From the DEPARTMENT OF
MEDICAL EPIDEMIOLOGY AND BIOSTATISTICS
Karolinska Institutet, Stockholm, Sweden

# SEX DIFFERENCES IN CANCER RISK AND SURVIVAL

Cecilia Radkiewicz


Karolinska Institutet

Stockholm 2019

Published by Karolinska Institutet.
Printed by Eprint AB 2019

ISBN 978-91-7831-573-4

Institutionen för medicinsk epidemiologi och biostatistik

# Sex differences in cancer risk and survival

AKADEMISK AVHANDLING

som för avläggande av medicine doktorsexamen vid Karolinska Institutet offentligen försvaras i föreläsningssal Petrén, Nobels väg 12B, Karolinska Institutet, Solna
**Fredagen den 6 december 2019, kl 09.00**

av


## Cecilia Radkiewicz


*Huvudhandledare:*
Docent Gustaf Edgren, MD, PhD
Karolinska Institutet
Institutionen för medicin, Solna
Enheten för klinisk epidemiologi

*Bihandledare:*
Professor Paul W Dickman, PhD
Karolinska Institutet
Institutionen för medicinsk epidemiologi och biostatistik

Professor Mats Lambe, MD, PhD
Karolinska Institutet
Institutionen för medicinsk epidemiologi och biostatistik

Doktor Anna LV Johansson, PhD
Karolinska Institutet
Institutionen för medicinsk epidemiologi och biostatistik

*Opponent:*
Doctor Michael B Cook, PhD
National Cancer Institute
Division of Cancer Epidemiology & Genetics

*Betygsnämnd:*
Docent Christofer Juhlin, MD, PhD
Karolinska Institutet
Institutionen för onkologi-patologi

Professor Michael Bretthauer, MD, PhD
University of Oslo
Faculty of Medicine, Institute of Health and Society

Docent Caroline Nordenvall, MD, PhD
Karolinska Institutet
Institutionen för molekylär medicin och kirurgi


Stockholm 2019

To Leon, for keeping me sane and focused in a chaotic world.

# ABSTRACT

*Aim:* The objective of **study I** was to delineate and quantify sex differences in cancer risk and survival together with assessing the potential gain achieved by eliminating the excess cancer risk in men. **Study II** and **III** aimed to in detail characterize the superior non-small cell lung cancer survival and the inferior urinary bladder cancer survival, in women, with the underlying objective to identify underlying drivers to these two phenomena. In **study IV** we wanted to explore to what extent taller body stature can explain the excess cancer risk in men.

*Methods:* All of the studies are Swedish population-based cohort studies. **Study I** included all incident cancer cases (n=872,397) recorded in the Swedish Cancer Register in 1970-2014 at age 15-84. The association between sex and cancer risk and sex and cancer survival was assessed by estimating male-to-female incidence rate ratios (IRRs) and excess mortality ratios (EMRs), respectively, using Poisson regression models adjusted for age and calendar year. All incident lung squamous cell carcinoma (n=10,325) and adenocarcinoma (n=23,465) cases recorded in the Swedish Lung Cancer Register in 2002-2016 formed the basis in **Study II**. Flexible parametric models were applied to compute adjusted female-to-male hazard ratios (HRs) and standardized survival proportions over follow-up, including; age, year, education, marital status, birth country, health care region, ECOG performance status, smoking history, comorbidity, TNM stage, and tumor location, in the final model. A subgroup analysis of lung adenocarcinoma, additionally adjusting for EGFR mutational status, was additionally performed. In **study III** we included all records of urothelial bladder cancer diagnosed in 1997-2014 at age 18-89 in the Swedish Urinary Bladder Cancer Register (n=36,344). We estimated empirical survival proportions and mortality rates in men and women as well as female-to-male adjusted HRs and standardized survival proportions, using flexible parametric models including; age, year, WHO grade, TNM stage, marital status, education, health care region, birth country, and comorbidity, in the fully-adjusted models. In **study IV** individual-level information on height from the Swedish Passport Register, the Conscription Register, and the Medical Birth Register (n=6,156,659) was linked to the Swedish Cancer Register where 285,778 cancer cases were identified. Contemporary mediation analysis was applied to assess the effect of male sex, explained by height, on cancer risk.

*Results:* In **study I** we found that men are at a higher risk of 34 of 39 malignancies, and have a poorer survival in 27 of 39. Except for smoking-associated malignancies, the excess risk in men is stable over calendar time. In male predominant sites, IRRs range from 1.05; 95% CI, 1.02-1.1 (lung adenocarcinoma) to 8.0; 95% CI, 7.5-85 (laryngeal cancer). Women with non-small cell lung cancer (**study II**) are younger, smoke less, and present better performance status, compared to men. Women with lung adenocarcinoma additionally present lower comorbidity burden, less advanced stage, and more often harbor activating EGFR mutations. Women with non-small cell lung cancer have a superior survival that is most consistent in lung adenocarcinoma where female-to-male HRs ranged from 0.69; 95% CI 0.63-0.76 (stage IA-IIB) to 0.94; 95% CI 0.88-0.99 (stage IIIB-IV). HR estimates remain largely unchanged after meticulous adjustments. Except for an unfavorable stage distribution in women, we

found sparse evidence of sex differences in clinical management or tumor aggressiveness, in urothelial bladder cancer (**study III**). Women, overall, have a poorer bladder cancer survival (adjusted HR 1.15; 95% CI 1.08-1.23) which is driven by muscle invasive tumors (adjusted HR 1.24; 95% CI 1.14-1.34) and restricted to the first two years from diagnosis. **Study IV** confirmed that a majority of investigated cancers are associated with male sex (here, 33 of 39) and body height (27 of 39). A fair proportion of the excess male cancer risk is explained by taller body stature, and ranges from 0.5% (laryngeal) to 100% (salivary, colon, melanoma, and AML). The effect of body height and the mediated effect through height on cancer risk are most consistent in cancers with weak or no established risk factors.

*Conclusion:* In **Study I** we found that male sex is a consistent risk as well as a negative prognostic factor for a majority of cancers. Identifying and eliminating underlying factors to the excess cancer risk in men could substantially reduce the global cancer burden. Men with lung adenocarcinoma have a consistently poorer survival that remained largely unchanged after adjustments for a range of prognostic factors, indicating sex differences in tumor biology (**study II**). The excess bladder cancer mortality in women is limited to muscle-invasive tumors, only noticeable within the first two years from diagnosis, and cannot be explained by the examined clinicopathological factors (**study III**). This warrants further investigation of sex differences in outcomes and complications to radical cystectomy. A large proportion of the excess cancer risk in men is explainable by height (**study IV**). This finding corroborate that a considerable proportion of cancer cases are a result of random processes during DNA replication (i.e., bad luck) rather than underlying hereditary and/or environmental factors.

# LIST OF SCIENTIFIC PAPERS

I. Radkiewicz C, Johansson ALV, Dickman PW, Lambe M, Edgren G. Sex differences in cancer risk and survival: A Swedish cohort study. European journal of cancer (Oxford, England : 1990). 2017;84:130-40.

II. Radkiewicz C, Dickman PW, Johansson ALV, Wagenius G, Edgren G, Lambe M. Sex and survival in non-small cell lung cancer: A nationwide cohort study. PloS one. 2019;14(6):e0219206.

III. Radkiewicz C, Edgren G, Johansson ALV, Jahnson S, Häggström C, Akre O, Dickman PW. Sex differences in urothelial bladder cancer survival. Clinical Genitourinary Cancer. In press 2019.

IV. Radkiewicz C, Edgren G, Sjölander E, Benyi E, Lambe E, Dickman E, Sävendahl L. Can body size explain the excess cancer risk in men? Manuscript.

# CONTENTS

# LIST OF ABBREVIATIONS

| | |
|---|---|
| ADC | adenocarcinoma |
| AIC | Akaike information criterion |
| ALK | anaplastic lymphoma kinase |
| ALL | acute lymphoblastic leukemia |
| AML | acute myeloid leukemia |
| aHR | adjusted hazard ratio |
| aOR | adjusted odds ratio |
| ATC | Anatomical Therapeutic Chemical |
| BladderBaSe | Bladder Cancer Data Base Sweden |
| BMI | body mass index |
| CANC/24.1 | World Health Organization Histological Classification of Neoplasms |
| CI | confidence interval |
| CCI | Charlson comorbidity index |
| CLL | chronic lymphocytic leukemia |
| CML | chronic myeloid leukemia |
| CT | computed tomography |
| CYP1A1 | cytochrome P450, family 1, subfamily A, polypeptide 1 |
| DAG | directed acyclic graph |
| df | degrees of freedom |
| ECOG | Eastern Cooperative Oncology Group |
| EGFR | endothelial growth factor receptor |
| EMR | excess mortality ratio |
| G1/G2/G3 | highly/intermediate/poorly differentiated urothelial cancer cells |
| GRPR | gastrin-releasing peptide receptor |
| GTSM1 | glutathione S transferase mu 1 |
| HPV | human papilloma virus |

| | |
|---|---|
| HR | hazard ratio |
| ICD | International Classification of Diseases |
| ICD-O | International Classification of Diseases for Oncology |
| IR | incidence rate |
| IRR | incidence rate ratio |
| LCBaSe | Lung Cancer Database Sweden |
| LISA | Longitudinal Integration Database for Health Insurance, Labor Market Studies |
| LMP | low malignant potential |
| MIBC | muscle invasive bladder cancer |
| MRI | magnetic resonance imaging |
| MSI | microsatellite instability |
| NHL | non-Hodgkin lymphoma |
| NMIBC | non-muscle invasive bladder cancer |
| NLCR | National Lung Cancer Register |
| NOS | not otherwise specified |
| NPR | National Patient Register |
| NRN | Swedish national registration number |
| NSCLC | non-small cell lung cancer |
| OR | odds ratio |
| PAR% | population attributable risk percent |
| PE | proportion explained by mediation |
| PET | positron emission tomography |
| RCC | Swedish Regional Cancer Center |
| RR | relative risk |
| RSR | relative survival ratio |
| SCC | squamous cell carcinoma |
| SCLC | small-cell lung cancer |

| | |
|---|---|
| SCR | Swedish Cancer Register |
| SD | standard deviation |
| SLL | small lymphocytic lymphoma |
| SNRUBC | Swedish National Register of Urinary Bladder Cancer |
| TNM | Tumor-Node-Metastasis |
| TURBT | transurethral resection of the bladder tumor |
| UBC | urothelial bladder cancer |
| UV | ultraviolet |

# 1 BACKGROUND

## 1.1 INTRODUCTION

Patient sex is simple to assess and a persistent predictor of disease risk, treatment response, and prognosis in multiple medical conditions, but is still rarely considered systematically in the clinical practice (1). Male or female sex are biological traits defined by the karyotype 46XY and 46XX, respectively, and are reflected by anatomical and physiological differences. Gender, on the other hand, indicates different behavior in men and women depending on social and cultural context (2). There is no validated tool to assess gender (1, 2). In other words, sex considers biological differences like sex hormone levels, reproductive organs, and secondary sex characteristics, while gender reflects behavior, including unequal environmental exposure and health care utilization (1, 2). Sex and gender are hereafter used according to these definitions, but the main focus throughout this thesis will be on biological differences between men and women affecting cancer risk and prognosis (2).

It is well-known that men, compared to women, are both at increased risk of and have a poorer prognosis in most malignancies affecting both sexes (3-11). The former has traditionally been attributed to gender differences in exposure to environmental risk factors, mainly tobacco smoking, alcohol consumption, and occupational carcinogens (12, 13). The male survival disadvantage has remained enigmatic, although a common belief is that men in general present with more advanced cancer stage at diagnosis resulting in poorer outcomes (8). The following section is a literature review with the purpose to give background to and motivate the performed studies as well as to justify some of the decisions made regarding definitions and stratifications described subsequently.

## 1.2 SEX AND CANCER RISK

### 1.2.1 Epidemiology

The excess cancer risk in men is a consistent finding in epidemiological studies and has been acknowledged for a long time (3-7, 14). The underlying drivers to the increased cancer susceptibility in men have however received less attention. Cancer registers usually report results stratified by sex and male sex is commonly mentioned as a risk factor for various cancers, without engaging in the rationale behind this phenomenon. Breast and genital cancer excluded, men have a higher probability of being diagnosed with a majority of malignancies, across geographic region (3, 5, 6, 15). Only a few sites seem to be more common in women; biliary, anal, meningioma, and thyroid cancer (3, 5-7). The remaining 30-40 cancers (the exact number depends on how finely sites are subdivided) are more common in men. The male relative risk for the sites with the largest male overbalance is 2- to 8-fold (Table 1) (3, 5-7). Even though the male excess risk in general and in particular in sites with a strong association to tobacco and alcohol appears to decrease over time, the risk remains elevated (5). Some malignancies, including urinary bladder, pleural, and the hematological malignancies, present a remarkably consistent pattern over calendar time and unequal

exposure to environmental risk factors can hardly explain the male overbalance in these sites (3, 5-7).

There are few comprehensive reports on sex and cancer risk by age. But, the male overbalance appears to be present before puberty, increase over age, and culminate in elderly when most malignancies peak (5, 7). Anti-oxidative properties of female sex hormones have been proposed to protect women from cancer (16). The observed continuous relative risk increase in men, many years after female menopause and rapidly declining female sex hormones, is inconsistent with this conjecture. Sex differences in cancer risk in children (mainly hematological malignancies) can hardly be explained by sex hormones or environmental factors since these do not differ substantially in boys and girls. That said, the drivers behind sex discrepancies in cancer risk can of course differ in children and adults.

*Table 1. Male-to-female incidence rate ratio (IRR) in cancer sites with the most consistently reported excess male relative risk. Data extracted from four publications on sex differences in cancer risk.*

| **Author, year:** | *Ashley, 1968 (4)* | *Cook, 2009 (5)* | *Biggar, 2009 (7)* | *Edgren, 2012 (6)* |
|---|---|---|---|---|
| **Geographic region:** | CI5 (60 countries) | SEER-9 (US) | Denmark | CI5 (60 countries) |
| **Calendar period:** | 1962-1964 | 1975-2004 | 1943-2003 | 1998-2002 |
| **Cancer site** | **male-to-female IRR[1]** | **male-to-female IRR[1]** | **male-to-female IRR[2]** | **male-to-female IRR[3]** |
| Larynx | 9.09 | 5.17 | 5.82 | 6.36 |
| Pleura | - | - | 3.87 | - |
| Esophagus adenocarcinoma | 2.78 | 7.64 | 2.92 | 3.72 |
| Esophagus squamous cell | | 2.57 | | |
| Lip | 11.11 | 7.16 | 8.08 | 4.25 |
| Urinary bladder | 3.57 | 3.92 | 3.91 | 4.12 |
| Pharynx | 2.63 | - | 3.03 | - |
| Hypopharynx | - | 4.13 | - | 5.75 |
| Tonsils | - | 3.07 | - | 3.98 |
| Lung total | 6.67 | 2.06 | 2.57 | 2.08 |
| Lung squamous cell | - | 3.50 | | - |
| Liver | - | 2.69 | 1.95 | 2.84 |
| Stomach | 2.04 | 2.19 | 2.05 | 2.22 |

[1]Adjusted for age at diagnosis. [2]Adjusted for age and calendar year at diagnosis. [3]Adjusted for age, gross domestic product, and geographic region.

### 1.2.2 Environmental risk factors

Gender imbalance in exposure to oxidative damage by known carcinogens, like tobacco, alcohol, chronic infections, UV/ionizing radiation, and carcinogenic substances including chemical agents, pharmaceuticals, food, metals, arsenic, dusts, and fibers (17-23), probably explain part of the excess cancer risk men (12, 13, 24). The effect of environmental risk factors is manifest for malignancies with a strong association with smoking and alcohol consumption, i.e., tumors in the respiratory tract, oral cavity, pharynx, esophagus, and liver. The male excess risk of these cancers have declined, but not completely vanished, in geographical areas where sex differences in smoking and alcohol consumption have levelled out (12). Urinary bladder cancer is closely related to cigarette smoking, but deviate from the

general pattern, and the 4-fold relative risk increase in men is remarkably stable over calendar time and geographic region (3, 5-7).

#### *1.2.2.1 Carcinogenic infections*

Cervical cancer is the fourth most common malignancy in women globally and nearly 100% is caused by infection with high-risk human papilloma virus (HPV) (23, 25, 26). HPV is also associated with squamous cell carcinoma of the oropharynx, anus, and penis. Other infectious agents with an established, strong association with cancer include *Helicobacter pylori* (gastric cancer), hepatitis B and C (liver cancer), Epstein-Barr virus (Burkitt, Hodgkin, non-Hodgkin lymphoma, and nasopharyngeal carcinoma), human herpes virus-8 (Kaposi's sarcoma), human T-cell leukemia virus-I (leukemia, lymphoma), liver flukes (biliary cancer), and schistosomiasis (bladder cancer) (23, 26). Human immunodeficiency virus increase the risk of multiple malignancies at various sites mainly through immunodeficiency (23, 26). Excluding cervical cancer, there is evidence of a male overbalance regarding the prevalence of carcinogenic infections in the general population as well as in cancer patients in total (26, 27). The fraction of incident non-sex-specific cancer cases attributable to infections are estimated to be higher in men (26, 27). This phenomenon may be a result of a lower immuno-surveillance in men and/or gender-related differences in transmission patterns (26-31).

### 1.2.3 Innate risk factors

Innate, or biological, determinants of sex differences in cancer risk cover anatomic, immunologic, hormonal, and/or metabolic differences in men and women.

#### *1.2.3.1 Stem cell number*

The lifetime risk of cancers of many different types correlates to the total number of cell divisions to keep organ homeostasis (32). Larger organs and consequently higher number of cell divisions and/or an increased metabolic rate in men compared to women is an interesting hypothesis that will be discussed in detail in *1.4 Height and cancer*.

#### *1.2.3.2 Immunology*

Recognition of the role of the immune system in carcinogenesis, cancer immunology, has evolved rapidly over the past decades. This interdisciplinary branch spans beyond malignancies associated with immunodeficiency conditions and carcinogenic infectious agents but is rather concerned with cancer immunosurveillance and immunoediting (33, 34). The former is regarded as a host protector property by CD8+ cytotoxic T cells recognizing and inhibiting cancer development (33, 34). Immunoediting is a process of interaction between the immune system and tumor cells (33, 34). Cancer immunotherapy is the application and a category of cancer drugs that acts by directing and/or enhancing the immune system to target tumor antigens and attack tumor cells (33, 34). Mounting evidence suggests a reduced innate and adaptive immune function in men compared to women (28, 30, 31, 35). This is likely the underlying reason for the higher relative risk of autoimmune diseases in women but also accounts for the increased susceptibility to infections acquired via

multiple routes and the reduced vaccine efficacy in men (28, 30, 31, 36). The explanation to this anomaly is believed to be due to a trade-off for a selection of sex characteristics affecting reproduction (28, 30). With newly gained understandings of the close relationship between the immune system and cancer, it is not farfetched to believe that immunological mechanisms at least partly account for sex differences in cancer risk as well as survival (28, 30, 31, 36).

#### *1.2.3.3 Other genetic factors*

Women harbor two copies of chromosome X, but only one copy of each allele is transcribed in each cell. Escape from X-chromosome inactivation could result in two copies of tumor suppressor genes on the X chromosome, contributing to the lower cancer incidence in women (7, 37). The effect of the Y chromosome in disease development and progression, including cancer, is unclear (38). Mosaic loss of chromosome Y is however not a rare event in elderly men and appears to be associated with an increased cancer risk (39). Sex hormones and sex-linked genes have been proposed to affect microRNA which has emerged as a potential regulatory molecule in several physiological aspects of disease (40). A study using molecular data in cancer patient cohorts from the Cancer Genome Atlas Project discovered extensive sex differences in gene expression signatures for some malignancies (41). In addition to gender differences in exposure to carcinogens, a different cellular response to oxidative stress in men and women have been proposed to contribute to the male cancer susceptibility (42). Men also appear to be more vulnerable to low serum concentrations of antioxidants (43).

## 1.3 SEX AND CANCER SURVIVAL

It is important to distinguish between cancer mortality in the population and survival/mortality among patients diagnosed with cancer. The former does not only reflect prognosis, but also cancer incidence (44). Sex differences in cancer incidence have been discussed previously and, to avoid confusion, this section will focus on sex differences in cancer patient survival, i.e., mortality after cancer diagnosis. Several partially interacting and/or overlapping patient, tumor, and external factors affect cancer patient survivorship. Patient characteristics comprise age, comorbidity burden, organ function, performance status, cancer-related symptoms, socio-economy, and ethnicity. Commonly used tumor prognostic factors include disease stage, within-site primary tumor location, number and location of metastasis, and tumor biology (histology, grade, lympho-vascular invasion, proliferation indices, and mutational status). Clinical management, like waiting times and treatment intensity, additionally affects survival.

### 1.3.1 Epidemiology

Men appear to have a poorer cancer prognosis in a majority of malignancies (8-11, 45). Survival inequalities are however not as consistent as when studying incidence. Estimated male-to-female hazard ratios for cancer sites with a consistently poorer survival in men, extracted from three publications, are summarized in Table 2 (9-11). Comparisons are hampered by the fact that different studies use different cancer classification systems with varying granularity, effect measures, and methodology. Still, the mortality in men diagnosed

with cancer is persistently 10-100% higher, compared to women. Only two sites, biliary and urothelial cancer, are typically associated with a poorer prognosis in women (9-11, 45). Few studies have reported results over different time scales, but, the female survival advantage appears be consistent over calendar time and age, and is most distinct during the female reproductive age (8, 46, 47).

*Table 2. Male-to-female excess hazard ratio in cancer sites with a consistently poorer survival in men. Data extracted from three publications on sex differences in cancer survival.*

| **Author, year:** | *Micheli, 2009 (11)* | *Cook, 2011 (9)* | *Jung, 2012 (10)* |
|---|---|---|---|
| **Geographic region:** | Eurocare-4[1] | Seer-9 (US) | Korea |
| **Calendar period:** | 1995-1999 | 1977-2006 | 2005-2009 |
| **Cancer site** | male-to-female hazard ratio[2] | male-to-female hazard ratio[3] | male-to-female hazard ratio[3] |
| Thyroid well-differentiated | 1.45 | | 1.89 |
| Skin | | 1.81 | |
| Skin melanoma | 2.04 | | 1.56 |
| Anus | | 1.07 | |
| Salivary glands | 1.75 | 1.52 | |
| Hodgkin lymphoma | 1.15 | 1.20 | |
| Non-Hodgkin lymphoma | 1.15 | | |
| Lung | 1.05 | 1.19 | 1.32 |
| Tongue | | 1.07 | |
| Nasal cavity/sinuses | 1.10 | 1.19 | 1.35 |
| Stomach | 1.09 | 1.04 | 1.01 |

[1]Austria, Denmark, England, Finland, Iceland, Ireland, Malta, Norway, Sweden, Scotland, Wales, Northern Ireland, Slovenia, Belgium, the Czech Republic, France, Germany, Italy, The Netherlands, Poland, Portugal, Spain, and Switzerland. [2]Adjusted for age and region. [3]Adjusted for age and calendar year.

### 1.3.2 Patient characteristics

Advanced age and comorbidity burden obviously affect life expectancy in general but cancer survivorship in particular (48, 49). Aging is associated with a decrease in physiological reserve, including cognitive, pulmonary, renal, hepatic, and bone marrow function. This results in a general frailty and limited tolerance to anticancer therapy, including surgery, in elderly (49). Advanced age and comorbidity burden connote an increased number of medications and potential drug interactions. History of cardiovascular disease is moreover a contraindication to several antineoplastic drugs due to acute and cumulative cardiotoxicity. Awareness of the elevated risk of treatment-related complications and toxicity in elderly affects cancer treatment decision making and impinges the probability of receiving standard of care (48, 49). Information on age is usually available and handled in epidemiological studies. There is no general difference in age distribution between men and women diagnosed with cancer, though some variation in-between sites do exist. Comorbidity burden is complex to measure, impossible to fully adjust for, and is discussed in detail in a separate section (*3.2.3 Measuring comorbidity*) (50, 51). Comorbidity is associated with poorer survival directly, but also through decreased cancer treatment intensity and increased treatment-related complications. Comorbidity burden, across all age groups, appears to be higher in men and sex differences are most evident in cardiovascular and pulmonary diseases and in smoking-related malignancies (48).

Physiological functions affect capacity to tolerate and/or metabolize anticancer treatment and could conceivably differ in men and women with cancer (52, 53). There are evidence suggesting that the clearance of anticancer drugs, including chemotherapy and monoclonal antibodies, is higher in men resulting in more toxicity but also higher efficacy in women (52, 53). Objective measures of specific organ functions using laboratory, physiologic, and cognitive tests, exist but are rarely used in population-based studies due to differences in test method, reporting, and variation between laboratories and/or over time. The most commonly used score to assess performance status in cancer patients is the Eastern Cooperative Oncology Group (ECOG), ranging from grade 0 (fully active and no restriction compared to pre-disease) to 5 (dead) (54). ECOG performance status has been shown to correlate with response and tolerability to chemotherapy and cancer survival, but is hampered by subjectivity, large inter-observer variability, and is also possibly affected by gender preconceptions (55).

The predictive value of cancer alarm symptoms, like hemoptysis, visible hematuria, dysphagia, and rectal bleeding, vary between cancer sites and association with cancer prognosis is usually stage-dependent. Low socioeconomic status has been shown to impinge cancer prognosis (56, 57). The drivers behind this vary between health care systems and populations but are probably a combination of risk factor exposure, stage distribution, comorbidity, and treatment-related inequalities (56, 57). The excess cancer mortality in black populations in the United States conceivably reflects socioeconomic inequalities while the survival advantage in lung and gastric cancer in Asian populations is believed to act through tumor biology (57). Both are examples of ethnic disparities in cancer outcome (57).

### 1.3.3 Tumor factors

The most widely used staging system for solid tumors is the tumor-node-metastasis (TNM) system by the American Joint Committee on Cancer and the International Union for Cancer Control (58, 59). The system is based on the size and invasion of the primary tumor (T), the extent of regional lymph node involvement (N), and the presence or absence of distant metastases (M) (58, 59). Logarithms are used to categorize patients into prognostic groups (stage I-IV). Stage I typically implies a small primary tumor (T1) with no lymph node engagement (N0) while stage IV connotes the existence of distant metastases (any T, any N, and M1). The extent of disease at diagnosis is crucial in treatment decision making and prognosis. Women typically present at an earlier stage at diagnosis, but the excess mortality in men remains elevated even after adjusting for stage (8-11, 47). Whether the more advanced stage at diagnosis in men reflects more aggressive, faster-spreading tumors and/or different health care seeking behavior in men and women, is unclear.

Several biomarkers with either prognostic or predictive value may affect treatment outcome and/or prognosis in men and women (49). Left-sided colon tumors are more common in men and right-sided in women, malignant melanoma located on the head, neck and trunk is more prevalent in men while lower extremities dominate in women (60-62). Men and women appear to have different histological distribution within the same anatomical site. Malignant

tumors typically harbor abnormal chromosome numbers, which is associated with poorer prognosis and seem to be more common in men (63). Free testosterone has been suggested to drive cancer aggressiveness and increase cancer mortality in both sexes, the mechanism of action behind this association is unclear (64). In colorectal cancer, high levels of microsatellite instability (MSI) is not only a positive prognostic factor, but also more common in colorectal tumors in women (65, 66).

The comprehension of the regulating role of immune cells in the tumor microenvironment is expanding (33, 34). A more active host immune response to malignant tumors in women offers a potential explanation to the female survival advantage (28). See *1.2.3.2 Immunology* for a more detailed background to immunological mechanisms in carcinogenesis.

### 1.3.4 Clinical management

Initial management of cancer alarm symptoms in primary health care is of vital importance. Reduced diagnostic delays can result in cancer diagnosis at an earlier stage and improve the chances of long-term cure. A handful of qualitative studies on health seeking behavior have been conducted and two meta-analysis synthesizing these results were identified (67, 68). In summary, a pattern of longer patient delay among male and socioeconomically deprived patients is apparent (67, 69). According to these studies, help-seeking pattern in men appears to be negatively affected by fear of embarrassment, weakness, and loss of masculinity (67, 69). Sanctioning of help-seeking from partners also seems to be more important to men than women (67). Gender differences in health care utilization were consistent over different cancer sites and geographic areas and it is likely that this phenomenon contribute to sex differences in survival, through delayed treatment and more advanced disease stage in men (67). Most women in western societies establish a contact to the healthcare system via maternity and obstetrics services and screening programs for breast and cervical cancer early in adult life. Being acquainted to healthcare from previous experience could potentially lower the threshold to seek medical advice when cancer alarm symptoms arise.

No reports of less optimal cancer care in men compared to women were identified. But unequal management could indeed apply to all levels, including sex differences in primary prevention and treatment of risk factors, detection of early symptoms, diagnosis, staging, and treatment of malignant disease. Reports on gender disparities in medicine rather suggest the opposite. Men seem to receive adequate management of myocardial infarctions, are prescribed modern, costly cardiovascular drugs, and undergo surgery of liver metastasis, more frequently compared to women (70-72).

### 1.3.5 Non-small cell lung cancer

#### *1.3.5.1 Epidemiology and etiology*

Lung cancer is the leading cause of cancer death in Sweden as well as globally (25, 73). Small cell and non-small cell lung cancer (NSCLC) are distinct disease entities with different clinical characteristics. The major histological cell types comprising NSCLC are

adenocarcinoma, squamous cell carcinoma, and large cell carcinoma (74, 75). Women are more commonly diagnosed with adenocarcinoma and men with squamous cell carcinoma. Lung adenocarcinoma is progressively becoming more common, while the other cell types are decreasing. Lung adenocarcinoma is currently the dominating cell type in both sexes in most parts of the world, the reasons behind this is not fully understood (76, 77). Despite therapeutic advances, NSCLC prognosis has remained poor with 5-year relative survival ranging from 17-18% (squamous cell carcinoma) to 18-24% (adenocarcinoma). Tobacco smoking accounts for 80% of incident lung cancer cases in Sweden and there is a clear dose-risk-relationship between tobacco smoking and all of the major lung cancer subtypes, the association is moderately weaker in adenocarcinoma (78, 79). Women and young patients are more likely to suffer non-smoking associated NSCLC (78).

#### *1.3.5.2 Symptoms and diagnostics*

Loco-regional lung cancer symptoms include cough, hemoptysis, dyspnea, thoracic pain, nerve entrapment, and cardiovascular manifestations. Distant metastases induce symptoms depending on location, in lung cancer typically the central nervous system, bone, liver and adrenal glands. General symptoms; anorexia, cachexia, fatigue, and fever, indicate advanced disease. Most lung cancers are histologically confirmed from needle biopsy or cytology via flexible bronchoscopy. Morphological criteria are combined with immunohistochemistry to distinguish cell type and determine mutational status. Large cell carcinoma is a diagnosis by exclusion. Computed tomography (CT) of the thorax and upper abdomen is used to determine TNM stage and positron emission tomography is routine when treatment intention is curative. Patients planned for surgery additionally undergo pre-operative pulmonary function tests. NSCLC staging relies on the TNM system by the American Joint Committee on Cancer. The TNM combination determines stage I-IV, which forms the basis for therapeutic decisions and predicts prognosis (58, 80).

#### *1.3.5.3 Treatment*

Standard treatment in stage I-IIIA (localized disease) NSCLC is radical thoracotomy with lobectomy or pneumonectomy followed by adjuvant platinum-based chemotherapy (stage IB-IIIA). Localized disease in patients unfit for surgery can be treated with hypo-fractioned or conventional radiotherapy, depending on tumor size and location. Curative radiotherapy with concomitant chemotherapy can be offered to stage IIIB patients in good performance status. Palliative radiotherapy can relieve local symptoms and palliative platinum-based chemotherapy prolongs life in stage IIIB-IV. Patients with activating EGFR mutations are recommended tyrosine kinase inhibitors first line, and EGFR negative/ALK positive tumors selective ALK inhibitors second line. Immune checkpoint inhibitors is revolutionizing NSCLC therapy, but is so far mainly recommended in second line advanced NSCLC (81, 82).

