Foto: Jacob Forsell

Karolinska Institutet

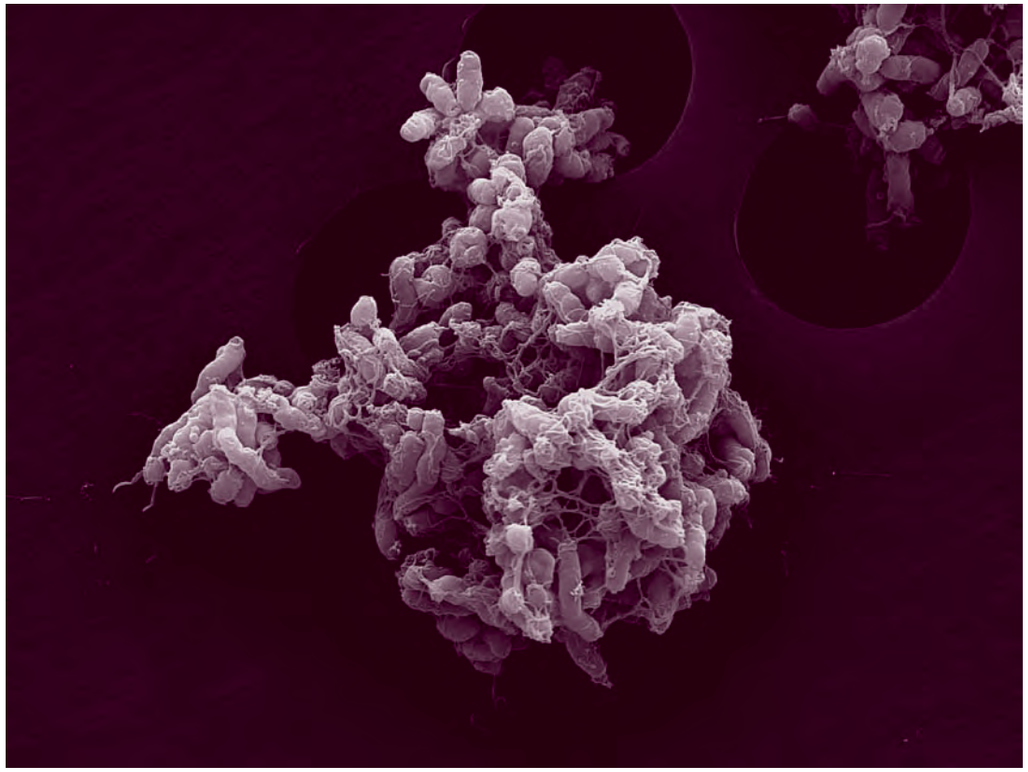
Thesis for doctoral degree (Ph.D.) 2009

ROLES OF *H. PYLORI* INFECTION, HOST GENETIC VARIATION, & ENVIRONMENTAL EXPOSURES IN GASTRIC CANCER

Christina Persson

Thesis for doctoral degree (Ph.D.)
2009

ROLES OF *HELICOBACTER PYLORI* INFECTION, HOST GENETIC VARIATION, AND OTHER ENVIRONMENTAL EXPOSURES IN GASTRIC CARCINOGENESIS



Christina Persson

Karolinska Institutet

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

ROLES OF *HELICOBACTER PYLORI INFECTION*, HOST GENETIC VARIATION, AND OTHER ENVIRONMENTAL EXPOSURES IN GASTRIC CARCINOGENESIS

Christina Persson

Karolinska Institutet

Stockholm 2009

Published by Karolinska Institutet.

ISBN 978-91-7409-692-7

Printed by

www.reproprint.se
Gårdsvägen 4, 169 70 Solna

ABSTRACT

Despite a general declining secular trend of incidence in most parts of the world, stomach cancer is still a major cause of cancer-related death worldwide due to its poor prognosis. Although *Helicobacter pylori* (*H. pylori*) infection is an established risk factor for stomach cancer, the true magnitude of the association is still not determined. Besides *H. pylori* infection, some other factors including host genetic variation and environmental exposures might to play a role in gastric carcinogenesis.

As acquisition of the infection almost invariably occurs before adulthood, the serostatus at age 16-40 should best reflect the lifetime occurrence of the infection. Study I is a case-control study nested within a historic cohort approximately of 400 000 individuals. These individuals donated sera before age 40 years to either of two large Swedish biobanks between 1967 and 2006 and were linked to complete nationwide registers. In total 59 incident cases of stomach adenocarcinoma (41 non-cardia tumours) were confirmed. The risk of non-cardia stomach adenocarcinoma among subjects with antibodies against both *H. pylori* cell surface antigens and CagA, relative to those who were seronegative, was 48.5 (95% CI, 5.8-407.4). This result provides strong evidence that *H. pylori* infection is a much stronger risk factor for non-cardia stomach adenocarcinoma than initial reports indicated.

The study II is composed of a population-based case-control study based in five Swedish counties and a hospital-based case-control study conducted at eight Swedish hospitals which provided a combined total of 351 gastric cancer cases and 539 controls for analysis. The *IL1B*-31, *IL1B*-511, and *IL1B*+3954 biallelic polymorphisms were genotyped. The risk of gastric cancer was unrelated to genotype in all of the studied polymorphic loci, and the absence of association was also evident when the population-based and hospital-based case-control studies were analyzed separately. These results do not support the hypothesis that human genetic polymorphisms linked to IL-1β production are associated with the risk of gastric cancer.

Study III is a population-based case-control study conducted in Warsaw, Poland, during 1994-1996. This study provided 273 stomach cancer cases and 377 controls in which 6, 8, and 14 tagSNPs in *MUC1, MUC5AC,* and *MUC6* genes, respectively, were genotyped. Each of the six tagSNPs tested across the *MUC1* region showed association with an increased risk of stomach cancer. Carriers of the haplotype ACTAA rare alleles of rs4971052, rs4276913, rs4971088, rs4971092, and rs4072037 had an approximate two-folded increase in risk (OR 1.93, 95% CI, 1.49-2.48) compared to the referent haplotype. Out of 8 tagSNPs across *MUC5AC* region, only the minor allele of rs868903 was significantly associated with an increased risk of stomach cancer (OR 1.80, 95% CI, 1.22-2.63). No significant association was found in tagSNPs across *MUC6* genes. In conclusion, some common variations in *MUC1* and *MUC5AC* genes likely contribute to an elevated risk of stomach cancer.

Study IV, was based in the JPHC Prospective Study which is composed of 36 745 subjects aged 40 to 69 years, who responded to a baseline questionnaire and provided blood during 1990-1995 and were subsequently followed until 2004. Plasma levels of carotenoids in 511 gastric cancer cases and 511 matched controls were measured. The plasma level of β-carotene was inversely associated with risk of gastric cancer (compared with lowest quartile: OR 0.63, 95% CI, 0.31-0.75; OR 0.48, 95% CI, 0.31-0.75; and OR 0.46, 95% CI, 0.28-0.75, for the 2nd, 3rd and 4th quartile respectively). Inverse associations were evident in men for alpha-carotene and beta-carotene, but not in women who had relatively higher plasma levels compared with men. No statistically significant association between lutein/zeaxanthin, lycopene, retinol, α- or γ-tocopherol with gastric cancer was found. These findings suggest that those who have low plasma levels of α- and β-carotene have an increased risk of gastric cancer.

LIST OF PUBLICATIONS

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

I Persson C, Jia Y, Pettersson H, Dillner J, Nyrén O, Ye W
***Helicobacter pylori* seropositivity before age 40 and subsequent risk of stomach cancer: A glimpse of the true relationship?**
[In preparation]

II Persson C, Engstrand L, Nyrén O, Hansson LE, Enroth H, Ekström AM, Ye W
Interleukin 1-β gene polymorphisms and risk of gastric cancer in Sweden.
Scand J Gastroenterol. 2009;44(3):339-45.

III Jia Y, Persson C, Hou L, Zheng Z, Yeager M, Lissowska J, Chanock SJ, Chow WH, Ye W
A comprehensive analysis of common genetic variation in *MUC1, MUC5AC, MUC6* genes and risk of stomach cancer
Cancer Causes and Control [In Press]

IV Persson C, Sasazuki S, Inoue M, Kurahashi N, Iwasaki M, Miura T, Ye W, Tsugane S; JPHC Study Group
Plasma levels of carotenoids, retinol and tocopherol and the risk of gastric cancer in Japan: a nested case-control study
Carcinogenesis. 2008, 29(5):1042-8.

CONTENTS

LIST OF ABBREVIATIONS

BMI	body mass index
bp	base pair
CagA	cytotoxin-associated gene A
CI	confidence interval
DNA	deoxyribonucleic acid
ELISA	enzyme linked immunosorbent assay
GC	gastric cancer
H. pylori	*Helicobacter pylori*
Hp-CSA	*Helicobacter pylori* cell surface antigens
IARC	the International Agency for Research on Cancer
HWE	Hardy-Weinberg equilibrium
ICD	International Classification of Diseases
IHC	immunohistochemistry
JPHC	Japan Public Health Center
kb	kilo base pair
LD	linkage disequilibrium
OR	odds ratio
PAI	pathogenicity island
PCR	polymerase chain reaction
PG	pepsinogen
PPI	proton pump inhibitor
RT-PCR	real time polymerase chain reaction
SNP	single nucleotide polymorphism
WHO	World Health Organization
VNTR	variable number of tandem repeats

1 INTRODUCTION

Gastric cancer is the fourth most common malignancy and the second most common cause of cancer death worldwide. Its incidence rates are high in Eastern Europe, Central and South America and Eastern Asia. Despite improvement in treatment modalities and the introduction of screening programs, the 5-year survival rate of this malignancy is rarely reported more than 30%. *Helicobacter pylori* is a well-established risk factor for gastric cancer, but the strength of the association is still under debate. Besides *H. pylori* infection, host genetic susceptibility and other environmental exposures like diet might also be associated with risk of gastric cancer.

This thesis aimed to study the roles of *H. pylori* infection, host genetic susceptibility, and other environmental exposures in the development of gastric cancer, by using several unique epidemiological study resources in Sweden, Japan, and Poland.

2 PHYSIOLOGY OF THE STOMACH

2.1 ANATOMY OF THE STOMACH

The stomach is the sack between the esophagus and the small intestine. The area surrounding the junction between the stomach and the esophagus is known as the *cardia* and is located in the most proximal section of the stomach. The upper rounded part of the stomach is the *fundus* followed by the part between the lesser and the greater curvature - the *corpus* (body). The *antrum*, in the bottom part of the body connects to *pylorus* which ends with duodenum, Figure 2-1.

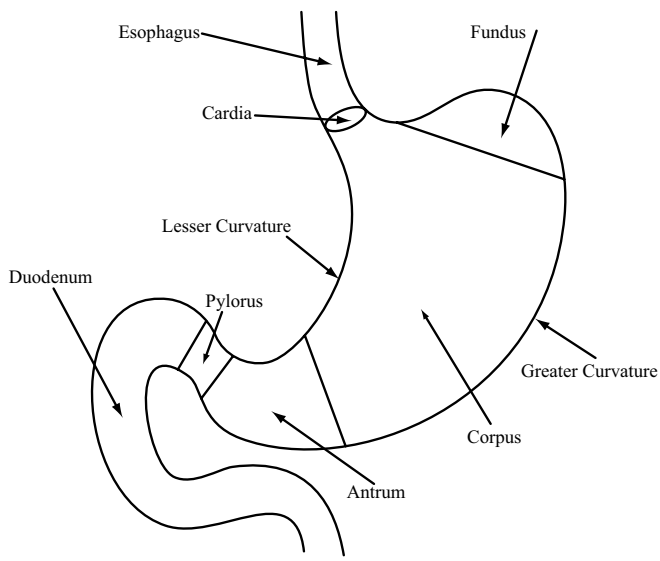


Figure 2-1. A schematic illustration of the stomach anatomy

The stomach epithelium consists of gastric pits each composed of 4-5 gastric glands. The structure of these glands varies with different parts of the stomach. The cardia contains less than 5% of the total number of gastric glands and includes mucous secreting cells and few parietal cells. The corpus and the fundus include acid secreting parietal cells and chief cells[1], which store and secret pepsinogen type A and C. In the pyloric area chief cells secrete pepsiongen C into the gastric lumen.

2.2 GASTRIC ACID SECRETION

Parietal cells in the oxyntic mucosa secrete hydrochloric acid into the lumen of the stomach. The acid secretion is stimulated by acetylcholine from the postganglionic enteric neurons, gastrin from antral G-cells, histamine from enterochromaffin-like cells, and together with sensory induction by thought, sight, smell, taste, and swallowing of food. The inhibitor of the acid secretion is somatostatin secreted from D-cells in the corpus and pyloric mucosa, Figure 2-2.

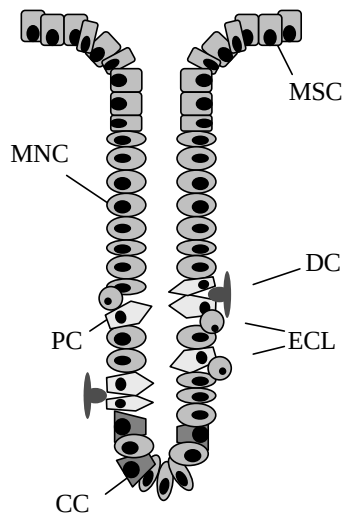


Figure 2-2. **A schematic illustration of the gastric oxyntic gland layout**
Schematically showing mucous surface cells (MSC), mucous neck cells (MNC), parietal cells (PC), D-cells (DC), enterochromaffin-like cells (ECL), and chief cells (CC)

Gastric acid facilitates digestion of protein and absorption of nutrients such as iron, calcium, and vitamin B_{12}. Gastric acid also further prevents bacterial overgrowth and enteric infection. Pepsin is activated by gastric acid and it metabolizes proteins into peptides. Pepsinogens (PG) are proenzymes of pepsin synthesized in chief cells and mucous neck cells. Pepsinogen in adults is detectable in two immunologic types: pepsinogen A (PGI) and pepsinogen C (PGII). Whereas PGI is only synthesized in the oxyntic mucosa of the corpus, PGII can be found in most parts of the stomach and part of the duodenum. When corpus atrophy develops, a direct correlation between loss of chief cells and serum PGI can be observed.

3 H. PYLORI

In the year 2005 the Nobel Prize in Physiology or Medicine was awarded to the Australian physicians Barry J. Marshall and J. Robin Warren, "for their discovery of the bacterium *Helicobacter pylori* (*H. pylori*) and its role in gastritis and peptic ulcer disease"[2].

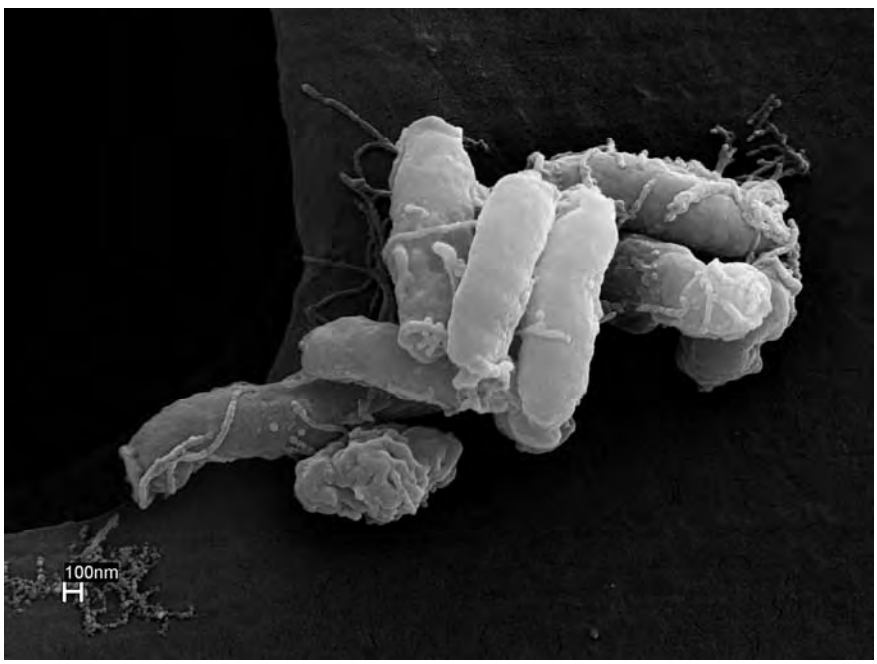


Figure 3-1. *H. pylori* with flagella imaged by an electron microscope on agar plate

This gram-negative, spiral shaped, micro-aerophilic bacterium (Figure 3-1) is highly adapted for colonization in the human stomach. The stomach environment is lethal to most other microbes due to factors such as the high acidity, limited availability of nutrients, and frequent gastric emptying. To overcome the stomach environmental barriers *H. pylori* have developed a number of mechanisms including the enzyme urease. Urease converts urea into ammonia and carbon dioxide, resulting in increased pH in the local environment around the bacteria, which is favorable for the survival and attachment of the bacteria. When *H. pylori* have established a chronic infection in the stomach, it can elicit a chronic inflammatory response. The *H. pylori* infection's effect on the gastric acid production depends on severity, anatomic location, and extension of the gastritis. The inflammation of the antrum leads to inhibition of somatostatin and stimulation of gastrin and acid secretion, while in some

individuals the infection further spreads to the corpus, followed by development of atrophy which destroys the parietal cells leading to hypochlorhydria that is further mediated by the elevated levels of inflammatory cytokines, such as IL1-β and TNF-α[3].

3.1 DETECTION OF *H. PYLORI* INFECTION

The gold standard for *H. pylori* detection is histopathologic analysis and/or culture of biopsy tissue obtained during upper gastrointestinal endoscopy. The gastric tissue biopsy can further be used for enzyme activity (urease test) detection indicating presence of *H. pylori*.

Endoscopy-based diagnosis is limited due to its expense in large studies, the time required for bacterial culture, and the need of endoscopy facilities. Further, sometimes in carcinogenesis process it can be difficult to accurately measure whether a patient was ever infected with *H. pylori* using this method of detection, as atrophy may cause spontaneous disappearance of *H. pylori*.

Therefore serological tests of measuring antibodies in the plasma or serum can be an alternative method for detection of *H. pylori* and may also be used to measure previous infection. The performance of commercial serology tests varies mainly due to strain heterogeneity in different populations. Therefore, validation of a serological test is necessary before it can be used diagnostically in a population. In addition to strain variation, other factors can cause misclassification and impair the performance of serological tests. For example, the test cut-off values might differ with the age of the patients, and also sample storage duration can affect the titers.

Serological tests usually include conventional immunoglobulin (IgG) enzyme-linked immunosorbent assay (ELISA) and immunoblotting assay. Some *H. pylori* strains contain the cytotoxin associated gene pathogenicity island (cagPAI) and the CagA effector protein[4], which is known to cause more extensive inflammation in the stomach mucosa. Compared to antibodies against *H. pylori* cell surface antigens (Hp-CSA) the CagA antibodies persist much longer after eradication and are therefore a better indicator of a past infection[5, 6].

3.2 EPIDEMIOLOGY OF *H. PYLORI* INFECTION

H. pylori chronically infects more than half the world's populations[7]. The infection is generally introduced during childhood through oral-fecal contacts. In developing countries the prevalence of *H. pylori* infection peaks at 80-100% during adolescence and then persists in

these individuals throughout life. In contrast, the infection is acquired later in developed countries and cleared in about 10% of the cases; its prevalence also peaks at 50-70% during young adulthood and this peak prevalence is declining[7]. The adult prevalence of infection is estimated to be 82% in Eastern Europe, 71% in Japan, 62% in China, 60-70% in Jamaica, 62% in Central America, 62% in Mexican Americans, 52% in non-Hispanic blacks in the U.S., and 26% among non-Hispanic whites in the U.S.[8]. Suggested risk factors for *H. pylori* infection include low socioeconomic status and presence of infected family members[9, 8].

4 STOMACH CANCER

4.1 DESCRIPTIVE EPIDEMIOLOGY OF GASTRIC CANCER

Although its incidence has declined in many countries, estimates in 2002 by the IARC indicate that gastric cancer is still the fourth most common type of cancer in the world following lung, breast, and colorectal cancers. In the year 2002, 934 000 new cases were estimated annually which represented 8.6% of all new cancer cases. Stomach cancer is the second most common cause of cancer related death with about 700 000 deaths annually. The overall age-standardized incidence rates were 22.0 per 100 000 person-years in men and 10.3 per 100 000 person-years in women with corresponding mortality rates of 14.3 and 8.3 per 100 000 person-years, in year 2002[8]. The 5-year survival rates varied between countries, ranging from 11 to 30%. In Japan the corresponding survival rate for stomach cancer is as good as 52%, mostly due to mass screening programs which were introduced in the 1960's[9].

Gastric cancer incidence rates vary geographically, and close to two-thirds of the cancers occur in developing countries. Eastern Europe, Central and South America, and Eastern Asia are the areas with highest incidence rates. The incidence of stomach cancer in Africa has been suggested to be low, but age standardized rates for women in central Africa has been found to be similar to those in Eastern Europe[8], Figure 4-1.

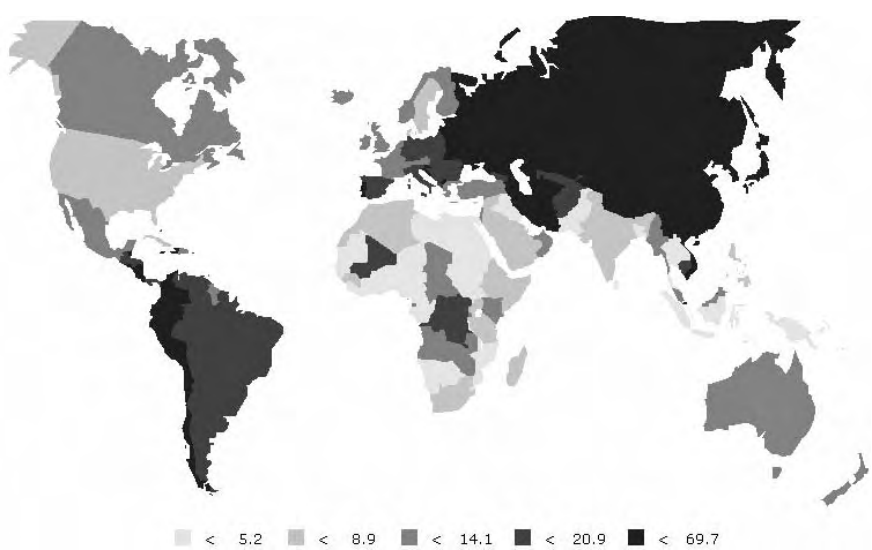


Figure 4-1. The global burden of stomach cancer among men in year 2002
Age-standardized incidence rate per 100 000 person-years, GLOBOCAN 2002 IARC

In the U.S. gastric cancer incidence in blacks has been reported to be double than in whites. In Singapore gastric cancer incidence rates among Malay and Chinese men are approximately 9 and 29 per 100 000 person-years[10], respectively.

