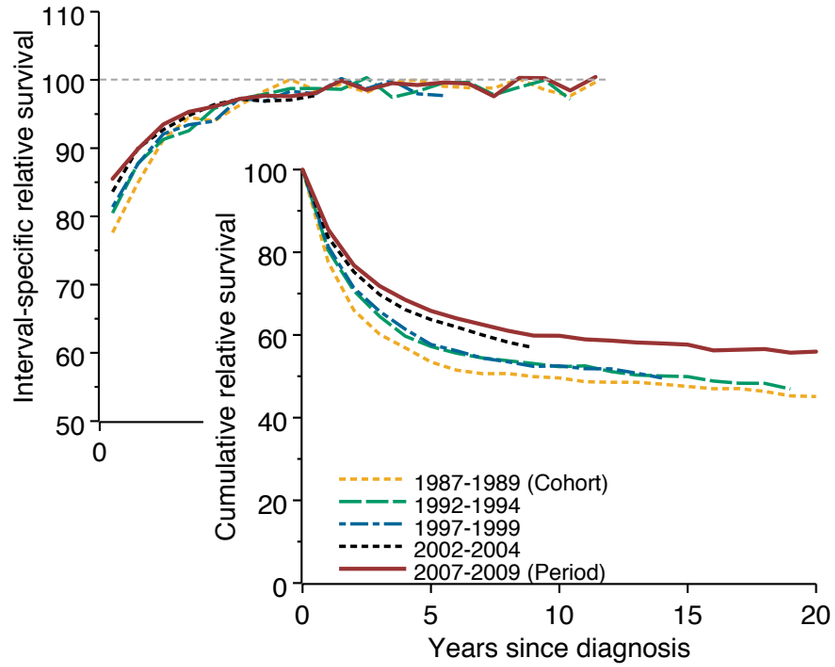
Thesis for doctoral degree (Ph.D.) 2011

Cancer Patient Survival in Sweden – theory and application

Mats Talbäck

Thesis for doctoral degree (Ph.D.)
2011

Cancer Patient Survival in Sweden – theory and application



Mats Talbäck

Karolinska Institutet

Karolinska Institutet

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

Cancer Patient Survival in Sweden

– theory and application

Mats Talbäck

Karolinska Institutet

Stockholm 2011

The figures on the front page represent the interval-specific and cumulative relative survival for females 55 – 74 years of age diagnosed with cancer of the colon.

Published by Karolinska Institutet.
Printed by Larserics Digital Print AB

ISBN 978-91-7457-530-9

"If you only learnt from success, you're not going to learn very much in life"

This quote is attributed to Bill Bowerman, co-founder of Nike.

I saw this quote in an ad for Nike running shoes in the spring of 2003. Since then, I have had it attached to the door of my refrigerator, with the intent of using it for my thesis. I have tried to verify it via Nike. They were not even remotely interested. I also tried the University of Oregon where Bowerman was a track and field coach. At the university they were helpful, but could not verify the quote.

ABSTRACT

Cancer patient survival in Sweden is generally increasing, and Sweden compares well in an international perspective. Despite these achievements, there are nevertheless socio-economic and regional differences in survival that need to be addressed to meet the intensions of the Swedish Health and Medical Service Act. The Act emphasises good health and access to care for the entire population where priority shall be given to those in the greatest need of care. The new organisational structure for oncological care that is now being implemented, with Regional Cancer Centres as central nodes in a network in their respective health care region, will hopefully be able to address these inequalities and take the past achievements in cancer patient survival even further.

The quality and completeness of the Swedish Cancer Register is high. There are nevertheless areas regarding procedures for reporting incident cases to the regional registries and registration that need to be reviewed as a certain degree of under-reporting does exist. The under-reporting appears to be systematic rather than random as it is site-specific, increases with age, and non-reported tumours are often without histopathological verification. For most uses in epidemiology, the degree of under-reporting will be without significant consequences, but to a varying degree, it will have implications for specific research questions. The accuracy of the cancer register should be monitored on a continuous basis rather than in the ad-hoc fashion that has been done so far. This is also the method recommended by the International Agency for Research on Cancer and in guidelines from the Council for Official Statistics at Statistics Sweden.

Two studies in this thesis evaluate period-based and cohort-based analyses ability to predict long-term survival for recently diagnosed cancer patients. Both studies show that period analysis gives better prediction of the future true survival, particularly when not all of the available information is used for the cohort analysis. The first of these two studies was the first systematic evaluation of period analysis that was independent of the researchers who proposed the method. Previous evaluations had mainly been performed using data from the Finnish Cancer Registry. This study had a significant role in demonstrating the utility of period analysis, which has since become an established method in population-based survival analysis. The second study is the first prospective evaluation of the ability of period analysis to predict future survival.

Relative survival is defined as the observed survival of the cancer patients divided by the expected survival of a comparable group from the general population, *free from the specific cancer under study*. However, as expected survival is usually calculated from general population life tables these estimates are biased. This bias is generally ignored since mortality among individuals with a specific cancer is regarded as a small negligible part of the total mortality of the general population. To estimate the size of this bias the Swedish computerised population registers were used to calculate expected survival both including and excluding individuals with cancer. A simple method to correct for this bias using cause of death statistics was also evaluated. The results show that the bias is often sufficiently small to be ignored for most applications, especially for cancers with high or low mortality and for younger age groups. However, for older age groups and for common cancers the bias can be greater than one percent unit, and even larger for all cancer sites combined. The proposed method to correct for this bias seems to work well, and it may often be sufficient to use cause of death statistics for one recent year to gain a satisfactory correction to the bias.

LIST OF PUBLICATIONS

I. Barlow L, Westergren K, Holmberg L, Talbäck M. The completeness of the Swedish Cancer Register: a sample survey for year 1998. Acta Oncol 2009;48:27-33.

II. Talbäck M, Stenbeck M, Rosén M, Barlow L, Glimelius B. Cancer survival in Sweden 1960 – 1998 – Developments across four decades. Acta Oncol 2003;42:637-59.

III. Talbäck M, Stenbeck M, Rosén M. Up-to-date long-term survival of cancer patients: an evaluation of period analysis on Swedish Cancer Registry data. Eur J Cancer 2004;40:1361-72.

IV. Talbäck M, Rosén M, Stenbeck M, Dickman PW. Cancer patient survival in Sweden at the beginning of the third millennium – predictions using period analysis. Cancer Causes Control 2004;15:967-76.

V. Talbäck M, Dickman PW. Predicting the survival of cancer patients recently diagnosed in Sweden and an evaluation of predictions published in 2004. Acta Oncol 2011; doi:10.3109/0284186X.2011.626444.

VI. Talbäck M, Dickman PW. Estimating expected survival probabilities for relative survival analysis – Exploring the impact of including cancer patient mortality in the calculations. Eur J Cancer 2011; doi:10.1016/j.ejca.2011.08.010.

TABLE OF CONTENTS

1 INTRODUCTION

Cancer is the combined name for some 200 diseases that have the common feature that cells in the body become abnormal and divide without control. Cancer cells invade nearby tissues and may spread through the bloodstream and lymphatic system to other parts of the body. Cancer is mostly regarded as a chronic disease and has a number of different causes. Many types of cancer are serious diseases with a high mortality. Cancer is the second most common cause of death in Sweden after cardiovascular diseases, accounting for 26 percent of fatalities for males and 22 percent for females[1]. Since mortality from cardiovascular diseases has decreased substantially during the last decades, cancer has increased its share of the total number of deaths, despite the fact that cancer mortality has also decreased throughout the last decades[1].

Fifty-three thousand persons living in Sweden were diagnosed with cancer in 2009, of which 45,000 were diagnosed for the first time. Indirectly, it was substantially more who were affected. Most of us know someone who has cancer, a family member, a relative, an acquaintance, or a colleague. Approximately one-third of the Swedish population is likely to get a diagnosis of cancer before their seventy-fifth birthday[2].

There were 411,000 persons living in Sweden at the end of 2009 who had previously been diagnosed with cancer. Of these, 172,000 had received their diagnosis within the previous five years. Under the assumption that patients with skin cancer require care for two years, prostate cancer patients for ten years, and other cancer patients for five years, the number of persons needing some form of care in relation to their cancer is expected to increase by 130 percent for males and 70 percent for females to the year 2030[3].

The Swedish population is getting older. The number of people aged 65 years or older has increased by 700,000 during the past four decades and is expected to increase with an additional 600,000 from today to the year 2030. In relative terms, the proportion of people 65 years or older compared to the total Swedish population, was 14 percent 40 years ago, is 19 percent today, and is estimated to be 23 percent in 2030. Compared to the working-age population (25 – 64 year olds), the proportion of people aged 65 years or older has increased from 28 percent 40 years ago, to 37 percent today, and is estimated to be 47 percent in 2030.[4] The older age structure of the population and the increased prevalence of cancer are in itself enough to increase the resources that need to be devoted to cancer. In 2004, the total cost of cancer morbidity for the Swedish society was estimated to 34.1 billion Swedish crowns annually. In current prices, this is expected to increase to 70 billion in 2030. Just over half of the total cost (17.6 billion) constitutes directs costs and can be divided into, primary prevention 0.9, medical care 14.5, medicine 2.0, and screening and other secondary prevention 0.2 billion. The indirect costs correspond to 16.5 billion; increased mortality 13.0, sick leave 1.6, and disability pension 1.9 billion.[5]

The burden of cancer is becoming a major economic expenditure worldwide, and the cost of cancer due to premature death and disability (not including direct medical costs) was estimated to 895 billion US Dollars in 2008. The cost associated with new cancer cases was estimated to at least 286 billion US Dollars in 2009. Medical costs accounted for more than half, and productivity losses for nearly a quarter of this amount. Cancer affects today worldwide an estimated 12 million new patients annually, and leads to more than 7.5 million deaths each year, and it is estimated that by 2030, there will be 27 million new cancer patients per year.[6]

Primary prevention is important, and in order to bring down the burden of cancer on the society, or at least slow down the rate of growth, it is crucial to initiate treatment as soon as possible after diagnosis, to reduce suffering for patients and the future burden for the society. Prevention is the most important and long-term means of reducing the morbidity and mortality from cancer. Approximately 30 percent of the cancer mortality in developed countries could be preventable through lifestyle changes[5]. Reducing the number of smokers is the single most effective way of cancer prevention, and a lower proportion of smokers in the society would also have a large impact on morbidity and mortality as a whole. The proportion of lung cancer mortality in the population attributable to smoking is, for example, estimated to 79 percent for males and 47 percent for females worldwide, and to 92 and 48 percent for males and females, respectively in Europe and Central Asia. For ischemic heart disease, smoking accounts for an estimated 18 percent for males and 6 percent for females worldwide, and to 30 and 5 percent, respectively in Europe and Central Asia. The attributable fraction depends on the prevalence of exposure, the attributable risk at a specific level of exposure, and their distributions, and will therefore vary between populations and regions.[7]

There are also other risk factors, such as obesity, alcohol consumption, and sun exposure that need to be reduced.[8] Vaccination is another possible method for primary prevention. Two vaccines against two types of human papilloma virus, which are behind around 60 – 70 percent of all cervical cancers, have been approved for use in Sweden[5, 9]. Vaccination against these viruses will be introduced into the Swedish national vaccination program for children. The vaccine will be offered to girls when they are in grade 5 or 6 and catch-up vaccination will be offered to girls born in 1999 or later who are now in a higher grade at school.[10] However, the vaccination program has so far been delayed on two occasions due to legal issues regarding the national procurement when the respective pharmaceutical company that lost the contract has appealed. On the first occasion the Administrative Court of Appeal ruled that the procurement had to be redone due to legal issues regarding the length of the contract period[11]. At the time of writing, it is too early to know the outcome of the second appeal that was made to the Administrative Court in late September 2011. The decision from this court can later be appealed to the Administrative Court of Appeal. Primary prevention programs, such as vaccination, are evaluated by monitoring changes in precancerous lesions, cancer incidence, and adherence to the vaccination program.

Cancers that are not prevented need to be diagnosed, and the affected patients treated. Cancer patient survival is a basic measure used to evaluate progress in this specific area of cancer control. Survival can be estimated as the observed survival, the probability of surviving all causes of death for a specified time period, but usually we are more interested in the net survival of the cancer patients. Net survival is a measure of the survival the cancer patients would have achieved in the absence of other cause of death. A new method for directly estimating net survival was presented this year although it can also be estimated using cause-specific survival or relative survival. Relative survival is the method most often used when analysing data from population-based cancer registers, as it does not require information of the cause of death and captures both the direct and indirect mortality of the cancer patients.

Estimates of relative survival from population-based cancer registers can be used to evaluate temporal trends in survival, and survival differences between various groups in society and across nations. Evaluation of specific treatments can most effectively be done in clinical controlled trials to which patients are individually selected and ran-

domly assigned a pre-specified treatment. Population-based estimates can nevertheless provide useful information of cancer treatment and care in the area covered by the cancer registry. Since all patients can be included in the analysis from population-based registers, regardless of treatment, or lack thereof, the real every day value of a program can be accessed and not only the theoretical effectiveness shown in clinical trials.

Population-based cancer registries provide data on incidence, prevalence, and survival for cancer patients in reference to their background population, and the registers are an important source of information for planning and assessing intervention programs against cancer, to establish these priorities, and to evaluate the need for future resources. The registers are thus of considerable value to policy makers at the ministry, government agencies, county councils, public health workers, non-governmental organisations, and not least for epidemiological research.

1.1 AIMS

Overall aim

The overall aim of Karolinska Institutet is to improve human health through research and education. In its extension, this is also the overall objective of postgraduate studies. Sweden has good preconditions for epidemiologic research through computerised population registers, health data registers, and personal identity numbers that enable easy and secure linkage between registers. Through my employment at the National Board of Health and Welfare, where I mainly have been working with the health data registers, it became natural to use these resources to make my (perhaps small) contribution towards improving human health through my education and research. The Swedish Cancer Register has, to a varying degree, always been a part of my daily work. My overall interest and the aim of my research have been to apply, evaluate, and if possible improve methods for cancer survival analysis.

Specific aims

- ***Paper I:*** To estimate the overall coverage of malignant cancer cases in the Swedish Cancer Register by measuring the number of patients diagnosed with cancer in the Hospital Discharge Register who should have been reported to the cancer registry in 1998, and to reveal possible reasons behind the non-reporting.

