importantly affected the validity of the assessment of hMLH1 protein expression given the high agreement with a marker of a closely related phenomenon (the microsatellite marker BAT-26) in a subset of the same sample[55], researchers should carefully consider these issues when planning a future study based on tissue; suggestions for the design of such studies embedded within a cohort study have been described elsewhere[54].

Value of Studies on Normal Tissue

In Chapter 2, we used normal rectal tissue obtained from people who underwent endoscopy to study a detoxification pathway. Studying effects on specific pathways relevant to colorectal carcinogenesis - such as detoxification and DNA repair - rather than focusing on the overall lengthy multistage carcinogenic process, may yield clues as to the role of diet, lifestyle and heritable factors. Consequently, the predictive value of processes that were studied in normal tissue with respect to colorectal carcinogenesis cannot be derived from our observations and remains to be inferred from other research.

Value of Studies on Adenomas

In Chapters 3 to 5, we studied adenomas, as precursor lesions of colorectal cancer. Epidemiological studies on adenomas may therefore point to risk factors that are important in the earlier stages of carcinogenesis, and many mutations and changes to the organization of the tissue are already present in adenomas, such as K-*ras* mutations that were studied in Chapter 3.

Cohort studies or intervention trials on adenoma occurrence or recurrence have advantages over such studies on colorectal cancer, because adenomas occur with higher frequency: adenomas are found in about 30-40% of people aged 65 years or older[3-7], whereas the lifetime risk of colorectal cancer is 5.6%[56] in Western countries; the recurrence rate at 3 years has been estimated to lie between 35% to 50%[57]; and adenomas occur, on average, earlier in life than carcinomas do[58, 59]. Hence, a shorter follow-up period and a relatively smaller study size would be needed to detect risk factors. Furthermore, as discussed earlier, studying adenomas rather than carcinomas may have the advantage of reduced recall bias in a case-control setting, because symptomless and chance-detected adenomas are unlikely to have affected the patients' diet.

The relevance of adenomas as precursors of colorectal cancer remains a topic of discussion. About thirty percent of the carcinomas might result from *de novo* carcinogenesis and the etiology of these tumors is not covered by adenoma studies[58]. Furthermore, only a subset of the adenomas develops further into carcinomas. Schatzkin and Gail[60] consider the possibility that a risk factor may solely affect a subgroup of adenomas that have no malignant potential. If so, the risk factor might be associated with adenomas but the association has no predictive value for its association with carcinomas. If a risk factor only affected the risk of adenomas with malignant potential, then the true effect upon all adenomas would be diluted[60].

Hence, associations between diet, lifestyle and heritable factors, and adenomas are not automatically equivalent to associations between diet, lifestyle and heritable factors, and carcinomas. Dietary and lifestyle factors that are associated with colorectal cancer will necessarily be associated with adenoma formation only if adenoma formation is the rate limiting step in the adenoma-carcinoma sequence[61]; this may not be the case as not all adenomas develop further into colorectal cancer and some may even regress[5, 61]. However, adenoma formation may be the rate-limiting step for at least a subgroup of colorectal cancers.

Value of Studies on Carcinomas

In Chapter 6, we studied carcinomas. Since one ultimately aims to reduce colorectal cancer incidence and mortality, this is the primary outcome of interest. However, inference from case-control studies on colorectal carcinomas may be hampered by recall bias and cohort studies need long follow-up periods for a sufficient number of colorectal cancer cases to occur.

In Chapter 6, we assessed hMLH1 protein expression in colorectal carcinomas, based on which we separated the tumors in two groups. By doing so, we aimed at distinguishing two postulated major etiological pathways. Indeed the associations with diet were different; but the complexities of tumor biology prohibit clear conclusions on the questions whether these are completely distinct pathways.

Thus, although the overall effect of risk factors in different phases of colorectal carcinogenesis results in occurrence of colorectal cancer, many transitions occur long before clinical detection. The role of exposures in early rate-limiting steps may no longer be detectable in studies with carcinoma as endpoint.

Studying Determinants of Transitional Stages in the Adenoma-carcinoma Sequence

We have just indicated that interpretation of results from studies on adenomas and carcinomas is not simply equivalent; the same may apply to other intermediate endpoints, such as different adenoma stages. In Chapter 3, this was assessed at the molecular level by K-*ras* mutational status. In Chapters 4 and 5, we evaluated this using histopathological and clinical adenoma characteristics. To be able to study whether associations of risk factors with advanced and non-advanced adenomas differed, we had to rely on studying different types of adenoma endpoints separately using adenomas detected by, and removed during, endoscopy. Although this enabled us to obtain indications as to whether transition of adenomas into carcinomas could be influenced by dietary and lifestyle factors, it remains a matter of inference, because the transitional stages cannot be observed directly in the same person.

7.4.2 Future Research

Handling and Storage of Tissue

Within the scope of a hypothetical study within the screening setting (see Section 7.2.2), it is feasible to collect fresh tumor tissue, which can be used for assessment of endpoints. Standardized embedding and storage conditions in the screening protocol should be incorporated as well. This is particularly important because the nature of clinical practice is selecting and treating diseased people. As a consequence, research based on such clinically obtained specimen brings the risk that selection bias according to current clinical opinion occurs.

Furthermore, heterogeneity is an inherit characteristic of most tumors. Hence, the mutation or gene expression status of interest may not always be present in the entire tissue sample. Therefore, care should be taken to collect as much tissue as possible from each lesion, because the classification of mutational status or expression profile is generally based on the worst characteristic. If heterogeneity with respect to the studied characteristic of interest occurs, consideration should be given to studying tumors in a later stage of development instead, as most tumors may not have yet developed the characteristic of interest.

Choice of the Endpoint

The choice of histopathological and molecular endpoints will be easier when a specific hypothesis is present regarding the exposure-disease relationship. However, this may be impossible if the underlying pathways are not sufficiently defined or are not well characterized by the endpoint. If the research hypotheses would be defined in terms of underlying differential

etiological pathways rather than based on a single and specific exposure-mutation association, this could provide a framework for integrating apparently unrelated specific research findings.

When one wishes to study whether exposures cause specific mutations in humans, normal tissue would be of interest to be used, but the frequency of specific mutations in cancer genes is rather low, and highly sensitive techniques are required to detect such mutations. *In vivo* or *in vitro* experiments might be better equipped to evaluate this.

In order to study specific pathways, a better approach than using tumors may be to conduct small-scale studies on the exposure of interest in relation to the pathway of interest. An example of such a study is described in Chapter 2 on glutathione *S*-transferase enzyme activity. Similar studies could be conducted for numerous pathways and enzymes that are thought to play a role in colorectal carcinogenesis, *e.g.* studying DNA repair capacity. Studies that link these markers to risk of colorectal cancer are necessary as well to study their predictive value[60].

Studying Transitional Stages in the Adenoma-Carcinoma Sequence

Associating risk factors across different stages of the adenoma-carcinoma sequence may improve understanding as to whether the role of risk factors varies across different stages, but studying adenoma and carcinoma progression more directly may be a more promising technique.

A few studies have evaluated whether growth of adenomas over a time interval of up to three years was associated with specific exposures[62-67]. These studies were observational and small. It would be of interest to repeat such a study using an intervention trial design, or conduct larger studies of such a sort under careful supervision. The feasibility of such a study may be improved when non-invasive digital imaging techniques that can monitor the growth of adenomas closely become more readily available. In the future, when non-invasive diagnostic techniques like virtual endoscopy will be sufficiently developed that reliable diagnosis is possible[68], adenoma growth and regression can be studied feasibly at a larger scale.

Modeling approaches may also be useful to improve understanding of factors that influence the transition of different stages. For example, in the EPIC study in which a multistage model was used to mathematically describe the effects of smoking with respect to lung cancer incidence, the data were best described with smoking dependence on the rates of malignant transformation and clonal expansion[69]. Such modeling approaches may either follow or precede biological literature on this topic, but in each case, confirmation is needed as most datasets are usually compatible with multiple models[70].

Furthermore, a risk score that reflects the probability of colorectal cancer of a certain adenoma could be developed, which can subsequently be related to expose factors. Such a risk score might be developed using models analogous to the Miscan Colon Model, which is used to evaluate the effect of screening[5, 71]; this model has been developed based on data from previous studies and clinical experience. As the model can be used to test hypotheses about the natural history of colorectal cancer, such as the duration of progressive adenomas, it might be possible to incorporate information on diet, lifestyle and heritable factors in the model at some stage; so that their effect on the development of colorectal cancer can be simulated.

Correlated Outcomes

Multivariate models, that evaluate multiple outcomes simultaneously, may also be of use when studying correlated responses at the same time. If we had used a multivariate categorical model, such as the multivariate Dale model[72], we might have been able to study associations between diet, lifestyle and heritable factors with different correlated tumor characteristics, such as *p53* and K-*ras* mutations, simultaneously.