Gastric cancer is two to three-fold more common among men than in women regardless of geographic area[9]. Most stomach cancer cases occur between 60 and 80 years of age. The female incidence rates are equivalent to male rates at a 10-year lag of age[9]. This 10-year time shift in incidence between men and women is not fully understood but has been suggested to be due to sex-specific hormonal protective effects[11, 12].

The incidence and mortality has declined in most countries over the last decades. The age-adjusted incidence rates for stomach cancer has declined approximately 15% compared to 1985[8]. A similar pattern in mortality rates has been observed, though in some populations mortality declined faster as a result of early detection, Figure 4-2.

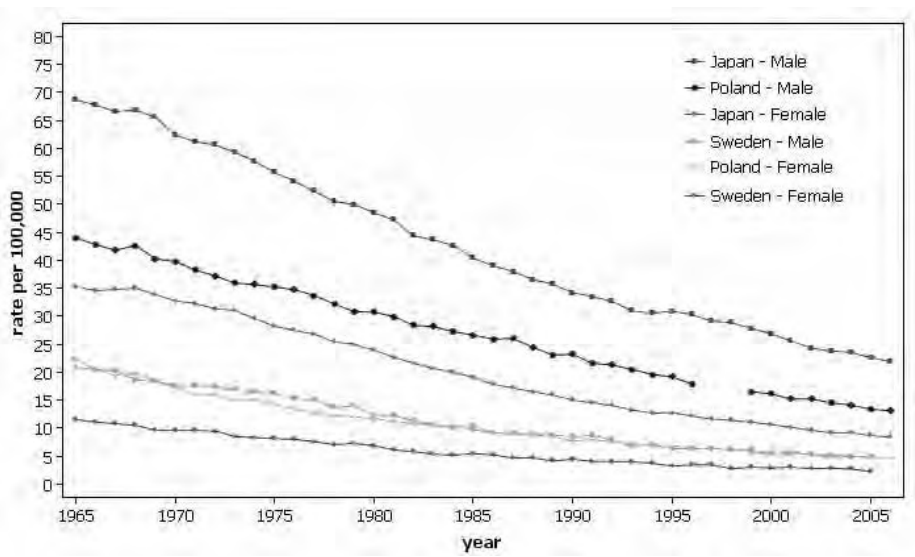


Figure 4-2. Age-standardized mortality rates for Japan, Poland and Sweden in 1965-2005
WHO-IARC, information missing 1997-98 from Poland

Notwithstanding the fact that stomach cancer is declining, assuming a growing world population and increased longevity, an overall 1.1 million gastric cancer cases can be expected in year 2010[8, 9].

4.2 SUBTYPE BY ANATOMIC SITE AND HISTOLOGY

Stomach cancer can be subdivided according to anatomic site into cardia vs. non-cardia (or distal) cancer. Adenocarcinoma is the most common histologic type of gastric caner; other uncommon histologic types include lymphoma (mucosa-associated lymphoid tissue), malignant gastric stromal tumor, and carcinoid tumors. This thesis focuses on adenocarcinoma, if not otherwise specified. Furthermore the gastric adenocarcinoma can be subdivided into two main categories, i.e. diffuse and intestinal type, based on Laurén's classification[13].

4.3 RISK FACTORS COVERED IN THIS THESIS

4.3.1 *H. pylori* infection

Most subjects with *H. pylori* infection are asymptomatic. However, some infected subjects develop chronic inflammation in the antrum with mild inflammation in the oxyntic mucosa (the acid-forming part of the gastric glands: the parietal cells) and a subset of these patients develop duodenal ulcer. In other infected subjects, development of corpus predominant gastritis results in the development of atrophic gastritis and intestinal metaplasia, and is associated with risk of gastric ulcer and gastric cancer[14]. Microbial presence in or near epithelia provides a stimulus for recruitment and activation of inflammatory cells from the blood stream. Cytokines, chemokines, and free radicals initiate and perpetuate inflammatory responses[14, 15]. The activation of inflammatory cells leads to a respiratory burst that releases free radicals, contributing to malignant transformation by peroxidating lipids and inducing genetic mutations. This type of damage to epithelial cells stimulates cell death, epithelial hyperproliferation, and further mutation, Figure 4-3.

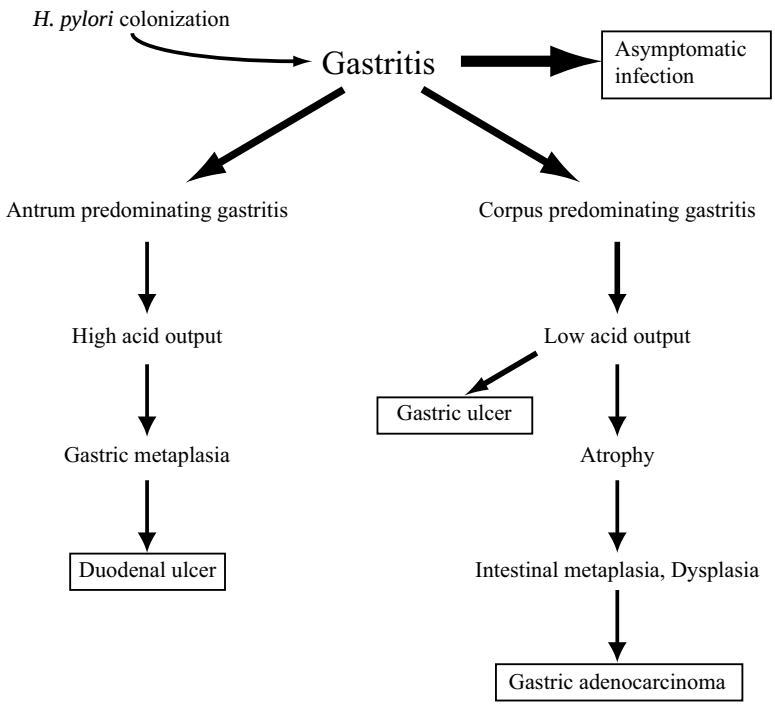


Figure 4-3. *H. pylori* colonization and subsequent outcomes by Correa[16]

Accumulation of mutations in key host cell regulatory genes eventually leads to changes in cellular phenotype, at which stage eliminating the initiators of inflammation (e.g. smoking or *H. pylori*), or abolition of the inflammation itself cannot prevent the progression toward cancer. Progression through these histological cancer precursors is accompanied by the accumulation of mutations, typical of a multi-step process of carcinogenesis. The modulation of gastric growth factors and peptides together with low acid secretion encourages the growth of nitrosating bacteria, other than *H. pylori*, which further compound these effects.

4.3.2 Host factors

4.3.2.1 Mucous barrier

The mucous layer in the stomach acts as a protective barrier; it protects the gastric glands from the acid environment. Heavily glycosylated glycoproteins such as mucins constitute the major components of the mucous protective layer[17], Figure 4-4.

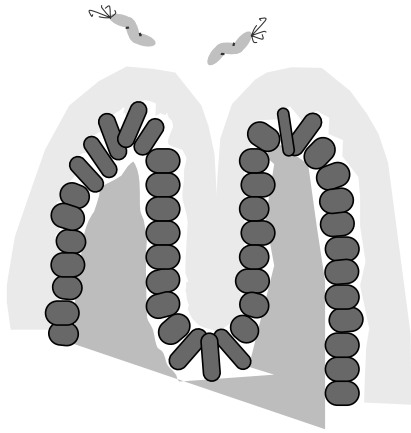


Figure 4-4. Sketch of the gastric pits covered with mucous layer

Mucins are differentially expressed in epithelia with cell type specificity[18-20]. Within the human body, *H. pylori* reside primarily in the gastric mucous layer[21]. The normal gastric mucosa express MUC1, MUC5AC, and MUC6, where MUC1 and MUC5AC are found in the superficial epithelia and the MUC6 in the deep glands[18, 22-26]. MUC5AC forms the major receptor for *H. pylori* in the human stomach[27, 28] and the infection further alters the expression of *MUC1, MUC5AC,* and *MUC6* genes. High levels of MUC1 mRNA have been detected in gastric carcinomas, whereas decreased levels have been reported for MUC5AC and MUC6[29]. The smaller size variable number tandem repeat (VNTR) polymorphism of *MUC1* is suggested to be associated with *H. pylori* infection[30] and in the Portuguese population, the smaller size *MUC1* VNTR alleles have been shown to be associated with an increased risk for gastric carcinoma[31] as well as chronic atrophic gastritis and intestinal metaplasia[32]. Similarly, the smaller size VNTR alleles of *MUC6* are associated with *H. pylori* infection[33] and an increased risk of stomach cancer[34].

4.3.2.2 The proinflammatory response

An association between single nucleotide polymorphisms (SNPs) in genes that regulate the host's inflammatory response and gastric cancer has been reported, although some studies have failed to replicate these findings. The inflammatory response related genes that have been most frequently investigated are the interleukins (IL) *IL1B, IL1RN, IL8, IL10* and tumor necrosis factor alpha (*TNFA*), coding for the proteins IL-1β, IL-1ra, IL-8, IL-10 and TNF-α respectively, Table 4-1.

Table 4-1. Recently reported proinflammatory gene variants related with gastric cancer risk

Gene loci	rsID	Chromosome position	Wild type	Homozygote*
IL1B-511	rs16944	2q	*C/C*	*T/T*
IL1B-31	rs1143627	2q	*T/T*	*C/C*
IL1B+3954	rs143634	2q	*C/C*	*T/T*
IL1RN §	VNTR	2q	*L/L*§	*2/2*
IL8-251	rs4073	4q	*T/T*	*A/A*
IL10-1082	rs1800896	1q	*A/A*	*G/G*
IL10-819	rs3021097	1q	*C/C*	*T/T*
IL10-592	rs1800872	1q	*C/C*	*A/A*
TNFA-308	rs1800629	6p	*G/G*	*A/A*
TNFA-238	rs361525	6p	*G/G*	*A/A*

*homozygote versus wild type, §L represents any long allele (allele 1, 3, 4, or 5)

These cytokines have been reported to be important mediators in the inflammatory process, being involving in, amongst other mechanisms stomach acidity, stimulation of proliferation of endothelial cells and down-regulation of the cytotoxic response. The proinflammatory cytokine IL-1β, is a potent suppressor of gastric acid secretion in cell cultures, animal models, and in humans[21, 35, 36]. The link to gastric cancer has been suggested to be through prolonged acid suppression leading to dispersal of *H. pylori* from the antrum to the gastric corpus. The *IL1* gene cluster is located on chromosome 2q and contains three related genes; *IL1A*, *IL1B*, and *IL1RN*, which correspondingly code for IL-1α, IL-1β and the IL-1 receptor antagonist (ra). IL-1ra binds to IL-1 receptors and inhibits binding of IL-1α and IL-1β [37]. Polymorphisms of the *IL1B* and *IL1RN* genes have been suggested to alter the production and mucosal levels of IL-1β and could thus potentially modulate the inflammatory response and play a role in the postulated sequence from gastric hypochlorhydria to gastric cancer[38-41], Figure 4-5.

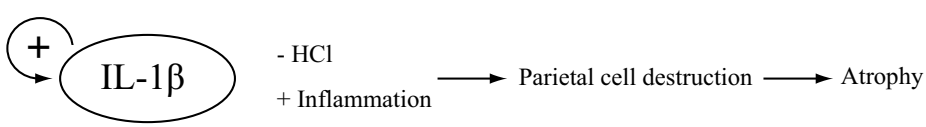


Figure 4-5. The potential effects of elevated IL-1β levels

The IL1B related genes, except for *IL1B*+3954 (exon 5) and *IL1RN* receptor antagonist (intron 2), have SNPs that are situated in the gene promoter region, playing an important role in the expression of the gene and thereby affecting the inflammatory response. There has been

inconsistency in the reported association between related SNPs and stomach cancer risk, which is potentially the result of variation in population allele frequencies, Table 4-2.

Table 4-2. Allele frequencies among controls in Asian and non-Asian populations[42]

Gene SNP allele	Asian (min-max %)	non-Asian (min-max %)
IL1B-511T	47-53	30-36
IL1B-31C	43-55	30-56
IL1B+3954T	2-7	8-25
*IL1RN*2	2-12	17-38

The originally reported SNPs[38] are three single nucleotide polymorphisms in the *IL1B* gene: a T-C base transition at *IL1B*-31 and C-T base transitions at *IL1B*-511 and *IL1B*-3954. *IL1B*-31 is located in the TATA-box of the promoter region[43, 44], and most likely affects the expression of the gene and has been found to be in almost total linkage disequilibrium with *IL1B*-511[45], Figure 4-6.

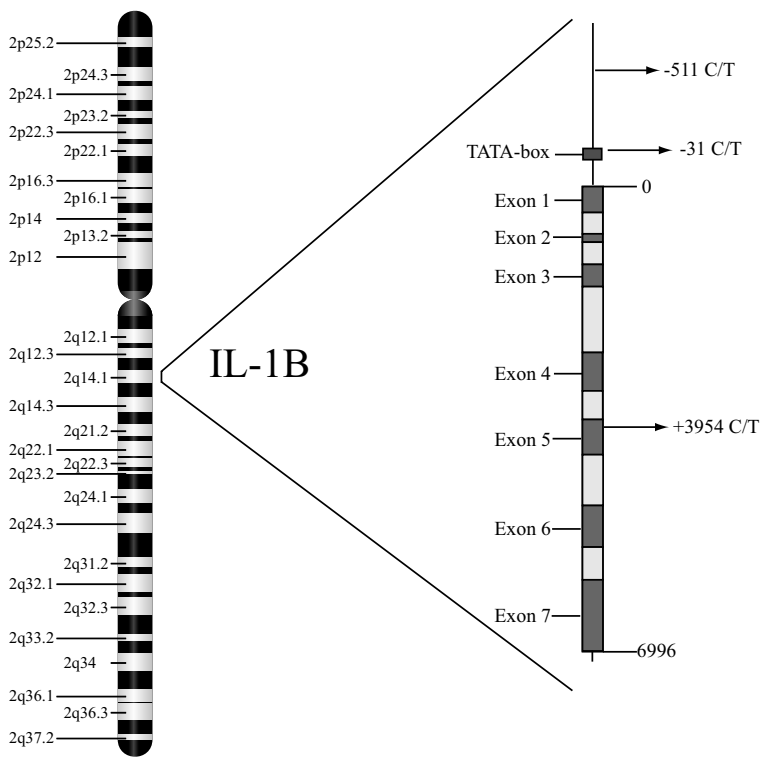


Figure 4-6. Three frequently reported SNPs in the *IL1B* gene localized in chromosome 2q14 region

The receptor antagonist *IL1RN* gene, contains a variable number of tandem repeats (VNTR) of 86 base pairs that vary from 2 to 6 copies[46].

After an initial study demonstrated an impressive elevated risk of gastric cancer among individuals with the proinflammatory and hypoacidity-related *IL1B*-31 C/T or T/T and *IL1RN* 2/2 genotypes, relative to those with the *IL1B*-31 C/C and *IL1RN* L/L genotypes, respectively[38], subsequent reports analyses have reported conflicting results. In some studies the presence of the *IL1B*-31 C, *IL1B*-511 T, *IL1B*-3954 T and/or *IL1RN* 2 was associated with an increased risk of gastric cancer[38, 47-50], but not in others[51-54].

In a meta-analysis[55], no association between the *IL1B*-511, *IL1B*-31, *IL1RN* polymorphisms and risk of developing gastric cancer, independent of the histology type, was found. On the other hand another meta-analysis[42] suggested that *IL1B*-511 T and *IL1RN* 2 carriers had an increased risk of developing intestinal subtype of gastric cancer.

4.3.3 Environmental factors

Besides *H. pylori* infection and host genetic polymorphisms there are other factors that might be involved in the development of stomach adenocarcinoma such as tobacco smoking, diet, sanitary conditions, and occupational exposures[56-58].

4.3.3.1 Dietary factors

The ingested food stays for at least 4 hours in the stomach; it is therefore likely that digested compounds would in some way affect the system and risks of disease development. Diet has been debated to play an important role, especially fruits and vegetables which have been suggested to be protective, owing to the effect of antioxidants[59-62]. A Japanese population-based prospective study reported fruit and vegetable intake decreased the risk of gastric cancer[63]. Fruits and vegetables contain antioxidants amongst which typically are carotenoids, retinol, and tocopherol[60, 62, 64-69]. Due to the long half-life[70] of carotenoids, they could be a stable indicator of antioxidants.

The worldwide decrease in age-adjusted incidence of gastric cancer has not only been related to the international program of *H. pylori* infection eradication but also related to a reduction of salt and salted foods intake, replaced by fresh fruit and vegetables and with the popularization of refrigerators[71]. The so-called healthy diet including fruit and vegetables compared to

traditional Japanese food or Western food, is associated with a reduced risk of gastric cancer[59]. Traditional Japanese food such as salted fish roe and salted fish preservatives, is high in salt, and positively associated with gastric cancer risk in a dose-dependent manner[72]. On the other hand, in a western population such as Norway the high intake of salt did not appear to increase the risk of gastric cancer[73].

4.3.3.2 Other factors

An increased risk of gastric cancer among first degree relatives of stomach cancer patients has been reported[74]. Further, the structure of the family has been reported to affect the risk of stomach cancer development, with higher birth orders compared to lower shown to be at increased risk[75].

Some working environments where individuals were exposed to herbicides with phenoxyacetic acids[76], or airborne exposures such as diesel exhaust, cement and quarts dust[77] have been reported to increase the risk of gastric cancer development.

Tobacco smoking has been reported to have a causal role in gastric cancer development. The risk of stomach cancer increased among individuals with higher and longer duration of cigarette smoking compared to non-smokers, and the estimates are also higher in men than women[8, 9].

5 AIMS

The aims of this thesis were to further investigate the roles of *H. pylori* infection, host genetic variation, and other environmental exposures in gastric carcinogenesis. Specifically:

- To estimate the true magnitude of the association between *H. pylori* infection and stomach cancer risk, based on a cohort of subjects whose blood samples were collected before age 40 years.

- To better understand the roles of *IL1B*-31, *IL1B*-511, *IL1B*+3954 and *IL1RN* polymorphisms in the development of stomach cancer, based on a population-based and a hospital-based case-control study in Sweden.

- To investigate the roles of variations in the *MUC1*, *MUC5AC*, and *MUC6* genes in the development of stomach cancer, in a population-based case-control study in Warsaw, Poland.

- To investigate the impact of plasma levels of carotenoids (β-cryptoxanthin, α- and β-carotene, lutein/zeaxanthin and lycopene), retinol and α- and γ-tocopherol, on the risk of gastric cancer using a case-control study nested within the JPHC-based Prospective Study.

6 SUBJECTS AND METHODS

6.1 SETTINGS

6.1.1 Swedish registries

6.1.1.1 Personal Identity Numbers

Since 1947, all those registered in Sweden, including immigrants, are given a personal identity number for identification. The number consists of ten digits since 1967, when an extra digit was added for checksum control. The personal identity number gives the date of birth of the person, a birth number, and a check digit. If, at a certain day of birth, all possible personal identity numbers are assigned, the day will be changed to the next following day in the month. When a person reaches the age of 100 years old, the hyphen between the day of birth and the birth number is changed to a plus sign. The only information that can be read from a personal identity number is date of birth and sex, Figure 6-1.

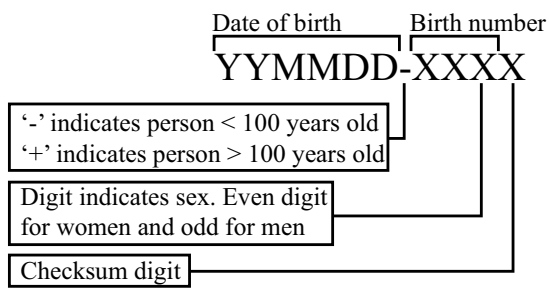


Figure 6-1. The built-up of personal identity numbers in Sweden

The personal identity number is kept the whole life, including when individuals migrate. The personal identity number is widely used in the Swedish society by authorities, healthcare and other administrations. Consequently, the number can be used as a link between registers to gain and merge information about a single individual[78].

6.1.1.2 The Total Population Registry and Cause of Death Registry

The Total Population Registry is administered by the Tax Agency in Sweden, and includes people born in Sweden or moving to Sweden. Frequent reports from the register are supplied by Statistics of Sweden, reporting information on sex, age, civil status, citizenship, country of birth, migration, birth and death[79].

The Swedish Cause of Death Registry is based on death certificates, with nationwide coverage since 1911. A death certificate must be issued before burial is permitted and is filled-in by the physician or surgeon in charge of the patient during the last hospitalization or, for those who die outside hospitals, by a family physician or specialist in forensic medicine. The Cause of Death Registry supplies data and statistics for research and provides information on causes of death to the general public. From the beginning of 1749, the reporting was the responsibility of the clergy, but since 1860 it became the responsibility of medical doctors to complete and submit death certificates in cities with medical officers. The death certificates are based on the International Statistical Classification of Diseases (ICD) designed by WHO. The registry reached 99.5% compliance in 1997 and since 1994 the National Board of Health and Welfare has held the responsibility for producing an annual report[80].

6.1.1.3 The Swedish Cancer Registry

The Swedish Cancer Registry supplies data for statistics (e.g. cancer incidence rate) and research. The register started in 1958 and is linked to the Cause of Death Registry. The Cancer Registry does not include cancers identified solely by a death certificate. All newly diagnosed cancers by law must be reported to the Cancer Registry by clinicians, pathologists and cytologists. Each Regional Oncology Center processes forms that are subsequently reported into the Cancer Registry in the National Board of Health and Welfare[81]. To assure the validity of the personal identification number, these numbers are compared and checked with the Total Population Register.

6.1.2 Polish registries

6.1.2.1 Polish Electric System of Residence Evidency

In Poland all permanent and temporary residents, are assigned a Polish Electric System of Residence Evidency (PESEL) number. The PESEL number has been used since 1979 and is built up in a similar way as the Swedish personal identity numbers, including date of birth and a personal identification number. The regional residence registries are updated monthly with close to 100% completeness.

6.1.2.2 Cancer Registry

In Poland a law in 1951was established to initiates cancer registries, and one year later eight local registries in the city areas of: Olsztyn, Gdańsk, Bydgoszcz, Poznań, Łódź, Warsaw, Lublin, and Białystok, were started. Data are supplied by eleven oncology units covering the

whole country, to the Institute of Oncology in Warsaw which is responsible for nationwide reporting.

6.1.3 Japanese registries

6.1.3.1 Residency and death registry

In Japan, residency and death registration are required by law. Residence status and survival is registered by residential registers and kept by each municipality. The resident register is open to the general public under the resident registration law. There are no personal identity numbers in Japan, resulting in linkage of information based on name and address of the individual. Further, information on the causes of death is provided through examining death certificates provided by the Ministry of Health, Labor, and Welfare with permission from the Ministry of Internal Affairs and Communications.

6.1.3.2 Cancer registry

There is no nationwide population-based register in Japan but regional registries based on prefecture have been active since the 1960's. Today active cancer registries are compose of: (i) population-based cancer registries - covering diagnosed cases in a well-defined population, (ii) hospital (institution) - based cancer registries - registering all cancer patients receiving cancer care at a hospital or institution, (iii) site specific cancer registers - conducted by clinicians, academic societies or research groups specializing in various cancer sites that cover major hospitals nationwide and collect detailed clinical information in[82].