- ***Paper II:*** To correlate changes in relative survival with the prevention and treatment measures applied in different periods. Changes in patient survival are discussed in relation to available knowledge of factors potentially influencing survival outcome.

- ***Paper III:*** To provide an empirical evaluation of period-based and conventional cohort-based survival analyses ability to predict the long-term relative survival for cancer patients, using data from the Swedish Cancer Register.

- ***Paper IV:*** To provide predictions of the long-term relative survival for cancer patient's diagnosed in 2000 – 2002 using period-based survival analysis, and to compare these estimates with the latest available cohort-based estimates.

- ***Paper V:*** To provide predictions of the long-term relative survival of cancer patients diagnosed in 2005 – 2009 using period-based analysis, to compare these estimates to the latest available cohort-based estimates, and to evaluate predictions published in 2004 (paper IV) regarding the future relative survival of cancer patients diagnosed in 2000 – 2002.

- ***Paper VI:*** To determine the impact on the relative survival ratio by erroneously including cancer patient mortality in the survival probabilities used to estimate the expected survival, and to evaluate a simple method to adjust survival probabilities from general population life tables for cancer patient mortality using only cause of death statistics.

2 A SHORT HISTORY OF POPULATION-BASED CANCER REGISTRATION

Cancer was first documented as a cause of death in 1629 in the Bills of Mortality, produced annually in England[12]. John Graunt a haberdasher in London used the Bills of Mortality in 1662 to describe the patterns of causes of death. This provided an outline for the systematic collection of data on disease and death, a predecessor of epidemiological researchers of today and the use of epidemiology to inform clinical practice.[13] The first systematic attempt to register patients diagnosed with cancer seems to have been in the form of censuses, and the first was the general census of cancer in London in 1728. The first reliable statistics were probably mortality data for the city of Verona from 1842.[14] Demands for improved statistics on the spread of cancer were seen around the year 1900, and an attempt was made to record the prevalence of all cancer patients under medical treatment in Germany on the 15th of October 1900. The survey approach was also used a few years later in other countries, among them Sweden in 1905/1906, and Denmark and Iceland in 1908.[15] However, these surveys were largely unsuccessful due to low participation rates.

The first example of a modern cancer registry is that of Hamburg, which was started with the idea that cancer control not only involves medical and scientific aspects but also public health and economic dimensions. An after-care organisation for cancer patients started on a private basis in 1926 and it obtained official status in 1929. Three nurses regularly visited hospitals and medical practitioners in Hamburg to record new cancer patients. The recorded patients were compared once a week to official death certificates.

Population-based cancer registration started in the state of Connecticut USA in 1935 at the Connecticut State Department of Health. The registry began operating on a statewide basis in 1941, and retrospectively registered cases back to 1935. Registries started in the state of New York (except New York City) in 1940, Denmark in 1942, Norway in 1952, Finland 1953, Iceland in 1954[16], and Sweden in 1958.

At a conference in Denmark in 1946, a group of experts in the field of cancer control recommended the worldwide establishment of cancer registries. In 1950, the World Health Organization established a subcommittee that provided the first methodological guidelines on the registration and statistical presentation of cases of cancer. The same year the need to record all new cases of cancer in defined geographical areas was recognised at an international symposium on geographical pathology and demography of cancer. In 1965, the International Agency for Research on Cancer was established as a cancer research centre of the World Health Organization and in 1966, the International Association of Cancer Registries was established.[16] Approximately 20 percent of the world population is today covered by population-based cancer registries.

When the first volume of Cancer Incidence in Five Continents was published in 1966, it covered a population of 80 million people from 32 registries in 24 countries, approximately 2.7 percent of the world population. Incidence was reported for a three-year period around the year 1961[17]. Volume nine, published in 2007, covers 704 million people from 225 registries in approximately 60 countries. For most registries, the incidence data pertain to the period 1998 – 2002. However, only 11 percent of the world population is covered in the latest volume, and the coverage is very unevenly distributed across the world, ranging from one percent of the population in Africa to 100 percent in the Nordic and some other European countries.[18]

3 CANCER REGISTERS IN CANCER CONTROL

The aim of a cancer control program is to reduce cancer incidence and the morbidity and mortality of cancer patients as well as to improve the quality of life of affected patients and the burden to the society. A population-based cancer register is an essential tool in an effective cancer surveillance program, in which cancer incidence reflects prevention, patient survival reflects the effectiveness of early diagnosis and access to treatment and other health care measures, and prevalence (and cancer mortality) reflects the combination of incidence and survival.

3.1 CANCER CONTROL

Monitoring of disease is an important part in the development and implementation of a coherent cancer control program, and population-based cancer registers are valuable in this context. Without means of identifying the main problems, rational planning will not be possible. Since resources are always going to be scarce, and in practice be insufficient to meet all expectations, there will constantly be a need to prioritise. It is necessary to define priorities between preventive programs and curative and palliative care, to evaluate if earlier goals have been met, and what has been achieved in relation to the resources expended. Thus, it is an important part of any cancer control program to continuously re-evaluate the on-going activities and to modify programs in relation to the achieved results.[19] Even the mere simple fact that the average lifespan is increasing is sufficient to increase the demand for resources for cancer control. Evaluation of the current situation provides a background for the launch of future programs and setting targets provide a means for these programs to be monitored and evaluated. Setting of targets not only involves a comprehensive knowledge of the present situation but also a good assessment and understanding of how it may evolve into the future. However, it is important to recognise that the set targets will inevitably direct resources towards meeting these goals (which of course in part is the objective), but it also means that resources will be drained from other areas. "What's measured gets done!"

3.2 NEW ORGANISATIONAL STRUCTURE FOR ONCOLOGICAL CARE IN SWEDEN

The organisational structure of oncological care in Sweden is currently undergoing structural changes with the implementation of Regional Cancer Centres[5, 20] as central nodes in a network that encompasses their respective health care region, as well as interregional cooperation. In a future perspective, the centres are also expected to develop cooperation with similar structures in other countries. In the field of research and innovations, the new centres should promote both national and international collaboration, develop structures for cooperation with research conducted both at universities and in the industry, and promote innovations in cancer care. The centres should also utilise the competitive advantages that Sweden has through available registers, health data registers, quality registers, and population registers.[20] A commission from the Government and the Swedish Association of Local Authorities and Regions into the quality registers and their future development[21, 22], is also a part in this wider context. The quality registers in Sweden usually contain extensive information about patient characteristics, stage distribution and treatment procedures. This makes it possible to analyse survival for different patient groups in relation to their specific treatment.

I like to stress that this thesis is not about the organisational structure of oncology care in Sweden, and this subject has not been a direct part of any of the papers included in this thesis. The organisation of cancer care is, however, relevant for this thesis since it impacts both the completeness and quality of data used for survival analyses (addressed in paper I), as well as how the results of such analyses can be translated into changes in clinical practice.

3.3 POPULATION-BASED CANCER REGISTERS

Population-based cancer registers are an important part of any cancer control program as they facilitate data on incidence, survival, and prevalence of cancer patients in reference to their background population. Mortality of cancer patients is estimated through cause of death statistics that, according to a World Health Organization agreement between its member states[23], should be available in most countries. The registers are of primary value to clinicians, policy makers at the ministry, government agencies, county councils, public health workers, non-governmental organisations, and not least for epidemiological research. For patients, the benefit is directed towards future patients, where present patients benefit from the past achievements.

All of these four dimensions are needed in order to gain as complete a picture as possible of the cancer burden in society. Incidence and mortality provide a measure of the number of diagnosed and deceased patients, respectively, during a specific period in relation to the size of the background population. Observed survival is an estimate of the actual percentage of patients alive at some specific point in time subsequent to cancer diagnosis, and net survival the proportion of patients alive in the absence of other causes of death. Among population-based cancer registries, relative survival has become the preferred measure of the theoretical concept of net survival.

Incidence and prevalence are the essential dimensions for the allocation of resources. Prevalence is a measure of the number patients living with cancer at a certain point in time. It is a function of incidence and observed survival and can, depending on the objective, be measured either as the number of patients who have previously received a diagnosis of cancer, or as the number of patients still requiring care. In the first case, all patients still alive are included and in the second only those patients who received their diagnosis at a more recent point in time (depending on the type of cancer, e.g., during the latest five years), who are still perceived as being in need of medical care. A statistical approach is to include patients in the prevalence estimates, if they are still alive, until they as a group reach the point of statistical cure (i.e., no longer experience excess mortality). However, an absence of excess mortality is not equivalent with an absence in the need of health care since individual cancer patients often require treatment and experience morbidity long past the statistical point of cure.

Population-based cancer registers can be used in the evaluation of treatment and intervention programs that have been implemented after their effectiveness has been established in randomised clinical trials. Patients are individually selected to clinical trials, are healthier than the average patient (e.g., lower prevalence of co-morbidity), and are thus not representative of cancer patients in general. With a population-based cancer register the real, every day, value of a program can be shown in practise, and not only the more theoretical effectiveness gained through the clinical trial.[19]

3.4 THE ROLE OF SURVIVAL ANALYSIS

Survival estimates from population-based cancer registers may have different interpretations for clinicians, policy makers, and the general public. Population-based survival is often lower than survival obtained from clinical trials. Patients are individually selected to clinical trials and thus not representative of cancer patients in general.

For clinicians, who are the ones primarily subjected to both type of estimates, the population-based estimates represent an unselected group of patients that can be used as an approximate benchmark for cancer patient survival. If the survival of patients treated by a specific clinician, clinic, or hospital differ from the population-based estimate it may be due to a different case-mix of patients or selection by socio-economic, demographic, or geographical region, and shall not unduly be regarded as a sign of better or worse than average performance in the treatment of their cancer patients.[24]

For policy makers and others responsible for allocating resources in the area of cancer, it is important to have reliable estimates of cancer incidence and prevalence in the general population. Prevalence can, as mentioned above, increase from two directions, by an increased incidence and by an improved survival. An increased incidence implies that more resources need to be channelled into initial treatment or, viewed from a long-term perspective, primary prevention. An improved survival implies that resources need to be allocated to subsequent treatment, monitoring, and palliative care for surviving cancer patients.[24]

For the general public, cancer survival becomes a subject of public knowledge and discussion as it is often mentioned in the media. It is up to the scientific community, and the media, to ensure that the right messages are communicated to the general public.[24] Cancer incidence and survival may appear more relevant to the general public than mortality since they relate to those who are still alive. Mortality relates to what has happened to others, whereas survival applies to what could happen to you or someone close to you.

Improvements in survival may be due to new or better diagnostic procedures or an implementation of a screening program, for example, extensive testing for prostate-specific antigen (PSA) resulting in possible over-diagnosis where non-lethal tumours are detected. These developments within the medical sciences are all associated with diagnosis at an earlier point in time. They favour slow-growing tumours at less advanced stages with an apparent improved survival, even if the actual lifespan is not affected. The gain may entirely be due to an increase in lead or length time bias. Despite these reservations, cancer patient survival can nevertheless be useful in assessing the extent to which overall advances in treatments have affected the population.

There are several factors that might distort the interpretation of differences between patients from various groups (e.g., geographical or socio-economic) and temporal trends in survival analyses[25]. Some of these factors can be accounted for in the statistical analysis, whereas other factors depend on the completeness, quality, and validity of the cancer register data. Cancer patients from different groups or time periods may be subject to co-morbidity and mortality to a varying degree. These differences can be accounted for by relative survival analysis. Younger cancer patients generally live longer compared to older patients and female patients longer than their male counterparts. This is generally true even after excluding differences in background mortality through relative survival analysis. Comparisons between groups and time periods therefore also require age standardisation to account for differences in age distribution.[26] Analyses are usually made separately for males and females. There is otherwise also a

need to take different sex distributions between the various populations into account, and to enable comparisons between males and females the same age-distribution must be used in the standardisation. Differences that cannot be accounted for in the statistical analysis may distort comparisons between different groups and time periods. These include; definition of cancer and its morphology, coding practices, treatment, definition of date of diagnosis, follow-up and correct censoring of patients, case mix distribution between and within groups of cancers, and the completeness and quality of registration[26]. However, it is nevertheless important and crucial that these differences are evaluated and discussed in conjunction with the statistical analysis in order for the right conclusions to be drawn.

It is sometimes[e.g.27-30] argued that mortality is a better indicator of the outcome of cancer morbidity than survival since it is generally available for entire populations, and because lead and length time bias may have an impact on survival estimates. However, mortality estimates can seldom in itself distinguish between preventive measures, earlier diagnosis, improved treatments, or changes in incidence. Their temporal trends are often distorted due to the time lag between the onset of disease and death. A time lag that for some forms of cancer can approach decades where the diagnosis of cancer may have occurred in any one of several years subsequent to death.[24]

Coding practices may also change over time, sometimes without any clear explanation. The instruction "[if] there is significant difference in the apparent seriousness of the conditions reported, prefer the more serious condition", in the Sixth Revision of International Classification of Diseases[31] from the World Health Organization, was interpreted in some countries as if cancer should be given priority over other diseases in classifying cause of death. This often misinterpreted instruction was later removed in the Seventh Revision of International Classification of Diseases, where instead the importance of strict adherence to selection rules was stressed.

An example of how changes in coding practices may occur without an apparent explanation is when new personnel were employed at Statistics Sweden as coding staff at the Cause of Death Registry at about the same time as the transition to the Eighth Revision of International Classification of Diseases was made. The new staff began, for some unknown reason, to apply the old selection rules from the sixth revision and overly accepting cancer as the underlying cause of death. They also seem to have applied these rules increasingly more meticulously with time, as they gradually began to rely more on their own judgment rather than on the selection rules in the eighth revision. It is likely that the new staff were not aware of the fact that the coding principles they applied were remnants from an earlier period. The research community began to question the inflated cancer mortality rates towards the late 1970's, but the practise of giving cancer priority over other possible causes of death did not end until the coding instructions for the Cause of Death Register were revised after a detailed audit in the early 1980's. However, it is not very well documented at Statistics Sweden exactly what happened and why.[32] That cancer was given priority over other causes of death is noticeable in the cause of death statistics for 1970's, for which years a marked increase in cancer morality can be seen for some forms of cancer[33]. The mortality is about the same after the audit as before the increase in the early 1970's. This suggests a coding artefact rather than a true increase in cancer mortality.