#### *1.3.5.4 Prognostic factors and sex*

Known prognostic factors in NSCLC include TNM stage, tumor biology (histology, differentiation grade, proliferation rate, pleural and vascular invasion, mutational status),

primary tumor location, number of resected lymph nodes, number and location of metastasis, age, smoking history, socioeconomic status, ethnicity, comorbidity, performance status, and presence of pulmonary symptoms and weight loss at diagnosis (83-94). Multiple studies from separate geographic regions have concluded that male sex is an independent, negative prognostic factor in NSCLC (83, 84, 86, 93, 95-105). Women are more often diagnosed with lung adenocarcinoma, having a slightly better prognosis *per se,* and several studies have noted that the female survival advantage seems to be limited to this cell type (100, 102-104, 106-109). Moreover, men with NSCLC have an unfavorable stage distribution, comorbidity burden, more often smoke, and are older at diagnosis, compared to women (86, 87, 98, 110-113). The perioperative mortality in men with NSCLC has been reported to be 4-fold compared to women, possibly reflecting unaccounted comorbidity in male patients (87, 107). Women undergo more partial NSCLC resections with lower risk of complications, presumably because of the higher incidence of adenocarcinomas, typically located in the peripheral lung tissue, in women. Reports on sex differences in chemotherapy toxicity and response rates have been inconsistent (52, 53, 95, 96, 114-117).

Sex differences in stage can be the effect of gender differences in health care utilization as well as an indication of more aggressive tumor behavior in men. Disparities in age, histology and smoking patterns in men and women indicate biologic sex differences. Activating EGFR mutations are both prognostic and predict response to treatment with tyrosine kinase inhibitors in lung adenocarcinoma. EGFR positive tumors are more common in non-smokers, Asian populations, and women. Other proposed sex differences in tumor biology include hormonal influences and gene expression and polymorphisms affecting DNA-repair capacity, p53, GRPR, CYP1A1, GTSM1, ALK, and tumor mutational load (112, 118-121).

The understanding of the biological rationale behind the observed sex differences in NSCLC epidemiology and the possible impact on NSCLC prognosis is improving but many questions remain unanswered (120). Many of the listed and discussed prognostic factors overlap, interact, and are associated with sex and/or can be hypothesized to affect men and women differently. It is indeed a complex task, entailing methodological challenges to disentangle the effect of sex from other components affecting NSCLC outcome.

### 1.3.6 Urinary bladder cancer

#### *1.3.6.1 Epidemiology and etiology*

Urothelial carcinoma arises from the transitional epithelium in the urinary tract and 95% originates in the urinary bladder (UBC). Urothelial cancer is the 7$^{th}$ most common cancer in Sweden and the 9$^{th}$ globally (25, 73). Tobacco smoking is believed to account for 50% of UBC (122). Other risk factors include arsenic in drinking water, occupational exposure to aromatic amines and polycyclic aromatic hydrocarbons, pelvic radiotherapy, chemotherapeutic agents, and (developing regions) urinary Schistosomiasis (123). UBC is 3-4 times more common in men and mean age at diagnosis is approximately 70 years in both

sexes (5, 8, 123, 124). The excess bladder cancer risk in men has traditionally been explained by an unequal exposure to carcinogens (smoking and occupational).

#### *1.3.6.2 Symptoms and diagnostics*

Visible hematuria is the most important alarm symptom and reported in approximately 75% of UBC patients. Weight loss, fatigue, anemia, and pain, are symptoms of advanced disease. Cystoscopy, transurethral resection of the bladder tumor (TURBT), and histologic examination of tumor material, form routine diagnostics (125). CT and/or magnetic resonance imaging (MRI) are used for staging (125). UBC is staged according to the TNM system by the American Joint Committee on Cancer (58, 126). Tumors are histologically classified into low malignant potential (LMP), high (G1), intermediate (G2), and poor (G3) cell differentiation (127).

#### *1.3.6.3 Treatment*

Approximately 75% of the tumors are non-muscle invasive (NMIBC) at diagnosis and typically cured after TURBT. Depending on a set of tumor risk factors (T stage, size, number, location, histology, radicality after re-resection, vascular invasion, concomitant carcinoma in situ) for recurrence or progression, TURBT is followed by intravesical instillations (125). The 5-year relative survival in non-metastasized muscle-invasive bladder cancer (MIBC) receiving standard of care, i.e., cisplatin-based chemotherapy followed by radical cystectomy, is around 50%. Advanced age, comorbid conditions, and impaired performance status in this frail patient group impinge capacity to undergo intensive, multimodal treatment and conventional radiotherapy forms an alternative in patients unfit for surgery (128, 129). Approximately 10-15% of UBC patients present with distant metastasis at diagnosis and have a median overall survival of 3-6 months if left untreated (124). First line palliative, cisplatin-based chemotherapy improves survival, but is associated with severe toxicity, including nephrotoxicity (125). Modern immunotherapy is paradigm shifting in urothelial carcinoma and several checkpoint inhibitors are now approved for treating advanced disease in Sweden (130).

#### *1.3.6.4 Prognostic factors and sex*

A shorter delay from symptom onset to diagnosis and treatment typically results in less advanced disease stage and consequently improved survival (131). Moreover, NMIBC and MIBC entail completely different standard of care and prognosis, and also seem to be characterized by different genetic and molecular changes (128, 132). Pathologic down-staging after preoperative chemotherapy is a favorable prognostic marker in MIBC (133). In the primary metastatic setting, the so-called Bajorin factors, i.e., ECOG performance status ≥1 and visceral metastasis, are validated negative prognostic markers (134).

Unlike the majority of cancers, women with UBC have poorer survival, compared to men (9, 135, 136). The female survival disadvantage is however not consistent across geographic region (137). A register-based study on sex differences in cancer survival in Estonia reported

similar UBC stage distribution and survival in men and women (8). Another, Canadian study restricted to UBC patients treated with radical cystectomy or radiotherapy, found no sex differences in stage, treatment, nor outcome (138). Women with UBC typically present with a more advanced stage at diagnosis (139-142). There are evidence that women with visible hematuria experience longer diagnostic delays and less frequently undergo imaging according to guidelines (139, 141, 142). Importantly, the female survival disadvantage persists after adjustment for both stage at diagnosis and treatment modality (135, 139, 140). Women seem to suffer from more cisplatin toxicity, the drug of choice in the preoperative and palliative setting, whether this affects bladder cancer outcome has not been evaluated (52). Different risk factor exposure, metabolism of carcinogens, and sex hormones have also been proposed to contribute to the poorer bladder cancer survival in women (141, 142).

## 1.4 HEIGHT AND CANCER RISK

### 1.4.1 Epidemiology

The association between adult body height and cancer risk is a well-known phenomenon in cancer epidemiology (143-154). Several studies from separate geographical regions have concluded that tall stature is associated with an increased risk of cancer overall, site-specific, and in both sexes (143-154). The risk appears to increase by 5-15% per 5-10 cm increase in body height for most of the studied cancer sites (Table 3) (143, 144, 146, 148).

A majority of the performed studies on cancer risk and body height have applied a cohort approach using cancer register, cause-of-death register, or administrative medical data to identify incident cancer cases (143-154). Many studies have been restricted to the most prevalent malignancies in the population due to small numbers and low statistical power. In combination with lack of information on histological cell type, this has limited the possibility to analyze less common subtypes. Ecologic studies imply a different epidemiological approach, and a positive association between average adult height and cancer incidence rates in different countries has been reported (149). Using death register data to measure cancer incidence obviously connotes multiple barriers and calls for cautious interpretations since only fatal cases are included and death certificate information is unreliable (152, 154).

Cancer risk and attained body height are both heavily dependent on socioeconomic status (57, 155, 156). Previous studies have found the height-cancer association to be remarkably stable after adjustments for a wide range of sociodemographic factors (143, 146-148, 151, 152, 154). Smoking-related tumors (esophageal and head and neck) pose an exception and appear to be inversely related to height. In studies containing data on socioeconomic status and/or smoking history, tall stature did increase the risk of these tumors in multivariable-adjusted models including socioeconomic factors and/or smoking history (143, 146, 147, 151, 152, 154). The height-cancer association is unexpectedly consistent in malignant melanoma, colorectal, kidney, and the hematological malignancies (Table 3).

*Table 3. Adjusted relative risk (RR) for cancer overall and selected sites, per unit increase in adult body height. Data extracted from four publications on body height and cancer risk.*

| **Author, year:** | *Green, 2011 (143)* | *Kabat, 2013 (146)* | *Sung, 2009 (148)* | | Wirén, 2014 (144) | |
|---|---|---|---|---|---|---|
| **Geographic region:** | Million Women Study (UK) | Women's Health Initiative (US) | Korea | | Me-Can (Norway, Sweden, Austria) | |
| **Calendar period:** | 1996-2001 | 1993-1998 | 1994-2003 | | 1972-2005 | |
| **Sex:** | women | women | men | women | men | women |
| **Cancer site** | **Relative risk[1] /10 cm** | **Relative risk[2] /10 cm** | **Relative risk[3] /5 cm** | | **Relative risk[4] /5 cm** | |
| Overall | 1.16 | 1.12 | 1.05 | 1.07 | 1.04 | 1.07 |
| Malignant melanoma | 1.32 | 1.15 | | | 1.13 | 1.17 |
| Kidney | 1.29 | 1.23 | | | 1.12 | 1.05 |
| Colon cancer | 1.25 | 1.14 | 1.04 | 1.08 | 1.09 | 1.11 |
| Rectal cancer | 1.14 | 1.26 | 1.06 | 1.00 | 1.06 | 1.09 |
| Non-Hodgkin lymphoma | 1.21 | 1.11 | | | 1.10 | 1.06 |
| Multiple myeloma | 1.13 | 1.30 | | | | |
| Leukemia | 1.26 | 1.04 | 1.02 | 1.21 | | |

[1]Adjusted for age, region, socio-economy, smoking, alcohol, BMI, exercise, age at menarche/first birth, and parity. [2]Adjusted for age, alcohol, pack-years, hormone replacement therapy, education, ethnicity, and randomization status. [3]Adjusted for age, BMI, smoking, alcohol, exercise, salary, occupation, and area. [4]Adjusted for date of birth and age.

### 1.4.2 Somatic mutations and cell number

The rationale behind the height-cancer relationship is debatable and not fully understood. But the consistency across geographic region, anatomical site, and sex suggests a common underlying mechanism. Somatic driver mutations in cancer occur during DNA replication (157). Organ-specific cancer risk is driven by the accumulation of mutations in proto-oncogenes which in turn is related to stem cell number and turnover rate within tissue (32, 157, 158). Many mutations occur by chance rather than are caused by extrinsic carcinogens or inherited mutations in oncogenes, tumor suppressors, or DNA-repair genes (32, 153, 158-160). A larger body reasonably consists of more cells and hence connotes a higher lifetime risk of developing a malignant tumor. This theorem is supported by the multistage model of carcinogenesis (32, 157). The age-incidence pattern, where most malignancies increase rapidly with age, is also consistent with the proposed effect of cumulative cell divisions on cancer risk (153, 157, 159-161). The exact relationship between height and organ-specific stem cell number is however not established and probably differs between organs and cell types.

### 1.4.3 Other pathways

Other causative pathways behind the height-cancer association have been outlined. Attained adult stature is basically determined by two factors: genes (i.e., parental height) and nutritional status during periods of growth; intrauterine, in childhood, and in adolescence (155, 156). Body height is a polygenic trait influenced by hundreds of genetic variants in at least 180 loci (162). Genetic determinants of height could theoretically promote carcinogenesis directly and not through increased cell numbers. Increased environmental exposure to carcinogens, for example through a higher basal metabolic rate, might contribute

to a higher cancer burden in taller individuals. Caloric intake in adult laboratory animals have been found to increase cancer risk (151). Energy expenditure increase with body size but an association with energy intake and human cancer is however not yet established. Evading apoptosis and thereby enhancing tumor cell survival is an important hallmark in carcinogenesis (157). Malnutrition during development do not only entail an increased risk of permanent stunting but also lower levels of insulin-like growth factors, a suspect promotor of carcinogenesis through down-regulation of apoptosis (152, 158).

### 1.4.4 Sex, height, and cancer

The uniform excess cancer risk in men over calendar time, age, and geographical region, described in *1.2 Sex and cancer risk*, indicates unknown, non-environmental drivers. Men are larger than women in general, and an interesting hypothesis is whether the universal male excess cancer risk is explained by a higher number of cumulative stem cell divisions? Walter et al explored the relationship between sex, cancer, and height using self-reported data on body height from the Vitamins and Lifestyle study, a cohort of approximately 65,000 men and women. Incident cancers were identified through the Surveillance, Epidemiology, and End Results cancer registries (163). Men had a 50% higher risk of non-sex specific malignancies and approximately one third of the male excess cancer risk was explained by sex differences in body height (163). This study is to our knowledge the first to explore whether the increased cancer susceptibility in men can be explained by body height.

# 2 AIMS

The global cancer burden is expanding rapidly, increasing costs for already strained health care systems worldwide (164). It is well-known that men, compared to women, suffer an increased cancer risk as well as poorer cancer survival. The underlying reasons are however not satisfactorily outlined.

The aims of this thesis are to:

Study I

- Delineate temporal and age trends in sex differences in cancer risk and survival using high-quality, longterm, population-based data
- Adress uncharted risk and prognostic factors underlying the disproportionate cancer burden in men
- Quantify the potential gain achieved by eliminating the excess cancer risk in men

Study II

- Characterize sex differences in a range of clinicopathological factors, including clinical management, affecting non-small cell lung cancer survival
- Quantify the male survival disadvantage in non-small cell lung cancer using absolute and relative effect measures, accounting for various distribution of prognostic factors in men and women

Study III

- Outline sex differences in clinicopathological factors, comorbidity burden, socioeconomy, and clinical management of urothelial bladder cancer
- Explore if these can explain the poorer outcome in women compared to men with bladder cancer
- Distinguish in which stage group and time-window the excess urinary bladder cancer mortality in women occurs

Study IV

- Explore the relationship between attained body height and a number of malignancies, using large, high-quality, population-based data
- Quantify to what extent taller body stature, as a proxy for stem cell number and cumulative turnover rate, can explain the excess cancer risk in men

# 3 MATERIAL AND METHODS

## 3.1 MATERIAL

The numerous Swedish population-based registers constitute a gold mine for epidemiological research owing to the governmental bodies responsible for the long-term data collection and maintenance (165). The Swedish national registration number (NRN) is a unique identifier assigned to all Swedish residents that enables individual-level data linkage between multiple registers, and ensures a long-term follow-up regardless of domestic migration (165, 166). The national health data registers are state funded and held by the Swedish National Board of Health and Welfare who coordinates register linkage in population-based research projects. The aim of this section is to give a more detailed background to the included registers and register-holders. Exact definitions and management, including groupings, of the included covariates are described in the attached publications and manuscripts, studies I-IV.

### 3.1.1 The Swedish Cancer Register

The Swedish Cancer Register (SCR) is maintained by the Swedish National Board of Health and Welfare. The overall purpose the SCR is to monitor cancer incidence and survival in the Swedish population for health care management and planning, international comparisons, and medical research. The SCR prospectively collects data on virtually all incident cancer cases diagnosed in Swedish residents since 1958. Not only are malignant tumors included but also benign, pre-malignant lesions, as well as conditions of unknown malignant potential. Data on basal cell carcinoma is collected in a separate register. Notification of clinical, morphological, and autopsy-based cancer diagnoses is mandatory by law which ensures a high national coverage of over 95% when validated using the National Patient Register (NPR) (167). Reliability is also deemed to be high and roughly 99% of reported cases are morphologically verified (168). Some issues still remain, for example under-reporting of malignancies in elderly has been noted in cancers with poor outcome from difficult to access anatomical locations (i.e., pancreas and lung) (167). To ensure valid comparisons over time, up-to-date cancer classification systems are supplemented with the historical revisions. The earliest used are the 7th revision of the International Classification of Diseases (ICD-7) for anatomical site and the WHO Histological Classification of Neoplasms for morphology (CANC/24.1) (169, 170). For some cancer sites, refined subdivision is not feasible using the historical classification systems, thus explaining why studies of these subtypes are limited to later time periods. In addition to medical data (anatomical site, morphology, method and date of diagnosis) the SCR comprises linked information on patient (sex, place of residence) and follow-up (cause and date of death and date of international migration).

### 3.1.2 The cancer quality registers

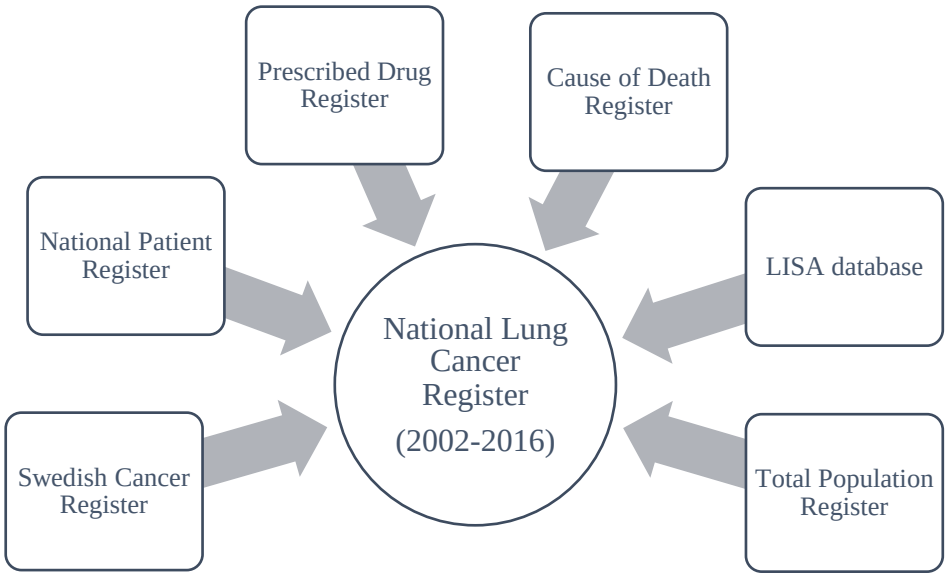
Swedish cancer registration is administered by the six Swedish Regional Cancer Centers (RCCs) who coordinate the registration, coding, verification, correction, and transfer of SCR data to the Swedish National Board of Health and Welfare. The RCCs are also responsible for administrating site-specific national cancer quality registers. Clinicians initiated the launching

of cancer quality registers for different anatomical sites with the overarching aim to ensure equal access to high-quality cancer care irrespective of geographic region in Sweden. The quality registers are funded by the Swedish state and county councils. Reporting is not mandatory by law but considered a quality measure when comparing clinics and regions. A passive patient consent is required but patients can actively opt-out participation. The quality registers contain various information on clinical parameters like disease stage, performed diagnostic examinations and results from these, first-line treatment, and sometimes also follow-up data like treatment response and second line treatment. Steering committees consisting of health professionals from the different geographical regions and different professional backgrounds cooperate with the RCCs to structure, maintain, and manage the cancer quality registers.

#### *3.1.2.1 The National Lung Cancer Register and the Lung Cancer DataBase Sweden*

The RCC Uppsala coordinates the Swedish National Lung Cancer Register (NLCR). The NLCR started in 2002 and aims to include all incident cases of invasive lung cancer according to the International Classification of Diseases for Oncology code C34. The NLCR does not cover autopsy-detected cases, carcinoma *in situ*, tracheal, or pleural tumors. Coverage is deemed to be high, at approximately 96% compared to the SCR (167, 171). The composition of reported variables has varied slightly over the years, but mainly consists of diagnostic procedures, staging methods, stage at diagnosis, histopathology, primary tumor location, location of distant metastases, mutational status, smoking history, performance status, planned treatment, crucial dates (of referral, performed diagnostic investigations, primary treatment decision) to monitor waiting times, 1-year follow-up status, and a quality of life questionnaire (171).

*Figure 1. The constituting Swedish registers forming the Lung Cancer DataBase Sweden (LCBaSe)*

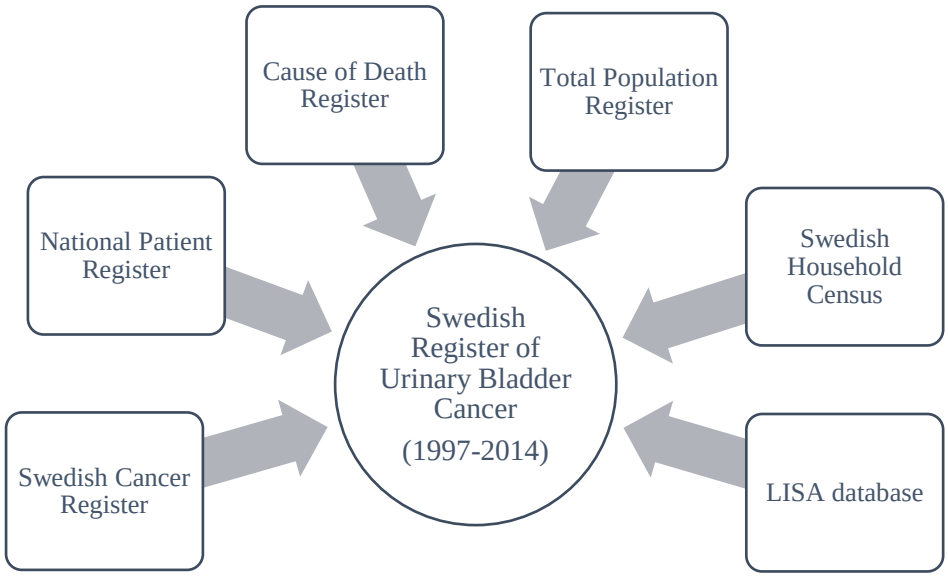


The Lung Cancer DataBase Sweden (LCBaSe) was initiated for research purposes and is based on all lung cancer cases registered in the NLCR in year 2002-2016 linked, using the NRN, to multiple nationwide registers contributing individual-level information like comorbidity burden, prescribed drugs, date and cause of death, and socioeconomic variables (Figure 1). Please see *3.1.3 Auxiliary national registers* for a detailed description of the constituting registers.

#### *3.1.2.2 The Swedish National Register of Urinary Bladder Cancer and the Bladder Cancer DataBase Sweden*

Various regional urinary bladder cancer registers started to form in Sweden in the early 1990s and in 1997 the Swedish National Register of Urinary Bladder Cancer (SNRUBC) was launched with a national coverage of over 90% from start that has now reached 97%, compared to the SCR (172, 173). The RCC Southeast is the responsible register holder. The register intends to include all incident cases of morphologically verified urinary bladder cancer according to the International Classification of Diseases for Oncology, 3rd revision code C67.0-C67.6 and C67.8-C67.9 in Swedish adults (age ≥18) (172). The original version consisted of a minimal set of variables (age, sex, date of diagnosis, tumor characteristics, primary treatment within 3 months from diagnosis), but have now expanded to include several forms covering detailed clinical diagnostic information (tumor location, size, grade, TNM stage, dates to estimate waiting times, and performed diagnostic investigations), pre- and postoperative data including surgical details in those undergoing radical cystectomy, and a 5-year follow up of non-muscle invasive tumors. Several quality indicators have been identified and these are continuously evaluated to compare and encourage regional and temporal improvements in the management of urinary bladder cancer.

*Figure 2. The constituting Swedish registers forming the Bladder Cancer DataBase Sweden (BladderBaSe).*



With the aim of performing epidemiological research to study bladder cancer, SNRUBC data from year of diagnosis 1997-2014 was linked to several national healthcare and demographic registers to construct the Bladder Cancer DataBase Sweden (BladderBaSe) in 2015 (Figure 2) (172). *3.1.3 Auxiliary national registers* provides a detailed description of the included registers.

### 3.1.3 Auxiliary national registers

If not otherwise stated, the Swedish National Board of Health and Welfare maintains the listed and described national health registers.

The Swedish *National Patient Register* (NPR) started as the Hospital Discharge Register already in 1964, but due to different recording practices and a staggered introduction, it did not reach complete national coverage of all in-patient care until 1987. In addition to demographic information similar to that in the SCR, the NPR contains hospital administrative data including dates of hospital admission and discharge as well as main and secondary discharge diagnoses and major interventions according to ICD-7 (year 1964-1967), ICD-8 (year 1968-1986), ICD-9 (year 1987-1996), and ICD-10 (year 1997-). The coverage and validity of the NPR is high with more than 99% of all somatic and psychiatric hospital discharges reported and roughly 90% of the diagnoses being valid (165, 174). Surgical daycare procedures and specialized outpatient visits (not primary care) were added to the register from the years 1997 and 2001, respectively. The national coverage of outpatient care is only around 80%, mainly due to under-reporting from private caregivers (174).

*The Swedish Cause of Death Register* was established in 1961, but death data for year 1952-1960 has been compiled retrospectively from medical records (175). The register contains date and underlying and contributing causes of death, and is used for official statistics and medical research (175). The proportion of missing death certificates is close to 0% (165, 175). Within three weeks from death, the death certificate including a version of the International Form of Medical Certificate of Cause of Death, has to be submitted by the responsible physician to the National Board of Health and Welfare. This form is used to identify underlying and contributing causes of death according to the rules from the ICD version currently in practice (175). Approximately 2.7% of all deaths lack a specific underlying cause of death but the proportion is higher in elderly with multiple chronic conditions (175). The quality of the register is dependent on the quality of the submitted death certificates. A 77% agreement with the expected cause of death from medical case summaries have been reported in Sweden (176). The agreement was however found to be higher in case of underlying malignant disease (176).

The *Swedish Prescribed Drug Register* was launched in 2005 and is one of the most recent nationwide healthcare registers (177). The register contains detailed information, including; substance, brand name, formulation, package, amount, and dosage, on prescribed and dispensed drugs in the Swedish population. Data is raised by the state-owned National Corporation of Swedish Pharmacies and transferred to the National Board of Health and

Welfare annually (177). Drugs are classified according to the Anatomical Therapeutic Chemical (ATC) classification system (178).

Statistics Sweden administers the *Total Population Register* as well as the *Longitudinal Database for Health Insurance and Labor Market studies* (LISA) which handles information on marital status, country of birth, dates of immigration and emigration in addition to individual- and group-level data on socioeconomic factors such as education, employment, and income (165, 179). Statistics Sweden additionally provides open-access data on population counts and mortality by sex, age, and calendar year, to use as denominators when estimating approximations of cancer incidence and excess mortality in the population, see *3.2.1 Relative survival* (165).

### 3.1.4 Data on body height

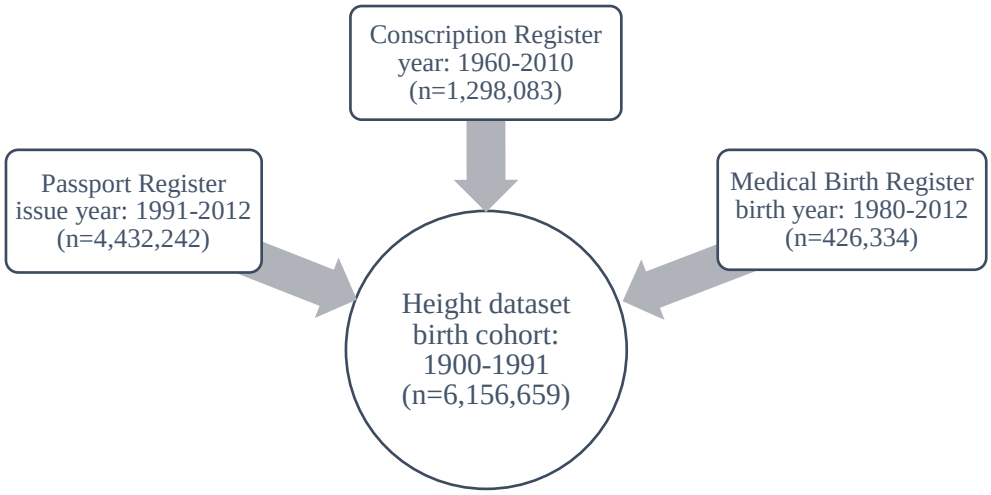
The *Swedish Passport Register* was computerized in 1991 and includes body height (self-reported or measured at time of application), sex, photo of passport holder, as well as administrative information including the NRN (180). Swedish citizens apply for a passport at the Swedish Police to facilitate international travel and for identification. Since 2005 an adult Swedish passport is valid for 5 years, but can also be renewed if lost or destroyed.

The *Swedish Conscription Register* is held by the Swedish National Archives and includes data on all conscripted Swedes since 1967. Conscription and military service was mandatory by law for all Swedish men until mid-2010, with a very small (2-3%) dropout rate, mainly due to severe medical conditions. The register contains information on height, weight, physical examination, together with cognitive tests at time of recruitment (usually at age 17-20).

The *Medical Birth Register* was established in 1973 for congenital malformation surveillance purposes. The register contains data collected at antenatal care visits including; maternal height, weight, concurrent disease, and smoking habits, as well as delivery including; birth weight, gestational age, and Apgar score. All mothers attending the public prenatal system are included in the register contributing to a coverage of approximately 97.0-99.5% of all births in Sweden (165).

The height dataset in study IV contains attained adult body height in 6,156,659 individuals. This database was created linking individual-level data on height from the passport register, the conscription register, and the medical birth register, using the NRN (Figure 3). In case of multiple, dissimilar height registrations within individual, the most frequent, tallest height was used and measured heights (conscription and medical birth registers) were prioritized rather than self-reported ditto (passport register).

*Figure 3. The Swedish registers contributing with information on adult height.*




## 3.2 METHODOLOGY

The final decision of which statistical methods to apply is not always straight-forward. In this doctoral project, methodology was partly defined from start (relative survival in study I). In study II and III multiple steps, taking both results from previously performed studies as well as findings in our own data into account, preceded the final decisions. Causal inference is an expanding area of focus in medical research. In study IV, this resulted in the application of a contemporary, not previously applied, method to perform mediation analysis of a time-to-event outcome (181). This section aims to motivate and give a more detailed background and description to the underlying methodology than can be fitted into a scientific publication. For details on data management, definitions, groupings of variables, and exact model composition, please see the attached publications and manuscripts, studies I-IV.

### 3.2.1 Relative survival

#### *3.2.1.1 Motivation*

Net mortality is the preferred measure when comparing cancer mortality in two populations with different non-cancer mortality, such as between countries, time periods, or men and women. Net survival is a hypothetical measure in a scenario where the disease of interest is the only possible cause of death. It can be estimated either in a cause-specific (using information on underlying cause of death) or in a relative survival setting (182). The relative survival in a cancer patient cohort is defined as the ratio between the observed all-cause survival in the cancer group and the expected survival from a comparable cancer-free group. The expected survival is commonly retrieved from publicly available population life tables, stratified on age, sex, and calendar year. Register-based cancer studies often present relative survival, circumventing the issue with non-reliability and/or non-availability of death certificates (44). Information from death certificates are often inaccurate and a brief background to the death certificate procedure in Sweden is described in *3.1.3 Auxiliary national registers*. Another advantage of relative survival is that the approach captures both

the direct and indirect effects of cancer on mortality (44). For example, lethal infections and cardiovascular deaths can occur as a result of malignant disease or treatment thereof, but also independently of cancer, making the true underlying cause of death in the individual hard to assess (183). Cancer mortality has consequently been shown to be over-estimated in older, comorbid populations when using a cause-specific approach (44, 183).

When applying a relative survival framework, it is a delicate task to assess how well the survival expected in the absence of the disease under study, is captured in relation to the patient group (44). As mentioned above, the expected survival is usually estimated using national life tables connoting the fundamental assumption that if the patient group did not have the disease of interest, they would have the same survival as the general population (44). This can be questioned, particularly in smoking-related cancers. Since smokers have a shorter life expectancy compared to the general population due to a higher comorbidity burden, a potential concern is that the cancer mortality in this group is over-estimated. This issue has however been shown to have a small impact on relative survival estimates, at least in malignancies with poor survival like lung cancer (184).

#### *3.2.1.2 Application (study I)*

In study I, we applied a relative survival framework to estimate cancer survival in men and women with the overarching aim to illustrate and quantify sex differences in cancer patient mortality (185). The 5-year relative survival ratio (RSR) was estimated as the observed survival in the cancer patient group 5 years from diagnosis divided by the corresponding observed survival in the Swedish population, using a cohort approach. The expected mortality, matched by age, sex, and calendar year, was estimated using the Ederer II method (186). Population-weighted survival estimates may be misleading since the age distribution of the cancer patients varies over time and age at diagnosis affects prognosis. We consequently age-standardized the RSR estimates according to the International Cancer Survival Standards using the age distribution proposed for malignancies increasing with age in broad classes (187).

