

From THE DEPARTMENT OF MEDICAL EPIDEMIOLOGY AND  
BIostatISTICS

Karolinska Institutet, Stockholm, Sweden

# **ETIOLOGY AND PROGNOSIS OF GASTROESOPHAGEAL CANCERS**

Isabella Ekheden



**Karolinska  
Institutet**

Stockholm 2020

All previously published papers were reproduced with permission from the publisher.

Published by Karolinska Institutet.

Printed by US-AB 2020

© Isabella Ekheden, 2020

ISBN 978-91-7831-777-6

# Etiology and prognosis of gastroesophageal cancers

## THESIS FOR DOCTORAL DEGREE (Ph.D.)

By

**Isabella Ekheden**

*Huvudhandledare:*

Professor Weimin Ye  
Karolinska Institutet  
Institutionen för Medicinsk Epidemiologi och  
Biostatistik

*Bihandledare:*

Professor emeritus Olof Nyrén  
Karolinska Institutet  
Institutionen för Medicinsk Epidemiologi och  
Biostatistik

Docent Amelie Plymoth  
Karolinska Institutet  
Institutionen för Medicinsk Epidemiologi och  
Biostatistik

Med. Dr. Pauline Raaschou  
Karolinska Institutet  
Institutionen för Medicin, Solna  
Enheten för Klinisk Epidemiologi

*Opponent:*

Professor Peter Malfertheiner  
Otto v. Guericke Universitetet Magdeburg  
Institutionen för Gastroenterologi, Hepatologi och  
Infektionssjukdomar

*Betygsnämnd:*

Docent Håkan Jonsson  
Umeå universitet  
Institutionen för Epidemiologi och global hälsa

Professor Per Hall  
Karolinska Institutet  
Institutionen för Medicinsk Epidemiologi och  
Biostatistik

Docent Peter Thelin Schmidt  
Karolinska Institutet  
Institutionen för Medicin, Solna  
Enheten för Klinisk Medicin



To my family



# ABSTRACT

Gastric and esophageal cancer – gastroesophageal cancers, belong to the most fatal malignancies in the world. Understanding the cause of these diseases is key to interventions such as primary prevention and/or surveillance with the potential of lowering the disease burden. Several important exposures have been identified or suggested but the causes of gastroesophageal cancers are still largely unknown. Further, only one in three gastric cancer patients and one in six esophageal cancer patients are still alive five years after their diagnosis. The prognosis is considerably improved if the tumor can be surgically removed, but 70-80% of the patients are not eligible for surgery due to advanced spread of the disease at diagnosis. Both curable and non-curable gastroesophageal cancer patients may be treated with chemotherapy – but reports about the survival outcome from real-world patients based on the treatment they receive, are very scarce. In this thesis we explored important biological determinants for gastroesophageal cancers to encourage further research on their etiology. Furthermore, we did a follow-up study on the survival of gastroesophageal cancer patients who were treated with various chemotherapy regimens to determine their effectiveness in a real-world setting.

We used population-based registers in Sweden and a case-control study in a high-risk region to explore the association between biological risk markers and gastroesophageal cancers. In study I we employed the nationwide Stomach Biopsy Cohort (SBC) study in Sweden to estimate the risk for gastric cancer associated with family history of gastric mucosal lesions. In study II we used the Epidemiology Strengthened by histoPathology Reports in Sweden (ESPRESSO) study to explore the risk for gastroesophageal cancers associated with esophageal lesions. In study III, we performed a case-control study in a high-risk area in China to study the association between gastric atrophy and esophageal squamous cell carcinoma, and examine its interaction with poor oral health, which might further increase the risk of esophageal squamous cell carcinoma. In study IV, we explored the influence of various chemotherapy regimens on the survival of gastroesophageal cancer patients based on a regional study on cancer chemotherapy at the Regional Cancer Center in Stockholm-Gotland.

In **study I**, we found that the excess risk was 50-60% higher for gastric non-cardia cancer among individuals who had a first-degree relative with gastric mucosal lesions (atrophic gastritis/intestinal metaplasia/dysplasia) compared to the general Swedish population. In **study II**, we demonstrated that non-dysplastic gastric and glandular metaplasia patients had a similar excess risk (Standardized Incidence Ratio, SIR 11.9; 95% Confidence Interval, CI, 9.9-14.1) for esophageal adenocarcinoma as intestinal metaplasia patients (SIR 10.8; 95% CI 7.8-14.6). In **study III**, we confirmed an association between gastric atrophy and esophageal squamous cell carcinoma (Odds Ratio 1.61; 95% CI 1.33-1.96), which was further increased in the presence of poor oral health (Relative Excess Risk due to Interaction 1.28; 95% CI 0.39-2.18). In **study IV**, we discovered that among patients who were intended to have curative treatment, esophageal cancer patients who received cisplatin-fluorouracil had better survival than those with carboplatin-fluorouracil (Hazard Ratio, HR, for carboplatin-fluorouracil vs cisplatin-fluorouracil 2.18; 95% CI 1.09-4.37), but gastroesophageal junction cancer patients treated with cisplatin-fluorouracil had worse survival than patients with fluorouracil-oxaliplatin (HR for fluorouracil-oxaliplatin vs cisplatin-fluorouracil 0.28; 95% CI 0.08-0.96).

We conclude that family history of gastric mucosal lesions can be employed for further risk stratification for non-cardia gastric cancer but needs to be evaluated regarding cost-effectiveness. Further, non-dysplastic columnar metaplasia patients may benefit from strengthened surveillance, but further validation studies are required. Moreover, gastric atrophy and its interaction with poor oral health are associated with esophageal squamous cell carcinoma in a high-risk region in China, thus future studies of the microbial alterations associated with gastric atrophy and poor oral health in the development of esophageal squamous cell carcinoma are warranted. Last, the choice of cisplatin-fluorouracil was associated with better survival outcome in esophageal but worse outcome in gastroesophageal junction cancer patients in Sweden. This finding needs to be further explored on a national level in Sweden.



# LIST OF SCIENTIFIC PAPERS

The following articles are referred to by their Roman numerals throughout, and are presented in full at the end of this thesis.

- I. Song H, Ekheden IG, Ploner A, Ericsson J, Nyren O, Ye W. **Family history of gastric mucosal abnormality and the risk of gastric cancer: a population-based observational study.** *International journal of epidemiology*. 2018;47(2):440-9.
- II. Ekheden I, Ludvigsson JF, Yin L, Elbe P, Ye W. **Columnar metaplasia elevates the risk for esophageal adenocarcinoma - a histopathology-based register study in Sweden.** (*Manuscript*)
- III. Ekheden I\*, Yang X\*, Chen H, Chen X, Yuan Z, Jin L, Lu M, Ye W. **Gastric atrophy and its interaction with poor oral health elevate the risk for esophageal squamous cell carcinoma in a high-risk region of China: a population-based case-control study.** *American Journal of Epidemiology* (2020).  
<https://doi.org/10.1093/aje/kwz283>

\*These authors contributed equally to this work

- IV. Ekheden I, Ebrahim F, Ólafsdóttir H, Raaschou P, Wettermark B, Henriksson R, Ye W. **Survival of esophageal and gastric cancer patients with adjuvant and palliative chemotherapy – a retrospective analysis of a register-based patient cohort.** *Eur J Clin Pharmacol* (2020).  
<https://doi.org/10.1007/s00228-020-02883-3>

# CONTENTS

|       |  |    |
|-------|--|----|
| 1     | INTRODUCTION .....   | 1  |
| 2     | BACKGROUND .....   | 3  |
| 2.1   | The burden of esophageal and gastric cancer .....                                | 3  |
| 2.1.1 | Classification .....   | 3  |
| 2.1.2 | Geographical distribution.....   | 4  |
| 2.1.3 | Age distribution.....  | 5  |
| 2.1.4 | Sex differences .....  | 5  |
| 2.1.5 | Environmental risk factors .....   | 5  |
| 2.1.6 | Familial aggregation and genetic risk factors.....                               | 8  |
| 2.1.7 | Esophageal precursor lesions.....  | 9  |
| 2.1.8 | Treatment of gastroesophageal cancers.....                                       | 9  |
| 3     | AIMS .....   | 11 |
| 4     | MATERIALS AND METHODS .....  | 13 |
| 4.1   | Research approach.....   | 13 |
| 4.2   | Data sources .....   | 13 |
| 4.2.1 | Registers .....  | 13 |
| 4.2.2 | Case-control study.....  | 20 |
| 4.3   | Statistical methods.....   | 21 |
| 4.3.1 | Measurements of risk, relative risk, and underlying assumptions.....             | 21 |
| 4.3.2 | Random and systematic error .....  | 23 |
| 4.4   | Research ethics .....  | 25 |
| 4.4.1 | Ethical principles and approvals .....   | 25 |
| 4.4.2 | Ethical considerations .....   | 26 |
| 5     | MAIN RESULTS .....   | 27 |
| 5.1   | High-risk individuals for gastroesophageal cancer .....                          | 27 |
| 5.1.1 | Family history of gastric mucosal changes .....                                  | 27 |
| 5.1.2 | Pre-malignant esophageal lesions.....  | 29 |
| 5.1.3 | Gastric atrophy .....  | 31 |
| 5.2   | Treatment factors associated with the prognosis of gastroesophageal cancer ..... | 32 |
| 5.2.1 | Choice of chemotherapy .....   | 32 |
| 6     | DISCUSSION .....   | 33 |
| 6.1   | From biological risk markers to etiology.....                                    | 33 |
| 6.1.1 | Familial clustering – is it due to shared environment or genetics?.....          | 33 |
| 6.1.2 | Gastroesophageal morphological lesions and the “point of no return” .....        | 35 |
| 6.2   | Personalized treatment – beyond the hype .....                                   | 39 |
| 6.3   | Strengths and limitations.....   | 41 |
| 7     | CONCLUSIONS.....   | 42 |
| 8     | FUTURE STUDIES.....  | 42 |
| 8.1   | Family-clustering.....   | 42 |

|     |  |    |
|-----|--|----|
| 8.2 | Precursor lesions.....                   | 43 |
| 8.3 | Personalized treatment .....             | 43 |
| 9   | POPULÄRVETENSKAPLIG SAMMANFATTNING ..... | 44 |
| 10  | ACKNOWLEDGEMENTS.....                    | 47 |
| 11  | REFERENCES.....                          | 49 |

## LIST OF ABBREVIATIONS

|                    |   |
|--------------------|---|
| AP                 | Attributable Proportion due to interaction                |
| BE                 | Barrett's Esophagus                                       |
| BMI                | Body Mass Index   |
| CI                 | Confidence Interval                                       |
| EC                 | Esophageal Cancer   |
| ESCC               | Esophageal Squamous Cell Carcinoma                        |
| EAC                | Esophageal Adenocarcinoma                                 |
| GAC                | Gastric Adenocarcinoma                                    |
| GC                 | Gastric Cancer  |
| GERD               | Gastroesophageal Reflux Disease                           |
| <i>H. pylori</i>   | <i>Helicobacter pylori</i>                                |
| HR                 | Hazard Ratio  |
| NSAID              | Non-Steroidal Anti-Inflammatory Drug                      |
| OR                 | Odds Ratio  |
| PGI, PGII          | Pepsinogen I, Pepsinogen II                               |
| PPI                | Proton Pump Inhibitor                                     |
| RERI <sub>OR</sub> | Relative Excess Risk due to Interaction (derived from OR) |
| S                  | Synergy index   |

# 1 INTRODUCTION

According to the most recent report from the International Agency for Research on Cancer (IARC), there were around 0.7 million deaths from stomach cancer and 0.4 million deaths from esophageal cancer in 2012. In terms of number of cancer deaths, these cancers rank third and sixth worldwide (1). In the same year, there were an estimated 0.95 million and 0.48 million incident cases. The high mortality to incidence ratio implies a poor prognosis, not only because of the aggressive growth that characterizes these malignancies, but also because these cancers are more common in developing countries, where most cases are first diagnosed at an advanced stage. Japan – a high-risk country for gastric cancer – is an exception to the high mortality: incidence ratio from gastric cancer. Early diagnosis of gastric cancer has proven effective in Japan, where the 5-year survival rate is above 90% mainly due to a screening program and early intervention. This is outstanding compared to other countries, including Europe where 5-year survival rates are much lower, 10-25% (2). The combination of high incidence and good prognosis could also be due to overdiagnosis at screening; however, there is no current standardized method of estimating this (3). Screening programs would, however, not be feasible for gastric and esophageal cancers in low-risk areas such as Sweden. Better understanding of the etiology of both cancers could help risk-stratifying patients and allow for primary prevention, targeted screening and surveillance of high-risk individuals. Such strategies would be potentially successful for helping more patients survive gastric and esophageal cancers in low-risk countries.



## 2 BACKGROUND

### 2.1 THE BURDEN OF ESOPHAGEAL AND GASTRIC CANCER

#### 2.1.1 Classification

Esophageal cancer displays two major histological types: esophageal squamous cell carcinoma (ESCC) and esophageal adenocarcinoma (EAC). ESCC is the predominant type, accounting for 88% of all new cases in 2012 worldwide (4).

Unlike esophageal cancer, the main histologic type of gastric cancer is adenocarcinoma, accounting for more than 90% of the cases (5). Other rarer histopathologic types are lymphoma, leiomyosarcoma (5) and neuroendocrine tumors (6).

By anatomic site, gastric cancer can also be divided into cardia and non-cardia gastric cancer (5). Although there is no present consensus on the anatomical definition, the proximal cardia cancer and the more distal non-cardia gastric cancer have different risk factors (5). The incidence of gastric cancer has declined from the most common cancer in the 1975 to the fifth most common cancer in 2018 (7). Declining incidence of non-cardia gastric cancer has been suggested to be mainly due to a decrease of *Helicobacter pylori* (*H. pylori*) infection (5). According to some previous studies, the incidence of gastric cardia cancer has continued to increase, which might partly be explained by improved classification (8).

Laurén's histopathological subdivision of gastric carcinomas into intestinal-type and diffuse-type, based on the histological appearance, is commonly used and has been linked to clinical characteristics, pathogenesis pathways and prognosis (9). The pathogenesis of intestinal-type gastric cancer is described as Correa's cascade and is well established (10). It is probably promoted by *H. pylori* infection and progresses from chronic superficial gastritis to atrophic gastritis, then intestinal metaplasia, dysplasia and ultimately gastric cancer. However, diffuse-type gastric cancer has not been associated with a similar pathological pathway as intestinal-type gastric cancer (11).

The most recent WHO definition of gastric adenocarcinomas (published in the fifth edition of the WHO Classification of Digestive System Tumors) uses yet another histological classification based on five main groups: papillary, tubular, mucinous, poorly cohesive and mixed.

Although the previous histopathological and anatomical classifications have been useful for prognosis assessment, new molecular definitions are gaining popularity to predict treatment response. Based on gene expression profiling, the Cancer Genome Atlas Research Group (TCGA) (12) and the Asian Cancer Research Group (ACRG) (13) have developed new molecular classification systems for gastric adenocarcinoma. By performing several molecular analyses (whole exome sequencing, somatic copy number analysis, DNA methylation profiling, messenger and microRNA sequencing and protein analysis) of chemotherapy-naïve gastric cancer samples, the TCGA system has identified four subtypes: EBV-positive (EBV),

microsatellite unstable (MSI), chromosomally unstable (CIN) and genomically stable (GS) (12). The subtypes identified in the TCGA system are gaining popularity in research and have been validated in subsequent patient cohorts (14, 15), but are not yet utilized as much in the clinical setting (16, 17).

The only exception to this is the assessment of an overexpression of the *HER2* gene which is currently the only routine diagnostic test used for both esophageal and gastric adenocarcinoma patients (16). This is mainly due to the significant survival benefit that was demonstrated in the ToGA trial in 2010 (18) for advanced gastroesophageal adenocarcinoma patients who received add-on treatment with trastuzumab (a *HER2* antibody) in the cisplatin-fluoropyrimidine-trastuzumab arm, and in particular within a subgroup of *HER2*-positive patients.

Molecular markers as a basis for gastroesophageal cancer classification therefore may have a brighter future potential than classical anatomical definitions to personalize treatments and improve the prognosis within subgroups of gastroesophageal cancer patients (19).

### **2.1.2 Geographical distribution**

The two subtypes of esophageal cancer display remarkably different incidence patterns and geographical distribution. ESCC is predominant in less developed countries, especially in the South-Eastern, Central Asian regions and China where over 79% of all ESCC cases occur. EAC however has shown rapidly increasing incidence in Northern and Western Europe, Northern America and Oceania where 46% of the global EAC cases occur (20). A dramatic increase in the incidence of EAC was noted in eight registries from Australia, North America, Europe and Asia with an estimated start sometime between 1960 and 1990, where calendar period rather than birth cohort was the more important determinant of incidence trend (21). Some countries such as the UK, the Netherlands, Iceland, Norway and Sweden currently have higher incidence rates of EAC than ESCC.

Gastric cancer is most common in Eastern Asia, Eastern Europe and South America, where the estimated age-standardized incidence rates in 2018 were highest with 12.7-32.1 in men and 6.9-13.2 in women per 100 000 person-years (7). About half of the total gastric cancer cases were located in Eastern Asia, and the age-standardized incidence rate was twice as high for males compared to females (especially in the Republic of Korea, Mongolia, Japan and China) (7, 11). Ethnicity has been associated with the incidence of gastric cancer as well. In the US, a low-risk country for gastric cancer with an overall age-adjusted incidence rate between 2012 and 2016 of 7.4/100 000, the incidence in Caucasians was 6.6/100 000 compared to almost the doubled rate of 10.3/100 000 in individuals with Afro-American origin (22). Individuals who have migrated from areas with high incidence of stomach cancer to low-risk areas display similar risk of gastric cancer as the country of origin, while subsequent generations have shown decreased incidence rates, demonstrating complex host-environmental interactions in the development of gastric cancer (11).

The most recent edition of the Textbook on Cancer Epidemiology from 2018 (23) states that the rather dramatic geographical variations in esophageal cancer incidence together with a low



explanatory proportion of genetic mutations may imply that external factors, in part enhanced by genetic susceptibility, play an important role in the pathogenesis of esophageal cancer. This description is applicable to gastric cancer as well. A recent analysis of the global burden of the five major gastrointestinal cancers from the GLOBOCAN 2018 data suggests that life-style choices are the driving factors behind the past and future burden of GI cancers (24).

### **2.1.3 Age distribution**

The incidence of gastroesophageal cancers increases with age and mainly affects patients 60-80 years of age (25, 26). Only a minority of patients are diagnosed with gastroesophageal cancers under the age of 40 (22, 27). The incidence of esophageal adenocarcinoma (28) and gastric non-cardia adenocarcinoma has been increasing in younger men (29, 30) which might indicate other etiological factors than among older patients.

### **2.1.4 Sex differences**

There is a marked male predominance in incidence of both ESCC (2-4 fold more common in males) and EAC (8-fold more common in males), but the molecular mechanisms for this difference are not yet clear (31).

Overall, there is a male predominance in incidence of gastric cancer, but limited data indicate this does not hold true for all age groups and subtypes of gastric cancer (29, 32, 33). The mechanism behind these differences is unknown, but previous studies have suggested female hormones such as estrogen are protective against the risk of gastric cancer (34). Unlike other sex-hormone dependent cancers such as breast cancer and prostate cancer, the role of sex hormone receptors has not yet been established in gastric cancer (32).