4 MATERIAL

The data that form the basis for the studies included in this thesis are all part of the Official Statistics of Sweden. The Swedish Cancer Register and the Cause of Death Register are held by the National Board of Health and Welfare and the annual population statistics are maintained by Statistics Sweden. Information on the cancer patients was obtained from the cancer register and the population statistics have been used to estimate the expected survival of the general population, to account for the background mortality in the relative survival analysis.

In Paper VI, cause of death statistics are introduced as a way to adjust the expected survival probabilities from the general population life tables to account for cancer patient mortality, when estimating the expected survival of the cancer patients.

The personal identity numbers that are assigned to each Swedish resident, or a running number in the cases where the identity numbers had been replaced, were used to combine the various registers on an individual level.

4.1 POPULATION REGISTRATION IN SWEDEN

The Swedish Tax Agency maintains the statutory population records, which are the core registration of those who are officially living in Sweden, the Swedish population. It records who lives in the country and where they live. Where you are registered is important for many rights and responsibilities you may have as a Swedish resident. The right to child and housing allowance, where you pay your taxes, and where you vote in the local elections, depends on where you are registered. Marital status, country of birth, and citizenship are, for example, also recorded in the population records. A register over all dwellings is now being established, and the dwelling numbers will be entered into the population register. The combined register will enable a register-based population and housing census, and the traditional censuses will as such no longer be needed.[34] This will be a huge improvement with new possibilities for register-based research, especially since Sweden had its last conventional census over 20 years ago, in 1990.

Sweden has a long history in registering its population. At the beginning it was handled by the church in each parish. The first time church records are mentioned in written documents was in 1571, and the oldest remaining records are from the early seventeenth century. The first nationwide regulation for church records came in 1686, and the priests were required to maintain records of the population residing in their parish.[35] The main objective behind the decision was to facilitate tax collection and to identify young men eligible for military service. In 1749, a nationwide reporting system for cause of death statistics was introduced, and the priests were additionally required to record the cause of death. After various changes over the years, the cause of death registration were adapted to international World Health Organization standards in 1951.[36]

In 1946, through a census reform, it was established that society's need to keep records of its citizens was the main objective of the population records. National registration became a tool to streamline tax collection and control, population and social statistics, and to regulate the labour market. In the 1960's the population records were computerised and in 1967 the check digit was added to the personal identity number that had been introduced in 1947. In 1991, population registration was reorganised and the church records were transferred from the parishes to the population records at the local tax administration within the Tax Agency.[35]

4.2 OFFICIAL STATISTICS OF SWEDEN

There is a distinction between official statistics and other public statistics in Sweden. Official statistics are statistics for public information, planning, and research purposes in specified areas that are produced and published by appointed public authorities in accordance with the Official Statistics Act[37] and the Official Statistics Ordinance[38] issued by the government. Official statistics should be continuous in nature, objective, and made available to the public free of charge at a pre-specified point in time, and must be published in a way that protects the privacy of individuals. It is the authority that is responsible for a specified area of statistics that decides what should be regarded as official statistics or as other public statistics. The National Board of Health and Welfare is responsible for official statistics in the areas of health and medical care and social welfare services, and Statistics Sweden is responsible for the official population statistics, and statistics in various other areas.[39]

The main distinction between official statistics and other public statistics is the requirement that official statistics must be made available to all persons at the same time at a pre-specified publication date, and that it should be free of charge. There is nothing in my experience from Statistics Sweden and the National Board of Health and Welfare that indicates that the public statistics that have not been defined as official statistics are produced with, for example, lower quality than the official statistics.

4.2.1 The Swedish Cancer Register

The Swedish National Cancer Register was established in 1958 and is held by the National Board of Health and Welfare. Cancer registration is statutory in Sweden and organisations operating within the health services should, according to act[40] and government ordinance[41], provide information to the national cancer register. This obligation applies to both the regional and local health services, i.e., counties, municipalities and private healthcare providers.[42] Physicians should report all malignant and certain benign tumours, and pathologists and cytologists should report, separately from the physicians, every tumour diagnosed from surgically removed tissues, biopsies, cytological specimens, bone marrow aspirates, and autopsies. The majority of tumours are thus notified at least twice in separate reports.

The reporting is mandatory to the National Board of Health and Welfare, but the Board has issued a general recommendation[42] that it can be carried out via the regional registries. In principle, all reporting to the national register is today through the regional registries. In October each year, the regional registries deliver data on newly registered cases from the preceding year and corrections concerning previously reported cases to the National Board of Health and Welfare[2].

In 1970, the National Board of Health and Welfare appointed a committee to evaluate oncological care in Sweden. A review of the organisational structure in this area had not been done following the enquiry into the regionalisation of medical care in the late 1950's, and the instigation of regionalised care in the early 1960's. In the instructions to the committee it was specified that any changes to the existing organisational structure that the committee wished to suggest must build on the Swedish system of regionalised health care. Following the report from the committee in 1972, the National Board of Health and Welfare proposed a new organisational structure for oncological care in 1974, which was based on regional oncological centres.[43] The cancer registration was gradually transferred to the oncological centres between 1977 and 1982[44]. The basic

coding and data registration is since then carried out at the regional cancer registries associated with each of the six medical regions. The regionalisation is intended to enable a close contact between the regional registry and the reporting physician, which in turn should facilitate the task of verifying and correcting the material.

Information about stage of disease at diagnosis became a part of the national cancer register towards the end of 2003, through an amendment to the regulations[42]. The regional cancer registries cannot use death certificates or information in the cause of death register to initiate the search for cancer cases due to legal grounds, and are therefore not able to register cases either initiated by, or based solely on, death certificate information.[2]

The purpose of the cancer registry is to monitor cancer incidence and its temporal trends in Sweden, and it aims to provide statistics to the public in a timely and accurate manner. Personal data in the cancer register may be used for production of statistics, monitoring and evaluation of interventions to prevent cancer morbidity, evaluation of screening procedures as well as research and epidemiological investigations[(fn 1) 41].

The national cancer register covers today the 9.4 million persons who have an official residency in Sweden. From its inception to the end of 2009, the register had accumulated information on 2.3 million tumours for 2 million persons. The non-reporting rate has been estimated at less than 2 percent based on death certificates for the late 1970's[45], and to 3.7 percent of the cancer cases reported in 1998 based on a comparison to the Hospital Discharge Register[46]. The overall completeness of the cancer register is high and comparable to other registers in Northern Europe. For most uses in epidemiology, the under-reporting will be without major implications. However, for specific research questions the under-reporting may have consequences as the degree of under-reporting is site specific, increases with age, and does not seem to be random, as diagnoses without histology or cytology verification are overrepresented.[46]

To ensure comparability regarding coding of cancer site and morphology between the regional cancer registries the National Board of Health and Welfare organises regular workshops and meetings. These meetings are attended by coding staff from the regional registries and staff from the National Board of Health and Welfare, including pathologists attached to the national registry. The pathologists answer questions that often have been submitted in advance, and they go through some more complex cases that are known to cause problems among the coding staff.

The regional registries use structured input routines and computerised controls to ensure quality and logic of the registered diagnoses and have procedures for investigating cases that lack either clinical or histopathological/cytological notification. The registries have regular contact with pathologists attached to the local organisation and actively investigate inconsistencies encountered during the registration process by referring questions back to the reporting organisation.[47]

4.2.2 The Cause of Death Register

The Cause of Death Register is held by the National Board of Health and Welfare. All persons who have an official residency in Sweden at their time of death should be included in the register, irrespective of whether the death occurred in Sweden or abroad.

[1] This is a direct translation of 3 § in Ordinance 2001:709 (reference 41). Epidemiological investigations in this context refer to investigations conducted at government agencies. Government agencies use of the cancer register, or health data registers in general, does not constitute research in legal terms.

The personal identity numbers on the cause of death certificate are verified against the population records at the Tax Agency.[36]

The cause of death statistics includes annually up to a hundred deaths that are not reported in the official numbers of deceased from Statistics Sweden. This is because Statistics Sweden produces the official population statistics at the beginning of February each year when some deaths are yet to be reported. For some years, the number missing from the death statistics are considerably larger than a hundred cases. This is usually due to some extraordinary event. According to the Swedish National Police Board there were 543 Swedes missing after the tsunami in Southeast Asia on the 29th of December 2004. Of those persons, the Cause of Death Register includes 522 persons who were registered as residents of Sweden at that time[48]. As they were pronounced dead after Statistics Sweden produced the official statistics for the year 2004, they are not included in the official number of deceased for this year, but are nevertheless included in the cause of death statistics from the National Board of Health and Welfare.

There is a considerable variation in the quality of the cause of death certificates received by the National Board of Health and Welfare. Some certificates are very extensive and minute investigations into the cause of death, whereas others are based only on an external examination of the body together with already available information from patient records. The accuracy of the cause of death registration varies with age, sex, and cause of death. It is more accurate for younger people, for causes of violent nature, and for diseases with a fast progression, than it is for the older people with chronic conditions. The numbers of missing death certificates or certificates with an inadequately specified cause of death have increased from below 1 percent in 1985, 2 percent in 1997, and 3 percent in 2008. The inadequately specified certificates have increased from a few isolated cases annually in the mid 1980's, 0.3 percent in 1995, and 0.8 percent in 2008. This is partly due to an increase in the proportion of older individuals who may suffer from several serious diseases, which makes it more problematic to determine the exact cause of death. The autopsy frequency, which is often used as a summary measure for the quality of the cause of death statistics, has dropped from approximately 40 percent in the mid 1980's to 12 percent 2009[1]. However, it is not certain that the cause of death statistics have become less accurate to the same degree as diagnostic methods have improved and a clinically correct diagnosis is now more likely to be established well before death making an autopsy less crucial.[36]

According to a study on the Cause of Death Register from 1995, made on an aggregated level with about 290 diagnostic groups overall and 30 groups for malignant tumours, the underlying cause was correct for 77 percent of the cases overall and for 90 percent of the malignant tumours. Malignant tumours, that had the highest proportion of correct causes, were over-reported overall by 4 percent, and the wrong tumour had been selected in 7 percent of the cases. Cancer could not be regarded as the underlying cause of death in 4 percent of the cases where a malignant tumour had been selected as the cause of death. As a comparison, ischemic heart disease was correct in 87 percent of the cases and cerebrovascular diseases in only 68 percent of the cases. Some of the inaccuracies recorded in the register cancel each other out, and in broad terms, the cause of death statistics can be regarded as correct. However, on an individual level, the underlying cause of death given on the death certificate appears to be more uncertain, which has consequences for research where the cause of death is needed on an individual level.[49]

The most recent study of the underlying cause of death was done on a random sample of 2.5 percent of the death certificates from 2002. On the three character ICD-10 level the underlying causes was estimated to be incorrect for 3.3 percent of the cases and for 5.4 percent of the cases on the four characters level.[48]

4.2.3 Population statistics in Sweden

Statistics Sweden is responsible for the official population statistics in Sweden. The population statistics are based on data from the population records at the Swedish Tax Agency. The quality of the population statistics is generally good, but some problems exist due to over-coverage, particularly for people born outside of Sweden. The annual population statistics are produced at the beginning of February each year. This means that only records received from the Tax Agency before this are included when the statistics are compiled.

Two weaknesses in the reporting of deaths and emigration result in an over-coverage of the Swedish population. This affects the annual population registers and potentially the studies included in this thesis. The rapid reporting procedure gives rise to a small under-coverage in the number of deceased due to unreported deaths, that for most years are assumed to be less 0.1 percent of the deceased (section 4.2.2), and a subsequent over-coverage in the population[50]. Emigrants cause over-coverage due to unreported migration. Studies from Statistics Sweden show that the end-of-year population registers contain a not insignificant number of people who no longer live in Sweden. This is also the most serious quality problem with the Swedish population records. The Tax Agency estimated in a report from 2006 that the over-coverage in the population records is likely to be in the range of 40,000 to 100,000 people[51]. In a report from Statistics Sweden in 2010, it is noted that the over-coverage of the population records has not declined in recent years, but on the contrary increased[52]. In connection with the establishment of the dwelling register now under way (section 4.1), some corrections might be possible that would reduce the over-coverage in the Swedish population records.

5 METHODS

The statistical methods used in the studies reported in this thesis are primarily in the field of survival analysis. The actuarial method has been used to estimate the observed survival of the cancer patients. Estimates of the expected survival for the general population have been used to relate the observed survival of the cancer patients to that of the general population to form estimates of the patients' relative survival. Period-based analysis has been used as a way to predict the future survival of recently diagnosed cancer patients.

In Paper VI, cause of death statistics are introduced as a way to adjust the expected survival probabilities from the general population life tables to account for cancer patient mortality when estimating the expected survival of the cancer patients.

5.1 RELATIVE SURVIVAL

Cancer patients are exposed to two forces of mortality. The mortality related to the specific cancer under study (the excess mortality) and the mortality from all other causes of death[53]. The excess mortality can be divided into two subgroups; causes directly and causes indirectly related to the diagnosis of cancer. Suicide[54] and treatment-related mortality are examples of the latter.

Relative survival is defined as the observed survival among the cancer patients divided by the expected survival among a comparable group from the general population with respect to the main factors affecting survival, but *free from the specific cancer under study*[53]. As such, the population from which the expected survival is estimated can be regarded as a reference population to the cancer patients. For the studies reported in this thesis, the cancer patients were matched to the expected survival in the general Swedish population on sex, age, and calendar year.

The net survival of the cancer patients is defined as the probability of surviving the cancer in the absence of other causes of death[55]. It is a theoretical concept that is not influenced by changes in mortality from other causes of death, and provides a measure for evaluating temporal trends and comparisons between different subgroups in the society, e.g., socio-economic and ethnic groups, or different regions.

For relative survival to estimate the net survival of the cancer patients it is assumed that the mortality in the general population, from which the expected survival of the cancer patients is estimated, adequately accounts for all causes of mortality and that the cancer-related mortality is negligible compared to all other causes of death. Net survival also require independence (conditional on covariates) between deaths from the specific cancer under study and deaths due to other causes.[55] Only when these conditions are met, can relative survival provide a good estimate of the net survival of the cancer patients.