To identify which tumor characteristics are influenced by a certain exposure, techniques such as path analysis and structural equation modeling may be used: hypothesized pathways may be tested against each other and new hypotheses may arise.

Heterogeneity of Mutational Patterns within Synchronous and Metachronous Adenomas

Studying variation in mutational patterns within people, with respect to both synchronous and metachronous tumors, deserves attention in future studies. This would add information relevant to the usual studies of associations between diet, lifestyle and heritable factors, and the occurrence of mutations. If the variation is high, it may suggest that the occurrence of the specific mutation is unlikely to be strongly related to exposure to diet, lifestyle or a heritable factor; linking the mutation to exposure factors may be like looking for a needle in a haystack, while being uncertain whether the needle is actually present. An alternative explanation is that the mutation has not yet occurred in most tumors, and that a later stage in the adenoma-carcinoma sequence should be studied to detect an association between exposure and mutations.

7.5 CONCLUDING REMARKS

One of the basic questions raised by our approach to further unravel the etiology of colorectal cancer is how different biological pathways can be distinguished and how transitions between different stages in such pathways can be studied in humans.

As these pathways and transitions may have different determinants, we studied presumed risk factors according to different clinical endpoints (normal tissue, adenoma, carcinoma) and at different levels of integration (mutation, histopathology, enzyme activity and protein levels). We observed that habitual levels of consumption may contribute to variations in detoxification capacity, as we observed for (citrus) fruits and (cruciferous) vegetables in relation to human rectal GST enzyme activity; we showed that different tumor subtypes as determined by their histopathological or molecular characteristics may indeed have a different etiology and that studying histopathological subtypes may reveal associations that would otherwise not have been detected (*e.g.* the association between fruits and hMLH1 protein-deficient colon cancer); we found indications that heritable factors may play a more important role in the early stages of colorectal carcinogenesis than in the later stages, and that associations with risk factors may also depend on the stage of adenoma development, as suggested for smoking and a high glycemic index in particular. Hence, studying specific histopathological and molecular endpoints, rather than disease occurrence per se, sheds further light on the roles of diet, lifestyle and heritable factors in colorectal carcinogenesis. However, as Loomis and Wing wrote, "while molecular techniques can directly address some of the weaknesses of 'black box' environmental theories they do not possess inherent scientific qualities that make them superior tools for understanding the determinants and distribution of disease in populations, despite their often-proclaimed advantages over other methods"[73].

When conducting our study, many practical issues had to be addressed and new questions were raised. We believe that studies in the directions mentioned below could yield substantial results to further elucidate the etiology of colorectal cancer.

First, we could study the origin of genetic aberrations of the DNA, as we aimed for in the study on K-*ras*. Measuring the characteristic mutational load (mutated copies per total number of copies) in non-tumerous tissue with high sensitivity, may generate new hypotheses[46, 47] regarding the etiology of specific mutations. This approach may work if there is a highly specific spectrum of mutations that is associated with the exposure, *e.g.* as for smoking and G:C to T:A transversions[46, 47]. However, for many exposures the DNA-damage may be less specific for a particular gene, and sufficiently sensitive techniques may not be available for all types of mutations. Furthermore, mutations occuring later in during the carcinogenic process may not be picked up when studying normal tissue, as it is possible that mutations occur in a specific order. Therefore, large databases on combinations of mutational aberrations in tumors could be established and used to identify patterns and formulate hypotheses on their common background; subsequently, these hypotheses can be evaluated in toxicological studies.

Knowledge of detoxification pathways at the biochemical level can provide leads for studying associations between presumed modifying factors of such pathways with parameters representing the pathway, as we did successfully in the study on fruits, vegetables and the glutathione *S*-transferase detoxification system. Additional short-term trials may evaluate whether modulation of the pathway results in lower levels of DNA damage (*e.g.* oxidative-stress related adducts), cell proliferation or, if sufficiently developed, early biomarkers of the carcinogenic process in tumor cells, urine or blood. Evaluating such associations in people at high risk of colorectal cancer (*e.g.* HNPCC or FAP patients) may provide further insight in the modulation of pathways that are probably relevant to carcinogenesis.

Second, adenoma occurrence could be studied in more detail. Since it is important to further delineate the predictive value of adenomas for colorectal cancer, it is important to study the dynamics of adenoma formation, growth, regression and recurrence. This can be done as we did in the Health Professionals Follow-Up Study (Chapters 4 and 5), although improvements in design of future studies and modeling approaches would be helpful. Also, studies on observing adenoma growth by careful surveillance of patients with adenomas[62, 66] could be repeated on a larger scale. Developments in virtual endoscopy may eventually lead to sensitive methods to detect new adenomas and follow their development until they require polypectomy. An intriguing question is whether multiple adenomas occurring in the same person have similar histopathological and molecular characteristics or whether they can be viewed as having independent etiologies. Their common characteristics may point at systematic factors in their etiology, either hereditary or environmental.

Third, cancer occurrence is the ultimate clinical endpoint we want to avoid. Family-based studies on cancer occurrence in younger people have provided leads for genetic factors (FAP & *APC*, HNPCC & MMR), and clinical prognosis of patients has also been linked to genetic factors. Our studies on K-*ras* and *hMLH1* were partly based on this idea because it has been suggested that the prognosis of colorectal cancer depends on their mutational status. Thus, determinants of tumor prognosis may indeed point at different etiologies. In fact, the same holds for adenoma recurrence studies, although it remains to be established to what extent recurrent adenomas and sporadic adenomas have similar or different etiologies.

Thus, many specific research questions shall be addressed through the application of future technological developments, and further studies. It is important that a framework be developed into which specific associations can be integrated, using markers of risk, signaling the molecular and biochemical pathways from normal tissue to malignant tumors. Such a framework, for the complex machinery of the human body, requires combining the findings - many seemingly unrelated, but fitting together with hindsight - from a large number of studies. Ideally, it would allow disease development to be described, and disease occurrence predicted, precisely. Even in its infancy, such a collated information base could be used as a solid foundation for preventative policies and actions.