Like for gastric cancer, estrogen has also been suggested as the factor protecting women and explaining the male predominance of esophageal cancer, particularly esophageal adenocarcinoma. There is, however, limited data showing a steady decline with age in the sex ratio for esophageal adenocarcinoma, which may indicate factors other than estrogen behind the male predominance of esophageal adenocarcinoma (34).

### **2.1.5 Environmental risk factors**

#### *2.1.5.1 Dietary factors*

Dietary factors have been hypothesized to play a major role in high-risk areas for ESCC. However, the highly-anticipated chemoprevention trial investigating supplementation of vitamins and minerals for risk reduction of ESCC has shown null effects (35).

A positive association between salt intake and gastric cancer was first reported in an observational study from Japan in 1959. Experimental studies support the link between high salt intake and gastric cancer development in interaction with *H. pylori* infection but not alone (36).

High intake of fruit and vegetables has been reported to reduce the risk of gastric cancer. A meta-analysis reported that the pooled relative risk (RR) for gastric cancer associated with consumption of fruits was 0.82 (95% CI 0.73-0.93), and even lower when the follow-up period was longer than 10 years (RR=0.66, 95% CI 0.52–0.83); for vegetables it was 0.88 (95% CI 0.69–1.13) and 0.71 (95% CI 0.53-0.94), respectively (37).

Whether high intake of red and processed meat is a risk factor for gastric cancer still remains inconclusive. A recent systematic review and meta-analysis reported a positive association in case-control studies for red meat and processed meat, where the summary relative risk of highest versus lowest consumption was 1.67 (95% CI 1.36-2.05), and 1.76 (95% CI 1.51-2.05), respectively. Cohort studies, however, reported no associations with gastric cancer (for red meat: RR=1.14 (95% CI 0.97-1.34); for processed meat: RR=1.23 (95% CI 0.98-1.55). Subtype analysis yielded similarly null results for both cardia cancer and non-cardia gastric cancer (38).

#### 2.1.5.2 *Helicobacter pylori*

In normal stomachs, or those affected only by superficial gastritis, the high acidic environment prohibits colonization of bacteria other than *H. pylori*. However, with the development of atrophic gastritis, the hypo- or achlorhydric stomach provides an environment that facilitates bacterial overgrowth. Some species in the microbiota may reduce nitrate to nitrite by nitrate reductase or produce enzymes capable of catalyzing N-nitrosation (39). We hypothesized that enhanced endogenous production of nitrosamines, which are transported to the esophagus by some still unknown mechanisms, might be the underlying mechanism for the observed association between gastric atrophy and ESCC. N-nitroso compounds have been proposed to play a role in the pathogenesis of upper gastrointestinal cancer, and show organ-specific effects in animal models.

To support this hypothesis, in a previous study we found that severe corpus atrophic gastritis was a strong risk factor for ESCC (40). This novel finding was later verified in studies with endoscopy data conducted in high-risk populations in Japan (41, 42), and in Brazil (43). One of these studies also reported an increasing risk for ESCC with increasing severity of atrophy; the multivariate-adjusted odds ratio with 95% confidence interval for histologic fundic atrophy was 4.2 (1.5-11.7), and 10.7 (2.3-50.4) for fundic intestinal metaplasia (42). Another study from Japan also reported an increased risk of ESCC in patients with profound hypochlorhydria (measured by endoscopic gastrin test) with an odds ratio of 6.0 (1.9-18.4) (44). In line with this, another small Japanese study (n=28) reported higher prevalence of hypochlorhydria (measured by 24-h pH-Impedance Monitoring) in the ESCC group, as well as a significantly higher number of total reflux and non-acid reflux episodes in the ESCC group versus non-ESCC group (45). A cohort study in the Netherlands with data from a histopathological register also reported an increased risk of ESCC in patients with gastric atrophy; however, they could not confirm that the excess risks increased with increasing severity of gastric atrophy, and concluded that confounding factors such as smoking could explain the observed association (46). Further, two other studies, one from Germany (47) – a low risk area for ESCC (measuring fundic gastric atrophy by serology and histology) – and the other from Linxian in China (48) –

a high-risk area for ESCC (measuring gastric atrophy by serology) – did not find a statistically significant association. Overall, the results regarding the association between ESCC and gastric atrophy are inconsistent. Variations in findings might be explained by differences in study design, reliability of the method for gastric atrophy measurement, whether atrophy was assessed before, during or after cancer diagnosis, cancer stage, appropriate control of potential sources of biases and confounding, etc. Further studies are therefore needed to explore the potential role of gastric atrophy in ESCC.

Moreover, poor oral health with an altered bacterial flora in the oral cavity might further contribute to the ESCC risk by adding more carcinogens to the esophagus. Poor oral health has been reported as an independent risk factor for ESCC in high-risk areas (49). This was recently confirmed by a study from our group in a high-risk population in China (50). In a study performed by our group in Iran, another high-risk area for ESCC, a possible interaction between gastric atrophy and poor oral health was reported, though the study was underpowered (51).

*H. pylori* infection seems to be inversely associated with EAC (52).

*H. pylori* was classified as a class I carcinogen for humans in 1994 (53). It affects about 50% of the global population (11), causing gastritis in most infected individuals, though many infected people are asymptomatic (53). Depending on complex interactions between environmental factors, bacterial virulence factors and host factors some *H. pylori* infections eventually lead to gastric cancer development (11).

*H. pylori* eradication has successfully reduced the risk of progression in individuals with lesions such as gastritis and gastric atrophy; meanwhile, results regarding the influence on progression of intestinal metaplasia have been inconsistent (11). Pooled data from placebo-controlled *H. pylori* eradication trials in Asia have shown moderate-quality evidence that *H. pylori* eradication in healthy asymptomatic patients has a protective effect against gastric cancer with a relative risk of 0.66 (95% CI 0.46-0.95) compared to no treatment (54).

#### *2.1.5.3 Gastroesophageal reflux and obesity*

For EAC and gastric cardia adenocarcinoma, gastroesophageal reflux and obesity have been established as risk factors (26, 55).

#### *2.1.5.4 Smoking and alcohol*

Risk factors for ESCC suggested in previous studies include tobacco smoking, alcohol consumption and history of head and neck cancer. The risk estimates for ESCC associated with smoking and alcohol consumption vary many-fold between different populations and is believed to contribute more to the excess risk in the economically developed countries than the economically developing. Interaction between heavy smoking and heavy alcohol drinking has been reported in many studies, but not all cohort studies have been able to verify this effect modification which might be due to selection bias (49).

There is sufficient evidence that smoking causes stomach cancer according to the IARC monographs in 2004 on tobacco smoking (56), and further confirmed in 2012 (57). The conclusion is based on numerous cohort and case-control studies summarized in several systematic reviews and meta-analyses (58-61). However, conflicting results were reported when subdividing gastric cancer by histological subtype or anatomical site, which might be due to misclassification of gastric cancer subtypes, scarcity of studies with focus on subtypes of gastric cancer and publication bias. The study selection in the different meta-analyses is quite different despite overlapping time-periods, which might also explain some of the heterogeneity.

#### *2.1.5.5 Socioeconomic status*

Low socioeconomic status (SES) with surrogate markers such as low education, low income, higher number of siblings, overcrowding and lower occupational activity, is an established risk factor for gastric cancer, especially non-cardia cancer. Low SES however correlates with possible biological exposures for gastric cancer such as higher possibility of *H. pylori* infection, lower consumption of fruits and vegetables, smoking and physical inactivity. These factors are particularly likely to explain the observed association between low SES and gastric cancer (62). Indeed, in a large European multicenter study, adjustment for *H. pylori* infection rendered a null association between SES and gastric cancer (63).

#### *2.1.5.6 Medication use*

Use of aspirin and Non-Steroidal Anti-Inflammatory Drugs (NSAIDs) have been associated with a risk reduction of esophageal and gastric cancer (64, 65) while proton pump-inhibitors have been hypothesized to increase the risk (66). When dealing with medication use, confounding by indication cannot be ruled out (67, 68).

### **2.1.6 Familial aggregation and genetic risk factors**

Familial aggregation has been reported to occur in roughly 10% of gastric cancer cases (69), wherein first-degree relatives of patients with gastric cancer have a 2-3 fold increased risk of developing gastric cancer compared to the general population (70, 71). Previous studies have also reported an increased prevalence of precancerous lesions such as atrophic gastritis (72), intestinal metaplasia (73) and dysplasia (74) in first-degree relatives of gastric cancer patients. However, to date, no previous study has reported the risk of gastric cancer in first-degree relatives of patients with precancerous lesions such as atrophic gastritis, intestinal metaplasia and dysplasia.

The mechanisms for familial aggregation of gastric cancer are not clear. Genetic susceptibility, shared environmental exposures and common lifestyle habits such as cigarette smoking, diet, bacterial virulence and gene-environment interactions have been suggested to play a role (75).

In line with this reasoning, there is growing evidence in support of non-cardia gastric cancer as a complex multi-cause disease where the development requires a long-term interplay between the host and the environment (including *H. pylori* related factors such as virulence and host genetics, as well as life-style factors) (76) already starting in childhood (77). This way of

understanding the disease could explain the higher elevated risk seen among siblings of gastric cancer patients (78). The complex interplay between *H. pylori* and host responses is somewhat illustrated by the identified host-related genetic polymorphisms of cytokines involved in the inflammatory response to *H. pylori*, which have been associated with both gastric pre-malignant lesions and gastric cancer (76). In addition, *H. pylori* genetic polymorphisms, primarily related to the virulence, have also been linked to gastric non-cardia cancer (79).

A minority (about 1-3%) of incident gastric cancer cases are considered to be hereditary (69). These are divided into three major syndromes: hereditary diffuse gastric cancer (HDGC), gastric adenocarcinoma and proximal polyposis of the stomach, and familial intestinal gastric cancer (69). HDGC is associated with mutations in the gene *CDH1* (encoding E-cadherin) which is a tumor suppressor gene, and the gene *CTNNA1* (encoding alpha-E-catenin) (69). Besides these three major syndromes, gastric cancer is also prevalent in other cancer-associated hereditary syndromes such as Lynch syndrome, Peutz-Jeghers syndrome and familial adenomatous polyposis (FAP) (69).

### **2.1.7 Esophageal precursor lesions**

EAC incidence is increasing and it is important to better understand the pathogenic mechanisms to find high-risk groups for targeted primary prevention. Barrett's esophagus (BE) is a condition where metaplastic columnar epithelium replaces the normal stratified squamous epithelium of the esophagus with gastroesophageal reflux as the main risk factor (80). Previous studies have estimated the risk of developing EAC in BE patients to be between 10 to 55 times higher than the general population (81). BE is the only known precursor lesion to EAC to date, but 95% of EAC patients did not have a previous diagnosis of BE in a previous study (82). BE incidence has been reported to increase in several studies from developed countries, but it is not clear if this is the main reason behind the increasing EAC incidence. BE patients are under surveillance by endoscopy in several countries, but the cost-effectiveness of such surveillance is undermined since most BE patients die from other causes than EAC (83). A randomized controlled trial is ongoing among BE patients in the UK to find out how endoscopic surveillance versus no endoscopic surveillance (endoscopy "at need") affects early mortality and malignant transformation to esophageal adenocarcinoma (84). The final study report after study completion in 2022 hopefully will shed light on this issue.

Squamous dysplasia is the established premalignant lesion for ESCC (85-87). However, previous studies were conducted mainly in high-risk areas, and studies from low-risk areas, such as Sweden, are still scarce.

### **2.1.8 Treatment of gastroesophageal cancers**

#### *2.1.8.1 Treatment guidelines*

Historically, the choice of treatment mainly depended on the localization and stage of the tumor. Histology has started to play a more significant role only recently. The current treatment guidelines for gastroesophageal cancers in Sweden have been established in collaboration

between Regional Cancer Centers. According to the current guidelines, early esophageal and gastric cancer (T1aN0M0) and intraepithelial neoplasia are treated with endoscopic mucosal resection with results comparable with traditional surgical resection.

Cervical or proximal tumors of the esophagus need certain considerations regarding surgery when it comes to saving the larynx, and curative radiochemotherapy without surgery is also an option. Middle and lower esophageal cancers and gastric cancer are treated in a similar fashion with neo-adjuvant chemotherapy, with or without radiotherapy, and surgical resection.

Current standard chemotherapy for ESCC is paclitaxel and carboplatin according to the CROSS study. For esophageal and gastric adenocarcinoma, the FLOT-regimen (5-Fu + kalciumfolinat + oxaliplatin + docetaxel) from the FLOT4 trial in 2016 is currently the standard chemotherapy.

Palliative treatment for advanced gastroesophageal cancer with distant metastases follows the same guidelines with possible surgical treatment, chemo- or radiotherapy in order to alleviate symptoms (25).

Although the efficacy is known for the chemotherapy regimens from previous clinical trials, head-to-head comparisons of commonly used perioperative chemotherapy from the post-marketing clinical setting are scarce. To our knowledge, a real-world comparison of neo-adjuvant chemotherapy in gastroesophageal cancer patients, separating patients who received surgery with curative intention from those with palliative treatment intention, has not been reported previously in Sweden.

#### *2.1.8.2 Studying oncology treatments in a real-world setting*

A “real-world setting” is, in our opinion, defined as an observational setting where data from existing health care registers and administrative registers are retrieved to study the effects of various exposures on an unselected population, in contrast to a clinical trial. However, several other definitions exist for real-world data (88).

Evaluations of the kind we performed in study IV are restricted in Sweden due to the limited access to inpatient chemotherapy data. Our study was only made possible by the temporary register “SALT” that contains inpatient chemotherapy data from 2008 to 2014. To gain access to comprehensive inpatient chemotherapy data before or after 2014 one needs to extract the drug information in the medical records systems. Before 2008 there were several electronic medical systems in the same region, but in 2014 most health-care providers in Stockholm-Gotland changed to TakeCare, following a regional decision to work in the same electronic medical records system. Studying chemotherapy regimens based on TakeCare would most likely require manual collection of chemotherapy data for each case. It is not unlikely that our situation with limited access to high-quality inpatient chemotherapy data is the same in many other countries, which probably can explain why this kind of evaluations are very rare.

Evaluations of this kind might however be more feasible in Sweden in the future. The quality register for gastroesophageal cancer (NREV) collects data about chemotherapy and radiotherapy given to this patient group. The completeness of these variables is quite poor for the period 2008-2014, but will hopefully improve from this time forth.

Performing a clinical trial is the gold-standard method to establish the benefit-risk ratio of new chemotherapy regimens in gastroesophageal cancer patients. Studying how these results translate to “real-world” settings is however pivotal for decision makers, care-givers and patients. The real world setting differs in many aspects from a clinical trial. Clinical trials typically do not include all ethnicities in the population, the frail and elderly, patients with all tumor stages, the palliative patients, those with co-morbidities, polypharmacy and so on. These patients, however, are to be treated in the real world and what the benefit-risk ratio is among them is therefore often unknown prior to treatment. In our study IV we included all patients with incident cancer except those who were not planned to have any tumor-specific treatment at diagnosis, since we wanted to study the effect of treatment in a real-world setting.

#### *2.1.8.3 Precision medicine*

The exact definition of precision medicine varies but usually entails individually tailored prevention and/or treatment taking into account individual variations in tumor molecular characteristics, and host genetic or other factors. The research community has very high hopes for precision medicine in gastroesophageal cancers, since the classical definitions and treatments have not been overwhelmingly successful yet. Although new molecular/genetic definitions have been introduced and could be used for new drug targets, very little of the research findings have made it into clinical practice due to various hindrances in data management, tumor- and patient-related factors (89-92).

### **3 AIMS**

Although some common epidemiological risk factors have been established, the causes of gastroesophageal cancers and the treatment factors determining the prognosis are not fully understood. Therefore, the overall aim of this thesis is **to elucidate biological factors associated with an elevated risk for gastroesophageal cancers, and evaluate treatments influencing the survival of gastroesophageal cancer patients in a real-world setting.**

More specifically, the aims are as follows:

- Estimate the risk for gastric cancer among first-degree relatives of patients with gastric premalignant lesions.
- Assess the gastroesophageal cancer risk associated with esophageal lesions.
- Validate the elevated risk for ESCC associated with gastric atrophy and explore its interaction with poor oral health.
- Compare the effectiveness on the survival of gastroesophageal cancer patients depending on the chemotherapy regimen they received.

These over-arching aims were addressed in the papers included in this thesis by the following research questions:

#### Paper I

- What is the risk for gastric cancer among parents, siblings and children to patients with gastric mucosal lesions based on a histopathology register study in Sweden?

#### Paper II

- What is the risk for esophageal adenocarcinoma associated with intestinal metaplasia in a population-based histopathology register study in Sweden?
- What is the risk for gastroesophageal cancers associated with other histopathological esophageal lesions in Sweden?

#### Paper III

- What is the risk for esophageal squamous cell carcinoma associated with gastric atrophy, measured by serum pepsinogen levels, in a case-control study in a high-risk region in eastern China? And does the interaction with poor oral health further increase the risk?

#### Paper IV

- How did the various chemotherapy regimens used in Stockholm-Gotland during 2008 to 2013 influence the survival of gastroesophageal cancer patients?



## 4 MATERIALS AND METHODS

### 4.1 RESEARCH APPROACH

The research approach in this thesis has been to address new questions about the pathways to gastroesophageal cancer and their prognosis by using tools from classical cancer epidemiology and biostatistics. Where it was suitable for the research question, we used existing health care registers in Sweden, sometimes complemented with our own data collection. The use of existing resources has the advantage of being both cost-effective and ethically sound. Furthermore, the national health care registers have nearly complete coverage and have been validated for research purposes. However, there are important limitations with administrative health care registers when used for research purposes. One limitation is the lack of reliable information on possible confounding factors, including smoking, alcohol consumption, BMI, diet, or biological confirmation such as histology. Furthermore, biochemical, molecular or genetic samples are seldom available. Another limitation is non-random missingness, such as TNM-stage. Selection bias is a potential issue, wherein the data from the register may not be representative of the source population of interest (for instance if studying alcoholism from the Patient Register or use of NSAIDs from the Prescribed Drug Register). Moreover, systematic error cannot be addressed easily in register-based studies since the researcher has no control and perhaps even limited insight with regards to the data collection process. Given these shortcomings, register-based research can be very useful if handled with caution and with close collaboration with those generating the data to be able to account for the limitations and sources of error that may be present.

The research questions in the first two studies were addressed using data from the Swedish histopathology registers linked through the Personal Identity Number (PIN) to several national health care registers and other administrative registers to obtain individual level data on outcomes and potential confounding factors. In the third study, we designed and performed a large scale case-control study in a high-risk area in China to validate the results from a previous study in Sweden and Iran. In the last study, we used a similar register-based approach as the first two studies, with the modification that this study was based on information about individual chemotherapy treatments delivered to hospital patients.

Below is a description of the key data sources and statistical methods employed, as well as the rationale for choosing these data and methods.

### 4.2 DATA SOURCES

#### 4.2.1 Registers

##### *4.2.1.1 Histopathology registers*

The pathology departments in Sweden receive pathology and cytology samples on a daily basis for health care purposes such as diagnosis and treatment. Residual paraffin-embedded specimens from this process are stored for future use in connection with data of care or

treatment for that patient. The pathology departments thereby maintain the largest tissue biobanks in Sweden. Each department has a separate data register using one of two pathology records (SymPathy or Safir) to keep track of the specimens in their biobank. The information stored is similar between the departments. The registers contain the personal identity number (PIN), date of the sample arrival, referral type and number, and the diagnosis based on the second edition of the SNOMED (Systemized Nomenclature of Medicine) system from the College of American Pathologists 1979. The first computerized records started in the 1970s, but large-scale computerization of records was seen close to the 1980s. By the late 1990s, all pathology departments had computerized their records. Besides their clinical use, the histopathology registers give unique possibilities for researchers to study certain clinical research questions that may not be feasible otherwise. The availability of these register data have allowed for the examination of, for example, the morphological, molecular and genetic alterations before and after disease diagnosis, as well as how different subgroups differ clinically, such as men versus women or younger versus older individuals. Biomedical researchers who want to use these biobank registers and/or specimens can apply to each department separately and ask for permission to do so. Applications are assessed by an independent group with medical professionals at each pathology department regarding ethical aspects, feasibility, patient benefit, how to cover costs and allocate resources. It is up to each pathology department to decide if and how they want to participate in research projects. The decision is always weighed against their top priority to maintain timely services to referring health care providers.