The assumption that the expected survival from the general population could, for all practical purposes, be regarded as emanating from a disease-free population with regard to the specific cancer under study is investigated in paper VI, while this premise is assumed to be met in papers II to V.

The cumulative relative survival ratio can be interpreted as the proportion of patients alive after a given time of follow-up in the hypothetical situation where the cancer in question is the only possible cause of death (the net survival). An interval-specific relative survival ratio of 100 percent indicates that, during this particular interval (year of

follow-up), mortality in the patient group was equivalent to that of the general population. If this level is maintained during subsequent years of follow-up, there is no longer evidence of an excess mortality due to cancer and the patients, as a group, can be considered as "statistically cured".[53]

The possibility to interpret the relative survival estimates are reduced when population life tables do not correctly represent the background mortality of the cancer patients under study. This problem occurs when mortality from other (non-cancer) causes of death among the cancer patients differ from that of the reference population, e.g., the general population life tables. Factors can be either positively (e.g., screening or higher socio-economic status), or negatively (e.g., smoking or lower socio-economic status) associated with mortality from other causes. Relative survival estimates require population life tables that are matched to the cancer population by age, sex, and calendar year, but also when required ethnicity, socio-economy, and region. Risk factors may also be taken into account, such as smoking when lung cancer is studied. Lung cancer patients will experience higher morbidity and mortality than the general population due to other smoking-related illnesses.

However, it is important to recognise that the requirement of the population life tables to correctly represent the background mortality does not make the relative survival estimates invalid if this requirement is not met as long as the estimates are interpreted in the correct way. If general population life tables are used the relative survival estimates can always be interpreted as the ratio of the observed survival of the cancer patients to the survival of the general population matched on, e.g., sex, age and calendar year. It is the interpretation of the ratio as the net survival of the cancer patients that can be problematic if the life tables do not accurately reflect the background mortality of the cancer patients under study.

Despite these potential weaknesses, relative survival has some considerable advantages over cause-specific survival for the analysis of population-based data. Relative survival does not require information on the cause of death and does as such not rely on the accuracy of the underlying causes of death assigned to the cancer patients. Even if some of the inaccuracies in the cause of death registration cancel each other out (section 4.2.2), and on an aggregated level can be regarded as more or less correct, it appears that the underlying cause of death is more uncertain on an individual level[49]. Cause-specific survival has not been used systematically in population-based survival analysis due to these concerns regarding the misclassification of the underlying cause of death. The exclusion of deaths from non-cancer causes that are a consequence of treatment, either directly or indirectly may not be cancer deaths in a strict medical sense, but they nonetheless reflect the consequences of cancer. They ought therefore to be considered as deaths due to cancer since the death would not have occurred if the patient had not had cancer.

However, an attempt has recently been made to develop a new classification method for "cancer death" that combines information from the cancer register with information from the cause of death registration. In cases where appropriate life tables were available, this new classification gave cause-specific survival estimates that were consistent with relative survival estimates. In situations when the population life tables does not appropriately reflect the background mortality patterns in the cancer population, this new classification may be a viable option to relative survival estimates and provide useful cause-specific estimates of cancer patient survival.[56]

Another advantage with relative survival compared to cause-specific survival is that information on the date of death is usually available much sooner than the underlying cause of death. In Sweden, the date of death is available within two months after the end of a calendar year whereas the underlying cause of death was, up until recently, only available after approximately one and a half year. The time-lag for the completion of the cause of death register has been gradually reduced and the cause of death statistics for the year 2010 were published in July 2011[57].

5.1.1 Estimating expected survival

There are traditionally three methods for estimating expected survival in relative survival analysis and these methods have been available for approximately 30 years or more. The methods are commonly referred to as Ederer I[53], Ederer II[58], and Hakulinen[59] after the respective lead authors. Earlier this year a method was proposed that directly estimates the net survival, and has so far been referred to as Pohar-Perme, again in reference to the lead author[60].

The three traditional methods differ in how long the matched (reference) individuals from the general population are considered to be under risk and included in the calculation of the expected survival. In the Ederer I estimate it is assumed that the cancer patient would be a member of the general population during the entire length of potential follow-up and the matched individuals are consequently considered to be under risk indefinitely, and the point in time at which a cancer patient is censored or dies has no effect on the expected survival. The Hakulinen method is similar to Ederer I but takes account of informative censoring and addresses the situation of heterogeneous potential follow-up time of the cancer patients. A matched individual is assumed to be under risk until the end of the potential follow-up period if the patient dies, and is censored if the cancer patient is censored. In the Ederer II estimates, the matched individuals are considered to be under risk until the corresponding cancer patient is censored or dies.

The method used to estimate the expected survival for the relative survival analysis in the papers included this thesis is the one that at the time of analysis was generally regarded as the method of choice. This means that the Hakulinen method was used in paper II to IV and VI. Although paper VI was accepted for publication in September 2011, the statistical analysis was performed in late 2005 and early 2006.

Earlier this year a study was published that compared the Ederer II and Hakulinen methods to the reference standard for the cumulative relative survival ratio, which is regarded to be the directly age-standardised relative survival ratio with weights proportional to the number of patients at the beginning of follow-up. Both theoretical and empirical results from this study showed good (and better) agreement between the Ederer II estimates and the reference standard. The authors conclude that there seems to be time to make a change from previous recommendations, and adopt the Ederer II method when reporting cumulative relative survival. The authors also note that the use of the Ederer II method is particularly important for long-term follow-up where estimates are more based on the younger patients.[61] In view of these results, the Ederer II estimates were used in the analysis for paper V, which was performed during the spring and summer of this year.

In the Pohar-Perme estimator, as for the Ederer II, the matched individuals are considered to be under risk until the corresponding cancer patient is censored or dies. To account for the proportion of patients that are expected to have been lost from mor-

tality due to other causes, the number of events and the number under risk are weighted by the inverse of the sum of the expected survival probabilities. This will increase the sample size at each time interval to the level that would have been if cancer truly were the only possible cause of death. This theoretically accounts for mortality due to other causes and the estimator will, in itself, be an estimate of the net survival.

However, even with this new estimator, the assumption that cancer-related mortality is negligible compared to all other causes of death has to be met and the expected survival probabilities should be calculated after excluding mortality due to the specific cancer under study. Otherwise, the estimator will overcompensate for mortality due to other causes and increase the number of events and the number under risk even with cancer-related mortality. By applying the new estimator with the adjustment to the expected survival probabilities proposed in paper VI, whenever necessary, one could possibly come even closer to the true net survival. Although the new estimator is theoretically promising, it still needs to be evaluated empirically before it can be used in regular analysis of cancer patient survival.

5.2 PERIOD ANALYSIS

Period-based life tables have been used for a long time in demography when estimating the remaining life expectancy of a population, and it is reported annually by the national statistical bureaus around the world. Life expectancy is one of the most basic measures of how public health develops, and it is often used when comparing different countries[33, 62]. Period-based analysis was not introduced into register-based cancer patient survival analysis until 1996[63].

The period life table is based on population and mortality statistics pertaining to a short narrow (recent) time period. The remaining life expectancy at a specific age represents the average number of years of life left for a group of persons who at that age were to live through their entire remaining lifespan with the same mortality patterns as in this particular time period.

The period life expectancy has historically been shown to be an underestimation for young and middle age persons, but a good estimate of the average remaining lifespan for older individuals. The older a person become, the better the period life expectancy will be as an estimate of that person's actual remaining lifespan, and the younger, the more underestimated the future lifespan is likely to become as we expect the mortality rates to decline in the future. Of course, if something should happen that changes the forces of mortality in the future, such as the spread of antibiotic-resistant bacteria[64], the tangible lifespan of those living today will be overestimated. On the other hand, if a common cure for cancer was found or if the mortality related to cardiovascular disease were to be reduced drastically, the average lifespan would likely be underestimated.

The alternative to use period life tables when estimating life expectancy would be to go back through the generations to the early 20th century and calculate the cohort life tables. Needless to say, if we wish to look forward, we know that young and middle-aged people living today will not live through the same mortality patterns experienced by previous generations, who have been affected by, for example, the Spanish flu, and the two world wars. What is certain is that the future will hold events unknown today that will affect the average lifespan of coming generations, but we do not know what those events will be.

5.2.1 Period cancer survival analysis

In period-based survival analysis only person-time at risk and events (death or censoring) occurring during a particular calendar period are considered. The estimates are obtained by left truncation of all observations at the beginning of the period and right censoring at the end of the period.[63] Cohort-based estimates represent the survival experience of a well-defined cohort of patients diagnosed during a specified calendar period. Period-based estimates, in contrast, do not represent the survival of any actual cohort of patients, instead they represent the survival that would be observed for a hypothetical cohort (a synthetic cohort) of patients if they were to experience the same interval-specific (conditional) survival estimates as the patients who were actually at risk during the specified calendar period. If prognosis improves over time, the period estimates are expected to be higher than those obtained by a corresponding cohort analysis. The opposite would be expected if survival was declining, and no difference between cohort and period estimates would be seen if survival was constant over time. Empirical studies comparing the two methods using historical data show that period-based estimates from a given time period in most cases predict, quite well, the long-term survival for patients diagnosed during that particular period[65-67].

Improvements in cancer patient survival between subsequent years of diagnosis are generally most pronounced during the first few years following diagnosis (see figure on front page). Period-based estimates of survival tend to be optimistic since period analysis systematically selects the interval-specific relative survival ratios from the "best" cohort of patients, i.e., recent cohorts for the early years of follow-up and early cohorts for the later years. However, this inherent overestimation is counterbalanced by the fact that cancer patient survival is generally improving over time, and the period estimates are therefore expected to be an underestimate of the subsequent true survival. This is the same as when period life tables are used to estimate the expectation of life.

Figure 5.1 shows period- and cohort-based interval-specific relative survival estimates for prostate cancer patients. The period-based survival curve has an unusual shape with declining estimates over time. The period estimates are based on a synthetic cohort of patients who are diagnosed during consecutive years in time, and may as such be diagnosed and treated with different methods at their respective time of diagnoses. The synthetic cohort will thus consist of patients with different (over time often improving) prognoses at their respective time of diagnosis. For prostate cancer, this is represented by recently diagnosed patients who in general have a good PSA-screening-related prognosis dominating the first years of follow-up for the synthetic cohort. Follow-up for the latter years are in contrast dominated by patients diagnosed earlier in time who in general have a much worse prognosis, i.e., less PSA-screening-related. The period estimates will as such be misleading when used for cancers with large improvements in survival.

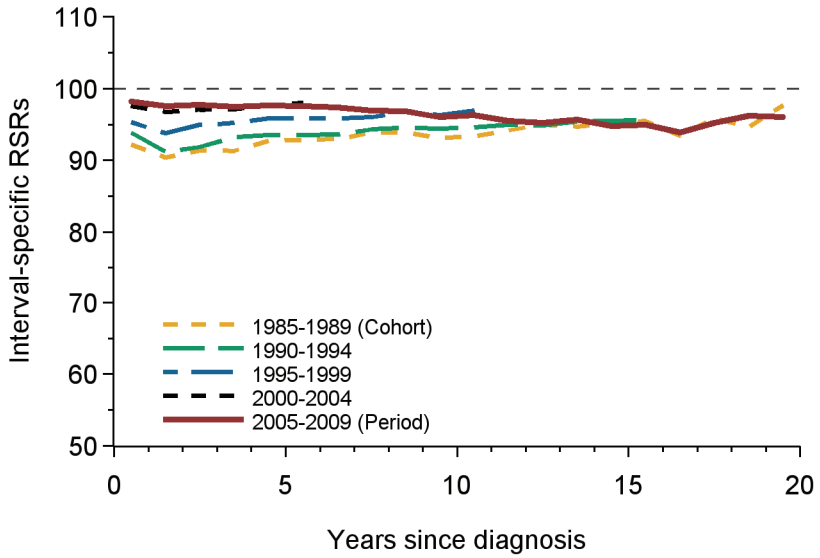


Figure 5.1 Interval-specific relative survival ratios (RSR) for males diagnosed with prostate cancer 1985 – 2009. Age-standardised RSRs for males 45 – 89 years of age at diagnosis. The figure illustrates the unusual shape of a declining interval-specific survival curve for the period-based estimates.

The example below illustrates the concept of period-based survival analysis, using females diagnosed with breast cancer, site and time period chosen for illustrational purpose. The period-based estimates for 1990 are calculated from the annual interval-specific survival estimates for the years 1986 – 1990, the top right to bottom left diagonal highlighted in blue shades in Table 5.1. A gradual improvement in relative survival can be seen for female breast cancer patients during these five years.

Table 5.1 Interval-specific and cumulative relative survival ratios for females diagnosed with breast cancer 1986 – 1990. Females 55 – 74 years of age at diagnosis.

	Interval-specific survival						
	Year of diagnosis						**Period estimate for 1990**
Year of follow-up	**1986**	**1987**	**1988**	**1989**	**1990**	**1986-1990**	
1	0.956	0.950	0.958	0.961	**0.970**	0.959	**0.970**
2	0.953	0.947	0.956	**0.964**	0.963	0.957	**0.964**
3	0.951	0.953	**0.961**	0.962	0.969	0.959	**0.961**
4	0.948	**0.966**	0.956	0.968	0.961	0.960	**0.966**
5	**0.957**	0.954	0.969	0.956	0.970	0.961	**0.957**
Cumulative	0.786	0.789	0.815	0.824	0.844	0.812	0.830

Survival for the first year after diagnosis is based on the interval-specific survival from the first year of follow-up for patients diagnosed in 1990, and survival for the second year, is based on the conditional interval-specific survival from the second year of follow-up for patients diagnosed in 1989. For the fifth year of follow-up, the survival is based on the conditional interval-specific survival from the fifth year of follow-up for patients diagnosed in 1986. These estimates are then multiplied consecutively, and constitute the cumulative relative survival for the period 1990.