The mortality analogue of relative survival, excess mortality, was estimated as the absolute difference between the all-cause mortality in the cancer cohort and the all-cause mortality in the general population (44, 186). To compare excess mortality between men and women, male-to-female excess mortality ratios (EMRs) with 95% confidence intervals (CIs) were estimated using a Poisson regression model adjusted for age and year of diagnosis, allowing for the effect of sex to vary over follow-up time.

All statistical analyses were performed using Stata Intercooled version 14.0 (StataCorp LP), the Stata command “strs” was applied when estimating relative survival (188).

### 3.2.2 Flexible parametric models

#### *3.2.2.1 Motivation*

A popular approach when contrasting the survival outcome in two groups, exposed versus unexposed, is to plot Kaplan-Meier survival proportions over follow-up together with one summary measure, the hazard ratio, estimated using a Cox proportional hazards regression model. This is in many cases adequate and Kaplan-Meier curves are, in the absence of competing risks, intuitive to interpret. In a randomized clinical trial, the exposure is the only determinant supposed to differ between the two groups and any difference in survival is interpreted as an effect of the exposure. In observational studies, like those presented in this thesis, we expect confounding of the exposure-outcome relationship, meaning that Kaplan-Meier curves do not represent and visualize the direct exposure effect.

Although Cox models provide a solution to adjusting for potential confounders, only one relative effect measure is estimated, namely the hazard ratio. The hazard ratio does not have a natural, authentic interpretation and ignores baseline risk, absolute differences (clinical relevance), and diversity among study subjects. Moreover the proportional hazards assumption, usually made in Cox models, implies that the ratio of the hazards (mortality rates) is constant over follow-up time. It is implausible to believe that the natural history of disease and underlying causes of clinical events constantly behave proportionally.

An alternative to the Cox model is the flexible parametric model suggested by Royston and Parmar (189). This model captures the different underlying hazard shapes without making strong assumptions about the functional forms. Furthermore, flexible parametric models yield absolute effect measures. Modeling is usually done on the cumulative hazard scale. The baseline hazard is modeled using natural cubic splines; mathematical functions defined by piecewise cubic polynomials with additional constraints, to produce smooth predictions. The model can estimate hazard ratios under proportional hazards, yielding close to identical estimates compared to the Cox model. But the proportional hazards assumption can also easily be relaxed by including interaction terms between the covariates and time scale in the model. With further extensions, the flexible parametric model readily predicts and visualizes smoothed hazard functions and hazard ratios over different time scales (189, 190).

User-written commands have been developed in Stata and are continuously updated with improved functionality and computational efficiency (189, 190). Flexible parametric models were, to a varying extent, applied in all of the included studies.

#### *3.2.2.2 Application (study II and III)*

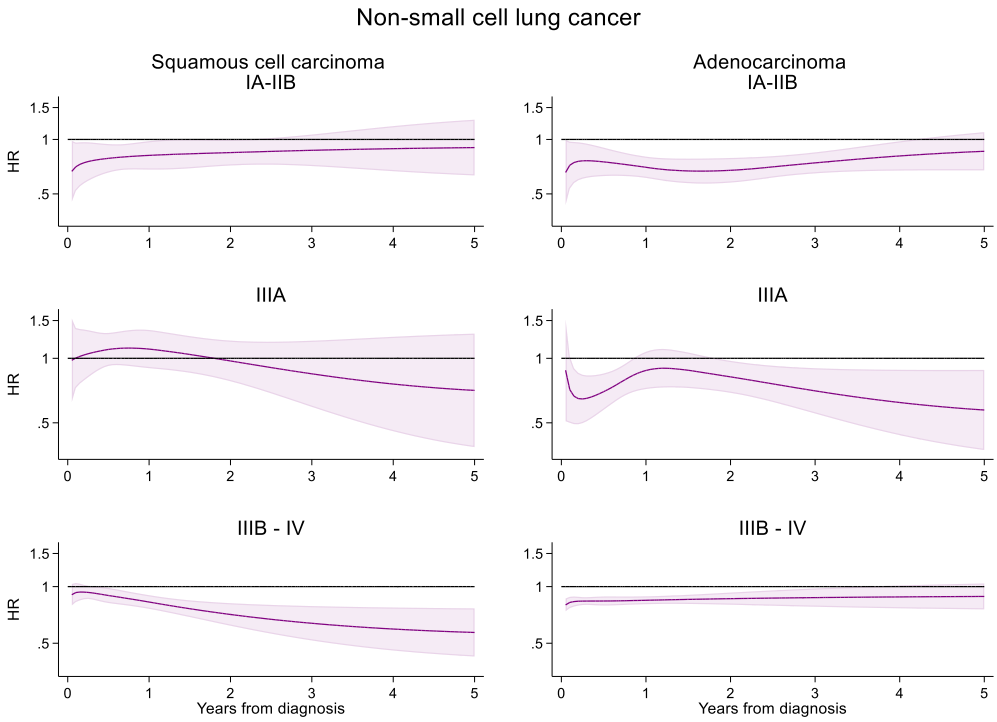
In study II, non-small cell lung cancer (NSCLC) and study III, urothelial bladder cancer (UBC), we applied flexible parametric models to study sex differences in cause-specific mortality and survival (190-192).

A reasonable starting point, in both studies II-III, was plotting the empirical survival proportions for men (unexposed) and women (exposed) over follow-up using the Kaplan-

Meier method. Visual inspection of the graphs confirmed a male survival disadvantage in NSCLC and an advantage in UBC. In study III (UBC) we additionally plotted the smoothed empirical mortality rates, i.e., number of deaths/person-years, over follow-up for men and women. This revealed that the majority of the excess deaths due to bladder cancer occurred within two years from diagnosis. To provide a non-parametric, absolute measure of cancer-specific mortality, we calculated the number and percentage of cancer-specific deaths within 5 years from NSCLC diagnosis and two years from UBC diagnosis, in men and women respectively.

With the purpose to further outline and quantify sex differences in cause-specific mortality, we subsequently applied flexible parametric models (190). Firstly, we evaluated the proportional hazards assumption for sex by estimating and plotting the female-to-male mortality rate ratio over follow-up time from a flexible parametric model with non-proportional hazards. The baseline hazard function was modelled with a restricted cubic spline with 5 degrees of freedom (df) and the time-dependent effect of sex included 3 df. Examining the graphs, we concluded that the proportionality assumption did not hold since this would have resulted in a flat line, see Figure 4, exemplifying the female-to-male hazard ratio over follow-up in NSCLC.

*Figure 4. Female-to-male hazard rate ratio (HR) including 95% confidence interval (shaded area) over follow-up, standardized over age, calendar year, education, marital status, birth country, health care region, ECOG performance status, smoking history, Elixhauser comorbidities, TNM stage, and primary tumor location.*




Since the purpose was to investigate the extent to which the measured covariates could explain sex differences in cancer mortality we continued by estimating the female-to-male hazard ratio, again employing flexible parametric models (190). To enable comparisons with previously reported findings, we started by fitting a univariate model, thereafter adding one

covariate at a time, evaluating the effect on the main outcome, the cancer-specific mortality ratio. As an additional sensitivity analysis to explore any potential effect modification, we tested models allowing for interactions between sex and a number of covariates. In the end, we presented two models in study II (NSCLC) and three models in study III (UBC), based on ours and previously reported findings (191, 192).

To verify the results in study II (NSCLC), we additionally applied a Cox proportional hazards model to estimate female-to-male cancer-specific mortality ratios, including the same covariates as in the fully adjusted flexible parametric survival model. As expected, we found the HR estimates to be almost identical to the output from the flexible parametric models assuming the effect of sex to be proportional over follow-up time (191).

To further outline and provide a more intuitive quantitative measure of sex differences in survival than hazard ratios, we plotted standardized survival proportions with 95% CIs, in men and women, over time since diagnosis. These were estimated from flexible parametric models, but with a slightly different approach in study II and III, respectively. In study II (NSCLC) we wanted to visualize the survival under two counterfactuals where the only thing differing between the two scenarios was patient sex (i.e., the remaining covariate distribution was kept identical in the two groups). We consequently predicted one survival curve for each individual and averaged over all included covariates. This allowed us to create two standardized curves where the only difference between the curves was that in one everyone was exposed (female sex) and in the other everyone was unexposed (male sex). The baseline hazard function was fitted using restricted cubic splines with 5 df and the time-dependent effect of sex was modelled using 3 df. To contrast the survival in men and women, we estimated the absolute difference in standardized survival at 1, 3, and 5 years from diagnosis. In study III (UBC) we predicted the survival in women standardized to the observed covariate distribution in men. This is interpreted as the survival in women if women had the same age, comorbidity burden, stage, etcetera as men. The df for the restricted cubic splines in the baseline hazard functions and time-varying effect of sex, were the same as in study II. The survival proportion in women standardized to the male covariate pattern was added to the traditional Kaplan-Meier survival curves for men and women, to evaluate a potential shift.

#### *3.2.2.3 Modelling and goodness-of-fit (study I)*

See *3.2.1 Relative survival* for a description of the background and advantages of applying a relative survival framework and definitions of the relative survival ratio (RSR) and excess mortality ratio (EMR).

*Figure 5. Non-Hodgkin lymphoma (A) 5-year relative survival in men and women and (B) male-to-female excess mortality ratio (EMR), comparing yearly, empirical estimates (dots) with modelled predictions using restricted cubic splines with three degrees of freedom, including 95% confidence interval (CI, shaded area).*

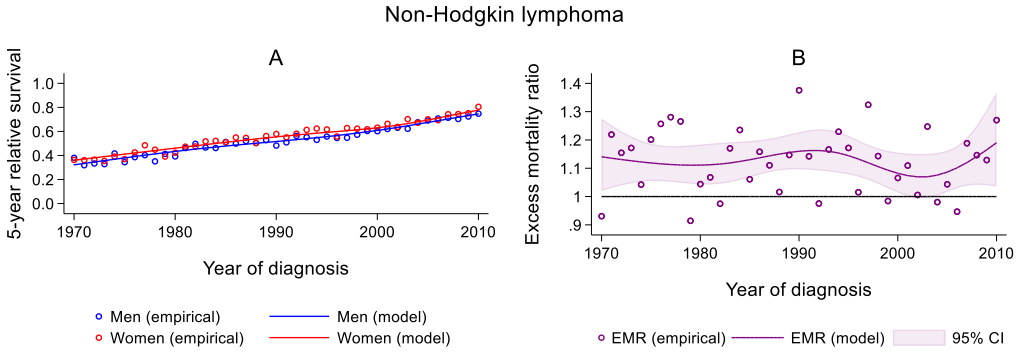



*Table 4. Applying the Akaike information criterion (AIC) to compare goodness-of-fit using 3 or 4 degrees of freedom (df), when estimating male-to-female excess mortality ratios by year of diagnosis.*

| Anatomical tract | Cancer site | AIC (3 df) | AIC (4 df) | 3 superior to 4 df |
|---|---|---|---|---|
| Head-neck | Lip | 9204 | 9205 | yes |
| | Tongue | 14088 | 14090 | yes |
| | Salivary glands | 8608 | 8599 | no |
| | Other oral cavity | 16931 | 16923 | no |
| | Pharynx | 15016 | 15019 | yes |
| | Tonsils | 8679 | 8677 | no |
| | Thyroid well-differentiated | 10274 | 10277 | yes |
| | Thyroid anaplastic | 6456 | 6456 | yes |
| Upper digestive | Esophagus adenocarcinoma | 10846 | 10842 | no |
| | Esophagus squamous cell | 17991 | 17994 | yes |
| | Stomach | 49273 | 49271 | no |
| | Liver primary | 21190 | 21191 | yes |
| | Biliary tract | 29704 | 29701 | no |
| | Pancreas | 34827 | 34820 | no |
| Lower digestive | Small intestine | 18284 | 18279 | no |
| | Colon | 66670 | 66653 | no |
| | Rectum | 55113 | 55112 | no |
| | Anus | 10063 | 10066 | yes |
| Respiratory | Nasal cavity/sinuses | 8448 | 8452 | yes |
| | Larynx | 14902 | 14894 | yes |
| | Lung | 61421 | 61401 | no |
| | Lung squamous cell carcinoma | 17253 | 17254 | yes |
| | Lung adenocarcinoma | 23205 | 23207 | yes |
| | Lung small cell carcinoma | 14352 | 14356 | yes |
| | Lung other non-small cell | 17461 | 17465 | yes |
| | Pleura mesothelioma | 10047 | 10051 | yes |
| Urinary | Urinary tract | 46154 | 46153 | no |
| | Kidney | 41370 | 41373 | yes |
| Skin | Skin melanoma | 40172 | 40143 | no |
| | Skin non-melanoma | 32852 | 32846 | no |
| CNS | Brain | 30810 | 30814 | yes |
| | Meninges | 12987 | 12987 | yes |
| Hematological | Non-Hodgkin lymphoma | 48583 | 48576 | no |
| | Chronic lymphocytic leukemia | 16796 | 16799 | yes |
| | Hodgkin lymphoma | 14529 | 14532 | yes |
| | Multiple myeloma | 35525 | 35526 | yes |
| | Acute lymphocytic leukemia | 7109 | 7109 | yes |
| | Acute myeloid leukemia | 20944 | 20949 | yes |
| | Chronic myeloid leukemia | 8143 | 8143 | yes |

In study I we fitted age-standardized 5-year RSR for men and women together with age-adjusted male-to-female EMR by calendar time. The EMRs were estimated using flexible parametric models with 3 df for the baseline cumulative hazard function (190). This final model was chosen after comparing goodness-of-fit using different degrees of freedom. We started by doing a visual, subjective, comparison, plotting yearly EMR point estimates and spline modelled estimates by calendar time, for all cancer sites, as exemplified with non-Hodgkin lymphoma, panel B in Figure 5. As an additional sensitivity analysis we used the Akaike information criterion (AIC) comparing 3 and 4 df for the baseline cumulative hazard function in the flexible parametric models, verifying that 3 df proved objectively better model fit (lower AIC value) than 4, for a majority of sites (Table 4).

#### *3.2.2.4 A counterfactual approach (study IV)*

In study IV flexible parametric models were again applied. This time arising from the precise counterfactual definitions of total and indirect effects described below in *3.2.4 Causal inference on time-to-event outcomes*. We plotted the cancer-free survival in men $S_{1M_1}(t)$ and in women $S_{0M_0}(t)$ using age as the time scale (190). To quantify and illustrate the effect of height on cancer-free survival, we predicted and plotted the counterfactual survival proportion in men if they had the same height distribution as women $S_{1M_0}(t)$ and women if they had the same height distribution as men $S_{0M_1}(t)$ using height-standardized flexible parametric models. The proportion of the excess cancer risk in men explained by height (PE) was estimated as the survival difference in men before and after height standardization divided by the total survival difference, at age 90:

$$PE = \frac{S_{1M_1}(t) - S_{1M_0}(t)}{S_{1M1}(t) - S_{0M0}(t)}$$

We used Stata Intercooled (StataCorp LP) version 14.0 (study I) and 15.1 (studies II-IV) and the Stata commands “stpm2” and “stpm2_standsurv” when applying flexible parametric models (190).

### 3.2.3 Measuring comorbidity

#### *3.2.3.1 Motivation*

As discussed previously in *1.3.2 Patient characteristics*, comorbidity poses a major issue and needs to be meticulously handled when studying different aspects of cancer survival (51). The innate complexity of comorbidity in itself, potentially interacting with basically every aspect of cancer mortality, entails no, single, optimal management. Information on comorbidity in observational population-based studies usually consists of administrative data extracted from in-patient medical records and is a rough, underestimate of real-world probabilities. Even when details on current medical history and symptoms are readily available in the routine clinical setting, it is hard to fully appraise the influence of concurrent

disease on cancer outcome. Relying on aggregated diagnostic codes without information on duration and severity of disease makes this task close to impossible. As often, when no gold standard exists, many approaches to assess the comorbidity burden in cancer patients have been developed. The two most commonly used; the Charlson Comorbidity Index (CCI) and the Elixhauser approach, together with a handful of alternative concepts will be briefly discussed in this section (51, 193, 194).

#### *3.2.3.2 The Elixhauser approach (study II)*

The Elixhauser approach is based on clinical experience and literature in the area together with empirical testing using administrative data on 1,779,167 adult acute care hospital patients (51, 194). This way, 30 medical conditions with impact on short-term outcome were identified (51, 194) The original version requires large datasets since the included conditions are treated as distinct, binary (yes/no) variables, using the underlying impact of each condition on mortality in the cohort under study (51, 194). The Elixhauser components can also be summarized and used as a count and methods to estimate a weighted summary score have been developed (51, 195). The method has been applied in a number of settings studying different malignancies and there is evidence of a predictive validity in most cancer sites as well as cancer in general (51). In study II we used data on main and secondary diagnoses at hospital discharge from the National Patient Register and data on other malignancies from the Swedish Cancer Register, recorded 15 years-1 month before date of lung cancer diagnosis to identify the Elixhauser disease entities (167, 174, 194). The ICD revision 9 was used for conditions registered in 1987-1996 and 10 in 1997–2016. When comparing the Elixhauser to the Charlson Comorbidity Index (CCI) groupings (using the same time frame and ICD revision), we found that the proportion of patients with score 0 (no comorbidity) was smaller; 35-44% compared to 48-60%, using the Elixhauser approach, see Table 5 (191). Comorbidity burden in this elderly, smoking-prevalent cohort of lung cancer patients was most probably underestimated using both methods, but less so applying the Elixhauser approach. Moreover, we had enough data to avoid potentially misestimated, outdated weights and could instead make use of the impact of each comorbid condition on morbidity in our cohort, as specified in the original Elixhauser approach (194).

#### *3.2.3.3 The Charlson Comorbidity Index (study III)*

The Charlson Comorbidity Index (CCI) is one of the earliest described and the most cited classification systems to assess concurrent disease (51, 193). Empirical analyses were applied to identify 17 disease conditions with prognostic potential in 608 patients receiving medical in-patient care in year 1984 (193). The original data consisted of clinical notes, but algorithms have been developed to use administrative data and even patient questionnaires (51). As originally specified, CCI is a weighted index where the weights correspond to the rounded, adjusted relative risk for mortality within one year from inpatient stay for each medical condition. The maximum weight was set to 6 and conditions with a relative risk < 1.2 were excluded (193). The CCI does not include alcohol abuse, obesity, drug abuse, angina, osteoporosis, non-diabetes endocrine disorders, and tuberculosis, among other conditions.

The approach has been validated in a number of studies and the score has been re-weighted and reduced to include 12 conditions, to match progress in medical care (51, 196). Multiple studies in different settings and on separate malignancies have applied the CCI and/or different adaptions using the same disease categorization (51). In study III we did not have access to source data from the NPR to estimate and compare different approaches to assess comorbidity. Information on the CCI for each study subject, according to the original version, was delivered by the BladderBaSe holders (172). The proportion of patients with CCI 0 was found to be surprisingly high, 52-69%, in this elderly patient cohort with a high smoking prevalence. Moreover the CCI increased with calendar time indicating that a longer time-window had been used in patients diagnosed in latter time periods. We concluded that the true comorbidity burden was definitely underestimated and that the underlying reasons are probably related to the method in itself and/or in combination with inaccurate data extraction and management (192).

*Table 5. Numbers (n) and proportions (%) of men and women diagnosed with lung squamous cell and adenocarcinoma, comparing different measures of comorbidity.*

| | Squamous cell carcinoma | | | | | Adenocarcinoma | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Men | | Women | | p-value | Men | | Women | | p-value |
| | n | % | n | % | | n | % | n | % | |
| **Charlson Comorbidity Index** | | | | | | | | | | |
| 0 | 3115 | 47.5 | 1953 | 51.8 | | 5652 | 52.4 | 7614 | 60.1 | |
| 1-2 | 2280 | 34.8 | 1248 | 33.1 | | 3215 | 29.8 | 3336 | 26.3 | |
| 3+ | 1161 | 17.7 | 568 | 15.1 | 0.000 | 1928 | 17.9 | 1720 | 13.6 | 0.000 |
| **Elixhauser approach** | | | | | | | | | | |
| 0 | 2261 | 34.5 | 1386 | 36.8 | | 4196 | 38.9 | 5557 | 43.9 | |
| 1-2 | 2628 | 40.1 | 1469 | 39.0 | | 4171 | 38.6 | 4842 | 38.2 | |
| 3-4 | 1151 | 17.6 | 628 | 16.7 | | 1703 | 15.8 | 1617 | 12.8 | |
| 5+ | 516 | 7.9 | 286 | 7.6 | 0.130 | 725 | 6.7 | 654 | 5.2 | 0.000 |
| **Pharmaceutical groups**[1] | | | | | | | | | | |
| 0-2 | 1658 | 35.0 | 862 | 29.9 | | 3292 | 37.9 | 3891 | 37.3 | |
| 3-6 | 2106 | 44.4 | 1340 | 46.4 | | 3817 | 43.9 | 4690 | 45.0 | |
| 7+ | 974 | 20.6 | 683 | 23.7 | 0.000 | 1578 | 18.2 | 1848 | 17.7 | 0.352 |
| **Outpatient visits**[2] | | | | | | | | | | |
| 0-1 | 3868 | 63.3 | 2207 | 61.5 | | 6606 | 63.6 | 7767 | 63.6 | |
| 2-4 | 1429 | 23.4 | 863 | 24.1 | | 2396 | 23.1 | 2821 | 23.1 | |
| 5+ | 818 | 13.4 | 518 | 14.4 | 0.184 | 1384 | 13.3 | 1632 | 13.4 | 0.997 |
| **Inpatient visits** | | | | | | | | | | |
| 0 | 5190 | 79.2 | 3024 | 80.2 | | 8816 | 81.7 | 10640 | 84.0 | |
| 1 | 807 | 12.3 | 436 | 11.6 | | 1164 | 10.8 | 1295 | 10.2 | |
| 2+ | 559 | 8.5 | 309 | 8.2 | 0.418 | 815 | 7.5 | 735 | 5.8 | 0.000 |

[1]Year of diagnosis 2006-2016. [2]Year of diagnosis 2003-2016.

### *3.2.3.4 Alternative concepts (study II)*

Using pharmaceutical data to calculate medication-based indices poses an attractive alternative or supplement to traditional comorbidity indices since this method bypasses potentially inaccurate recording of diagnoses. Prescribed medications also captures diseases managed in out-patient clinics and to some extent reflects disease severity. The method will however only include conditions for which regular drugs are prescribed and is vulnerable to utilization and prescribing habits. No medication-based index to measure comorbidity in

cancer patients has been developed (51). Since LCBaSe contains linked data on prescribed and dispensed medications we attempted to make use of this information (177). We created counts of number of different classes of prescribed drugs, grouped according to the first three positions of the Anatomical Therapeutic Chemical (ATC) classification system, dispensed 6-18 months before lung cancer diagnosis (Table 5). This yielded a straightforward approximation of number of medications of different conditions, as a rough additional proxy of concurrent comorbidity. The method does not consider dosage and pharmaceuticals indicating health awareness and/or resource consumption were assigned the same weight as drugs indicated for severe conditions.

We additionally explored the number of out- and inpatient medical consultations in specialized (non-primary) health care 6-18 months prior to lung cancer diagnosis as a measure of comorbidity. The results were however hard to interpret since we did not have information on the cause for consultation, and visits due to pre-malign symptoms and investigations of these are doubtlessly included. Unlike the other explored comorbidity approaches, women and men had an equally high disease burden, measured as number of consultations (Table 5). We believe that this reflects a higher health care consumption rather than comorbidity burden in women, compared to men (191).

### 3.2.4 Causal inference on time-to-event outcomes

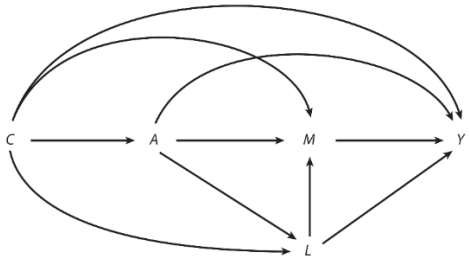
#### *3.2.4.1 Motivation*

Methodology for causal mediation analysis in epidemiological research is a rapidly advancing research field (197). The purpose of mediation analysis is to quantify possible mechanisms, i.e., pathways, through which an exposures executes or mediates effect on the outcome under study (198). The total outcome effect is separated into indirect and direct effects, where the indirect effect works through the mediator(s) under study. A mediator is on the causal pathway between the exposure and the outcome. The ratio of the indirect effect and the total effect is the proportion mediated and a relative assessment of the pathway of interest (198).

So far, most methods have been developed to disentangle indirect and direct effects on point estimates rather than time-to-event outcomes. We wanted to quantify to what extent the excess cancer risk in men is mediated by height, as a proxy of cumulative number of stem cell divisions. A newly developed methodology (study IV), the regression coefficient method was consequently applied to assess to what extent body height (mediator) explains the increased cancer risk (outcome) observed in men (exposure) (181).

A causal interpretation of indirect and direct effects, irrespective of methodology, requires fairly strong assumptions regarding confounding (181, 197, 198). Exposure-outcome (applies to all observational studies), mediator-outcome, and exposure-mediator confounders can all be controlled for by including the covariates of interest in the regression models (197). A potential mediator-outcome confounder affected by the exposure is more problematic since this confounder in itself acts as a mediator, see Figure 6 (197-199).

*Figure 6. Directed acyclic graph, DAG; A, exposure; Y, outcome; M, mediator; C, classic confounder; L, mediator-outcome confounder affected by the exposure.*




### *3.2.4.2 The regression coefficient method and application (study IV)*

In addition to the general confounder assumptions, conclusions from the regression coefficient method can only be made presuming that the exposure (male sex) and mediator (height), measured at baseline, remain unchanged during follow-up (181). Moreover, the mediator is assumed to fit a linear regression model and the outcome (time to cancer diagnosis) is supposed to be rare and follow a Cox proportional hazards regression model (181).

Cancer-free survival (i.e., time-to-cancer) was computed from date of height measurement until date of site-specific cancer diagnosis and was censored at date of death, emigration, or end of follow-up (December 31, 2011), whichever occurred first. Birth year and educational level are both associated with height as well as cancer risk, see *1.4 Height and cancer risk*, and were, if not otherwise stated, included in the outcome models.

An established counterfactual approach to define total and indirect effects were applied (200-202). The proportion of subjects remaining cancer-free at time $t$ was denoted $S_a(t)$ if the exposure was set to $a$ (1=male or 0=female) for everyone and the total effect was consequently defined as the difference in cancer-free survival in men and women:

$$S_1(t) - S_0(t)$$

Furthermore the counterfactual equivalent, $S_{aM_{a^*}}(t)$, implied the proportion remaining cancer-free if the exposure was set to $a$ and the mediator (height), for each subject, was set to the value it would have had if the exposure simultaneously was set to $a^*$ (not necessarily equal to $a$). The indirect effect of male sex (the pathway mediated through height) was defined as the difference in cancer-free survival in men at actual (male) height and men at counterfactual (female) height, the only thing differing between these two populations being the height distribution:

$$S_{1M_1}(t) - S_{1M_0}(t)$$

The ratio of the indirect and total effect define the proportion explained by mediation (PE) (181, 197, 198, 202-204):

$$PE = \frac{S_{1M_1}(t) - S_{1M_0}(t)}{S_1(t) - S_0(t)}$$

Under the previously described assumptions the proportion explained by mediation may be approximated as a function of the standard regression coefficients (181):

$$PE \approx \frac{e^{\gamma_1}(e^{\beta\gamma_2} - 1)}{e^{\gamma_1+\beta\gamma_2} - 1}$$

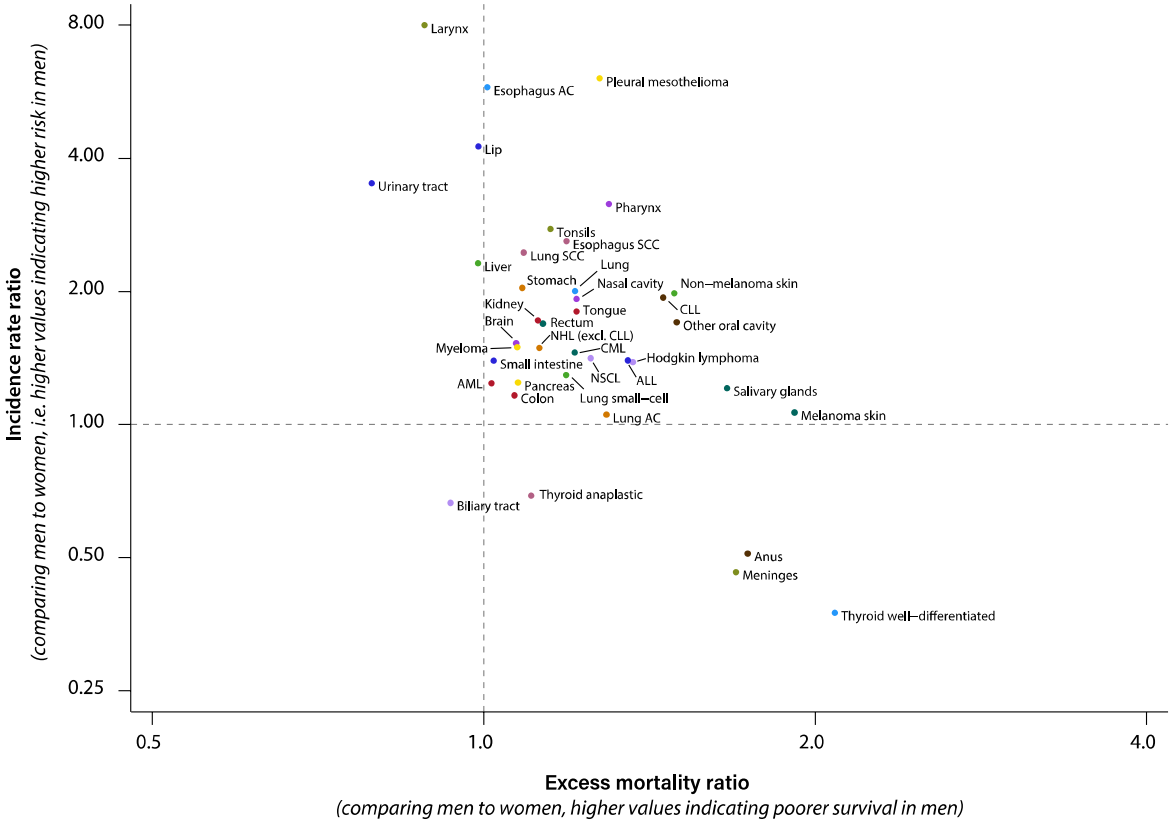
The regression coefficient $\gamma_1$ measures the direct effect of the exposure on the outcome, if $\gamma_1 = 0$, the proportion mediated is equal to 1, i.e., 100% of the exposure effect on outcome is executed through height. The regression coefficient $\beta$ measures the effect of the exposure on the mediator in the linear regression model and $\gamma_2$ the effect of the mediator on the outcome in the Cox regression model. If $\beta = 0$ or $\gamma_2 = 0$ the proportion explained by mediation is equal to 0, i.e., none of the effect is mediated by height. (181). The PE was estimated by fitting the mediator and outcome models and then plugging the regression coefficient estimates into the expression. The delta method was used to construct a 95% confidence interval for the proportion mediated (181).

A user-written function in R statistical software v.3.5.1 (R Foundation for Statistical Computing, Vienna, Austria) was used to estimate proportions explained by mediation (181).

# 4 RESULTS

The main results are all presented and discussed in detail in the attached studies I-IV. The purpose of this section is not to duplicate previously published and/or presented figures and tables. The aim is instead to add supporting material that did not fit into the final publications or manuscripts, but still illustrates how and why we ended up with the final results and conclusions. Some of the figures and tables have been included in the manuscripts including appendices and some have never been made public previously.