Although most departments are very helpful and willing to contribute to research, the process of doing multi-center research with all the pathology departments in Sweden involves long lead-times and is resource-consuming for both the pathology department and the researcher. This has restricted the use of the histopathology registers in scientific research and they are therefore not validated as thoroughly as compared to many other national health registers in Sweden.

A strength of the histopathology register is the use of the same coding system; the SNOMED II system across the whole country since the 1970s, enabling nationwide studies over many decades. However, the SNOMED II system is not maintained centrally by the SNOMED organization since it was replaced by newer SNOMED systems such as SNOMED III and later the SNOMED CT system. Consequently, a limitation with the SNOMED II system is that there is no centralized update of the system. Further, the coding system is no longer entirely uniform between counties and departments. Although efforts are being made by the Swedish Association for Pathology to spread standard diagnosis criteria and codes across the country, local variations between the different pathology departments exist.

#### *4.2.1.2 Population-based health registers and population registers*

Besides the National Board of Health and Welfare (Swedish: Socialstyrelsen) and Statistics Sweden (Swedish: Statistiska Centralbyrån) which are the two main authorities that own and maintain many of the population-based health and demographic registers in Sweden, there is a

large number of National Quality Registers that are supported by the state and the Swedish Association of Local Authorities and Regions, but maintained by various care-givers. Currently, just over one hundred National Quality Registers are active. A brief overview of the population-based health registers used in this thesis is given below.

#### 4.2.1.2.1 The Swedish Cancer Register

The Swedish Cancer Register was founded in 1958 to follow the incidence and changes over time for cancer diseases. All health care providers are obliged by law to report cancer cases among registered individuals in Sweden to the authority. The register is therefore nationwide and in general, more than 99% of all cancer cases are reported, albeit with some delay. Since the 1980s, the reports from health care providers first go to one of six regional cancer centers who code, register and control the quality before reporting their statistics annually to the Swedish Cancer Register. The most recent study of the completeness of the Swedish Cancer Register reported a 90% coverage of digestive organ cancers compared to the hospital discharge records in 1998 (93). Based on the hospital journals, the study found that underreporting was not random. Instead, underreporting was more likely to occur if there was a missing pathology verification, if the patient was of older age, and if health care providers were not part of University hospitals. Still, the authors concluded that the completeness of the Swedish Cancer Register was comparable to other high-quality Northern European cancer registers and that the underreporting should not have major influence on most epidemiological studies (93).

In the Swedish Cancer Register, one can find detailed information, including the personal identity number, the date of cancer diagnosis, sex and age at diagnosis, the location of the cancer (according to the International Classification of Diseases (ICD) systems from 1958 and onwards), morphology (according to SNOMED II), histopathological diagnosis (PAD) and TNM-status (and if it is clinical or pathological staging) of the tumor at diagnosis. The TNM-status registration began in 2004 and the completeness is unfortunately poor. Moreover, high-grade dysplasia or cancer in situ and pathological grade of the tumor are generally not registered.

#### 4.2.1.2.2 The Cause of Death Register

The Cause of Death Register in Sweden holds electronic records from 1952 to 2012 of deaths and underlying causes of death among those registered in Sweden (94). From 2012 and onwards it includes all deaths, no matter if the individual was registered in Sweden or not (94). The completeness of death reports is considered full since a death certificate is mandatory for a burial to take place (94). Until 1991 a report of the cause of death was also mandated by law for a burial (94). This requirement was abolished in 1991, which meant from 1991 until 1997, only those reports with a cause of death were included in the register (94). After 1997, all death reports, even those with a missing report of the cause of death were included in the register (94). The cause of death is coded according to the international ICD system, unlike the rest of the health care system that uses the Swedish version of the ICD coding system (94).

The quality of the cause of death reports is influenced by the reporting physician's knowledge of the patient history and understanding of the logic underlying the report. The means and aids in the reporting system and the measures undertaken by responsible authorities to control the quality and correctness of the register are also important. The most recent assessment of the register among a random sample of 1,094 deaths, reported a 77% match between cause of death reports and the cause expected from hospital discharge records (95). For malignant neoplasm cases, there was a 90% match, while COPD and other pulmonary diseases had the lowest match of 47% (95).

Overall, considering the high degree of completeness and high quality of the reports of underlying cause of death, the temporal changes, the heterogeneity between practitioners and loss of underlying cause of death reports should not have a major impact on the quality of malignant cases in the Cause of Death Register.

#### 4.2.1.2.3 The National Patient Register

The National Patient Register consists of the National Inpatient Register (also called the Hospital Discharge Register) and the Outpatient Register. The National Inpatient Register was initiated in 1964 but full-scale registrations, i.e., reaching nationwide coverage, were not achieved until 1987. From 2001 and onwards, specialized outpatient visits are included in the National Patient Register. According to a review of validation studies concerning the National Inpatient Register, the positive predictive value (PPV) was found to be 85-95% for most evaluated diagnoses (96). The sensitivity for COPD was in the lower range: 27% in one study (96).

The National Patient Register does not seem to have been validated for the diagnosis of alcoholism, but it is likely that the PPV is low considering that the validity for alcoholism in the Cause of Death Register was low in a previous study (97). We could not find a validation of esophagectomy, gastrectomy or anti-reflux surgery in the National Patient Register, but it could be similar to other surgical procedures such as amputations, appendectomies and surgery of inguinal hernia with a PPV of over 90% in previous validation studies (96).

A major limitation of the National Patient Register is that primary health care patients and patients not treated by physicians are not included in the register. As a consequence, certain diagnoses not requiring specialized health care or certain procedures such as endoscopy may be incomplete.

#### 4.2.1.2.4 The Prescribed Drug Register

The Prescribed Drug Register began in July 2005 and includes all prescribed drugs dispensed at pharmacies in Sweden since then. It is updated monthly. Since it is an automated administrative register regulated by law, the reporting is nationwide and the loss of data is very low. However, there are many challenges to be aware of when using this register. The first thing to keep in mind is that drugs administered within the health care system, such as ambulatory care, inpatient and outpatient care, day-care, elderly care (except ApoDos which is

delivered by pharmacies) and so on are not registered here. Since there are regional health care system differences, some drugs may be provided by the care-giver in some regions and dispensed at the patient's own expense in other regions. Furthermore, over-the-counter drugs and drugs prescribed but never collected at a pharmacy, do not enter the register. Some drugs, such as pain-killers or anti-reflux medicines can both be prescribed and bought over-the-counter.

In an attempt to standardize the definition of drug dose, the World Health Organization (WHO) provides the Defined Daily Dose (DDD) system. Certain drugs such as chemotherapy do not have a DDD.

Another major limitation to be aware of is that the dispensed drugs are not the same as the drugs taken by the patient. A previous validation study of asthmatic adolescents (11-14 years of age) found that 30% of patients with parent-reported use of asthma medications did not have dispensed asthma drugs during the preceding 18 months (98). Moreover, even if the drug was taken by the patient, it is no guarantee that the drug dose corresponds to the drug exposure in that individual. For instance, inhalation asthma medications are frequently administered with inadequate technique, resulting in an under-exposure of the medicine.

#### 4.2.1.2.5 The Total Population Register and LISA

The Total Population Register (99) and LISA (100) from Statistics Sweden complement many medical epidemiology studies with demographical data on sex, date of birth, country of birth, immigration and emigration as well as socioeconomic information such as civil status, education, income, occupation, use of social benefits, etc. The demographical data is nearly complete, while socioeconomic variables are less complete. Older patients generally have more missing data on education, occupation and use of social benefits typically due to retirement. Other limitations are the inadequate registration of educational level of immigrants, and that co-habitation—which is more common than marriage—is not included in the variable civil status. For migration data to be registered, the individual has to report this to the Swedish tax authority (Skatteverket). Some, especially those with shorter periods of emigration, may not register immigration/emigration status.

#### 4.2.1.2.6 The Multi-Generation Register

The Multi-Generation Register is a part of the Total Population Register system and includes individuals who were registered in Sweden any time from 1961 and onwards and who were born 1932 or later. These individuals constitute the “index persons”. For every index person there is a link to the biological or adoptive parents. The register is updated annually with new “index persons” that have immigrated or been born. For more information please read the information available in the report about the Multi-Generation Register from Statistics Sweden in 2016 ([www.scb.se](http://www.scb.se)).

#### 4.2.1.2.7 The National Quality Register for Esophageal and Gastric Cancer (NREV)

This National Quality Register was founded in 2006 and is maintained by one of the Regional Cancer Centers. The primary aim of the register is to provide comparisons of indicators of quality of care and outcomes in Sweden to enable improvements. The register is however increasingly used for research purposes as well. In 2018, the completeness of reported esophageal cancer patients and gastric cancer patients was 95% and 91%, respectively, compared to the Swedish Cancer Register. The register comprises records of incident esophageal and gastric cancer patients, with information on high-grade dysplasia/cancer in situ. Patients diagnosed and examined outside the country, but treated in Sweden, are also included in the register. NREV initially only collected data on surgical procedures. Since 2008, the register also began collecting data on oncological treatments (NREV Annual report 2018 accessed from [www.cancercentrum.se](http://www.cancercentrum.se)). The mean lead-time from diagnosis to treatment (oncological or surgical) was 46 days for esophageal cancer and 45 days for gastric cancer in 2017-2018. At diagnosis, the treatment intention—which can be either curative, palliative or no tumor-specific treatment—is recorded. Health care providers generally aim for 40% of all gastroesophageal cancer patients to receive curative treatment, but what this treatment consists of varies between the regions due to currently unknown reasons. In Stockholm during 2017 to 2018, for patients who had a curative treatment intention at diagnosis, the distribution of the planned therapeutic strategy was as follows: about half of the patients were to receive a combination of curative resection and oncological treatment, about a third were to receive only curative resection, and the rest were to receive only oncological treatment. The resectability, i.e. the proportion of patients planned for resection that actually received the treatment, has varied between 80-90% for the whole country since the start of the register. Palliative treatment or no tumor-specific treatment intention at diagnosis was recorded for 37% of the esophageal cancer patients and 32% of gastric cancer patients; the majority in this patient group received only oncological treatment.

The overall 5-year survival rate recorded in NREV is surprisingly similar for esophageal and gastric cancer patients. In the most recent annual report from NREV the overall 5-year survival rate was 15%. Among those with curative resection the 5 year-survival rate was 40%. The survival is strongly dependent on the tumor stage at diagnosis. Furthermore, the mode of palliative treatment was also associated with different survival rates.

NREV is a powerful register to follow, compare and improve the health care practices for gastroesophageal cancer patients and provides many research opportunities in this context. The register is very extensive and many variables of interest for epidemiological studies can be found in this register. Many of the treatment-related variables for gastroesophageal cancer patients only exist in this register besides the medical charts, or have better completeness than other registers, such as TNM-stage. On the other hand, the completeness of some variables is not yet high enough for research purposes, such as smoking and the choice of oncological treatment. Other limitations with the register is the relatively short duration since the start in 2006 and that several variables have been included only recently, making the follow-up time even shorter. Inclusion of information about precursor lesions from the Correa cascade or

Barrett's esophagus would have been useful, but are probably not feasible considering the way data are currently collected. With that said, NREV is an extraordinary and valuable source for research about gastroesophageal cancers in Sweden.

#### *4.2.1.3 Regional or local registers*

##### 4.2.1.3.1 The electronic medical records system TakeCare

TakeCare is the current electronic medical records system in Stockholm-Gotland and contains all medical charts written in the region, as well as reports on all examinations (blood samples, pathology/cytology samples, x-ray, etc.) and all drugs and medical devices ordered or prescribed in the region. Inpatient and outpatient chemotherapy orders, dose and administration details (registered by the nurse) can be retrieved from this system.

The limitations regarding daily routine work, and research purposes for that part, are unfortunately exhausting. The main issues with using TakeCare for research is the structure of the system, redundant and false information, the lack of validation and the extensive manual work that has to be spent to retrieve reliable data and clean it for use in research. Furthermore, some clinics, such as the intensive care unit, use other parallel or add-on software for prescribing and/or documentation.

##### 4.2.1.3.2 The VAL-database

The VAL-database is an administrative individual-level register for the Stockholm County Council that follows all health care delivered to the inhabitants of the region for the purposes of planning and following up the need, quality and efficiency of the health care in the region. The database started with only inpatient care, but has since grown to encompass a number of other administrative registers. Since 1995, it also contains outpatient data including primary health care data. The VAL-database can be used to retrieve additional information about all drug prescriptions in the region and drug requisitions from health care providers. It also contains some information about over-the-counter drugs, though not all pharmacies report this data. The drug information in the VAL-database comes from several different sources, including the service for pharmacies (ASAB), pharmacies and drug providers. The VAL-database does not register health care given by providers without a contract with the Stockholm County Council.

The added value of the VAL-database information on drugs compared to the Prescribed Drug Register, is that VAL also includes both inpatient (hospital) and outpatient (primary health care) drug prescriptions and drug requisitions which might not be registered in the Prescribed Drug Register.

##### 4.2.1.3.3 The SALT register

The SALT register was a nationwide inpatient drug register that was active during 2008 to 2014, and was upheld by Apoteket AB, a state-owned pharmacy. Following the deregulation of the pharmacy market, the SALT register ceased registration of new data. The SALT register

contains information about the chemotherapy dose prepared for inpatient care at the individual level and the details of delivery to the health care provider.

#### **4.2.2 Case-control study**

A case-control study was carried out from 2010 to 2014 by our group in collaboration with colleagues at Fudan University and Shandong University in China for the purpose of studying several suspected risk factors for esophageal squamous cell carcinoma and gastric cancer. The location was chosen due to the high incidence of ESCC, with a raw incidence rate as high as 60/100,000 person-years (101). Given that Taixing, a city in China, has a population of 1.2 million, we estimated an annual total number of 680 eligible incident ESCC cases (95% of total number, after excluding those older than 79 years, and those residing in the area less than 10 years). A previous pilot study indicated that 90% of ESCC cases could be identified prior to treatment, and that 90% of these were willing to join the study.

The rationale for using a case-control study design in this setting is mainly that a case-control study is more efficient than a cohort study in the context of studying an exposure with a long latency period such as gastric atrophy, and a cancer disease with low incidence in the population. The drawback is that the odds ratios estimate derived from the logistic regression analysis in this study is not a direct measure of the relative risk even if it is a close approximation of the relative risk derived from a cohort study, in our context. Furthermore, when designing case-control studies, one needs to carefully consider and plan for how to reduce the effects of random and systematic errors.

To begin with, we tried to reduce the effect of random error by having a large sample size. Previous studies reported a prevalence of 15% for gastric atrophy. Based on different assumptions where the prevalence of gastric atrophy was 10%, 15% and 20% with an alpha level of 5%, we estimated that to detect an odds ratio of 1.5, the recruitment of 1100 cases and 1600 controls would be required to provide the statistical power of 0.92, 0.98 and 0.99, respectively. Furthermore, assuming a 15% prevalence of corpus atrophy and a dichotomized variable of decayed, missing or filled teeth (our main variable for poor oral health), the estimated statistical power to detect a gamma for additive interaction of 2 would be 0.85, under the assumption that  $RR_{01}$  and  $RR_{10}$  are equal to 2.

Cases were enrolled through a rapid case ascertainment system established at the four largest hospitals in the region where patients undergoing endoscopy were invited to participate in the study. We also complemented with cases from the local Cancer Registry. The controls were sex- and age-matched (by 5-years intervals) according to the Total Population Register and invited to participate in the study every twelve months. The collected data underwent continuous quality controls, and dialogue was kept with our collaborators to correct errors during the data collection.

A field station was established and local staff were trained to perform interviews of cases and healthy controls and register the answers in a validated electronic questionnaire. Blood samples were collected and analyzed for Pepsinogen I and II as biomarkers for gastric atrophy, and *H. pylori* serology.



ELISA of Pepsinogen I and II were performed at Qilu Hospital of Shandong University, China with Pepsinogen I and II Kits from Biohit HealthCare (Helsinki, Finland). The inter-assay coefficients of variation were 11.3 % and 14.9 %, respectively. Immunoblotting with IgG antibodies directed against *H. pylori* were quantified using *Helicobacter pylori* IgG Antibody Detection Kit from Syno Gene Digital Technology, Taizhou, China.

Study III in this thesis was a sub study of the above mentioned case-control study and was focused on the risk for esophageal squamous cell carcinoma associated with gastric atrophy and the interaction with poor oral health. The gold standard method for clinical diagnosis of gastric atrophy is endoscopy. It should be noted that gastric atrophy has a patchy distribution, which may result in atrophic sites being missed at endoscopy. Gastric atrophy, however, affects the pepsinogen producing glands and is therefore associated with low serum pepsinogen I level. Furthermore, collecting serum samples is less invasive and the risk of complications is considerably lower than with endoscopy. These concerns, paired with the ethical and practical aspects in a study with such a large sample size, led us to choose serum samples instead of endoscopy as the means of measuring the exposure for gastric atrophy. We used the pepsinogen cut-off values to determine gastric atrophy based on a previous validation study from our group in a high-risk population in northern Iran. It would have been ideal to perform a validation study of the cut-off value for pepsinogen in the Taixing population as well, but unfortunately this was not within the scope of this study. Other than this limitation, we made considerable efforts to decrease the risk for systematic errors such as bias, confounding and misclassification of the outcome and exposures which are described in paper III.

## **4.3 STATISTICAL METHODS**

### **4.3.1 Measurements of risk, relative risk, and underlying assumptions**

#### *4.3.1.1 Incidence rate and cumulative risk*

Crude incidence rates were calculated by dividing the number of observed cases with the accumulated person-years for our different exposure groups. We used this disease frequency estimate in paper I and II for our exposure groups that consisted of the first record of a gastric or esophageal biopsy taken at endoscopy and diagnosed with histopathology. Individuals in each exposure group could have several findings at the same endoscopic examination. We therefore chose to use the most severe finding to categorize each patient. The rationale for this categorization was to focus on the prognostic value of the first biopsy which we believe is mostly affected by the most severe finding, though other cut-off points as specified on a statistical, clinical or operational basis could also have been an alternative (102). We also chose to only count person-time until the first diagnosis of gastroesophageal cancer. Since a majority of gastroesophageal cancers are diagnosed in an advanced stage, many patients are likely to have lived with their cancer for some time before the diagnosis. The incidence rates for cancer in this study thus reflect the time until diagnosis, and not the actual occurrence of disease. As these cancers progress rapidly, the time lapse between actual occurrence of the disease and the subsequent diagnosis should not be a major issue.

Cumulative incidence rate was calculated using the Nelson-Aalen estimator in paper I. Since the Nelson-Aalen estimator is non-parametric, it can be used to calculate the cumulative incidence rate function of right-censored and left truncated survival data as in our study. The estimator can also be used in several other situations such as counting relative mortality, cumulative infection rate or Markovian multi-state models (103). On the other hand, the Kaplan-Meier estimator can only be used for survival data. We used the Kaplan-Meier estimator in paper IV to illustrate cumulative risk by different exposures groups.