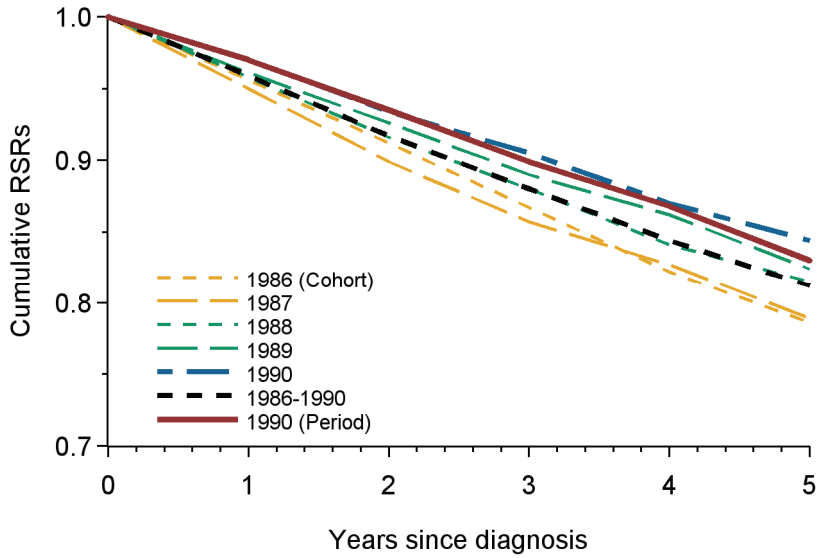


Figure 5.2 Cumulative relative survival ratios (RSR) for females diagnosed with breast cancer 1986 – 1990. Females 55 – 74 years of age at diagnosis.

As shown in Table 5.1, and graphically illustrated in Figure 5.2, the cumulative relative survival estimate for the cohort of patients diagnosed in 1986 – 1990 is 1.8 percent units below the period-based estimate after five years of follow-up. The period-based estimate is in itself 1.4 percent units below the cohort of patients diagnosed in 1990, since it is based on information that is up to five years old. Compared to patients diagnosed in 1986 (five years earlier), the period-based estimate is nonetheless 4.4 percent units higher.

5.3 ADJUSTMENT OF EXPECTED SURVIVAL PROBABILITIES

Relative survival is defined as the observed survival of the cancer patients divided by the expected survival of a comparable group from the general population, but *free from the specific cancer under study*.[53] As the general population survival probabilities from ordinary published life tables reflect the mortality from all causes of death, it would seem necessary to adjust these probabilities for cancer patient mortality if they are to be regarded as free from the specific cancer under study. However, this is rarely, if ever done, since the mortality from a specific form of cancer often is regarded as a small negligible part of the total mortality and that the ordinary life table estimates as such should provide a satisfactory proxy for a disease-free reference population with respect to cancer.

For younger patients and less common forms of cancer this seems plausible, but it could be questioned for older age groups, common forms of cancer, and for all sites combined. There are some references to adjustment of the population survival probabilities in a seminal article by Ederer et al from 1961[53], and in an academic dissertation from the University of Tampere in Finland from 1998[68]. Ederer et al conclude in their article that, "since we are usually concerned with analysing survival of patients with specific forms of cancer, it appears that we do not need to make an adjustment in estimating expected survival from population life tables".

It is possible to adjust the ordinary survival probabilities from general population life tables, and to account for cancer patient mortality if the proportion of cancer deaths in the population is known. In this way, it would thus be possible to attain an estimate of

the expected survival that is closer to the theoretical definition when calculating relative survival.

The mortality in the general population can be decomposed as

$$\lambda_{gp} = \lambda_c + \lambda_o$$

where λ_c is the mortality due to the specific cancer under study and λ_o is the mortality due to all other causes of death. For relative survival we require an estimate of λ_o, i.e., the expected mortality of a population free from the specific cancer under study. The expression above can be rearranged and λ_o written as

$$\begin{aligned}\lambda_o &= \lambda_{gp} - \lambda_c \\ &= \lambda_{gp}(1-\propto)\end{aligned}$$

where $\propto$ is the proportion of deaths in the general population due to the specific cancer under study.

On the probability scale this can be written as

$$\begin{aligned}\pi_o &= exp[-\lambda_o] \\ &= exp\left[-\lambda_{gp}(1-\propto)\right] \\ &= exp\left[-\lambda_{gp}\right]^{(1-\propto)} \\ &= \pi_{gp}^{(1-\propto)}\end{aligned}$$

where π_o is the probability of death in a population *free from the specific cancer under study* and π_{gp} the probability of death in the general population.

In practise, these calculations are done specifically for each combination of sex, age, and calendar year, and if required, e.g., socio-economic status, ethnicity, or region. Subscripts have been omitted in order to simplify the notation.

In paper VI[69], the proportion of cancer deaths in the general population $(\propto)$ is estimated from cause of death statistics. The cause of death registration is not perfect (section 4.2.2), but as it in broad terms can be regarded as correct on an aggregated level[49], the hypothesis on the outset was that the cause of death statistics could be used to give a usable estimate of the proportion of cancer deaths in the general population. Cause of death statistics should be readily available in most countries according to a World Health Organization agreement with its member states on the compilation and publication of statistics with respect to diseases and causes of death[23].

6 SUMMARY OF PAPERS

6.1 PAPER I. THE COMPLETENESS OF THE SWEDISH CANCER REGISTER – A sample survey for year 1998

Introduction

The Swedish Cancer Register is extensively used as an important tool to monitor cancer incidence and survival, and for research purposes. Users of the register depend on the completeness and quality of registration for their analysis. The cancer register is generally considered to be of good quality as approximately 99 percent of the cases are morphologically verified. The proportion of cases not reported was estimated to be less than 2 percent in the late 1970's based on information from death certificates.

Aim

The aim of the study was to estimate the overall coverage of malignant cancer cases in the Swedish Cancer Register by measuring the number of patients diagnosed with cancer in the Hospital Discharge Register that should have been reported to the cancer registry in 1998, and to reveal possible reasons behind the non-reporting.

Material and methods

All malignant cancer cases in the Hospital Discharge Register from 1998 were selected and compared to those recorded in the cancer register as an incident case for the year 1998. There were 43,761 hospital discharges for 42,010 individuals of whom 3,429 individuals were not recorded in the cancer register. From these 3,429 records, a simple random sample of 202 patients was selected for review of their medical records to determine whether they should have been included in the cancer register as an incident cases for 1998. The medical records were obtained from the hospitals where the patients had been treated and were reviewed by a trained medical secretary with 25 years of experience of cancer registration. Sex, age, type of tumour, mode of diagnosis, and type of hospital were also studied in order to identify possible amendable problems with reporting.

Results

Approximately half of the 202 randomly selected cases (93 malignant and 8 benign) should have been reported as an incident case for the year 1998. This translates into an additional 1,579 malignant cases (95% CI 1,349 -1,808), or 3.7% of the cases reported in 1998. The crude incidence rate for males and females combined would increase from 493 per 100,000 to 511 (95% CI 508 - 514) if these cases were taken into account.

Conclusion

The overall completeness of the Swedish Cancer Register is high and comparable to other cancer registers in Northern Europe. For most uses in epidemiological or public health surveillance, the under-reporting will be without major impact. However, for specific research questions the findings may have implications, as the degree of under-reporting is site specific, increases with age, and does not seem to be random, as diagnoses without histology or cytology verification are overrepresented.

6.2 PAPER II. CANCER SURVIVAL IN SWEDEN 1960 – 1998

– Developments across four decades

Introduction

This study was conducted as a follow-up on an earlier series of articles on the survival of cancer patients in Sweden diagnosed in 1961 – 1989 that was published as a supplement to Acta Oncologica in 1995[70]. The study extends previous analyses by nine years and summarises the relative survival for cancer patients diagnosed in 1960 – 1998, i.e., during the first four decades of national cancer registration in Sweden.

Aim

The aim was to correlate changes in relative survival with the prevention and treatment measures applied at different time periods. The changes are commented on and discussed in relation to available knowledge of factors potentially influencing survival outcome.

Material and methods

This study was based on all cancer cases in the Swedish Cancer Register diagnosed in patients less than 90 years of age in 1960 – 1998. At the time of analysis, the register had been updated with dates of death, censoring, and confirmed continued residency in Sweden up to the end of 2000. Patients were followed for a maximum of 30 years after diagnosis. Forty different forms of cancer and all sites combined were analysed. Cumulative and interval-specific relative survival was estimated using the Hakulinen and Ederer II methods, respectively. Survival was evaluated as three-year moving averages.

The aim of the study was to correlate chances in relative survival with the prevention and treatment measures applied at different time periods in the past, and cohort-based rather than period-based analysis was thus deemed most appropriate. Period-based analyses would in this case lead to questionable interpretations since cohorts diagnosed and treated at different time periods contribute to the overall period survival estimates.

Results

The analyses showed that the relative survival of people diagnosed with cancer continued to improve, and the improvements during the 1990's were in general of about the same magnitude as in earlier decades, in both short- and long-term. The expectation of life of a person diagnosed with cancer (all sites combined) was determined to be about seven years longer than for a person diagnosed during the mid-1960's. Approximately, three years were gained due to diagnosis later in life and about four years due to improved survival that was partly due to site migration. Sites with poorer survival had decreased and sites with better survival had increased their relative shares of the total incidence. The increase in life expectancy was determined to be due to improved health in general, and to better supportive care, on the one hand, and to earlier and enhanced diagnostic and treatment methods, on the other.

Three ultimate success stories were identified; the improved treatments for Hodgkin's lymphoma, non-seminoma testicular cancer, and acute lymphocytic leukaemia, all of which had led to a major increase in relative survival. At the time of analysis, it was deemed that nearly all of these patients survived and except for those diagnosed with Hodgkin's lymphoma, patients had - after a few years - the same death rates as the gen-

eral population. For both seminoma and non-seminoma testicular cancer, stabilisation of the interval-specific relative survival occurred after approximately five years during all four decades. This was taken as an indication that lead-time bias had not played a major part in this improvement. In paper IV and V it is shown that survival is now virtually identical for patients diagnosed with seminoma and non-seminoma testicular cancer.

Conclusion

This analysis provided a comprehensive overview of the changes in cancer patient survival during the first four decades of national cancer registration in Sweden. However, cohort-based cancer patient survival evaluates historical data, and will be of limited interest to the health service if lead-time, staging procedures, and treatments have changed over time. There is really no good solution to overcome this time lag if we aim to correlate decisions taken and methods used at different time periods with their subsequent impact on survival. Period-based analysis would in this case not be suitable as its estimate pertain from patients diagnosed and treated in different years. Cohort-based analysis should be used if patients diagnosed within the same time interval in the past is of primary interest.

The changes in relative survival for various cancers were discussed in relation to developments in the Swedish health-care system over the decades. However, this may have led to an overemphasis of the importance of treatment since this was the main focus of the paper, e.g., the general health status of cancer patients compared to earlier periods. A group of experts in the general field of oncology was invited, and before they were shown preliminary results of the analysis, they were asked to describe factors of relevance and in what ways they might have influenced survival. A surprisingly good concurrence was observed between the expert's prior assessments and the temporal trends in survival.

It was concluded that longer follow-up period than the traditional five years were generally needed to evaluate trends in survival. To ensure that no short- or long-term survival improvement is overlooked the entire survival curve should be considered, with no particular concentration on any specific point in time.

6.3 PAPER III. UP-TO-DATE LONG-TERM SURVIVAL OF CANCER PATIENTS

– An evaluation of period analysis on Swedish Cancer Registry data

Introduction

The natural developments of cancers as well as the measures to fight the diseases are often long processes that might require decades of follow-up. Information of long-term survival from cohort-based analyses will thus often appear out-dated, may pertain to clinical methods no longer in use, and will as such be regarded as irrelevant. The time lag between diagnosis and long-term survival can be reduced with period-based analysis that was introduced into cancer patient survival analysis in 1996. Empirical evaluations had at the time of analysis mainly been performed on data from the Finnish Cancer Registry and by the main promoters of period-based survival analysis.

Aim

The aim of this study was to provide an empirical evaluation for data in the Swedish Cancer Register of period-based and conventional cohort-based survival analysis ability to predict the long-term survival for cancer patients.

Material and methods

This study was based on all cancer cases in the Swedish Cancer Register diagnosed in patients less than 90 years of age in 1960 – 1998. At the time of analysis, the register had been updated with dates of death, censoring, and confirmed continued residency in Sweden up to the end of 2001. Forty different forms of cancer and all sites combined were analysed. This is the same data material that was used for the second paper with one additional year of follow-up.

Relative survival was estimated using the Hakulinen method. The 5-, 10-, and 15-year relative survival actually observed for different cohorts of patients were calculated and compared to the most up-to-date cohort estimates available at that particular cohort's time of diagnosis, using two conventional estimates of cohort-based survival and period-based survival. "Cohort analysis" evaluates survival for cohorts of patients diagnosed at close proximity in time to each other (e.g., within the same calendar year). "Complete analysis" additionally includes patients diagnosed in later years, providing a mixture of patients with short- and long-term potential follow-up time. Relative survival was estimated for one-year intervals, and for three- and five-year moving averages. Single year estimates were used to evaluate common cancers. For less common cancer, three- or five-year moving averages were used, making a necessary trade-off between up-to-date information and precision. The mean difference and the mean squared difference between the relative survival actually observed and the latest cohort, complete, and period estimates available at the time of diagnosis were calculated. The mean difference is a measure of systematic over- or underestimation of the relative survival and the mean squared difference quantifies the degree of deviation.

Results

The analyses showed that period relative survival in general is a better estimate of the subsequent observed survival compared to the cohort and complete estimate. The variability between the period and observed estimates are in most cases much smaller than

for the corresponding cohort and complete estimates. The difference between the estimated and observed survival increase with increasing length of follow-up, and the estimated survival are in most cases an underestimation of the subsequent observed survival. The latter indicates that for many forms of cancer there had been on-going improvements in survival during the past decades. According to the cohort, complete, and period estimates that were comparable in time to the latest available observed estimate, some forms of cancer appeared to have had an on-going improvement in survival, whereas for other forms of cancer, the improvements in survival appeared to have levelled off.

Figure 6.3.1 show a graphical representation of the differences between the cohort-, complete-, and period-based relative survival estimates available at the time of diagnosis and the subsequent observed survival for males and females 0 – 89 years of age at diagnosis. Estimates close to the diagonal line indicate a good agreement between the estimates. Overestimation, compared to the observed survival, is indicated by estimates in the upper left hand triangle, whereas underestimation is indicated by estimates in the lower right hand triangle.