Table 6-1. Overview of the studies described in this thesis

Characteristics	Study I	Study II[83]	Study III	Study IV[84]
Design	Nested case-control	Case-control	Case-control	Nested case-control
Source-population	Subjects with blood samples collected in two biobanks	a) 5 Swedish counties b) 8 hospitals	Warsaw population	JPHC cohort
Participants	Swedish, 16-40 years at baseline	Swedish, 40-79 years at diagnose	Polish, 21-79 years at diagnose	Japanese, 40-69 years at baseline
Study period	1967-2006	a) 1989-1995 b) 1997-1997	1994-1996	1990-2004
Outcomes	GC by cancer register and medical record review	Histologically confirmed GC	Histologically confirmed GC	GC by cancer/hospital register
Exposures	*H. pylori* serology	SNPs in *IL1B*-511, -31, +3954 VNTR in *IL1RN*	tagSNPs in *MUC1, MUC5AC, MUC6*	Plasma levels of carotenoids, retinol, α-, γ-tocopherol
Matching variables	Age, sex, year of serum collection, biobank	a) Frequency: residence, age, sex b) Individually: hospital, age, sex	Frequency: age, sex	Age, sex, study area, blood donation date, fasting time before sample collection
Adjustments	-	a) Residence area, age, sex b) Hospital, age, sex	Age, sex	Family history of GC, *H. pylori*, smoking, BMI, salt intake, highly salted foods consumption
Statistical analysis	Conditional logistic regression	Unconditional logistic regression	Unconditional logistic regression	Conditional logistic regression

6.2 STUDY POPULATION

6.2.1 Swedish biobanks

In Study I, the study cohort consists of subjects enrolled in two Swedish biobanks, each including more than 900 000 individuals. The two biobanks are the Swedish Institute for Infectious Disease Control (SIIDC) Biobank and the Malmö Microbiology (MM) Biobank. From both biobanks, a subset of individuals whose blood samples were collected between 16 and 40 years of age was selected.

6.2.1.1 Swedish Institute for Infectious Disease Control Biobank

The Swedish Institute for Infectious Disease Control (SIIDC) has performed a series of population-based, nationwide investigations on the immunity against infections in the Swedish population. SIIDC Biobank has served for decades as a reference laboratory for many microbiological analyses. As part of the quality control and documentation system, samples have been stored at -20ºC and have been partly collected since 1957 and completely since 1977.

6.2.1.2 The Malmö Microbiology Biobank

Since 1969, the Malmö Microbiology (MM) Biobank has stored samples submitted to the Department of Clinical Microbiology at MAS University Hospital in Malmö, and all samples are stored at -20ºC. The majority of the samples are sera that have been submitted for diagnosis of blood-borne viral infections (such as hepatitis virus) (over 500 000 samples). Part (100 000 samples) of the cohort are samples stored after the population-based serology screening for viral infections and rubella immunity during pregnancy. These samples are scheduled to be taken during week 14 of pregnancy and virtually all pregnant women in the population participate in this screening program. The MM Biobank serves the entire Skåne region, but there are also a large number of samples submitted from the adjacent counties in southern Sweden, notably Blekinge.

6.2.2 Two Swedish case-control studies on gastric cancer

The Study II was based on materials collected in a population-based case-control study[76] and a hospital-based case-control study[85].

6.2.2.1 Population-based case-control study

The study base for this population-based case-control study included all individuals aged 40-79 years born in Sweden, and living in one of the five counties in northern and central Sweden

from February 1989 through January 1995. Clinicians at all hospitals were contacted for continuous monitoring of whole incoming material at every pathology departments in the study areas. Regular checks with the regional and national cancer registration assured rapid and complete case ascertainment[86]. Controls were randomly selected from the population register covering the entire source population, frequency matched for age and sex distribution among the cases. Face-to-face interviews were carried out with 567 out of 908 eligible cases, and 1 165 out of 1 534 randomly selected controls. Collection of blood was added to the protocol during the latter half of the study; samples were obtained from 298 cases and 244 controls. Cases and controls who donated blood were similar to the interviewed subjects in age, sex, place of residence at inclusion, socio-economic status during childhood, number of siblings and education, as well as for cancer type and site among case patients[87]. DNA was extracted from 286 cases and 242 controls.

6.2.2.2 Hospital-based case-control study

The hospital-based case-control study[85] included 72 patients diagnosed with gastric adenocarcinoma upon gastroscopy at eight Swedish hospitals between September 1995 and August 1997. The next five consecutive patients without gastric adenocarcinoma undergoing endoscopy in the same hospital were selected as controls (n=324) and individually matched to the cases for age (10-year age groups) and sex. DNA samples were extracted from 65 cases and 297 controls.

6.2.3 A Polish population-based case-controls study on gastric cancer

In Study III the cases included were patients diagnosed with stomach cancer and aged 21-79 years between March 1994 and April 1996 in Warsaw, Poland. The cases were identified by collaborating physicians in the 22 hospitals serving the area. In total, 72 clinics and endoscopy departments within these hospitals and eight private endoscopy units were covered. To ensure completeness of case ascertainment Swedish Cancer Registry files were reviewed. Controls were randomly selected from a register of all Warsaw residents with PESEL, frequency matched to cases by sex and age in 5-year groups[34]. Blood samples were obtained from 305 (66%) and 427 (89%) of the participating cases and controls, respectively. DNA samples were available from 273 cases and 377 controls.

6.2.4 The Japan Public Health Center-based Prospective Study

In Study IV, the study population included subjects recruited into the Japan Public Health Center-based Prospective (JPHC) Study. The JPHC was launched in January 1990 for Cohort I

(subject age range 40 to 59 years) and in 1993 for Cohort II (40 to 69 years). The JPHC consisted of 11 public health center areas throughout Japan with a total of 140 420 middle-aged individuals. One public health area was not included in the present study due to lack of cancer incidence data. At the baseline of the study, a self-administered questionnaire survey on various lifestyle factors, including demographic characteristics, personal medical history, family history, smoking and drinking habits, dietary habits, physical activity and other lifestyle factors was conducted in each cohort, Figure 6-2.

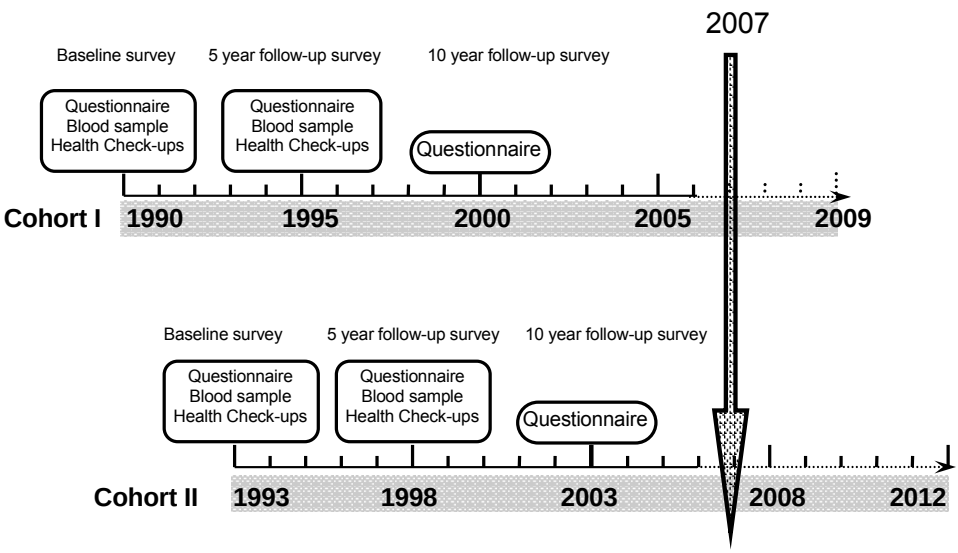


Figure 6-2. Flowchart of the JPHC study

A total of 99 808 participants responded, giving an 81% response rate. Blood was provided during health checkup. After exclusion of subjects with self-reported cancer at baseline, non-Japanese race, and those who had moved away before baseline survey, 46 803 men and 50 841 women remained eligible. Among these, 13 467 (29%) men and 23 278 (46%) women donated blood samples at baseline, collected 1990-1992 in Cohort I and 1993-1995 in Cohort II. The same protocol for specimen collection was used in all centers[88].

6.3 FOLLOW-UP AND CASE ASCERTAINMENT

6.3.1 Study I

Personal identity numbers in the two biobanks, the SIIDC and the MM Biobank, provided the possibility to link the databases to the nationwide registers: Migration Register, Cancer Register, Total Population Register, and Cause of Death Register. If a personal identity

number could not be found in Total Population Register, Migration Register or Cause of Death Register, the personal identity number was considered not valid. The gastric cancer cases were identified through the Cancer Register and their medical records were collected from corresponding treatment units. The medical records were reviewed by a professional, blinded to earlier reported ICD code in the Swedish Cancer Register as well as *H. pylori* status, Figure 6-3.

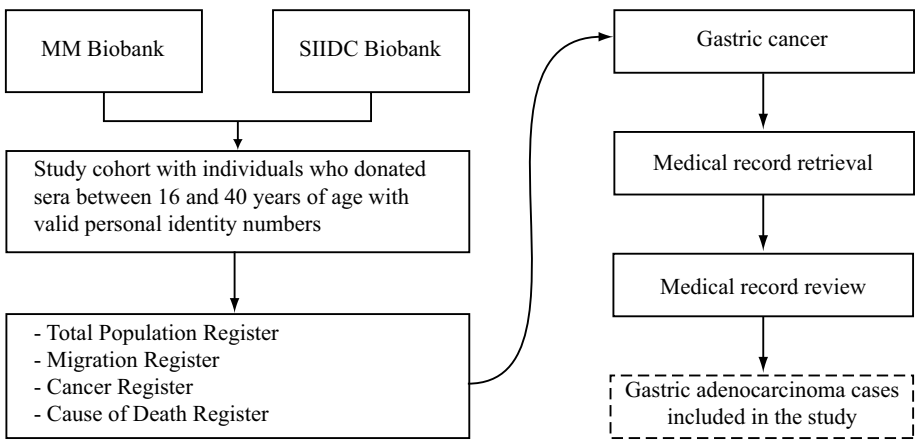


Figure 6-3. Flowchart of the study procedure for selection of stomach cancer cases in Study I

6.3.2 Study II

In Study II all gastric cancer cases were confirmed histologically and classified according to Laurén's classification[13] as intestinal or diffuse type, and by anatomic site as cardia and non-cardia stomach adenocarcinoma. In 27 of the population-based cases, classification of histology was not specified.

6.3.3 Study III

In Study III, all cases were pathologically confirmed as gastric cancer cases after identification through collaborating physicians in 22 hospitals and 8 private endoscopy units serving the study area. The cases were also confirmed by regular reviews of the Cancer Register to ensure completeness of case ascertainment. Controls were randomly selected among Warsaw residents using a computerized registry of all legal residents in Poland. Controls were frequency-matched to cases by sex and age in 5-year strata. Information on demographic background and lifestyle factors was collected through interviews. Among eligible cases (n = 515) and controls (n = 549), successful interviews were conducted for

464 cases (90%) (324 direct and 140 next-of kin interviews for deceased cases) and 480 (87%) controls.

6.3.4 Study IV

In Study IV, subjects were followed from January 1, 1990 to December 31, 2004 in Cohort I and from January 1, 1993 to December 31, 2004 in Cohort II. Among subjects, 10% moved outside the study area, and 0.3% was lost to follow-up during the study period. Incident cancer cases were identified by active patient notification from the local major hospitals in the study area and data linkage with population-based cancer registries. Candidate patients were linked by name, address, and date of birth to the cancer registry for the JPHC study, when the date of birth and residence fulfilled cohort inclusion criteria.

Death certificate data were used as a supplementary information source. Data from histologic verification were available for 94% of gastric cancer cases and in 3.1% death certificated were the only source of information. The anatomic site of each case was coded using the International Classification of Diseases for Oncology, Third Edition (ICD-O-3). A tumor located in the lower side of the stomach was classified as distal gastric cancer (non-cardia site) (code C16.2-16.7) and in the upper side of the stomach as proximal gastric cancer (cardia site) (code C16.0-16.1). Tumors that could not be classified because of overlapping lesions (code C16.8) or no available information (code C16.9) were categorized as unclassified. Histological classification was based on review of the record from each hospital[89], and subdivided as follows: differentiated type (corresponding to intestinal type by Laurén's classification[13], including papillary adenocarcinoma and well/moderately-differentiated type) and undifferentiated type (corresponding to diffuse type by Laurén's classification[13], including poorly differentiated adenocarcinoma). Adenocarcinoma, not otherwise specified, and other types of carcinoma such as adenosquamous carcinoma, squamous cell carcinoma, carcinoid tumor, undifferentiated carcinoma and miscellaneous tumors were considered unclassified types. For each case, one control matched for sex, age (±3 years), study area, blood donation date (±2 months), and time since last meal at blood donation (±5 hours) was selected. The final analysis included 511 sets of matched cases and controls.

6.4 LABORATORY ANALYSIS

All laboratory analyses were carried out by trained personnel blinded to case-control status.

6.4.1 Measurement of antibodies against *H. pylori*

In Study I and Study II, status of *H. pylori* infection was assayed through serum IgG antibodies against Hp-CSA using ELISA, according to manufacturer's instructions.

Furthermore, measurement of antibodies against CagA was performed in Study I and Study II use of immunoblot assay (Helicoblot 2.1 and Helicoblot 2.0 respectively, Genelabs Diagnostics, Singapore) according to manufacturer's instructions, Figure 6-4.

Figure 6-4. Immunoblotting strips detecting *H. pylori* antibodies
Stripe 1 positive control, stripe 2 negative control, stripe 3 and 4 patients, the line to the very right in strip 1 and 4, indicates CagA positivity

In Study II faint CagA bands were considered as positive and in Study I positivity was assessed according to manufacturer's instructions. In Study II, for the population-based study *H. pylori* infection was considered positive when ELISA or immunoblotting showed the presence of antibodies against *H. pylori* and/or CagA[87], while in the hospital-based study *H. pylori* infection was considered positive when any of the following showed presence of *H. pylori* infection: culture, ELISA, immunohistochemistry or immunoblotting[90]. In Study IV, serum IgG antibodies against whole *H. pylori* and CagA were measured with an ELISA kit (E Plate "Eiken" *H. pylori* Antibody, Eiken Kagaku Co. Ltd., Tokyo, Japan)[91]. *H. pylori* infection was regarded as positive when antibodies against either whole *H. pylori* or CagA were positive.

6.4.2 Measurement of pepsinogen

In Study I, presence of severe or moderate corpus atrophic gastritis was identified by measurements of the serum PGI and PGII levels, using ELISA method (Biohit, Finland) according to manufacturer's instructions. Individuals with PGI below 70 µg/l in combination with the ratio of PGI to PGII less than 3 were considered as positive for severe or moderate corpus atrophic gastritis.

6.4.3 DNA preparation and genotyping

6.4.3.1 DNA purification

In Study II, for the population-based study, DNA was purified from 250 µl serum. The serum samples were centrifuged at 6000g for 1 h; 24 µl of the bottom solution was kept and the remaining supernatant discarded and 175 µl phosphate buffered saline and 10 µl Protinease K (Qiagen, Hilden, Germany) were added. The samples were heated to 56°C during gentle agitation for 30 min, followed by DNA preparation using the BioRobot M48 Workstation (Qiagen, Hilden, Germany) and the MagAttract DNA Blood M48 kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. In the hospital-based study, DNA was purified from whole blood using QIAamp96 DNA Blood Kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. In Study III, DNA was extracted from buffy coats by standard procedures.

6.4.3.2 Selection of SNPs

In Study II, the selection of SNPs, *IL1B*-511, *IL1B*-31, *IL1B*+3954 and *IL1RN*, was based upon those reported earlier[38].

In Study III, for the selection of tagSNPs, information from HapMap Phase 2 data[*] was used with LD-based method by Carlson *et al*[92]. The HapMap2 database contains 26, 18, and 57 SNPs in the *MUC1, MUC5AC,* and *MUC6* genes and their adjacent regions, respectively. The region analyzed included 20 kb upstream of the first exon and 10 kb downstream of the termination of the last exon. LD threshold $r^2 > 0.8$ and a minor allele frequency (MAF) > 5% were used in tagSNP selection.

6.4.3.3 Genotyping

In Study II, the DNA-fragments containing the polymorphic regions in *IL1B*-31 and *IL1B*-511 were amplified using a semi-nested polymerase chain reaction (PCR) in 50 µl reactions with HotStarTaq DNA polymerase (Qiagen, Hilden, Germany) using 0.5-1 µl of the primary PCR product. The *IL1B*+3954 fragment was amplified by PCR. The SNP detection was performed with pyrosequencing[93], in a PSQ96 sample setting (Biotage AB, Uppsala, Sweden), according to the manufacturer's instructions. The genotyping results of *IL1B*-31 and *IL1B*-511 were confirmed using real-time PCR assays with Taqman probes among 10% randomly selected subjects, according to standard procedure.

[*] http://www.hapmap.org

Further in Study II, the VNTR regions of *IL1RN* were identified by PCR amplification in 30µl reactions using DyNAzyme II polymerase (Finnzymes, Espoo, Finland)[46]. The fragment sizes were determined by 2% agarose gel electrophoresis according to standard procedures. The sizes of the PCR products were correlated to numbers of repeated units (86 bp) conventionally coded: 2/2 (240 bp, 2 repeats), 2/4 (240 bp-410 bp, 2-4 repeats), which is further explained in Figure 6-5.

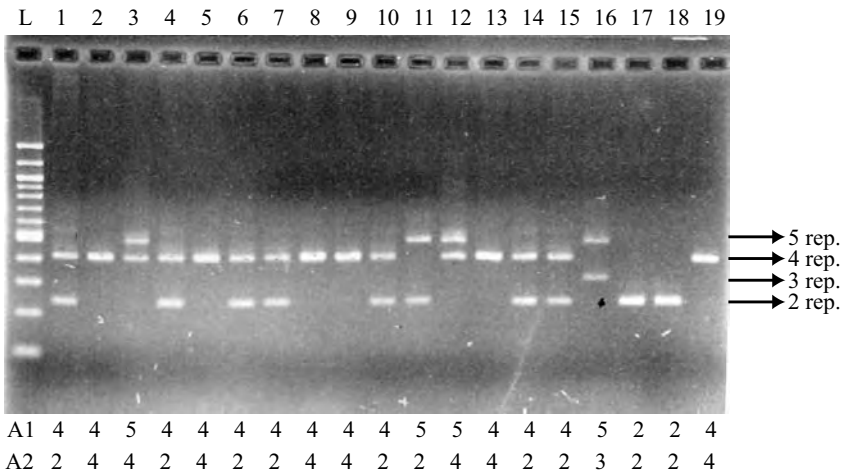


Figure 6-5. The VNTR of *IL-1RN*
The size of Allele 1 (A1) and Allele 2 (A2), showing genotype of 19 unrelated individuals (1-19), compared with 100 bp-distances ladder (L) (New England Biolabs Inc., Beverly, USA)

In Study III, genotyping was performed by SNPlex method at Core Genotyping Facility of National Cancer Institute, MD, USA. The SNPlex reactions were analyzed using an Applied Biosystem 3730 DNA Sequencer. Thirteen SNPs were excluded due to the assay design or genotyping failure, which left 6, 8, and 14 tagSNPs in *MUC1*, *MUC5AC*, and *MUC6* genes, respectively, were genotyped. These 28 tagSNPs span a region of 39.3 kb, 27.1 kb, and 110.5 kb across *MUC1*, *MUC5AC*, and *MUC6* gene, respectively. In total 273 cases with gastric cancer and 377 controls were successfully genotyped. Of the 28 tagSNPs, three (rs6427184 in *MUC1* gene, rs11602663 and rs7119740 in *MUC6* gene) did not conform to fitness for Hardy-Weinberg proportion among controls, thus these three SNPs were excluded from haplotype analyses.

6.4.4 Measurement of plasma levels of carotenoids, retinol, α- and γ-tocopherol

In Study IV, plasma levels of the carotenoids (β-cryptoxanthin, α- and β-carotene, lutein/zeaxanthin, lycopene), retinol, α- and γ-tocopherol were determined by reverse-phase high-performance liquid chromatography (HPLC) with a photodiode array detector[94]. In order to correct for recoveries, internal standards were run with each sample. Standard curves for β-cryptoxanthin, α- and β-carotene, lutein/zeaxanthin, lycopene, retinol, α- and γ-tocopherol displayed linearity. The detection limits were 0.1 μg/dl for β-cryptoxanthin, 0.1 μg/dl α- and β-carotene, 1.1 μg/dl for lutein/zeaxanthin, 0.5 μg/dl for lycopene, 1.4 μg/dl for retinol, 18.8 and 2.6 μg/dl for α- and γ-tocopherol among the plasma samples.

6.5 STATISTICAL ANALYSIS

6.5.1 Conditional logistic regression

In Study I, odds ratios (ORs) and their corresponding 95% confidence intervals (95% CIs) of developing stomach adenocarcinoma associated with *H. pylori* infection were estimated using conditional logistic regression models. Controls were individually matched to the cases with regard to age, sex, calendar year of serum collection, and biobank, thus these factors were inherently adjusted for.

In Study IV, ORs and 95% CIs were calculated using multivariable conditional logistic regression models, adjusted for family history of gastric cancer, smoking status, *H. pylori* infection, salt intake, consumption of highly salted foods (salted fish guts and roe) and body mass index (BMI). Further adjustments for serum levels of cholesterol were done for α- and γ-tocopherol. Smoking status was divided into never, past, current ≤20 cigarettes per day, and current ≥21 cigarettes per day. BMI was divided into ≤19.9 kg/m^2, 20-24.9 kg/m^2, and ≥25 kg/m^2. Family history was considered positive if at least one parent or sibling had gastric cancer. Plasma levels of β-cryptoxanthin, α- and β-carotene, lutein, lycopene, retinol, α- and γ-tocopherol were divided into quartiles according to plasma levels in the controls overall or by sex. Tests for trend were done by including continuous variables in the conditional logistic regression models. The statistical analyses in Study I and Study IV were performed using STATA/SE 9.2 (StataCorp, College Station, TX, USA).

6.5.2 Unconditional logistic regression

In Study II, the population-based study, ORs and 95% CIs were estimated using unconditional logistic regression models separately for each polymorphism, adjusting for age, sex and

county of residence and in the hospital-based study adjusted for age, sex and hospital. In the analysis of the *IL1RN* VNTR the presence of the short allele (2) was compared with the long allele (L-allele including 1, 3, 4 or 5).

In the supplementary analyses, relative risks for intestinal and diffuse histological types of gastric cancer were separately estimated, as were relative risks for cardia and distal stomach cancers. Analyses were further stratified by age, *H. pylori* infection status and family history of gastric cancer. All statistical analyses in Study II were performed using STATA/SE 9.2 (StataCorp, College Station, TX, USA).