#### *4.3.1.2 Standardized incidence ratio (SIR)*

For the purpose of calculating standardized incidence ratios (SIRs) in papers I and II we retrieved incidence rates for gastroesophageal cancers stratified by sex and age group (5-year intervals) for the general population from the Swedish Cancer Register. The Swedish Cancer Register encompasses all primary tumors (if the primary tumor is unknown, metastases can be reported). All in all, the Swedish Cancer Register contains records for about 3 million tumors from 2.4 million individuals residing in Sweden during the period of 1958-2016.

Standardized incidence ratios (SIRs) were calculated by dividing the observed number of cancer cases in each exposure group by the expected number. The expected number was derived by multiplying the incidence rate for the general population with the sum of person-years for each exposure group. Since the incidence rate for the general population is including repeated gastroesophageal cancers in the same individual (although this occurs rarely), the expected number in our SIR calculation may be overestimated, resulting in an underestimation of the relative risk. This should, however, not have a major impact since the survival is very poor for gastroesophageal cancer patients and the proportion of patients with repeated gastroesophageal cancer in the cancer register is low (< 1% in our dataset). Another aspect of the incidence rate from the Swedish Cancer Register is that cancer cases diagnosed post-mortem through autopsy are also recorded. We did not include those cases in our study population and therefore the SIR may be a slight underestimation. Even so, this should not have a major influence given the low proportion (about 2.5% of the cancer cases in our study population were identified first at autopsy) and that autopsies have become less and less frequent with time.

#### *4.3.1.3 Logistic regression and interactions*

We used logistic regression to calculate odds ratios (ORs). The rationale for using logistic regression instead of linear regression was that the outcome was binary and that logistic regression is more convenient for handling multivariate models. Although we frequency-matched the controls on group level by sex and age-group (5-year intervals), we used unconditional logistic regression. This was decided since the matching was done at such an aggregated level that standard analysis adjustment for the matching factors was the most appropriate method to use, which is in line with the reasoning in a previous article on matched logistic regression analysis (104). Some of the assumptions underlying a logistic regression model is that the independent variables should not correlate with each other. Another

assumption is that the independent variables and the dependent variable are linearly related on the log odds scale (105).

To test for multiplicative effects between our main exposure of gastric atrophy together with poor health, we included an interaction term in our multivariate logistic regression model.

Additive interaction was tested using the following measures: relative excess risk due to interaction (RERI<sub>OR</sub>), synergy index (S), and attributable fraction due to interaction also called attributable proportion (AP). The measures for additive or so called biologic interaction rely on a model by Rothman (106) for causal mechanisms and assume no confounding. Based on this model, Rothman argues that biologic interaction should be assessed as departure from additivity of effect and not multiplicativity.

As explained by Knol et al the RERI is “part of the total effect that is due to interaction”, S is the “ratio between combined effect and individual effects”, and AP is the “proportion of the combined effect that is due to interaction” (107). The equations for the additive interaction tests are described in previous articles on the subject (107, 108).

#### *4.3.1.4 Cox regression model*

Hazard ratios (HRs) were estimated by using Cox proportional hazards regression models. The Cox regression model is called a semi-parametric model because it does not make a parametric assumption about the baseline hazard function but instead, assumes that the hazard rate ratio is constant. The advantage of using Cox regression instead of a parametric model was that we did not know the underlying distribution of the hazard for death or disease outcome and did not have a need to estimate it. If it would have been important to estimate the baseline hazard rate, we could have employed a flexible parametric model. The proportional hazards assumption in our studies was checked using Schoenfeld residuals and significance tests. When there was indication of violation of the proportional hazards assumption, we stratified the model by the co-variate in question. Another way of dealing with non-proportional hazards is to use the interval Poisson model by defining short intervals of time that includes an interaction term with time. The interval Poisson regression model differs from the Cox regression model in that the Poisson model assumes a constant hazard within the time intervals. However, in a scenario where there are shorter intervals of a couple of months at a time, the Poisson model yields similar estimates to that from a Cox model. Alternatively, we could have introduced an interaction term with time for the non-proportional co-variate. Given that we did not need to estimate the effect of the non-proportional co-variables, model stratification was ultimately more efficient and the method of choice for dealing with non-proportional hazards.

### **4.3.2 Random and systematic error**

#### *4.3.2.1 Random error*

Random error refers to a random deviation from the true observation and can push the observed value away from the true value in any direction. Random error affects the variability but not

the average. Random error can be due to sampling error, or non-sampling error. Sampling error occurs due to taking a sample and not studying the entire source population and is usually influenced by the sample size and the variation of the variable of interest in the source population. Two widely used methods for appraising the impact of sampling error are hypothesis testing and estimating the confidence interval. An appropriate sample size can help decrease the impact of sampling error on study estimates. To arrive at a suitable sample size when planning a study a power calculation can be employed. On the other hand, random error can also occur due to non-sampling error, such as measurement error which can be due to difficulties measuring the variable of interest or variation with time. It is more difficult to assess the impact of non-sampling random errors than sampling errors and requires careful considerations when planning and conducting the study regarding the measurement methods and how they are executed, reported, processed and analyzed.

In study I and II we virtually used the entire source population in Sweden during the study period and the number of outcome is reasonably large, so the sampling error should be small but non-sampling errors are likely due to measurement error and variations with time since we could not influence the collection of retrospective data. In study III we tried to reduce the impact of sampling error by a prior power calculation and reaching an appropriate sample size. However, random measurement error is still likely. Study IV was hypothesis-generating and not experimental since we lacked prior knowledge about the prevalence of exposure to each specific chemotherapy regimen and the effect size we could expect. The impact of random sampling error is considerable, but the measurement error should be low due to high completeness of the registers used.

#### *4.3.2.2 Systematic error*

Systematic error, or bias, influences the estimate in a predicable direction, either under- or overestimation of the “true” population estimate. To reduce systematic error careful attention needs to be paid to the study design, gathering and analysis of the data. There are numerous sources of bias and each study requires prior planning to reduce the specific biases for each situation, as we described these in each paper I-IV. Below is a brief description followed by how we tried to decrease the major sources of bias. The goal with reducing systematic error is to achieve accurate estimates, close to the “true” population value.

##### *4.3.2.2.1 Selection bias*

Selection bias is when the study sample is not representative of the source or target population. In our register-based cohort studies (study I-II), selection bias is mainly due to reverse causation or loss of follow up. To reduce selection bias in study I and II we excluded the first two years of follow-up. The high-quality registers guaranteed that loss to follow-up is negligible. In study III, we tried to decrease selection bias by inviting all potential cases at the major hospitals in the region to participate. We also complemented with missing cases from the local Cancer register. Non-respondent cases and controls were not statistically different from respondents

regarding age at interview and sex, but we could not assess differences regarding other sociodemographic or tumor-related variables.

#### 4.3.2.2.2 Information bias

Information bias is concerned with systematic misclassification of the exposure or outcome. We believe that the risk for differential misclassification of the exposure should be low in study I and II. However, some histopathology diagnoses are known to have a high inter-observer variability so gastric atrophy and esophageal metaplasia with low-grade dysplasia are more prone to misclassification compared to intestinal metaplasia or high-grade dysplasia. In study III we tried to assess the impact of misclassification of the exposure by employing other cut-off criteria for Pepsinogen I, II and their ratio, resulting in a similar magnitude of association, although not always statistically significant. Further, in study III, the histopathology samples of suspected cases were carefully reviewed by one pathologist to reduce the risk of misclassification of the outcome. Differential misclassification of the exposure or outcome in study IV should not be a major issue given the high-quality registers employed for data collection.

#### 4.3.2.2.3 Confounding

A confounding factor is related to both the exposure and outcome and explains all or some of the estimated association between the exposure and the outcome. We have accounted for the potential confounding factors we considered in the separate papers I-IV. In addition to adjusting for potential confounding factors in our multivariate models we also frequency matched cases and controls by age (5-year intervals) and sex in study III and restricted or stratified the study sample to handle potential confounding factors in study I-IV. Despite our best efforts to adjust for potential confounding factors, there is always the risk for residual unmeasured confounding. In study I residual confounding could be life-style factors that we did not control for; in study II it could be some currently unknown but widespread environmental exposure; in study III it could be the socioeconomic status and other unknown factors; and in study IV it could be confounding by indication.

## 4.4 RESEARCH ETHICS

### 4.4.1 Ethical principles and approvals

All studies were conducted according to the WMA Declaration of Helsinki - Ethical Principles for Medical Research Involving Human Subjects and were approved by the regional ethics committee in Stockholm or the Swedish Ethical Review Authority according to the following: study I (Dnr 2010/819-31-3, 2013/1244-32, 2015/1469-32, 2016/247-32, 2016/525-32), study II (2014/1287-314, 2020-00382), study III (Dnr 2018/357-31, and the Institutional Review Board of School of Life Sciences, Fudan University), and study IV (2012/1236-31-4, 2012/1726-32, 2014/849-32, 2017/597-32).

#### **4.4.2 Ethical considerations**

Ethical considerations specific for epidemiological studies generally concern the interests of the society versus the interest of the individual. It is often assumed that what lies in the interest of society is also beneficiary for the individual. Many ethical dilemmas however arise when these interests are mutually exclusive. In extreme cases, the need of the society to identify or limit a disease might impose a threat to the integrity, well-being or liberty of the individual. It is the responsibility of the researcher to ensure that the research they conduct is ethically and morally justified. Briefly, research ethics encompasses four main aspects with regards to the research subjects involved: autonomy, non-maleficence, beneficence and justice.

The respect for autonomy comprehends preserving human dignity, human rights and freedoms. Further, it requires accessible and objective information about the study details, and that all participation in research is voluntary, where individuals are allowed to withdraw at any time. This is usually ensured by written informed consent. In some register-based epidemiological studies informed consent can be difficult to achieve and might inflict more harm than benefit to the individual, but the respect for the autonomy can still be uphold if individuals are sharing their individual data voluntarily or can opt-out from participation in health-care registers or administrative registers. We could not obtain informed consent from all the study participants in our register-based studies I, II and IV but we believe the benefits outweigh the risks in these studies. In study III participation was voluntary and patients were only included if they gave informed consent.

The next aspect is non-maleficence and many times entails avoiding unnecessary invasive or unsafe methods of diagnosis or interventions. In study III this was one of the reasons we chose to use the less invasive method of “serological biopsy” by measuring Pepsinogen I and II in blood samples instead of performing the more invasive method of diagnosis through endoscopy and histopathology, especially for controls. We also handled sensitive data with great care in all studies to ensure that it was only accessible by the necessary researchers involved in the study.

We aimed to design and conduct our studies so that they would be beneficial to either the study participants themselves or future patients. And lastly, the aspect of justice in our studies concerned studying gastroesophageal cancers that are much less studied but more fatal than more common cancers such as breast, lung or prostate cancer and furthermore including patient groups that are otherwise often excluded or disregarded from clinical trials such as patients with advanced tumor stage or palliative treatment in study IV.

Besides the ethical considerations in relation to the study participants there are numerous other ethical dilemmas in research ethics concerning undeclared conflicts of interest, plagiarism, authorship conflicts and other forms of scientific misbehavior. It is the interest of the research community as well the society to keep these ethical issues at a minimum since it undermines the trust in the scientific community.

## 5 MAIN RESULTS

### 5.1 HIGH-RISK INDIVIDUALS FOR GASTROESOPHAGEAL CANCER

#### 5.1.1 Family history of gastric mucosal changes

Our main finding in paper I showed that being a first-degree relative (parent, sibling or child) to a person with a gastric mucosal lesion was associated with an increased risk for gastric cancer, and more so for gastric non-cardia cancer than cardia cancer (**Table 1**) (77). The excess risk for gastric non-cardia cancer increased incrementally with the severity of the gastric mucosal lesion (except for “other, unspecified changes”), and the increase of excess risk was statistically significant. The excess risk for gastric non-cardia cancer was 30% among individuals with a family history of gastritis, and 50-60% among patients with atrophic gastritis, intestinal metaplasia or dysplasia compared to the general Swedish population.

**Table 1.** Standardized incidence ratios (SIRs) with 95% confidence intervals (CIs) for relatives of biopsy patients, grouped by the mucosal changes of index biopsy patient

| Family history of gastric mucosal changes*   | All gastric cancer |                       |                           | Non-cardia gastric cancer |                       |                           | Cardia gastric cancer |                       |                           |
|--|--------------------|-----------------------|---------------------------|---------------------------|-----------------------|---------------------------|-----------------------|-----------------------|---------------------------|
|  | Obs                | Crude IR <sup>±</sup> | SIR <sup>§</sup> (95% CI) | Obs                       | Crude IR <sup>±</sup> | SIR <sup>§</sup> (95% CI) | Obs                   | Crude IR <sup>±</sup> | SIR <sup>§</sup> (95% CI) |
| Normal/minor changes                         | 357                | 10.2                  | 1.01 (0.94-1.12)          | 268                       | 7.7                   | 1.00 (0.88-1.13)          | 89                    | 2.5                   | 1.04 (0.83-1.28)          |
| Gastritis                                    | 592                | 12.0                  | 1.24 (1.15-1.35)          | 475                       | 9.6                   | 1.31 (1.19-1.43)          | 117                   | 2.4                   | 1.04 (0.86-1.25)          |
| Other unspecified changes                    | 145                | 11.6                  | 1.09 (0.92-1.28)          | 99                        | 7.9                   | 0.99 (0.80-1.20)          | 46                    | 3.7                   | 1.40 (1.02-1.86)          |
| Atrophic gastritis                           | 47                 | 13.2                  | 1.31 (0.96-1.75)          | 40                        | 11.2                  | 1.47 (1.05-2.01)          | 7                     | 2.0                   | 0.81 (0.33-1.67)          |
| Intestinal metaplasia                        | 27                 | 12.7                  | 1.28 (0.84-1.86)          | 25                        | 11.7                  | 1.59 (1.03-2.35)          | 2                     | 1.0                   | 0.37 (0.04-1.33)          |
| Dysplasia                                    | 29                 | 13.1                  | 1.22 (0.82-1.75)          | 28                        | 12.6                  | 1.53 (1.02-2.21)          | 1                     | 0.5                   | 0.18 (0.00-1.01)          |
| Gastric cancer                               | 105                | 22.7                  | 2.17 (1.78-2.63)          | 85                        | 18.4                  | 2.33 (1.86-2.89)          | 20                    | 4.3                   | 1.68 (1.03-2.60)          |
| Chi <sup>2</sup> test for trend <sup>¶</sup> |                    |                       | <i>p</i> =0.0028          |                           |                       | <i>p</i> =0.0018          |                       |                       | <i>p</i> =0.2818          |

Obs, observed cases. IR, Incidence Rate.

\* Defined by the gastric cancer family history known at baseline (Cancer register) or the mucosal change diagnosis of the index biopsy patient

<sup>±</sup> Per 100 000 person-years

<sup>§</sup> Observed to expected number of GC cases, based on age- (5-year strata), calendar year- (5-year strata) and sex-specific incidence rates in the total Swedish population. Ninety-five percent CIs of SIRs were calculated by assuming that observed cancer occurrence followed a Poisson distribution.

<sup>¶</sup> Excluded the 'other unspecified changes' category

An interesting finding was that having a sibling with gastric lesions had a higher excess risk than having an affected parent or children. Those having siblings afflicted with severe gastric mucosal lesions (atrophic gastritis, intestinal metaplasia or dysplasia) had a more than 2-fold risk (HR ranged from 2.3 to 2.7) for gastric non-cardia cancer compared to those having a sibling with normal or minor mucosal changes (**Table 2**).

**Table 2.** Hazard ratios (HRs) and 95 % confidence intervals (CIs) for non-cardia gastric cancer among relatives of biopsy patients with different pathological changes in the stomach compared to relatives of patients with normal gastric mucosa, by classes of first-degree relatives

| Family history of gastric mucosal changes | Parents (n=195 704) |                   | Siblings (n=283 371) |                   | Children (n=424 262) |                   |
|---|---------------------|-------------------|----------------------|-------------------|----------------------|-------------------|
|   | Number of cases     | HRs and 95 % CIs* | Number of cases      | HRs and 95 % CIs* | Number of cases      | HRs and 95 % CIs* |
| Normal/minor changes                      | 204                 | Reference         | 55                   | Reference         | 9                    | Reference         |
| Gastritis                                 | 267                 | 1.16 (0.96-1.39)  | 184                  | 1.70 (1.25-2.31)  | 24                   | 1.26 (0.58-2.74)  |
| Other unspecified changes                 | 52                  | 0.91 (0.67-1.24)  | 46                   | 1.70 (1.15-2.52)  | 1                    | 0.19 (0.02-1.50)  |
| Atrophic gastritis                        | 20                  | 1.17 (0.73-1.85)  | 20                   | 2.48 (1.48-4.15)  | 0                    | -                 |
| Intestinal metaplasia                     | 10                  | 1.43 (0.75-2.71)  | 13                   | 2.71 (1.47-5.00)  | 2                    | 1.72 (0.36-8.14)  |
| Dysplasia                                 | 11                  | 1.11 (0.60-2.05)  | 14                   | 2.34 (1.29-4.24)  | 3                    | 2.62 (0.69-9.90)  |
| Gastric cancer                            | 25                  | 1.67 (1.10-2.55)  | 56                   | 3.83 (2.63-5.58)  | 4                    | 1.69 (0.52-5.55)  |

\* Using attained age as underlying time scale, estimated by Cox proportional hazards regression model, adjusted for sex, family size, year of birth, and stratified by pathology department.

Furthermore, we discovered that the risk for gastric non-cardia cancer could be further increased among individuals with various gastric lesions if they also had a first-degree relative with a gastric lesion. The excess risks ranged from 140-280% for those with gastritis, other changes, severe lesions (atrophic gastritis, intestinal metaplasia, and dysplasia) and with first-degree relatives ever diagnosed with normal/minor changes in the stomach, while the corresponding figures were 240% for gastritis and 580% for severe lesions among those with first-degree relatives ever diagnosed with a gastric severe lesion. The risks were consistently higher for both first-degree relatives of gastric cancer patients (Tables 1, 2) and those with gastric mucosal lesions who had family history of gastric cancer (**Table 3**) (77).