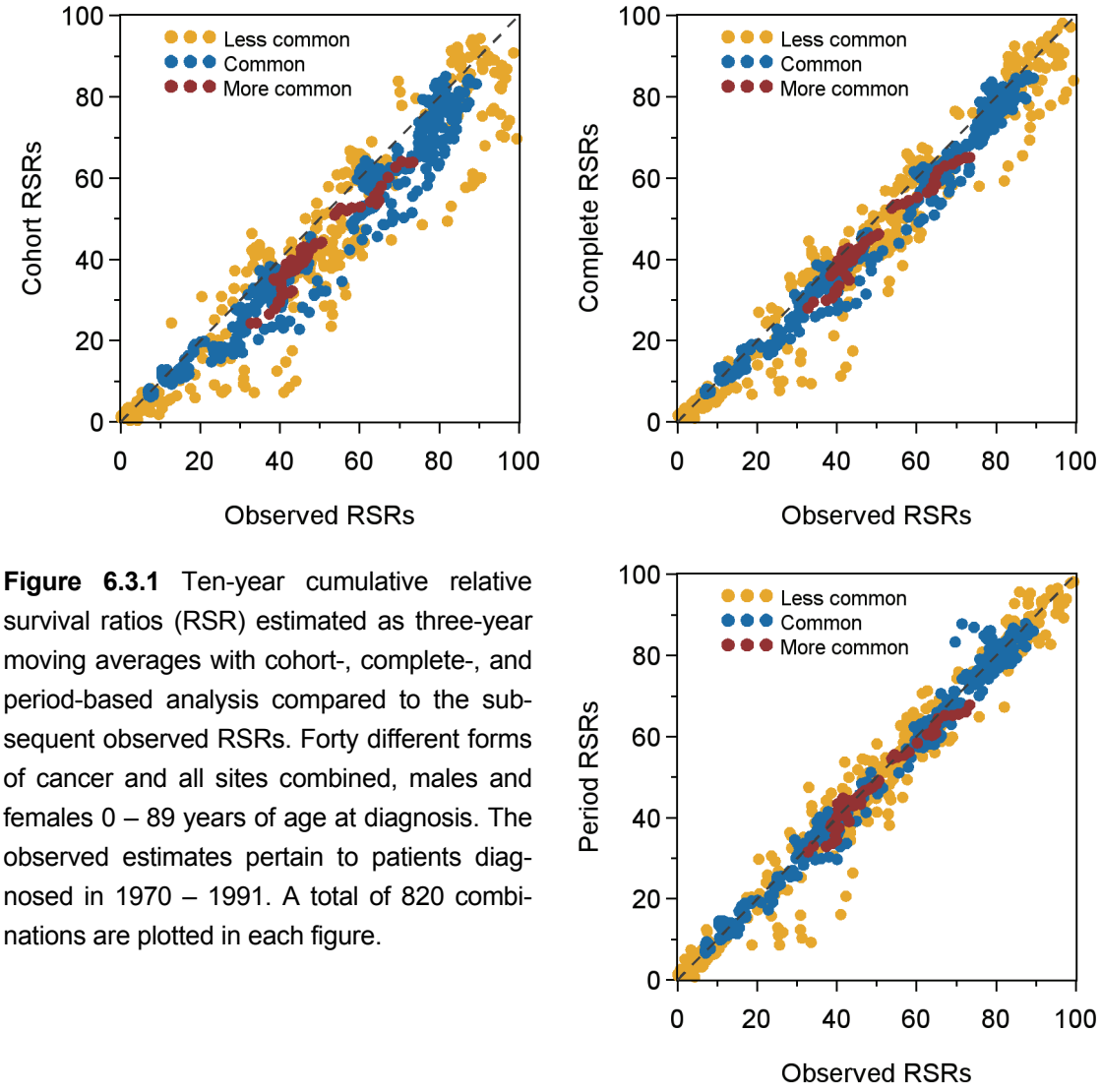


Figure 6.3.1 Ten-year cumulative relative survival ratios (RSR) estimated as three-year moving averages with cohort-, complete-, and period-based analysis compared to the subsequent observed RSRs. Forty different forms of cancer and all sites combined, males and females 0 – 89 years of age at diagnosis. The observed estimates pertain to patients diagnosed in 1970 – 1991. A total of 820 combinations are plotted in each figure.

A larger difference can be seen between the cohort and the observed relative survival ratios compared to the difference between the complete, period and observed estimates. It is also evident from the upper left hand figure that the cohort estimates underestimate the subsequent observed survival to a much larger extent than the complete and period estimates. This is especially true for common or more common forms of cancer, which usually become underestimated by cohort-based analysis. The smallest differences can be seen between the period and the observed estimates both in terms of absolute proximity to the diagonal line and in terms of underestimation of the observed survival.

Conclusion

This study confirmed earlier results from the Finnish Cancer Registry that period-based analysis in general provides more up-to-date estimates of long-term cancer survival than cohort-based analysis, and that period-based survival estimates from a given time period in most cases quite accurately predict the long-term survival for patients diagnosed at that point in time.

6.4 PAPER IV. CANCER PATIENT SURVIVAL IN SWEDEN AT THE BEGINNING OF THE THIRD MILLENNIUM

– Predictions using period analysis

Introduction

This paper builds on the comprehensive retrospective evaluation of period-based analysis made on Swedish Cancer Register data in paper III, and extends the analysis into predictions of the future long-term survival for cancer patients diagnosed in 2000 – 2002.

Aim

The aim of this study was to provide predictions using period-based analysis of the long-term relative survival for cancer patients diagnosed in the early 21th century. The period-based estimates were compared with the latest available cohort-based estimates in order to detect on-going improvements in survival.

Material and methods

This study was based on all cancer cases in the Swedish Cancer Register diagnosed in patients less than 90 years of age in 1980 – 2001. At the time of analysis, the register had been updated with dates of death, censoring, and confirmed continued residency in Sweden up to the end of 2002. Patients were followed for a maximum of 20 years after diagnosis. Forty different forms of cancer and all sites combined, and all sites combined excluding breast cancer for females and prostate cancer for males were analysed. Patient survival was estimated for males and females separately as well as for both sexes combined and for different age groups. The published article reports only results for males and females and for the age group 0 – 89 years. Detailed results for each site, stratified by age at diagnosis are available at the National Board of Health and Welfare web site (http://www.socialstyrelsen.se/statistics/cancersurvival).

Cumulative and interval-specific relative survival was estimated using the Hakulinen and Ederer II methods, respectively. Period-based analysis was made for the period 2000 – 2002, and cohort-based analysis was made for patients diagnosed in 1995 – 1997, 1990 – 1992, 1985 – 1987, and 1980 – 1982. The latter provide observed relative survival for the latest available corresponding 5-, 10-, 15-, and 20-year cohort-based estimates that are comparable to the period-based estimates. Approximate ratio and year of stabilisation of the interval-specific relative survival curves was determined by visual inspection.

Results

The analyses showed that there had been improvement in relative survival for many forms of cancer during the past two decades. For some sites survival seemed to have stabilised at a relatively high level, whereas other sites showed a continued poor survival that had remained essentially unchanged for the past two decades. Females had in general a better survival than males. The observed improvements in survival over time were primarily due to improvements during the first few years after diagnosis.

For all sites combined, the differences between the interval-specific survival for the period and cohort estimates were negligible approximately six years following diagnosis for males and approximately four years following diagnosis for females. For both

the period and cohorts analysed, the interval-specific relative survival curves stabilised for males at 97 percent after eight years of follow-up, and for females at 98 percent after seven years of follow-up. This suggests that cancer patients in general have an excess mortality for a considerable length of time after diagnosis. The fact that the interval-specific relative survival does not quite reach 100 percent during the 20 years of follow-up is probably because the general population mortality underestimates the expected mortality of the cancer patients. This is primarily due to the higher prevalence of smoking among cancer patients.

The interval-specific relative survival for prostate cancer and breast cancer exhibited an unusual pattern (Figure 6.4.1). Excess mortality for most forms of cancer is usually highest immediately following diagnosis and the level generally decreases with increasing follow-up time until the patients reach the point of statistical cure. In contrast, prostate and breast cancer patients experienced an approximately constant excess mortality throughout the first 20 years of follow-up.

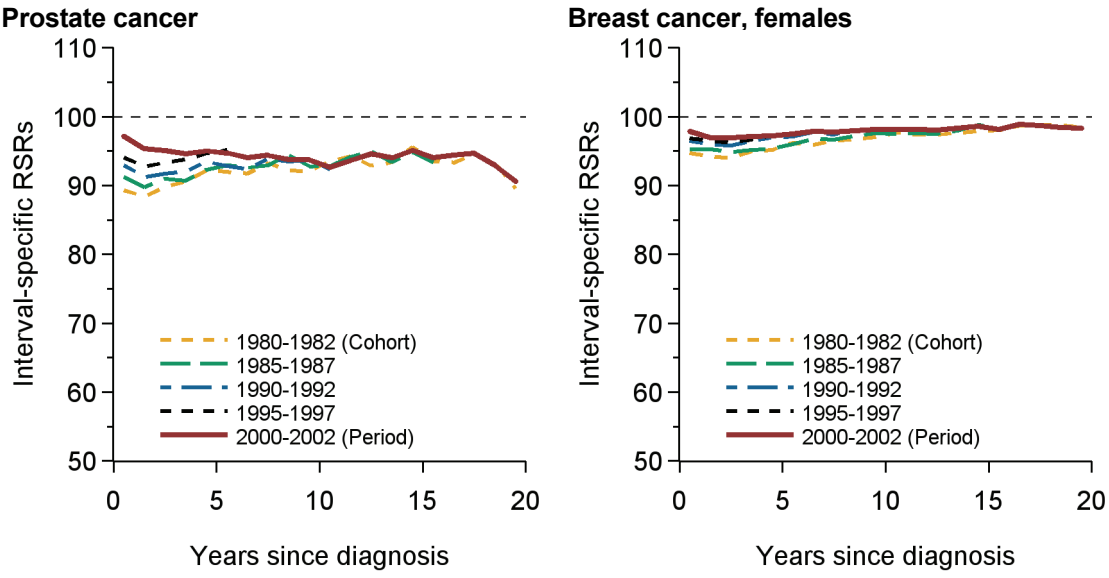


Figure 6.4.1 Interval-specific relative survival ratios (RSR) for prostate cancer and breast cancer for females. Males and females 0 – 89 years of age at diagnosis. Figure illustrating the unusual shape of the interval-specific relative survival curves for the period and cohort estimates.

Excess mortality for prostate cancer patients was slightly lower during the first two years following diagnosis. After seven years, the interval-specific relative survival curves levelled off at approximately 93 percent for the period and cohorts analysed. Female breast cancer survival exhibited a pattern similar to that of prostate cancer and the period estimate stabilised at 98 percent after six years of follow-up.

Patients diagnosed with some forms of cancer were shown to have a very favourable prognosis where only a small excess mortality could be seen for the first few years following diagnosis. This has been true for many years for seminoma testicular cancer. For non-seminoma testicular cancer, that had previously had a poor survival compared to seminoma, the cumulative relative survival was now closer to that for seminoma. Patients diagnosed with neurinoma have had a good survival since the latter part of the

1970's and now virtually no survival disadvantage could be seen for recently diagnosed patients.

However, there remained several cancer sites for which patient survival continued to be poor and for which only some long-term improvements on a low level could be seen. Fortunately, these sites also showed short-term improvements. For example, patients diagnosed with liver cancer had experienced an increase in 1- and 2-year cumulative relative survival of approximately ten percent units during the past decade.

Conclusion

This study showed that patient survival for many forms of cancer could be expected to be higher than previously estimated by cohort-based survival analysis. It was deemed likely that the lead-time introduced by prostate cancer screening would cause the period-based analysis to overestimate the 10-year cumulative relative survival for clinically detected prostate cancers by 20 percent, but for PSA-detected cancers, the period estimate was likely to be an underestimate of the true survival.

It was concluded that it is important to consider both short- and long-term survival in order to obtain a complete picture of trends in patient survival. There is otherwise a risk that short-term improvements will be missed if attention is directed at long-term survival at fixed intervals such as 5-, and 10-years. Cancer sites with low and constant long-term survival should be considered at shorter follow-up intervals than the five years traditionally used. However, it does not matter for short-term survival whether cohort-based or period-based analysis is used since, depending on the years included, these estimates are essentially the same.

6.5 PAPER V. PREDICTING THE SURVIVAL OF CANCER PATIENTS RECENTLY DIAGNOSED IN SWEDEN AND AN EVALUATION OF PREDICTIONS PUBLISHED IN 2004

Introduction

In paper IV, predictions were published of the long-term relative survival of cancer patients diagnosed in 2000 – 2002. The predictions were made by period-based analysis and as the 5-year follow-up for patients diagnosed in 2002 is now available, it is possible to evaluate the extent to which the predictions came true. An evaluation of the results from paper IV showed that the long-term relative survival for less common cancer sites were not stable over time even with survival estimates averaging over three years. Survival estimates averaging over five years are used in this study for all cancer sites. This is the first prospective evaluation of period-based survival analysis on its ability to predict future survival.

Aim

The first aim of this study was to provide up-to-date predictions of the long-term relative survival for cancer patients diagnosed in 2005 – 2009, and the second aim was to evaluate the predictions for 5-year relative survival made in paper IV seven years ago.

Material and methods

This study was based on all cancer cases in the Swedish Cancer Register diagnosed in patients less than 90 years of age in 1985 – 2009. At the time of analysis, the register had been updated with dates of death, censoring, and confirmed continued residency in Sweden up to the end of 2009. Patients were followed for a maximum of 20 years after diagnosis. Forty different forms of cancer were analysed.

Cumulative and interval-specific relative survival was estimated using the Ederer II method. Period-based analysis was made for the period 2005 – 2009, and cohort-based analysis was made for patients diagnosed in 2000 – 2004, 1995 – 1999, 1990 – 1994 and 1985 – 1989. The latter provide observed relative survival for the latest available corresponding 5-, 10-, 15-, and 20-year cohorts-based estimates that are comparable to the period-based estimates. Approximate ratio and year of stabilisation of the interval-specific relative survival curves was determined by visual inspection. The survival estimates of the main analyses were age-standardised according to the site-specific age structure of patients diagnosed during the period 2005 – 2009. The same weights were used for males and females.

The evaluation of the predictions published in 2004 was made with the Hakulinen method and was not age-standardised, as this was not done in the earlier study.