In Study III, haplotypes were imputed by Expectation-Maximization method using all 5 tagSNPs for *MUC1* and 8 tagSNPs for *MUC5AC*. The haplotypes of *MUC6* were constructed based on the LD blocks derived from the HaploView 4.0 program[95]. Rare haplotypes (with frequency less than 2% among the controls) were combined as one group. The probabilities on an individual-subject basis of having certain haplotypes, inferred from the Expectation-Maximization algorithm, were used as weights in weighted logistic regression models with Sandwich covariance[96]. ORs and their corresponding 95% CIs were adjusted for age and sex. To account for type I errors of multiple testing, statistical significance was assessed by empirical P values derived from the Westfall & Young permutation (n=10000) using minimum P values and the step-down method. Global P values derived from the regression models were used to assess the difference in overall haplotype profiles between comparison groups. All statistical analyses were performed using SAS 9.2 (SAS institute, Cary, NC).

6.5.3 Linkage and HWE

In Study II and Study III, Hardy-Weinberg equilibrium assumption (HWE) was tested among the controls for all polymorphisms using the Pearson's $\chi 2$ goodness-of-fit test with 1 df. In both studies homozygotes for the more common allele served as reference. In Study II, pair-wise linkage was investigated using Kappa statistics.

6.6 ETHICAL CONSIDERATIONS

Study I

The study was approved by the Regional Research Ethics Vetting Board in Stockholm, Sweden.

Study II

The study was approved by the Ethics Committee of Medical Faculty, Uppsala University, and the Regional Research Ethics Vetting Board in Stockholm, Sweden.

Study III

The study was approved by Institutional Review Boards at the US National Cancer Institute, the Cancer Center and Institute of Oncology of Health in Warsaw, Poland, and the Regional Research Ethics Vetting Board in Stockholm, Sweden.

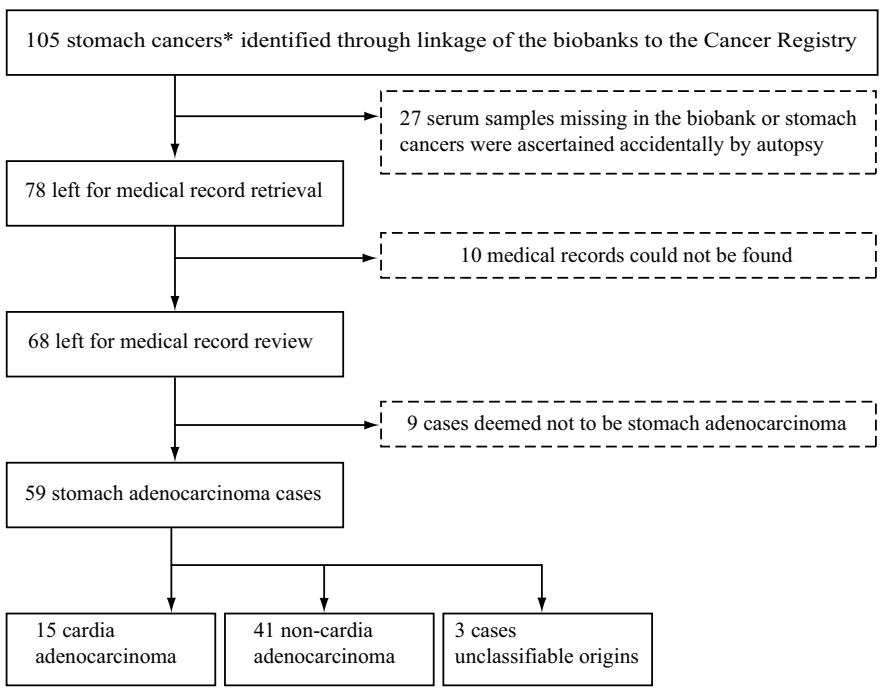
Study IV

The study was approved by the Institutional Ethical Review Board of the National Cancer Center, Tokyo, Japan.

7 RESULTS

7.1 *H. PYLORI* AND STOMACH ADENOCARCINOMA

In Study I a total of 105 stomach cancers (ICD-7 code 151), which occurred at least five years after blood sample collection, were identified during follow-up. Of these, 27 were excluded due to either missing serum samples or the fact that the cancers were found incidentally at autopsy. Of the remaining 78 cases, medical records were retrieved for 68. Nine were deemed to be tumors other than stomach adenocarcinoma. Among the remaining 59 stomach adenocarcinoma cases, 15 cases were classified as cardia adenocarcinoma, 41 as non-cardia stomach adenocarcinoma, and three of unclassifiable origins, Figure 7-1.



** Stomach cancer cases occurred 5 years after serum sampling*

Figure 7-1. Flowchart for selection of stomach cancer cases

No differences in the sex distribution and mean age at blood sample collection between cases and controls were observed, Table 7-1.

Table 7-1. Baseline characteristics of the case-controls in Study I

Characteristics	Stomach cancer, *n* (%) (*n*: 59)	Controls, *n* (%) (*n*: 117)
Men	30 (50.9)	60 (51.2)
Age at sample collection, mean (range)	30.8 (16-40)	30.9 (16-40)
Age at cancer diagnosis, mean (range)	47.3 (25-68)	
Years between serum collection and cancer diagnosis, mean (range)	16.5 (5-33)	
By anatomic site		
Cardia	15 (25)	
Non-cardia	41 (70)	
Unclassifiable origins	3 (5)	
Severe or moderate corpus atrophy*	5 (9)	1 (0.9)

*Defined as serum PGI <70 µg/l and PGI:II ratio <3 at the time of initial serum collection

The mean age at diagnosis of stomach adenocarcinoma was 47.3 years (range 25-68). The mean time between serum collection and stomach adenocarcinoma diagnosis was 16.5 years (range 5-33). Presence of severe or moderate corpus atrophic gastritis, based on PGI and PGII levels at the time of initial serum collection, was more common among stomach adenocarcinoma cases (9%) as compared with matched controls (0.9%).

ORs for all stomach adenocarcinoma among subjects with Hp-CSA antibodies (regardless of CagA serostatus) and antibodies against CagA (regardless of Hp-CSA serostatus) were 4.1 (95% CI, 1.9-8.5) and 3.5 (95% CI, 1.7-7.1), respectively (Table 7-2). In further analysis, data were combined for Hp-CSA and CagA antibodies, and the seronegative subjects for both were used as the unexposed reference group. Compared with this reference group, those seropositive with either Hp-CSA or CagA antibodies had a more than three fold increased risk of developing stomach adenocarcinoma. The excess risk was most pronounced among the group seropositive with both Hp-CSA and CagA antibodies (Table 7-2).

Table 7-2. Antibodies against Hp-CSAs and CagA and the risk of stomach adenocarcinoma

	Stomach adenocarcinoma			Cardia adenocarcinoma			Non-cardia adenocarcinoma		
Serology tests	**Case, *n***	**Control, *n***	**OR (95% CI)**	**Case, *n***	**Control, *n***	**OR (95% CI)**	**Case, *n***	**Control, *n***	**OR (95% CI)**
Hp-CSAs $^{-}$	20	76	Reference	12	21	Reference	6	51	Reference
Hp-CSAs $^{+}$	39	41	**4.1 (1.9-8.5)**	3	9	0.5 (0.1-2.8)	35	30	**17.1 (4.0-72.9)**
CagA^{-}	19	70	Reference	11	19	Reference	7	49	Reference
CagA^{+}	40	47	**3.5 (1.7-7.1)**	4	11	0.6 (0.2-2.5)	34	32	**10.9 (3.2-36.9)**
Hp-CSAs^{-} and CagA^{-}	15	61	Reference	10	18	Reference	4	41	Reference
Hp-CSAs^{+} or CagA^{+}	44	56	**3.3 (1.6-6.7)**	5	12	0.8 (0.2-2.7)	37	40	**9.7 (2.9-32.9)**
Hp-CSAs^{+} and CagA^{+}	35	32	**5.5 (2.3-12.9)**	2	8	0.3 (0.03-2.6)	32	22	**48.5 (5.8-407.4)**

*ORs were derived from conditional logistic regression models

In the analysis limited to the 41 cases of non-cardia stomach adenocarcinoma, the relative risks reached 17-fold and close to 11-fold among carriers of antibodies to Hp-CSAs and CagA, respectively. The corresponding results for cardia adenocarcinoma were 0.5 (95% CI, 0.1-2.8) and 0.6 (95% CI, 0.2-2.5), respectively, though not statistically significant.

Compared with the 'clean' reference group (those seronegative for both Hp-CSA and CagA antibodies) seropositive subjects with either Hp-CSA or CagA antibodies had a close to 10-fold increased risk that further increased to approximately 50-fold among those seropositive with both antibodies. The corresponding results for cardia adenocarcinoma were 0.8 (95% CI, 0.2-2.7) and 0.3 (95% CI, 0.03-2.6), respectively, though not statistically significant.

7.2 PROINFLAMMATORY INTERLEUKIN-1B AND STOMACH CANCER

In Study II, the male to female ratio was about 2:1. Close to 60% of the tumours in both the population-based and the hospital-based study were of intestinal histological type, and approximately one-third were of the diffuse type according to the Laurén's classification[13]. Cardia cancers were more common in the population-based study than in the hospital-based study (17% versus 9%), Table 7-3.

Table 7-3. Characteristics of the subjects in Study II

	Population-based study		Hospital-based study	
Characteristics	**Stomach cancer, *n* (%)** (*n:* 286)	**Control, *n* (%)** (*n:* 242)	**Stomach cancer, *n* (%)** (*n:* 65)	**Control, *n* (%)** (*n:* 297)
Age (SD)	67.9 (9.1)	67 (9.8)	72.7 (11.9)	70.9 (12.2)
Men	195 (68)	172 (1)	46 (71)	188 (63)
Family history of stomach cancer	49 (17)	22 (9)	8 (12)	41 (14)
H. pylori* positive	203 (71)	133 (55)	53 (82)	196 (66)
By histology type				
Intestinal	170 (59)	-	37 (57)	-
Diffuse	97 (34)	-	22 (34)	-
Mix or missing	19 (7)	-	6 (9)	-
By anatomic site				
Cardia	49 (17)	-	6 (9)	-
Non-cardia	210 (73)	-	59 (91)	-
Unclassified	27 (10)	-	-	-

*Presence of serum IgG antibodies against surface antigens of *H. pylori* or CagA

The genotype distribution of *IL1B*+3954, but not *IL1B*-31 and *IL1B*-511, deviated significantly among the controls from the HWE in both the hospital-based and population-based studies. The *IL1B*-31 and *IL1B*-511 were in linkage disequilibrium confirmed by pair-wise linkage analysis (population-based study, 96% agreement, Kappa=0.93, and hospital-based study, 92% agreement, Kappa=0.87).

In neither the population-based nor in the hospital-based study, no evidence in favour of an association between the three polymorphisms (*IL1B*-511, -31, +3954) and the risk of gastric cancer was found (Figure 7-2 and Figure 7-3). *IL1RN* (VNTR) polymorphisms were also analysed in the hospital-based case-control study, and no effect on risk of gastric cancer was observed.

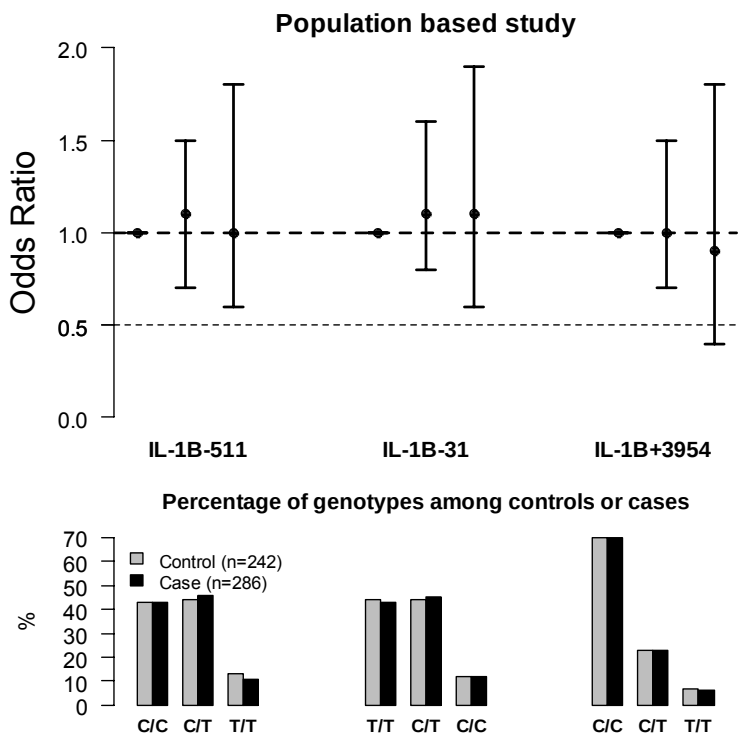


Figure 7-2. *IL1B* polymorphisms in the population-based study and stomach cancer risk
Showing ORs with 95% CI and the SNPs distribution in each genotype

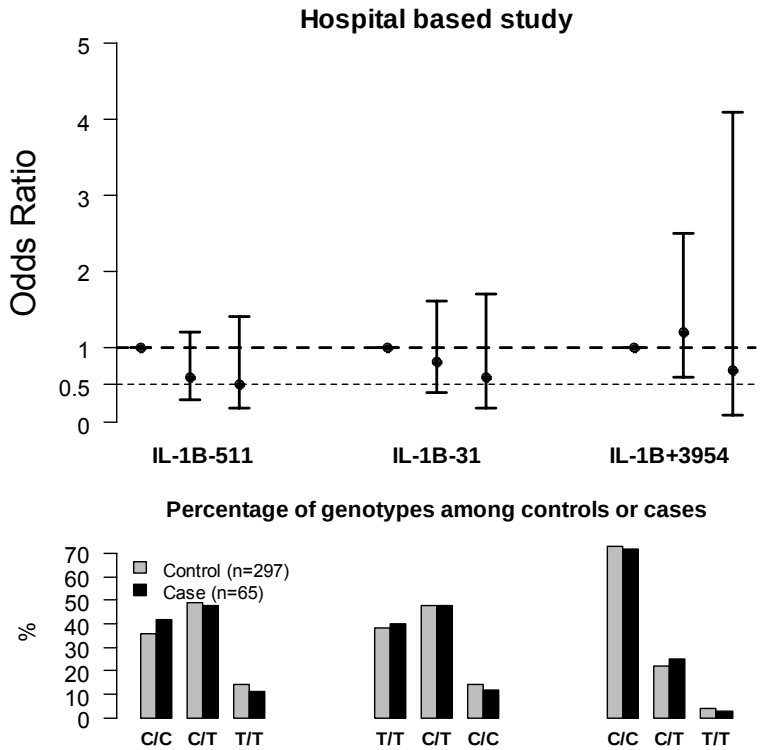


Figure 7-3. *IL1B* polymorphisms in the hospital-based study and stomach cancer risk
Showing ORs with 95% CI and SNPs distribution in each genotype

Further, the combined effect of numbers of polymorphisms was analysed, using subjects without any pro-inflammatory genotypes in the three studied *IL1B* loci as reference. ORs for those with one, two and three proinflammatory genotypes were 1.2 (95% CI, 0.7-2.1), 1.1 (95% CI, 0.7-1.7), and 1.0 (95% CI, 0.6-1.7), respectively.

None of the polymorphisms studied alone or in combination had any significant effect on the risk of gastric cancer.

No significant associations were found in separated analyses by Laurén's classification[13], nor by anatomic site (cardia and non-cardia cancers). Further stratification by age (<60 versus ≥60 years), *H. pylori* infection status, and family history of gastric cancer did not reveal any statistically significant association.

The results of the two studies were pooled using meta-analysis method (fixed-effect model). Homozygotes for *IL1B*-31 C had an adjusted OR of 0.9 (95% CI, 0.6-1.5). The corresponding ORs for *IL1B*-511 T and *IL1B*+3954 T were 0.8 (95% CI, 0.5-1.4) and 0.9 (95% CI, 0.5-1.7), respectively.

7.3 GENETIC VARAITION IN GENES CODING FOR MUCINS AND STOMACH CANCER RISK

Cases and controls in Study III were similar with respect to the distribution of age and sex, as selection of controls were frequency-matched on theses factors. Cases were more likely than controls to report a history of smoking and a family history of stomach cancer. Compared to controls, there were more former drinkers and less current drinkers among the cases. The majority of the cases had intestinal histological type (66%) and non-cardia cancer (72%), Table 7-4.

Table 7-4. Characteristics of the cases and controls in Study III

Characteristics	**Stomach cancer, *n* (%)** (*n*: 273)	**Control, *n* (%)** (*n*: 377)
Men	185 (68)	245 (65)
Family history of stomach cancer	34 (3)	16 (4)
H. pylori* positive	230 (84)	320 (85)
By histology type		
Intestinal	181 (66)	
Diffuse	46 (17)	
Mix or missing	46 (17)	
By anatomic site		
Cardia	34 (12)	
Non-cardia	197 (72)	
Mix or unclassified	42 (15)	

*Presence of IgG against antibodies Hp-CSA and/or CagA

There was no significant difference in the distribution of age, sex, smoking, *H. pylori* status, cancer phenotype, or alcohol use between the study subjects who were included in the genetic analysis from those who were not.

In Study III, the six tagSNPs tested across the *MUC1* region were found to be associated with an increased risk of stomach cancer, after adjusting for multiple tests, Figure 7-4.

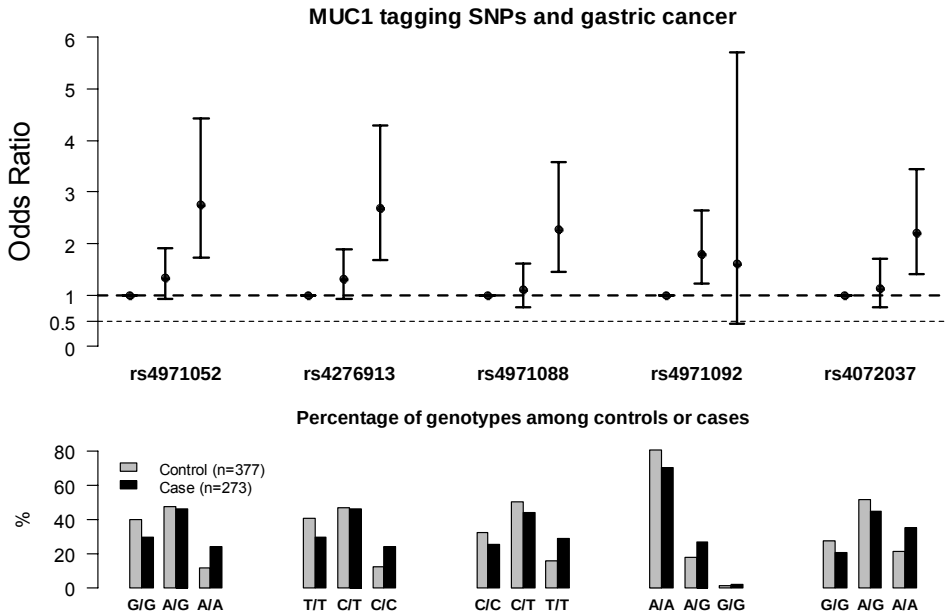


Figure 7-4. TagSNPs in *MUC1* and stomach cancer risk
Showing ORs with 95% CI and numbers SNPs distribution in each genotype

Furthermore, a gene-dosage effect was observed for all tagSNPs, with higher ORs significantly associated with an increasing number of minor alleles, except for rs4971092 and rs6427184. For rs4971092, homozygotes of minor allele were rare (only 5 cases and 5 controls), and rs6427184 had no minor allele homozygotes.

Haplotype analysis revealed 5 common haplotypes with frequency ≥ 5%, which had a cumulative frequency of 97% in controls. A statistically significant global association of haplotypes with stomach cancer risk was observed (P_{global}<0.0001). Carriers of the haplotype ACTAA containing the rare alleles of rs4971052, rs4276913, rs4971088, rs4971092, and rs4072037 with a frequency of 34.4% in controls, displayed a significantly increased risk of stomach cancer (OR 1.93 95% CI, 1.49-2.48) compared to the referent haplotype GTAAG, Table 7-5.

Table 7-5. Association between haplotypes of the *MUC1* gene and risk of stomach cancer

Haplotypes *MUC*[a]	Case (%)	Control (%)	OR (95% CI)[b, c]
GTAAG	31.4	46.2	Reference
ACTAA	45.2	34.4	**1.93 (1.49-2.48)**
GTTAA	4.0	6.3	0.92 (0.53-1.59)
GTAGG	8.8	5.3	**2.46 (1.59-3.81)**
GTAGA	6.6	4.8	**2.06 (1.25-3.40)**
Others	4.1	3.0	**2.02 (1.13-3.61)**

[a] SNP order: rs4971052, rs4276913, rs4971088, rs4971092, and rs4072037

[b] ORs adjusted for age and sex

[c] P_{global} value < 0.0001

The other two less common haplotypes (around 5% in controls), GTAGG and GTAGA, were also significantly associated with increased risks of the stomach cancer.

Of the eight tagSNPs across *MUC5AC* region, the minor allele of rs868903 was significantly associated with an increased risk of stomach cancer, Figure 7-5. There was also a tendency of gene-dosage effect (P_{trend}=0.0055), with ORs increasing from 1.74 for heterozygotes to 1.92 for homozygote carriers of the minor allele when compared to the common allele homozygotes. In addition, minor allele homozygotes of rs2014486 and rs2735733 were associated with an increased risk of stomach cancer when compared with their respective referent group, but neither were significant after adjustment for multiple testing.

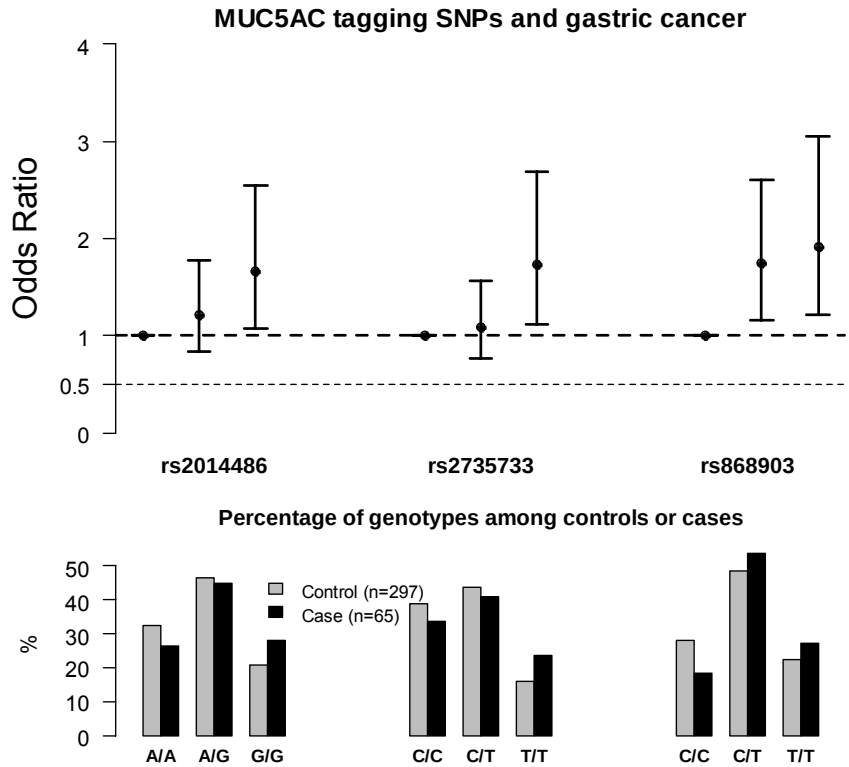


Figure 7-5. TagSNPs in *MUC5AC* and stomach cancer risk
Showing ORs with 95% CI and numbers SNPs distribution in each genotype

Ten haplotypes with a frequency ranging from 2% to 30% accounted for 93.4% of the observed haplotypes in controls (Table 7-6). An overall association with stomach cancer risk was observed (P_{global} =0.026).