**Table 3.** Observed number, crude incidence rate of non-cardia gastric cancers (GC) and standardized incidence ratios (SIRs) with 95% confidence intervals (CIs) for biopsy patients, grouped by family history of gastric mucosal changes

| Mucosal status at baseline   | Exposure group<br>Family history of gastric mucosal changes | Biopsy patients ( <i>n</i> = 240 101) |                       |                          |
|------------------------------|---|---------------------------------------|-----------------------|--------------------------|
|                              |   | Obs <sup>*</sup>                      | Crude IR <sup>±</sup> | SIR <sup>§</sup> (95%CI) |
| Normal/minor mucosal changes | No/minor changes detected                                   | 66                                    | 9.9                   | 1.13 (0.90-1.41)         |
|                              | Gastritis   | 6                                     | 7.7                   | 1.07 (0.39-2.33)         |
|                              | AG/ IM/ dysplasia   | 1                                     | 5.2                   | 0.75 (0.02-4.20)         |
|                              | GC  | 0                                     | 0.0                   | -                        |
| Gastritis                    | No/minor changes detected                                   | 238                                   | 25.0                  | 2.40 (2.10-2.72)         |
|                              | Gastritis   | 25                                    | 20.0                  | 2.08 (1.35-3.70)         |
|                              | AG/ IM/ dysplasia   | 11                                    | 33.2                  | 3.42 (1.71-6.12)         |
|                              | GC  | 13                                    | 43.4                  | 3.32 (1.77-5.68)         |
| Other unspecified diagnoses  | No/minor changes detected                                   | 85                                    | 31.1                  | 2.64 (1.11-3.26)         |
|                              | Gastritis   | 6                                     | 18.3                  | 1.76 (0.65-3.83)         |
|                              | AG/ IM/ dysplasia   | 2                                     | 21.4                  | 2.04 (0.25-7.36)         |
|                              | GC  | 4                                     | 37.0                  | 2.81 (0.77-7.20)         |
| AG/IM/ dysplasia             | No/minor changes detected                                   | 68                                    | 64.8                  | 3.78 (2.99-4.71)         |
|                              | Gastritis   | 13                                    | 88.8                  | 5.41 (2.96-9.07)         |
|                              | AG/ IM/ dysplasia   | 4                                     | 97.1                  | 6.83 (2.22-9.07)         |
|                              | GC  | 5                                     | 115.4                 | 7.92 (2.57-18.48)        |

AG, atrophic gastritis; IM, intestinal metaplasia; GC, gastric cancer; Obs, observed cases; IR, Incidence Rate

<sup>\*</sup> The first two years of observation and corresponding events were excluded.

<sup>±</sup> Per 100 000 person-years.

<sup>§</sup> Observed to expected number of GC cases, based on age- (5-year strata), calendar year- (5-year strata) and sex-specific incidence rates in the total Swedish population. Ninety-five percent CIs of SIRs were calculated by assuming that observed cancer occurrence followed a Poisson distribution.

### 5.1.2 Pre-malignant esophageal lesions

In paper II we demonstrated that individuals who were diagnosed with non-dysplastic intestinal metaplasia, at their first esophageal biopsy, had an 11-fold increased risk for esophageal adenocarcinoma compared to the general Swedish population (SIR 10.8; 95% CI 7.8-14.6) (**Table 4**).

Surprisingly, the excess risk for esophageal adenocarcinoma was similar among patients with gastric or glandular metaplasia (SIR 11.9; 9.9-14.1) and other metaplasia types (mainly unspecified and squamous) (SIR 9.8; 5.5-16.2) (**Table 4**).

The excess risk for esophageal adenocarcinoma was almost three times as high among columnar metaplasia (intestinal, gastric/glandular) patients with low-grade dysplasia as patients with non-dysplastic columnar metaplasia (SIR 30.9; 21.0-43.8) (**Table 4**).

**Table 4.** Incidence rate (IR) and standardized incidence ratio (SIR) of gastroesophageal cancers among patients with esophageal biopsies in Sweden

| Group by histopathology            | Esophagus      |     |                  |                         |     |                | Stomach               |     |                 |                           |     |               |
|------------------------------------|----------------|-----|------------------|-------------------------|-----|----------------|-----------------------|-----|-----------------|---------------------------|-----|---------------|
|                                    | Adenocarcinoma |     |                  | Squamous cell carcinoma |     |                | Cardia adenocarcinoma |     |                 | Non-cardia adenocarcinoma |     |               |
|                                    | Obs            | IR  | SIR (95%CI)      | Obs                     | IR  | SIR (95%CI)    | Obs                   | IR  | SIR (95%CI)     | Obs                       | IR  | SIR (95%CI)   |
| <b>Normal morphology</b>           | 18             | 0.1 | 1.9 (1.1,2.9)    | 14                      | 0.1 | 1.6 (0.9,2.7)  | 9                     | 0.1 | 1.0 (0.4,1.8)   | 23                        | 0.1 | 0.8 (0.5,1.2) |
| <b>Minor lesions</b>               |                |     |                  |                         |     |                |                       |     |                 |                           |     |               |
| Minor/other                        | 26             | 0.3 | 4.8 (3.1,7.0)    | 16                      | 0.2 | 3.3 (1.9,5.3)  | 16                    | 0.2 | 3.0 (1.7,4.8)   | 10                        | 0.1 | 0.6 (0.3,1.1) |
| Ulcer/hemorrhage                   | 28             | 0.4 | 5.0 (3.4,7.3)    | 4                       | 0.1 | 0.8 (0.2,2.1)  | 13                    | 0.2 | 2.3 (1.2,4.0)   | 9                         | 0.1 | 0.5 (0.2,0.9) |
| Inflammation/hyperplasia           | 63             | 0.2 | 2.7 (2.1,3.4)    | 39                      | 0.1 | 2.0 (1.4,2.7)  | 29                    | 0.1 | 1.3 (0.9,1.9)   | 46                        | 0.1 | 0.7 (0.5,0.9) |
| <b>Severe lesions</b>              |                |     |                  |                         |     |                |                       |     |                 |                           |     |               |
| Barrett's esophagus                |                |     |                  |                         |     |                |                       |     |                 |                           |     |               |
| Non-dysplastic columnar metaplasia | 173            | 0.9 | 11.6 (9.9,13.4)  | 12                      | 0.1 | 1.1 (0.6,1.9)  | 47                    | 0.3 | 3.6 (2.6,4.8)   | 26                        | 0.1 | 0.8 (0.5,1.2) |
| Intestinal metaplasia              | 43             | 0.9 | 10.8 (7.8,14.6)  | 3                       | 0.1 | 1.1 (0.2,3.2)  | 11                    | 0.2 | 3.3 (1.7,5.9)   | 8                         | 0.2 | 1.1 (0.5,2.1) |
| Gastric/glandular metaplasia       | 130            | 0.9 | 11.9 (9.9,14.1)  | 9                       | 0.1 | 1.1 (0.5,2.1)  | 36                    | 0.3 | 3.7 (2.6,5.1)   | 18                        | 0.1 | 0.7 (0.4,1.1) |
| Columnar metaplasia + LGD          | 31             | 3.0 | 30.9 (21.0,43.8) | 2                       | 0.2 | 2.9 (0.4,10.5) | 12                    | 1.2 | 14.1 (7.3,24.5) | 2                         | 0.2 | 1.0 (0.1,3.5) |
| Other metaplasia                   | 15             | 0.8 | 9.8 (5.5,16.2)   | 2                       | 0.1 | 1.7 (0.2,6.0)  | 3                     | 0.2 | 2.1 (0.4,6.2)   | 5                         | 0.3 | 1.2 (0.4,2.9) |

Obs, observed number of cancer cases; py, person-years.

\* (1/1000 py)

Moreover, the excess risk for esophageal adenocarcinoma among non-dysplastic columnar metaplasia patients decreased with age at entry, follow-up duration and calendar year at entry (**Table 5**).

**Table 5.** Incidence rate (IR) and standardized incidence ratio (SIR) of esophageal or cardia adenocarcinoma among patients with non-dysplastic columnar metaplasia according to sex, age at entry, follow-up duration and calendar year at entry

| Characteristics                  | Esophageal Adenocarcinoma |                            |                      | Gastric cardia adenocarcinoma |                            |                    |
|----------------------------------|---------------------------|----------------------------|----------------------|-------------------------------|----------------------------|--------------------|
|                                  | Obs                       | IR (1/1000 py)<br>(95% CI) | SIR<br>(95% CI)      | Obs                           | IR (1/1000 py)<br>(95% CI) | SIR<br>(95% CI)    |
| <b>Sex</b>                       |                           |                            |                      |                               |                            |                    |
| Men                              | 153                       | 1.3 (1.1,1.6)              | 11.6 (9.8,13.6)      | 41                            | 0.4 (0.3,0.5)              | 3.7 (2.7,5.0)      |
| Women                            | 20                        | 0.3 (0.2,0.4)              | 11.6 (7.1,17.8)      | 6                             | 0.1 (0.0,0.2)              | 3.0 (1.1,6.6)      |
| <b>Age group at entry</b>        |                           |                            |                      |                               |                            |                    |
| 30-49 yrs                        | 7                         | 0.3 (0.1,0.7)              | 58.6<br>(23.6,120.8) | 2                             | 0.1 (0.0,0.4)              | 12.2<br>(1.5,44.0) |
| 50-59 yrs                        | 22                        | 0.6 (0.4,1.0)              | 20.4<br>(12.8,30.9)  | 6                             | 0.2 (0.1,0.4)              | 5.7 (2.1,12.4)     |
| 60-69 yrs                        | 53                        | 1.0 (0.7,1.3)              | 12.4 (9.3,16.2)      | 18                            | 0.3 (0.2,0.5)              | 4.9 (2.9,7.8)      |
| ≥ 70 yrs                         | 91                        | 1.2 (0.9,1.4)              | 9.6 (7.7,11.8)       | 21                            | 0.3 (0.2,0.4)              | 2.6 (1.6,3.9)      |
| <i>P</i> for trend               |                           |                            | <0.01                |                               |                            | <0.01              |
| <b>Follow-up duration, years</b> |                           |                            |                      |                               |                            |                    |
| 0-5                              | 76                        | 3.2 (2.5,4.0)              | 36.2<br>(28.5,45.3)  | 22                            | 0.9 (0.6,1.4)              | 12.2<br>(7.6,18.5) |
| 5-10                             | 55                        | 1.0 (0.7,1.3)              | 11.8 (8.9,15.4)      | 15                            | 0.3 (0.1,0.4)              | 3.7 (2.1,6.2)      |
| 10-15                            | 22                        | 0.4 (0.2,0.5)              | 4.6 (2.9,7.0)        | 6                             | 0.1 (0.0,0.2)              | 1.4 (0.5,3.1)      |
| ≥ 15                             | 20                        | 0.4 (0.3,0.7)              | 5.8 (3.5,9.0)        | 4                             | 0.1 (0.0,0.2)              | 1.3 (0.4,3.3)      |
| <i>P</i> for trend               |                           |                            | <0.01                |                               |                            | <0.01              |
| <b>Calendar year at entry</b>    |                           |                            |                      |                               |                            |                    |
| 1981-1999                        | 67                        | 1.8 (1.4,2.3)              | 22.1<br>(17.1,28.1)  | 16                            | 0.4 (0.2,0.7)              | 5.2 (3.0,8.5)      |
| 2000-2006                        | 70                        | 0.8 (0.7,1.1)              | 10.6 (8.2,13.4)      | 16                            | 0.2 (0.1,0.3)              | 2.8 (1.6,4.6)      |
| 2007-2016                        | 36                        | 0.6 (0.4,0.8)              | 6.8 (4.8,9.4)        | 15                            | 0.2 (0.1,0.4)              | 3.5 (2.0, 5.8)     |
| <i>P</i> for trend               |                           |                            | <0.01                |                               |                            | 0.89               |

Obs, observed number of cancer cases; py, person-years.

### 5.1.3 Gastric atrophy

Our case-control study in Taixing, China described in paper III, revealed that gastric atrophy (defined as serum pepsinogen I < 55 µg/l) was associated with a 60% increased risk for esophageal squamous cell carcinoma (OR 1.61; 95% CI 1.33-1.96) when fully adjusted for the following potential confounding factors: age (continuous), sex, education, marital status, occupation, family wealth score, body mass index 10 years prior to the interview, tea drinking, history of esophageal cancer among first-degree relatives, smoking status, alcohol drinking status, *H. pylori* serology-status, sum of missing and filled teeth, and frequency of tooth brushing per day (Table 6) (109).

**Table 6.** The odds ratios (ORs) and 95% confidence intervals (CI) for esophageal squamous cell carcinoma in association with gastric atrophy defined by PGI <55 µg/l, and *H. pylori* seropositivity in a case-control study, Taixing, China during 2010-2014.

| Variables                           | Controls        |      | Cases           |      | Age/sex-adjusted <sup>b</sup> |            |         | Fully-adjusted (except MFT and tooth brushing) <sup>c</sup> |            |         | Fully-adjusted <sup>d</sup> |            |         |
|-------------------------------------|-----------------|------|-----------------|------|-------------------------------|------------|---------|---|------------|---------|-----------------------------|------------|---------|
|                                     | No <sup>a</sup> | %    | No <sup>a</sup> | %    | OR                            | 95% CI     | P-value | OR  | 95% CI     | P-value | OR                          | 95% CI     | P-value |
| <b>PGI (µg l<sup>-1</sup>)</b>      |                 |      |                 |      |                               |            | <0.01   |   |            | <0.01   |                             |            | <0.01   |
| PGI ≥ 55                            | 1592            | 83.3 | 841             | 75.1 | 1.00                          | Ref.       |         | 1.00  | Ref.       |         | 1.00                        | Ref.       |         |
| PGI < 55                            | 319             | 16.7 | 279             | 24.9 | 1.60                          | 1.34, 1.90 |         | 1.63  | 1.35, 1.97 |         | 1.61                        | 1.33, 1.96 |         |
| <b><i>H. pylori</i> sero-status</b> |                 |      |                 |      |                               |            | 0.22    |   |            | 0.21    |                             |            | 0.25    |
| <i>H. pylori</i> -                  | 622             | 32.5 | 340             | 30.3 | 1.00                          | Ref.       |         | 1.00  | Ref.       |         | 1.00                        | Ref.       |         |
| <i>H. pylori</i> +                  | 1291            | 67.5 | 782             | 69.7 | 1.10                          | 0.95, 1.29 |         | 1.11  | 0.94, 1.32 |         | 1.11                        | 0.93, 1.31 |         |

No, number; ORs, Odds ratios; 95% CI, 95% Confidence intervals; PGI, Pepsinogen I; *H. pylori*, *Helicobacter pylori*; MFT, Missing and filled teeth.

<sup>a</sup> Complete observations in the fully-adjusted model.

<sup>b</sup> Adjusted for age (continuous) and sex.

<sup>c</sup> Adjusted for age (continuous), sex, education, marital status, occupation, family wealth score, body mass index 10 years ago, tea drinking, history of esophageal cancer among first-degree relatives, smoking status, alcohol drinking status and *H. pylori* sero-status.

<sup>d</sup> Additionally adjusted for sum of missing and filled teeth and frequency of tooth brushing per day.

In paper III, we also showed an additive interaction between gastric atrophy and poor oral health, which was measured as number of tooth brushing per day (RERI<sub>OR</sub> 1.28; 0.39-2.18) (109).

## 5.2 TREATMENT FACTORS ASSOCIATED WITH THE PROGNOSIS OF GASTROESOPHAGEAL CANCER

### 5.2.1 Choice of chemotherapy

The main findings in paper IV, based on the fully-adjusted Cox model, showed that the choice of cisplatin-fluorouracil versus carboplatin-fluorouracil was associated with a better survival among esophageal cancer patients with curative treatment intention at diagnosis (HR for death in the carboplatin-fluorouracil group compared to cisplatin-fluorouracil was 2.18; 95% CI 1.09-4.37) (**Table 7**) (110). Among gastroesophageal junction cancer patients, however, cisplatin-fluorouracil was associated with worse survival than fluorouracil-oxaliplatin (HR for death among fluorouracil-oxaliplatin compared to cisplatin-fluorouracil was 0.28; 0.08-0.96) (**Table 7**). Other chemotherapy regimens as compared to the reference group of most common regimens that were used during the study period did not have a statistically significant influence on the survival of gastroesophageal cancer patients in the fully-adjusted Cox model (**Table 7**).

**Table 7.** Cohort size and hazard ratios for chemotherapy with curative intention within six months from diagnosis with cancer in the esophagus, gastroesophageal junction or stomach (n=279).

| <b>Chemotherapy groups by cancer site</b> | <b>Cohort N</b> | <b>Adjusted HR<sup>a</sup></b> | <b>P-value</b> | <b>Adjusted HR<sup>b</sup></b> | <b>P-value</b> |
|---|-----------------|--------------------------------|----------------|--------------------------------|----------------|
| <b>Esophagus, N</b>                       | 132             |                                |                |                                |                |
| Cisplatin-fluorouracil                    | 85              | Ref.                           | Ref.           | Ref.                           | Ref.           |
| Fluorouracil-oxaliplatin                  | 23              | 1.53 (0.90-2.60)               | 0.12           | 1.28 (0.70-2.35)               | 0.43           |
| Carboplatin-fluorouracil                  | 14              | 2.33 (1.24-4.38)               | 0.01           | 2.18 (1.09-4.37)               | 0.03           |
| Other chemotherapy                        | 10              | 2.77 (1.34-5.73)               | 0.01           | 2.23 (1.02-4.91)               | 0.05           |
| <b>Gastroesophageal junction, N</b>       | 59              |                                |                |                                |                |
| Cisplatin-fluorouracil                    | 34              | Ref.                           | Ref.           | Ref.                           | Ref.           |
| Fluorouracil-oxaliplatin                  | 13              | 0.45 (0.16-1.25)               | 0.12           | 0.28 (0.08-0.96)               | 0.04           |
| Epirubicin-oxaliplatin-capecitabine       | 7               | 0.76 (0.27-2.11)               | 0.60           | 0.34 (0.07-1.73)               | 0.20           |
| Other chemotherapy                        | 5               | 1.00 (0.25-4.06)               | 1.00           | 0.72 (0.15-3.46)               | 0.68           |
| <b>Stomach, N</b>                         | 88              |                                |                |                                |                |
| Epirubicin-oxaliplatin-capecitabine       | 71              | Ref.                           | Ref.           | Ref.                           | Ref.           |
| Fluorouracil-irinotecan                   | 8               | 2.64 (1.13-6.18)               | 0.03           | 2.26 (0.92-5.53)               | 0.07           |
| Other chemotherapy                        | 9               | 0.45 (0.15-1.36)               | 0.16           | 0.45 (0.14-1.40)               | 0.17           |

<sup>a</sup> Adjusted for age (continuous) , sex and TNM-stage.

<sup>b</sup> Additionally adjusted for radiotherapy, comorbidity, marital status, education, income and country of birth.

## 6 DISCUSSION

The goal with this thesis has been to further expand the understanding of the causes behind gastroesophageal cancers and inspire future research about the causal pathways of gastroesophageal cancers. We hope that this can ultimately pave the way for surveillance and/or primary prevention among high-risk individuals. In the long run, we hope that the research findings presented in this thesis can assist in decreasing the disease burden. There are a number of strategies to reduce the gastroesophageal cancer burden. To date, extensive research efforts have been undertaken to reduce the gastroesophageal cancer fatality by novel chemotherapies and/or new chemotherapy regimens in combination with radiotherapy. This is understandable given that these cancers are characterized by high rates of fatality. Still, to radically reduce gastroesophageal cancer mortality rates, it seems extraordinary changes are required, either in our understanding of the diseases and/or in the treatment strategies. A focus of this thesis is biological risk markers which will be discussed first. How do our findings relate to the etiology of gastroesophageal cancers?

### 6.1 FROM BIOLOGICAL RISK MARKERS TO ETIOLOGY

#### 6.1.1 Familial clustering – is it due to shared environment or genetics?

What good is it to know that there is familial clustering of gastric mucosal lesions that is associated with an increased risk for gastric non-cardia cancer? To begin with, this information supports the theory of a pathogenic pathway described as the Correa cascade. Secondly, it can be useful for specialist physicians in gastroenterology or upper GI-surgeons when deciding which patients to include in surveillance systems for gastric non-cardia cancer. Although the cost-effectiveness of surveillance of patients with family history of gastric mucosal lesions has

not been established, the most recent guidelines from the Swedish Society of Gastroenterology (2018) based on existing literature including two studies from our group (77, 111) recommends that “patients with extensive chronic atrophic gastritis and/or intestinal metaplasia or dysplasia in the gastric mucosa – and particularly if they also have heredity for gastric cancer (or severe pre-cancerous mucosal lesions) or belong to an ethnic group with high risk – should be included in endoscopic surveillance”. Thirdly, in addition to other established risk factors such as male sex, smoking and *H. pylori* infection among others, the factor of familial clustering could be useful for primary health care physicians who need to make a decision on which patients to investigate further for non-cardia gastric cancer.