Results

The analyses showed that the period-based analysis from 2004 predict the future true survival quite well with more common, larger sites, showing a better agreement than smaller, less common, sites (Figure 6.5.1).

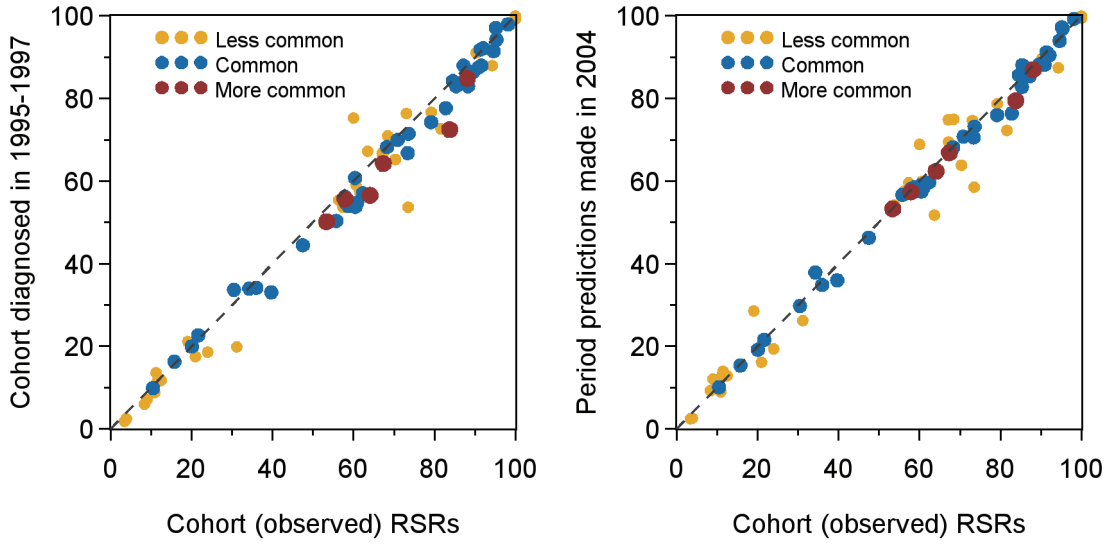


Figure 6.5.1 Five-year cumulative relative survival ratios (RSR) estimated for the latest available cohort-based analysis for patients with 5-years of follow-up in 2002 (patients diagnosed in 1995 – 1997) and from the period 2000 – 2002 in the analysis publicised in 2004[71], and the subsequent cumulative RSRs later observed for the actual cohort of patients diagnosed in 2000 – 2002. Estimates for both males and females are plotted in each graph. Males and females 0 – 89 years of age at diagnosis.

Points on the diagonal line represent perfect agreement between the earlier cohort-based estimates, the period-based predictions, and the subsequent true cohort-based survival, respectively. More common sites (red dots), are represented by all sites combined, all sites combined excluding prostate cancer for males and breast for females, and prostate and breast cancer for males and females, respectively. Common sites (blue dots), are represented by sites reported in paper III with a period/cohort width of three years, less common sites (yellow dots), are represented by sites reported with a period/cohort width of five years in paper III.

In general, most of the results shown in the paper IV still hold in paper V, even after the seven years that separate the main periods of analysis in the two studies. This is despite the differences in analytical methods applied, not age-standardised verses age-standardised estimates, and Hakulinen versus Ederer II. Based on a comparison of the period and cohort estimates there is still evidence of improvement in relative survival for many forms of cancer during the past two decades. As before, survival seems to have stabilised at a relatively high level for some sites, whereas other sites show a continued poor survival that has remained essentially unchanged for a long time.

The unusual pattern for the interval-specific relative survival estimates for prostate and breast cancer patients seen in paper IV has become more pronounced, especially for prostate cancer patients for whom the excess mortality was slightly lower during the first years following diagnosis. For prostate cancer patients, the interval-specific relative survival now levels off at some 96 percent after approximately 10 years of follow-up for both the period and cohorts analysed. For female breast cancer patients, the interval-specific relative survival level off at some 98 percent after approximately five years of follow-up.

Conclusion

The evaluation of the period-based analysis from 2004 presented in this study for up to 5-year relative survival shows once again that period-based analysis in most cases provide better predictions of the long-term future survival for recently diagnosed cancer patients compared to cohort-based analysis. Period-based, rather than cohort-based, survival analysis should be used if the primary aim is to predict the future survival of recently diagnosed patients. Recently diagnosed cancer patients can expect an improved survival for many forms of cancers compared to patients diagnosed only a few years earlier.

6.6 PAPER VI. ESTIMATING EXPECTED SURVIVAL PROBABILITIES FOR RELATIVE SURVIVAL ANALYSIS

– Exploring the impact of including cancer patient mortality in the calculations

Introduction

Relative survival is the preferred measure of cancer patient survival used by population-based cancer registries. It is defined as the observed survival of the cancer patients divided by the expected survival of a comparable group from the general population, but *free from the specific cancer under study*. However, the ordinary population-based survival probabilities for the general population reflect the mortality from all causes of death. It would therefore seem necessary to adjust these probabilities for cancer patient mortality if they are to be regarded as free from the specific cancer under study. This is however, rarely, if ever done, since the mortality from a specific form of cancer is regarded as a small negligible part of the total mortality, and the ordinary life table estimates should as such provide a satisfactory proxy for a disease-free reference population with regard to cancer. This seems plausible for cancers where the prevalence is low, as for younger patients, less common forms of cancer, or when the mortality is high, but it could be questioned for older age groups, common forms of cancer, and of course, for all cancer sites combined. There are some references to adjustment of the population survival probabilities in a seminal article by Ederer et al from 1961 on the relative survival rate, and in an academic dissertation from the University of Tampere in Finland from 1998. Ederer et al conclude, "[that] since we are usually concerned with analysing survival of patients with specific forms of cancer, it appears that we do not need to make an adjustment in estimating expected survival from population life tables". In the article, they evaluate lung cancer and conclude that the small differences between the relative survival ratios calculated with and without adjustment for smoking habits are due to the low survival of the lung cancer patients.

Aim

The first aim of this study was to determine the impact on the relative survival ratios by including cancer patient mortality in the expected survival probabilities calculated from general population life tables. The second aim was to evaluate a simple method to adjust the survival probabilities from general population life tables for cancer patient mortality using only cause of death statistics.

Material and methods

This study was based on the individual end-of-year population records from the Swedish Total Population Register for the years 1986 – 2002, and all cancer cases in the Swedish Cancer Register diagnosed in patients 1958 – 2001. At the time of analysis, the cancer register had been updated with dates of death, censoring, and confirmed continued residency in Sweden up to the end of 2002. Complete follow-up was available for over 99% of the reported cases.

The use of the individual population records for the total Swedish population enabled censoring of the cancer patients on an individual basis at their time of diagnosis. This contributes a disease-free population with respect to cancer from which to calculate population survival probabilities that is unaffected by cancer patient mortality (the

reference). The relative survival ratios estimated with these disease-free survival probabilities were compared to the relative survival ratios estimated with both ordinary unadjusted and adjusted expected survival probabilities from general population life tables.

The survival probabilities from the general population life tables were adjusted with the following equation, to more correctly reflect the survival of a disease-free population with respect to the specific cancer under study. Only information regarding the numbers and causes of death is needed to do the adjustment. Subscripts for sex, age, and period (year) have been omitted to simplify the notation.

$P_{\overline{C}} = P_{gp}^{(1-\propto)}$, where

$P_{\overline{C}}$ = Probability of surviving at least one additional period (year), for individuals not diagnosed with the specific cancer under study.

P_{gp} = Probability of surviving at least one additional period (year) in the general population, e.g., expected survival probabilities from general population life tables.

$\propto = \frac{D_C}{D_T}$, Proportion of deaths in the general population due to the specific cancer under study.

D_T = Total number of deaths in the general population, from cause of death statistics.

D_C = Number of deaths in the general population due to the specific cancer under study, from cause of death statistics.

Five common forms of cancer and all sites combined were analysed. The individual sites were chosen to represent different survival patterns among cancer patients. Colorectal cancer represent a common cancer with an intermediate cancer patient survival, lung cancer represent a cancer with a short survival, skin cancer represent a cancer with a more favourable survival, breast cancer is the most common female cancer, and prostate cancer is the most common male cancer. Prostate cancer also represents a cancer common in the older ages. All sites combined were included as it is often sought after, analysed, and reported.

Results

The analyses showed that there in general is an overestimation of the relative survival ratios when survival probabilities from general population life tables are used compared to the reference standard where the survival probabilities do not include cancer patient mortality. The overestimation of the relative survival ratios indicate that the survival probabilities become too low when the cancer patients (excess) mortality are included in the survival probabilities from the general population. The differences increase with the length of follow-up. The results also show that it is quite simple to adjust the survival probabilities from general population life tables for cancer patient mortality. The only additional information that is needed is cause of death statistics, and information from a single (recent) year is often sufficient to gain an approximate adjustment.

Conclusion

This study shows that it is possible to use cause of death statistics to adjust survival probabilities from general population life tables for cancer patient mortality. The bias of the unadjusted estimates is sufficiently small to be ignored for most applications, notably for cancers with high or low mortality and for younger age groups (e.g., < 60 years). However, for common cancer sites, the bias in the relative survival ratios can for older age groups be greater than one percent unit and up to five percent units after ten years for all cancer sites combined. The bias in 10-year relative survival for men aged 75+ diagnosed with prostate cancer was 2.6 percent units, which I think is of sufficient magnitude to warrant an adjustment.

7 DISCUSSION

Cancer patient survival in Sweden is generally increasing and has been doing so for decades[70-72]. In international comparisons, Swedish cancer patients have in general a comparatively good survival[73-76]. Despite the temporal improvements and in an international perspective good survival, there are nevertheless socio-economic and regional differences in survival among cancer patients in Sweden[e.g.77, 78]. In an unpublished analysis of ten common forms of cancer regarding the socio-economic differences in relative survival for patients diagnosed 1987 – 2001, I found that the absolute differences in survival decreased over time. However, the relative differences remained and even appeared to become larger over time. In a report from the National Board of Health and Welfare to the Government in early 2011[77], we found larger socio-economic differences in period-based relative survival for 12 common forms of cancer compared to county or regional differences. The report concludes that these differences reflect a multi-factorial process, but does not venture into suggesting any specific causes. The report calls for an analysis of the possible causes to see if there are any interventions that could be implemented that would significantly improve the survival of the disadvantaged groups. According to the report, the next step should be to analyse and identify variations in survival depending on disease stage at diagnosis. That special attention needs to be taken regarding differences between population groups are emphasised in the memorandum outlining the criteria that shall characterise the new Regional Cancer Centres[20]. An overall aim of cancer patient care is that it should be equal and of high quality across the country, and special measures should be taken at the centres to achieve greater equality between population groups based on gender, age, socio-economic conditions, and place of residence[20].

In an analysis of data from EUROCARE-3, it was found that the marked differences in survival for breast cancer across Europe were mainly due to differences in cancer stage at diagnosis[79]. Patients residing in areas with better diagnostic facilities for staging are likely to have an improved stage-specific survival compared to patients from areas where diagnostics are less developed, or for other reasons is not performed as thoroughly. An improved diagnosis and staging also leads to more appropriate treatment with a subsequent better survival. More advanced staging techniques, or more extensive diagnosis, will detect metastases for tumours that otherwise would have been classified as having a less advanced stage, stage migration. Because the survival of patients who migrate to a group with a more advanced stage, although worse than that for other members of the group with a less advanced stage, their survival will be better than for other members of the group with a more advanced stage. Survival will thus increase in both groups without any change in survival for the individual patient or for the group as a whole. This is often referred to as Will Rogers's phenomenon[80] in the field of epidemiology. Data for Sweden was not included in the analysis of the EUROCARE-3 data, but it is likely that this result also should be valid for cancer patients in Sweden. A stage-specific analysis, would at least in part take lead-time into consideration, which is likely to account for some of the observed survival differences between the socio-economic groups. As for the observed geographical differences in survival, at least some of these should have its origin in socio-economic differences between the counties, and possibly also for some of the differences between the medical regions.

In paper I[46], investigating the completeness of the Swedish Cancer Register, we concluded, based on studies made in other countries, that the completeness of the Swedish

Cancer Register is high and comparable to the other cancer registers in Northern Europe. Together with five other countries, Swedish Cancer Register data were also selected to be part of the International Cancer Benchmarking Partnership for its long-standing and high quality population-based cancer registration[75].

A search in Medline via PubMed on the 23rd of August 2011 for articles with a title or abstract referring to the completeness or quality of the Swedish Cancer Register[(fn 2)], revealed 67 articles published between 1971 and 2011. However, only 14 of these articles[45, 46, 81-92] were studying the completeness or quality of the Swedish Cancer Register. In most of these 14 studies patients' medical records were reviewed, eight were done on a national level whereas six referred to one of the regional cancer registers. Ten studies focused on some specific cancer or related cancers, three studies involved all sites or groups of sites, and one study investigated multiple primaries. In two studies cause of death certificates were used to find potentially additional cancer cases and hospital records were used in three studies. According to these 14 studies it seems that the overall coverage of the Swedish Cancer Register is high and in parity with other registers. It also appears that the under-reporting that does exist has diminished in magnitude over time, is site specific, increases with age, and that non-reported tumours are often without histopathological or cytological verification. There might also be a diagnostic disparity in the clinical setting[e.g.88] that is carried over to the cancer register.

In 2007[93], an Official Government Inquiry proposed that cause of death certificates should be used to supplement the current notification system for cancer registration. However, this was neither included in an update of the Official Secrecy Act[94] in 2008, nor exempt from the general rule of absolute confidentiality in the Public Access to Information and Secrecy Act[95], that succeeded the Official Secrecy Act in 2009.

The importance of using cause of death certificates in cancer registration to diminish the non-notification "as practised in other Nordic countries" was stressed in one of the earlier studies of the completeness of the Swedish Cancer Register that was published in 1984. It is also mentioned in this article that cause of death certificates will be collected and used as a supplement to the cancer notifications on a regional basis.[45] However, as far as I have been able to determine this seem never to have been realised. In an article regarding the completeness and accuracy in registration of acute leukaemia in adults diagnosed 1987 – 1992 it was found that the combined use of the cancer and cause of death registries gave "acceptable coverage"[87].