*Figure 7. Adjusted male-to-female incidence rate ratios (y-axis) and excess mortality ratios (x-axis) for all included cancer sites. Both axes are plotted on a logarithmic scale with base 2.*




*ALL, Acute lymphoblastic leukemia; AML, Acute myeloid leukemia; CLL, Chronic lymphocytic leukemia; CML, Chronic myeloid leukemia; NHL, Non-Hodgkin lymphoma; AC, Adenocarcinoma, NSCL, Lung other non-small cell; SCC, Squamous cell carcinoma.*

## 4.1 STUDY I

Figure 7 from study I provides a comprehensive overview of our main findings and is based on 872,397 cancer cases diagnosed in Sweden in 1970-2014 and at age 15-84 (185). Figure 7 illustrates the male-to-female incidence rate ratio (IRR) together with the male-to-female excess mortality rate ratio (EMR) for all included non-sex specific cancer sites. Both estimates are adjusted for age and year of diagnosis. An IRR > 1 indicates higher cancer risk in men and an EMR > 1 poorer cancer survival in men. It is apparent that a majority of sites are aggregated in the upper right corner, being both more common and deadlier in men. Figure 7 is however constructed from aggregated data, incidence is based on cases diagnosed in year 1970-2014 and mortality on those diagnosed in 1995-2014. Consistency across calendar year and age at diagnosis is relevant to draw conclusions and generate hypotheses on

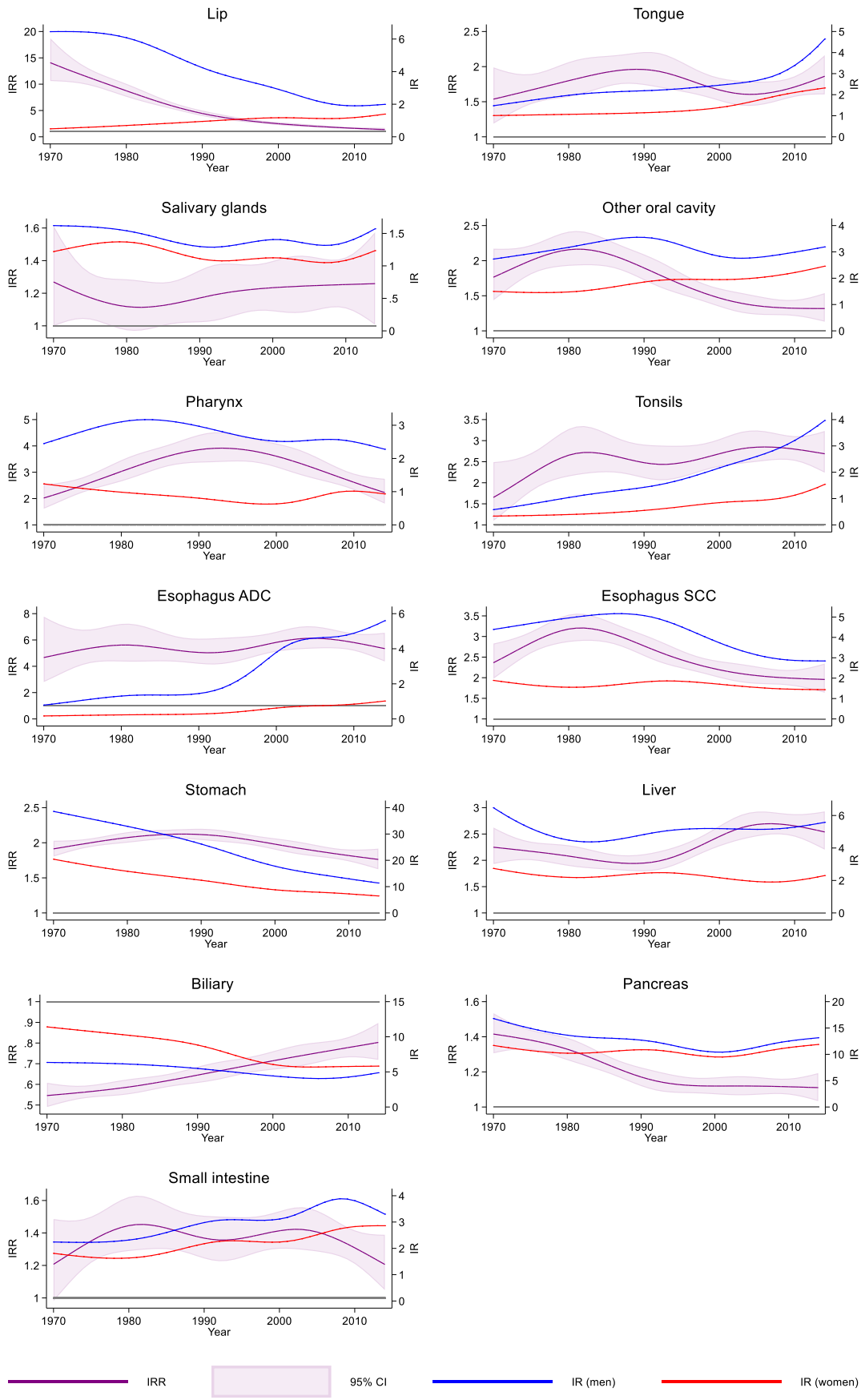
underlying drivers. Furthermore, incidence and relative survival in men and women over calendar time is interesting from a public health perspective to evaluate the effect of preventive work and advances in cancer care. We therefore decided to plot the sex-specific incidence rates (IRs) together with the male-to-female IRRs by year and age at diagnosis as well as the 5-year relative survival ratios (RSRs) together with the male-to-female EMRs by year of diagnosis for all cancer sites, please see Figure 8, Figure 9, and Figure 10, respectively. Estimates plotted over calendar time were adjusted for age and estimates by age were adjusted for calendar time.

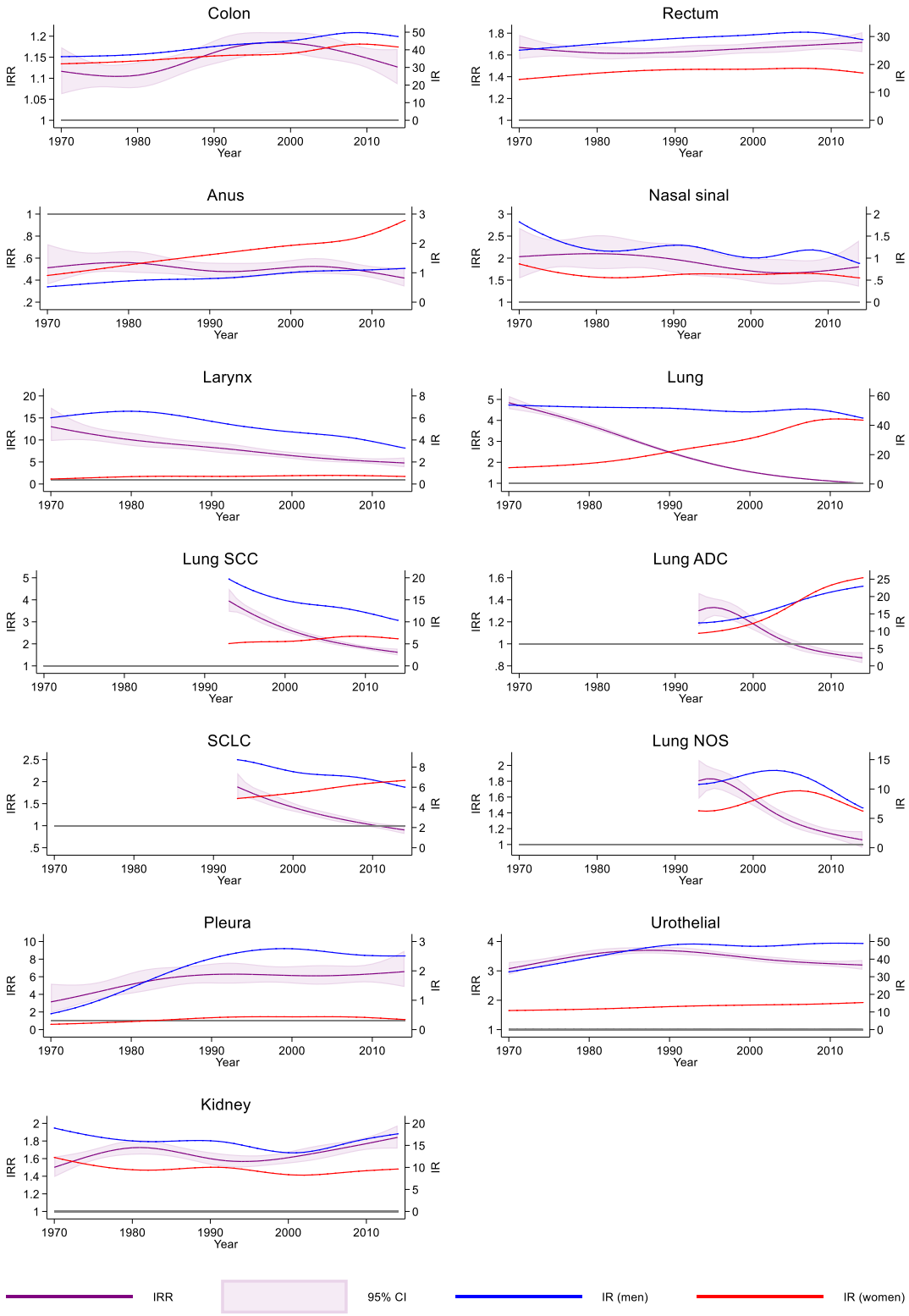
The incidence of most cancer sites increase over calendar time in both sexes while the excess cancer risk in men is more or less stable (Figure 8, page 34). The incidence of malignancies associated with smoking and/or alcohol, i.e., respiratory and head and neck tumors respectively, are decreasing slightly in both sexes. The slope is however steeper in men, resulting in a male-to-female IRR approaching, but not reaching 1. Pulmonary adenocarcinoma and small-cell lung cancer are the only exceptions and these tumors are nowadays more common in women than men (IRR < 1). Urothelial carcinoma is strongly associated with smoking, but incidence is not decreasing and moreover remains 3-4 times more common in men across the whole study period. Malignant melanoma incidence is the fastest increasing malignancy in both sexes and was more common in women in 1970's but is thereafter consistently 10% more common in men.

A majority of cancer sites increase with age in both sexes (Figure 9, page 37). The increase is more pronounced in men and the IRRs increase over age and peak around age 65-75. Some sites are more common in women up to age 40-50; salivary, colon, lung adenocarcinoma, small cell lung, other non-small cell lung, and malignant melanoma, suggesting different biological mechanisms in men and women. Biliary cancer is more common in men up to age 40, but is thereafter surpassed by women. The remaining female predominated sites; anus, meninges, and thyroid, are consistently more common in women across age.

Figure 10, page 40 shows 5-year RSR in men and women together with the male-to-female EMR by year of diagnosis. Survival has improved over calendar time for most malignancies and in both sexes. The female survival disadvantage is still more or less noticeable across the whole study period. The pattern is however far from as consistent as when studying incidence. For some sites; tongue, pharynx, gastric, pancreas, small intestine, colon, rectum, kidney, brain, multiple myeloma, ALL, and AML, survival differences have evened out over the last 5-10 years, indicating temporal changes in patient behavior and/or clinical management. Interestingly the largest survival difference is seen in sites that are both more common in women and have a very good prognosis (5-year relative survival > 80%), i.e., well-differentiated thyroid, meninges, salivary, melanoma, skin non-melanoma, NHL, CLL, and Hodgkin lymphoma. Perhaps indicating that pre-malign and/or earlier stage tumors are more frequently diagnosed in women. Biliary cancer stands out as exception and is both more common and deadlier in women, but also implies a very poor prognosis (5-year relative survival < 10%).

*Figure 8. Age-adjusted male-to-female incidence rate ratio (IRR) and age-standardized incidence rate (IR) per 100,000 person-years in men and women, by calendar year.*




Colon
Rectum
Anus
Nasal sinal
Larynx
Lung
Lung SCC
Lung ADC
SCLC
Lung NOS
Pleura
Urothelial
Kidney
IRR
IR
Year
1970
1980
1990
2000
2010
IRR
95% CI
IR (men)
IR (women)

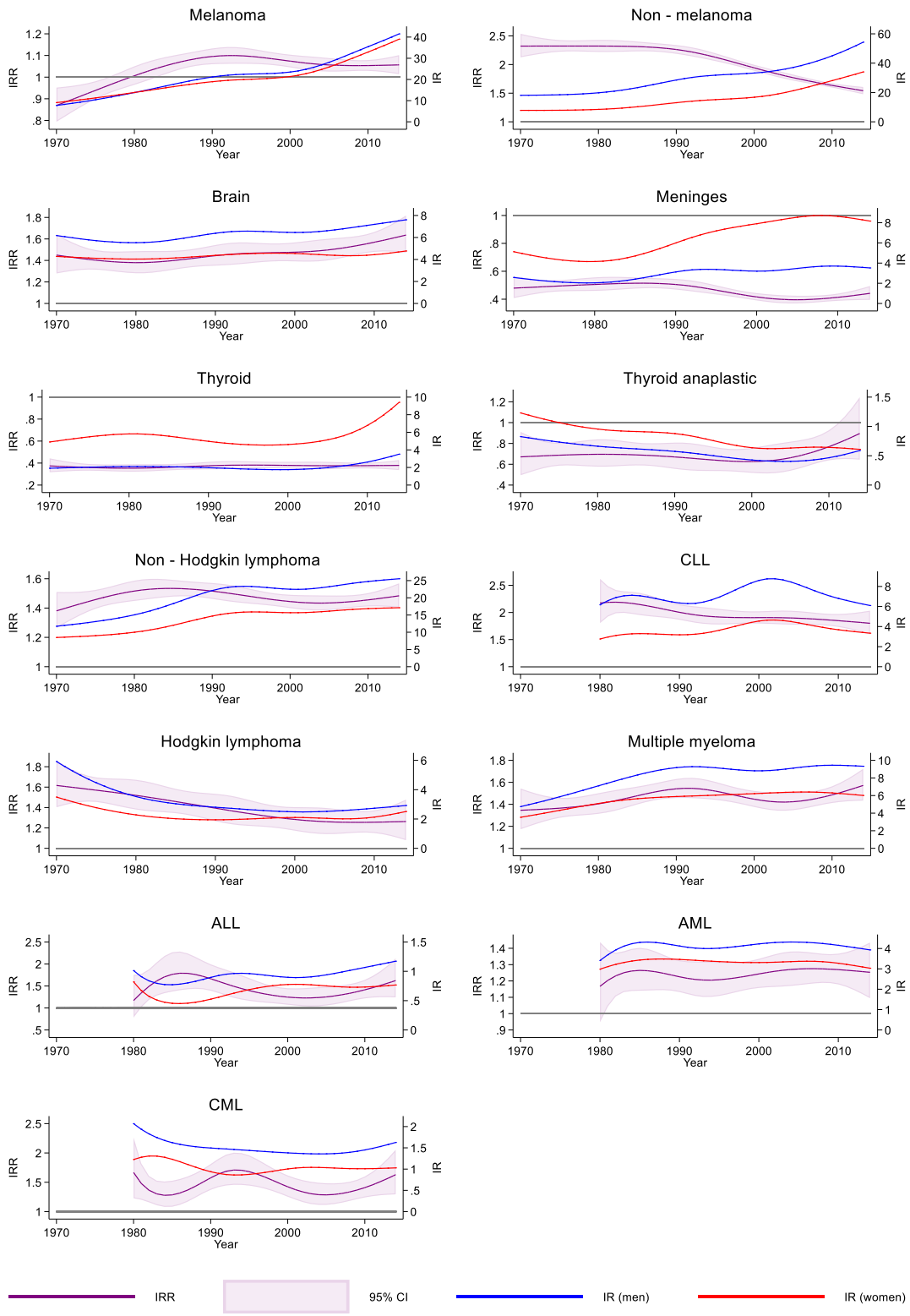
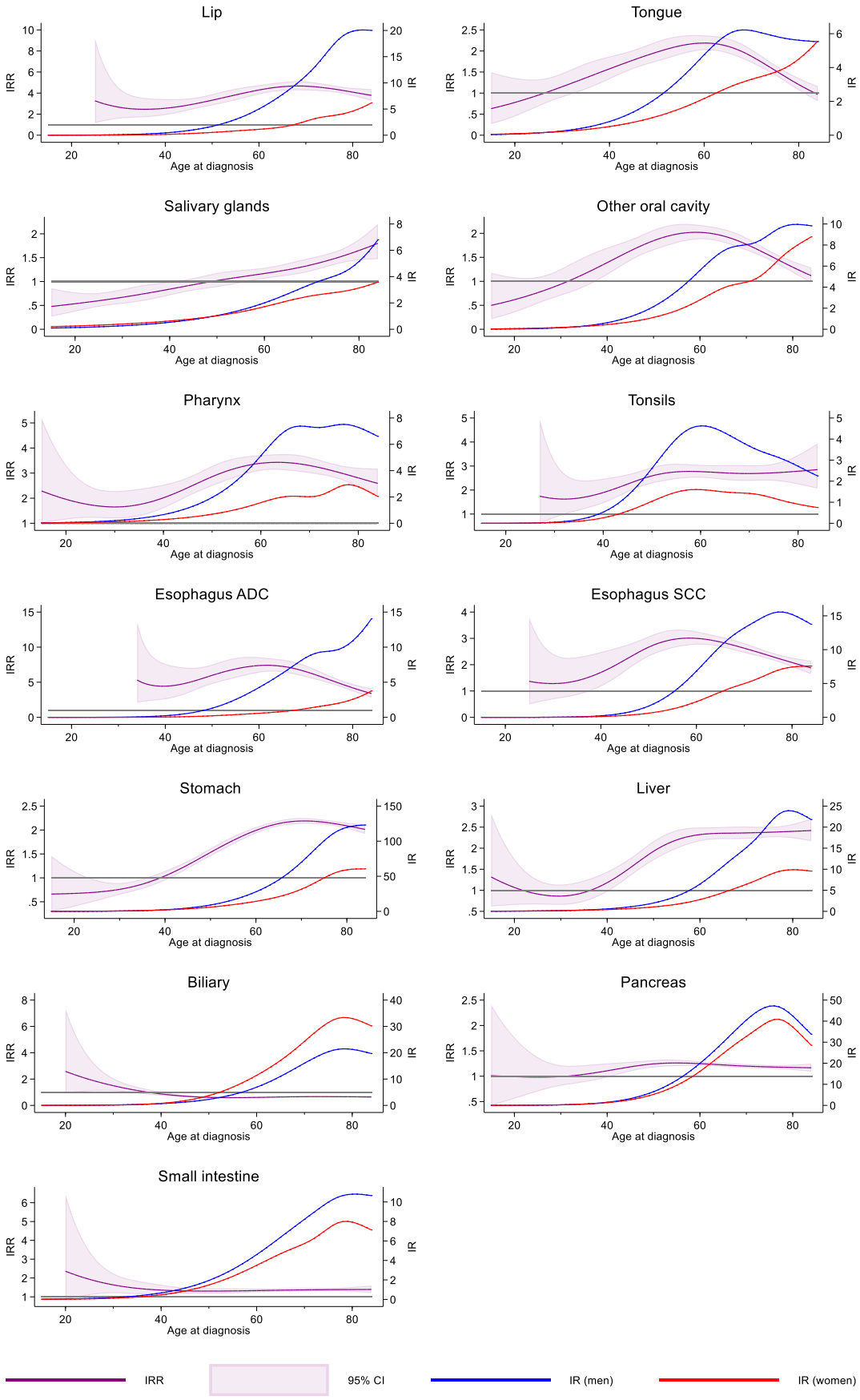
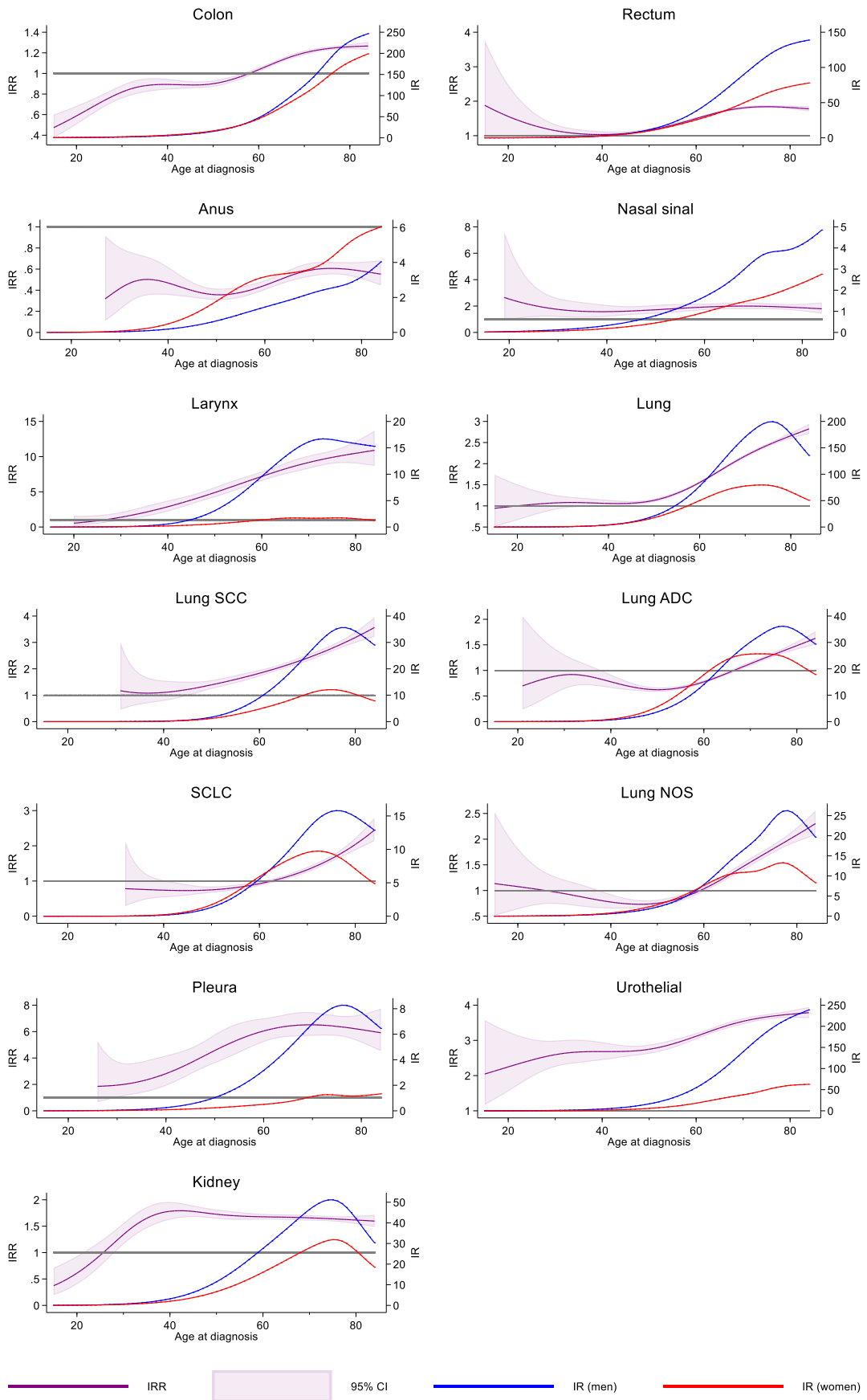

Melanoma
Non - melanoma
Brain
Meninges
Thyroid
Thyroid anaplastic
Non - Hodgkin lymphoma
CLL
Hodgkin lymphoma
Multiple myeloma
ALL
AML
CML
IRR
IR
Year
1970
1980
1990
2000
2010
IRR
95% CI
IR (men)
IR (women)

*Figure 9. Male-to-female incidence rate ratio (IRR) adjusted for year of diagnosis and incidence rate (IR) per 100,000 person-years in men and women by age at diagnosis.*




Colon
Rectum
Anus
Nasal sinal
Larynx
Lung
Lung SCC
Lung ADC
SCLC
Lung NOS
Pleura
Urothelial
Kidney
IRR
IR
Age at diagnosis
IRR
95% CI
IR (men)
IR (women)

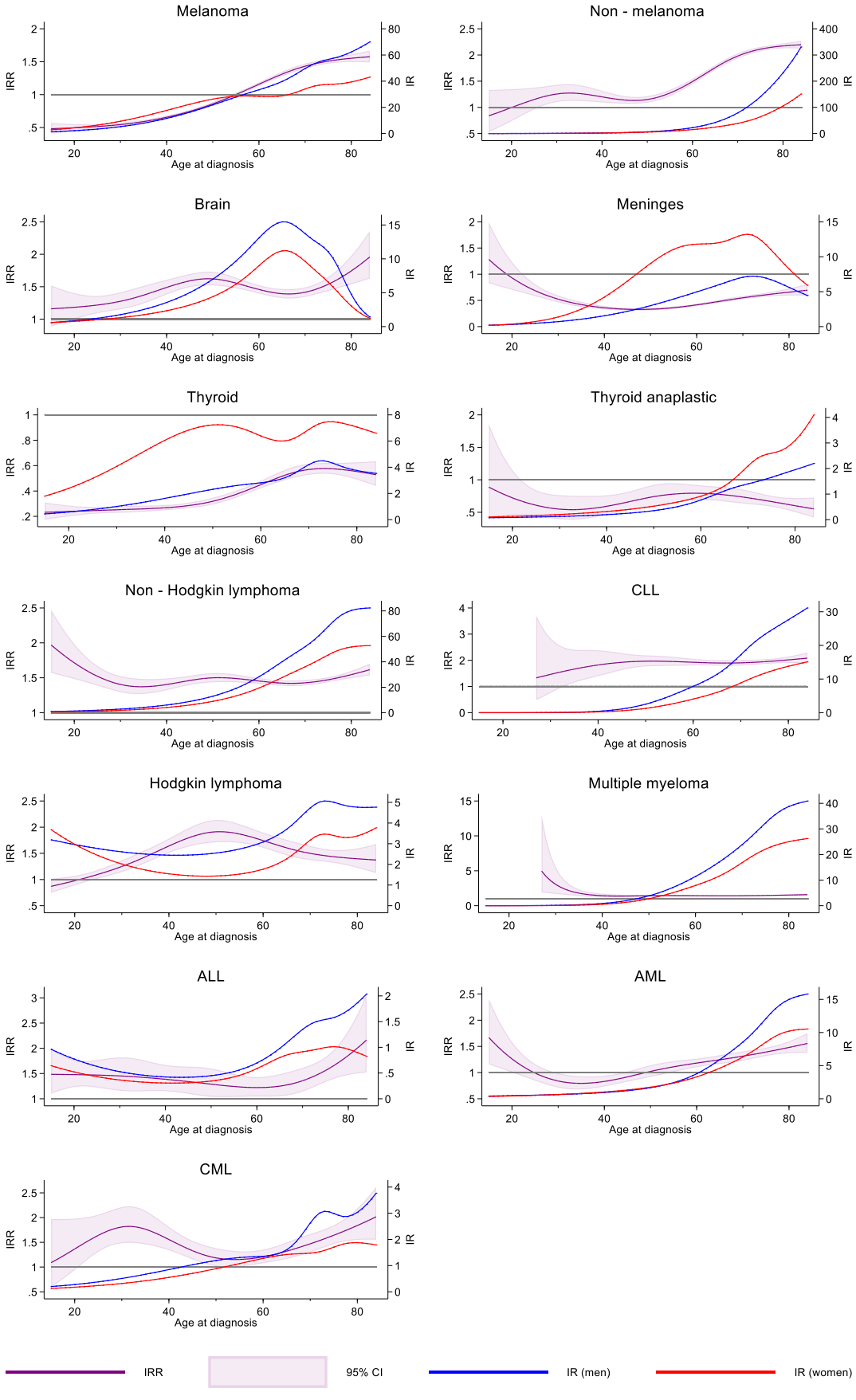
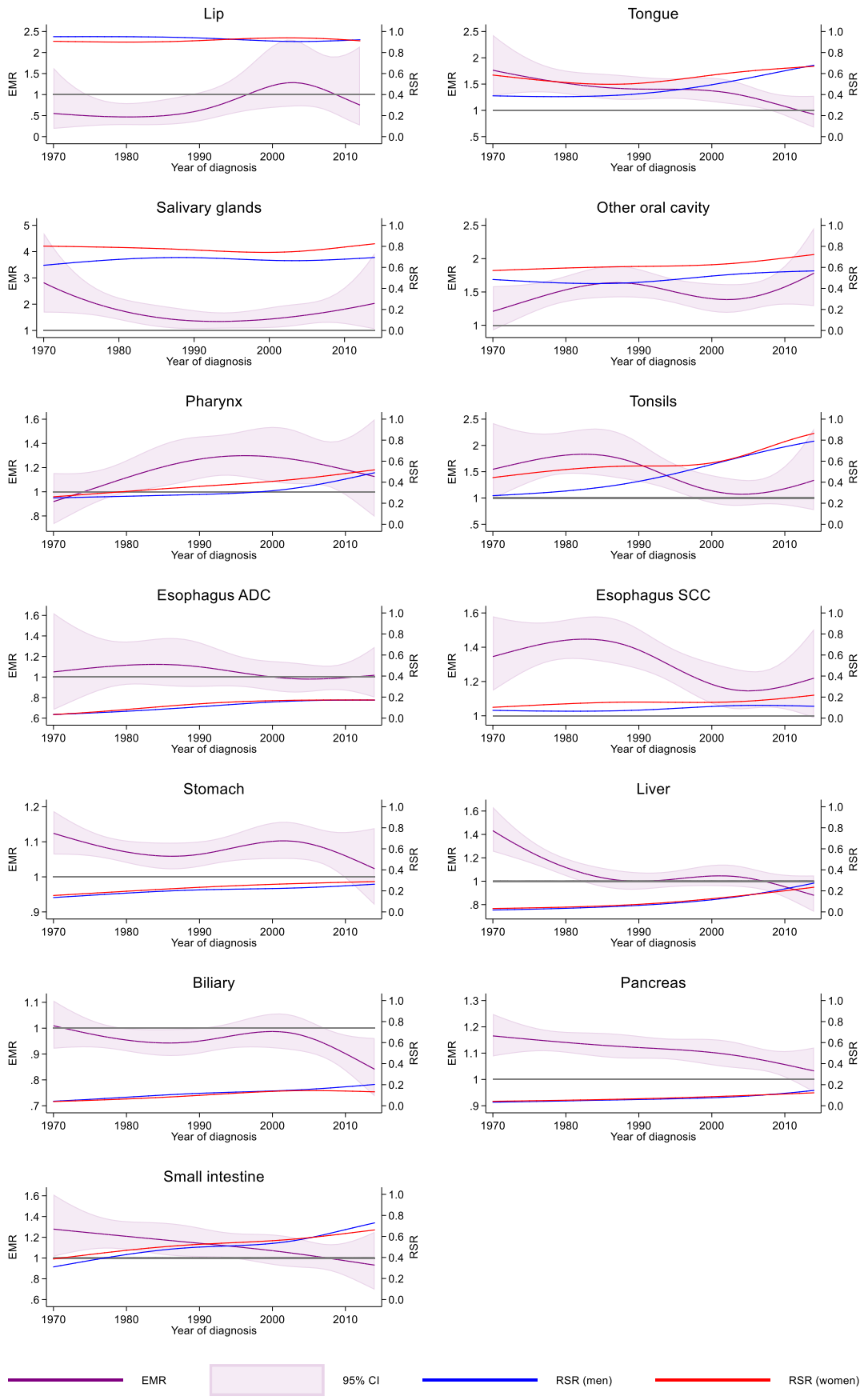
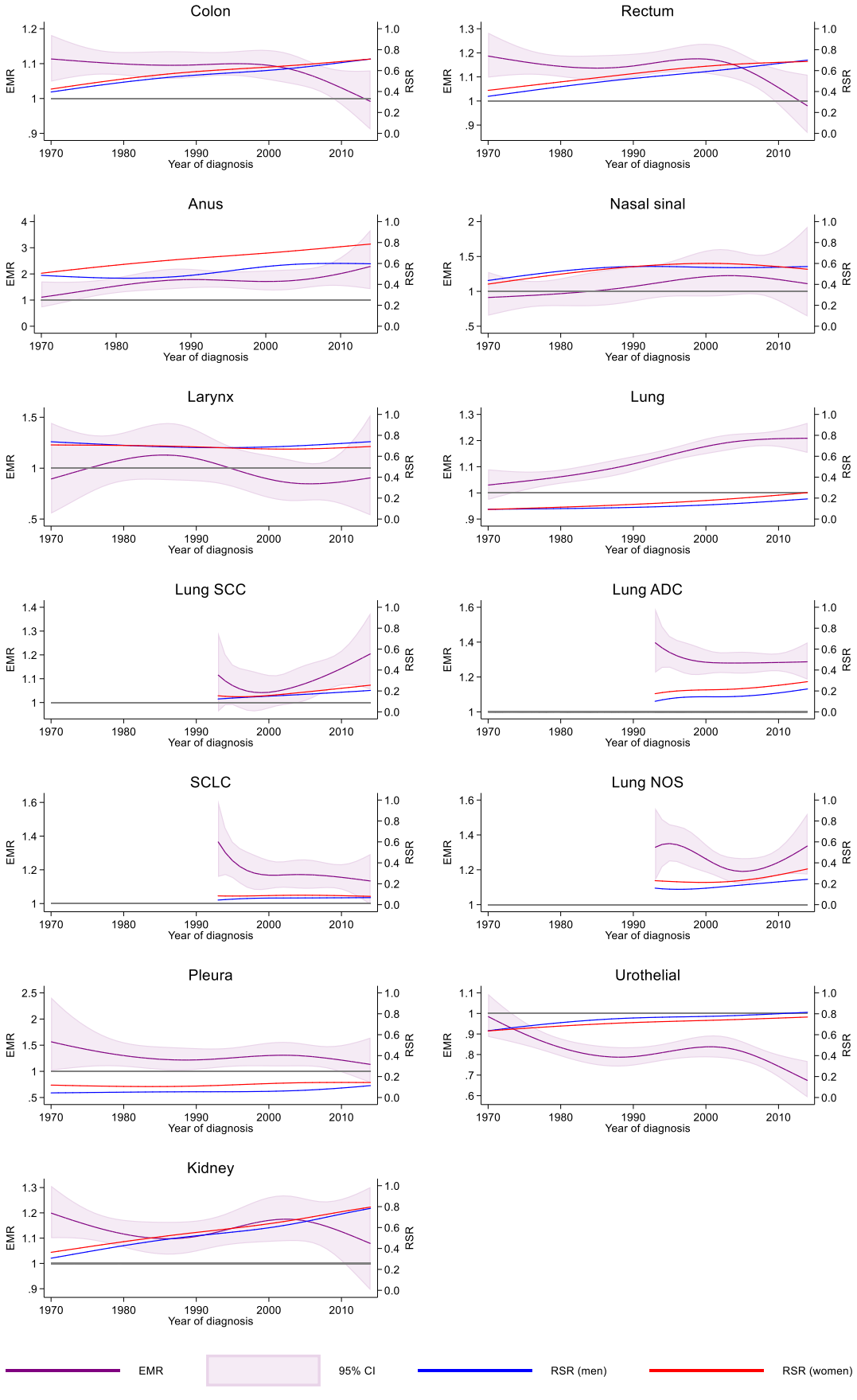
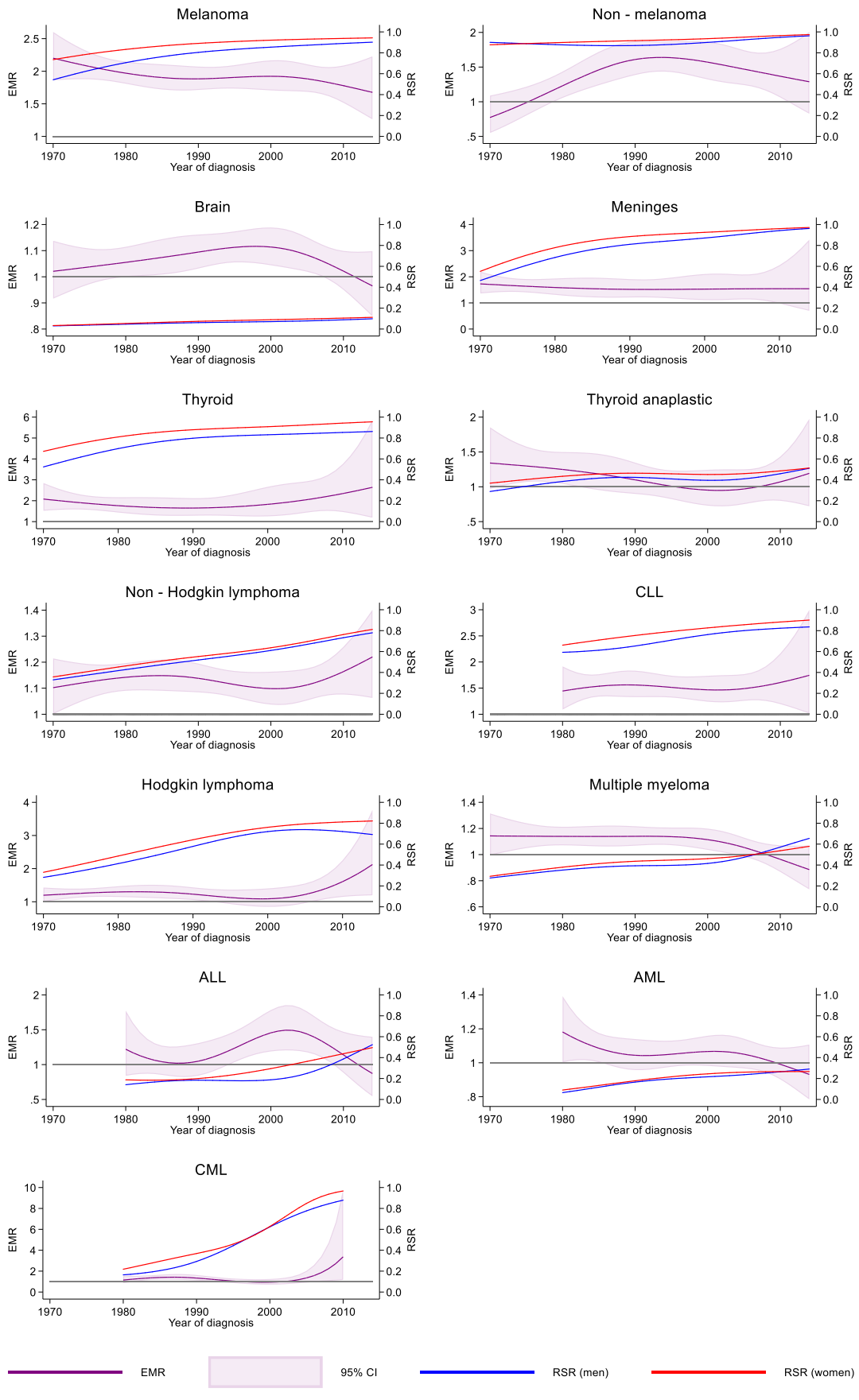