An examination of the relative contribution of environmental or genetic factors in the causation of sporadic gastric cancers was published previously in a Nordic twin registry collaboration (112), where it was reported that genetic heritability accounted for 28% of the variance in susceptibility, shared environmental factors explained 10%, and the remaining 62% could be attributed to non-shared environmental factors. Although the majority of gastric cancers are believed to be sporadic, about 10-20% are reported to have familial clustering while a minority of about 2-5% have hereditary forms of the disease (113). As our study encompassed gastric cancers with familial clustering, it is possible that the proportion of variance in gastric cancers explained by genetic, shared and non-shared environmental factors may differ in our study population compared to that based on a study population of sporadic cases. Our study on familial gastric cancers could therefore, in a broader sense, help untangle the environmental-genetic relationship among family-clustered gastric cancers.

Our study’s findings indicated a higher risk for gastric non-cardia cancer among siblings to index patients with severe gastric mucosal lesions (atrophic gastritis, intestinal metaplasia and dysplasia) than parents or children of index patients. The analysis among children was limited due to very few cancer cases as gastric cancer occurring at young age is rare. Still, the higher elevated risk among siblings than parents, could reflect that shared early-life environmental factors, such as life-style, *H. pylori* infection or *H. pylori* virulence in connection with susceptibility genes, are more likely causal factors for the majority of family-clustered non-cardia gastric cancer cases than hereditary dominant genes or non-shared environmental factors.

In summary, our study adds additional individuals to the group of high-risk individuals for gastric non-cardia cancer and has made a contribution to the surveillance recommendations in Sweden.

In contrast to non-cardia cancer, first-degree relatives to index patients with severe gastric mucosal lesions (atrophic gastritis, intestinal metaplasia and dysplasia) did not display a statistically significant increase in risk for gastric cardia cancer. Such a finding would support the hypothesis that gastric cardia cancer cases are mainly of sporadic origin, wherein environmental or life-style related factors are key. However, family clustering of gastroesophageal reflux and overweight/obesity which are associated with the risk for gastric cardia cancer also demonstrates possible relation with genetic and/or shared environmental

factors. For instance, a recent statistical modelling study suggests a total elimination of overweight and obesity in 2016 would reduce the number of gastric cardia cancer cases in the Nordic countries by 11.5% during the period of 2016-2045 (114). Similar studies could help reveal the relative importance of eradicating or preventing other environmental risk factors such as gastroesophageal reflux disease and smoking. Besides informing the general population about the risks associated with certain life-style factors, the health-care system has a responsibility to pay attention to treatment of high-risk conditions. This strategy is well established within the health care system. If found cost-effective, measures such as primary prevention of gastroesophageal cancers by treating gastroesophageal reflux disease or obesity could be introduced into the health-care system. Similar to the way high blood pressure is managed today.

### **6.1.2 Gastroesophageal morphological lesions and the “point of no return”**

#### *6.1.2.1 Gastric cancer*

The Correa cascade is the most established pathway among gastroesophageal cancers. The pathogenic pathway is initiated by *H. pylori* and followed by chronic inflammation, gastric atrophy, intestinal metaplasia, dysplasia and gastric non-cardia cancer (intestinal type). Two recent studies from our group (77, 111) support the epidemiological strength of this cascade, but many questions still remain unanswered regarding the etiology of non-cardia gastric cancer. Is it so that all cancer cases go through all stages of the Correa cascade and at the same pace, or can this differ between patients? What determines the progression from one lesion to another? Is the Correa cascade reversible and at what stages? Since most of the patients with changes according to the Correa cascade never progress to gastric cancer, one of the current challenges is to understand which patients are at high risk of progression, and how to identify such patients in order to offer the right patients surveillance in low-risk areas.

Intestinal metaplasia has been suggested as the “point of no return” for gastric non-cardia cancer and could be a candidate condition as an impetus for starting surveillance. The reason is that the cancer risk associated with intestinal metaplasia is considerable. Further, most intervention studies suggest that *H. pylori* eradication therapy is not as successful for patients with manifested intestinal metaplasia or any of the subsequent conditions in the Correa cascade (115). The existence of a “point of no return” is under debate and far from established since there are a few reports of regression in some intestinal metaplasia patients with or without *H. pylori* eradication (115). The current controversy set aside, the concept of a “point of no return”, if verified, could be of great clinical value to select high-risk individuals to include in surveillance programs in low-risk populations of gastric cancer where population-based screening is not cost-effective.

There is a need for high-quality prospective studies mapping cellular and molecular changes of intestinal metaplasia to understand which of these patients might progress to cancer and which of these are stable or might regress at follow-up. So far, a prospective study from China on intestinal metaplasia patients found genomic and epigenomic alterations associated with

progression to gastric cancer among patients with severe intestinal metaplasia (116). If confirmed in other populations or studies, these molecular markers could be useful for identifying sub-populations of intestinal metaplasia patients that would benefit from screening or surveillance (116). Another interesting prospective study from the Netherlands and Norway (n=279), of patients with previous diagnosis of atrophic gastritis, intestinal metaplasia or dysplasia that were surveilled for a median of 4.7 years, reported that low serum pepsinogens (PGI/PGII $\leq$ 3) and/or advanced OLGIM stage (OLGIM stage III/IV) was associated with malignant progression (117). If verified, these markers could also be of use in clinical practice.

#### 6.1.2.2 *Esophageal adenocarcinoma*

The pathway to esophageal adenocarcinoma is also believed to start with inflammatory changes due to chronic gastroesophageal reflux leading to Barrett's esophagus, with the potential of developing into low-grade dysplasia, high-grade dysplasia and esophageal adenocarcinoma.

There have been some recent updates regarding the esophageal adenocarcinoma risk associated with Barrett's esophagus. Currently, the research on esophageal adenocarcinoma etiology is focused on determining the sub-populations of Barrett's patients that are at high risk of malignant progression, much like for gastric non-cardia cancer.

The risk estimates for non-dysplastic intestinal metaplasia patients in paper II are in line with the results from a Danish histopathology register study (118), but lower for columnar metaplasia patients with low-grade dysplasia than most previous studies. The latter discrepancy could be explained by successful preventive treatment and surveillance efforts, or an effect of unverified low-grade dysplasia. The inter-observer variability for low-grade dysplasia is considerable and verification of the diagnosis by at least one additional pathologist or with a follow-up biopsy is required in specialist clinics. It is common that a suspected low-grade dysplasia diagnosis is not verified at a second review and the diagnosis is no dysplasia or indefinite for dysplasia. The number of verified low-grade dysplasia diagnoses has been associated with an increasing risk for esophageal adenocarcinoma. Unfortunately, we could not assess whether the low-grade dysplasia cases in our study were verified or not. We tried to assess the EAC risk among low-grade dysplasia patients with repeated low-grade dysplasia diagnoses but there were too few patients to allow a meaningful analysis. We can therefore not rule out that the observed lower risk estimate for columnar metaplasia patients with low-grade dysplasia was due to unverified low-grade dysplasia diagnoses.

The clinical challenge with esophageal adenocarcinoma is that the proportion of patients with a prior diagnosis of Barrett's esophagus is low, only 7% in a previous observational study (119). It is an open question if the remaining cancer cases were also preceded by Barrett's esophagus but not diagnosed due to lack of symptoms, or if there are alternative pathways to esophageal adenocarcinoma.

In current health care practice in Sweden, many Barrett's esophagus patients with low-grade dysplasia are surveilled; if the diagnosis is verified during follow-up, preventive treatment is considered. Barrett's esophagus with low-grade dysplasia that has been verified at one or



several follow-up endoscopies is managed as a form of “point of no return” due to the significant risk increase, but much is left to learn about the etiology of this disease.

Epidemiologic studies of the increased incidence of esophageal adenocarcinoma might help in generating hypotheses about the etiology. For instance, a global assessment of esophageal adenocarcinoma incidence suggested that some environmental exposure introduced in the 1950s could explain the dramatic increase of esophageal adenocarcinoma (21).

Study II contributes to the field by adding some epidemiological pieces to the esophageal adenocarcinoma puzzle. Our study demonstrates that the relative risk for esophageal adenocarcinoma is similar for those with intestinal metaplasia and those with gastric/glandular metaplasia. Whether gastric and glandular/cardia metaplasia is a risk factor for esophageal adenocarcinoma is currently under debate, where some countries include these patients in surveillance alongside intestinal metaplasia patients while others do not. The inconsistency is also seen in research, where some previous studies have included gastric metaplasia in their definition of Barrett’s esophagus while others did not. No previous population-based study of this sample size has separated these two groups and so the risk associated with gastric and glandular metaplasia has been unclear before.

Furthermore, other metaplasia types were also at a corresponding level of excess risk for esophageal adenocarcinoma. The implications of these findings could be quite substantial due to the vast number of patients with columnar lined epithelium in comparison with intestinal metaplasia. The most recent prevalence study of the general population in Sweden reported that the prevalence of columnar lined epithelium was about 10% and intestinal metaplasia 1.6% (120). If all columnar-lined epithelium patients were to be incorporated into surveillance program that would stress the health-care system substantially and might not even be possible due to limited number of doctors who can perform endoscopy. Besides the ethical consideration, practical and cost-effectiveness issues might be formidable. We hope that the largest impact will not be on clinical surveillance practices but instead on the etiological research field. Adding gastric and glandular patients to the population at risk for esophageal adenocarcinoma might lead to new discoveries about the pathogenic pathway. Before considering to include gastric and glandular metaplasia patients in clinical surveillance programs, it would be valuable to study how often and for how long these patients need follow-up. The results in study II suggest that the risk elevation is substantial up to ten years after the initial biopsy, but how often these patients should be examined was out of the scope of our study.

It was also noted that the risk for gastric cardia adenocarcinoma was elevated according to a similar pattern as for esophageal adenocarcinoma, but with a lower magnitude. This could be an indication that these cancers share common etiological pathways. It is not unreasonable considering that esophageal adenocarcinoma and gastric cardia cancer already share several “up-stream” risk factors such as gastroesophageal reflux, overweight/obesity and hiatal hernia. On the other hand, molecular and genetic differences are also present between esophageal adenocarcinoma and gastric cardia cancer, so hopefully the advent of new molecular and

genetic classification systems will correlate better with etiological processes, progression and treatment than current anatomical/histologic criteria.

#### *6.1.2.3 Esophageal squamous cell carcinoma*

The pathogenic pathway to esophageal squamous cell carcinoma remains largely unknown. Squamous dysplasia is the only established precursor lesion in high-risk areas, but studies in low-risk populations are limited. The results in study II showed that the relative risk was about twice as high for esophageal squamous cell carcinoma among patients with squamous dysplasia (in the group with minor/other lesions) and inflammation/hyperplasia, but no associations were found with the other esophageal lesions. The risk associated with inflammation and squamous dysplasia was low in our study population, much lower than the estimates reported in a previous study from Linxian, China (121); could these findings vindicate a different pathogenesis for esophageal squamous cell carcinoma in low-risk areas than high-risk areas?

There are some indications supporting this theory in the previous literature. Smoking and alcohol consumption seem to explain a larger proportion of the ESCC cases in Western countries than in Asia. Drinking hot beverages and using shallow water sources have been reported as risk factors in high-risk areas, but probably not feasible to study due to the low proportion exposed in low-risk areas.

Could gastric atrophy be a risk factor for ESCC even in low-risk areas? In study III the association between gastric atrophy and ESCC was confirmed in a high-risk area in China. Study III was however preceded by a histopathology register study in the Netherlands that also reported an association but could not demonstrate an increased risk with increasing severity of gastric atrophy. The authors of that study therefore explained the association they found with confounding, possibly smoking. The association was however first reported by our group in a study from Sweden, showing that gastric atrophy (measured by serum pepsinogen) was associated with an increased risk for ESCC in a study enrolling 85 incident ESCC patients and 499 randomly selected controls (40).

The reason for the discrepancy between the previous serology and histopathology studies could be due to methodological differences. The histopathological diagnosis of atrophic gastritis suffers from large inter-observer variation in a previous study from the Netherlands (122). Furthermore, 44% of the patients with atrophic gastritis had regressed at follow-up biopsy (122). In addition to this, the endoscopy inter-observer variation is already moderate in experienced endoscopists in high-risk areas (123), potentially increasing the risk for misclassification of the exposure. It is therefore reasonable to assume that there is a risk of misclassification of gastric atrophy assessed by retrospective histopathology registers. The serologic diagnosis of gastric atrophy in patients with already manifest ESCC however leaves the question open about the causal pathway between gastric atrophy and ESCC. The biological mechanism for this remains to be elucidated and some suggestions for future studies are listed in the section “Future studies” in this thesis.

## 6.2 PERSONALIZED TREATMENT – BEYOND THE HYPE

The investigation of the comparative effectiveness of chemotherapy regimens for gastroesophageal cancer patients in the Stockholm-Gotland area of Sweden in study IV revealed that the choice of cisplatin-fluorouracil was associated with better survival of esophageal cancer patients compared to those who were treated with carboplatin-fluorouracil, but worse survival in gastroesophageal junction cancer patients compared to those treated with fluorouracil-oxaliplatin among patients who were planned to receive curative treatment at diagnosis. We could not compare the relative effectiveness of the same chemotherapy regimens among palliative cancer patients due to different choice of chemotherapy regimens.

How can the knowledge we acquired in study IV contribute to personalized treatment of gastroesophageal cancer patients?

### 6.2.1.1 *The influence of chemotherapy among early and late stage cancer patients*

First, the aggregated effect of chemotherapy vs. no chemotherapy in gastroesophageal cancer patients with early or late stage tumors in the curative and palliative treatment groups was explored separately. The long-term survival was better for patients without chemotherapy in the group with early-stage tumors in the curative treatment intention at diagnosis. The median time to surgery was shorter among those without chemotherapy which could partly explain this finding. Another explanation could be confounding by indication; patients without a need for chemotherapy might have had more favorable patient and/or tumor characteristics such as better performance status, younger age, a favorable size and/or location of the tumor. Patients with early stage tumors who were in need of chemotherapy displayed a better relative survival than those without chemotherapy for up to about 200 days from chemotherapy, but this effect waned thereafter. Among those with early stage tumors that were treated with chemotherapy there is a minor proportion of patients who did not receive curative surgery due to the location of the tumor. This group with “definitive chemotherapy” (only chemotherapy, no surgery) could also have influenced the diminished long-term survival in the group with chemotherapy treatment.

Study IV therefore contributes with the knowledge that curative management of patients with early gastroesophageal cancers seems to influence the survival of these patients significantly. Time to surgery seems to be a key prognostic factor.

More finely stratified analyses according to histology, location and specific tumor stage could not be performed, but this would be feasible with the material from NREV. If the results from study IV are confirmed, it could be worthwhile to stress the importance of shortening time to surgery among early stage gastroesophageal cancer patients with curative treatment intention.

Further, the survival among patients who were intended to have curative treatment at diagnosis but who were found to have late stage tumors was analyzed. Among these patients, the difference between those with versus those without chemotherapy was not statistically significant. In this group there should be a minority of patients who had worsening of their

health status or tumor progression from the date of diagnosis until the planned start of treatment, which made them unsuitable for curative treatment. Compared to late stage palliative patients, the late stage curative patients had a much better survival. The late stage curative patients without chemotherapy even had a better survival than late stage palliative patients with chemotherapy. This difference should largely be attributed to the effect of surgery. It is likely that sub-groups of late stage curative patients had statistically significant better survival among chemotherapy patients than those without chemotherapy, but we did not have a sample size large enough for investigating this further.

Moreover, there were relatively few patients in the palliative group with early stage tumors and they did not have a statistically different survival probability in the chemotherapy vs. no chemotherapy group. Their survival was comparable to late stage palliative patients with chemotherapy treatment. This group most likely had factors making them unfit for curative surgery and this strongly affects their survival probability. The reasons for not offering this group curative surgery could not be assessed, but credible reasons could be patient-related factors such as co-morbidities, low performance status or otherwise frail health.

Many of the patients in study IV were late stage palliative patients (43%) and in this group the survival was significantly better for patients with chemotherapy than without. Late-stage palliative patients without chemotherapy had the worst survival in our study cohort. In this large group of patients there are most certainly individuals with aggressive tumors. With continued research on the etiology of these tumors it might be possible to find biomarkers to detect these patients earlier and provide primary prevention and/or develop targeted therapy based on their molecular/genetic setup.

In summary, chemotherapy seems to have a significant influence on the survival of early stage patients in the curative intention group and late stage patients in the palliative intention group compared to no chemotherapy.

#### *6.2.1.2 The choice of first line chemotherapy*

In study IV there was some variation in the choice of first line chemotherapy treatment among patients with the same treatment intention at diagnosis. The study period from 2008 to 2013 was before the pivotal clinical trials that have shaped current treatment recommendations. In a similar study on the period after 2013 one would expect to see less variation in the choice of treatment. The basis for the variation is unknown, but it seems it was not random. To begin with, there was a “preferred” choice of chemotherapy in all three sites: esophageal (cisplatin-fluorouracil), gastroesophageal junction (cisplatin-fluorouracil) and gastric cancer (epirubicin-oxaliplatin-capecitabine), which more than half of the patients were given in the curative treatment group. This pattern was most pronounced in the curative gastric cancer group. A similar pattern but not as distinct and with different chemotherapy regimens was noted in the palliative treatment group. A preference for cisplatin-fluorouracil in the curative group was noticed, but the combination oxaliplatin-fluorouracil was more common in the palliative group. This could be due to the “milder” toxicity profile of oxaliplatin compared to cisplatin (124).

### *6.2.1.3 Large variation in dose and duration*

In study IV, the chemotherapy dose and duration and subsequent lines of chemotherapy were assessed but not reported due to the small sample size that did not allow for stratification on those variables. As expected, the inter-individual variation was high.

The dose and duration of the treatment is generally based on subgroup characteristics such as age, sex, body surface area or BMI, renal function, co-morbidities, co-medications etc. The goal with adjusting the dose and duration of the treatment based on subgroup characteristics is to reach the exposure necessary for treatment effect without unacceptable toxicity. Based on these parameters the inter-individual variation in dose and duration is generally quite high.

Currently, the individual exposure to gastroesophageal chemotherapy drugs is measured mainly by following the effect on the tumor and organ-specific toxicities by monitoring the patient with blood samples and asking the patient at follow-up visits. Adjustments to the dose and duration are made along the way if the effect is undesirable or the patient cannot tolerate the treatment.

## **6.3 STRENGTHS AND LIMITATIONS**

The main strength of this thesis is the broad scope which includes both esophageal and gastric cancer. The rationale for studying both cancers in the same thesis was the knowledge from previous studies or guidelines that esophageal and gastric cancer patients share some common epidemiological risk factors and treatment strategies, but in this thesis we explored common biological pathways to cancer development, and the survival effects following chemotherapy treatment. These comparisons have been complex because gastroesophageal cancers are not just esophageal or gastric cancer but many different cancers within those two organs. Nevertheless fruitful conclusions can be drawn from comparing biological markers for these diseases. An altered gastric mucosal microenvironment is associated with an increased risk for esophageal squamous cell carcinoma in the middle and lower third of the esophagus, but esophageal mucosal lesions do not appear to be associated with an increased risk for gastric non-cardia adenocarcinoma and only a moderate increase in the risk for gastric cardia cancer.