References

1. Wacholder S, Silverman DT, McLaughlin JK, Mandel JS. Selection of controls in case-control studies. II. Types of controls. *Am J Epidemiol.* 1992;135(9):1029-1041.
2. Wacholder S, McLaughlin JK, Silverman DT, Mandel JS. Selection of controls in case-control studies. I. Principles. *Am J Epidemiol.* 1992;135(9):1019-1028.
3. Neugut AI, Jacobson JS, De Vivo I. Epidemiology of colorectal adenomatous polyps. *Cancer Epidemiol Biomarkers Prev.* 1993;2(2):159-176.
4. Peipins LA, Sandler RS. Epidemiology of colorectal adenomas. *Epidemiol Rev.* 1994;16(2):273-297.
5. Loeve F, Boer R, Zauber AG, et al. National Polyp Study data: evidence for regression of adenomas. *Int J Cancer.* 2004;111(4):633-639.
6. Lieberman DA, Weiss DG, Bond JH, Ahnen DJ, Garewal H, Chejfec G. Use of colonoscopy to screen asymptomatic adults for colorectal cancer. Veterans Affairs Cooperative Study Group 380. *N Engl J Med.* 2000;343(3):162-168.
7. Koretz RL. Malignant polyps: are they sheep in wolves' clothing? *Ann Intern Med.* 1993;118(1):63-68.
8. Savitz DA, Pearce N. Control selection with incomplete case ascertainment. *Am J Epidemiol.* 1988;127(6):1109-1117.
9. Hoppin JA, Tolbert PE, Taylor JA, Schroeder JC, Holly EA. Potential for selection bias with tumor tissue retrieval in molecular epidemiology studies. *Ann Epidemiol.* 2002;12(1):1-6.
10. Schroeder JC, Weinberg CR. Use of missing-data methods to correct bias and improve precision in case-control studies in which cases are subtyped but subtype information is incomplete. *Am J Epidemiol.* 2001;154(10):954-962.
11. Thompson ML, Myers JE, Kriebel D. Prevalence odds ratio or prevalence ratio in the analysis of cross sectional data: what is to be done? *Occup Environ Med.* 1998;55(4):272-277.
12. Pearce N. Effect measures in prevalence studies. *Environ Health Perspect.* 2004;112(10):1047-1050.
13. Ioannidis JP. Large scale evidence and replication: insights from rheumatology and beyond. *Ann Rheum Dis.* 2005;64(3):345-346.
14. Hertz-Picciotto I, Neutra RR. Resolving Discrepancies among Studies - the Influence of Dose on Effect Size. *Epidemiology.* 1994;5(2):156-163.
15. Martinez ME, Maltzman T, Marshall JR, et al. Risk factors for Ki-ras protooncogene mutation in sporadic colorectal adenomas. *Cancer Res.* 1999;59(20):5181-5185.
16. Jacques PF, Tucker KL. Are dietary patterns useful for understanding the role of diet in chronic disease? *Am J Clin Nutr.* 2001;73(1):1-2.
17. Lindsey JC, Ryan LM. Tutorial in biostatistics methods for interval-censored data. *Stat Med.* 1998;17(2):219-238.
18. Bellamy SL, Yi L, Ryan LM, Lipsitz S, Canner MJ, Wright R. Analysis of clustered and interval censored data from a community-based study in asthma. *Statistics in Medicine.* 2004;23(23):3607-3621.
19. Giovannucci E, Rimm EB, Stampfer MJ, Colditz GA, et a. A prospective study of cigarette smoking and risk of colorectal adenoma and colorectal cancer in U.S. men. *J.Natl.Cancer Inst.* 1994;86:183-191.
20. Willett W. *Nutritional Epidemiology.* 2nd ed. New York, NY: Oxford University Press; 1998.
21. Ocké MC, Bueno de Mesquita HB, Goddijn HE, et al. The Dutch EPIC food frequency questionnaire. I. Description of the questionnaire, and relative validity and reproducibility for food groups. *Int J Epidemiol.* 1997;26 Suppl 1:S37-S48.
22. Chasan-Taber S, Rimm EB, Stampfer MJ, et al. Reproducibility and validity of a self-administered physical activity questionnaire for male health professionals. *Epidemiology.* 1996;7(1):81-86.
23. Feskanich D, Rimm EB, Giovannucci EL, et al. Reproducibility and validity of food intake measurements from a semiquantitative food frequency questionnaire. *J Am Diet Assoc.* 1993;93(7):790-796.
24. Rimm EB, Giovannucci EL, Stampfer MJ, Colditz GA, Litin LB, Willett WC. Reproducibility and validity of an expanded self-administered semiquantitative food frequency questionnaire among male health professionals. *Am J Epidemiol.* 1992;135(10):1114-1126; discussion 1127-1136.
25. Goldbohm RA, Van den Brandt PA, Brants HA, et al. Validation of a dietary questionnaire used in a large-scale prospective cohort study on diet and cancer. *Eur J Clin Nutr.* 1994;48(4):253-265.
26. Randall E, Marshall JR, Graham S, Brasure J. High-risk health behaviors associated with various dietary patterns. *Nutr Cancer.* 1991;16(2):135-151.
27. Hu FB. Dietary pattern analysis: a new direction in nutritional epidemiology. *Curr Opin Lipidol.* 2002;13(1):3-9.
28. Newby PK, Tucker KL. Empirically derived eating patterns using factor or cluster analysis: a review. *Nutr Rev.* 2004;62(5):177-203.
29. Jurek AM, Greenland S, Maldonado G, Church TR. Proper interpretation of non-differential misclassification effects: expectations vs observations. *Int J Epidemiol.* 2005;34(3):680-687.
30. Murff HJ, Spigel DR, Syngal S. Does this patient have a family history of cancer? An evidence-based analysis of the accuracy of family cancer history. *Jama.* 2004;292(12):1480-1489.

31. Ocké MC, Bueno-de-Mesquita HB, Pols MA, Smit HA, Van Staveren WA, Kromhout D. The Dutch EPIC food frequency questionnaire. II. Relative validity and reproducibility for nutrients. *Int J Epidemiol.* 1997;26 Suppl 1: S49-58.

32. Goldbohm RA, Van 't Veer P, Van den Brandt PA, et al. Reproducibility of a food frequency questionnaire and stability of dietary habits determined from five annually repeated measurements. *Eur J Clin Nutr.* 1995;49(6): 420-429.

33. Hu FB, Manson JE, Liu S, et al. Prospective study of adult onset diabetes mellitus (type 2) and risk of colorectal cancer in women. *J Natl Cancer Inst.* 1999;91(6):542-547.

34. Ben-Shlomo Y, Kuh D. A life course approach to chronic disease epidemiology: conceptual models, empirical challenges and interdisciplinary perspectives. *Int J Epidemiol.* 2002;31(2):285-293.

35. Magnus P, Irgens LM, Haug K, Nystad W, Skjaerven R, Stoltenberg C. Cohort profile: The Norwegian Mother and Child Cohort Study (MoBa). *Int J Epidemiol.* 2006.

36. Pembrey M. The Avon Longitudinal Study of Parents and Children (ALSPAC): a resource for genetic epidemiology. *Eur J Endocrinol.* 2004;151 Suppl 3:U125-129.

37. Golding J. The Avon Longitudinal Study of Parents and Children (ALSPAC)--study design and collaborative opportunities. *Eur J Endocrinol.* 2004;151 Suppl 3:U119-123.

38. Riboli E, Kaaks R. The EPIC Project: rationale and study design. European Prospective Investigation into Cancer and Nutrition. *Int J Epidemiol.* 1997;26 Suppl 1:S6-14.

39. Smith-Warner SA, Spiegelman D, Ritz J, et al. Methods for pooling results of epidemiologic studies: the Pooling Project of Prospective Studies of Diet and Cancer. *Am J Epidemiol.* 2006;163(11):1053-1064.

40. Michels KB. Nutritional epidemiology--past, present, future. *Int J Epidemiol.* 2003;32(4):486-488.

41. Slimani N, Kaaks R, Ferrari P, et al. European Prospective Investigation into Cancer and Nutrition (EPIC) calibration study: rationale, design and population characteristics. *Public Health Nutr.* 2002;5(6B):1125-1145.

42. Steenland K, Greenland S. Monte Carlo sensitivity analysis and Bayesian analysis of smoking as an unmeasured confounder in a study of silica and lung cancer. *Am J Epidemiol.* 2004;160(4):384-392.

43. Living Tomorrow Amsterdam. http://www.livingtomorrow.nl. Accessed 10 september, 2005.

44. Pearce N, de Sanjose S, Boffetta P, Kogevinas M, Saracci R, Savitz D. Limitations of biomarkers of exposure in cancer epidemiology. *Epidemiology.* 1995;6(2):190-194.

45. Shih W, Chetty R, Tsao MS. Expression profiling by microarrays in colorectal cancer (Review). *Oncol Rep.* 2005;13(3):517-524.

46. Hainaut P, Olivier M, Pfeifer GP. TP53 mutation spectrum in lung cancers and mutagenic signature of components of tobacco smoke: lessons from the IARC TP53 mutation database. *Mutagenesis.* 2001;16(6):551-553; author reply 555-556.

47. Hussain SP, Harris CC. p53 mutation spectrum and load: the generation of hypotheses linking the exposure of endogenous or exogenous carcinogens to human cancer. *Mutat Res.* 1999;428(1-2):23-32.

48. Khrapko K, Coller HA, Andre PC, Li XC, Hanekamp JS, Thilly WG. Mitochondrial mutational spectra in human cells and tissues. *Proc Natl Acad Sci U S A.* 1997;94(25):13798-13803.

49. Noori P, Hou S, Jones IM, Thomas CB, Lambert B. A comparison of somatic mutational spectra in healthy study populations from Russia, Sweden and USA. *Carcinogenesis.* 2005;26(6):1138-1151.

50. Smith GD, Ebrahim S. Mendelian randomization: prospects, potentials, and limitations. *Int J Epidemiol.* 2004;33(1): 30-42.

51. Clayton D, McKeigue PM. Epidemiological methods for studying genes and environmental factors in complex diseases. *Lancet.* 2001;358(9290):1356-1360.

52. Brennan P. Commentary: Mendelian randomization and gene-environment interaction. *Int J Epidemiol.* 2004;33(1): 17-21.

53. Thomas DC, Conti DV. Commentary: the concept of 'Mendelian Randomization'. *Int J Epidemiol.* 2004;33(1): 21-25.

54. Rundle AG, Vineis P, Ahsan H. Design options for molecular epidemiology research within cohort studies. *Cancer Epidemiol Biomarkers Prev.* 2005;14(8):1899-1907.

55. Lüchtenborg M, Weijenberg MP, Wark PA, et al. Mutations in APC, CTNNB1 and K-ras genes and expression of hMLH1 in sporadic colorectal carcinomas from the Netherlands Cohort Study. *BMC Cancer.* 2005;5:160.

56. Ries LAG, Eisner MP, Kosary CL, et al., eds. *SEER Cancer Statistics Review, 1975-2002.* Bethesda, MD: National Cancer Institute. http://seer.cancer.gov/csr/1975_2002/ based on November 2004 SEER data submission, posted to the SEER web site 2005.

57. Schoen RE, Papachristou GI. Screening intervals for colonic neoplasia. *Curr Opin Gastroenterol.* 2003;19(1):51-56.

58. Chen CD, Yen MF, Wang WM, Wong JM, Chen TH. A case-cohort study for the disease natural history of adenoma-carcinoma and de novo carcinoma and surveillance of colon and rectum after polypectomy: implication for efficacy of colonoscopy. *Br J Cancer.* 2003;88(12):1866-1873.