Melanoma
Non - melanoma
Brain
Meninges
Thyroid
Thyroid anaplastic
Non - Hodgkin lymphoma
CLL
Hodgkin lymphoma
Multiple myeloma
ALL
AML
CML
IRR
IR
Age at diagnosis
IRR
95% CI
IR (men)
IR (women)

*Figure 10. Age-adjusted male-to-female excess mortality ratio (EMR) and age-standardized 5-year relative survival ratio (RSR) for men and women, by calendar year.*




Colon
Rectum
Anus
Nasal sinal
Larynx
Lung
Lung SCC
Lung ADC
SCLC
Lung NOS
Pleura
Urothelial
Kidney
EMR
RSR
Year of diagnosis
1970
1980
1990
2000
2010
EMR
95% CI
RSR (men)
RSR (women)


Melanoma
Non - melanoma
Brain
Meninges
Thyroid
Thyroid anaplastic
Non - Hodgkin lymphoma
CLL
Hodgkin lymphoma
Multiple myeloma
ALL
AML
CML
EMR
RSR
Year of diagnosis
1970
1980
1990
2000
2010
EMR
95% CI
RSR (men)
RSR (women)

From a public health perspective, it is appealing to quantify the impact of the excess cancer risk in men in the population. With this in mind we calculated the population attributable risk percent (PAR%), which is defined as the proportion of incident cancers in the total population that can be attributed to sex differences in cancer risk (Table 6). The PAR% estimates ranged from negative values for cancer sites that are more common in women (biliary, anal, meningeal, thyroid) through 2% (lung adenocarcinoma) to 77% (laryngeal).

*Table 6. Proportion of cases attributed to male sex, i.e., population attributable risk percent (PAR%) with 95% confidence interval (CI).*

| Anatomical tract | Cancer site | PAR% (95% CI) |
|---|---|---|
| Head and neck | Lip | 61 (60-63) |
| | Tongue | 28 (26-31) |
| | Salivary glands | 9 (6-12) |
| | Other oral cavity | 26 (23-28) |
| | Pharyngeal | 51 (49-54) |
| | Tonsillar | 46 (44-49) |
| Upper digestive tract | Esophagus adenocarcinoma | 70 (68-72) |
| | Esophagus squamous cell carcinoma | 44 (42-46) |
| | Gastric | 34 (33-34) |
| | Liver primary | 39 (37-41) |
| | Biliary tract | -20 (-21--18) |
| | Pancreatic | 11 (10-12) |
| Lower digestive tract | Small intestine | 16 (14-18) |
| | Colon | 7 (7-8) |
| | Rectal | 25 (25-26) |
| | Anal | -32 (-34--29) |
| Respiratory organs | Nasal cavity/sinuses | 31 (28-35) |
| | Laryngeal | 77 (76-79) |
| | Lung (all) | 33 (32-33) |
| | Lung squamous cell carcinoma | 41 (40-43) |
| | Lung adenocarcinoma | 2 (1-4) |
| | Lung small cell carcinoma | 13 (11-15) |
| | Lung other non-small cell | 17 (15-18) |
| | Pleura mesothelioma | 71 (69-73) |
| Urinary | Urothelial | 55 (55-56) |
| | Renal | 26 (25-27) |
| Skin, central nervous system, thyroid | Skin melanoma | 3 (2-4) |
| | Skin non-melanoma | 31 (31-32) |
| | Brain | 21 (19-22) |
| | Meningeal | -36 (-37--34) |
| | Thyroid well-differentiated | -44 (-46--43) |
| | Thyroid anaplastic | -18 (-22--14) |
| Hematological malignancies | Non-Hodgkin lymphoma | 19 (19-20) |
| | Chronic lymphocytic leukemia | 32 (31-34) |
| | Hodgkin lymphoma | 16 (14-18) |
| | Multiple myeloma | 20 (18-21) |
| | Acute lymphocytic leukemia | 16 (12-21) |
| | Acute myeloid leukemia | 10 (8-12) |
| | Chronic myeloid leukemia | 18 (15-22) |

## 4.2 STUDY II

Study II aimed to explore potential factors driving the observed male survival disadvantage in non-small cell lung cancer, see *1.3.5 Non-small cell lung cancer*, for a detailed introduction and background. Since pulmonary squamous cell carcinoma and adenocarcinoma are considered different disease entities with different epidemiology, clinical management, and prognosis, all analysis were stratified on cell type.

We identified 33,790 cases of lung squamous cell carcinoma (n = 10,325) and adenocarcinoma (n = 23,465) diagnosed in Sweden in 2002-2016, at age ≥ 20 (see *3.1.2.1 The National Lung Cancer Register and the Lung Cancer DataBase Sweden*). Tables 1 and 2 in study II provide an overview of sex differences in clinicopathological and socioeconomic factors, and treatment intensity, respectively (191). Men with NSCLC were older, less educated, and presented with poorer performance status and at a more advanced stage, upon diagnosis. Women with lung adenocarcinoma additionally presented with less comorbidity and were more often never-smokers. No, or minor sex differences, were found comparing treatment intensity in models adjusted for age and calendar time nor fully-adjusted. Our main finding, presented as adjusted lung cancer-specific hazard ratios and standardized survival curves, study II, table 3 and figure 2, respectively, concluded that men with NSCLC have a consistently poorer prognosis compared to women which was most pronounced in lung adenocarcinoma, and that cannot be explained by a range of prognostic factors (191).

To fully explore if and how clinical management differs in men and women, we assessed investigational intensity (Table 7). We found no or very minor sex differences regarding clinical management of lung squamous cell carcinoma. In lung adenocarcinoma, thoracentesis was more often performed in men and positron emission tomography (PET) scans more frequently in women, conceivably reflecting different symptomatology and disease spread in men and women. In lung adenocarcinoma stage IIIB-IV, year of diagnosis 2010-2016, EGFR testing was slightly more common in women.

We additionally compared different waiting times in men and women diagnosed with lung squamous cell carcinoma and adenocarcinoma by plotting the cumulative proportion of patients by days from referral-to-diagnosis, diagnosis-to-treatment, and referral-to-treatment (Figure 11). No sex differences was obvious at visual inspection of the graphs. It was however striking that only 36-37% of NSCLC patients received a final treatment decision and/or initiated treatment within 28 days from referral (Table 7), a strong recommendation according to Swedish guidelines (79).

A recently discovered favorable prognostic factor, as well as predictive of response to tyrosine kinase inhibitor therapy, in lung adenocarcinoma is activating EGFR mutations. EGFR positive tumors have also been shown to be more common in women compared to men. EGFR status is included in the NLCR from year 2010 and onwards. We performed a subgroup analysis of lung cancer specific mortality in patients diagnosed in 2010-2016 and tested for EGFR, adding a third model additionally adjusted for EGFR status (Table 8). The

female-to-male hazard ratios remained close to identical after adjusting for EGFR status in all stage groups and consequently EGFR status fails to explain the superior survival in women with lung adenocarcinoma.

As a sensitivity analysis we explored models including an interaction term between sex and subset of variables, i.e., the effect of sex was allowed to vary over different covariate categories (Table 9). The presented p-value compares model fit with and without interaction, a p-value < 0.05 indicates that the latter model is superior. With the exception of smoking history in advanced stage lung adenocarcinoma, we did not find any evidence of a consistent trend together with significantly better model fit, when allowing for interaction. The female survival advantage was more pronounced in non-smokers (never or former) compared to smokers. The interaction between sex and birth country was based on very few non-Scandinavian study participants and considered to be a random finding.

*Table 7. Numbers (n), percentages (%) of men and women diagnosed with non-small cell lung cancer and female-to male odds ratios (ORs) with 95% confidence intervals (CIs), undergoing diagnostic procedures, by histological subtype.*

| | **Squamous cell carcinoma** | | | | | | |
|---|---|---|---|---|---|---|---|
| | **Men** | | | **Women** | | | |
| | n | % | OR (95% CI) | n | % | OR (95% CI)[1] | aOR (95% CI)[2] |
| Bronchoscopy | 5413 | 82.6 | 1.00 (ref.) | 3011 | 79.9 | 0.84 [0.76,0.94] | 0.86 [0.76,0.96] |
| CT thorax | 6398 | 97.6 | 1.00 (ref.) | 3676 | 97.5 | 0.97 [0.73,1.28] | 0.89 [0.66,1.22] |
| US/CT abdomen | 5813 | 88.7 | 1.00 (ref.) | 3317 | 88.0 | 0.95 [0.84,1.08] | 0.97 [0.84,1.11] |
| Thoracentesis | 421 | 6.4 | 1.00 (ref.) | 205 | 5.4 | 0.84 [0.71,1.00] | 0.83 [0.68,1.00] |
| Transthoracic biopsy | 1459 | 22.3 | 1.00 (ref.) | 1075 | 28.5 | 1.37 [1.25,1.51] | 1.32 [1.19,1.47] |
| CT/MRI brain[3] | 1021 | 23.6 | 1.00 (ref.) | 607 | 22.9 | 0.93 [0.83,1.05] | 0.94 [0.82,1.07] |
| PET scan[3] | 1978 | 45.7 | 1.00 (ref.) | 1305 | 49.2 | 1.10 [0.99,1.22] | 1.07 [0.94,1.22] |
| Multidisciplinary case conference | 4207 | 64.2 | 1.00 (ref.) | 2506 | 66.5 | 1.03 [0.94,1.13] | 1.00 [0.90,1.12] |
| Treatment-on-time[4] | 2375 | 36.2 | 1.00 (ref.) | 1331 | 35.3 | 0.98 [0.90,1.07] | 1.02 [0.92,1.12] |
| | **Adenocarcinoma** | | | | | | |
| | **Men** | | | **Women** | | | |
| | n | % | OR (95% CI) | n | % | OR (95% CI)[1] | aOR (95% CI)[2] |
| Bronchoscopy | 7487 | 69.4 | 1.00 (ref.) | 8886 | 70.1 | 1.03 [0.97,1.09] | 1.01 [0.95,1.08] |
| CT thorax | 10531 | 97.6 | 1.00 (ref.) | 12343 | 97.4 | 0.84 [0.71,1.00] | 0.82 [0.67,0.99] |
| US/CT abdomen | 9590 | 88.8 | 1.00 (ref.) | 11341 | 89.5 | 1.06 [0.97,1.15] | 1.07 [0.98,1.18] |
| Thoracentesis | 1967 | 18.2 | 1.00 (ref.) | 1903 | 15.0 | 0.82 [0.76,0.88] | 0.83 [0.76,0.90] |
| Transthoracic biopsy | 3191 | 29.6 | 1.00 (ref.) | 3722 | 29.4 | 0.99 [0.93,1.05] | 0.96 [0.90,1.02] |
| CT/MRI brain[3] | 2041 | 25.2 | 1.00 (ref.) | 2424 | 24.8 | 0.95 [0.89,1.02] | 1.01 [0.94,1.10] |
| PET scan[3] | 3367 | 41.6 | 1.00 (ref.) | 4460 | 45.6 | 1.15 [1.09,1.23] | 1.04 [0.95,1.12] |
| Multidisciplinary case conference | 6869 | 63.6 | 1.00 (ref.) | 8455 | 66.7 | 1.08 [1.02,1.14] | 1.00 [0.93,1.07] |
| Treatment-on-time[4] | 4034 | 37.4 | 1.00 (ref.) | 4630 | 36.5 | 0.94 [0.89,0.99] | 0.98 [0.92,1.04] |
| *EGFR* testing[5] | 3263 | 53.4 | 1.00 (ref.) | 4235 | 57.3 | 1.11 [1.03,1.19] | 1.10 [1.01,1.20] |

[1]Adjusted for age and calendar year of diagnosis. [2]Additionally adjusted for level of education, marital status, country of birth, health care region, ECOG performance status, smoking history, Elixhauser comorbidity categories, TNM stage, and primary tumor location. [3]Year of diagnosis 2007-2016. [4]Treatment within 28 days from referral. [5]Stage IIIB-IV, year of diagnosis 2010-2016.

*Figure 11. Proportion of non-small cell lung cancer patients waiting (days) from referral-to-diagnosis, diagnosis-to-treatment, and referral-to treatment (with dashed vertical line recommended waiting time < 28 days), by cell type.*

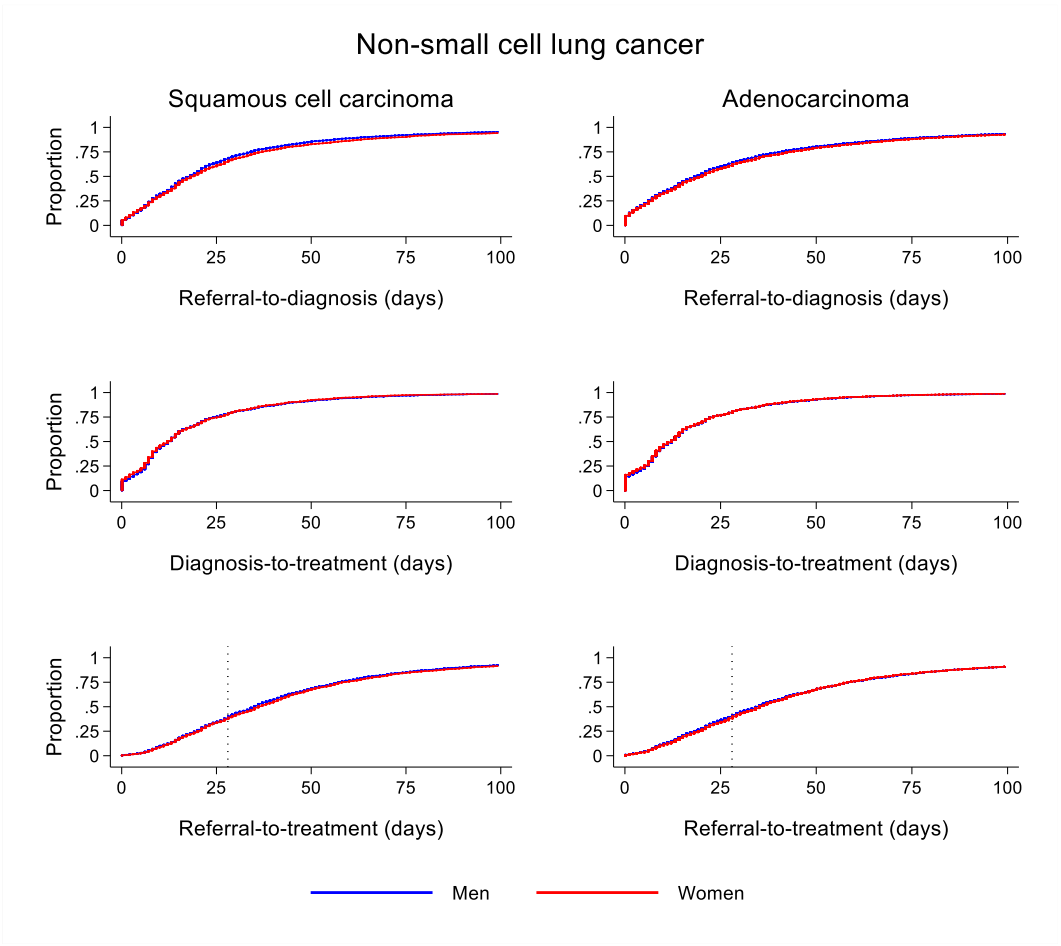



*Table 8. Subgroup analysis of lung adenocarcinoma diagnosed in 2010-2016 and tested for EGFR, by stage group. Numbers (n), percentages (%) of lung cancer deaths and female-to-male hazard ratios (HRs) with 95% confidence intervals (CIs).*

| | n | % | HR (95% CI)[1] | HR (95% CI)[2] | HR (95% CI)[3] |
|---|---|---|---|---|---|
| **Stage IA-IIB** | | | | | |
| Men | 149 | 22.6 | 1.00 | 1.00 | 1.00 |
| Women | 157 | 16.8 | 0.64 [0.51,0.80] | 0.65 [0.50,0.85] | 0.65 [0.50,0.85] |
| **Stage IIIA** | | | | | |
| Men | 138 | 49.3 | 1.00 | 1.00 | 1.00 |
| Women | 177 | 46.0 | 0.88 [0.70,1.10] | 0.91 [0.69,1.18] | 0.93 [0.71,1.21] |
| **Stage IIIB-IV** | | | | | |
| Men | 1665 | 73.0 | 1.00 | 1.00 | 1.00 |
| Women | 2012 | 70.0 | 0.85 [0.79,0.90] | 0.88 [0.82,0.94] | 0.90 [0.84,0.97] |

[1]Adjusted for age and calendar year of diagnosis. [2]Additionally adjusted for level of education, marital status, country of birth, health care region, ECOG performance status, smoking history, Elixhauser comorbidity categories, TNM stage, and primary tumor location. [3]Additionally adjusted for EGFR status.

*Table 9. Adjusted* female-to-male hazard ratios (HRs) by non-small cell lung cancer subtype and stage, exploring interaction between female sex and selected covariates and model fit (p-value) compared to the original model.*

| | Squamous cell carcinoma | | | | | | Adenocarcinoma | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | stage IA-IIB | | Stage IIIA | | Stage IIIB-IV | | stage IA-IIB | | stage IIIA | | stage IIIB-IV | |
| | HR | p-value | HR | p-value | HR | p-value | HR | p-value | HR | p-value | HR | p-value |
| **Age** | | | | | | | | | | | | |
| 20-59 | 0.75 | | 1.45 | | 0.99 | | 0.71 | | 0.87 | | 0.96 | |
| 60-69 | 0.90 | | 0.91 | | 0.78 | | 0.69 | | 0.85 | | 0.78 | |
| 70-79 | 0.78 | | 0.93 | | 0.88 | | 0.72 | | 0.73 | | 0.80 | |
| 80-89 | 0.73 | | 1.32 | | 0.89 | | 0.76 | | 0.73 | | 0.97 | |
| 90+ | 1.93 | 0.455 | 1.88 | 0.150 | 1.70 | 0.057 | 1.26 | 0.876 | 0.00 | 0.676 | 0.62 | 0.000 |
| **Year of diagnosis** | | | | | | | | | | | | |
| 2002-2006 | 0.85 | | 0.94 | | 0.82 | | 0.75 | | 0.68 | | 0.85 | |
| 2007-2011 | 0.81 | | 1.09 | | 0.91 | | 0.72 | | 0.73 | | 0.82 | |
| 2012-2016 | 0.73 | 0.646 | 1.07 | 0.648 | 0.86 | 0.349 | 0.69 | 0.791 | 0.93 | 0.114 | 0.85 | 0.701 |
| **Education** | | | | | | | | | | | | |
| low | 0.79 | | 1.29 | | 0.87 | | 0.73 | | 0.80 | | 0.87 | |
| middle | 0.87 | | 0.93 | | 0.82 | | 0.68 | | 0.78 | | 0.82 | |
| high | 0.67 | | 0.67 | | 0.97 | | 0.78 | | 0.80 | | 0.82 | |
| missing | 0.78 | 0.706 | 0.63 | 0.010 | 0.90 | 0.409 | 0.80 | 0.832 | 0.71 | 0.994 | 0.74 | 0.241 |
| **Origin** | | | | | | | | | | | | |
| Scandinavian | 0.78 | | 1.01 | | 0.86 | | 0.73 | | 0.80 | | 0.84 | |
| European | 1.24 | | 1.15 | | 0.99 | | 0.77 | | 0.63 | | 0.76 | |
| Non-European | 3.24 | 0.017 | 2.41 | | 0.56 | | 0.14 | | 0.77 | 0.690 | 0.87 | |
| missing | - | | 1.76 | 0.456 | 0.98 | 0.323 | 0.61 | 0.058 | - | | 1.02 | 0.576 |
| **Performance status** | | | | | | | | | | | | |
| 0 | 0.67 | | 0.86 | | 0.80 | | 0.71 | | 0.81 | | 0.86 | |
| 1 | 0.81 | | 1.06 | | 0.88 | | 0.70 | | 0.81 | | 0.84 | |
| 2 | 0.79 | | 0.99 | | 0.94 | | 0.92 | | 0.79 | | 0.85 | |
| 3 | 1.14 | | 1.12 | | 0.89 | | 0.61 | | 0.49 | | 0.83 | |
| 4 | 1.02 | | 2.53 | | 0.51 | | 0.23 | | 1.66 | | 0.86 | |
| missing | 0.72 | 0.418 | 1.27 | 0.513 | 0.93 | 0.001 | 0.49 | 0.106 | 0.42 | 0.264 | 0.68 | 0.303 |
| **Smoking history** | | | | | | | | | | | | |
| Smoker | 0.79 | | 1.18 | | 0.85 | | 0.75 | | 0.96 | | 0.79 | |
| Former smoker | 0.80 | | 0.89 | | 0.87 | | 0.69 | | 0.65 | | 0.84 | |
| Never smoker | 0.62 | | 0.98 | | 0.90 | | 0.72 | | 0.79 | | 0.95 | |
| missing | 1.88 | 0.179 | 1.20 | 0.263 | 1.07 | 0.753 | 0.66 | 0.879 | 1.11 | 0.031 | 0.93 | 0.010 |
| **Elixhauser comorbidities** | | | | | | | | | | | | |
| 0 | 0.77 | | 0.98 | | 0.87 | | 0.86 | | 0.81 | | 0.84 | |
| 1-2 | 0.82 | | 1.01 | | 0.86 | | 0.66 | | 0.77 | | 0.83 | |
| 2-3 | 0.85 | | 0.94 | | 0.86 | | 0.75 | | 0.71 | | 0.90 | |
| 5+ | 0.80 | 0.962 | 1.73 | 0.167 | 0.85 | 0.987 | 0.55 | 0.057 | 1.02 | 0.658 | 0.84 | 0.563 |

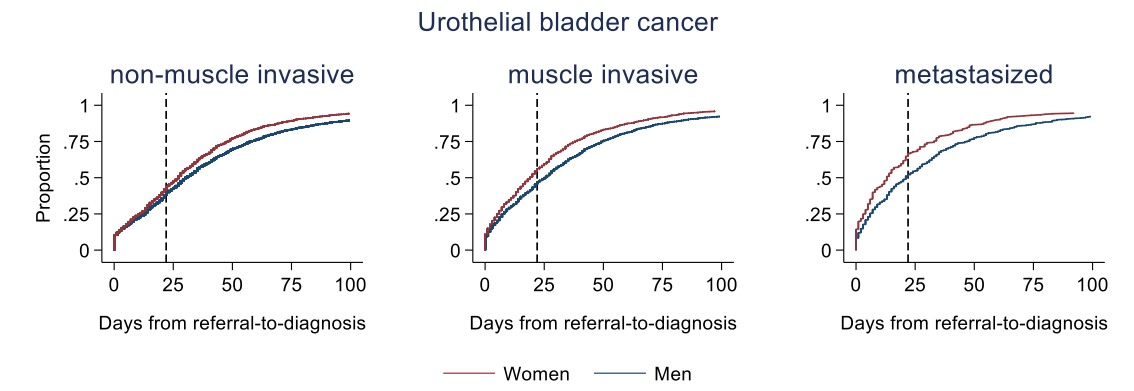
*age, calendar year, educational level, marital status, birth country, health care region, ECOG performance status, smoking history, Elixhauser comorbidity groups, TNM stage, and primary tumor location.

## 4.3 STUDY III

Urinary bladder cancer is one of few malignancies where women have a consistently poorer prognosis, see *1.3.6 Urinary bladder cancer* for a detailed background. With the use of a comprehensive research database, BladderBaSe, we identified 36,344 Swedish men and women diagnosed with urothelial bladder cancer (UBC), at age 18-89, in year 1997-2014 (see *3.1.2.2 The Swedish National Register of Urinary Bladder Cancer and the Bladder Cancer DataBase Sweden*). Due to distinct clinical management, treatment and prognosis, we stratified all analyses into three mutually exclusive stage groups at diagnosis; non-muscle invasive (NMIBC: T0/Tis/Ta/T1, any N, M0), muscle invasive (MIBC: T2/T3/T4, any N, M0), and primarily metastasized (M1) UBC, see *1.3.6 Urinary bladder cancer*.

Main findings are presented in detail in the attached manuscript, study III (192). In summary we found the female survival disadvantage to be limited to MIBC, only noticeable within the first two years from diagnosis, and robust for adjustments for a range of prognostic factors. With the exception of an adverse stage distribution in women, we did not find any evidence of an inferior management of women with UBC, compared to men (192).

*Figure 12. Proportion of urothelial bladder cancer patients waiting (days) from referral-to-diagnosis (with dashed vertical line recommended waiting time < 22 days), by stage group.*




In addition to diagnostic and treatment intensity (table 2, study III), we investigated time from referral from primary care to UBC diagnosis, i.e., date of transurethral resection of the bladder tumor (TURBT). We plotted the cumulative proportion diagnosed by days from referral, noting that women experienced slightly shorter waiting times compared to men (Figure 12). More strikingly, only 43% and 48% of the men and women, respectively underwent TURBT within 22 days from referral, a strong recommendation according to Swedish guidelines (124).