Table 7-6. Association between haplotypes of the *MUC5AC* gene and risk of stomach cancer

Haplotypes *MUC5AC* [a]	Case (%)	Control (%)	OR (95% CI)[b, c]
GACACCCA	24.1	30.4	Reference
GACACCTA	17.8	14.5	**1.60 (1.19-2.15)**
GGTATCTA	14.7	13.1	1.43 (1.00-2.05)
GGTATCCA	10.1	9.4	1.37 (0.96-1.96)
AGCGTCTA	9.2	6.6	**1.86 (1.20-2.89)**

[a]SNP order: rs1541314, rs2014486, rs2075859, rs2672785, rs2735733, rs7118568, rs868903, rs4963049

[b]ORs adjusted for age and sex

[c]P_{global} =0.026

Haplotype GACACCTA, containing the minor allele of rs868903, had a frequency of 14.5% in controls and was associated with an increased risk of stomach cancer (OR 1.60, 95% CI, 1.19-2.15). Haplotype AGCGTCTA, with a frequency of 6.6% in controls, was also significantly associated with an increased risk (OR 1.86, 95% CI, 1.20-2.89).

No evidence of association with gastric cancer risk and 14 tagSNPs across the *MUC6* region was found. To minimize uncertainty in haplotype inference due to inclusion of a large number of tagSNPs, the estimated haplotype effects separately in the LD blocks. Similarly, no significant association was observed with any of the haplotypes in the two blocks.

When the analyses were restricted to non-cardia stomach cancer or subjects with *H. pylori* infection defined by positivity in anti-*H. pylori* or/and anti-CagA assays, the results did not change materially.

7.4 CAROTENOIDS, RETINOL, TOCOPHEROL AND RISK OF STOMACH CANCER

In Study IV, *H. pylori* infection, family history of gastric cancer and ever being a smoker were

Table 7-7. Baseline characteristics of the study subjects in Study IV

Characteristics	**Stomach cancer, *n* (%)** (*n*: 511)	**Control, *n* (%)** (*n*: 511)
Age, mean (SD)	57.4 (7.23)	57.4 (7.25)
Men	342 (67)	342 (67)
Family history of gastric cancer	36 (7.1)	16 (3.1)
***H. pylori* infection**[†]	505 (98.8)	460 (90.0)
Smoking status		
Never	234 (45.8)	253 (49.5)
Past	94 (18.4)	103 (20.2)
Current ≤20 daily cigarettes	140 (27.4)	113 (22.1)
Current ≥21 daily cigarettes	43 (8.4)	42 (8.2)
Alcohol intake		
Never	208 (41.9)	195 (39.7)
1-3 times per month	26 (5.2)	37 (7.5)
1-2 times per week	33 (6.7)	31 (6.3)
≥3 times per week	229 (46.2)	228 (46.4)
Body mass index, mean (SD)	23.0 (2.9)	23.4 (2.8)
Total cholesterol (mg/dl)[*], mean (SD)	193.2 (34.9)	197.5 (36.1)
Salt intake (g/day), mean (SD)	5.2 (2.6)	5.0 (2.3)
Highly salted foods (g/day), mean (SD)	12.5 (17.5)	12.1 (19.1)

[†]Including *Hp* IgG-positive or CagA-positive, [*]plasma levels of cholesterol

more common in patients with gastric cancer than in controls, Table 7-7.
Cholesterol levels were higher in the control group than in the gastric cancer cases.

The associations between plasma levels of lutein/zeaxanthin, β-cryptoxanthin, α- and β-carotene, lycopene, retinol, α- and γ-tocopherol and risk of gastric cancer were examined by using the lowest quartile as reference. A statistically significant inverse association for α-carotene in second and third quartiles, and for β-carotene in quartiles 2nd, 3rd, and 4th, was found. No statistically significant association was found for lutein/zeaxanthin, lycopene, retinol, α- or γ-tocopherol and stomach cancer risk, Table 7-8.

Table 7-8. Plasma levels of lutein/zeaxanthin, retinol, tocopherol, and the risk of gastric cancer

	Plasma level			
Variable	**median (µg/dl)**	**min (µg/dl)**	**max (µg/dl)**	**OR* (95% CI)**
Lutein/Zeaxanthin				
1st quartile	22.7	5.5	28.9	Reference
2nd quartile	34.1	29	39.5	0.91 (0.63-1.31)
3rd quartile	45.7	39.6	54.8	0.73 (0.49-1.10)
4th quartile	69.9	55	168.4	0.76 (0.49- 1.19)
P_{trend}				0.14
Lycopene				
1st quartile	0	0	3	Reference
2nd quartile	4.9	3	7.2	1.45 (1.00- 2.12)
3rd quartile	9.9	7.2	13.6	0.93 (0.60-1.44)
4th quartile	19.2	13.6	74.9	1.02 (0.66- 1.58)
P_{trend}				0.55
Retinol				
1st quartile	57	6.7	66	Reference
2nd quartile	73.4	66.1	82.4	1.29 (0.87-1.91)
3rd quartile	90.3	82.6	98.2	1.05 (0.68- 1.62)
4th quartile	113.1	98.3	229.3	1.39 (0.90- 2.16)
P_{trend}				0.25
α-tocopherol ‡				
1st quartile	973.1	89.1	1108.6	Reference
2nd quartile	1200.1	1111.2	1296.7	0.81 (0.54- 1.21)
3rd quartile	1409.6	1298.7	1540.2	0.93 (0.59- 1.47)
4th quartile	1760.5	1541.7	4948.2	1.01 (0.60- 1.68)
P_{trend}				0.91
γ-tocopherol ‡				
1st quartile	72.3	12.8	94	Reference
2nd quartile	110.9	94	127.9	0.99 (0.67-1.46)
3rd quartile	148.4	129	179.6	1.05 (0.71-1.58)
4th quartile	222.7	179.9	1130.2	0.94 (0.58-1.51)
P_{trend}				0.92

*ORs adjusted for family history of gastric cancer, *H. pylori* status, smoking, BMI, salt intake, and consumption of highly salted foods, ‡Further adjusted for plasma levels of cholesterol

The plasma carotenoid levels were found to vary by sex, which further lead to stratified analyses by sex, using sex specific quartiles for β-cryptoxanthin, α- and β-carotene. In men, an increased level of α- and β-carotene was associated with a reduced risk of gastric cancer, whereas no statistically significant association was found in women, Tables 7-9 and 7-10.

Table 7-9. Plasma levels of α-carotene and the risk of gastric cancer, stratified by sex, in Study IV

	Plasma level			
α-carotene	**median (µg/dl)**	**min (µg/dl)**	**max (µg/dl)**	**OR* (95% CI)**
Men				
1st quartile	0.7	0.0	1.2	Reference
2nd quartile	1.5	1.2	1.9	0.78 (0.49-1.24)
3rd quartile	2.4	1.9	3.1	0.66 (0.40-1.10)
4th quartile	4.1	3.1	23.0	**0.60 (0.36-1.00)**
P_{trend}				**0.04**
Women				
1st quartile	1.7	0.4	2.0	Reference
2nd quartile	2.5	2.0	2.9	0.69 (0.32-1.49)
3rd quartile	3.3	2.9	4.1	1.00 (0.50-2.02)
4th quartile	5.5	4.1	7.3	1.15 (0.56-2.40)
P_{trend}				0.49

*ORs adjusted for family history of gastric cancer, *H. pylori* status, smoking, BMI, salt intake, and consumption of highly salted foods

Table 7-10. Plasma levels of β-carotene and the risk of gastric cancer, stratified by sex, in Study IV

	Plasma level			
β-carotene	**median (µg/dl)**	**min (µg/dl)**	**max (µg/dl)**	**OR* (95% CI)**
Men				
1st quartile	3.3	0.0	6.2	Reference
2nd quartile	8.4	6.3	11.9	0.84 (0.54-1.32)
3rd quartile	15.7	12.1	20.4	**0.35 (0.21-0.60)**
4th quartile	28.5	20.5	119.0	**0.47 (0.27-0.81)**
P_{trend}				**< 0.01**
Women				
1st quartile	11.7	0.0	15.6	Reference
2nd quartile	19.4	15.6	24.0	1.00 (0.46-2.18)
3rd quartile	30.7	24.4	37.4	1.21 (0.54-2.68)
4th quartile	48.5	37.4	161.0	0.76 (0.35-1.70)
P_{trend}				0.59

*ORs adjusted for family history of gastric cancer, *H. pylori* status, smoking, BMI, salt intake, and consumption of highly salted foods

Furthermore, no statistically significant association for β-cryptoxanthin plasma level was found in neither men nor women, although its plasma level was more than twice as high in women than men, as were those of α- and β-carotene. For men, analysis stratified by smoking status, revealed a decreased risk with a statistically significant trend for β-carotene in both smokers and non-smokers as well as for β-cryptoxanthin in non-smokers.

No statistically significant association was found in stratified analyses by anatomic site or histological type, except that in men, carotene and β-cryptoxanthin seemed to confer a protective effect on distal differentiated gastric cancer, whereas a high β-carotene level was associated with a decreased risk of distal undifferentiated gastric cancer.

No statistically significant association for lutein/zeaxanthin, lycopene, retinol, α- or γ-tocopherol and risks of gastric cancer subtypes by anatomic subsite or histological type was found.

8 DISCUSSION

8.1 METHODOLOGICAL CONSIDERATIONS

8.1.1 Study design

In epidemiological research both observational and experimental study design are employed. Most epidemiological researches are based on observational study design including cohort study and case-control study. The timing of events divides cohort studies into retrospective (looking back in time) and prospective studies (looking forward in time). Two of the studies included in this thesis were of nested case-control study design, based on a retrospective cohort study (Study I) and a prospective cohort study (Study IV). In Study I, the advantage of the Swedish healthcare system and the national registration number assigned to every Swedish resident enabled linkage of historical records in biobanks with nationwide registers. Study IV is based on a prospective population-based cohort study with baseline exposure information and detailed information on potential confounders.

Cohort studies are considered less sensitive to recall bias since outcomes have not occurred when exposures are measured. In general the major disadvantage of retrospective cohort is usually the limited information on other key variables due to that studies typically rely on existing records which were not designed for the research purposes. In contrast, in a prospective cohort study more detailed information can be obtained on exposures and other variables due to that information is gathered from the participants directly.

In a case-control study, a smaller sample of the population is ascertained, among whom risk factors for a trait are investigated through comparing prevalence of exposures between ill individuals with those that are not ill. The cases are identified at diagnosis and then corresponding controls are selected from source population.
In Study II and Study III, the population-based case-control studies enrolled the cases from the hospitals were also ascertained via registries. The controls were randomly selected from the study area. Although this is a preferred study design, sometimes hospital-based case-control study might be the only choice, e.g. the endoscopy room-based study in the Study II where also the controls were selected.

8.1.2 Bias

Bias is defined as a systematic error, which results in an incorrect estimate of association between exposure and risk of disease. There are several types of bias, which can be grouped into two broad categories: selection bias and information bias.

8.1.2.1 Selection bias

Selection bias is a systematic error that occurs when the probability of being included in a study, often of case-control design, is related to the exposure. However, selection bias can also occur in cohort studies, when loss to follow-up is highly or directly related to the exposure and outcome under study or when the choice of exposed and unexposed individuals is related to developing the outcome of interest.

The loss of follow-up is ignorable in the Study I as use of personal registration number is widespread and of also high quality nationwide registers exist in Sweden. In the Study IV also rate of loss to follow-up was only 0.3%. In the Study III, 140 cases died before DNA samples could be taken. However, if the studied genetic variants were associated with survival this could be a matter of survival bias. The characteristics of cases who were interviewed and those with information provided by next of kin were compared. No differences in distributions by sex, age, education, and smoking status was found. Further the distribution of the genetic variants across tumor stages did not reveal any pattern of association.

8.1.2.2 Information bias

Information bias, occurs when the precision of a measurement differs between the comparison groups. The measurement error, also called the classification error or misclassification, classified into differential or non-differential misclassifications. Bias caused by differential misclassification can either overestimate or underestimate an effect, while the non-differential, also known as random misclassification, more likely dilutes the effect.

In the Study II, Study III, and Study IV, the stomach cancer cases were pathologically or histologically confirmed, most likely ruling out misclassification of the disease. In Study I, the stomach cancer cases were ascertained through the Swedish Cancer Registry. To ensure correct diagnosis as well as correct classification of tumor anatomic subsite, medical records were collected from the corresponding treatment or diagnostic units. Nine out of 68 tumor-cases with retrieved medical records were deemed to malignancies other than stomach adenocarcinoma.

Non-differential misclassification

In all the four studies, the laboratory analyses have been handled by professional lab personnel who were blinded to the case/control status. Further, in nested case-control studies, case and controls are usually matched on year of blood collection, and always measured in the same plate. These precautions might assure that, if misclassification existed, it would most likely be non-differential between case group and control group. In Study II and Study III, to assure no genotyping errors have occurred, the quality of the genotyping was evaluated by a randomly selected subsample repeatedly reanalyzed with 100% consistence.

In the Study IV, plasma carotenoid levels are measured to reflect food intake, based on that the half-life of carotenoids are couple of weeks[70]. Therefore, in this study, plasma levels are assumed to better reflect the exposure at a specific point in time than food frequency questionnaires due to recall bias. Though still, measurement at a single time point might possibly not reflect habitual intake over a longer period of time. To rule out variation in diet due to season and time of the day the samples were also matched for blood donation date, and fasting time before sample collection.

Differential misclassification

The *H. pylori* infection status is a source of differential misclassification, due to the disappearance of the bacteria during cancer development. To deal with this concern, Study I was designed to be based on a cohort of subjects whose blood samples were collected from young adults until age 40, which is a time window with best reflecting life-time infection. The prevalence of severe/moderate corpus atrophy among cases and controls was estimated by measuring baseline serum pepsinogen I and II levels. Although the prevalence rates are generally low, 9% of stomach cancer cases had an indication of severe/moderate corpus atrophy as compared with 0.9% in control. Thus, the possibility of underestimating the prevalence of *H. pylori* infection among cases did exist.

In the hospital-based case-control study in the Study II, controls were selected among patients scheduled for gastroscopy, thus possibly subjects with *H. pylori*-induced gastritis and atrophic gastritis are over-represented compared to the source population that generated the cancer cases. This might result in underestimation of the true association between the infection and stomach cancer risk.

8.1.3 Confounding

In short, a confounder is a risk factor associated with the exposure and at the same time with the outcome. In the Study IV, the observed inverse association between carotenoids and stomach cancer risk might be due to confounding by smoking, as smoking decreased the plasma levels of carotenoids in blood[98, 99] through inhibition of carotenoid uptake[60, 100, 101], and smoking is a risk factor for stomach cancer, Figure 8-1. However, smoking is only moderately associated with an increased risk of gastric cancer, and the inverse association with gastric cancer remained clear after adjustment for smoking.

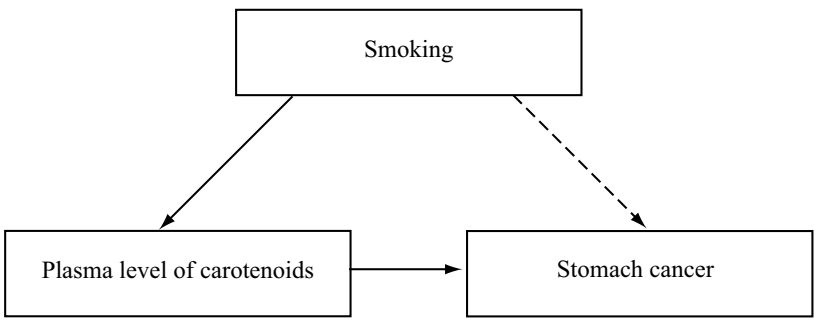


Figure 8-1. Confounding by smoking

Also alcohol intake has been reported to lower the amount of carotenoids in the blood[102], in the Study IV, regular drinkers (≥3 times per week) were found to have lower levels of carotenoids as compared with nondrinkers. However, since alcohol is not associated with gastric cancer risk, alcohol is unlikely to be an important confounder for the observed inverse associations in the Study IV.

Interestingly, *H. pylori* infection has been found to decrease the absorption of many nutrients, e.g. concentration of β-carotene is reduced in the gastric juice (though not in the plasma), and α-tocopherol reduced in the corpus[103]. In the Study IV, plasma levels of β-cryptoxanthin, α- and β-carotene, and lycopene were lower in *H. pylori* positive subjects as compared with reference individuals, but the results remained almost unchanged after adjustment for *H. pylori* infection status.

9 FINDINGS AND INTEPRETATION

9.1 *H. PYLORI* AND STOMACH ADENOCARCINOMA

Although *H. pylori* has been classified as a class I carcinogen for stomach cancer by the IARC in 1994[104], the strength of the association remains uncertain. In Study I, the aim was to investigate whether the association between *H. pylori* infection and gastric cancer risk has been underestimated. In an early meta-analysis, with serostatus of *H. pylori* assayed at time of cancer diagnosis, the summary odds ratio was only around 2[105], probably because of the late life (55+) development of this cancer, and spontaneous disappearance of the bacteria due to the histological changes in the stomach during cancer development. The time window of *H. pylori* infection and antibody presence in the blood is further described in Figure 9-1.

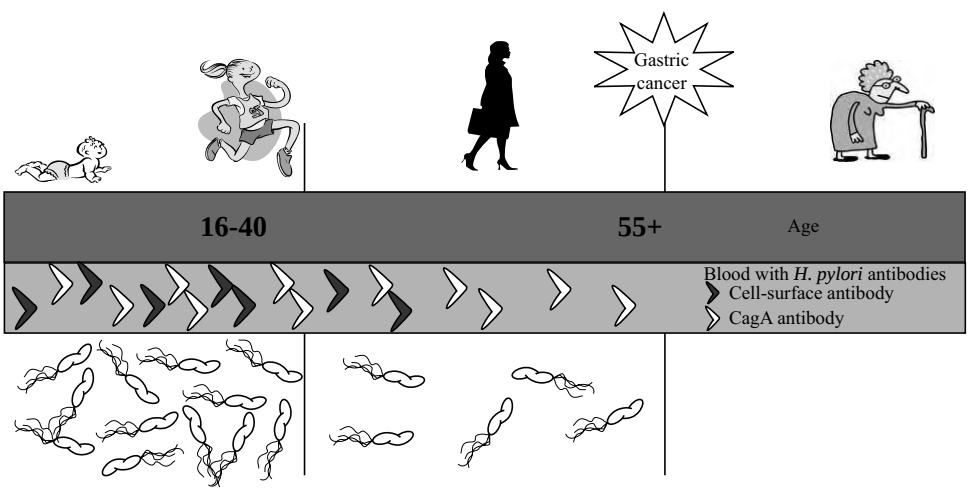


Figure 9-1. *H. pylori* over time in gastric cancer development

Already in 1994 a combined analysis of three prospective studies found that excess risks increased with increasing time window between sample collection and cancer diagnosis[106]. This finding was corroborated later by a pooled analysis of 12 prospective studies[107], in which a six-fold increased risk for non-cardia adenocarcinoma was found in patients whose samples were collected 10 years or more before cancer diagnosis, while the relative risk within 10 years after serotesting was 2.4. Similarly in the Alpha-Tocopherol, Beta-Carotene Cancer Prevention (ATBC) prospective cohort study, a close to 8-fold relative risk for non-cardia stomach cancer associated with *H. pylori* infection was reported[109]. In a Swedish population based case-control study[87], after the individuals with a presumed "burnt-out" *H. pylori* infection were removed from the reference group, the seropositive subjects, compared with the

'clean' reference group, had a 21-68 fold risk of non-cardia stomach cancer. Similarly, in a recent Japanese cohort study among middle-aged subjects[108], the group seronegative with H. pylori antibodies and with low PG I:II ratio had a more than 100-fold excess risk of stomach cancer[108].

Study I was based on a cohort of subjects whose blood samples were collected in early adulthood till 40 years of age; a critical time window to accurately measure *H. pylori* infection status before the infection changes the stomach environments. Further excluded, the first five years of observation to avoid possible reverse causality. The results suggest a strong association between *H. pylori* infection and risk of non-cardia stomach adenocarcinoma (OR=48.5, 95% CI=5.8-407.4, for the $CagA^+$ and $Hp\text{-}CSA^+$ stratum compared to the $CagA^-$ and $Hp\text{-}CSA^-$). An inverse association between cardia cancer and *H. pylori* infection was also found although not statistically significant. Interestingly, the ATBC prospective cohort study also reported an inverse association between *H. pylori* and cardia gastric cancer (OR 0.31 95% CI, 0.11-0.89).

However, Study I does not provide clear support for the hypothesis that *H. pylori* is a necessary cause of non-cardia stomach adenocarcinoma but the strength of the association between *H. pylori* infection and risk of non-cardia cancer was greater than some earlier results reported among older subjects.

9.2 IL1B GENETIC POLYMORPHISMS AND GASTRIC CANCER RISK

H. pylori have been shown to promote the production of inflammatory mediators such as IL-1β and TNF-α, both potent suppressors of stomach acid secretion. Worldwide, the relation between polymorphisms in genes associated with inflammatory response and the risk of gastric cancer has been extensively investigated since the landmark paper from El-Omar *et al*[38], which was published in the year 2000 and followed by a tremendous amount of studies in this field. Polymorphisms in several other genes related to inflammation have also been investigated in different populations and studies of variable qualities.

In a recent systematic review and meta-analysis, 40 studies (Figure 9-2) were found to have had examined the relationship between *IL1B*-511 T and the risk gastric cancer[110] between 1990 and 2006 with the majority of studies conducted after year 2000.