59. Ponz de Leon M, Antonioli A, Ascari A, Zanghieri G, Sacchetti C. Incidence and familial occurrence of colorectal cancer and polyps in a health-care district of northern Italy. *Cancer.* 1987;60(11):2848-2859.

60. Schatzkin A, Gail M. The promise and peril of surrogate end points in cancer research. *Nat Rev Cancer.* 2002;2(1): 19-27.

61. Hill MJ, Davies GJ, Giacosa A. Should we change our dietary advice on cancer prevention? *Eur J Cancer Prev.* 2001;10(1):1-6.

62. Hofstad B, Almendingen K, Vatn M, et al. Growth and recurrence of colorectal polyps: a double-blind 3-year intervention with calcium and antioxidants. *Digestion.* 1998;59(2):148-156.

63. Almendingen K, Hofstad B, Vatn MH. Does a family history of cancer increase the risk of occurrence, growth, and recurrence of colorectal adenomas? *Gut.* 2003;52(5):747-751.

64. Almendingen K, Hofstad B, Vatn MH. Does high body fatness increase the risk of presence and growth of colorectal adenomas followed up in situ for 3 years? *Am J Gastroenterol.* 2001;96(7):2238-2246.

65. Almendingen K, Hofstad B, Vatn MH. Does intake of alcohol increase the risk of presence and growth of colorectal adenomas followed-up in situ for three years? *Scand J Gastroenterol.* 2002;37(1):80-87.

66. Almendingen K, Hofstad B, Vatn MH. Dietary habits and growth and recurrence of colorectal adenomas: results from a three-year endoscopic follow-up study. *Nutr Cancer.* 2004;49(2):131-138.

67. Hoff G, Moen IE, Trygg K, et al. Colorectal adenomas and food. A prospective study of change in volume and total mass of adenomas in man. *Scand J Gastroenterol.* 1988;23(10):1253-1258.

68. Bond JH. Progress in refining virtual colonoscopy for colorectal cancer screening. *Gastroenterology.* 2005;129(6): 2103-2106.

69. Schollnberger H, Manuguerra M, Bijwaard H, et al. Analysis of epidemiological cohort data on smoking effects and lung cancer with a multi-stage cancer model. *Carcinogenesis.* 2006;27(7):1432-1444.

70. Vineis P. Epidemiological models of carcinogenesis: the example of bladder cancer. *Cancer Epidemiol Biomarkers Prev.* 1992;1(2):149-153.

71. Loeve F, Boer R, Van Oortmarssen GJ, Van Ballegooijen M, Habbema JD. The MISCAN-COLON simulation model for the evaluation of colorectal cancer screening. *Comput Biomed Res.* 1999;32(1):13-33.

72. Van Steen K, Tahri N, Molenberghs G. Introducing the Multivariate Dale Model in Population-Based Genetic Association Studies. *Biometrical Journal.* 2004;46(2):187-202.

73. Loomis D, Wing S. Is molecular epidemiology a germ theory for the end of the twentieth century? *Int J Epidemiol.* 1990;19(1):1-3.

Summary

SUMMARY

Colorectal cancer is the third most common cancer worldwide. Many risk factors for colorectal cancer have been postulated, which include a positive family history of colorectal cancer, high intake of alcohol, and consumption of red meat and processed meat. However, the role of many factors has not yet been established, in particular those thought to have a weaker effect.

Most colorectal cancers appear to result from adenomas, which are benign lesions. During the gradual development of colorectal cancer, several changes occur to the DNA, the cell and the structure of the tissue. The resulting cancers may have different clinical and pathological characteristics, and they may have resulted from (partially) distinct pathways that are differently influenced by diet, lifestyle and heritable factors.

The aim of this thesis is to evaluate whether distinguishing colorectal tissue by its histopathological and molecular characteristics sheds further light on the role of dietary, lifestyle and heritable factors that are possibly involved in colorectal carcinogenesis. We did so by selecting five specific research questions. They referred to the study of associations between a variety of risk factors and the occurrence of specific changes in healthy tissue, adenoma tissue or colorectal cancer tissue that are thought to be related to colorectal carcinogenesis.

Specific Research Question 1:

Are parameters reflecting the activity of the rectal glutatione S-transferase detoxification system associated with consumption of fruits and vegetables at habitual exposure levels? If so, do genetic variations in glutathione S-transferase genes GSTM1 and GSTT1 modify this association?

As an example of a study that used characteristics of healthy tissue as endpoint, we focused on the activity of the glutathione (GSH)/glutathione *S*-transferase (GST) detoxification system in human rectal mucosa. Previous studies suggested that fruits and vegetables, in particular citrus fruits, brassica and allium vegetables, may increase the activity of the GSH/GST detoxification system. Whether habitual patterns and levels of consumption are also associated with higher activity was not yet known. In rectal biopsies from 94 people who underwent endoscopy, we measured GST enzyme activity, GST isoenzyme levels of GST-alpha (A1-1, A1-2 and A2-2), -mu (M1-1) and -pi (P1-1), and GSH levels, and we associated consumption of fruits and vegetables, as assessed by questionnaires, with these markers. We confirmed the study hypotheses that consumption of citrus fruits and brassica vegetables were positively associated with GST enzyme activity. The association with brassica was confined to individuals with the *GSTM1* genotype, which corresponds with the observed association between consumption of brassica vegetables and GSTM1-1 isoenzyme level. Consumption of allium vegetables was not associated with GST enzyme activity, but was negatively with GSTP1-1. Associations were similar among people with the *GSTT1*-plus and *GSTT1*-null genotype.

Specific Research Question 2:

Do associations between presumed modifiable risk factors for colorectal adenomas depend on K-ras mutational status of the adenoma?

As an example of a study that separates the endpoint adenomas according to mutational status, we focused on mutations in the growth-related K-*ras* oncogene present in adenomas. K-*ras* mutation-positive (K-*ras*$^+$) and -negative (K-*ras*$^-$) colorectal adenomas may have different pathological features, and their clinical behaviour may also differ. As environmental compounds may cause mutations in K-*ras* or affect clonal selection depending on mutational status, we evaluated whether the etiology of K-*ras*$^+$ and K-*ras*$^-$ adenomas differs. K-*ras* mutations in codons 12 and 13 were assessed in colorectal adenoma tissue (K-*ras*$^+$: n= 81, K-*ras*$^-$: n=453). Dietary and lifestyle data were collected through questionnaires that were also administered to 709 polyp-free controls. In spite of the observation that intake of vitamin B2 and monounsaturated fat were differently associated with risk of K-*ras*$^+$ and K-*ras*$^-$ adenomas, and that we found potential differences in risk for red meat, total dietary and polyunsaturated fat, dairy products, calcium, protein, tea and smoking, the overall scientific literature remains inconsistent with respect to associations between modifiable factors and K-*ras* mutational status.

Specific Research Question 3:

Are dietary and lifestyle risk factors for colorectal cancer differently associated with adenomas that are thought to have a high risk of malignancy (advanced adenomas) and adenomas that are thought to have a low risk of becoming malignant?

Third, we focused on adenomas again but this time we separated them according to their potential risk of malignancy and stage of development, arguing that the role of risk factors may vary according to the stage. In a prospective study among 26,769 American men who underwent an endoscopy, we evaluated whether risk factors for advanced (≥1cm or with any villous characteristics) and non-advanced colorectal adenomas differed after 17 years of follow-up. We observed that most dietary and lifestyle factors were similarly associated with advanced and non-advanced adenomas. Smoking was more strongly associated with advanced adenomas than with non-advanced adenomas, and a high glycemic index was inversely associated with advanced adenomas but not with non-advanced adenomas.

Specific Research Question 4:

Do people with a family history of colorectal cancer have a higher risk of advanced adenomas or do they mostly have a higher likelihood of getting multiple adenomas?

Fourth, we focused on adenoma advancement and adenoma multiplicity in relation to a positive family history of colorectal cancer as a marker for inherited genetic susceptibility to colorectal cancer. Data from the same prospective cohort study as used for research question 3 were analyzed. Associations with family history of colorectal cancer were stronger for men with multiple distal adenomas than for men with a single distal adenoma at first diagnosis.

Associations between family history and advanced and non-advanced adenomas were of similar strength, but a tendency towards a somewhat stronger association with non-advanced adenomas was found.

Specific Research Question 5:

Is consumption of fruits and vegetables differentially associated with colon cancer with and without defects in the mismatch repair system?