Secondly, a strength of exploring biological markers is that we helped to set the scene for more goal-oriented future studies on biomarkers for pre-malignant lesions of gastroesophageal cancers.

Further, this thesis encompasses studies in both high- and low-risk areas of gastroesophageal cancers which is necessary when looking for etiological biomarkers and not just associations or population-specific risk factors.

Last but not least, in this thesis study designs with high precision and validity were employed. The large sample size in study I-III enabled the exploration of associations with better statistical power than previously. The high completeness of the registers used in study I, II, IV and the good response rate in study III is another strength. In addition, extensive efforts were made to control for possible sources of bias and confounding in all studies I-IV.

In spite of all efforts to deliver precise and valid estimates, there are clear limitations with this thesis. In the register-based studies the main limitation was the lack of validation of histopathology codes. Re-assessments of the histopathology samples to validate the accuracy of the diagnosis would have been preferable but was not feasible within the scope of this thesis. Future validation studies of the histopathology diagnoses for gastric and esophageal lesions are therefore warranted.

The main limitation with our case-control study (study III) was the lack of a validation study of the serological gastric atrophy cases with endoscopy.

## **7 CONCLUSIONS**

First, we conclude that family history of gastric mucosal lesions can be used for further risk stratification of gastric non-cardia adenocarcinoma among healthy individuals or individuals with various gastric lesions.

Second, we conclude that the risk for esophageal adenocarcinoma associated with gastric/glandular metaplasia is similar to the risk among patients with intestinal metaplasia.

Third, we confirm the association between gastric atrophy and esophageal squamous cell carcinoma, and its interaction with poor oral health to further increase the risk, in a high-risk region in China.

Last, we conclude that the choice of chemotherapy regimen for esophageal and gastroesophageal junction cancer patients may predict the survival among patients in the curative intention treatment group.

## **8 FUTURE STUDIES**

### **8.1 FAMILY-CLUSTERING**

Future studies of molecular changes underlying the Correa cascade should include patients with a family-history of the gastric mucosal lesions, including gastric atrophy, intestinal metaplasia and dysplasia. Particularly patients with siblings who also have gastric mucosal lesions could enable meaningful insights into the pathogenesis. Furthermore, studies exploring biomarkers for family-clustered gastric non-cardia cancer may also include healthy individuals or individuals afflicted with various gastric lesions with family history of gastric mucosal lesions.

Future genome-wide association studies should map host and *H. pylori* related genetic polymorphisms in family-clustered non-cardia gastric cancer. It might help understanding the necessary causal factors among high-risk individuals. In the long run, this might enable the

identification of a more specific group of high-risk individuals in need of surveillance and/or primary prevention.

There is also the question of when to start the surveillance of first-degree relatives and how often. Future epidemiological studies could help to improve risk stratification by considering geographical region, age-group and number of relatives affected. But modelling cost-effectiveness studies and ultimately clinical trials could help answer this question definitely.

## **8.2 PRECURSOR LESIONS**

The findings in study II regarding the association between columnar metaplasia and esophageal adenocarcinoma should be tested in other populations, preferably based on histopathological data. A validation of the Barrett's esophagus diagnosis in the Swedish histopathology registers is also warranted.

A very exciting approach to continue exploring the findings in study II regarding the biological mechanisms of esophageal adenocarcinoma development based on histopathology would be to study molecular/genetic/epigenetic differences between non-dysplastic Barrett's esophagus samples in patients who progressed to low-grade dysplasia vs. those who did not progress. Another approach could be to perform a nested case-control among patients with esophageal biopsies stored in Swedish pathology departments several years prior to their cancer diagnosis. Future studies could determine molecular/genetic differences between EAC patients with a prior diagnosis of Barrett's esophagus, compared to patients without a prior diagnosis of Barrett's esophagus.

The continued research regarding the association between gastric atrophy, poor oral health and esophageal squamous cell carcinoma should focus on testing the biological mechanism underlying this association. Preferably by studying the role of altered oral or gastric microbiota in the developments of esophageal squamous cell carcinoma among patients with gastric atrophy verified by endoscopy and histopathology.

## **8.3 PERSONALIZED TREATMENT**

There is substantial work to do within the area of personalized medicine in gastroesophageal cancer patients.

To begin with, a validation study of the findings in study IV using the nation-wide NREV register linked to nationwide data from the SALT register would be feasible and justified. Different treatment strategies and chemotherapy choices across the country might not make it possible to validate our findings from the Stockholm-Gotland region, but important conclusions could be drawn regarding the survival outcomes from past chemotherapy treatment. In such a study, a better control for delivered surgery and oncology treatments could hopefully be reached.

The INCA register is managed at the Regional Cancer Centers in Sweden and contains data about health care delivered to cancer patients, including the delivery of new chemotherapies

reported from the care-givers since 2018. The INCA register provides an excellent platform for future comparisons of the survival outcomes of gastroesophageal cancer patients who received various chemotherapy regimens.

There is a major gap between the current knowledge of dose optimization in chemotherapy and the clinical practice. Future clinical trials evaluating the survival benefit of therapeutic drug monitoring for certain chemotherapy drugs given to gastroesophageal cancer patients are much needed. In addition, it would be a major achievement to find clinically useful biomarkers that can predict responders and non-responders to chemotherapy prior to treatment. Prospective clinical validation studies of suggested biomarkers from research studies could potentially be very valuable.

## **9 POPULÄRVETENSKAPLIG SAMMANFATTNING**

Cancer i magsäcken och matstrupen hör till de dödligaste cancerformerna i världen. Det är bara en av tre patienter med magsäckscancer och en av sex patienter med matstrupscancer som lever fem år efter sin diagnos. Överlevnaden är avsevärt bättre bland patienter där man kan operera bort tumören, men majoriteten av patienterna, så många som 70-80% har så pass spridd cancer att detta inte är möjligt. Därför får de flesta patienter med spridd tumör endast lindrande eller ingen behandling. Både patienter som har en utsikt att bli botade och patienter som planeras för lindrande behandling kan få cellgifter. Tyvärr följer man sällan upp överlevnaden bland patienter som fått olika cellgifter.

Det är högst önskvärt att minska det lidande som cancer i magsäcken och matstrupen orsakar. En hörnsten i att minska sjukdomsburden är att förstå vad som orsakar dessa sjukdomar. Även om forskningen har ökat förståelsen för nödvändiga omständigheter för cancerutveckling så vet vi fortfarande inte orsaken till cancer i magsäcken och matstrupen.

Målet med den här avhandlingen har varit att utforska förstadier till cancer i magsäcken och matstrupen samt studera hur valet av cellgiftsbehandling påverkar överlevnaden hos patienter med cancer i magsäcken och matstrupen. De fyra studier som ingår i avhandlingen baserades på hälsoregister i Sverige (studie I, II och IV) samt en fältstudie i ett högriskområde i Kina (studie III). I studie I undersöktes hur stor risk familjemedlemmar löper att få magsäckscancer om de har en släkting med förstadier till cancer i magsäcken. I studie II beräknades risken att utveckla cancer i magsäcken och matstrupen hos patienter som hade tagit vävnadsprover från matstrupen. I studie III prövades sambandet mellan slemhinneförtvining i magsäcken och risken för skivepitelcancer i matstrupen. I sista studien jämfördes överlevnaden hos patienter med cancer i magsäcken och matstrupen beroende på vilken cellgiftsbehandling de fått.

Risken för cancer i magsäcken var 50-60 % förhöjd bland individer med familjemedlemmar (föräldrar, syskon eller barn) som hade förstadier till cancer i magsäcken än den allmänna befolkningen i Sverige. Nästa upptäckt var att patienter som har cylindercellsförändringar i matstrupen, löpte en tio gånger ökad risk att utveckla körtelcellcancer i matstrupen jämfört med



den allmänna befolkningen i Sverige, vilket var i nivå med ett sedan tidigare känt förstadium till körtelcellcancer som kallas Barrett-slemhinna. I den tredje studien var risken för skivepitelcancer i matstrupen 60 % högre bland de patienter som hade slemhinneförsvinnning i magsäcken i ett högriskområde i Kina. I den sista studien som genomfördes i Stockholm-Gotland, upptäcktes att patienter som fick botande cellgiftsbehandling med cellgiftskombinationen cisplatin-fluorouracil hade bättre överlevnad om de hade cancer i matstrupen (jämfört med karboplatin-fluorouracil), men sämre överlevnad om de hade cancer i övergången mellan matstrupen och magsäcken (jämfört med fluorouracil-oxaliplatin).

Slutsatsen är därför först och främst att man kan använda uppgifter om förstadium till cancer i magsäcken för att identifiera personer med hög risk att utveckla magsäckscancer i sjukvården. Vidare kan det vara värdefullt att följa upp patienter med cylindercellsförändringar i matstrupen, oavsett diagnosen Barrett-slemhinna eller inte. Vidare, bekräftades det omstridda sambandet mellan slemhinneförsvinnning i magsäcken och risken för skivepitelcancer i matstrupen som ytterligare förstärks av undermålig munhälsa. Sist men inte minst verkar valet av cisplatin-fluorouracil vara förknippat med bättre överlevnad hos patienter med cancer i matstrupen, men sämre överlevnad bland patienter med cancer i övergången mellan matstrupen och magsäcken. Samtliga fynd bör studeras ytterligare innan de omsätts till användning i hälso- och sjukvården.



## 10 ACKNOWLEDGEMENTS

To begin with I would like to thank my supervisors, mentor and Karolinska Institute for making my thesis possible.

Weimin Ye, my main supervisor, thank you for your time, scientific guidance and support.

Olof Nyrén, Amelie Plymoth and Pauline Raaschou, my co-supervisors and Magnus Nilsson, my mentor, for your time, kind advice and support.

Karolinska Institute, and the MD/PhD grant (later CSTEP) for financially supporting part of my PhD.

Furthermore I would like to give a special thanks to my co-authors for sharing your time, resources, knowledge and enthusiasm with me during our projects.

Thank you to my fellow PhD students and colleagues at MEB who have made my PhD journey brighter.

My warmest thanks to my colleagues and friends at Karolinska University Hospital, Clinical Pharmacology for being so supportive of my PhD and interest in research.

My parents and siblings, thank you for being by my side.

My husband and son, thank you for your love.



## 11 REFERENCES

1. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *International journal of cancer Journal international du cancer*. 2015;136(5):E359-86.
2. Asaka M, Mabe K. Strategies for eliminating death from gastric cancer in Japan. *Proc Jpn Acad Ser B Phys Biol Sci*. 2014;90(7):251-8.
3. Hamashima C. Overdiagnosis of gastric cancer by endoscopic screening. *World journal of gastrointestinal endoscopy*. 2017;9(2):55-60.
4. Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A. Global cancer statistics, 2012. *CA Cancer J Clin*. 2015;65(2):87-108.
5. Karimi P, Islami F, Anandasabapathy S, Freedman ND, Kamangar F. Gastric cancer: descriptive epidemiology, risk factors, screening, and prevention. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*. 2014;23(5):700-13.
6. Morita M, Taguchi K, Kagawa M, Nakanoko T, Uehara H, Sugiyama M, et al. Treatment strategies for neuroendocrine carcinoma of the upper digestive tract. *Int J Clin Oncol*. 2020.
7. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2018;68(6):394-424.
8. Corley DA, Kubo A. Influence of site classification on cancer incidence rates: an analysis of gastric cardia carcinomas. *Journal of the National Cancer Institute*. 2004;96(18):1383-7.
9. 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*. 1965;64:31-49.
10. Correa P. A human model of gastric carcinogenesis. *Cancer research*. 1988;48(13):3554-60.
11. Ang TL, Fock KM. Clinical epidemiology of gastric cancer. *Singapore Med J*. 2014;55(12):621-8.
12. Comprehensive molecular characterization of gastric adenocarcinoma. *Nature*. 2014;513(7517):202-9.
13. Cristescu R, Lee J, Nebozhyn M, Kim KM, Ting JC, Wong SS, et al. Molecular analysis of gastric cancer identifies subtypes associated with distinct clinical outcomes. *Nat Med*. 2015;21(5):449-56.
14. Setia N, Agoston AT, Han HS, Mullen JT, Duda DG, Clark JW, et al. A protein and mRNA expression-based classification of gastric cancer. *Mod Pathol*. 2016;29(7):772-84.
15. Ahn S, Lee SJ, Kim Y, Kim A, Shin N, Choi KU, et al. High-throughput Protein and mRNA Expression-based Classification of Gastric Cancers Can Identify Clinically Distinct Subtypes, Concordant With Recent Molecular Classifications. *Am J Surg Pathol*. 2017;41(1):106-15.

16. Baraniskin A, Van Laethem JL, Wyrwicz L, Guller U, Wasan HS, Matysiak-Budnik T, et al. Clinical relevance of molecular diagnostics in gastrointestinal (GI) cancer: European Society of Digestive Oncology (ESDO) expert discussion and recommendations from the 17th European Society for Medical Oncology (ESMO)/World Congress on Gastrointestinal Cancer, Barcelona. *European journal of cancer* (Oxford, England : 1990). 2017;86:305-17.
17. Van Laethem JL, Carneiro F, Ducreux M, Messman H, Lordick F, Ilson DH, et al. The multidisciplinary management of gastro-oesophageal junction tumours: European Society of Digestive Oncology (ESDO): Expert discussion and report from the 16th ESMO World Congress on Gastrointestinal Cancer, Barcelona. *Digestive and liver disease : official journal of the Italian Society of Gastroenterology and the Italian Association for the Study of the Liver*. 2016;48(11):1283-9.
18. Bang Y-J, Van Cutsem E, Feyereislova A, Chung HC, Shen L, Sawaki A, et al. Trastuzumab in combination with chemotherapy versus chemotherapy alone for treatment of HER2-positive advanced gastric or gastro-oesophageal junction cancer (ToGA): a phase 3, open-label, randomised controlled trial. *The Lancet*. 2010;376(9742):687-97.
19. Salem ME, Puccini A, Xiu J, Raghavan D, Lenz HJ, Korn WM, et al. Comparative Molecular Analyses of Esophageal Squamous Cell Carcinoma, Esophageal Adenocarcinoma, and Gastric Adenocarcinoma. *Oncologist*. 2018;23(11):1319-27.
20. Hongo M, Nagasaki Y, Shoji T. Epidemiology of esophageal cancer: Orient to Occident. Effects of chronology, geography and ethnicity. *Journal of gastroenterology and hepatology*. 2009;24(5):729-35.
21. Edgren G, Adami HO, Weiderpass E, Nyren O. A global assessment of the oesophageal adenocarcinoma epidemic. *Gut*. 2013;62(10):1406-14.
22. Howlader N NA, Krapcho M, Miller D, Brest A, Yu M, Ruhl J, Tatalovich Z, Mariotto A, Lewis DR, Chen HS, Feuer EJ, Cronin KA (eds). SEER Cancer Statistics Review, 1975-2016, National Cancer Institute. Bethesda, MD, [https://seer.cancer.gov/csr/1975\\_2016/](https://seer.cancer.gov/csr/1975_2016/), based on November 2018 SEER data submission, posted to the SEER web site, April 2019.
23. Adami HO, Hunter DJ, Lagiou P, Mucci L. *Textbook of Cancer Epidemiology*. 3 ed: Oxford University Press; 2018.
24. Arnold M, Abnet CC, Neale RE, Vignat J, Giovannucci EL, McGlynn KA, et al. Global Burden of 5 Major Types Of Gastrointestinal Cancer. *Gastroenterology*. 2020.
25. Sitarz R, Skierucha M, Mielko J, Offerhaus GJA, Maciejewski R, Polkowski WP. Gastric cancer: epidemiology, prevention, classification, and treatment. *Cancer Manag Res*. 2018;10:239-48.
26. Zhang Y. Epidemiology of esophageal cancer. *World J Gastroenterol*. 2013;19(34):5598-606.
27. De B, Rhome R, Jairam V, Ozbek U, Holcombe RF, Buckstein M, et al. Gastric adenocarcinoma in young adult patients: patterns of care and survival in the United States. *Gastric Cancer*. 2018;21(6):889-99.
28. Zeng Y, Ruan W, Liu J, Liang W, He J, Cui F, et al. Esophageal cancer in patients under 50: a SEER analysis. *Journal of thoracic disease*. 2018;10(5):2542-50.
29. Islami F, DeSantis CE, Jemal A. Incidence Trends of Esophageal and Gastric Cancer Subtypes by Race, Ethnicity, and Age in the United States, 1997-2014. *Clinical*

gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association. 2019;17(3):429-39.

30. Sung H, Siegel RL, Rosenberg PS, Jemal A. Emerging cancer trends among young adults in the USA: analysis of a population-based cancer registry. *Lancet Public Health*. 2019;4(3):e137-e47.
31. Domper Arnal MJ, Ferrandez Arenas A, Lanas Arbeloa A. Esophageal cancer: Risk factors, screening and endoscopic treatment in Western and Eastern countries. *World J Gastroenterol*. 2015;21(26):7933-43.
32. Kim HW, Kim JH, Lim BJ, Kim H, Kim H, Park JJ, et al. Sex Disparity in Gastric Cancer: Female Sex is a Poor Prognostic Factor for Advanced Gastric Cancer. *Ann Surg Oncol*. 2016;23(13):4344-51.
33. Anderson WF, Rabkin CS, Turner N, Fraumeni JF, Jr., Rosenberg PS, Camargo MC. The Changing Face of Noncardia Gastric Cancer Incidence Among US Non-Hispanic Whites. *Journal of the National Cancer Institute*. 2018;110(6):608-15.
34. Rutegard M, Shore R, Lu Y, Lagergren P, Lindblad M. Sex differences in the incidence of gastrointestinal adenocarcinoma in Sweden 1970-2006. *European journal of cancer (Oxford, England : 1990)*. 2010;46(6):1093-100.
35. Qiao YL, Dawsey SM, Kamangar F, Fan JH, Abnet CC, Sun XD, et al. Total and cancer mortality after supplementation with vitamins and minerals: follow-up of the Linxian General Population Nutrition Intervention Trial. *Journal of the National Cancer Institute*. 2009;101(7):507-18.
36. Wang X-Q, Terry P-D, Yan H. Review of salt consumption and stomach cancer risk: epidemiological and biological evidence. *World J Gastroenterol*. 2009;15(18):2204-13.
37. Lunet N, Lacerda-Vieira A, Barros H. Fruit and vegetables consumption and gastric cancer: a systematic review and meta-analysis of cohort studies. *Nutrition and cancer*. 2005;53(1):1-10.
38. Zhao Z, Yin Z, Zhao Q. Red and processed meat consumption and gastric cancer risk: a systematic review and meta-analysis. *Oncotarget*. 2017;8(18):30563-75.
39. Calmels S, Dalla Venezia N, Bartsch H. Isolation of an enzyme catalysing nitrosamine formation in *Pseudomonas aeruginosa* and *Neisseria mucosae*. *Biochem Biophys Res Commun*. 1990;171(2):655-60.
40. Ye WM, Held M, Lagergren J, Engstrand L, Blot WJ, McLaughlin JK, et al. *Helicobacter pylori* infection and gastric atrophy: Risk of adenocarcinoma and squamous-cell carcinoma of the esophagus and adenocarcinoma of the gastric cardia. *Journal of the National Cancer Institute*. 2004;96(5):388-96.
41. Akiyama T, Inamori M, Iida H, Endo H, Hosono K, Yoneda K, et al. Macroscopic extent of gastric mucosal atrophy: increased risk factor for esophageal squamous cell carcinoma in Japan. *BMC Gastroenterol*. 2009;9:6.
42. Iijima K, Koike T, Abe Y, Inomata Y, Sekine H, Imatani A, et al. Extensive gastric atrophy: An increased risk factor for superficial esophageal squamous cell carcinoma in Japan. *Am J Gastroenterol*. 2007;102(8):1603-9.
43. Almodova ED, de Oliveira WK, Machado LFA, Grejo JR, da Cunha TR, Colaiacovo W, et al. Atrophic gastritis: Risk factor for esophageal squamous cell carcinoma in a Latin-American population. *World J Gastroenterol*. 2013;19(13):2060-4.