As a sensitivity analysis, we explored the interaction between sex and a subset of prognostic factors in MIBC (Table 10). We found the effect of sex to be relatively stable across age groups, educational level, comorbidity, and N stage. There was no evidence of a superior model fit compared to the fully adjusted model without the interaction term. T stage was the only exception and the excess female mortality was driven by T4 tumors. This finding urged us to stratify survival analysis of MIBC into T2, T3, and T4, discovering that women

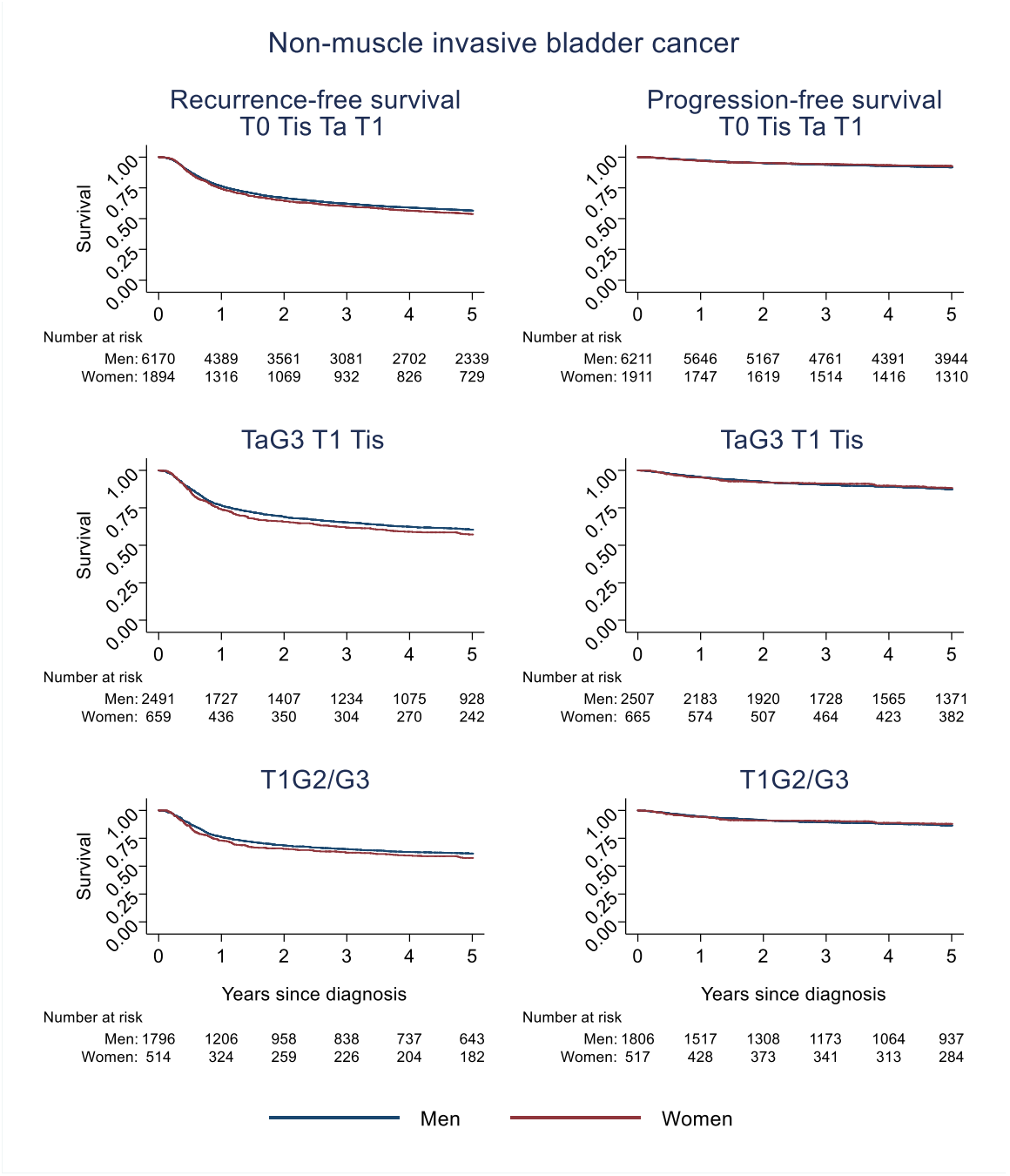
diagnosed with non-metastasized T4 UBC had close to the same mortality as primarily metastasized UBC!

*Table 10. Adjusted female-to-male hazard ratios (HRs) with 95% confidence intervals (CIs), exploring interaction between sex and selected covariates and comparing model fit (p-value) to the fully adjusted original model (aHR2 in table 3, study III) in muscle invasive bladder cancer.*

| | | HR | 95% CI | | p-value |
|---|---|---|---|---|---|
| **Age groups** | | | | | |
| | 18-59 | 1.24 | 0.97 | 1.59 | |
| | 60-69 | 1.13 | 0.95 | 1.34 | |
| | 70-79 | 1.19 | 1.05 | 1.36 | |
| | 80-89 | 1.26 | 1.12 | 1.43 | 0.7211 |
| **Educational level** | | | | | |
| | low | 1.28 | 1.15 | 1.42 | |
| | medium | 1.20 | 1.05 | 1.38 | |
| | high | 1.10 | 0.88 | 1.37 | |
| | missing | 0.90 | 0.65 | 1.25 | 0.1503 |
| **Charlson Comorbidity Index** | | | | | |
| | 0 | 1.19 | 1.07 | 1.32 | |
| | 1 | 1.26 | 1.07 | 1.48 | |
| | 2 | 1.34 | 1.09 | 1.66 | |
| | 3+ | 1.13 | 0.90 | 1.43 | 0.6367 |
| **T stage** | | | | | |
| | T2 | 1.17 | 1.06 | 1.30 | |
| | T3 | 1.09 | 0.93 | 1.26 | |
| | T4 | 1.59 | 1.33 | 1.90 | 0.0028 |
| **N stage** | | | | | |
| | N0 | 1.22 | 1.08 | 1.39 | |
| | N+ | 1.26 | 1.03 | 1.54 | |
| | NX | 1.19 | 1.07 | 1.33 | 0.8858 |

It has been hypothesized that UBC tumor behavior is more aggressive in women and that sex differences in tumor biology underlie the adverse stage distribution and poorer UBC survival in women. We therefore decided to explore recurrence-free and progression-free survival in men and women diagnosed with NMIBC by plotting Kaplan-Meier survival curves (Figure 13). This resulted in essentially overlapping curves and did not provide any evidence sex difference in non-muscle invasive UBC tumor biology.

*Figure 13. Recurrence- and progression-free survival in men and women diagnosed with non-muscle invasive (T0 Tis Ta T1) urothelial bladder cancer and by risk group.*



## 4.4 STUDY IV

Since both male sex and tall body stature are associated with increased cancer risk and since men are taller than women, we decided to apply mediation analysis to explore to what extent body height can explain the excess cancer risk in men. The cancer-free survival in men and women was compared using age as the underlying time scale. Cancer-free survival time was counted from date of first adult (age ≥18) height measurement to date of cancer diagnosis. All models were adjusted for birth year and educational level. See *3.2.4 Causal inference on time-to-event outcomes* for a detailed background to the methodology.

Our main results are presented in the attached manuscript, study IV panel plot E (205). This figure is restricted to cancer sites that were found to be both more common in men as well as associated with increased body height. The proportion of the excess cancer risk in men explained by taller body stature ranged from 0.5% (laryngeal) to 100% (salivary gland, colon, melanoma, AML).The following tables and figures provide additional support and guidance to how and why we ended up with the final results.

From panel plot B and C (study IV) it is clear that men are taller than women and that height has increased over calendar time in both sexes. We additionally wanted to assess whether height is associated with socioeconomic status (i.e., educational level) and if this relationship is consistent over calendar time. Mean height by sex and educational level was subsequently plotted over birth year 1900-1992 (Figure 14). Aside from instable estimates due to few individuals in the oldest cohort born before year 1910, we observed a remarkable consistency where the most highly educated men and women were approximately 5 cm taller than the least educated across birth year.

*Figure 14. Mean body height (cm) in men and women over year of birth, by educational level.*

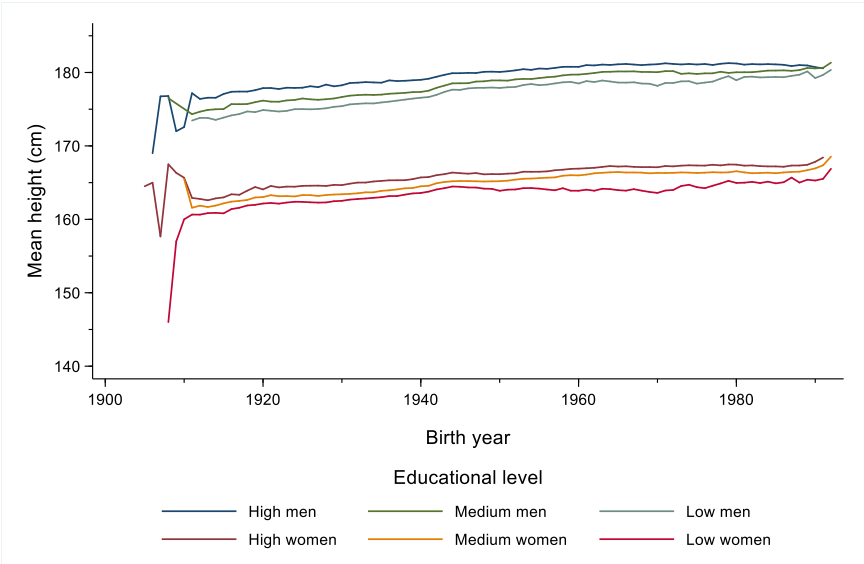



Table 11 provides an overview of numbers and percentages of men and women diagnosed with cancer, male-to-female hazard ratios (HRs), and relative risk (RR) of cancer per 10 cm

height increase in both sexes, both estimates adjusted for birth year and educational level. Out of 39 cancer sites, men were at significantly increased risk of 33, compared to women. The only exceptions were biliary, anal, lung adenocarcinoma, meningeal, and both subtypes of thyroid cancer. Most malignancies were also found to be associated with height, whereof 27 statistically significant.

*Table 11. Numbers (n) and percentages (%) of men and women, male-to-female hazard ratio (HR), and relative cancer risk (RR) per 10 centimeter height increase, 95% confidence interval (CI), both estimates adjusted for birth year and education.*

| **Cancer site** | Men | | Women | | Male-to-female | Men and women |
|---|---|---|---|---|---|---|
| | n | % | n | % | HR [95% CI] | RR [95% CI] |
| Lip | 1048 | 63 | 619 | 37 | 1.94 [1.75,2.14] | 1.06 [ 0.98 , 1.15 ] |
| Tongue | 1167 | 61 | 751 | 39 | 1.61 [1.46,1.76] | 0.98 [ 0.92 , 1.06 ] |
| Salivary glands | 657 | 55 | 539 | 45 | 1.25 [1.12,1.41] | 1.25 [ 1.14 , 1.37 ] |
| Other oral cavity | 1140 | 56 | 906 | 44 | 1.35 [1.24,1.48] | 0.97 [ 0.90 , 1.04 ] |
| Pharyngeal | 920 | 74 | 325 | 26 | 2.92 [2.57,3.32] | 0.91 [ 0.83 , 1.00 ] |
| Tonsillar | 1215 | 74 | 422 | 26 | 2.91 [2.61,3.25] | 1.10 [ 1.02 , 1.19 ] |
| Esophageal ADC | 1561 | 85 | 274 | 15 | 6.38 [5.61,7.26] | 1.04 [ 0.97 , 1.12 ] |
| Esophageal SCC | 1189 | 65 | 651 | 35 | 2.04 [1.85,2.25] | 1.06 [ 0.98 , 1.14 ] |
| Gastric | 5988 | 63 | 3452 | 37 | 1.97 [1.89,2.05] | 0.94 [ 0.91 , 0.97 ] |
| Hepatic | 1827 | 71 | 751 | 29 | 2.70 [2.48,2.94] | 0.97 [ 0.91 , 1.03 ] |
| Biliary | 1689 | 42 | 2354 | 58 | 0.79 [0.74,0.84] | 1.14 [ 1.09 , 1.20 ] |
| Pancreatic | 4475 | 52 | 4220 | 49 | 1.15 [1.10,1.20] | 1.11 [ 1.07 , 1.15 ] |
| Small intestine | 1524 | 57 | 1144 | 43 | 1.44 [1.33,1.55] | 1.15 [ 1.08 , 1.22 ] |
| Colon | 19236 | 51 | 18274 | 49 | 1.18 [1.16,1.21] | 1.18 [ 1.16 , 1.20 ] |
| Rectal | 12363 | 60 | 8390 | 40 | 1.64 [1.59,1.68] | 1.10 [ 1.08 , 1.13 ] |
| Anal | 449 | 31 | 1021 | 70 | 0.47 [0.42,0.53] | 1.22 [ 1.12 , 1.33 ] |
| Nasal | 434 | 60 | 289 | 40 | 1.59 [1.37,1.85] | 1.11 [ 0.99 , 1.25 ] |
| Laryngeal | 1772 | 85 | 318 | 15 | 6.08 [5.39,6.85] | 1.02 [ 0.95 , 1.09 ] |
| Lung (all) | 17049 | 54 | 14686 | 46 | 1.28 [1.25,1.30] | 1.09 [ 1.07 , 1.11 ] |
| Lung SCC | 4316 | 66 | 2242 | 34 | 2.19 [2.08,2.30] | 1.04 [ 1.00 , 1.09 ] |
| Lung ADC | 6345 | 48 | 7001 | 53 | 1.00 [0.96,1.03] | 1.11 [ 1.07 , 1.14 ] |
| SCLC | 2363 | 52 | 2169 | 48 | 1.20 [1.13,1.27] | 1.08 [ 1.03 , 1.14 ] |
| Lung (other) | 3983 | 54 | 3406 | 46 | 1.28 [1.22,1.34] | 1.11 [ 1.07 , 1.15 ] |
| Pleural | 1074 | 85 | 191 | 15 | 6.28 [5.38,7.32] | 1.20 [ 1.10 , 1.32 ] |
| Urothelial | 20316 | 75 | 6684 | 25 | 3.49 [3.39,3.58] | 1.10 [ 1.07 , 1.12 ] |
| Renal | 6405 | 62 | 3856 | 38 | 1.78 [1.71,1.85] | 1.23 [ 1.19 , 1.27 ] |
| Melanoma | 15392 | 51 | 14721 | 49 | 1.06 [1.03,1.08] | 1.31 [ 1.29 , 1.34 ] |
| Skin | 17232 | 59 | 11824 | 41 | 1.74 [1.70,1.78] | 1.24 [ 1.22 , 1.27 ] |
| Brain | 3229 | 62 | 1988 | 38 | 1.59 [1.50,1.68] | 1.19 [ 1.14 , 1.24 ] |
| Meningeal | 1728 | 29 | 4242 | 71 | 0.40 [0.38,0.43] | 1.03 [ 0.99 , 1.08 ] |
| Thyroid well diff | 1154 | 28 | 2956 | 72 | 0.35 [0.32,0.37] | 1.20 [ 1.14 , 1.26 ] |
| Thyroid anaplastic | 206 | 41 | 295 | 59 | 0.70 [0.59,0.84] | 1.26 [ 1.09 , 1.46 ] |
| NHL | 10353 | 58 | 7544 | 42 | 1.47 [1.43,1.52] | 1.16 [ 1.13 , 1.19 ] |
| CLL | 3198 | 63 | 1906 | 37 | 1.87 [1.77,1.98] | 1.22 [ 1.17 , 1.28 ] |
| Hodgkin | 1584 | 64 | 885 | 36 | 1.35 [1.24,1.47] | 1.21 [ 1.14 , 1.29 ] |
| Myeloma | 3719 | 57 | 2759 | 43 | 1.50 [1.43,1.58] | 1.12 [ 1.07 , 1.16 ] |
| ALL | 459 | 62 | 284 | 38 | 1.41 [1.21,1.64] | 1.04 [ 0.92 , 1.16 ] |
| AML | 1759 | 54 | 1518 | 46 | 1.23 [1.15,1.32] | 1.27 [ 1.20 , 1.34 ] |
| CML | 726 | 59 | 502 | 41 | 1.38 [1.23,1.54] | 1.11 [ 1.01 , 1.21 ] |

The regression coefficient method applied to estimate the proportion of the excess cancer risk in men explained by height, relies on multiple assumptions, see section *3.2.4 Causal inference on time-to-event outcomes*, and among these the assumption of proportional hazards. We applied flexible parametric models allowing for age-varying effects of sex to test this assumption as well as to illustrate the cancer-free survival proportion in men (S1M1) and women (S0M0) over adult age for cancer sites associated with male sex and tall body stature.

In addition to the actual cancer-free survival proportions, we plotted the counterfactual survival in men at women's height (S1M0) and the counterfactual survival in women at men's height (S0M1) and estimated the proportion of the excess cancer risk in men explained by height (PE) at age 90, see *3.2.2 Flexible parametric models* for a description of the methodology.

*Figure 15. Cancer-free survival proportion in women, S0M0; men, S1M1; men at women's height, S1M0; women at men's height, S0M1; and the proportion of the excess cancer risk in men explained by height at age 90, PE.*

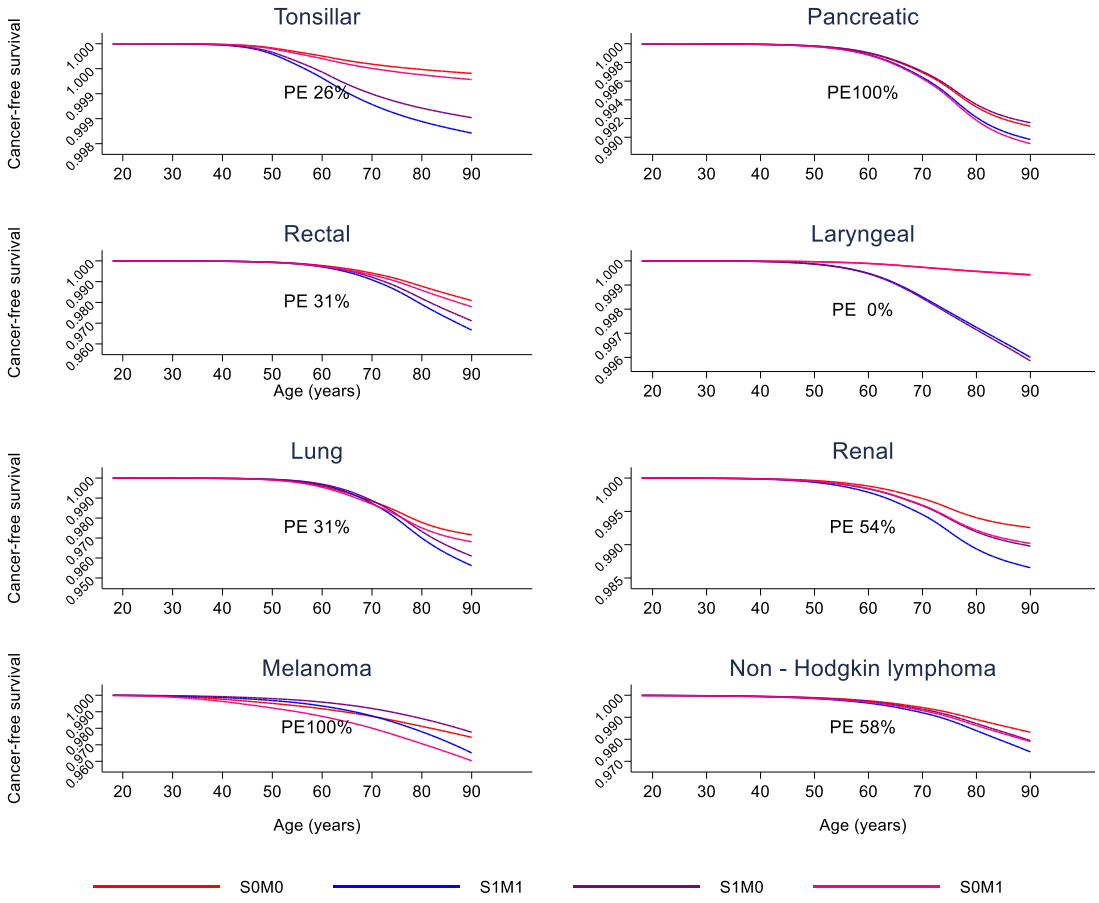



Figure 15 exemplifies 8 male-predominated malignancies, covering different anatomical tracts and illustrating various effects of the height-standardization. The proportions explained by height, estimated using flexible parametric models allowing for time-varying effects of sex, were identical or at least very close to those estimated using the, unadjusted, regression coefficient method (not presented). Tonsillar, pancreatic, rectal, renal cancer, and non-Hodgkin lymphoma were all more common in men across age, and a substantial proportion of the excess male cancer risk at age 90 (26-100%) was explained by height. The effect of the height standardization was evaluated by assessing a potential shift of the counterfactual male cancer-free survival function (S1M0) compared to the actual (S1M1).

Laryngeal cancer was found to be 6 times more common in men but the relative risk did not increase significantly with height (Table 11). Consequently, height standardization did not alter the cancer-free survival function in neither men nor women, and none of the excess male

laryngeal cancer risk was thus explained by body height (Figure 15). In pancreatic cancer, height-standardization caused the counterfactual survival in men at women's height (S1M0) to practically overlay the actual cancer-free survival in women (S0M0). Body height thus accounted for 100% of the excess pancreatic cancer risk in men. Men are at a higher risk for malignant melanoma at age 90, but melanoma is more common in young women than men. This results in crossing actual cancer-free survival functions in men (S1M1) and women (S0M0) suggesting a potential problem when interpreting the output from Cox proportional hazards regression models.

# 5 DISCUSSION

A balanced report of strengths, limitations and weaknesses, together with comparisons of ours with previously published results, can be found in the attached publications and manuscripts, studies I-IV.

## 5.1 SEX AND CANCER RISK

In studies I and IV, both large population-based cohort studies based on Swedish Cancer Register data, we outlined the persistent excess cancer risk in men and investigated whether this can be explained be attained height. We found that male sex is a consistent risk factor for 34 of 39 studied malignancies affecting both men and women. The excess risk in men decreases over time in cancers with established, strong, environmental risk factors but remains elevated for a majority of sites. Most cancers display a similar age pattern where the excess cancer risk in men culminates when cancer incidence peaks, around the age of 70. The population attributable risk percent estimates confirmed that biological and/or environmental factors related to sex account for a substantial fraction of all incident cancer cases. Moreover, we were able to demonstrate that a substantial proportion of the excess cancer risk in men is mediated by body height. The height effect was most consistent and pronounced in cancer sites with few or no known environmental risk factors and less so in sites with an established association to smoking and/or alcohol.

No one can argue against the fact that a large proportion of the observed sex differences in cancer incidence is due to a historically disproportionate exposure to mainly tobacco smoking (respiratory tract and esophageal squamous cell), but also alcohol (oral cavity, pharynx, larynx, and liver) and/or occupational carcinogens (pleura, urothelial). The proportion of daily smokers is however larger in Swedish women compared to men since the late 1980s. By that, excess smoking in men cannot fully account for the prevailing male predominance in these sites, even when considering birth cohort effects and latency periods. Moreover, environmental factors offer no explanation to the consistent excess male cancer risk in sites with weak or largely no known risk factors (e.g., small intestine, CNS, and hematological malignancies) nor in sites where environmental carcinogens contribute moderately. The consistency across anatomical tract, calendar time, and geographical region, rather indicates underlying innate, biological processes spanning over immune function, hormonal regulation, gene expression, response to oxidative damage, metabolic, and/or anatomic mechanisms (14, 32, 120). The increasing male-to-female cumulative incidence over age, does not necessarily reflect different biology in different age groups, but could also imply increasing life-time cumulative effects. Exceptions do however always exist, and the deviant age pattern observed in tumors originating from the salivary glands, stomach, colon, and malignant melanoma, is highly interesting and could possibly reflect the effect of female sex hormones or other factors related to the female reproductive age (120).

Recognition of the intimate interplay between immunological processes and carcinogenesis is rapidly advancing. This research area focus on host protector properties and/or interaction

between the immune system and tumor cells (33, 34). Immunological processes act differently in men and women resulting in sex differences in risk of and outcomes from not only autoimmune conditions (more common in women), infections, (higher vulnerability in men), but conceivably also malignant diseases (26, 28-31). Immunological mechanisms, including a reduced innate and adaptive immune function in men, are likely to at least partly account for sex differences in cancer risk as well as survival (29, 33, 34, 206). Among multiple possible mechanisms women seem to have an increased number of activated T cells and a higher cytotoxic T cell activity in peripheral blood while men express higher levels of regulatory T cells (28-30, 206). Immunologic mechanisms may indeed act differently in different tumor types. Promotor elements of several immune-related genes have androgen and estrogen response elements, perhaps offering an explanation to some tumors being more common in young women (28-31, 33, 34, 206).

Another not well-investigated hypothesis is the relationship between sex, cancer risk, and height (161, 163). Somatic cell division and DNA replication are continuously ongoing processes to maintain tissue homeostasis. Somatic driver mutations in cancer occur spontaneously or are induced by carcinogens in the environment. It is well-established that cancer risk in different organs diverges radically, but it was not until recently suggested that this is driven by stem cell number and turnover rate within tissue, predicting the accumulation of mutations in proto-oncogenes (32, 153, 158, 161). It has been estimated that only approximately one third of the variation in cancer risk between different organs is explained by environmental and/or hereditary risk factors (32, 159). The remaining variation is an effect of randomly acquired deleterious mutations (32, 159). The positive association between tall stature and cancer rate in both sexes supports this concept (143, 151, 152, 161), while the absent relationship between body size and cancer risk comparing mammal species (Peto's paradox) is contradictory (207). The later phenomenon probably has basic, evolutionary explanations and it has been shown that larger, long-lived animals have evolved mechanisms to suppress carcinogenesis. The African elephant, for example, have 20 copies of the tumor suppressor gene TP53, while the human genome typically harbors one (208). Within species, the idea of increasing cancer risk with stem cell numbers to target (body size) and lifespan seems to hold. Larger body size in men is probably a trade-off between increased access to resources, mating, and predator avoidance and an increased cancer risk with age (209). In mammals, body size is dependent on cell number, i.e., larger individuals consists of more and not larger cells (210). An exception is tissue consisting of slowly dividing cells, like nerve, muscle, and white fat, which seems to increase in volume. Malignant tumors originating from these types of tissue are however very rare. (210), Body size, in this context, is not to be mixed up with overweight and obesity. The latter implies excessive body fat and other causative pathways to increased cancer risk.

Previous studies on height and cancer have been small and presented discrepant results due to limited number of incident cancer cases and consequently low statistical power (143, 151, 152). To our knowledge, only one has studied to what extent body height can explain the association between male sex and cancer risk (163). Study IV is, to our knowledge, by far the

largest to investigate the relationship between sex, cancer risk, and height (163, 205). The population-based approach and individual-level record linkage of multiple nationwide registers, yielded high-quality data on incident cancers as well as an unbiased long-term follow-up. None of the previously performed studies on height and cancer had enough data and/or the same level of granularity to categorize malignancies into equally many subtypes. Some were based solely on one sex and/or cancer mortality rates in the population (143, 145, 147, 151-153). Study IV is moreover the first to use state-of-the-art mediation analysis in a time-to-event setting for this specific research question (163, 181, 205).

The notion of cumulative cell divisions and cancer risk is consistent with the multistage model of carcinogenesis and the general cancer age-incidence pattern (153, 157, 159-161). This concept is however not uncontroversial. Other research groups have estimated that environmental factors account for more than 90% of all incident cancer cases, even in sites without established associations to environmental carcinogens, like osteosarcoma (> 81%) and glioblastoma (> 95% ) (160, 211). The correlation between height and stem cell number probably varies between organs and cell types (161). Together with various environmental risk factors this offers an explanation to the discrepant effect of height in malignancies with different anatomical and histological origin. Lung cancer is an interesting example, basically all subtypes are strongly associated with smoking, but somewhat to lesser extent in lung adenocarcinoma which also happens to be the subtype with the strongest association to height. Small intestine cancer is very rare (only 2,668 cases in our cohort) while colon cancer is one of the most common malignancies (37,510 cases), the RR per 10 cm height increase is however very similar, 15% and 18%, respectively. The disparate cancer incidence in small intestinal and colon mucosa, despite similar exposure to carcinogens from dietary intake, is believed to be due to a higher stem cell turnover rate, in the latter (32). Alternative explanations to the height-cancer association, than through cell number, have been proposed. Genetic determinants of height may also be directly associated with cancer. Environmental exposure, like adult caloric intake due to a higher energy expenditure and/or basal metabolic rate in taller individuals may accelerate cancer risk (151). Malnutrition during growth results in stunting and conceivably lower levels of insulin-like growth factors, which are suspected to speed carcinogenesis (152, 158). There is seldom one single answer to complex questions and the drivers behind the observed relationship between height and cancer can obviously vary and interact differently over time, sex, and by cancer site.

The strong biological effect of attained educational level is appealing. It was a pre-specified hypothesis to use educational level as a proxy for socioeconomic confounders associated with cancer risk and height. We did however not expect that low education is such a strong risk factor for basically every cancer site (results not presented). This association is most probably driven by various exposure to environmental risk factors due to behavioral differences in educational groups, but the topic deserves to be studied further (57). The strength and consistency of association between educational level and height was also unexpected (155, 156, 212). Those with the highest attained educational level was found to be approximately 5 cm taller than those with the lowest education, in men and women as well as over birth

cohort. Adult stature is mainly determined by two factors; genes and malnutrition during growth in uterus as well as in childhood and adolescence (155, 156, 212). Even slightly preterm children are believed to develop a shorter adult stature and but also to have a lower probability of attaining higher education, offering a feasible hypothesis to this conundrum (213). Maternal risk factors for preterm labor include concurrent disease, low education and income, emotional distress, and substance abuse, all possible contributors to a lower education in their offspring (214).

## 5.2 SEX AND CANCER SURVIVAL

In study I we aimed to delineate sex differences in cancer survival for all non-sex-specific cancers and in study II and III we searched for possible explanations to the observed survival inequality in non-small cell lung cancer (NSCLC) and urothelial bladder cancer (UBC), respectively. We found that male sex is a persistent negative prognostic factor in malignant disease (study I). But, with the exception of lung cancer, malignant melanoma, non-Hodgkin and Hodgkin lymphoma and CLL, sex differences in survival have diminished and/or evened out over calendar time. Biliary and urothelial cancer represent intriguing exceptions and are the only two sites were women do consistently worse. When investigating NSCLC closer (study II), men were found to have a consistently poorer prognosis across cell type and stage group that remained unaltered after careful adjustments for a range of clinicopathological and socioeconomic prognostic factors. In UBC (study III), the female survival disadvantage was limited to muscle invasive tumors and only evident within the first two years from diagnosis. The increased bladder cancer mortality in women with muscle invasive tumors was robust following adjustments for multiple prognostic factors.

Disentangling potential, underlying causes to the observed sex differences in cancer survival is indeed challenging and multiple interacting factors must be considered. The inconsistent pattern across calendar time suggests environmental and/or behavioral factors that have equalized over the years (215). In both NSCLC and UBC we found that men had a higher comorbidity burden. This is, at least partly, likely to be driven by a higher tobacco and alcohol consumption in men. We did find a smaller proportion of never-smokers in men diagnosed with NSCLC, supporting this notion. The rationale behind the higher smoking prevalence in male patients while the proportion of smokers in the Swedish population is higher in women than men, is not fully understood. But this could be a birth cohort phenomenon indicative of latency and/or reflecting different tumor biology in men and women. Smoking in itself affects lung capacity, wound healing, and reduces the chance of undergoing and surviving, not only pulmonary, but all major surgery, due to risks associated with anesthesia but also of postoperative morbidity and mortality. Comorbidity, discussed in detail in previous sections, affects the ability to undergo and tolerate intensive cancer therapy, an important predictor of long-term survival. We found no sex differences in the clinical management, including treatment, of NSCLC and UBC, after adjusting for factors like age, comorbidity, and (in NSCLC) performance status and smoking history. Comorbidity is however complex to measure (see *3.2.3 Measuring comorbidity*) and self-reported smoking

habits at one time point is an unreliable measure and does not capture, for example, pack-years. Even if men and women have the same likelihood of undergoing cancer surgery, different smoking history and comorbidity burden may alter their chance of recovering from the procedure.

In NSCLC we found an adverse stage distribution in men, while the opposite was noted in UBC. The former could reflect a common belief in the medical community corroborated by results from qualitative studies, that men, in general, endure a higher threshold to seek medical attention (67-69). Despite higher age and comorbidity burden in men, men and women diagnosed with NSCLC where found to have an equivalent health care utilization, measured as number of outpatient visits, supporting an increased health awareness in women. The excess mortality in women diagnosed with UBC is traditionally believed to be a result of the first early symptoms of UBC, i.e., visible hematuria, being dismissed by female patients as well as their general practitioners. We did not find any evidence of an unequal clinical management including waiting times of men and women diagnosed with NSCLC nor UBC after referral to specialist care. This finding does however not exclude previous patient's and/or doctor's delay, in the primary care setting. According to the few studies on gender inequality in health care that exist, men seem to be more likely to undergo curative cancer surgery, receive up-to-date pharmaceuticals in cardiovascular disease, and adequate management of myocardial infarctions (70-72, 216). Stage is a very strong, independent prognostic factor. Despite the careful handling and adjustments for sex differences in stage, and (in NSCLC) also primary tumor location, we most probably have an issue with residual confounding, i.e., more advanced stage within reported stage category.