Last, but not least, we focused on different pathways in colorectal carcinogenesis by distinguishing colon carcinomas with and without expression of the mismatch repair gene hMLH1, which have different clinical and pathological characteristics. As previous animal and *in vitro* studies suggest that fruits, vegetables, folate, and antioxidants are associated with colonic expression of mismatch repair genes, we evaluated associations between consumption of fruits and vegetables and hMLH1 protein-deficient and -proficient colon cancer in a cohort study among 120,852 individuals ages 55 to 69 years who completed a questionnaire at baseline. Using immunohistochemistry, hMLH1 protein expression was assessed in colon cancer tissue obtained from 441 patients who were identified over 7.3 years of follow-up, excluding cases occurring in the initial 2.3 years. hMLH1 protein expression was absent in 54 tumors (12.2%) and present in 387 tumors. We observed an inverse association between fruit consumption and colon cancer, which was confined to the subgroup with hMLH1 protein-deficiency. Total consumption of vegetables was not associated with either type of tumor and no associations were observed for folate, fiber, antioxidants, or subgroups of vegetables.

In this thesis, we showed that distinguishing colorectal tissue by its histopathological and molecular characteristics may indeed shed further light on the role of dietary, lifestyle and heritable factors that are possibly involved in colorectal carcinogenesis. We observed that studying histopathological and molecular endpoints can address some of the weaknesses of traditional epidemiology, but incorporating them in the design brings an additional layer of complexity. When clearly defined hypotheses and outcomes are studied within a larger framework of carcinogenesis, the approach is likely to add further knowledge. Such a framework, however, has to originate from integration of knowledge on the carcinogenic process within DNA, at the cellular level, and at the tissue level; such integration can be enhanced by improved methodology and a broad overview of the possibilities and limitations of scientific research.

Samenvatting

SAMENVATTING

Dikkedarmkanker is de op drie na meest voorkomende soort kanker in de wereld. Bepaalde voedings- en leefgewoonten kunnen de kans op dikkedarmkanker mogelijk verhogen, waaronder het drinken van grote hoeveelheden alcohol en het eten van rood of bewerkt vlees. Ditzelfde geldt voor bepaalde erfelijke factoren. De rol van de meeste van deze factoren is echter nog niet precies duidelijk. Dit geldt met name voor de factoren waarvan gedacht wordt dat ze een kleine invloed hebben.

De meeste dikkedarmkankers lijken voort te komen uit adenomateuze dikkedarmpoliepen ("adenomen"), wat goedaardige afwijkingen zijn. Tijdens de geleidelijke ontwikkeling tot dikkedarmkanker treden vele veranderingen op in het DNA, de cel en de structuur van het darmweefsel. De resulterende kankers kunnen verschillende klinische en pathologische kenmerken hebben. Dit kan komen doordat ze (deels) langs een verschillende weg zijn ontstaan. Ze kunnen bijvoorbeeld op een andere manier zijn beïnvloed door voedings-, leefstijl- en erfelijke factoren.

Het doel van dit proefschrift is om na te gaan of het indelen van darmweefsel op basis van histopathologische en moleculaire kenmerken meer inzicht verschaft over de rol van voedings-, leefstijl- en erfelijke factoren bij het ontstaan van dikkedarmkanker. Om dat te doen hebben we vijf specifieke onderzoeksvragen geformuleerd. In de bijbehorende onderzoeken associeerden we een verscheidenheid van mogelijke risicofactoren met de aan- of afwezigheid van specifieke veranderingen in gezond weefsel, adenoomweefsel en dikkedarmkankerweefsel. De specifieke veranderingen die we bestudeerden, worden van belang geacht voor het ontstaansproces van dikkedarmkanker.

Onderzoeksvraag 1:

Zijn merkers die de activiteit van het glutathion-S-transferase-ontgiftingssysteem in de endeldarm reflecteren, gerelateerd aan de gebruikelijke hoeveelheden van consumptie van groente en fruit? Zo ja, zijn deze verbanden afhankelijk van de genetische variatie in de glutathion-S-transferase-genen GSTM1 and GSTT1?

Als voorbeeld van een onderzoek waarin kenmerken van gezond weefsel als eindpunt zijn beschouwd, hebben we ons gericht op de activiteit van het 'glutathion (GSH)/glutathion-*S*-transferase (GST)'-ontgiftingssysteem in het slijmvlies van de endeldarm. Enzymen zoals glutathion-*S*-transferase zorgen ervoor dat schadelijke stoffen worden omgezet in onschadelijke stoffen en ons lichaam zo snel mogelijk verlaten. Eerder onderzoek suggereerde dat groenten en fruit - met name citrusvruchten, alliumgroente (zoals ui, prei, knoflook) en brassicagroenten (koolsoorten, zoals broccoli, bloemkool, spruitjes) - de activiteit van het GSH/GST-systeem kunnen verhogen. Dit eerdere onderzoek was gebaseerd op ongewoon grote hoeveelheden groenteconsumptie en het was nog niet bekend of dit ook geldt voor meer gebruikelijke hoeveelheden en patronen van consumptie. In stukjes weefsel van de endeldarm van 94 mensen die een endoscopie ondergingen, hebben we de activiteit van het enzym GST alsmede niveaus

van de enzymeiwitten GST-alpha (A1-1, A1-2 and A2-2), -mu (M1-1) en -pi (P1-1) en GSH gemeten. Inderdaad vonden we dat mensen met een hoge consumptie van brassicagroenten en citrusvruchten hogere waardes van GST-enzymactiviteit hadden. Het verband tussen brassicagroenten en GST-enzymactiviteit was echter alleen aanwezig bij mensen met het zogenaamde *GSTM1*-genotype. Ook vonden we dat mensen die veel brassicagroenten aten, hogere niveaus van het enzymeiwit GSTM1-1 hadden. Voor alliumgroenten konden we geen verband met GST-enzymactiviteit aantonen. Mensen die veel alliumgroenten aten, hadden echter lagere waarden van het enzymeiwit GSTP1-1. De verbanden hingen niet af van het *GSTT1* genotype.

Onderzoeksvraag 2:

Hangen verbanden tussen veronderstelde modificeerbare risicofactoren voor dikkedarmkanker en het vóorkomen van adenomen af van de aan- of afwezigheid van K-ras mutaties in de tumor?

Als een voorbeeld van een onderzoek dat adenomen indeelt op basis van de aan- of afwezigheid van mutaties in adenomen, richtten we ons op het K-*ras*-gen. Als er een verandering in het K-*ras*-gen optreedt in een tumor, bevordert dit zijn groei. Adenomen met en zonder een mutatie in het K-*ras*-gen kunnen verschillende pathologische kenmerken hebben en klinisch kunnen ze zich ook anders gedragen. Aangezien omgevingsfactoren mutaties in het K-*ras*-gen kunnen veroorzaken, probeerden we aanwijzingen te krijgen of adenomen met en zonder een K-*ras*-mutatie op een andere manier ontstaan. We verzamelden adenoomweefsel, waarin we nagingen of er een K-*ras*-mutatie in codons 12 en 13 aanwezig was; dat was het geval bij 81 patiënten en niet het geval bij 453 patiënten. Deze patiënten - en tevens 709 mensen die ook een endoscopie ondergingen, maar geen poliepen hadden - vulden voor ons vragenlijsten in, waarmee we informatie verkregen over hun voedings- en leefwijze. We observeerden dat de inname van vitamine B2 en enkelvoudig onverzadigd vet anders samenhing met adenomen mét en zonder een K-*ras*-mutatie. Dat was mogelijk ook het geval voor rood vlees, vetiname (totaal en meervoudig onverzadigd vet), zuivelprodukten, calciuminname, eiwitinname, theeconsumptie en roken. In de wetenschappelijke literatuur vonden we echter grote verschillen in de resultaten van vergelijkbare onderzoeken naar risicofactoren en tumoren met of zonder een K-*ras*-mutaties. Het algemene beeld met betrekking tot de relatie tussen voedings- en leefstijlfactoren en K-*ras*-mutaties blijft daardoor inconsistent.

Onderzoeksvraag 3:

Zijn voedings- en leefstijlfactoren anders geassocieerd met adenomen waarvan vermoed wordt dat ze een hoge kans hebben om zich tot kanker te ontwikkelen (gevorderde adenomen) dan met adenomen waarvan gedacht wordt dat ze een lage kans hebben om door te groeien tot kanker?

In het derde onderzoek richtten we ons opnieuw op adenomen, maar dit keer definieerden we groepen adenomen die een hogere of lagere kans hadden om door te groeien tot kanker. Het idee hierachter was dat de rol van risicofactoren per ontwikkelingsstadium van een tumor kan verschillen. In een onderzoek onder 26.769 Amerikaanse mannen, die een endoscopie

ondergingen en 17 jaar werden gevolgd, gingen we na of risicofactoren voor gevorderde adenomen (≥1cm en/of met een villeus kenmerk of een carcinoom in situ) en niet-gevorderde adenomen van elkaar verschilden. We observeerden dat de meeste voedings- en leefstijlfactoren op dezelfde manier geassocieerd waren met gevorderde en niet-gevorderde adenomen. Roken liet een sterker verband zien met gevorderde adenomen dan met niet-gevorderde adenomen. Een hoge glycemische index daarintegen liet een invers verband zien met gevorderde adenomen, maar niet met niet-gevorderde adenomen.