44. Iijima K, Koike T, Abe Y, Yamagishi H, Ara N, Asanuma K, et al. Gastric Hyposecretion in Esophageal Squamous-Cell Carcinomas. *Dig Dis Sci*. 2010;55(5):1349-55.
45. Uno K, Iijima K, Hatta W, Koike T, Abe Y, Asano N, et al. Direct Measurement of Gastroesophageal Reflux Episodes in Patients With Squamous Cell Carcinoma by 24-h pH-Impedance Monitoring. *Am J Gastroenterol*. 2011;106(11):1923-9.
46. de Vries AC, Capelle LG, Looman CWN, van Blankenstein M, van Grieken NCT, Casparie MK, et al. Increased risk of esophageal squamous cell carcinoma in patients with gastric atrophy: Independent of the severity of atrophic changes. *International Journal of Cancer*. 2009;124(9):2135-8.
47. Venerito M, Kohrs S, Wex T, Adolf D, Kuester D, Schubert D, et al. Helicobacter pylori infection and fundic gastric atrophy are not associated with esophageal squamous cell carcinoma: a case-control study. *European journal of gastroenterology & hepatology*. 2011;23(10):859-64.
48. Ren JS, Kamangar F, Qiao YL, Taylor PR, Liang H, Dawsey SM, et al. Serum pepsinogens and risk of gastric and oesophageal cancers in the General Population Nutrition Intervention Trial cohort. *Gut*. 2009;58(5):636-42.
49. Abnet CC, Arnold M, Wei WQ. Epidemiology of Esophageal Squamous Cell Carcinoma. *Gastroenterology*. 2018;154(2):360-73.
50. Chen X, Yuan Z, Lu M, Zhang Y, Jin L, Ye W. Poor oral health is associated with an increased risk of esophageal squamous cell carcinoma - a population-based case-control study in China. *International journal of cancer Journal international du cancer*. 2017;140(3):626-35.
51. Nasrollahzadeh D, Malekzadeh R, Aghcheli K, Sotoudeh M, Merat S, Islami F, et al. Gastric atrophy and oesophageal squamous cell carcinoma: possible interaction with dental health and oral hygiene habit. *British journal of cancer*. 2012;107(5):888-94.
52. Ye W, Held M, Lagergren J, Engstrand L, Blot WJ, McLaughlin JK, et al. Helicobacter pylori infection and gastric atrophy: risk of adenocarcinoma and squamous-cell carcinoma of the esophagus and adenocarcinoma of the gastric cardia. *Journal of the National Cancer Institute*. 2004;96(5):388-96.
53. Infection with Helicobacter pylori. *IARC Monogr Eval Carcinog Risks Hum*. 1994;61:177-240.
54. Ford Alexander C, Forman D, Hunt R, Yuan Y, Moayyedi P. Helicobacter pylori eradication for the prevention of gastric neoplasia. *Cochrane Database of Systematic Reviews* [Internet]. 2015; (7). Available from: <http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD005583.pub2/abstract>  
<http://onlinelibrary.wiley.com/store/10.1002/14651858.CD005583.pub2/asset/CD005583.pdf?v=1&t=igktc59s&s=cdb0c8425b246d714b84adbfcf2763d8defe1b12>.
55. Carr JS, Zafar SF, Saba N, Khuri FR, El-Rayes BF. Risk factors for rising incidence of esophageal and gastric cardia adenocarcinoma. *J Gastrointest Cancer*. 2013;44(2):143-51.
56. Tobacco smoke and involuntary smoking. *IARC Monogr Eval Carcinog Risks Hum*. 2004;83:1-1438.
57. Personal habits and indoor combustions. Volume 100 E. A review of human carcinogens. *IARC Monogr Eval Carcinog Risks Hum*. 2012;100(Pt E):1-538.



58. Tredaniel J, Boffetta P, Buiatti E, Saracci R, Hirsch A. Tobacco smoking and gastric cancer: review and meta-analysis. *International journal of cancer Journal international du cancer*. 1997;72(4):565-73.
59. Ladeiras-Lopes R, Pereira AK, Nogueira A, Pinheiro-Torres T, Pinto I, Santos-Pereira R, et al. Smoking and gastric cancer: systematic review and meta-analysis of cohort studies. *Cancer causes & control : CCC*. 2008;19(7):689-701.
60. La Torre G, Chiaradia G, Gianfagna F, De Lauretis A, Boccia S, Mannocci A, et al. Smoking status and gastric cancer risk: an updated meta-analysis of case-control studies published in the past ten years. *Tumori*. 2009;95(1):13-22.
61. Nishino Y, Inoue M, Tsuji I, Wakai K, Nagata C, Mizoue T, et al. Tobacco smoking and gastric cancer risk: an evaluation based on a systematic review of epidemiologic evidence among the Japanese population. *Japanese journal of clinical oncology*. 2006;36(12):800-7.
62. de Martel C, Forman D, Plummer M. Gastric cancer: epidemiology and risk factors. *Gastroenterology clinics of North America*. 2013;42(2):219-40.
63. Nagel G, Linseisen J, Boshuizen HC, Pera G, Del Giudice G, Westert GP, et al. Socioeconomic position and the risk of gastric and oesophageal cancer in the European Prospective Investigation into Cancer and Nutrition (EPIC-EURGAST). *International journal of epidemiology*. 2007;36(1):66-76.
64. Lindblad M, Lagergren J, Garcia Rodriguez LA. Nonsteroidal anti-inflammatory drugs and risk of esophageal and gastric cancer. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*. 2005;14(2):444-50.
65. Liao LM, Vaughan TL, Corley DA, Cook MB, Casson AG, Kamangar F, et al. Nonsteroidal anti-inflammatory drug use reduces risk of adenocarcinomas of the esophagus and esophagogastric junction in a pooled analysis. *Gastroenterology*. 2012;142(3):442-52.e5; quiz e22-3.
66. Brusselaers N, Lagergren J, Engstrand L. Duration of use of proton pump inhibitors and the risk of gastric and oesophageal cancer. *Cancer Epidemiol*. 2019;62:101585.
67. Bateman DN, Colin-Jones D, Hartz S, Langman M, Logan RF, Mant J, et al. Mortality study of 18 000 patients treated with omeprazole. *Gut*. 2003;52(7):942-6.
68. Lagergren J. Etiology and risk factors for oesophageal adenocarcinoma: possibilities for chemoprophylaxis? Best practice & research *Clinical gastroenterology*. 2006;20(5):803-12.
69. Oliveira C, Pinheiro H, Figueiredo J, Seruca R, Carneiro F. Familial gastric cancer: genetic susceptibility, pathology, and implications for management. *Lancet Oncol*. 2015;16(2):e60-70.
70. Zanghieri G, Di Gregorio C, Sacchetti C, Fante R, Sassatelli R, Cannizzo G, et al. Familial occurrence of gastric cancer in the 2-year experience of a population-based registry. *Cancer*. 1990;66(9):2047-51.
71. La Vecchia C, Negri E, Franceschi S, Gentile A. Family history and the risk of stomach and colorectal cancer. *Cancer*. 1992;70(1):50-5.

72. El-Omar EM, Oien K, Murray LS, El-Nujumi A, Wirz A, Gillen D, et al. Increased prevalence of precancerous changes in relatives of gastric cancer patients: critical role of H. pylori. *Gastroenterology*. 2000;118(1):22-30.
73. Oh S, Kim N, Yoon H, Choi YJ, Lee JY, Park KJ, et al. Risk factors of atrophic gastritis and intestinal metaplasia in first-degree relatives of gastric cancer patients compared with age-sex matched controls. *Journal of cancer prevention*. 2013;18(2):149-60.
74. You WC, Ma JL, Liu W, Gail MH, Chang YS, Zhang L, et al. Blood type and family cancer history in relation to precancerous gastric lesions. *International journal of epidemiology*. 2000;29(3):405-7.
75. Choi YJ, Kim N. Gastric cancer and family history. *Korean J Intern Med*. 2016;31(6):1042-53.
76. Negovan A, Iancu M, Fulop E, Banescu C. Helicobacter pylori and cytokine gene variants as predictors of premalignant gastric lesions. *World J Gastroenterol*. 2019;25(30):4105-24.
77. Song H, Ekheden IG, Ploner A, Ericsson J, Nyren O, Ye W. Family history of gastric mucosal abnormality and the risk of gastric cancer: a population-based observational study. *International journal of epidemiology*. 2018;47(2):440-9.
78. Palli D, Galli M, Caporaso NE, Cipriani F, Decarli A, Saieva C, et al. Family history and risk of stomach cancer in Italy. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*. 1994;3(1):15-8.
79. Whitmire JM, Merrell DS. Helicobacter pylori Genetic Polymorphisms in Gastric Disease Development. *Advances in experimental medicine and biology*. 2019;1149:173-94.
80. Schoofs N, Bisschops R, Prenen H. Progression of Barrett's esophagus toward esophageal adenocarcinoma: an overview. *Ann Gastroenterol*. 2017;30(1):1-6.
81. Theron BT, Padmanabhan H, Aladin H, Smith P, Campbell E, Nightingale P, et al. The risk of oesophageal adenocarcinoma in a prospectively recruited Barrett's oesophagus cohort. *United European Gastroenterol J*. 2016;4(6):754-61.
82. Dulai GS, Guha S, Kahn KL, Gornbein J, Weinstein WM. Preoperative prevalence of Barrett's esophagus in esophageal adenocarcinoma: a systematic review. *Gastroenterology*. 2002;122(1):26-33.
83. Sikkema M, de Jonge PJ, Steyerberg EW, Kuipers EJ. Risk of esophageal adenocarcinoma and mortality in patients with Barrett's esophagus: a systematic review and meta-analysis. *Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association*. 2010;8(3):235-44; quiz e32.
84. Old O, Moayyedi P, Love S, Roberts C, Hapeshi J, Foy C, et al. Barrett's Oesophagus Surveillance versus endoscopy at need Study (BOSS): protocol and analysis plan for a multicentre randomized controlled trial. *J Med Screen*. 2015;22(3):158-64.
85. Taylor PR, Abnet CC, Dawsey SM. Squamous dysplasia--the precursor lesion for esophageal squamous cell carcinoma. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*. 2013;22(4):540-52.

86. Jain S, Dhingra S. Pathology of esophageal cancer and Barrett's esophagus. *Ann Cardiothorac Surg*. 2017;6(2):99-109.
87. Shimizu M, Ban S, Odze RD. Squamous dysplasia and other precursor lesions related to esophageal squamous cell carcinoma. *Gastroenterology clinics of North America*. 2007;36(4):797-811, v-vi.
88. Makady A, de Boer A, Hillege H, Klungel O, Goettsch W. What Is Real-World Data? A Review of Definitions Based on Literature and Stakeholder Interviews. *Value Health*. 2017;20(7):858-65.
89. Matsuoka T, Yashiro M. Precision medicine for gastrointestinal cancer: Recent progress and future perspective. *World journal of gastrointestinal oncology*. 2020;12(1):1-20.
90. Bonelli P, Borrelli A, Tuccillo FM, Silvestro L, Palaia R, Buonaguro FM. Precision medicine in gastric cancer. *World journal of gastrointestinal oncology*. 2019;11(10):804-29.
91. Ho SWT, Tan P. Dissection of gastric cancer heterogeneity for precision oncology. *Cancer Sci*. 2019;110(11):3405-14.
92. Kang X, Chen K, Li Y, Li J, D'Amico TA, Chen X. Personalized targeted therapy for esophageal squamous cell carcinoma. *World J Gastroenterol*. 2015;21(25):7648-58.
93. Barlow L, Westergren K, Holmberg L, Talback M. The completeness of the Swedish Cancer Register: a sample survey for year 1998. *Acta oncologica (Stockholm, Sweden)*. 2009;48(1):27-33.
94. Brooke HL, Talback M, Hornblad J, Johansson LA, Ludvigsson JF, Druid H, et al. The Swedish cause of death register. *Eur J Epidemiol*. 2017;32(9):765-73.
95. Johansson LA, Bjorkenstam C, Westerling R. Unexplained differences between hospital and mortality data indicated mistakes in death certification: an investigation of 1,094 deaths in Sweden during 1995. *J Clin Epidemiol*. 2009;62(11):1202-9.
96. Ludvigsson JF, Andersson E, Ekbom A, Feychting M, Kim JL, Reuterwall C, et al. External review and validation of the Swedish national inpatient register. *BMC public health*. 2011;11:450.
97. Agren G, Jakobsson SW. Validation of diagnoses on death certificates for male alcoholics in Stockholm. *Forensic Sci Int*. 1987;33(4):231-41.
98. Dahlen E, Almqvist C, Bergstrom A, Wettermark B, Kull I. Factors associated with concordance between parental-reported use and dispensed asthma drugs in adolescents: findings from the BAMSE birth cohort. *Pharmacoepidemiol Drug Saf*. 2014;23(9):942-9.
99. Ludvigsson JF, Almqvist C, Bonamy AK, Ljung R, Michaelsson K, Neovius M, et al. Registers of the Swedish total population and their use in medical research. *Eur J Epidemiol*. 2016;31(2):125-36.
100. Ludvigsson JF, Svedberg P, Olen O, Bruze G, Neovius M. The longitudinal integrated database for health insurance and labour market studies (LISA) and its use in medical research. *Eur J Epidemiol*. 2019;34(4):423-37.
101. Cai L, Mu LN, Lu H, Lu QY, You NC, Yu SZ, et al. Dietary selenium intake and genetic polymorphisms of the GSTP1 and p53 genes on the risk of esophageal squamous cell carcinoma. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*. 2006;15(2):294-300.

102. Coggon D, Rose G, Barker DJP. Epidemiology for the uninitiated. 5 ed. London: BMJ Books; 2003.
103. Borgan Ø. Nelson–Aalen Estimator. In: Armitage P, Colton T, editors. Encyclopedia of Biostatistics. 2 ed: John Wiley & Sons, Ltd.; 2005.
104. Pearce N. Analysis of matched case-control studies. BMJ (Clinical research ed). 2016;352:i969.
105. Vittinghoff E, Shiboski SC, Glidden DV, McCulloch CE. Regression Methods in Biostatistics: Linear, Logistic, Survival, and Repeated Measures Models. Gail M, Krickeberg K, Samet J, Tsiatis A, Wong W, editors. New York, NY: New York, NY: Springer New York; 2005.
106. Rothman KJ. Epidemiology: An introduction. USA: Oxford University Press; 2002.
107. Knol MJ, VanderWeele TJ, Groenwold RHH, Klungel OH, Rovers MM, Grobbee DE. Estimating measures of interaction on an additive scale for preventive exposures. European journal of epidemiology. 2011;26(6):433-8.
108. Andersson T, Alfredsson L, Kallberg H, Zdravkovic S, Ahlbom A. Calculating measures of biological interaction. Eur J Epidemiol. 2005;20(7):575-9.
109. Ekheden I, Yang X, Chen H, Chen X, Yuan Z, Jin L, et al. Gastric Atrophy and its Interaction with Poor Oral Health Elevate the Risk for Esophageal Squamous Cell Carcinoma in a High-risk Region of China: a Population-based Case-control Study. Am J Epidemiol. 2020.
110. Ekheden I, Ebrahim F, Olafsdottir H, Raaschou P, Wettermark B, Henriksson R, et al. Survival of esophageal and gastric cancer patients with adjuvant and palliative chemotherapy-a retrospective analysis of a register-based patient cohort. Eur J Clin Pharmacol. 2020.
111. Song H, Ekheden IG, Zheng Z, Ericsson J, Nyren O, Ye W. Incidence of gastric cancer among patients with gastric precancerous lesions: observational cohort study in a low risk Western population. BMJ (Clinical research ed). 2015;351:h3867.
112. Lichtenstein P, Holm NV, Verkasalo PK, Iliadou A, Kaprio J, Koskenvuo M, et al. Environmental and heritable factors in the causation of cancer--analyses of cohorts of twins from Sweden, Denmark, and Finland. The New England journal of medicine. 2000;343(2):78-85.
113. Lott PC, Carvajal-Carmona LG. Resolving gastric cancer aetiology: an update in genetic predisposition. Lancet Gastroenterol Hepatol. 2018;3(12):874-83.
114. Andersson TM, Weiderpass E, Engholm G, Lund AQ, Olafsdottir E, Pukkala E, et al. Avoidable cancer cases in the Nordic countries - The impact of overweight and obesity. European journal of cancer (Oxford, England : 1990). 2017;79:106-18.
115. Koulis A, Buckle A, Boussioutas A. Premalignant lesions and gastric cancer: Current understanding. World journal of gastrointestinal oncology. 2019;11(9):665-78.
116. Huang KK, Ramnarayanan K, Zhu F, Srivastava S, Xu C, Tan ALK, et al. Genomic and Epigenomic Profiling of High-Risk Intestinal Metaplasia Reveals Molecular Determinants of Progression to Gastric Cancer. Cancer Cell. 2018;33(1):137-50.e5.

117. den Hollander WJ, Holster IL, den Hoed CM, Capelle LG, Tang TJ, Anten MP, et al. Surveillance of premalignant gastric lesions: a multicentre prospective cohort study from low incidence regions. *Gut*. 2019;68(4):585-93.
118. Hvid-Jensen F, Pedersen L, Drewes AM, Sorensen HT, Funch-Jensen P. Incidence of adenocarcinoma among patients with Barrett's esophagus. *The New England journal of medicine*. 2011;365(15):1375-83.
119. Bhat SK, McManus DT, Coleman HG, Johnston BT, Cardwell CR, McMenamin U, et al. Oesophageal adenocarcinoma and prior diagnosis of Barrett's oesophagus: a population-based study. *Gut*. 2015;64(1):20-5.
120. Ronkainen J, Aro P, Storskrubb T, Johansson SE, Lind T, Bolling-Sternevald E, et al. Prevalence of Barrett's esophagus in the general population: an endoscopic study. *Gastroenterology*. 2005;129(6):1825-31.
121. Dawsey SM, Lewin KJ, Wang GQ, Liu FS, Nieberg RK, Yu Y, et al. Squamous esophageal histology and subsequent risk of squamous cell carcinoma of the esophagus. A prospective follow-up study from Linxian, China. *Cancer*. 1994;74(6):1686-92.
122. den Hoed CM, Holster IL, Capelle LG, de Vries AC, den Hartog B, Ter Borg F, et al. Follow-up of premalignant lesions in patients at risk for progression to gastric cancer. *Endoscopy*. 2013;45(4):249-56.
123. Miwata T, Quach DT, Hiyama T, Aoki R, Le HM, Tran PL, et al. Interobserver and intraobserver agreement for gastric mucosa atrophy. *BMC Gastroenterol*. 2015;15:95.
124. Montagnani F, Turrisi G, Marinozzi C, Aliberti C, Fiorentini G. Effectiveness and safety of oxaliplatin compared to cisplatin for advanced, unresectable gastric cancer: a systematic review and meta-analysis. *Gastric Cancer*. 2011;14(1):50-5.