Sex differences in tumor biology causing more aggressive tumor behavior in men with NSCLC and women with UBC, have been suggested. Estrogen, the primary female sex hormone, has been suggested to improve cancer prognosis in women in reproductive age through inhibitory effects on distant metastasizing (46). And testosterone, the primary male sex hormone, has been hypothesized to drive cancer aggressiveness (64). We had limited data to explore biomarkers of tumor behavior in our material. Pulmonary adenocarcinomas in women more often harbor certain genetic alterations (activating EGFR mutations and ALK translocations) predicting response to tyrosine-kinase inhibitors but also independently, indicative of a more favorable prognosis. The very same genetic alterations have been shown to be more common in Asian and non-smoking lung adenocarcinoma patients. EGFR mutational status was recorded in patients diagnosed in 2010-2016, whereof approximately 55-60% were tested. We found that activating EGFR mutations were more common in women, compared to men. In a subgroup analysis of lung adenocarcinoma patients tested for EGFR, we additionally adjusted for EGFR mutational status and found that the male excess mortality diminished slightly. The interpretation of this finding is hampered by small numbers and short follow-up in this subgroup. The observed sex discrepancies in age distribution, smoking history, EGFR status, and performed diagnostics (thoracentesis and PET scans) indicative of variant disease spread, together with the robust superior female

survival across stage group, following adjustments, support the notion of sex differences in lung adenocarcinoma tumor biology.

Non-muscle invasive UBC in women were slightly more often WHO low-grade, indicating lower risk of recurrence and/or progression. In addition to bladder cancer-specific survival proportions, we plotted recurrence- and progression-free survival over follow-up in men and women with non-muscle invasive UBC and found no evidence of faster advancing bladder tumors in women. Female UBC cases were discussed at multidisciplinary conferences and women with muscle invasive tumors underwent radical cystectomy at high-volume hospitals slightly more frequently, even after adjustments for stage. This indicates awareness of the poorer prognosis and/or expectations of more complications in women, amongst clinicians working with UBC. Exploring the interaction between sex and selected covariates in muscle invasive UBC led to a remarkable finding: women with T4 tumors had close to the same mortality as patients with primarily metastasized UBC. We used information on T stage from histopathological examination of diagnostic TURBT specimens and clinical N stage from radiological examinations, both can diverge largely from pathological staging (217). T4b bladder tumors (extension to pelvic wall and/or adjacent organs other than prostate, vagina, or uterus) are commonly considered unresectable. We do not know if T4b tumors are more common in women. Comparing clinical and, postoperative, pathological T, but also N, stage in men and women could help us understand the drivers behind the excess UBC mortality in women with T4 tumors. Studying urinary bladder physiology in men and women in relation to tumor invasion, T stage, and conceivably also surgical complications, could also enlighten this finding. In conclusion, we found no evidence of a more aggressive UBC tumor behavior in women, our results rather point towards sex differences regarding the only, potentially curative treatment of muscle-invasive UBC, namely radical cystectomy.

NSCLC and UBC are both malignancies where modern immunotherapy is revolutionizing cancer prognosis among responders. Checkpoint inhibitors were not approved for use in neither of these two malignancies during the studied time period. Studies on immunotherapy in NSCLC and UBC have so far mainly consisted of comparisons with chemotherapy in the palliative situation. From our results, we know that in the “standard” treatment situation (including best supportive care) women with advanced stage non-small cell lung (both cell types) have a survival advantage while men and women with advanced bladder cancer seem to do equally poorly. Meta-analyses studying potential sex differences in beneficial effect or efficacy from immunotherapy have demonstrated conflicting results (206, 218, 219). Studies demonstrating effect estimates favoring men have in general been based on study cohorts composed of less than 20% women (219). Performed meta-analysis have relied on the presented clinical trial hazard ratios and not on patient-level data. Moreover, the focus has been on comparing sex-specific hazard ratios comparing traditional treatment (chemotherapy) and immunotherapy, not taking the general female cancer survival advantage into account (218, 219). As discussed previously, cancer immunology does not only cover therapeutic advances, but also the role of the immune system in tumor progression and formation of metastasis (33, 220). This includes immunosuppressive components in the tumor

microenvironment like down-modulation of antigen presentation, recruitment of suppressor immune cells, and production of immunosuppressive factors. Many of these pathways have not only been shown to differ between the sexes but are also prognostic of cancer outcome and could provide an answer to the consistently superior female survival in most malignancies (28, 30, 33, 34, 206).

# 6 CONCLUSIONS

Sex differences in cancer risk

- Male sex is an independent risk factor for a majority of common-site cancers and the fraction of cases attributable to factors related to male sex is considerable.
- The consistency over time and age indicates underlying biological drivers.
- A fair proportion of the excess cancer risk in men is explained by sex differences in body height, strengthening the evidence behind the stochastic effects of DNA replication with increasing cell numbers on cancer risk.
- Preventive measures to reduce the excess male cancer risk in sites where a majority of the effect is mediated through body height are not worthwhile and resources are better spent elsewhere.

Sex differences in cancer survival

- Male sex is associated with poorer cancer prognosis, but the male survival disadvantage has leveled out over time for a majority of cancer sites.
- Our findings suggest fundamental, but modifiable, behavioral gender differences underlying the poorer cancer outcome in men.

Sex differences in non-small cell lung cancer survival

- The female non-small cell lung cancer survival advantage is persistent over calendar time.
- This effect is strongest and most consistent in pulmonary adenocarcinoma and robust to adjustments for multiple clinicopathological factors, including stage, at diagnosis.
- Women with lung adenocarcinoma are younger, more often never-smokers, and harbor activating EGFR mutations more frequently, compared to men.
- These findings suggest sex differences in tumor biology, contributing to the female survival advantage in lung adenocarcinoma.

Sex differences in urothelial bladder cancer survival

- The excess bladder cancer mortality in women is limited to muscle-invasive but not yet metastasized tumors and is largely driven by T4 tumors, at diagnosis.
- The female excess mortality is only noticeable within two years from diagnosis.
- This indicates sex differences in treatment, including complications, of muscle-invasive bladder cancer.

# 7 FUTURE PERSPECTIVES

In the era of personalized medicine it is revolting that such a simple variable as patient or cell sex (in vitro studies) is not consistently reported. Evidence to differential treatment tolerability and long- and short-term response in men and women are momentous and costs for newly developed medicines in health care are growing exponentially. This thesis is mainly focusing on and investigating the effect of sex as a biological variable on cancer risk and survival. But the pathways through which the effect of sex are executed are most probably reflected by complex interactions between genetic, anatomic, hormone, immunologic, and environment factors, where the latter includes gender. Gender research in medicine has mainly focused on women's right to equal access to and utilization of health care resources. Our main conclusion, from a cancer preventive and treatment perspective, is that men do substantially worse compared to women.

Studying the exceptions to the general pattern is attractive and can occasionally generate new hypotheses. The excess risk of anal cancer in women is believed to be due to HPV infection, and the effect of the national HPV vaccination program on anal cancer incidence in Sweden will be disclosed in the near future. The remaining female-dominated cancer sites; biliary tract, meninges, and thyroid, deserves to be studied further. Meningioma and thyroid cancer are of benign character and over-diagnosis of subclinical disease that will never cause symptoms due to health seeking behavior and/or investigations of other conditions, like autoimmune thyroid disease in women, are possible explanations. Biliary cancer is rare but aggressive, and survival has remained very poor. Except for in young ages when men dominate, biliary cancer is more common in women. The most well-known risk factor in young, western populations is primary sclerosing cholangitis associated with inflammatory bowel disease. This relationship and the possible increasing trend in young adults in Sweden (results not presented) deserves to be studied further as well as the underlying drivers behind the poorer survival in women. An aberrant incidence pattern over age was observed for tumors originating from the salivary glands, stomach, colon, lung cancer (all major subtypes except squamous cell), and malignant melanoma of the skin. These sites were all found to be more common in women before menopause. The underlying reasons for this phenomenon deserve to be studied further. Pregnancy is a state of temporary immunosuppression in women potentially triggering carcinogenesis. Exploring incidence patterns in relation to pregnancy in cancers that are more common in women compared to men during female reproductive age could provide a clue to this inconsistency.

The association between height and cancer risk is highly interesting and deserves to be studied further. The continuous increase in body height in Swedish men and women during the 20$^{th}$ century are most probably due to improved living conditions. This has resulted in a decreased prevalence of diseases and conditions during pregnancy, childhood and adolescence, affecting attained adult body height through malnutrition during periods of growth. This is supported by the notion that the height increase seems to have flattened out in individuals born in the 1980s and 1990s. What predicts an individual's attained height in

Sweden today is consequently mainly the target height of your parents, i.e., genetic, and not environmental factors. One way to further test the hypothesis of the stochastic effects of DNA replication with increasing cell numbers on cancer risk would be to study the association between cancer and known genetic determinants of height using existent genome-wide association study datasets.

It is estimated that tobacco smoking accounts for half of all incident cases of UBC. Despite this, UBC incidence continues to increase in both sexes and has remained consistently 3-4 times higher in men. UBC is a malignancy of the elderly, indicating cumulative carcinogen exposure, i.e., a long latency period. Urinary bladder function is commonly impaired in men through benign prostatic hyperplasia causing urinary retention and, in the long run, bladder muscle wall thickening. Women more often suffer from overactive bladder, aggravated by dry mucous membranes and repeated lower urinary tract infections after menopause and/or pelvic floor impairment after pregnancy and childbirth. Benign prostate hyperplasia, through prolonged exposure to urine carcinogens, could potentially account for the excess UBC risk and, through bladder wall thickening, contribute to the superior survival, in men. Benign prostate hyperplasia does seem to increase UBC risk, but it is not known to what extent this accounts for the excess risk in men compared to women (221). Tumor invasion into the bladder wall ought to take longer in case of bladder wall thickening, resulting in less advanced stage at diagnosis in men. Delicate bladder walls in women might increase complication rates after radical cystectomy. Studying sex differences in muscle-invasive tumors closer and with a special regard to preoperative treatment and radical cystectomy details, such as duration, type of urinary diversion, lymph node dissection, clinical and pathological staging, and short- and long-term surgical complications, could improve our understanding of the female survival disadvantage.

Personalized cancer medicine involves genetic testing of cancer and normal cells to achieve more effective, customized strategies for cancer prevention, screening, and treatment. Genetic and/or hormonal differences in men and women have been shown to affect elimination rate of anticancer drugs (52, 53). The clearance of chemotherapy but also monoclonal antibodies seems to be more effective in men compared to women (52, 53). In the routine clinical setting, these drugs are still identically dosed and administered, using body surface area rather than drug elimination rate, in men and women. This leads to unpredictable effect variation where overdosing probably accounts for the repeated reports of higher chemotherapy toxicity rates in women, while unrecognized underdosing may underlie poorer response rates and cancer survival in men (53). Clinical cancer studies should collect data on pharmacokinetics, like expression of metabolic enzymes and transporters in liver and kidney, which probably differs between anticancer substances and protocols, to form the basis of future dosing algorithms. Retrospective subgroup analyses on already performed clinical studies containing information on side effects, dose reductions, and short- and long-term cancer patient outcome, could readily support clinical decision making and prompt dose-escalations or dose-reductions in patients experiencing unexpectedly low or high toxicity, respectively, regardless of sex.

The high cost of and inconsistent response to immunotherapeutic drugs have urged a search for factors that predict treatment response. Various biomarkers have been proposed, including protein death-ligand 1, tumor mutational burden, microsatellite instability, and tumor-infiltrating lymphocytes (216). Many of these tests have also been shown to be prognostic, regardless of therapy (216). To redo analyses stratified by sex on already existing datasets, could delineate if these alterations are more or less common in women and whether the strength of association to prognosis and/or response to immunotherapy differs between the sexes. In addition to a potential improvement of personalized medicine this could add valuable information on the underlying biological drivers behind sex differences in cancer risk as well as prognosis (191, 213, 214).

Studies on amenable factors associated with disease risk and prognosis are simple and straightforward to motivate. But biological drives, like attained height and immune competence, can be equally interesting. Results from studies on sex differences in cancer risk and survival can enable us to guide limited health care resources to where they are most needed and have the greatest chance of postponing and/or reducing the number of preventable cancer deaths. Instead of focusing on men and women receiving the exact same management, we ought to aim towards treating men and women to achieve the most effective prevention of cancer death in both sexes. To accomplish this, we need to improve our knowledge of the determinants of sex on cancer risk and treatment outcome. Therefor patient sex should always be considered in studies on cancer risk and survival.

# 8 SWEDISH SUMMARY

Det är ett väletablerat faktum att män, jämfört kvinnor, har större risk att insjukna i de flesta cancerformer som drabbar bägge könen. Att män dessutom även har sämre prognos efter insjuknande i cancer är inte lika känt. Orsakerna till det först nämnda har traditionellt ansetts vara en högre exponering för cancerframkallande faktorer såsom tobaksrökning, alkohol, UV-strålning och kemikalier inom industri och verkstadsarbete. Könsskillnader i canceröverlevnad är styvmoderligt studerat, men en allmän uppfattning har varit att män söker vård i ett mer framskridet stadium där sjukdomen inte lika ofta går att bota.

I studie I ville vi kartlägga könsskillnader i cancerrisk och -överlevnad över tid och ålder samt kvantifiera hur stor andel av alla cancerfall som kan förklaras av den manliga canceröverrisken. Med svenska cancerregisterdata identifierade vi 39 icke-könsspecifika tumörformer, diagnosticerade år 1970-2014, vid 15-84 års ålder. Med undantag för tumörer med en mycket stark koppling till rökning och/eller alkohol (luftvägar, huvud-hals, matstrupe och lever), kunde vi visa att den manliga överrisken är stabil över kalendertid. När vi studerade insjuknande över ålder var det tydligt att överrisken bland män var som störst i den åldersgrupp där cancerrisken i befolkningen kulminerar (65-75 år). Trender i överlevnad var inte lika slående och för flera tumörformer kunde vi konstatera att män och kvinnor numera har samma förväntade överlevnad. Sammantaget talar våra fynd för att något mer än miljöfaktorer ligger bakom den generellt ökade cancerrisken bland män.

Canceröverlevnad är multifaktoriellt och det finns troligtvis inte en förklaringsmodell för alla tumörformer. Sannolikt spelar skillnader i beteende mellan män och kvinnor en viktig roll. För att på detaljnivå studera könsskillnader i överlevnad valde vi i studie II och III att fokusera på två vanliga tumörer; icke-småcellig lungcancer respektive cancer i urinblåsa. Svenska kvalitetsregisterdata länkade till flera andra populationsbaserade register gav oss detaljerad information om tumörtyp/-stadium, cancerbehandling, väntetider, samsjuklighet, socioekonomi, dödsorsak, med mera.

Det är välkänt att kvinnor med icke-småcellig lungcancer har bättre prognos än män, orsakerna till detta är inte klarlagda. Icke-småcellig lungcancer kan delas upp i två olika celltyper; skivepitelcancer och adenocarcinom, där epidemiologi, behandling och prognos skiljer sig radikalt. Vi valde därför att analysera dessa två celltyper separat. Liksom tidigare studier på området, kunde vi konstatera att kvinnliga patienter har en generellt bättre överlevnad i alla stadier av sjukdomen, men detta var mest uppenbart bland de med lung-adenocarcinom. Kvinnor i denna grupp var yngre, hade mindre samsjuklighet, högre utbildningsnivå och var oftare icke-rökare, jämfört med män. Vi hittade inga eller mycket små skillnader i klinisk handläggning; genomförda undersökningar, väntetider och behandling, mellan män och kvinnor. Våra resultat talar för skillnader i tumörbiologi mellan män och kvinnor med adenocarcinom i lunga.

Urinblåstumörer är tre till fyra gånger vanligare bland män. Kvinnor har dock, till skillnad från de flesta andra tumörformer, sämre prognos. Det förstnämnda tros vara relaterat till

rökning och yrkesexponering för carcinogener. Orsakerna bakom den ojämlika blåscancerdödligheten är inte klarlagda, men tros bero på att kvinnor med synligt blod i urinen inte handläggs lika skyndsamt som män vilket leder till mer avancerade tumörstadium och sämre prognos. Ytligt växande tumörer i urinblåsa har en mycket god prognos och de flesta är botade efter lokalbehandling. Muskelinvasiv och metastaserad (spridd till andra organ) blåscancer har däremot ett aggressivt förlopp med förväntad 5-årsöverlevnad på 50 respektive 5 %, vid optimal behandling. Vi valde därför att gruppera överlevnadsanalyserna i dessa tre stadier. Utöver att andelen kvinnor ökade med tumörens utbredningsgrad kunde vi inte hitta några tecken till ojämlik klinisk handläggning. Överdödligheten i blåscancer bland kvinnor var begränsad till de första två åren efter diagnos och sågs endast i subgruppen med muskelinvasiv, men ännu icke spridd, blåscancer. Skillnaderna kvarstod efter justering för en rad prognostiska faktorer. Sammantaget talar detta för skillnader i klinisk handläggning och behandling (radikal kirurgi) av män och kvinnor med muskelinvasiva tumörer. Den ogynnsamma stadiefördelningen och den uttalade mortaliteten bland kvinnor med potentiellt botbar muskelinvasiv blåscancer fordrar en ökad kännedom i befolkning och primärvård om att synligt blod i urinen är ett alarmsymptom på blåscancer som kräver omedelbar utredning på specialiserad enhet, även hos kvinnor.

Flera studier har visat att längre individer har en förhöjd cancerrisk under livet. En stor andel av alla cancerfall beror antagligen inte på miljö- eller ärftliga faktorer, utan är en följd av slumpmässiga fel i arvsmassan (mutationer) som sker i samband med celldelning. Längre individer har fler celler och därmed fler celldelningar vilket skulle kunna förklara detta fenomen. I studie IV ville vi undersöka om canceröverrisken hos män kan förklaras av att män generellt är längre än kvinnor. Vi använde oss av data på vuxen kroppslängd hos drygt sex miljoner svenska män och kvinnor extraherade från pass-, mönstrings- och mödravårdsregistret, länkat till cancerregistret. Med nyutvecklad metodologi, tid-till-utfall mediationsanalys, kunde vi påvisa att kroppslängd spelar en avgörande roll för den förhöjda cancerrisken hos män för ett flertal av de undersökta cancertyperna.

# 9 ACKNOWLEDGEMENTS

*Gustaf Edgren*, main supervisor. Thank you for giving me “frihet under eget ansvar” to explore my own ideas, taking my time, and always protecting my scientific integrity.

*Paul Dickman*, co-supervisor. Your door has always been open for me, you have spoiled me with endless patience, teaching skills, and good humor. It has been luxurious to be a PhD student under your wings!

*Mats Lambe*, co-supervisor. Thank you for mental support and strategic input.

*Anna Johansson*, co-supervisor. For keeping me organized and teaching me the importance of rehearsing every presentation as if it was the premiere at the opera!

A big thanks to all of my hardworking, *colleagues and fellow doctoral students at MEB and KEP*. Many of you are friends for life!

*Caroline Weibull, Hannah Bower, Elisabeth Dahlqwist*, for invaluable input on statistical methods, you are all “horse whisperers” (i.e., statisticians that can talk statistics to clinicians)!

*Arvid Sjölander*, co-author. For sharing your wisdom and common sense.

*Christel Häggström*, co-author. For guiding me through the variable jungle of BladderBaSE, always with patience and a genuine kindness.

*Mats Lindblad*, head of upper GI surgery, Karolinska Universitetssjukhuset. For believing and implementing that everything is possible!

*Mariann Iristo,* former head of the Oncology Department, Södersjukhuset. Tough but fair and always making ends meet, I miss you.

*Svetlana Bajalica Lagercrantz,* former chief of residents at the Oncology Department, Karolinska Universitetssjukhuset. For calling me a fighter in the exact right moment!

*Maria Helde Frankling,* former chief of residents at the Oncology Department, Karolinska Universitetssjukhuset. Thank you for starting me up!

Representatives and data holders at the Swedish *Cancer Quality Registers, LCBaSe*, and *BladderBaSe*. Without maintaining high-quality health data registers, this research would not have been accomplished.

*Clinival colleagues*; doctors, nurses, administrators, and friends at every Oncology Department in Stockholm. “I will remember that there is art to medicine as well as science, and that warmth, sympathy, and understanding may outweigh the surgeon's knife or the chemist's drug.”

*Friends and family*. I know that you will always have my back.

*My mother.* Probably the only one who have read the whole kappa, twice…

# 10 REFERENCES


1. Legato MJ, Johnson PA, Manson JE. Consideration of Sex Differences in Medicine to Improve Health Care and Patient Outcomes. Jama. 2016;316(18):1865-6.

2. Clayton J, Tannenbaum C. Reporting sex, gender, or both in clinical research? Jama. 2016;316(18):1863-4.

3. Ashley DJ. A male-female differential in tumour incidence. British journal of cancer. 1969;23(1):21-5.

4. Ashley DJ. Sex differences in the incidence of tumours at various sites. British journal of cancer. 1969;23(1):26-30.

5. Cook MB, Dawsey SM, Freedman ND, Inskip PD, Wichner SM, Quraishi SM, et al. Sex disparities in cancer incidence by period and age. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology. 2009;18(4):1174-82.

6. Edgren G, Liang L, Adami HO, Chang ET. Enigmatic sex disparities in cancer incidence. European journal of epidemiology. 2012;27(3):187-96.

7. Biggar RJ, Bergen AW, Poulsen GN. Impact of x chromosome genes in explaining the excess risk of cancer in males. American journal of epidemiology. 2009;170(1):65-71.

8. Innos K, Padrik P, Valvere V, Aareleid T. Sex differences in cancer survival in Estonia: a population-based study. BMC cancer. 2015;15(1):72.

9. Cook MB, McGlynn KA, Devesa SS, Freedman ND, Anderson WF. Sex disparities in cancer mortality and survival. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology. 2011;20(8):1629-37.

10. Jung KW, Park S, Shin A, Oh CM, Kong HJ, Jun JK, et al. Do female cancer patients display better survival rates compared with males? Analysis of the Korean National Registry data, 2005-2009. PloS one. 2012;7(12):e52457.

11. Micheli A, Ciampichini R, Oberaigner W, Ciccolallo L, de Vries E, Izarzugaza I, et al. The advantage of women in cancer survival: an analysis of EUROCARE-4 data. European journal of cancer (Oxford, England : 1990). 2009;45(6):1017-27.

12. Lortet-Tieulent J, Renteria E, Sharp L, Weiderpass E, Comber H, Baas P, et al. Convergence of decreasing male and increasing female incidence rates in major tobacco-related cancers in Europe in 1988-2010. European journal of cancer (Oxford, England : 1990). 2015;51(9):1144-63.

13. McCartney G, Mahmood L, Leyland AH, Batty GD, Hunt K. Contribution of smoking-related and alcohol-related deaths to the gender gap in mortality: evidence from 30 European countries. Tobacco control. 2011;20(2):166-8.

14. Dorak MT, Karpuzoglu E. Gender differences in cancer susceptibility: an inadequately addressed issue. Frontiers in genetics. 2012;3:268.

15. Benigni R. Social sexual inequality and sex difference in cancer incidence. Environmental research. 2007;104(1):128-34.

16. Vahter M, Akesson A, Liden C, Ceccatelli S, Berglund M. Gender differences in the disposition and toxicity of metals. Environmental research. 2007;104(1):85-95.

17. Grosse Y, Baan R, Straif K, Secretan B, El Ghissassi F, Bouvard V, et al. A review of human carcinogens--Part A: pharmaceuticals. The Lancet Oncology. 2009;10(1):13-4.

18. Bouvard V, Baan R, Straif K, Grosse Y, Secretan B, El Ghissassi F, et al. A review of human carcinogens--Part B: biological agents. The Lancet Oncology. 2009;10(4):321-2.

19. Straif K, Benbrahim-Tallaa L, Baan R, Grosse Y, Secretan B, El Ghissassi F, et al. A review of human carcinogens--Part C: metals, arsenic, dusts, and fibres. The Lancet Oncology. 2009;10(5):453-4.

20. El Ghissassi F, Baan R, Straif K, Grosse Y, Secretan B, Bouvard V, et al. A review of human carcinogens--part D: radiation. The Lancet Oncology. 2009;10(8):751-2.

21. Secretan B, Straif K, Baan R, Grosse Y, El Ghissassi F, Bouvard V, et al. A review of human carcinogens--Part E: tobacco, areca nut, alcohol, coal smoke, and salted fish. The Lancet Oncology. 2009;10(11):1033-4.

22. Baan R, Grosse Y, Straif K, Secretan B, El Ghissassi F, Bouvard V, et al. A review of human carcinogens--Part F: chemical agents and related occupations. The Lancet Oncology. 2009;10(12):1143-4.

23. de Martel C, Franceschi S. Infections and cancer: established associations and new hypotheses. Critical reviews in oncology/hematology. 2009;70(3):183-94.

24. Slade T, Chapman C, Swift W, Keyes K, Tonks Z, Teesson M. Birth cohort trends in the global epidemiology of alcohol use and alcohol-related harms in men and women: systematic review and metaregression. BMJ open. 2016;6(10):e011827.

25. 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.

26. Plummer M, de Martel C, Vignat J, Ferlay J, Bray F, Franceschi S. Global burden of cancers attributable to infections in 2012: a synthetic analysis. The Lancet Global health. 2016;4(9):e609-16.

27. Sonawane K, Suk R, Chiao EY, et al. Oral human papillomavirus infection: Differences in prevalence between sexes and concordance with genital human papillomavirus infection, nhanes 2011 to 2014. Annals of Internal Medicine. 2017.

28. Klein SL, Flanagan KL. Sex differences in immune responses. Nature reviews Immunology. 2016;16(10):626-38.

29. Lin PY, Sun L, Thibodeaux SR, Ludwig SM, Vadlamudi RK, Hurez VJ, et al. B7-H1-dependent sex-related differences in tumor immunity and immunotherapy responses. Journal of immunology (Baltimore, Md : 1950). 2010;185(5):2747-53.

30. vom Steeg LG, Klein SL. SeXX Matters in Infectious Disease Pathogenesis. PLoS pathogens. 2016;12(2):e1005374.

31. Klein SL, Jedlicka A, Pekosz A. The Xs and Y of immune responses to viral vaccines. The Lancet Infectious diseases. 2010;10(5):338-49.

32. Tomasetti C, Vogelstein B. Variation in cancer risk among tissues can be explained by the number of stem cell divisions. Science. 2015;347(6217):78-81.

33. Yang Y. Cancer immunotherapy: harnessing the immune system to battle cancer. The Journal of clinical investigation. 2015;125(9):3335-7.

34. Miller JF, Sadelain M. The journey from discoveries in fundamental immunology to cancer immunotherapy. Cancer cell. 2015;27(4):439-49.

35. Libert C, Dejager L, Pinheiro I. The X chromosome in immune functions: when a chromosome makes the difference. Nature reviews Immunology. 2010;10(8):594-604.

36. Whitacre CC, Reingold SC, O'Looney PA. A gender gap in autoimmunity. Science. 1999;283(5406):1277-8.

37. Dunford A, Weinstock DM, Savova V, Schumacher SE, Cleary JP, Yoda A, et al. Tumor-suppressor genes that escape from X-inactivation contribute to cancer sex bias. Nature genetics. 2017;49(1):10-6.

38. Kido T, Lau YF. Roles of the Y chromosome genes in human cancers. Asian journal of andrology. 2015;17(3):373-80.

39. Forsberg LA, Rasi C, Malmqvist N, Davies H, Pasupulati S, Pakalapati G, et al. Mosaic loss of chromosome Y in peripheral blood is associated with shorter survival and higher risk of cancer. Nature genetics. 2014;46(6):624-8.

40. Sharma S, Eghbali M. Influence of sex differences on microRNA gene regulation in disease. Biology of sex differences. 2014;5(1):3.

41. Yuan Y, Liu L, Chen H, Wang Y, Xu Y, Mao H, et al. Comprehensive Characterization of Molecular Differences in Cancer between Male and Female Patients. Cancer cell. 2016;29(5):711-22.

42. Ali I, Hogberg J, Hsieh JH, Auerbach S, Korhonen A, Stenius U, et al. Gender differences in cancer susceptibility: role of oxidative stress. Carcinogenesis. 2016;37(10):985-92.

43. Higdon JV, Frei B. Is there a gender difference in the effect of antioxidants on cancer risk? The British journal of nutrition. 2005;94(2):139-40.

44. Dickman PW, Adami HO. Interpreting trends in cancer patient survival. J Intern Med. 2006;260(2):103-17.

45. Najari BB, Rink M, Li PS, Karakiewicz PI, Scherr DS, Shabsigh R, et al. Sex disparities in cancer mortality: the risks of being a man in the United States. The Journal of urology. 2013;189(4):1470-4.

46. Adami HO, Bergstrom R, Holmberg L, Klareskog L, Persson I, Ponten J. The effect of female sex hormones on cancer survival. A register-based study in patients younger than 20 years at diagnosis. Jama. 1990;263(16):2189-93.

47. Oberaigner W, Siebert U. Do women with cancer have better survival as compared to men after adjusting for staging distribution? European journal of public health. 2011;21(3):387-91.

48. Janssen-Heijnen ML, Houterman S, Lemmens VE, Louwman MW, Maas HA, Coebergh JW. Prognostic impact of increasing age and co-morbidity in cancer patients: a population-based approach. Critical reviews in oncology/hematology. 2005;55(3):231-40.

49. Pal SK, Hurria A. Impact of age, sex, and comorbidity on cancer therapy and disease progression. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2010;28(26):4086-93.

50. de Groot V, Beckerman H, Lankhorst GJ, Bouter LM. How to measure comorbidity. a critical review of available methods. Journal of clinical epidemiology. 2003;56(3):221-9.

51. Sarfati D. Review of methods used to measure comorbidity in cancer populations: no gold standard exists. Journal of clinical epidemiology. 2012;65(9):924-33.

52. Kim H-I, Lim H, Moon A. Sex Differences in Cancer: Epidemiology, Genetics and Therapy. Biomol Ther (Seoul). 2018;26(4):335-42.

53. Schmetzer O, Florcken A. Sex differences in the drug therapy for oncologic diseases. Handbook of experimental pharmacology. 2012(214):411-42.

54. Oken MM, Creech RH, Tormey DC, Horton J, Davis TE, McFadden ET, et al. Toxicity and response criteria of the Eastern Cooperative Oncology Group. American journal of clinical oncology. 1982;5(6):649-55.

55. Kelly CM, Shahrokni A. Moving beyond Karnofsky and ECOG Performance Status Assessments with New Technologies. Journal of oncology. 2016;2016:6186543.

56. Woods LM, Rachet B, Coleman MP. Origins of socio-economic inequalities in cancer survival: a review. Annals of Oncology. 2005;17(1):5-19.

57. Singh GK, Jemal A. Socioeconomic and Racial/Ethnic Disparities in Cancer Mortality, Incidence, and Survival in the United States, 1950-2014: Over Six Decades of Changing Patterns and Widening Inequalities. Journal of environmental and public health. 2017;2017:2819372.

58. Egner JR. Ajcc cancer staging manual. Jama. 2010;304(15):1726-7.

59. Edge SB, Compton CC. The American Joint Committee on Cancer: the 7th edition of the AJCC cancer staging manual and the future of TNM. Annals of surgical oncology. 2010;17(6):1471-4.

60. Derwinger K, Gustavsson B. Variations in demography and prognosis by colon cancer location. Anticancer research. 2011;31(6):2347-50.

61. Gamba CS, Clarke CA, Keegan TH, Tao L, Swetter SM. Melanoma survival disadvantage in young, non-Hispanic white males compared with females. JAMA dermatology. 2013;149(8):912-20.

62. Joosse A, Collette S, Suciu S, Nijsten T, Lejeune F, Kleeberg UR, et al. Superior outcome of women with stage I/II cutaneous melanoma: pooled analysis of four European Organisation for Research and Treatment of Cancer phase III trials. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2012;30(18):2240-7.

63. Schulze S, Petersen I. Gender and ploidy in cancer survival. Cellular oncology (Dordrecht). 2011;34(3):199-208.

64. Orsted DD, Nordestgaard BG, Bojesen SE. Plasma testosterone in the general population, cancer prognosis and cancer risk: a prospective cohort study. Annals of oncology : official journal of the European Society for Medical Oncology. 2014;25(3):712-8.

65. Seppala TT, Bohm JP, Friman M, Lahtinen L, Vayrynen VMJ, Liipo TKE, et al. Combination of microsatellite instability and BRAF mutation status for subtyping colorectal cancer. British journal of cancer. 2015;112(12):1966-75.

66. Bonneville R, Krook MA, Kautto EA, Miya J, Wing MR, Chen H-Z, et al. Landscape of Microsatellite Instability Across 39 Cancer Types. JCO Precision Oncology. 2017(1):1-15.

67. Smith LK, Pope C, Botha JL. Patients' help-seeking experiences and delay in cancer presentation: a qualitative synthesis. The Lancet. 2005;366(9488):825-31.