Onderzoeksvraag 4:

Hebben mensen met een familiegeschiedenis van dikkedarmkanker een hogere kans op gevorderde adenomen of hebben zij vooral een hogere kans om meerdere adenomen te krijgen?

Ten vierde richtten we ons op de ontwikkelingstadia van adenomen en op het aantal adenomen in relatie tot het hebben van een eerstegraadsfamilielid met dikkedarmkanker, als een merker voor erfelijke gevoeligheid voor dikkedarmkanker. Hiervoor maakten we gebruik van dezelfde prospectieve cohortstudie als voor vraagstelling 3. Het hebben van een familielid met dikkedarmkanker leek samen te hangen met het hebben van meerdere adenomen in het laatste, distale, gedeelte van de darm bij de eerste diagnose. We konden niet aantonen dat de verbanden tussen het hebben van een familiegeschiedenis van dikkedarmkanker en de kans op gevorderde en niet-gevorderde adenomen anders waren, al leken de verbanden met niet-gevorderde adenomen wat sterker.

Onderzoeksvraag 5:

Is het verband tussen groente- en fruitconsumptie en de kans op colonkanker afhankelijk van de aan- of afwezigheid van een goed functionerend mismatchrepairsysteem?

Als laatste richtten we onze aandacht op verschillende routes van het ontstaan van dikkedarmkanker door onderscheid te maken van dikkedarmkanker op basis van de aan- of afwezigheid van het mismatchrepaireiwit hMLH1. Mismatchrepaireiwitten, zoals hMLH1, spelen een rol bij het herstellen van fouten die optreden tijdens de replicatie van het DNA. Tumoren waarin zo'n eiwit wordt aangetoond en tumoren waarin zo'n eiwit niet wordt aangetoond, zien er anders uit, ze reageren ook anders op behandeling en ze hebben een ander klinisch verloop. In dierexperimenteel onderzoek en onderzoek met celkweken zijn aanwijzingen gevonden dat groenten, fruit, foliumzuur en antioxidanten de expressie van de mismatchrepaireiwitten kunnen beïnvloeden. Daarom gingen we na of consumptie van groente en fruit samenhing met het vóorkomen van colonkanker waarin het hMLH1-eiwit werd aangetoond en met het vóorkomen van colonkanker waarin dit eiwit niet werd aangetoond. We volgden 120.852 mensen in de leeftijd van 55 tot 69 jaar die bij aanvang van het onderzoek een vragenlijst hadden ingevuld over voedingsgewoonten en risicofactoren voor kanker. Zeven (7.3) jaar na aanvang van het onderzoek verzamelden we weefsel van 441 patiënten bij wie colonkanker werd vastgesteld, waarbij we patiënten die in de eerste 2.3 jaar werden gediagnosticeerd niet meenamen. Met behulp van kleuringstechnieken ("immunohistochemie") bepaalden we of het hMLH1-eiwit

al dan niet aanwezig was in het colonkankerweefsel: in 54 (12.2%) van de tumoren was dit eiwit afwezig en in 387 tumoren aanwezig. We vonden dat mensen die veel fruit aten een lagere kans hadden op colonkanker waarin het hMLH1-eiwit afwezig was, maar niet op colonkanker waarin dit eiwit aanwezig was. In dit onderzoek konden we echter niet aantonen dat groenten, foliumzuur, vezel en antioxidanten de kans op colonkanker met en zonder aanwezigheid van dit eiwit beïnvloedden.

In dit proefschrift hebben we laten zien dat onderscheid maken in darmweefsel op basis van de histopathologische en moleculaire kenmerken inderdaad meer duidelijkheid kan verschaffen over de rol van voedings- en leefgewoonten en erfelijke factoren die mogelijk een rol spelen bij het ontstaan van dikkedarmkanker. Het bestuderen van histopathologische en moleculaire eindpunten kan bepaalde zwakke punten van traditioneel epidemiologisch onderzoek verbeteren, maar het meenemen van zulke eindpunten maakt het onderzoek op bepaalde punten ook extra lastig. Als duidelijk gedefinieerde onderzoeksvraagstellingen en eindpunten worden bestudeerd binnen een groter kader van de ontstaanswijze van dikkedarmkanker, dan kan deze benadering verdere kennis toevoegen. Zo'n kader moet kennis bevatten over de ontstaanswijze van darmkanker op het niveau van het DNA, de cel en het weefsel, gekoppeld aan de wijze waarop dit in epidemiologisch onderzoek zichtbaar gemaakt kan worden; een dergelijke. integratie kan worden bevorderd door verbeterde onderzoeksmethodologie en van het hebben van een breed overzicht van de mogelijkheden en beperkingen van wetenschappelijk onderzoek.

Consideration

"As we express our gratitude, we must never forget that the highest appreciation is not to utter words, but to live by them."

John F. Kennedy

Acknowledgements

Any epidemiological study relies on a team of researchers, study participants, administrative and medical staff, and institutions, especially when it involves multidisciplinary research. The studies described in this thesis are no exception, as illustrated by the large number of co-authors, people and institutions listed in the acknowledgements in chapters 2 to 6. I am grateful to all these people and institutions, as without them this thesis would never have seen completion. In particular, I would like to thank the following people:

- Prof. Dr. Pieter van 't Veer, for being my intellectual guide; for challenging discussions of methodology; and his never-ending enthusiasm and encouragement.
- Prof. Dr. Frans Kok, for his broad overviews of science; and challenging the limits of my perceptions.
- Dr. Matty Weijenberg, for accurate and precise feedback on manuscripts; and facilitating communication between Maastricht and Wageningen.
- Prof. Dr. Edward Giovannucci, for being my scientific mentor; for inspiration; and thorough and timely feedback.
- Dr. Ellen Kampman, for shaping my thinking about the biological mechanisms involved in carcinogenesis.
- The *POLIEP*-study team at the Division of Human Nutrition of Wageningen University, in particular Dr. Ellen Kampman, Dr. Brenda Diergaarde, Ms Mariken Tijhuis, Dr. Edine Tiemersma, Dr. Maureen van den Donk and Mrs Maria van Vugt, for jointly producing a superb dataset and a pleasant work environment.
- The team of the Netherlands Cohort Study on diet and cancer (NLCS) at Maastricht University and TNO Quality of Life, in particular Prof. Dr. Piet van den Brandt, Dr. Sandra Bausch-Goldbohm, Dr. Margreet Lüchtenborg and Dr. Matty Weijenberg, for giving me the opportunity to work with them and their inspiring dataset.
- Dr. Ton de Goeij, Dr. Mirian Brink, Mr Guido Roemen, Prof. Dr. Adriaan de Bruïne, and their colleagues at the Department of Pathology of Maastricht University, for collecting and handling an impressive number of tumor blocks from the NLCS; and their scientific input.
- Dr. Goos van Muijen, Ms Hanneke Braam and Prof. Dr. Han van Krieken at the Department of Pathology of the Radboud University Nijmegen Medical Centre, for coordination of retrieval of adenoma tissue, conducting many laboratory assessments, reading numerous slides and supervising the hMLH1 assessment; and for being a joy to work with.
- Mr Jan Harryvan and Ms Annelies Bunschoten at the Division of Human Nutrition of Wageningen University, for efficiently handling many blood samples and genotyping.
- Dr. Wilbert Peters and his colleagues at the Department of Gastroenterology of the Radboud University Nijmegen Medical Centre, for their engaging support with the GST enzyme paper.

- Miss Wieke van der Kuil, Miss Janneke Ploemacher, Miss Colette Smit, Miss Olga Souverein and Miss Gerda van Wijhe, for conducting their MSc research within the framework of my PhD project; for their stimulating and valuable input; and help with data analysis.
- Prof. Dr. Walter Willett, Prof. Dr. Edward Giovannucci, Dr. Kana Wu, Ms Barbara Vericker, and their colleagues for welcoming me to their department at the Harvard School of Public Health, Boston, MA; and for making my stay there a scientifically enlightening, and memorable, experience.
- Prof. Julian Peto and Dr. Isabel dos Santos Silva at the Non-Communicable Disease Epidemiology Unit of the London School of Hygiene and Tropical Medicine, for providing a flexible environment which allowed me to complete this PhD thesis while broadening my scope by working on other projects with them.
- Miss Irene van der Meer and Miss Wieke Ormel, for being my paranymphs; and for their care, friendship and support.
- Mr Rick Yagodich, for proofreading, editing and linguistic advice; for the professional layout of this thesis; and commitment, encouragement and friendship.
- Miss Eva Radix, for impressive graphic design work on the cover.
- Mr Lewis Palmer, for kindly proofreading the introduction.
- My colleagues and co-workers at the Division of Human Nutrition of Wageningen University; the Department of Pathology of the Radboud University Nijmegen Medical Centre and the Department of Nutrition of the Harvard School of Public Health, for making each a pleasant place to work.
- My family and friends, for believing in me.
- And last, but by no means least, the hospitals, physicians, nursing and administrative staff, registries and funding bodies, and the many participants involved in the studies described throughout this thesis.