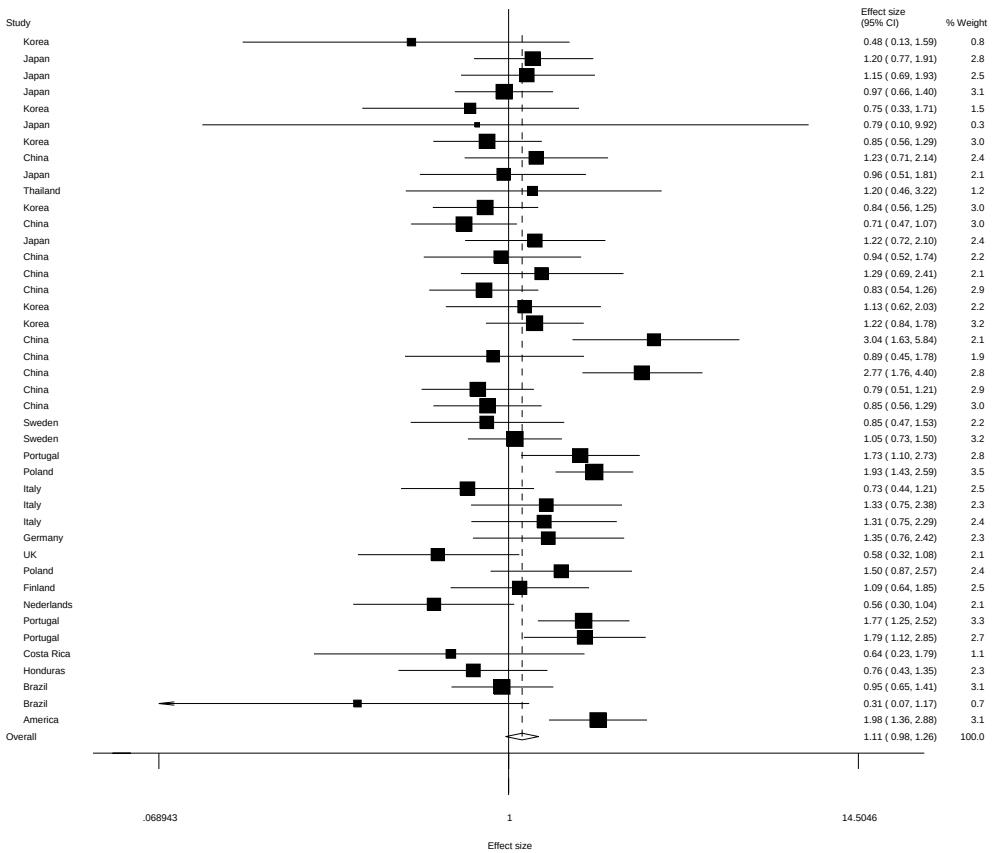


Figure 9-2. A cumulative analysis regarding *IL1B*-511 T SNP and stomach cancer risk between 2000 and 2006[110]

The inconsistent results from different studies might be due to various reasons. Definitions of histologic type and anatomic site of tumor might differ in Japan from the western world.

A cumulative meta-analysis on studies published since year 2000 revealed that an initially strong positive association diminished over time and showed overall null association in recent years, Figure 9-3. This implies that publication bias might exist.

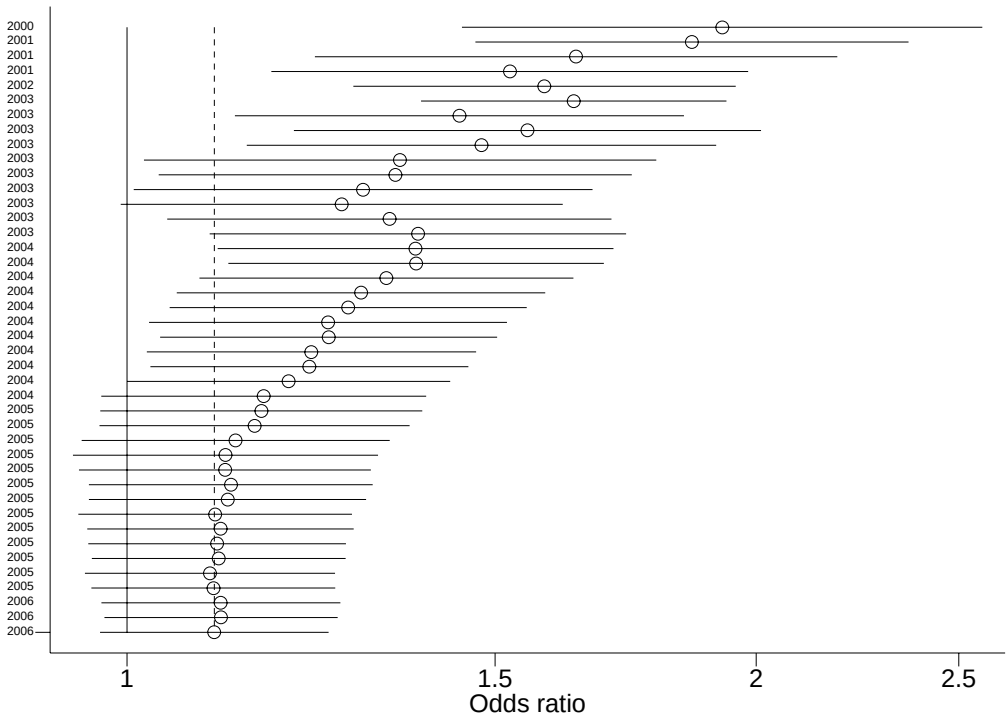


Figure 9-3. A cumulative analysis including *IL1B*-511 T SNP results between 1990 and 2006[110]

In Study II, the role of *IL1B*-31, *IL1B*-511, *IL1B*+3954 and *IL1RN* polymorphisms on gastric cancer risk was investigated in Sweden using a population-based case-control study and a hospital-based case-control study. The study revealed a statistically significant linkage between *IL1B*-31 and *IL1B*-511 but no linkage between *IL1B*-511 and *IL1-RN* or *IL1B*+3954, in accordance with other studies[38, 111]. The observed deviation from HWE for *IL1B*+3954 has been reported earlier[112]. No evidence was found for any important associations between the selected genotype variations in loci -31, -511 or +3954 of *IL1B* or the VNTR of *IL1RN* and the risk of gastric cancer, either combined or by anatomic or histological subtypes. These findings were consistent with a previous meta-analysis[55]. However, another recent meta-analysis[110] that performed stratified analysis by Asian and non-Asian population and study quality[42], reported

that among Asian populations, *IL1B*-31 C carrier was observed to be associated with a reduced overall risk which was also confirmed after restriction to high quality studies, Table 9-1.

Table 9-1. Risk for gastric cancer development in carriers of *IL1B* and *IL1RN* polymorphism[110]

	Overall studies (*N*)		Asian studies (*N*)		non-Asian studies (*N*)	
SNP	**All studies**	**HQs**	**All studies**	**HQs**	**All studies**	**HQs**
IL1B-511 T	-	-	-	§	-	-
IL1B-31 C	-	-	↓ (17)*‡	↓ (5)*‡	-	-
IL1B+3954 T	-	-	-	§	-	-
IL1RN 2	↑ (39)*‡	↑ (13)*‡	-	-	↑ (21)*	↑ (9)*‡

HQs-high quality studies, § less than 5 studies, ↑ increased risk, ↓ decreased risk, * also in *H. pylori* positive, ‡ also homozygote

In Study II no association was found between *IL1RN* 2 and gastric cancer risk, though a positive association was found in the meta-analysis[110]. This lack of association can be due to the small number of *IL1RN* 2/2 carriers in the hospital-based study. The *IL1RN* 2/2 genotype has been reported to be associated with high circulating IL1ra and IL1β levels, possibly resulting in severe and prolonged inflammatory response.

9.3 MUCIN GENETIC POLYMORPHISMS AND GASTRIC CANCER RISK

H. pylori dwell primarily in the gastric mucous layer that is basically built up by mucins[21]. Mucins are expressed in epithelia with cell type specificity, in the mucosa and the deep glands[18, 18-20, 22-26]. MUC1 is a highly polymorphic mucin expressed on the surface of many epithelia, including gastric mucosa. The expression of MUC5AC is reduced in carcinomas compared to normal tissue independently on sex, age, tumor staging, or tumor grading and correlated with worse survival of gastric cancer[113]. In addition, individuals carrying small VNTR alleles of *MUC6* were found to have an increased possibility of being infected with *H. pylori*[32]. Therefore, the aim of Study III was to investigate the roles of genetic variation in the *MUC1*, *MUC5AC*, and *MUC6* genes in the development of stomach cancer, based on a tagSNP approach.

The results revealed that multiple SNPs in *MUC1* and *MUC5AC* genes, as well as their haplotypes, were associated with an increased risk of stomach cancer. SNP rs4072037, located in exon 2 of *MUC1* gene, controls alternative splicing at the boundary between exon 1 and exon 2. Its A allele is associated with smaller size VNTRs. The study showed that homozygotes of A/A, likely with smaller size VNTR alleles, had an excess risk of stomach

cancer. The other five tagSNPs are all located in 3' flanking region of *MUC1* gene. Additional work is required to characterize the functional aspects of these SNPs and also to determine whether they are themselves the high risk alleles or in LD with a causal variant. Further, in eight of the *MUC5AC* tagSNPs, three (rs2014486, rs2735733, and rs868903) showed statistically significant association with stomach cancer risk (although the latter two did not reach significance level after adjustment for multiple testing), and an overall haplotype association with stomach cancer was also observed. Although MUC5AC forms the major receptor for *H. pylori* in the human stomach, no significant association by *H. pylori* status was found, possibly due to the small sample size, given that 84% of the cases and controls were tested *H. pylori* positive.

The availability of comprehensive SNP frequency data through the HapMap consortium lead to assessment of genomic regions of interest rather than simply genotyping SNPs of theoretical *a priori* significance, as investigated in Study II. Though, due to assay design or genotyping failure, not all tagSNPs in Study III were suitable for investigation, which might result in limitation of power in the gene-wide association test. Since the selection of SNPs in this study was based on a tagging approach rather than putative function, it is not possible to fully understand the functional significance of these findings. The exact nature of the functional alterations associated with each tagSNP will require further exploration.

9.4 CAROTENOIDS AND GASTRIC CANCER

The intake of fruit, as well as yellow vegetables (such as carrot, pumpkin) and white vegetables (such as Chinese cabbage, radish, tomato, and cucumber) has been reported to be inversely related to gastric cancer incidence[62, 63]. The protective mechanism of this effect is not fully understood but suggested to be regulated by antioxidants[62]. The effect of carotenoids on gastric cancer development has been studied using plasma levels[66, 68, 69, 115, 116] and food frequency questionnaires[64, 65, 67, 117-120], with varying results.

The aim of Study IV was to investigate the impact of plasma levels of carotenoids (β-cryptoxanthin, α- and β-carotene, lutein/zeaxanthin and lycopene), retinol, α- and γ-tocopherol on the risk of gastric cancer. In this study some carotenoids were found to be protective against gastric cancer development, in consistence with some studies[67, 68] but inconsistent with others[69, 115, 116]. The deficient or very low intake of β-carotene conferred an increased risk of gastric cancer development, consistent with results from previous intervention trials in China (Linxian study)[121, 122], Finland (ATBC study)[123], and the US (CARET study, PHS study)[124 98],

where subjects with sufficient levels do not benefit from further intake. Interestingly, a high dose might even increase the risk of cancer development[98]. Furthermore, plasma levels of α-carotene appear to be extremely low in the Japanese population in Study IV, especially among men, compared to men in Europe[69], and possibly the excess risk of gastric cancer could be detected because this study contained a group of men with low plasma levels of carotene.

The inverse association between plasma carotenoids and gastric cancer was found only in men This sex discrepancies can be explained by the different intake of carotenoids, and also by the minor influence of smoking and alcohol in Japanese women. Smoking was uncommon among women (less than 3%), while almost half of men (48%) smoked. Alcohol intake was also low among women, with about 16% drinking alcohol, whereas more than 65% of the men consumed alcohol at least three times a week. Women had thus higher or sufficient plasma levels than men. Therefore this difference explains the absence of an association with gastric cancer in women.

Furthermore, the small sample size in Study IV might explain the non-statistically significant findings in stratified analysis by anatomic site or histological type. With regard to the proximal site of gastric cancer, a high dietary intake of retinol has been reported to decrease risk and a high intake of α- and γ-tocopherol to increase it, but for the distal stomach carcinoma decreased risks have been reported with high intake of fruit, vitamin C, lycopene, α- and γ-tocopherol[60].

The findings suggest null effect of plasma levels in lutein/zeaxanthin, β-cryptoxanthin, lycopene, or retinol on gastric cancer, which was consistent with previous studies[65, 68, 115, 117, 119]. One possible reason might be due to that the individuals in Study IV had sufficient levels, therefore obviating the detection of any effect, despite other studies have shown an effect[66, 116].

10 CONCLUSIONS

- The results add confidence to the conclusion that *H. pylori* infection is a much stronger risk factor for non-cardia stomach adenocarcinoma than initially realized, but it is probably not a necessary cause. However, since misclassification of *H. pylori* status may have occurred also in this study, the association might be even stronger than indicated by the data.

- Human genetic polymorphisms linked to the production of IL-1β might not be associated with the risk of gastric cancer in the Swedish population.

- Polymorphisms in genes (*MUC1* and *MUC5AC*) encoding for mucins, that build up the gastric protective mucous layer, might contribute to an elevated risk of gastric cancer in the Polish population.

- Subjects with very low plasma levels of α-carotene and β-carotene are at a higher risk of gastric cancer.

11 FUTURE PERSPECTIVES

Although stomach cancer incidence is decreasing in most countries, it remains the fourth in cancer incidence and the second leading cause of cancer-related death in the world. *H. pylori* is a well-established risk factor for stomach cancer and has been classified as a definite human carcinogen by the IARC based on epidemiological and animal studies. Nearly half of the world's population is infected with *H. pylori*. The connection to cancer is convincing, but still more studies are needed to confirm or refute whether *H. pylori* infection is a necessary cause for distal stomach cancer.

In this thesis, the host genetic variety was also investigated for its role in the stomach cancer development. The employed approach was a traditional way, i.e. functional genes based on carcinogenesis model. Future genome wide association studies might provide new clues for susceptible genetic loci.

Besides human genetic susceptibility, exploration of the genetic variation in *H. pylori* genome might also pave a new way for understanding gastric carcinogenesis. Genes in *H. pylori* related to processes such as virulence, adhesion and survival of the bacteria might be of particular interests. The role of the changed microbiota in the atrophic gastric milieu is another interesting topic. Further studies are needed to explore interactions between bacterial infection, host genetic susceptibility, and other environmental exposures, and hopefully this integrated approach might eventually help to find a subgroup of particular high risk for distal stomach cancer, on whom primary prevention might be targeted.

12 POPLUÄRVETENSKAPLIG SAMMANFATTNING

12.1 ALLMÄN BAKGRUND

Epidemiologi innebär att undersöka olika faktorer som ur ett populationsperspektiv påverkar utvecklingen av en sjukdom. Etiologi i sin tur är läran om kausalitet, eller orsakssamband, som inom medicinen handlar om bakomliggande orsaker till vissa sjukdomstillstånd. I mitten av 1980-talet bekräftades att bakterien *Helicobacter pylori* (*H. pylori*) är den vanligaste orsaken till magsår. Upptäckten av det sambandet kom att bli avgörande för behandlingen av drabbade patienter. Några år senare gjordes även kopplingen mellan *H. pylori*-infektion och cancer i magsäcken – en av de cancerformer som årligen skördar flest liv i världen.

Ungefär hälften av världens befolkning beräknas vara infekterade med *H. pylori*, men endast ett fåtal utvecklar magcancer. Den här avhandlingen syftar till att på förhand kunna identifiera vilka individer som kommer att utveckla cancer i magen under inverkan av *H. pylori*-infektion. De olika delarbetena står för olika infallsvinklar – (i) skillnader i *H. pylori*s levnadsbetingelser i magsäcken, (ii) mänskliga genetiska variationer och (iii) variationer i kosten. Populationer i Sverige, Polen och Japan har studerats i epidemiologiskt utformade studier. Dessutom har laborativa metoder använts som redskap för att ge svar på frågeställningarna.

12.2 DELARBETE I

Delarbete I utreder möjligheten att *H. pylori*-infektionens koppling till cancer i magen är underskattad. Antagandet bygger på att bakterien spontant försvinner när cancer utvecklas, eftersom dess levnadsbetingelser i magen då förändras. Studien bygger på en så kallad historisk kohort, med redan existerande prover från individer valda ur två svenska biobanker. Individerna i studien har lämnat blodprov till biobanken innan de fyllt 40 år inom ramen för en nationell immunitetsutredning och senare utvecklat cancer i magen. Tumörens placering i magsäcken har visat sig vara avgörande för kopplingen till *H. pylori*. Därför skiljer studien på tumörer nära den övre magmunnen och tumörer som återfunnits på andra ställen i magsäcken. Studien visar ingen koppling mellan *H. pylori*-infektion och tumörer i övre magmunnen, men däremot en starkt ökad risk för utveckling av cancer i övriga magsäcken.

12.3 DELARBETE II

Delarbete II undersöker huruvida individuella genetiska variationer i gener som kodar för komponenter i kroppens immunförsvar påverkar risken att utveckla cancer i magen. Studien

undersöker Interleukin-1β, ett proinflammatoriskt cytokin som produceras av immunförsvaret och är en viktig del i kroppens försvar mot infektioner. Förutom funktioner i immunförsvaret reglerar *Interleukin-1β* också produktionen av saltsyra i magen och därmed också miljön för *H. pylori*. Genetiska variationer i Interleukin-1β-genen har påvisats minska produktionen av saltsyra, vilket gör att toxiska biprodukter från bland annat födoämnen ackumuleras i magsäcken och i förlängningen att ett cancerförlopp i magen kan inledas.

För att få en tydligare bild av hur genetiska variationer påverkar sannlikheten för magcancer inkluderar studien såväl ett sjukhus- som ett populationsbaserat material från svenska individer. Studien påvisar inte det samband mellan genetisk variation i den proinflammatoriska *Interleukin-1β*-genen och magcancer som tidigare rapporterats.

12.4 DELARBETE III

Delarbete III undersöker genvariationer relaterade till muciner, de molekyler som bland annat bygger upp slemmet i magen och därmed påverkar magslemhinnans egenskaper. Slemhinnor och slem är en del av kroppens immunförsvar och studien undersöker genvariation som är kopplad till slem som binder *H. pylori*. För att öka möjligheten att upptänka en eventuell koppling bygger studien på en högriskpopulation i Polen. Arbetet visar att det finns en rad genetiska variationer bakom mucinproduktionen, som är förknippade med en högre risk att utveckla cancer i magen.

12.5 DELARBETE IV

Avslutningsvis undersöktes i delarbete IV hur miljöfaktorer som exempelvis antioxidanter i kosten påverkar utvecklingen av magcancer. Undersökningen bygger på en japansk population. I Japan är majoriteten av befolkningen infekterad med *H. pylori* och förekomsten av cancer i magen bland de högsta i världen. I studien prövas antagandet att antioxidanter, till exempel karoten, förhindrar utvecklingen av cancer i magen under inverkar av *H. pylori*.

Studien visar att män med lägre karotenhalt i blodet löper högre risk att utveckla cancer i magsäcken än män med högre halt. Materialet visar också att förekomsten av karoten i blodet varierar mellan kvinnor och män, vilket sannolikt beror på olika levnadsvanor. Studien visar ingen effekt av karotenhalten hos kvinnor, vilket leder till slutsatsen att kvinnors karotenhalt (som generellt är högre än mäns) är tillräcklig för att förhindra utveckling av cancer i magen.

13 POPULAR SCIENCE SUMMARY

13.1 GENERAL BACKGROUND

Epidemiology consists of investigating different aspects of populations affecting the development of a disease. *Etiology*, on the other hand, is the science of causality, links of cause and effect, which within the field of medicine translates to the causes underlying disease. In the mid-nineteen eighties, it was confirmed that the bacterium *Helicobacter pylori* (*H. pylori*) is the most common cause of stomach ulcers. The discovery of that link was to become decisive for the treatment of ulcer patients. A few years later, the connection was made between *H. pylori* infection and cancer of the stomach - one of the forms of cancer that reaps the most lives each year in the world. As estimated half of the world's population is infected by *H. pylori*, but only a fraction develop stomach cancer. This thesis aims to identify beforehand which individuals will develop stomach cancer when infected by the *H. pylori* bacterium. The different parts of the thesis represent different facets - (i) differences in the living conditions of *H. pylori* in the stomach, (ii) human genetic variations, and (iii) variations in diet. Populations in Sweden, Poland, and Japan have been studied in epidemiologically designed studies. Furthermore, laboratory methods have been used as tools in order to give answers to the questions at issue.

13.2 PART I

Part I investigates whether the link between the *H. pylori* bacterium and cancer in the stomach might be underestimated. The assumption is based on the fact that the bacterium disappears spontaneously when cancer develops, since its living conditions change. The study is based on a so-called *retrospective cohort*, with already existing samples from individuals chosen from two Swedish biobanks. The individuals in the study have given blood samples to the biobank before they turned 40 (during a national immunity investigation), and later developed cancer in the stomach. The location of the tumor in the stomach has turned out to be crucial for the link to *H. pylori*. Therefore, the study distinguishes tumors in the upper part of the stomach (cardia) and tumors not found in the cardia (non-cardia stomach cancers). The study doesn't show any correlation between *H. pylori* infection and tumors in the cardia, but shows, on the other hand, a strongly increased risk of the development of non-cardia stomach cancer.

13.3 PART II

Part II investigates whether genetic variations in genes coding for components in the body's immune defense affect the risk of developing cancer in the stomach. The study investigates *Interleukin-1*β, a proinflammatory cytokine protein produced by the immune defense which forms an important part of the body's defense against infections. Besides being a part of the immune defense, *Interleukin-1*β also regulates the production of hydrochloric acid in the stomach, and thereby also the environment for *H. pylori*. Genetic variations in the *Interleukin-1*β gene have been shown to decrease the production of hydrochloric acid, causing toxic by-products from (among other things) foodstuffs to be accumulated in the stomach, eventually leading to the formation of cancer. In order to get a clearer picture of how genetic variations affect the probability of stomach cancer, the study includes hospital data as well as population data from Swedish individuals. The study does not show the previously reported correlation between genetic variation in the proinflammatory *Interleukin-1*β gene and stomach cancer.

13.4 PART III

Part III investigates genetic variations in connection with *mucins*, molecules which among other things form the mucus in the stomach and thereby affect the properties of the mucous membrane of the stomach. Mucous membranes and mucus are a part of the body's immune defense, and the study investigates genetic variation connected to mucus which binds *H. pylori*. In order to increase the chances of discovering a possible connection, the study is based on a high-risk population in Poland. The results show that there are a number of genetic variations behind the production of mucins that are linked to a higher risk of developing cancer in the stomach.

13.5 PART IV

Lastly, in part IV it was investigated how environmental factors, for example antioxidants in the diet, influence the development of stomach cancer. The examination is based on a Japanese population. In Japan, the majority of the population is infected by *H. pylori* and the incidence of cancer in the stomach is among the highest in the world. In the study, the assumption being tested is that antioxidants, for example carotene, prevent the development of cancer in the stomach under the influence of *H. pylori*. The study shows that men with lower carotene content in their blood run a higher risk of developing cancer in the stomach than men with a higher content. The data also shows that the occurrence of carotene in the blood varies between women and men, probably due to different lifestyles. The study does not show an effect of the carotene content on women, which leads to the conclusion that women's carotene

content (which is generally higher than men's) is sufficient to prevent the development of cancer in the stomach.