68. Walter F, Webster A, Scott S, Emery J. The Andersen Model of Total Patient Delay: a systematic review of its application in cancer diagnosis. Journal of health services research & policy. 2012;17(2):110-8.

69. McCutchan GM, Wood F, Edwards A, Richards R, Brain KE. Influences of cancer symptom knowledge, beliefs and barriers on cancer symptom presentation in relation to socioeconomic deprivation: a systematic review. BMC cancer. 2015;15:1000.

70. Smolina K, Ball L, Humphries KH, Khan N, Morgan SG. Sex Disparities in Post-Acute Myocardial Infarction Pharmacologic Treatment Initiation and Adherence: Problem for Young Women. Circulation Cardiovascular quality and outcomes. 2015;8(6):586-92.

71. Noren A, Eriksson HG, Olsson LI. Selection for surgery and survival of synchronous colorectal liver metastases; a nationwide study. European journal of cancer (Oxford, England : 1990). 2016;53:105-14.

72. Khan E, Brieger D, Amerena J, Atherton JJ, Chew DP, Farshid A, et al. Differences in management and outcomes for men and women with ST-elevation myocardial infarction. The Medical journal of Australia. 2018;209(3):118-23.

73. Bergman O, Hont G, Johansson E. Cancer i siffror 2013. 6 ed. Stockholm: Socialstyrelsen, Cancerfonden; 2013.

74. Gazdar AF. Should we continue to use the term non-small-cell lung cancer? Annals of oncology : official journal of the European Society for Medical Oncology. 2010;21 Suppl 7:vii225-9.

75. Travis WD, Brambilla E, Noguchi M, Nicholson AG, Geisinger KR, Yatabe Y, et al. International association for the study of lung cancer/american thoracic society/european respiratory society international multidisciplinary classification of lung adenocarcinoma. Journal of thoracic oncology : official publication of the International Association for the Study of Lung Cancer. 2011;6(2):244-85.

76. Devesa SS, Bray F, Vizcaino AP, Parkin DM. International lung cancer trends by histologic type: male:female differences diminishing and adenocarcinoma rates rising. International journal of cancer Journal international du cancer. 2005;117(2):294-9.

77. Houston KA, Henley SJ, Li J, White MC, Richards TB. Patterns in lung cancer incidence rates and trends by histologic type in the United States, 2004-2009. Lung cancer. 2014;86(1):22-8.

78. Wakelee HA, Chang ET, Gomez SL, Keegan TH, Feskanich D, Clarke CA, et al. Lung cancer incidence in never smokers. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2007;25(5):472-8.

79. Nationellt vårdprogram Lungcancer. Uppsala, Sweden: Regionalt cancercentrum: Regional cancercentrum Uppsala Örebro; 2015.

80. Goldstraw P, Crowley J, Chansky K, Giroux DJ, Groome PA, Rami-Porta R, et al. The IASLC Lung Cancer Staging Project: proposals for the revision of the TNM stage

groupings in the forthcoming (seventh) edition of the TNM Classification of malignant tumours. Journal of thoracic oncology : official publication of the International Association for the Study of Lung Cancer. 2007;2(8):706-14.

81. Mok TSK, Loong HH. Are we ready for immune checkpoint inhibitors for advanced non-small-cell lung cancer? The Lancet. 2016;387(10027):1488-90.

82. Herzberg B, Campo MJ, Gainor JF. Immune Checkpoint Inhibitors in Non-Small Cell Lung Cancer. The oncologist. 2017;22(1):81-8.

83. Chansky K, Sculier JP, Crowley JJ, Giroux D, Van Meerbeeck J, Goldstraw P. The International Association for the Study of Lung Cancer Staging Project: prognostic factors and pathologic TNM stage in surgically managed non-small cell lung cancer. Journal of thoracic oncology : official publication of the International Association for the Study of Lung Cancer. 2009;4(7):792-801.

84. Moore R, Doherty D, Chamberlain R, Khuri F. Sex differences in survival in non-small cell lung cancer patients 1974-1998. Acta oncologica. 2004;43(1):57-64.

85. Pirker R, Pereira JR, Szczesna A, von Pawel J, Krzakowski M, Ramlau R, et al. Prognostic factors in patients with advanced non-small cell lung cancer: data from the phase III FLEX study. Lung cancer. 2012;77(2):376-82.

86. Radzikowska E, Glaz P, Roszkowski K. Lung cancer in women: age, smoking, histology, performance status, stage, initial treatment and survival. Population-based study of 20 561 cases. Annals of oncology : official journal of the European Society for Medical Oncology. 2002;13(7):1087-93.

87. Rena O, Massera F, Boldorini R, Papalia E, Turello D, Davoli F, et al. Non-small cell lung cancer in surgically treated women. Tumori. 2013;99(6):661-6.

88. Takifuji N, Fukuoka M, Negoro S, Takada M, Kusunoki Y, Matsui K, et al. [Prognostic factors affecting survival and response in patients with advanced non-small cell lung cancer treated with combination chemotherapy]. Gan to kagaku ryoho Cancer & chemotherapy. 1990;17(3 Pt 2):429-34.

89. Werner-Wasik M, Scott C, Cox JD, Sause WT, Byhardt RW, Asbell S, et al. Recursive partitioning analysis of 1999 Radiation Therapy Oncology Group (RTOG) patients with locally-advanced non-small-cell lung cancer (LA-NSCLC): identification of five groups with different survival. International journal of radiation oncology, biology, physics. 2000;48(5):1475-82.

90. Asmis TR, Ding K, Seymour L, Shepherd FA, Leighl NB, Winton TL, et al. Age and comorbidity as independent prognostic factors in the treatment of non small-cell lung cancer: a review of National Cancer Institute of Canada Clinical Trials Group trials. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2008;26(1):54-9.

91. Ichinose Y, Yano T, Asoh H, Yokoyama H, Yoshino I, Katsuda Y. Prognostic factors obtained by a pathologic examination in completely resected non-small-cell lung cancer. An analysis in each pathologic stage. The Journal of thoracic and cardiovascular surgery. 1995;110(3):601-5.

92. Kawaguchi T, Matsumura A, Fukai S, Tamura A, Saito R, Zell JA, et al. Japanese ethnicity compared with Caucasian ethnicity and never-smoking status are independent favorable prognostic factors for overall survival in non-small cell lung cancer: a collaborative epidemiologic study of the National Hospital Organization Study Group for

Lung Cancer (NHSGLC) in Japan and a Southern California Regional Cancer Registry databases. Journal of thoracic oncology : official publication of the International Association for the Study of Lung Cancer. 2010;5(7):1001-10.

93. Ou SH, Zell JA, Ziogas A, Anton-Culver H. Prognostic factors for survival of stage I nonsmall cell lung cancer patients : a population-based analysis of 19,702 stage I patients in the California Cancer Registry from 1989 to 2003. Cancer. 2007;110(7):1532-41.

94. Paesmans M, Sculier JP, Libert P, Bureau G, Dabouis G, Thiriaux J, et al. Prognostic factors for survival in advanced non-small-cell lung cancer: univariate and multivariate analyses including recursive partitioning and amalgamation algorithms in 1,052 patients. The European Lung Cancer Working Party. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 1995;13(5):1221-30.

95. Albain KS, Crowley JJ, LeBlanc M, Livingston RB. Survival determinants in extensive-stage non-small-cell lung cancer: the Southwest Oncology Group experience. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 1991;9(9):1618-26.

96. Cerfolio RJ, Bryant AS, Scott E, Sharma M, Robert F, Spencer SA, et al. Women with pathologic stage I, II, and III non-small cell lung cancer have better survival than men. Chest. 2006;130(6):1796-802.

97. Ferguson MK, Skosey C, Hoffman PC, Golomb HM. Sex-associated differences in presentation and survival in patients with lung cancer. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 1990;8(8):1402-7.

98. Fu JB, Kau TY, Severson RK, Kalemkerian GP. Lung cancer in women: analysis of the national Surveillance, Epidemiology, and End Results database. Chest. 2005;127(3):768-77.

99. Nakamura H, Ando K, Shinmyo T, Morita K, Mochizuki A, Kurimoto N, et al. Female gender is an independent prognostic factor in non-small-cell lung cancer: a meta-analysis. Annals of thoracic and cardiovascular surgery : official journal of the Association of Thoracic and Cardiovascular Surgeons of Asia. 2011;17(5):469-80.

100. Pitz MW, Musto G, Navaratnam S. Sex as an independent prognostic factor in a population-based non-small cell lung cancer cohort. Canadian respiratory journal. 2013;20(1):30-4.

101. Kawaguchi T, Takada M, Kubo A, Matsumura A, Fukai S, Tamura A, et al. Performance status and smoking status are independent favorable prognostic factors for survival in non-small cell lung cancer: a comprehensive analysis of 26,957 patients with NSCLC. Journal of thoracic oncology : official publication of the International Association for the Study of Lung Cancer. 2010;5(5):620-30.

102. Batevik R, Grong K, Segadal L, Stangeland L. The female gender has a positive effect on survival independent of background life expectancy following surgical resection of primary non-small cell lung cancer: a study of absolute and relative survival over 15 years. Lung cancer. 2005;47(2):173-81.

103. Chen KY, Chang CH, Yu CJ, Kuo SH, Yang PC. Distribution according to histologic type and outcome by gender and age group in Taiwanese patients with lung carcinoma. Cancer. 2005;103(12):2566-74.

104. Hsu LH, Chu NM, Liu CC, Tsai SY, You DL, Ko JS, et al. Sex-associated differences in non-small cell lung cancer in the new era: is gender an independent prognostic factor? Lung cancer. 2009;66(2):262-7.

105. Yoshino I, Baba H, Fukuyama S, Kameyama T, Shikada Y, Tomiyasu M, et al. A time trend of profile and surgical results in 1123 patients with non-small cell lung cancer. Surgery. 2002;131(1 Suppl):S242-8.

106. de Perrot M, Licker M, Bouchardy C, Usel M, Robert J, Spiliopoulos A. Sex differences in presentation, management, and prognosis of patients with non-small cell lung carcinoma. The Journal of thoracic and cardiovascular surgery. 2000;119(1):21-6.

107. Alexiou C, Onyeaka CV, Beggs D, Akar R, Beggs L, Salama FD, et al. Do women live longer following lung resection for carcinoma? European journal of cardio-thoracic surgery : official journal of the European Association for Cardio-thoracic Surgery. 2002;21(2):319-25.

108. Chang JW, Asamura H, Kawachi R, Watanabe S. Gender difference in survival of resected non-small cell lung cancer: histology-related phenomenon? The Journal of thoracic and cardiovascular surgery. 2009;137(4):807-12.

109. Debieuvre D, Oster JP, Riou R, Berruchon J, Levy A, Mathieu JP, et al. The new face of non-small-cell lung cancer in men: Results of two French prospective epidemiological studies conducted 10 years apart. Lung cancer. 2016;91:1-6.

110. Harichand-Herdt S, Ramalingam SS. Gender-associated differences in lung cancer: clinical characteristics and treatment outcomes in women. Seminars in oncology. 2009;36(6):572-80.

111. Sakurai H, Asamura H, Goya T, Eguchi K, Nakanishi Y, Sawabata N, et al. Survival differences by gender for resected non-small cell lung cancer: a retrospective analysis of 12,509 cases in a Japanese Lung Cancer Registry study. Journal of thoracic oncology : official publication of the International Association for the Study of Lung Cancer. 2010;5(10):1594-601.

112. Donington JS, Colson YL. Sex and gender differences in non-small cell lung cancer. Seminars in thoracic and cardiovascular surgery. 2011;23(2):137-45.

113. Barrera-Rodriguez R, Morales-Fuentes J. Lung cancer in women. Lung Cancer (Auckland, NZ). 2012;3:79-89.

114. O'Connell JP, Kris MG, Gralla RJ, Groshen S, Trust A, Fiore JJ, et al. Frequency and prognostic importance of pretreatment clinical characteristics in patients with advanced non-small-cell lung cancer treated with combination chemotherapy. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 1986;4(11):1604-14.

115. Wakelee HA, Wang W, Schiller JH, Langer CJ, Sandler AB, Belani CP, et al. Survival differences by sex for patients with advanced non-small cell lung cancer on Eastern Cooperative Oncology Group trial 1594. Journal of thoracic oncology : official publication of the International Association for the Study of Lung Cancer. 2006;1(5):441-6.

116. Wheatley-Price P, Le Maitre A, Ding K, Leighl N, Hirsh V, Seymour L, et al. The influence of sex on efficacy, adverse events, quality of life, and delivery of treatment in National Cancer Institute of Canada Clinical Trials Group non-small cell lung cancer chemotherapy trials. Journal of thoracic oncology : official publication of the International Association for the Study of Lung Cancer. 2010;5(5):640-8.

117. Yamamoto H, Sekine I, Yamada K, Nokihara H, Yamamoto N, Kunitoh H, et al. Gender differences in treatment outcomes among patients with non-small cell lung cancer given a combination of carboplatin and paclitaxel. Oncology. 2008;75(3-4):169-74.

118. Belani CP, Marts S, Schiller J, Socinski MA. Women and lung cancer: epidemiology, tumor biology, and emerging trends in clinical research. Lung cancer. 2007;55(1):15-23.

119. Holm B, Mellemgaard A, Skov T, Skov BG. Different impact of excision repair cross-complementation group 1 on survival in male and female patients with inoperable non-small-cell lung cancer treated with carboplatin and gemcitabine. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2009;27(26):4254-9.

120. Planchard D, Loriot Y, Goubar A, Commo F, Soria JC. Differential expression of biomarkers in men and women. Seminars in oncology. 2009;36(6):553-65.

121. Xiao D, Pan H, Li F, Wu K, Zhang X, He J. Analysis of ultra-deep targeted sequencing reveals mutation burden is associated with gender and clinical outcome in lung adenocarcinoma. Oncotarget. 2016;7(16):22857-64.

122. Freedman ND, Silverman DT, Hollenbeck AR, Schatzkin A, Abnet CC. Association between smoking and risk of bladder cancer among men and women. Jama. 2011;306(7):737-45.

123. Burger M, Catto JW, Dalbagni G, Grossman HB, Herr H, Karakiewicz P, et al. Epidemiology and risk factors of urothelial bladder cancer. European urology. 2013;63(2):234-41.

124. Nationellt vårdprogram för cancer i urinblåsa, urinledare, urinrör och njurbäcken. Malmö, Sweden: Regionalt cancercentrum syd; 2015.

125. Bellmunt J, Orsola A, Leow JJ, Wiegel T, De Santis M, Horwich A. Bladder cancer: ESMO Practice Guidelines for diagnosis, treatment and follow-up†. Annals of Oncology. 2014;25(suppl_3):iii40-iii8.

126. Cheng L, Montironi R, Davidson DD, Lopez-Beltran A. Staging and reporting of urothelial carcinoma of the urinary bladder. Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc. 2009;22 Suppl 2:S70-95.

127. WHO Classification of Tumours of the Urinary System and Male Genital Organs: WHO/IARC; 2016.

128. Stein JP, Lieskovsky G, Cote R, Groshen S, Feng A-C, Boyd S, et al. Radical Cystectomy in the Treatment of Invasive Bladder Cancer: Long-Term Results in 1,054 Patients. Journal of Clinical Oncology. 2001;19(3):666-75.

129. James ND, Hussain SA, Hall E, Jenkins P, Tremlett J, Rawlings C, et al. Radiotherapy with or without Chemotherapy in Muscle-Invasive Bladder Cancer. New England Journal of Medicine. 2012;366(16):1477-88.

130. Netto GJ. Role for anti-PD-L1 immune checkpoint inhibitor in advanced urothelial carcinoma. The Lancet. 2016;387(10031):1881-2.

131. Wallace DM, Bryan RT, Dunn JA, Begum G, Bathers S. Delay and survival in bladder cancer. BJU international. 2002;89(9):868-78.

132. Knowles MA, Hurst CD. Molecular biology of bladder cancer: new insights into pathogenesis and clinical diversity. Nature reviews Cancer. 2015;15(1):25-41.

133. Neoadjuvant chemotherapy in invasive bladder cancer: a systematic review and meta-analysis. Lancet. 2003;361(9373):1927-34.

134. Bajorin DF, Dodd PM, Mazumdar M, Fazzari M, McCaffrey JA, Scher HI, et al. Long-term survival in metastatic transitional-cell carcinoma and prognostic factors predicting outcome of therapy. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 1999;17(10):3173-81.

135. Tracey E, Roder D, Luke C, Bishop J. Bladder cancer survivals in New South Wales, Australia: why do women have poorer survival than men? BJU international. 2009;104(4):498-504.

136. Thorstenson A, Hagberg O, Ljungberg B, Liedberg F, Jancke G, Holmang S, et al. Gender-related differences in urothelial carcinoma of the bladder: a population-based study from the Swedish National Registry of Urinary Bladder Cancer. Scandinavian journal of urology. 2016;50(4):292-7.

137. Donsky H, Coyle S, Scosyrev E, Messing EM. Sex differences in incidence and mortality of bladder and kidney cancers: national estimates from 49 countries. Urologic oncology. 2014;32(1):40.e23-31.

138. Patafio FM, Robert Siemens D, Wei X, Booth CM. Is there a gender effect in bladder cancer? A population-based study of practice and outcomes. Canadian Urological Association journal = Journal de l'Association des urologues du Canada. 2015;9(7-8):269-74.

139. Dobruch J, Daneshmand S, Fisch M, Lotan Y, Noon AP, Resnick MJ, et al. Gender and Bladder Cancer: A Collaborative Review of Etiology, Biology, and Outcomes. European urology. 2016;69(2):300-10.

140. Lucca I, Klatte T, Fajkovic H, de Martino M, Shariat SF. Gender differences in incidence and outcomes of urothelial and kidney cancer. Nature reviews Urology. 2015;12(10):585-92.

141. Marks P, Soave A, Shariat SF, Fajkovic H, Fisch M, Rink M. Female with bladder cancer: what and why is there a difference? Translational andrology and urology. 2016;5(5):668-82.

142. Shariat SF, Sfakianos JP, Droller MJ, Karakiewicz PI, Meryn S, Bochner BH. The effect of age and gender on bladder cancer: a critical review of the literature. BJU international. 2010;105(3):300-8.

143. Green J, Cairns BJ, Casabonne D, Wright FL, Reeves G, Beral V. Height and cancer incidence in the Million Women Study: prospective cohort, and meta-analysis of prospective studies of height and total cancer risk. The Lancet Oncology. 2011;12(8):785-94.

144. Wiren S, Haggstrom C, Ulmer H, Manjer J, Bjorge T, Nagel G, et al. Pooled cohort study on height and risk of cancer and cancer death. Cancer causes & control : CCC. 2014;25(2):151-9.

145. Gunnell D, Okasha M, Smith GD, Oliver SE, Sandhu J, Holly JM. Height, leg length, and cancer risk: a systematic review. Epidemiologic reviews. 2001;23(2):313-42.

146. Kabat GC, Anderson ML, Heo M, Hosgood HD, 3rd, Kamensky V, Bea JW, et al. Adult stature and risk of cancer at different anatomic sites in a cohort of postmenopausal women. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology. 2013;22(8):1353-63.

147. Kabat GC, Kim MY, Hollenbeck AR, Rohan TE. Attained height, sex, and risk of cancer at different anatomic sites in the NIH-AARP diet and health study. Cancer causes & control : CCC. 2014;25(12):1697-706.

148. Sung J, Song YM, Lawlor DA, Smith GD, Ebrahim S. Height and site-specific cancer risk: A cohort study of a korean adult population. American journal of epidemiology. 2009;170(1):53-64.

149. Jiang Y, Marshall RJ, Walpole SC, Prieto-Merino D, Liu DX, Perry JK. An international ecological study of adult height in relation to cancer incidence for 24 anatomical sites. Cancer causes & control : CCC. 2015;26(3):493-9.

150. Wang F, Xu X, Yang J, Min L, Liang S, Chen Y. Height and lung cancer risk: A meta-analysis of observational studies. PloS one. 2017;12(9):e0185316.

151. Albanes D, Jones DY, Schatzkin A, Micozzi MS, Taylor PR. Adult stature and risk of cancer. Cancer Res. 1988;48(6):1658-62.

152. Batty GD, Shipley MJ, Langenberg C, Marmot MG, Davey Smith G. Adult height in relation to mortality from 14 cancer sites in men in London (UK): evidence from the original Whitehall study. Annals of oncology : official journal of the European Society for Medical Oncology. 2006;17(1):157-66.

153. Nunney L. Size matters: height, cell number and a person's risk of cancer. Proceedings of the Royal Society B: Biological Sciences. 2018;285(1889).

154. Adult height and the risk of cause-specific death and vascular morbidity in 1 million people: individual participant meta-analysis. International journal of epidemiology. 2012;41(5):1419-33.

155. Perkins JM, Subramanian SV, Davey Smith G, Ozaltin E. Adult height, nutrition, and population health. Nutrition reviews. 2016;74(3):149-65.

156. Jelenkovic A, Sund R, Hur YM, Yokoyama Y, Hjelmborg JV, Moller S, et al. Genetic and environmental influences on height from infancy to early adulthood: An individual-based pooled analysis of 45 twin cohorts. Scientific reports. 2016;6:28496.

157. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell. 2011;144(5):646-74.

158. Albanes D, Winick M. Are cell number and cell proliferation risk factors for cancer? Journal of the National Cancer Institute. 1988;80(10):772-4.

159. Noble R, Kaltz O, Hochberg ME. Peto's paradox and human cancers. Philosophical transactions of the Royal Society of London Series B, Biological sciences. 2015;370(1673).

160. Noble R, Kaltz O, Nunney L, Hochberg ME. Overestimating the Role of Environment in Cancers. Cancer prevention research (Philadelphia, Pa). 2016;9(10):773-6.

161. Lin ZW. Analysis of the height dependence of site-specific cancer risk in relation to organ mass. Annals of translational medicine. 2016;4(5):88.

162. Lango Allen H, Estrada K, Lettre G, Berndt SI, Weedon MN, Rivadeneira F, et al. Hundreds of variants clustered in genomic loci and biological pathways affect human height. Nature. 2010;467(7317):832-8.

163. Walter RB, Brasky TM, Buckley SA, Potter JD, White E. Height as an explanatory factor for sex differences in human cancer. Journal of the National Cancer Institute. 2013;105(12):860-8.

164. 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: a cancer journal for clinicians. 2018;68(6):394-424.

165. 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. European journal of epidemiology. 2016;31(2):125-36.

166. Ludvigsson JF, Otterblad-Olausson P, Pettersson BU, Ekbom A. The Swedish personal identity number: possibilities and pitfalls in healthcare and medical research. European journal of epidemiology. 2009;24(11):659-67.

167. Barlow L, Westergren K, Holmberg L, Talback M. The completeness of the Swedish Cancer Register: a sample survey for year 1998. Acta oncologica. 2009;48(1):27-33.

168. Cancer Incidence in Sweden 2011. 54 ed: Socialstyrelsen; 2012.

169. Organization WH. Manual of the international statistical classification of diseases, injuries, and causes of death : based on the recommendations of the seventh revision Conference, 1955, and adopted by the ninth World Health Assembly under the WHO Nomenclature Regulations. Geneva: World Health Organization; 1957.

170. Organization WH. Histological Classification of Neoplasms. Paris: The Committee on Tumour Nomenclature and Statistics of the International Union Against Cancer, Statistics ECOH; 1956. Contract No.: 24.1.

171. Årsrapport från Nationella Lungcancerregistret 2015. Uppsala, Sweden: Regionalt cancercentrum, Uppsala Örebro; 2016 October 2016.

172. Haggstrom C, Liedberg F, Hagberg O, Aljabery F, Strock V, Hosseini A, et al. Cohort profile: The Swedish National Register of Urinary Bladder Cancer (SNRUBC) and the Bladder Cancer Data Base Sweden (BladderBaSe). BMJ open. 2017;7(9):e016606.

173. Urinblåse- och urinvägscancer 2016. In: Samverkan RCi, editor. Årsrapport Nationellt Kvalitetsregister. Lund: Regionalt Onkologiskt Centrum Lund; 2017.

174. 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.

175. Brooke HL, Talbäck M, Hörnblad J, Johansson LA, Ludvigsson JF, Druid H, et al. The Swedish cause of death register. European journal of epidemiology. 2017;32(9):765-73.

176. 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. Journal of clinical epidemiology. 2009;62(11):1202-9.

177. Wettermark B, Hammar N, Fored CM, Leimanis A, Otterblad Olausson P, Bergman U, et al. The new Swedish Prescribed Drug Register--opportunities for pharmacoepidemiological research and experience from the first six months. Pharmacoepidemiology and drug safety. 2007;16(7):726-35.

178. Guidelines for ATC classification and DDD assignment 2019. In: Methodology WCCfDS, editor. Oslo Norway2019.

179. 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. European journal of epidemiology. 2019;34(4):423-37.

180. Benyi E. Tall stature : morbidity, mortality and treatment outcomes. Stockholm: Karolinska Institutet 2016.

181. Sjölander A. An approximate expression for the proportion explained by mediation in survival analysis. Epidemiology (Cambridge, Mass). 2019;in press.

182. Ederer F, Axtell LM, Cutler SJ. The relative survival rate: a statistical methodology. National Cancer Institute monograph. 1961;6:101-21.

183. Welch HG, Black WC. Are deaths within 1 month of cancer-directed surgery attributed to cancer? Journal of the National Cancer Institute. 2002;94(14):1066-70.

184. Hinchliffe SR, Rutherford MJ, Crowther MJ, Nelson CP, Lambert PC. Should relative survival be used with lung cancer data? British journal of cancer. 2012;106(11):1854-9.

185. Radkiewicz C, Johansson ALV, Dickman PW, Lambe M, Edgren G. Sex differences in cancer risk and survival: A Swedish cohort study. European journal of cancer (Oxford, England : 1990). 2017;84:130-40.

186. Dickman PW, Coviello E. Estimating and modelling relative survival. The Stata Journal. 2015;15(1):186-215.

187. Corazziari I, Quinn M, Capocaccia R. Standard cancer patient population for age standardising survival ratios. European journal of cancer (Oxford, England : 1990). 2004;40(15):2307-16.

188. Dickman PW, Sloggett A, Hills M, Hakulinen T. Regression models for relative survival. Statistics in medicine. 2004;23(1):51-64.

189. Royston P, Parmar MK. Flexible parametric proportional-hazards and proportional-odds models for censored survival data, with application to prognostic modelling and estimation of treatment effects. Statistics in medicine. 2002;21(15):2175-97.

190. Lambert PC, Royston P. Further development of flexible parametric models for survival analysis. The Stata Journal. 2009;9(2):265-90.

191. Radkiewicz C, Dickman PW, Johansson ALV, Wagenius G, Edgren G, Lambe M. Sex and survival in non-small cell lung cancer: A nationwide cohort study. PloS one. 2019;14(6):e0219206.

192. Radkiewicz C, Edgren G, Johansson ALV, Jahnson S, Häggström C, Akre O, et al. Sex differences in urothelial bladder cancer survival. Clinical Genitourinary Cancer In press. 2019.

193. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. Journal of chronic diseases. 1987;40(5):373-83.

194. Elixhauser A, Steiner C, Harris DR, Coffey RM. Comorbidity measures for use with administrative data. Medical care. 1998;36(1):8-27.

195. van Walraven C, Austin PC, Jennings A, Quan H, Forster AJ. A modification of the Elixhauser comorbidity measures into a point system for hospital death using administrative data. Medical care. 2009;47(6):626-33.

196. Quan H, Li B, Couris CM, Fushimi K, Graham P, Hider P, et al. Updating and validating the Charlson comorbidity index and score for risk adjustment in hospital discharge abstracts using data from 6 countries. American journal of epidemiology. 2011;173(6):676-82.

197. VanderWeele TJ. Mediation Analysis: A Practitioner's Guide. Annual review of public health. 2016;37:17-32.

198. Lee H, Herbert RD, McAuley JH. Mediation Analysis. Jama. 2019;321(7):697-8.

199. Sjolander A. Regression standardization with the R package stdReg. European journal of epidemiology. 2016;31(6):563-74.

200. Robins JM, Greenland S. Identifiability and exchangeability for direct and indirect effects. Epidemiology (Cambridge, Mass). 1992;3(2):143-55.

201. Proceedings of the 17th Conference in Uncertainty in Artificial Intelligence2001. 647235: Morgan Kaufmann Publishers Inc.

202. VanderWeele T. Explanation in Causal Inference: Oxford University Press Inc; 2015.

203. Nevo D, Liao X, Spiegelman D. Estimation and Inference for the Mediation Proportion. The international journal of biostatistics. 2017;13(2).

204. Hafeman DM. "Proportion explained": a causal interpretation for standard measures of indirect effect? American journal of epidemiology. 2009;170(11):1443-8.

205. Radkiewicz C, Edgren G, Sjölander A, Benyi E, Lambe M, Dickman PW, et al. Can body size explain the excess cancer risk in men? [Original Research]. In press 2019.

206. Wang S, Cowley LA, Liu XS. Sex Differences in Cancer Immunotherapy Efficacy, Biomarkers, and Therapeutic Strategy. Molecules (Basel, Switzerland). 2019;24(18).

207. Peto R. Epidemiology, multistage models, and short-term mutagenicity tests. In: Hiatt HH WJ, Winstein JA, editor. The origins of human cancer New York: Cold Spring Harbor Laboratory; 1977. p. 1403-28.

208. Sulak M, Fong L, Mika K, Chigurupati S, Yon L, Mongan NP, et al. TP53 copy number expansion is associated with the evolution of increased body size and an enhanced DNA damage response in elephants. eLife. 2016;5.

209. Tollis M, Boddy AM, Maley CC. Peto's Paradox: how has evolution solved the problem of cancer prevention? BMC biology. 2017;15(1):60.

210. Savage VM, Allen AP, Brown JH, Gillooly JF, Herman AB, Woodruff WH, et al. Scaling of number, size, and metabolic rate of cells with body size in mammals. Proceedings of the National Academy of Sciences of the United States of America. 2007;104(11):4718-23.

211. Wu S, Powers S, Zhu W, Hannun YA. Substantial contribution of extrinsic risk factors to cancer development. Nature. 2016;529(7584):43-7.

212. Silventoinen K. Determinants of variation in adult body height. Journal of biosocial science. 2003;35(2):263-85.

213. Lindstrom K, Winbladh B, Haglund B, Hjern A. Preterm infants as young adults: a Swedish national cohort study. Pediatrics. 2007;120(1):70-7.

214. Di Renzo GC, Giardina I, Rosati A, Clerici G, Torricelli M, Petraglia F. Maternal risk factors for preterm birth: a country-based population analysis. European journal of obstetrics, gynecology, and reproductive biology. 2011;159(2):342-6.

215. Report no 114: Use of alcohol and tobacco. Stockholm: Statistics Sweden, Unit of Social Welfare; 2007.

216. Wang Y, Freemantle N, Nazareth I, Hunt K. Gender differences in survival and the use of primary care prior to diagnosis of three cancers: an analysis of routinely collected UK general practice data. PloS one. 2014;9(7):e101562.

217. Cheng L, Neumann RM, Weaver AL, Cheville JC, Leibovich BC, Ramnani DM, et al. Grading and staging of bladder carcinoma in transurethral resection specimens. Correlation with 105 matched cystectomy specimens. American journal of clinical pathology. 2000;113(2):275-9.

218. Conforti F, Pala L, Bagnardi V, De Pas T, Martinetti M, Viale G, et al. Cancer immunotherapy efficacy and patients' sex: a systematic review and meta-analysis. The Lancet Oncology. 2018;19(6):737-46.

219. Wallis CJD, Butaney M, Satkunasivam R, Freedland SJ, Patel SP, Hamid O, et al. Association of Patient Sex With Efficacy of Immune Checkpoint Inhibitors and Overall Survival in Advanced Cancers: A Systematic Review and Meta-analysis. JAMA oncology. 2019;5(4):529-36.

220. Chu KC, Miller BA, Feuer EJ, Hankey BF. A method for partitioning cancer mortality trends by factors associated with diagnosis: an application to female breast cancer. Journal of clinical epidemiology. 1994;47(12):1451-61.

221. Dai X, Fang X, Ma Y, Xianyu J. Benign Prostatic Hyperplasia and the Risk of Prostate Cancer and Bladder Cancer: A Meta-Analysis of Observational Studies. Medicine (Baltimore). 2016;95(18):e3493-e.