Thanks to everybody who was involved in one way or another, from the bottom of my heart.

Petra Wark.

Curriculum Vitae

ABOUT THE AUTHOR

Petra Ariane Wark was born on 3 September, 1976 in Amsterdam, the Netherlands. After completing grammar school (Stedelijk Gymnasium Schiedam) in 1995, she pursued a MSc degree in Biomedical Health Sciences from the University of Nijmegen (now: Radboud University), specialising in epidemiology, with minors in oncology, nutrition and psychology. Her MSc included two training periods during which she gained insight into the diagnosis, treatment and clinical course of cancer, and the workings of the Netherlands Cancer Registry. Furthermore, she contributed to the writing of two grant proposals, and conducted research on colorectal adenomas (using data from the case-control study referenced in Chapter 3) at the Division of Human Nutrition of Wageningen University, which led to two MSc theses and a second-authorship.

Following her graduation in May 2000, Petra reprised her relationship with Wageningen to begin working towards her PhD. She worked on a project involving the Division of Human Nutrition of Wageningen University, the Departments of Epidemiology and Pathology of Maastricht University and the Department of Pathology of the Radboud University Nijmegen Medical Centre; the core of which was funded by the Netherlands Organisation for Health Research and Development. Working four days per week on her PhD, Petra dedicated the fifth to pursuing 16 courses comprised in the MSc programs in Applied Statistics and Biostatistics at the Limburg Universitair Centrum (now: Hasselt University) in Diepenbeek, Belgium. During her PhD studies, she obtained a grant from the Dutch Cancer Society to work at the Department of Nutrition, Harvard School of Public Health, Boston, MA as guest researcher for nine months (April-December 2003); which work is also included in this thesis. Petra was a member of the organising committee of a 10-day scientific study tour for PhD students to Germany, Switzerland and Italy, and a member of the International Student Committee of Wageningen University's Division of Human Nutrition.

Since October 2005, Petra has worked as a research fellow in the Cancer Research UK Epidemiology and Genetics Group within the Non-Communicable Disease Epidemiology Unit of the London School of Hygiene and Tropical Medicine in the United Kingdom.

LIST OF PUBLICATIONS

Peer-Reviewed Publications

- Wark PA, Van der Kuil W, Ploemacher J, Van Muijen GNP, Mulder CJJ, Weijenberg MP, Kok FJ, Kampman E. Diet, lifestyle and risk of K-ras mutation-positive and -negative colorectal adenomas. *Int J Cancer. 2006*; 119(2)*:* 398-405.
- Wark PA, Weijenberg MP, Van 't Veer P, Van Wijhe G, Lüchtenborg M, Van Muijen GNP, De Goeij AFPM, Goldbohm RA, Van den Brandt PA. Fruits, vegetables and hMLH1 protein-deficient and -proficient colon cancer: the Netherlands Cohort Study. *Cancer Epidemiol Biomarkers Prev.* 2005;14(7):1619-25.
- Tijhuis MJ, Wark PA, Aarts JMMJG, Visker MHPW, Nagengast FM, Kok FJ, Kampman E. GSTP1 and GSTA1 polymorphisms interact with cruciferous vegetable intake in colorectal adenoma risk. *Cancer Epidemiol Biomarkers Prev.* 2005;14:2943-51.
- Lüchtenborg M, Weijenberg MP, De Goeij AFPM, Wark PA, Brink M, Roemen GMJM, Lentjes MHFM, De Bruïne AP, Goldbohm RA, Van 't Veer P, Van den Brandt PA. Meat and fish consumption, APC gene mutations and hMLH1 expression in colon and rectal cancer: a prospective cohort study (the Netherlands). *Cancer Causes Control.* 2005;16(9): 1041-54.
- Lüchtenborg M, Weijenberg MP, Wark PA, Saritas AM, Roemen GMJM, Van Muijen GNP, De Bruïne AP, Van den Brandt PA, De Goeij AFPM. Mutations in APC, CTNNB1 and K-ras genes and expression of hMLH1 in sporadic colorectal carcinomas from the Netherlands Cohort Study. *BMC Cancer.* 2005;5:160.
- Wark PA, Grubben MJAL, Peters WHM, Nagengast FM, Kampman E, Kok FJ, Van 't Veer P. Habitual consumption of fruits and vegetables: associations with human rectal glutathione *S*-transferase. *Carcinogenesis.* 2004; 25(11):2135-42.
- Tiemersma EW, Wark PA, Ocké MC, Bunschoten A, Otten MH, Kok FJ, Kampman E. Alcohol consumption, alcohol dehydrogenase 3 polymorphism, and colorectal adenomas. *Cancer Epidemiol Biomarkers Prev.* 2003; 12(5):419-25.

Selected Abstracts and Other Publications

- Wark PA, Wu K, Van 't Veer P, Fuchs CS, Willett WC, Giovannucci EL. Diet, lifestyle, and advanced and non-advanced colorectal adenomas in men. *J Epidemiol Community Health.* 2004; 58 Suppl I: A60: #197 & *Proceedings of the AACR Frontiers in Cancer Prevention Conference.* 2004, Seattle, WA:, p133: #B91
- Wark PA, Souverein OW, Van Steen K, Van 't Veer P. [Evaluating etiological heterogeneity of cancer using multinomial logistic regression] [In Dutch]. In: *Abstractboek WEON 2002. Het individu in de populatie: heterogeniteit, kwantiteit en toepasbaarheid.* ISBN 90-9015597-x, p. 36.

- Bakker-Zierikzee A, Geelen A, Mars M, Pellis L, Rutten R, Wark P. [Diversity within European research on nutrition]. [In Dutch]. *Voeding Nu* 2002;9:26-7.
- Weijenberg MP, Brink M, Lüchtenborg M, Wark PA, De Goeij AFPM, De Bruïne AP, Van 't Veer P, Kampman E, Van Muijen GNP, Goldbohm RA, Van den Brandt PA. Dietary factors, genetic susceptibility and somatic mutations in colorectal cancer: a prospective study. *IARC Sci Publ.* 2002; 156:503-4.
- Wark PA, Lüchtenborg M, Van den Brandt PA , Weijenberg MP, Van Muijen GNP, Arends JW, De Goeij AFPM , Kampman E, Goldbohm RA, Kok FJ, Van 't Veer P. Dietary factors and genetic susceptibility as determinants of mutation of cancer genes in colorectal adenomas and carcinomas. *Eur J Cancer Prev.* 2000; 9(6):444.

OVERVIEW OF COMPLETED TRAINING ACTIVITIES
GRADUATE SCHOOL VLAG

Discipline-Specific Activities

- Frontiers in Cancer Prevention Conference, AACR, Seattle, WA, US, 2004
- European Congress of Epidemiology, IEA - EEF, Porto, Portugal, 2004
- Pre-conference Course 'Life Course of Epidemiology', Porto, Portugal, 2004
- Course 'Causal Modeling', Erasmus Summer Programme, Rotterdam, 2004
- Guest researcher at Harvard School of Public Health, Boston, MA, US, April-December 2003
- Symposium of the Environmental Statistics Program, Harvard School of Public Health, Sheffield, MA, US, 2003
- Masterclass 'From Nutrigenomics to Healthy Food', VLAG/NUTRIM, Maastricht, 2002
- Short Course 'Advanced Linkage Analysis', Limburgs Universitair Centrum, Diepenbeek, Belgium, 2002
- Symposium 'the Future of Molecular Epidemiology', VLAG, Wageningen, 2002
- Working with immunohistochemistry, Radboud University Nijmegen Medical Centre, Nijmegen, 2002
- Course 'Methodological aspects of case-control and case-cohort studies', RIVM, 2001
- Meeting of the 3 Country Corner (Local group of Royal Statistical Society) on Statistical Genetics, Diepenbeek, Belgium, 2001
- Annual meeting of the workgroup 'Nutrition' of the Netherlands Organisation for Health Research and Development, Papendal, 2001
- First Belgian-Dutch Biometrical conference, Belgian and Dutch Regions of the International Biometric Society, Peer, Belgium, 2001
- Wageningen Biometrics Colloquium on Epidemiology, Biometrics, Wageningen, 2001
- 3rd International Course on Molecular Epidemiology, IARC, Lyon, France, 2000
- 18th symposium of the ECP `Precancerous lesions of the digestive tract', Maastricht, 2000
- Experts meeting at IARC, Lyon, France, 2000
- Annual meetings of the Netherlands Epidemiological Society, 2000-02, 2004, 2005
- Annual meetings of the Netherlands Organisation for Health Research and Development (on priority program Nutrition and Chronic Diseases), 2000-04