14 研究の要旨

14.1 背景

疫学とは、人間集団において、疾病の発生に影響を及ぼす様々な状態・特性を研究する学問である。一方、病因学とは、原因と結果をつなげる因果関係の学問であり、医学分野においては、疾病の原因を究明する学問と言い換えることができる。80年代半ばには、ヘリコバクター・ピロリ（ピロリ菌）は胃潰瘍に最も共通する原因菌であることがつきとめられた。この因果関係の発見は、胃潰瘍患者の治療法に決定的な道筋をつくった。その数年後には、ピロリ菌感染が、世界中で最も多くの生命を奪っているがんの一つである胃がんの原因であることが裏付けられた。ピロリ菌には世界の約50%が感染していると推計されているが、そのうち胃がん発生に至るのは、ほんのひと握りである。本学位研究では、既にピロリ菌感染している人のうち、どのような人がその後に胃がんを発生するのかを事前に同定することを最終的な目的とする。各パートでは、関連する以下の側面、すなわち、(i)胃におけるピロリ菌の生活環境における差違、(ii)ヒトにおける関連遺伝子型の差違、(iii)食事要因の差違、について述べる。スウェーデン人、ポーランド人、及び日本人集団を用いて疫学的手法により研究を実施した。さらに、課題に対する解答を得るために、実験的研究手法を用いた検証を加える。

14.2 パートⅠ

パートⅠでは、ピロリ菌感染と胃がんとの関連が過小評価されているかどうかを検証する。この仮定は、胃がんが発生すると、ピロリ菌の生活環境が変化して、自然に消滅してしまうことによるものである。この研究は、いわゆる後ろ向きコホート研究によるものであり、２つのスウェーデン人バイオバンクから抽出した既存生体試料を利用して研究を実施した。研究対象者は、40歳誕生日前に（国の免疫検査期間を利用して）血液試料をバイオバンクに提供し、その後胃がんと診断された者である。胃がんの原発亜部位はピロリ菌との関連にとって非常に重要であることが既に示されている。そのため、本研究では上位部（噴門部）胃がんと噴門部以外の胃がんとに分類した。結果としては、ピロリ菌感染と噴門部胃がんとの間には何の関連も見られなかったが、噴門部以外の胃がん発生のリスクを大きく増加させていた。

14.3 パートⅡ

パートⅡでは、体内の免疫防御機能にかかわる遺伝子コード化機能の遺伝的差違が胃がん発生リスクに影響を与えるかについて研究する。本研究は、人体の感染防御にとって重要な役割を担っている免疫防御機能によって産生した炎症性サイトカインの一つであるインターロイキン-

1βというタンパク質について検討している。インターロイキン-
1βは、免疫防御機能を担う他、胃塩酸の産生制御、ひいてはピロリ菌の生息環境の制御にかかわっている。インターロイキン-
1βの遺伝子型によっては、胃塩酸の産生を低下させ、胃内に蓄積された食餌由来の毒性副産物が産生され、結果として、胃発がんに結びつくことが示されている。どのような遺伝子型の差違が、どのように胃がん発生に影響するのか、さらに明確な答えを得るため、スウェーデン人集団とその病院データを用いた研究を実施した。その結果、以前に報告のある炎症性インターロイキン-
1βの遺伝子型の差違と胃がん発生との関連は見られなかった。

14.4 パートⅢ

パートⅢでは、胃粘液と関連して、胃粘液を産生し、胃粘膜の特性に影響を与える分子の遺伝子型について検討する。粘膜と粘液は人体の免疫防御の一部であるが、本研究では、ピロリ菌に結合する粘液の遺伝子型について検討する。この関連を確認できる可能性を高めるために、高危険集団であるポーランド人集団において本研究を実施した。結果として、粘液産生にかかわる様々な遺伝子型が胃がん発生リスクの上昇と関連していた。

14.5 パートⅣ

最後に、パートⅣでは、例えば食事中の抗酸化物質など、環境因子がどのように胃がんの進展に影響を与えているかについて検討した。本研究は、日本人集団において実施した。日本では、集団の大半がピロリ菌に感染しており、胃がん発生が、世界で最も高い国の一つである。本研究は、カロテン等の抗酸化物質がピロリ菌感染下でも胃がんの発生を予防できるかどうかを検証するものである。その結果、男性では、血中カロテン濃度が低い方が、胃がんの発生リスクがより大きかった。また、生活習慣の違いから血中カロテン濃度は男女で異なっていることも示された。本研究では、女性においてカロテンと胃がんとの関連はみられなかったが、これは、（男性よりも通常高値である）女性の血中カロテン濃度が胃がんの発生を予防するのに十分な量であったことによるものと結論づけた。

15 研究背景

流行病学研究人群中各种能够影响疾病发生发展的特性。病因学，从另外一个角度讲，是研究因果关系即起因和效果之间联系的科学。在医学中，起因指的是疾病的病因。在二十世纪八十年代中期，幽门螺旋杆菌被确认为导致胃溃疡的最常见的病因。这个发现很大程度上决定了溃疡病人的治疗。几年后，幽门螺旋杆菌和胃癌-
一个每年吞噬世界上众多生命的癌症-
之间的关联也被得以确认。据估计，将近一半的世界人口感染幽门螺旋杆菌，但是只有一小部分的人群最终被诊断为胃癌。这本论文主旨于区分出在感染幽门螺旋杆菌的人群中最终会发展为胃癌的小部分人。本论文的不同部分研究不同的方向: 1)
幽门螺旋杆菌在人体胃部的不同生存状况; 2) 人体的基因变异; 3)
饮食差异。应用不同的流行病学研究方法，我们对来源于瑞典，波兰，和日本的不同人群进行了研究。为了回答我们提出的研究问题，我们同时应用了先进的实验室技术作为辅助手段。

15.1 第一部分

第一部分研究旨在探讨现有的研究结果是不是低估了幽门螺旋杆菌和胃癌之间的关系。这个假设是基于在胃癌发展过程中幽门螺旋杆菌渐渐消失(可能是由于它们的生存状况发生了变化)的发现。这个研究包括从两个原有的瑞典生物样本库中选择的研究个体，组成了所谓的回顾性队列研究。这些研究个体在他们40岁之前在一个瑞典全国性的免疫调查中向这两个生物样本库贡献了血液样本，并且后来被诊断为胃癌。癌症在胃部的位置被证明在幽门螺旋杆菌和胃癌的关联中举足轻重。所以，在此次研究中，我们严格区分了在胃上部(贲门)的癌症和其他部位的癌症。本研究并未发现幽门螺旋杆菌和贲门部胃癌的任何关联，但是却观察到了幽门螺旋杆菌和非贲门部胃癌的强关联。

15.2 第二部分

第二部分研究探讨人体中免疫防御机能相关基因的变异和胃癌发生的关系。□细**胞介素**-
1β是一个机体免疫防御产生的炎性细胞因子，是机体防御感染的重要组成部分。除此之外，□细**胞介素**-

1β也能够调节胃内盐酸的产生，从而影响幽门螺旋杆菌的生存状况。□细胞介素-1β基因的变异能够降低盐酸生成，延缓食物分解，导致食物在胃部的残留以及有毒物质的产生，最终逐步导致癌症发生。为了更清楚的表明基因变异和胃癌发生的关联，本次研究应用了对不同瑞典胃癌病例的住院数据和人口数据。本次研究未能重复以往研究中所报道的炎性□细胞介素-1β基因突变和胃癌的关联。

15.3 第三部分

第三部分探讨与mucins相关的基因变异。Mucins是形成胃粘液并且影响胃粘膜性态的分子。胃粘膜和胃粘液是机体免疫防御系统的一部分。本次研究探讨和胃粘液相关的能够整合幽门螺旋杆菌的基因变异。为了能够更有效的发现一组可能的相关，本次研究基于波兰的一个高危人群。本次研究表明确实存在不少影响mucins产量的并且和高胃癌发生率相关的变异。

15.4 第四部分

第四部分探讨环境因素，例如饮食中的抗氧化因子，影响胃癌的发生。该研究在一个日本人群中进行。在日本，绝大多数的人口是感染幽门螺旋杆菌的，而且日本也是世界上胃癌发病率最高的国家之一。本次研究的假设是看氧化物，例如胡萝卜素，可以在幽门螺旋杆菌感染的人群中预防胃癌发生。本次研究表明，相比较于拥有较高血液胡萝卜素含量的人群，血液中胡萝卜素含量较低人群的胃癌发生率要高。这组关联在男性和女性人群中表现不同，可能由不同的生活方式导致。在女性人群中，这组关联不成立。可能的解释包括由于女性拥有比较高的血液胡萝卜素含量（平均高于男性），即使是比较低的水平都已经能够预防胃癌发生了，所以更高的血液胡萝卜素含量并不带来明显的预防效果。

16 НАУЧНО-ПОПУЛЯРНОЕ РЕЗЮМЕ

16.1 ОБЩИЕ ПОЛОЖЕНИЯ

Эпидемиология – наука о факторах, влияющих на развитие заболеваний среди населения. Этиология – наука о причинных связях патологических состояний. В середине 80-х годов было установлено, что бактерия *Helicobacter pylori (H. pylori)* является самой распространенной причиной язвы желудка. Выявление этой причинной связи стало решающим фактором в назначении лечения конкретным пациентам. Через несколько лет была установлена связь между *H. pylori* – инфекцией и раком желудка, одной из самых смертоносных форм рака. Считается, что примерно половина населения земли инфицирована *H. pylori*, но только у части инфицированных развивается рак желудка. Задачей данной работы является определение, у каких индивидуумов может развиться рак желудка под влиянием инфицирования бактериями *H. pylori*. В главах диссертации проблема рассматривается в различных аспектах: (i) различные условия развития *H. pylori* в желудке, (ii) генетические различия, (iii) различия в питании. Эпидемиология развития рака желудка изучалась у населения Швеции, Польши и Японии. Кроме того, для получения ответов на вопросы исследования использовались лабораторные методы.

16.2 ГЛАВА I

Первая глава исследует возможность того, что связь между *H. pylori*–инфекцией и раком желудка недооценена. Предположение базируется на том, что бактерия самопроизвольно исчезает при развитии рака, так как ее среда обитания в желудке изменяется. Исследование основано на так называемой ретроспективной когорте, включающей образцы, ранее собранные в двух шведских биобанках. Обследуемые сдавали образцы крови в биобанк до наступления возраста 40 лет в рамках национального иммунологического исследования. Выбирались образцы от лиц, у которых после сдачи анализов выявлялся рак желудка. Решающим для обнаружения связи с *H. pylori* оказался учёт расположения опухоли в желудке. Поэтому проводилось изучение различий между опухолями находящимися в верхней части желудка (cardia) и опухолями в других частях желудка (вне cardia). Исследование не показало какой-либо корреляции между инфекцией *H. pylori* и

опухолями расположенными в cardia. Однако, в других частях желудка обнаруживается значительное повышение риска развития опухоли.

16.3 ГЛАВА II

Вторая глава исследует, влияют ли индивидуальные генетические вариации в генах, кодирующих компоненты иммунной системы, на вероятность развития рака желудка. Исследование рассматривает протовоспалительный цитокин *Interleukin-1β* – важный фактор в борьбе организма с инфекцией, который вырабатывается иммунной системой. Кроме этих функций, *Interleukin-1β* регулирует выработку соляной кислоты в желудке, и следовательно, влияет на среду для *H. pylori.* Вариации в гене *Interleukin-1β* уменьшают производство соляной кислоты. Последнее приводит к аккумуляции токсических побочных продуктов в желудке и, со временем, возможному развитию ракового процесса.

Чтобы получить более полную картину влияния генетических изменений на вероятность развития рака желудка, в исследовании использовались как клинические материалы, так и обследование индивидуумов и групп населения Швеции. Наше исследование не подтвердило существование связи генетических вариаций протовоспалительного гена *Interleukin-1β-* с раком желудка, утверждавшееся ранее.

16.4 ГЛАВА III

Третья глава исследует генетические изменения, вызванные муцинами – молекулами, в числе прочего вырабатывающими слизь в желудке и соответственно влияющими на свойства его слизистой оболчки. Слизистая оболочка и слизь являются компонентами иммунной защиты организма и данная работа изучает генетические изменения, ассоциированные со слизью, которая связывает *H. pylori.* Чтобы увеличить возможность установления данной связи, исследование проводилось среди населения Польши, где фактор риска высок. Работа показывает, что существует ряд генетических изменений в выработке муцина, связанных с повышенным риском развития рака в желудке.

16.5 ГЛАВА I V

Наконец, в четвертой главе исследуются факторы окружающей среды, как например, антиоксиданты в продуктах питания, влияющие на развитие рака желудка. Исследование основано на изучении населения Японии, где большинство популяции инфицировано *H. pylori* и заболеваемость раком желудка одна из самых высоких в мире. Рассматривается предположение, что антиоксиданты (например, каротин) предотвращают развитие рака в желудке под влиянием *H. pylori.*

Исследование показывает, что люди с низким содержанием каротина в крови подвержены повышенному риску развития рака желудка, по сравнению с индивидуумами с высоким содержанием каротина. Материал также демонстрирует, что содержание каротина в крови различно у мужчин и женщин, что скорее всего зависит от различий в образе жизни. Исследование не обнаружило эффекта уровня каротина у женщин, что приводит к заключению, что уровень каротина у них (в общем более высокий, чем у мужчин) достаточен, чтобы предотвратить развитие рака желудка.

17 ACKNOWLEDGEMENTS

A PhD thesis is never built up by one person alone; it lay on the foundations of good mentorship and interesting discussions over time. If you are holding this thesis in your hand you have most likely been involved or crossed my scientific road in one way or another. Thank you!

My first contact with *H. pylori* infection was during an introduction course in microbial pathogenesis at the University of Linköping. The teachers were very dedicated, and when I started my PhD studies H. pylori came up on to of my mind when choosing subject.

I would like to express an extra gratitude to:

Weimin Ye, for your never ending energy and contagious interest in science that you have shown as my supervisor. Your knowledge and unlimited support over the years have always impressed me. Your fantastic network has taken me to places I didn't expect to visit during my PhD. Thank you for giving me the opportunity to build up a world wide network of my own. You are a great tutor.

Olof Nyrén, the co-supervisor who introduced me into epidemiology. Thank you for giving me this chance and for your deep knowledge in so many areas.

Lars Engstrand, the co-supervisor who, with never ending enthusiasm, encouraged me in my laboratory work at Swedish Institute for Infectious Disease Control where the limits are very few.

David Forman, for extraordinary supervision at Leeds University during my Marie Curie fellowship. You made my stay in Leeds memorable and educative.

Shoichiro Tsugane, for letting me do research with your group at National Cancer Center in Tokyo, and for the valuable discussions of research and life in Japan.
Manami Inoue, for your fantastic support and being a great tutor in research.
Shizuka Sasazuki, for being helpful and very kind in all matters.

The lost and found mentor **Cecilia Dahlberg.** I appreciate all your advice and your compassion. You are a true mentor.

My excellent coauthors: Stephen J. Chanock, Wong-Ho Chow, Joakim Dillner, Anna-Mia Ekström, Lars Engstrand, Helena Enroth, Lars-Erik Hansson, Lifang Hou, Motoki Iwasaki, Yanbin Jia, Jolanta Lissowska, Norie Kurahashi, Tsutomu Miura, Olof Nyrén, Helena Pettersson, Meredith Yeager and Zongli Zheng.

Zongli Zheng, **Yanbin Jia** and **Dariush Nasrollahzadeh** for being great people to work with both in the lab as well as in the scientific discussions in coauthorships. I appreciate you very much and have enjoyed all the chats we have had.

Special thanks to Marie Curie fellowship, Sweden Japan foundation, Edith and Erik Fernström foundation, Svenska läkarsällskapet, KI foundation, GroupGrant-in-Aid for Cancer Research and Third-Term Comprehensive Control Research for Cancer from the Ministry of Health Labor and Welfare Japan, for the financial support over the years.

Department of Medical Epidemiology and Biostatistics, Karolinska Institutet Sweden –to all the people that put an extra effort in making MEB a nice place to work at.
Hans-Olov Adami for building up a top of the line research department, and **Nancy Pedersen** and **Henrik Grönberg** for their continuous work in forming an inspiring research environment.

Swedish Institute for Infectious Disease Control-colleagues, for nice talks and encouragements.

Epidemiology department at Leeds University UK for supporting me in my research and taking very good care of me during my stay.

Epidemiology and Prevention Division, Research Center for Cancer Prevention and Screening at National Cancer Center Tokyo, for supporting me in my research and taking very good care of me during my stay.

The present members of **Prof Ye working group,** Zongli, Dariush, Fang, Caroline, Maria-Pia, Jiaqi, Anna, and Ulrika.

An extra gratitude to past and present members of the **MEB-PhD group**, you are all brilliant people and I hope that one day our roads will cross again.

***Pylori*-Pinglorna**: Elin, Annika, Hedvig, Mathilda, Anna, Farzad, Helene, Maria, Christina, Lena, Kristina, Marianne, Karin, Anders, Sönke, Annelie, Maria, Cecilia, Philippe, Britta, Annika and Cia.

My dear friends at MEB, especially Alexander, Lisa, Maria, Martin, Sara, Therese and Thomas. I can not express words. In short, thank you for everything! I will miss you very much!

The Sunaga house, for making my stay in Tokyo very special, for your fantastic energy, all support and help and introducing me in your fantastic network, where I have also made new friends.

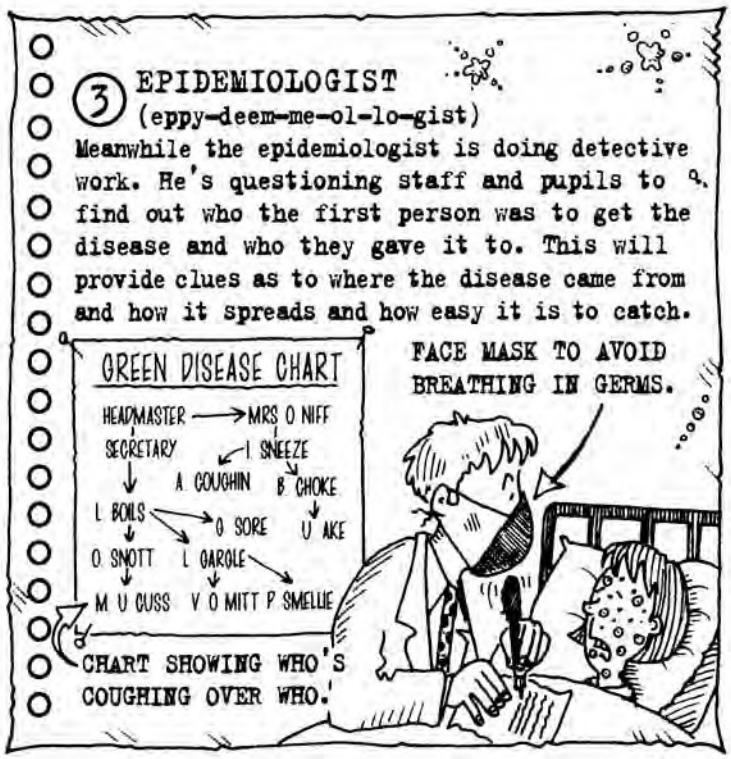
The Leeds Vic Friday club members, you are simply the best! An extra thank you to those of you who took me hiking, let me share house with you or joined me on my Betty's tour. True friends like you are hard to find, I miss you all.

The Fujimotto club members, the fantastic welcoming atmosphere, leading me into adventures and discussions I never thought I would experience.

My scientific quadruplets; Annika Nilsson, Kristina Nyström, Ulrika Zagai and Helena Ågerstam for everything, support in scientific and in private enjoyments.

My dear friends and family, for always being there in one way or another and reminding me about life outside work. Many of you have visited me and taken care of me during my stays abroad. Ni får mig att trivas lite extra i Stockholm, Östergötland, Skåne eller var ni än dyker upp. **Familjen Persson**, för att med er ändlösa energi finna ut lösningar på det mesta och för att det alltid blir en annorlunda avkoppling att åka söder ut. **Familjen Wennerström,** för att ni tar väl hand om mig.

Sven, for being around. From now on I intend to be your Dr Persson, Mrs Wennerström



With kind permission from Tony De Saulles

18 LIST OF FIGURES

19 REFERENCES

1. McColl,K.E. Cancer of the gastric cardia. *Best. Pract. Res. Clin. Gastroenterol.* **20**, 687-696 (2006).

2. Marshall,B.J. & Warren,J.R. Unidentified curved bacilli in the stomach of patients with gastritis and peptic ulceration. *Lancet* **1**, 1311-1315 (1984).

3. Schubert,M.L. Gastric secretion. *Curr. Opin. Gastroenterol.* **21**, 636-643 (2005).

4. Censini,S. *et al.* cag, a pathogenicity island of Helicobacter pylori, encodes type I-specific and disease-associated virulence factors. *Proc. Natl. Acad. Sci. U. S. A* **93**, 14648-14653 (1996).

5. Crabtree,J.E. *et al.* Helicobacter pylori induced interleukin-8 expression in gastric epithelial cells is associated with CagA positive phenotype. *J. Clin. Pathol.* **48**, 41-45 (1995).

6. Klaamas,K., Held,M., Wadstrom,T., Lipping,A., & Kurtenkov,O. IgG immune response to Helicobacter pylori antigens in patients with gastric cancer as defined by ELISA and immunoblotting. *Int. J. Cancer* **67**, 1-5 (1996).

7. Mbulaiteye,S.M., Hisada,M., & El-Omar,E.M. Helicobacter Pylori associated global gastric cancer burden. *Front Biosci.* **14**, 1490-1504 (2009).

8. Parkin,D.M., Bray,F., Ferlay,J., & Pisani,P. Global cancer statistics, 2002. *CA Cancer J. Clin.* **55**, 74-108 (2005).

9. Forman,D. & Burley,V.J. Gastric cancer: global pattern of the disease and an overview of environmental risk factors. *Best. Pract. Res. Clin. Gastroenterol.* **20**, 633-649 (2006).

10. Parkin,D.M., Pisani,P., Ferlay,J., Raymond,L., & Young,J. Cancer Incidence in Five Continents (International Agency for Research on Cancer, Lyon, 1997).

11. Sipponen,P. & Correa,P. Delayed rise in incidence of gastric cancer in females results in unique sex ratio (M/F) pattern: etiologic hypothesis. *Gastric. Cancer* **5**, 213-219 (2002).

12. van Blankenstein,M., Looman,C.W., Johnston,B.J., & Caygill,C.P. Age and sex distribution of the prevalence of Barrett's esophagus found in a primary referral endoscopy center. *Am J Gastroenterol* **100**, 568-76 (2005).

13. Lauren,P. The two histological main types of gastric carcinoma: diffuse and so-called intestinal-type carcinoma an attempt at a histo-clinical classification. *Acta Pathol. Microbiol. Scand.* **64**, 31-49 (1965).