General Courses

- WGS course 'Career perspectives', Wageningen, 2004
- Postdoctoral Fall Seminar Series, Harvard School of Public Health, Boston MA, USA, 2003
- Organizing and supervising MSc thesis work, Wageningen, 2002
- PhD student week VLAG, Nijmegen, 2000

Optional Activities

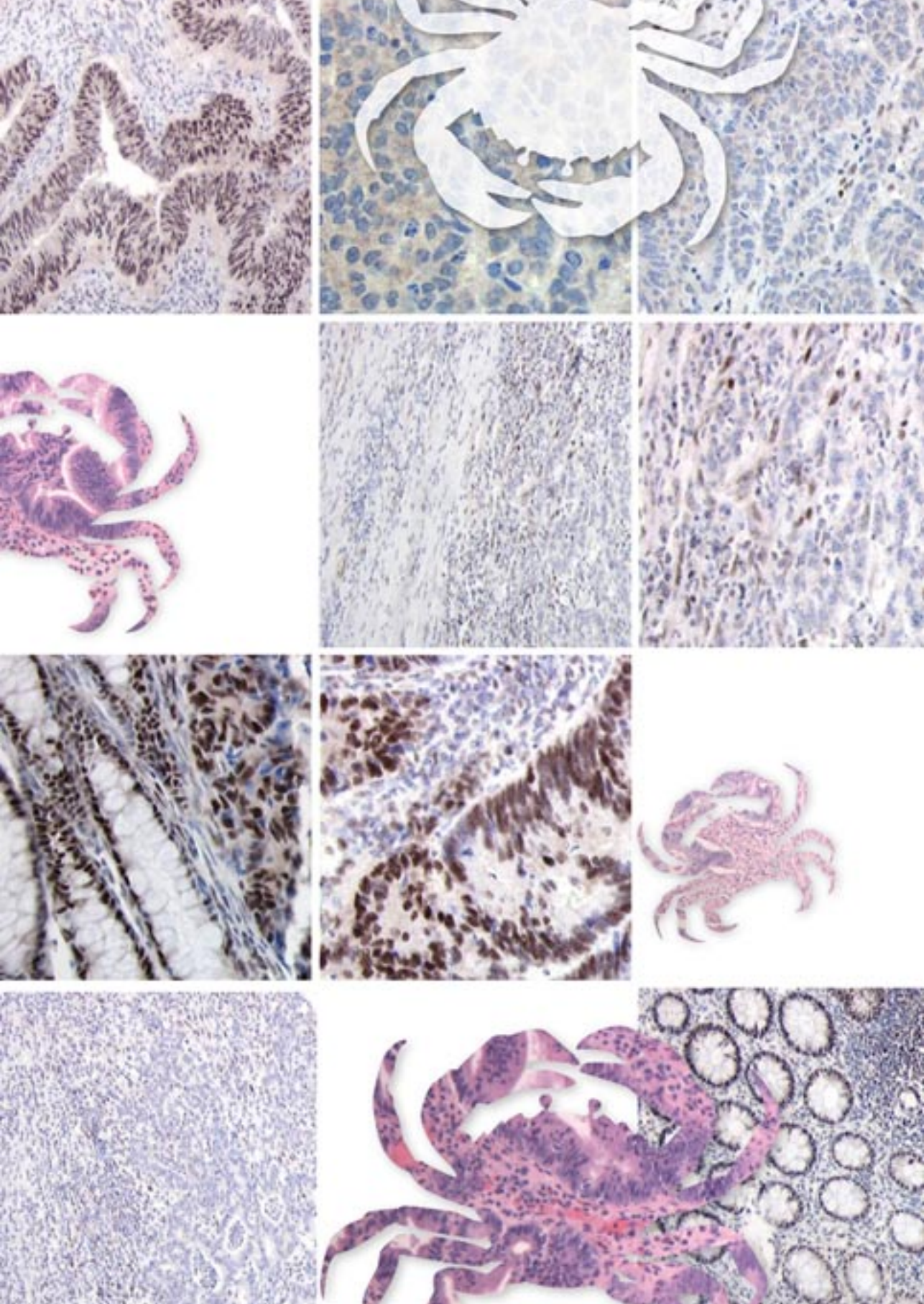
- Various general scientific meetings at Wageningen University, Radboud University Nijmegen Medical Centre and Harvard School of Public Health, 2000-2005
- Mini-symposium `Nutrition and Antisocial Behaviour among Young Adult Prisoners', Division of Human Nutrition, Wageningen, 2004
- The Harvard Forums on Health 'Focus on Obesity', Boston MA, USA, 2003
- PhD study tour to Switzerland, Italy and Germany, Division of Human Nutrition, Wageningen University, 2001
- Preparation PhD research proposal, Wageningen, 2000

The research described in this thesis was part of a broader research project that was funded by the Netherlands Organisation for Health Research and Development (grant number 980-10-026). Further financial support for the studies in this thesis was provided by the Dutch Cancer Society, the Netherlands Digestive Diseases Foundation and the NCI.

Financial support, by Wageningen University and the Netherlands Organisation for Health Research and Development, for the publication of this thesis is gratefully acknowledged.

Cover design: Miss Eva Radix, Tilburg, the Netherlands
Layout: Mr Rick Yagodich, Twickenham, United Kingdom
Photos: Prof. Dr. J.H.J.M. van Krieken and Miss Gerda van Wijhe, Nijmegen, the Netherlands

Printing: Ponsen & Looijen B.V., Wageningen, the Netherlands



STELLINGEN

1. Erfelijke factoren spelen een prominentere rol in de vroege stadia van de colorectale carcinogenese dan in de latere stadia.
(*dit proefschrift*)

2. Als pathologisch onderzoek meer gebruik zou maken van een epidemiologische benadering, dan zou dit de voedingsepidemiologie vooruit kunnen helpen.
(*dit proefschrift*)

3. De vooruitgang van de medische wetenschap heeft weliswaar bijgedragen aan een hogere levensverwachting, maar heeft ook de beleving van wat gezond is, veranderd.

4. Het spelen van strategische spellen, lezen en het oplossen van puzzels moet worden gestimuleerd bij volwassenen, aangezien dergelijke activiteiten geassocieerd zijn met het behoud van cognitief functioneren.
(n.a.v. Rundek T, *Neurology.* 2006;66(6):794–5)

5. De invulling die in verschillende landen en instituten aan een PhD-opleidingsprogramma wordt gegeven, zou moeten uitmonden in vergelijkbare eindtermen die uit een minimum set van algemeen academische en vakinhoudelijke vaardigheden en competenties bestaan.

6. "The distinction between 'the social' and 'the scientific' is itself an artful contrivance of scientists."
(Shaplin S, *Med Hist.* 1981;25(3):342)

7. "I am not an outlier; I just haven't found my distribution yet."
Ronan M. Conroy

Behorend bij het proefschrift:
"Diet, Lifestyle, Heritable Factors and Colorectal Carcinogenesis:
Associations with Histopathological and Molecular Endpoints"

Te verdedigen door Petra A. Wark op 16 januari 2007 in Wageningen.

PROPOSITIONS

1. Heritable factors play a more prominent role in the earlier, rather than later, stages of colorectal carcinogenesis.
(*this thesis*)

2. If pathological research would more extensively incorporate an epidemiological approach, it may contribute to the progress of nutritional epidemiology.
(*this thesis*)

3. Notwithstanding that the progress of medical science has contributed to a prolonged lifespan, it has also changed the perception of healthiness.

4. Playing strategy games, reading and solving puzzles should be encouraged in adults, given that such activities are associated with the maintenance of cognitive function.
(based on Rundek T, *Neurology.* 2006;66(6):794–5)

5. The structure of various PhD-training programs should, irrespective of country and institution, result in comparable evaluation criteria that consist of a minimum set of general academic, and subject-specific, skills and competencies.

6. "The distinction between 'the social' and 'the scientific' is itself an artful contrivance of scientists."
(Shaplin S. *Med Hist.* 1981;25(3):342)

7. "I am not an outlier; I just haven't found my distribution yet."
Ronan M. Conroy

Belonging to the thesis:
"Diet, Lifestyle, Heritable Factors and Colorectal Carcinogenesis:
Associations with Histopathological and Molecular Endpoints"

To be defended by Petra A. Wark in Wageningen on January 16, 2007.