14. Dixon,M.F., Genta,R.M., Yardley,J.H., & Correa,P. Classification and grading of gastritis. The updated Sydney System. International Workshop on the Histopathology of Gastritis, Houston 1994. *Am. J. Surg. Pathol.* **20**, 1161-1181 (1996).

15. Moss,S.F. & Blaser,M.J. Mechanisms of disease: Inflammation and the origins of cancer. *Nat. Clin. Pract. Oncol.* **2**, 90-97 (2005).

16. Correa,P. Helicobacter pylori and gastric carcinogenesis. *Am. J. Surg. Pathol.* **19 Suppl 1**, S37-S43 (1995).

17. Kim,Y.S., Gum,J., Jr., & Brockhausen,I. Mucin glycoproteins in neoplasia. *Glycoconj. J.* **13**, 693-707 (1996).

18. de,B.C., Garrido,M., & Real,F.X. MUC6 apomucin shows a distinct normal tissue distribution that correlates with Lewis antigen expression in the human stomach. *Gastroenterology* **109**, 723-734 (1995).

19. Audie,J.P. *et al.* Expression of human mucin genes in respiratory, digestive, and reproductive tracts ascertained by in situ hybridization. *J. Histochem. Cytochem.* **41**, 1479-1485 (1993).

20. Gendler,S.J. & Spicer,A.P. Epithelial mucin genes. *Annu. Rev. Physiol* **57**, 607-634 (1995).

21. Peek,R.M., Jr. & Blaser,M.J. Helicobacter pylori and gastrointestinal tract adenocarcinomas. *Nat. Rev. Cancer* **2**, 28-37 (2002).

22. Reis,C.A. *et al.* Intestinal metaplasia of human stomach displays distinct patterns of mucin (MUC1, MUC2, MUC5AC, and MUC6) expression. *Cancer Res.* **59**, 1003-1007 (1999).

23. Reis,C.A. *et al.* Immunohistochemical study of MUC5AC expression in human gastric carcinomas using a novel monoclonal antibody. *Int. J. Cancer* **74**, 112-121 (1997).

24. Carrato,C. *et al.* Differential apomucin expression in normal and neoplastic human gastrointestinal tissues. *Gastroenterology* **107**, 160-172 (1994).

25. Babu,S.D., Jayanthi,V., Devaraj,N., Reis,C.A., & Devaraj,H. Expression profile of mucins (MUC2, MUC5AC and MUC6) in Helicobacter pylori infected pre-neoplastic and neoplastic human gastric epithelium. *Mol. Cancer* **5**, 10 (2006).

26. Wang,R.Q. & Fang,D.C. Effects of Helicobacter pylori infection on mucin expression in gastric carcinoma and pericancerous tissues. *J. Gastroenterol. Hepatol.* **21**, 425-431 (2006).

27. Kocer,B. *et al.* A confirmatory report for the close interaction of Helicobacter pylori with gastric epithelial MUC5AC expression. *J. Clin. Gastroenterol.* **38**, 496-502 (2004).

28. Van de Bovenkamp,J.H. *et al.* The MUC5AC glycoprotein is the primary receptor for Helicobacter pylori in the human stomach. *Helicobacter.* **8**, 521-532 (2003).

29. Carvalho,F. *et al.* MUC1 gene polymorphism and gastric cancer--an epidemiological study. *Glycoconj. J.* **14**, 107-111 (1997).

30. Vinall,L.E. *et al.* Altered expression and allelic association of the hypervariable membrane mucin MUC1 in Helicobacter pylori gastritis. *Gastroenterology* **123**, 41-49 (2002).

31. Silva,F. *et al.* MUC1 gene polymorphism in the gastric carcinogenesis pathway. *Eur. J. Hum. Genet.* **9**, 548-552 (2001).

32. Nguyen,T.V. *et al.* Short mucin 6 alleles are associated with H pylori infection. *World J. Gastroenterol.* **12**, 6021-6025 (2006).

33. Garcia,E., Carvalho,F., Amorim,A., & David,L. MUC6 gene polymorphism in healthy individuals and in gastric cancer patients from northern Portugal. *Cancer Epidemiol. Biomarkers Prev.* **6**, 1071-1074 (1997).

34. Chow,W.H. *et al.* Risk of stomach cancer in relation to consumption of cigarettes, alcohol, tea and coffee in Warsaw, Poland. *Int. J. Cancer* **81**, 871-876 (1999).

35. Takashima,M., Furuta,T., Hanai,H., Sugimura,H., & Kaneko,E. Effects of Helicobacter pylori infection on gastric acid secretion and serum gastrin levels in Mongolian gerbils. *Gut* **48**, 765-73 (2001).

36. Santtila,S., Savinainen,K., & Hurme,M. Presence of the IL-1RA allele 2 (IL1RN*2) is associated with enhanced IL-1beta production in vitro. *Scand. J. Immunol.* **47**, 195-198 (1998).

37. Dinarello,C.A. Biologic basis for interleukin-1 in disease. *Blood* **87**, 2095-147 (1996).

38. El-Omar,E.M. *et al.* Interleukin-1 polymorphisms associated with increased risk of gastric cancer. *Nature* **404**, 398-402 (2000).

39. Furuta,T. *et al.* Interleukin 1beta polymorphisms increase risk of hypochlorhydria and atrophic gastritis and reduce risk of duodenal ulcer recurrence in Japan. *Gastroenterology* **123**, 92-105 (2002).

40. Hwang,I.R. *et al.* Effect of interleukin 1 polymorphisms on gastric mucosal interleukin 1beta production in Helicobacter pylori infection. *Gastroenterology* **123**, 1793-1803 (2002).

41. Kato,S. *et al.* Association of the interleukin-1 beta genetic polymorphism and gastric cancer risk in Japanese. *J. Gastroenterol.* **36**, 696-699 (2001).

42. Camargo,M.C. *et al.* Interleukin-1beta and interleukin-1 receptor antagonist gene polymorphisms and gastric cancer: a meta-analysis. *Cancer Epidemiol. Biomarkers Prev.* **15**, 1674-1687 (2006).

43. El-Omar,E.M. The importance of interleukin 1beta in Helicobacter pylori associated disease. *Gut* **48**, 743-7 (2001).

44. Rocha,G.A. *et al.* IL1RN polymorphic gene and cagA-positive status independently increase the risk of noncardia gastric carcinoma. *Int. J Cancer* **115**, 678-683 (2005).

45. Cox,A., Camp,N.J., Nicklin,M.J., di Giovine,F.S., & Duff,G.W. An analysis of linkage disequilibrium in the interleukin-1 gene cluster, using a novel grouping method for multiallelic markers. *Am. J. Hum. Genet.* **62**, 1180-1188 (1998).

46. Tarlow,J.K. *et al.* Polymorphism in human IL-1 receptor antagonist gene intron 2 is caused by variable numbers of an 86-bp tandem repeat. *Hum Genet* **91**, 403-4 (1993).

47. Chen,A. *et al.* Risks of interleukin-1 genetic polymorphisms and Helicobacter pylori infection in the development of gastric cancer. *Aliment. Pharmacol. Ther.* **20**, 203-211 (2004).

48. El-Omar,E.M. *et al.* Increased risk of noncardia gastric cancer associated with proinflammatory cytokine gene polymorphisms. *Gastroenterology* **124**, 1193-1201 (2003).

49. Garza-Gonzalez,E. *et al.* Role of the polymorphic IL-1B, IL-1RN and TNF-A genes in distal gastric cancer in Mexico. *Int. J. Cancer* **114**, 237-241 (2005).

50. Machado,J.C. *et al.* Interleukin 1B and interleukin 1RN polymorphisms are associated with increased risk of gastric carcinoma. *Gastroenterology* **121**, 823-829 (2001).

51. Glas,J. *et al.* Allele 2 of the interleukin-1 receptor antagonist gene is associated with early gastric cancer. *J Clin. Oncol.* **22**, 4746-4752 (2004).

52. Perri,F. *et al.* Cytokine gene polymorphisms in gastric cancer patients from two Italian areas at high and low cancer prevalence. *Cytokine* **30**, 293-302 (2005).

53. Zambon,C.F. *et al.* Pro- and anti-inflammatory cytokines gene polymorphisms and Helicobacter pylori infection: interactions influence outcome. *Cytokine* **29**, 141-152 (2005).

54. Zhang,W.H., Wang,X.L., Zhou,J., An,L.Z., & Xie,X.D. Association of interleukin-1B (IL-1B) gene polymorphisms with risk of gastric cancer in Chinese population. *Cytokine* **30**, 378-381 (2005).

55. Kamangar,F., Cheng,C., Abnet,C.C., & Rabkin,C.S. Interleukin-1B polymorphisms and gastric cancer risk--a meta-analysis. *Cancer Epidemiol. Biomarkers Prev.* **15**, 1920-1928 (2006).

56. Uemura,N. *et al.* Helicobacter pylori infection and the development of gastric cancer. *N. Engl. J. Med.* **345**, 784-789 (2001).

57. Malfertheiner,P. *et al.* Helicobacter pylori eradication has the potential to prevent gastric cancer: a state-of-the-art critique. *Am. J. Gastroenterol.* **100**, 2100-2115 (2005).

58. Sasazuki,S., Sasaki,S., & Tsugane,S. Cigarette smoking, alcohol consumption and subsequent gastric cancer risk by subsite and histologic type. *Int. J. Cancer* **101**, 560-566 (2002).

59. Kim,M.K., Sasaki,S., Sasazuki,S., & Tsugane,S. Prospective study of three major dietary patterns and risk of gastric cancer in Japan. *Int. J. Cancer* **110**, 435-442 (2004).

60. Nouraie,M. *et al.* Fruits, vegetables, and antioxidants and risk of gastric cancer among male smokers. *Cancer Epidemiol Biomarkers Prev* **14**, 2087-92 (2005).

61. Riboli,E. & Norat,T. Epidemiologic evidence of the protective effect of fruit and vegetables on cancer risk. *Am. J. Clin. Nutr.* **78**, 559S-569S (2003).

62. Serafini,M., Bellocco,R., Wolk,A., & Ekstrom,A.M. Total antioxidant potential of fruit and vegetables and risk of gastric cancer. *Gastroenterology* **123**, 985-991 (2002).

63. Kobayashi,M., Tsubono,Y., Sasazuki,S., Sasaki,S., & Tsugane,S. Vegetables, fruit and risk of gastric cancer in Japan: a 10-year follow-up of the JPHC Study Cohort I. *Int. J. Cancer* **102**, 39-44 (2002).

64. Ito,L.S. *et al.* Dietary factors and the risk of gastric cancer among Japanese women: a comparison between the differentiated and non-differentiated subtypes. *Ann. Epidemiol.* **13**, 24-31 (2003).

65. Botterweck,A.A., van den Brandt,P.A., & Goldbohm,R.A. Vitamins, carotenoids, dietary fiber, and the risk of gastric carcinoma: results from a prospective study after 6.3 years of follow-up. *Cancer* **88**, 737-748 (2000).

66. Abnet,C.C. *et al.* Prospective study of serum retinol, beta-carotene, beta-cryptoxanthin, and lutein/zeaxanthin and esophageal and gastric cancers in China. *Cancer Causes Control* **14**, 645-655 (2003).

67. Larsson,S.C., Bergkvist,L., Naslund,I., Rutegard,J., & Wolk,A. Vitamin A, retinol, and carotenoids and the risk of gastric cancer: a prospective cohort study. *Am. J. Clin. Nutr.* **85**, 497-503 (2007).

68. Yuan,J.M. *et al.* Prediagnostic levels of serum micronutrients in relation to risk of gastric cancer in Shanghai, China. *Cancer Epidemiol. Biomarkers Prev.* **13**, 1772-1780 (2004).

69. Jenab,M. *et al.* Plasma and dietary carotenoid, retinol and tocopherol levels and the risk of gastric adenocarcinomas in the European prospective investigation into cancer and nutrition. *Br. J. Cancer* **95**, 406-415 (2006).

70. Rock,C.L., Swendseid,M.E., Jacob,R.A., & McKee,R.W. Plasma carotenoid levels in human subjects fed a low carotenoid diet. *J. Nutr.* **122**, 96-100 (1992).

71. Tsugane,S. Salt, salted food intake, and risk of gastric cancer: epidemiologic evidence. *Cancer Sci* **96**, 1-6 (2005).

72. Tsugane,S., Sasazuki,S., Kobayashi,M., & Sasaki,S. Salt and salted food intake and subsequent risk of gastric cancer among middle-aged Japanese men and women. *Br. J. Cancer* **90**, 128-134 (2004).

73. Sjodahl,K. *et al.* Salt and gastric adenocarcinoma: a population-based cohort study in Norway. *Cancer Epidemiol. Biomarkers Prev.* **17**, 1997-2001 (2008).

74. Shin,C.M. *et al.* Stomach Cancer Risk in Gastric Cancer Relatives: Interaction Between Helicobacter pylori Infection and Family History of Gastric Cancer for the Risk of Stomach Cancer. *J. Clin. Gastroenterol.*(2009).

75. Blaser,M.J., Nomura,A., Lee,J., Stemmerman,G.N., & Perez-Perez,G.I. Early-life family structure and microbially induced cancer risk. *PLoS. Med.* **4**, e7 (2007).

76. Ekstrom,A.M. *et al.* Occupational exposures and risk of gastric cancer in a population-based case-control study. *Cancer Res.* **59**, 5932-5937 (1999).

77. Sjodahl,K. *et al.* Airborne exposures and risk of gastric cancer: a prospective cohort study. *Int. J. Cancer* **120**, 2013-2018 (2007).

78. Swedish Tax Agency (Skatteverket). Personal Identity Numbers. 2009. http://www.skatteverket.se/

79. Statistics Sweden (Statistiska centralbyrån). Total population registry. 2009. http://www.scb.se

80. National Board of Health and Welfare (Socialstyrelsen). Cause of Death Registry (Dödsorsaksregistret). 2009. http://www.socialstyrelsen.se/register/halsodataregister

81. National Board of Health and Welfare (Socialstyrelsen). Cancer Registry **(**Cancerregistret), 2009. http://www.socialstyrelsen.se/register/halsodataregister

82. Sobue,T. Current activities and future directions of the cancer registration system in Japan. *Int. J. Clin. Oncol.* **13**, 97-101 (2008).

83. Persson,C. *et al.* Interleukin 1-beta gene polymorphisms and risk of gastric cancer in Sweden. *Scand. J. Gastroenterol.* **44**, 339-345 (2009).

84. Persson,C. *et al.* Plasma levels of carotenoids, retinol and tocopherol and the risk of gastric cancer in Japan: a nested case-control study. *Carcinogenesis* **29**, 1042-1048 (2008).

85. Enroth,H., Kraaz,W., Engstrand,L., Nyren,O., & Rohan,T. Helicobacter pylori strain types and risk of gastric cancer: a case-control study. *Cancer Epidemiol. Biomarkers Prev.* **9**, 981-985 (2000).

86. Ekstrom,A.M. *et al.* Evaluating gastric cancer misclassification: a potential explanation for the rise in cardia cancer incidence. *J. Natl. Cancer Inst.* **91**, 786-790 (1999).

87. Ekstrom,A.M., Held,M., Hansson,L.E., Engstrand,L., & Nyren,O. Helicobacter pylori in gastric cancer established by CagA immunoblot as a marker of past infection. *Gastroenterology* **121**, 784-791 (2001).

88. Tsugane,S. & Sobue,T. Baseline survey of JPHC study--design and participation rate. Japan Public Health Center-based Prospective Study on Cancer and Cardiovascular Diseases. *J. Epidemiol.* **11**, S24-S29 (2001).

89. Japanese Gastric Cancer Association Japanese Classification of Gastric Carcinoma - 2nd English Edition -. *Gastric. Cancer* **1**, 10-24 (1998).

90. Ye,W. *et al.* Histology and culture results among subjects with antibodies to CagA but no evidence of Helicobacter pylori infection with IgG ELISA. *Scand. J. Gastroenterol.* **40**, 312-318 (2005).

91. Sasazuki,S. *et al.* Effect of Helicobacter pylori infection combined with CagA and pepsinogen status on gastric cancer development among Japanese men and women: a nested case-control study. *Cancer Epidemiol. Biomarkers Prev.* **15**, 1341-1347 (2006).

92. Carlson,C.S. *et al.* Selecting a maximally informative set of single-nucleotide polymorphisms for association analyses using linkage disequilibrium. *Am. J. Hum. Genet.* **74**, 106-120 (2004).

93. Ronaghi,M. Improved performance of pyrosequencing using single-stranded DNA-binding protein. *Anal. Biochem.* **286**, 282-288 (2000).

94. Talwar,D., Ha,T.K., Cooney,J., Brownlee,C., & O'Reilly,D.S. A routine method for the simultaneous measurement of retinol, alpha-tocopherol and five carotenoids in human plasma by reverse phase HPLC. *Clin. Chim. Acta* **270**, 85-100 (1998).

95. Barrett,J.C., Fry,B., Maller,J., & Daly,M.J. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics.* **21**, 263-265 (2005).

96. French,B. *et al.* Simple estimates of haplotype relative risks in case-control data. *Genet. Epidemiol.* **30**, 485-494 (2006).

97. Iwasaki,M. *et al.* Background characteristics of basic health examination participants: the JPHC Study Baseline Survey. *J. Epidemiol.* **13**, 216-225 (2003).

98. Goodman,G.E. *et al.* The Carotene and Retinol Efficacy Trial (CARET) to prevent lung cancer in high-risk populations: pilot study with cigarette smokers. *Cancer Epidemiol. Biomarkers Prev.* **2**, 389-396 (1993).

99. Peng,Y.M. *et al.* Concentrations and plasma-tissue-diet relationships of carotenoids, retinoids, and tocopherols in humans. *Nutr. Cancer* **23**, 233-246 (1995).

100. Gabriel,H.E. *et al.* A comparison of carotenoids, retinoids, and tocopherols in the serum and buccal mucosa of chronic cigarette smokers versus nonsmokers. *Cancer Epidemiol. Biomarkers Prev.* **15**, 993-999 (2006).

101. Handelman,G.J., Packer,L., & Cross,C.E. Destruction of tocopherols, carotenoids, and retinol in human plasma by cigarette smoke. *Am. J. Clin. Nutr.* **63**, 559-565 (1996).

102. Tsubono,Y., Tsugane,S., & Gey,K.F. Differential effects of cigarette smoking and alcohol consumption on the plasma levels of carotenoids in middle-aged Japanese men. *Jpn. J. Cancer Res.* **87**, 563-569 (1996).

103. Salgueiro,J. *et al.* Review article: is there a link between micronutrient malnutrition and Helicobacter pylori infection? *Aliment. Pharmacol. Ther.* **20**, 1029-1034 (2004).

104. IARC Schistosomes, liver flukes and Helicobacter pylori. IARC Working Group on the Evaluation of Carcinogenic Risks to Humans. Lyon, 7-14 June 1994. *IARC Monogr Eval Carcinog Risks Hum* **61**, 1-241 (1994).

105. Danesh,J. Helicobacter pylori infection and gastric cancer: systematic review of the epidemiological studies. *Aliment. Pharmacol. Ther.* **13**, 851-856 (1999).

106. Forman,D., Webb,P., & Parsonnet,J. H pylori and gastric cancer. *Lancet* **343**, 243-244 (1994).

107. Gastric cancer and Helicobacter pylori: a combined analysis of 12 case control studies nested within prospective cohorts. *Gut* **49**, 347-353 (2001).

108. Yanaoka,K. *et al.* Risk of gastric cancer in asymptomatic, middle-aged Japanese subjects based on serum pepsinogen and Helicobacter pylori antibody levels. *Int. J. Cancer* **123**, 917-926 (2008).

109. Kamangar,F. *et al.* Opposing risks of gastric cardia and noncardia gastric adenocarcinomas associated with Helicobacter pylori seropositivity. *J. Natl. Cancer Inst.* **98**, 1445-1452 (2006).

110. Persson,C., Canedo,P., Machado,J.C., El-Omar,E., & Forman,D. Polymorphisms in Inflammatory Response Genes and their Association with Gastric Cancer: Systematic Review and Meta-analyses: A HuGE Review. 2009. In manuscript.

111. Hamajima,N. *et al.* Interleukin 1 polymorphisms, lifestyle factors, and Helicobacter pylori infection. *Jpn. J. Cancer Res.* **92**, 383-389 (2001).

112. Pyo,C.W. *et al.* Polymorphisms of IL-1B, IL-1RN, IL-2, IL-4, IL-6, IL-10, and IFN-gamma genes in the Korean population. *Hum. Immunol.* **64**, 979-989 (2003).

113. Baldus,S.E. *et al.* Correlation of MUC5AC immunoreactivity with histopathological subtypes and prognosis of gastric carcinoma. *Ann. Surg. Oncol.* **9**, 887-893 (2002).

114. Xu,Q. *et al.* [The relationship between the polymorphism of MUC1 and susceptibility to gastric cancer in Liaoning region]. *Yi. Chuan* **30**, 1163-1168 (2008).

115. Ito,Y. *et al.* Cancer mortality and serum levels of carotenoids, retinol, and tocopherol: a population-based follow-up study of inhabitants of a rural area of Japan. *Asian Pac. J. Cancer Prev.* **6**, 10-15 (2005).

116. Taylor,P.R. *et al.* Prospective study of serum vitamin E levels and esophageal and gastric cancers. *J. Natl. Cancer Inst.* **95**, 1414-1416 (2003).

117. Chen,H. *et al.* Nutrient intakes and adenocarcinoma of the esophagus and distal stomach. *Nutr. Cancer* **42**, 33-40 (2002).

118. Garcia-Closas,R., Gonzalez,C.A., Agudo,A., & Riboli,E. Intake of specific carotenoids and flavonoids and the risk of gastric cancer in Spain. *Cancer Causes Control* **10**, 71-75 (1999).

119. Harrison,L.E., Zhang,Z.F., Karpeh,M.S., Sun,M., & Kurtz,R.C. The role of dietary factors in the intestinal and diffuse histologic subtypes of gastric adenocarcinoma: a case-control study in the U.S. *Cancer* **80**, 1021-1028 (1997).

120. Lissowska,J. *et al.* Diet and stomach cancer risk in Warsaw, Poland. *Nutr. Cancer* **48**, 149-159 (2004).

121. Blot,W.J. *et al.* Nutrition intervention trials in Linxian, China: supplementation with specific vitamin/mineral combinations, cancer incidence, and disease-specific mortality in the general population. *J. Natl. Cancer Inst.* **85**, 1483-1492 (1993).

122. Blot,W.J. *et al.* The Linxian trials: mortality rates by vitamin-mineral intervention group. *Am. J. Clin. Nutr.* **62**, 1424S-1426S (1995).

123. Albanes,D. *et al.* Alpha-Tocopherol and beta-carotene supplements and lung cancer incidence in the alpha-tocopherol, beta-carotene cancer prevention study: effects of base-line characteristics and study compliance. *J. Natl. Cancer Inst.* **88**, 1560-1570 (1996).

124. Hennekens,C.H. *et al.* Lack of effect of long-term supplementation with beta carotene on the incidence of malignant neoplasms and cardiovascular disease. *N. Engl. J. Med.* **334**, 1145-1149 (1996).