It is of course not only the completeness of the register that is of importance. For most applications in epidemiology and public health surveillance, it is likely to be more important that the information in the register is correct than if some tumours are missing. Strictly, this is only true if the under-reporting can be regarded as completely at random, and since we did not find this to be the case in paper I, the under-reporting will, to a varying degree, have implications for specific research questions.

To me, it seems preferable that the completeness of the cancer register and the quality of registration is monitored on a continuous basis rather than in an ad-hoc fashion as has been done so far. This is also the method advocated by the International Agency for Research on Cancer[96]. In the guidelines from the Council for Official Statistics at Statistics Sweden it is stated that quality studies should be planned and carried out regu-

[2] Search question: [(cancer[Title/Abstract]) AND (registry[Title/Abstract] OR register[Title/Abstract]) AND (sweden[Title/Abstract] OR swedish[Title/Abstract]) AND (completeness[Title/Abstract] OR quality[Title/Abstract] OR validation[Title/Abstract])]

larly and that annual monitoring of the quality of the statistics should be carried out with the main users of the statistics[97]. None of these guidelines is, at least to the extent that I have been able to determine, adhered to today or has previously been followed. There are nevertheless systems in place at the regional cancer registries to ensure that the registered information in itself is logical and correct, but there are no procedures to ensure completeness and quality in a broader sense.

According to volume nine of Cancer Incidence in Five Continents, the overall reporting to the Swedish Cancer Register for the years 1998 – 2002 is estimated to be 96 percent of all diagnosed cases[18]. This is in good agreement with the overall estimate of the under-reporting of 3.7 percent of the reported cases for the year 1998 reported in paper I. The proportion of microscopically verified[fn 3)] cases is, according to Cancer Incidence in Five Continents, around 98 percent for major cancer sites. However, for pancreatic cancer the proportion is only around 87 percent and the mortality/incidence ratio 149 percent for males and 163 percent for females. A high mortality/incidence ratio is an indication that some incident cases are not recorded in the cancer register[98]. In a steady-state environment there would be a constant ratio between incidence and mortality. For highly fatal cancers, such as pancreatic cancer, and in the absence of reporting errors, this ratio would be expected to be 100 percent or slightly below. In a study of the reporting and long-term survival of patients diagnosed with pancreatic cancer in 1987 – 1999 it was found that there was a considerable under-reporting to the Swedish Cancer Register and that the under-reporting increased with age at diagnosis and was more pronounced during the second half of the period[92]. For lung cancer, the mortality/incidence ratio is 106 for males and 102 for females while the ratio for all cancer sites combined is around 50 percent and between 50 and 25 percent for other major cancer sites. The large mortality/incidence ratio for cancers with a high fatality is also in agreement with the findings in paper I in which it was found that many of the patients in the reviewed sample had a severely deteriorated general condition. This is emphasised by that 55 percent of the patients also died in 1998 and an additional 18 percent died in 1999. In the annual incidence report from the Swedish Cancer Registry[e.g.2] it is noted the number of persons in the Cause of Death Register with cancer stated as the underlying cause of death who had not been recorded in the Cancer Register. The largest differences are seen for older persons with fatal cancers as their underlying cause of death, i.e., lung and pancreatic cancer. However, the vast majority of these cases had not undergone an autopsy.

In paper III[65], an evaluation on historical cancer register data of the ability for the period- and cohort-based survival approaches to predict the long-term future survival for recently diagnosed cancer patients was made. This was also evaluated in paper V[99], using a prospective approach for the five-year survival of patients diagnosed in 2000 – 2002. Both studies show that period-based analysis in general gives better predictions of the subsequent true survival, especially when not all the available information is used in the cohort analysis[65]. That period-based analysis in general gives better predictions of the subsequent true survival can clearly be seen in Figure 6.3.1 where a large proportion of the survival estimates are underestimated when cohort-based analysis are used. These results are consistent with other evaluations, which for large data material

[3] Percentage of cases for which the diagnosis was based upon microscopic verification of a tissue specimen. This includes, in addition to histological confirmation of diagnosis, those based upon exfoliative cytology specimens, and diagnoses of leukaemia based on haematological examination (without examination of bone marrow).

otherwise have mostly been performed on data from the Finnish Cancer Registry[100-102], but also, for example, on data from the Canadian Cancer Registry[103].

In a seminal article from 1961, Ederer et al defined relative survival as the observed survival of the cancer patients divided by the expected survival of a comparable group from the general population, but free from the specific disease under study. However, they concluded "[that] since we are usually concerned with analysing survival of patients with specific forms of cancer, it appears that we do not need to make an adjustment in estimating expected survival from population life tables".[53] Since then, this statement has more or less been taken for granted.

Using the computerised population registers that exist in Sweden, I had the unique opportunity to calculate expected survival both including and excluding individuals diagnosed with cancer, and thereby estimate the size of the bias introduced into relative survival estimates by using survival probabilities from general population life tables. To my knowledge, a systematic evaluation of this bias has never been published before. For most applications, the statement made by Ederer et al fifty years ago seems nevertheless to be more or less correct. However, it can be questioned when the cancer prevalence is high in the background population and the excess mortality induced by the disease simultaneously is sufficiently large to make an impact on cancer patient survival. One also has to recognise that the situation is different today than it was in the early 1960's. As mentioned in the introduction, cancer has become a more common cause of death in the last decades, not least due to improved treatment for cardiovascular diseases and an aging population. As such, it may very well be that when the statement was made it was truer than it is today, especially since it is more or less accurate even today.

In the study carried out in paper VI[69], the requirement of a sufficiently high and simultaneous prevalence and mortality seem to be met for breast and prostate cancer and for all sites combined. Cancers with very high fatality (e.g., lung cancer) or with comparatively low mortality (e.g., skin cancer) do not meet these criteria since the prevalence and the mortality, respectively, are too low. For younger age groups even a large proportion of deaths due to cancer will only have a tiny impact on the expected survival since the general probability of death is so low in these groups. On the other hand, it is no disadvantage in doing the adjustment that I propose in the paper even when it is not strictly necessary since the unadjusted and adjusted survival probabilities in those cases will, for all practical purposes, be essentially the same. Since the proposed adjustment is easy to do, it seems reasonable that it is made whenever the combination of prevalence and mortality is believed to have a noticeable impact on the survival estimates. In many cases, it also seems sufficient to do the adjustment with cause of death statistics from a single (recent) year. This will in any case bring down the bias considerably. The only foreseeable issue is that unless it is done universally it will hamper comparisons between different studies.

In a sensitivity analysis on Finnish Cancer Registry data it was shown for common cancers and for older age groups that the proportion of deaths due to a specific cancer would need to reach at least two percent before important differences occurred between adjusted and unadjusted relative survival ratios. Specific cancers rarely reach such proportions even in the older age groups due to the large amount of competing causes of deaths.[104] In paper VI, it is shown that regardless of the magnitude of cancer-related deaths in the general population, there is no need to adjust the survival probabilities from the general population life tables if the patients are below 60 years of age.

8 CONCLUSIONS

Cancer patient survival in Sweden is generally increasing and in an international perspective Swedish cancer patients have a comparatively good survival. However, it is not enough to be satisfied with past achievements and the good work that previously have accomplished these advances should continue. The new organisational structure with the Regional Cancer Centres[5, 20] will hopefully take the past achievements even further.

The overall completeness and quality of the Swedish Cancer Register is high and comparable to cancer registers in the other Nordic countries. For most uses in epidemiology, public health surveillance, and planning the under-reporting that does exist is likely to be without significant implications. However, there seems to be some aspects of the reporting and registration process that ought to be reviewed, since the degree of under-reporting is site specific, increases with age, and does not otherwise appear to be random. I believe that monitoring of the completeness and quality should be as close as possible to the source of registration and be structured as a continuous on-going process at the regional cancer registries and the organisation in which they reside.

One area that needs special attention is the socio-economic and regional differences in cancer patient survival. The Swedish Health and Medical Service Act[105, section 2], emphasises good health and access to care on equal terms for the entire population, where priority for health and medical care shall be given to those in the greatest need of care. There are clearly some improvements that need to be made by the county councils, together with the new centres, in order to live up to these requirements, and to elevate the survival of the disadvantaged groups. As primarily responsible for health care in Sweden and according to the Health and Medical Service Act[105, section 3], the county councils shall offer good health and medical services to persons living within its boundaries. The National Board of Health Welfare's responsibility in this area is to compile, analyse and develop knowledge, set standards and guidelines, and to supervise health and medical services and their staff.[106]

A promising area for Sweden is to utilise its competitive advantages with the population-based registers. By linking data on patient characteristics, diagnosis, and treatment in the quality registers with the health data and population registers, there are good opportunities for research that could improve cancer treatments in the future.

Regarding the question of whether to use cohort- or period-based survival analysis my recommendation from paper III, eight years ago, still holds. I am more convinced today, than I was at that time, that period-based analysis is the right approach if the main purpose is to predict future survival or when attention is not primarily directed to a specific cohort of patients diagnosed during a particular time period. However, I see no good reason to use the period approach when analysing cohorts of patients that have already completed their required time of follow-up, or when the main purpose is to correlate survival estimates to prevention and treatment measures applied during a specific time period.

It is reassuring to know that the statement made by Ederer et al fifty years ago "[that] it appears that we do not need to make an adjustment in estimating expected survival from population life tables", by and large, appears to be to be correct, and thus previously published results are reliable. My proposed adjustment for this bias is easy to implement, and the cause of death statistics required should already be available to most cancer registries.

9 ACKNOWLEDGEMENTS

As I have been working on my doctoral studies on and off for many years I might forget to recognise someone who in the past was important for part of this work. Nevertheless, there are some who cannot be forgotten and definitely needs to be mentioned specially. To the rest of you, whom I have not mentioned by name or affiliation, please do not feel forgotten, you are included as well.

First, of cause, **Paul Dickman**, my main supervisor, who never lost faith in me despite my not always so tenacious approach to my doctoral studies. Of course, the fact that he has been working on his book for so many years might have kept him at bay. Your encouragement and support over the years made me eventually pull this off.

My co-supervisor **Måns Rosén**, my former boss at the National Board of Health and Welfare, who after many attempts eventually managed to persuade me to register as a doctoral student. However, I must confess that, there have been times over the years when I have not been so pleased with the decision I made.

Magnus Stenbeck, colleague at the National Board of Health and Welfare, who gave me my first introduction into relative survival analysis. Your help was instrumental when I started working on the analysis that became my first article in this area of research and the second paper of this thesis.

My **present** and **former colleagues** at the National Board of Health and Welfare, and the **faculty**, **staff**, and **fellow doctoral students** in the biostatistics group at the Department of Medical Epidemiology and Biostatistics. To my colleagues at the National Board, thank you for having had patience with me as I have not been there to do my part of the work this year. To the people in the biostatistics group, you always make me feel at home when I enter your corridor.

I extend special thanks to **Pinelopi**, **Gunilla**, **Therese**, **Gudrun**, **Maria**, and **Sandra** who lent me some of your valuable time by reading the final draft of my manuscript. Your comments sometimes forced me to go back and think of what I really meant with what I had just put in writing.

I would also like to take the opportunity to thank the **administrative staff** at the Department of Medical Epidemiology and Biostatistics for the work they perform. Your help makes life as a doctoral student so much easier.

10 REFERENCES

1. Dödsorsaker 2009 (In English: Causes of Death 2009) [art. nr. 2011-3-22]. National Board of Health and Welfare, Stockholm, Sweden, 2011.

2. Cancer Incidence in Sweden 2009 - Cancerförekomst i Sverige 2009 [art. nr. 2010-12-17]. National Board of Health and Welfare, Stockholm, Sweden, 2010.

3. Wigzell K, A National Cancer Strategy for the Future – Summary. Report of the Commission of Inquiry on A National Cancer Strategy [SOU 2009:11]. Swedish Government Inquiries, Government Offices of Sweden, Stockholm, Sweden, 2009.

4. Statistikdatabasen (In English: The Statistical database) [Internet database]. Statistics Sweden, 2007. [cited 1 June 2011]. Available from: http://www.ssd.scb.se/databaser/makro/start.asp.

5. Wigzell K, En nationell cancerstrategi för framtiden (In English: A National Cancer Strategy for the Future). Report of the Commission of Inquiry on A National Cancer Strategy [SOU 2009:11]. Swedish Government Inquiries, Government Offices of Sweden, Stockholm, Sweden, 2009.

6. Sullivan R, Peppercorn J, Sikora K, Zalcberg J, Meropol NJ, Amir E, Khayat D, Boyle P, Autier P, Tannock IF, Fojo T, Siderov J, et al. Delivering affordable cancer care in high-income countries. Lancet Oncol 2011;12:933-80.

7. Ezzati M, Vander Hoorn S, Lopez AD, Danaei G, Rodgers A, Mathers CD, Murray CJL. Comparative Quantification of Mortality and Burden of Disease Attributable to Selected Risk Factors. In: Lopez AD, Mathers CD, Ezzati M, Jamison DT, Murray CJL, eds. Global Burden of Disease and Risk Factors. The International Bank for Reconstruction and Development, Washington (DC), USA, 2006.

8. Cancerfondsrapporten 2011 (In English: Swedish Cancer Society - Annual report 2011). Swedish Cancer Society, Stockholm, Sweden, 2011.

9. Allmän barnvaccination mot HPV 16 och 18 i syfte att förebygga livmoderhals-cancer (In English: Public childhood vaccination against HPV 16 and 18 to prevent cervical cancer) [SBU Alert-rapport 2008-01]. The Swedish Council on Technology Assessment in Health Care, Stockholm, Sweden, 2008.

10. The Swedish vaccination program - Vaccination against HPV for girls. National Board of Health and Welfare, Stockholm, Sweden, 2010.

11. Hedlund M. Upphandling av vaccin måste göras om (In English: Procurement of vaccine must be redone) [Internet]. Dagens Nyheter, Stockholm, Sweden, 2011. [published 23 March 2011; cited 26 September 2011]. Available from: http://www.dn.se/sthlm/upphandling-av-vaccin-maste-goras-om.

12. Clive RE. Introduction to cancer registration. In: Hutchison CL, Menck HR, Burch M, Gottschalk R. Cancer Registry Management - Principles and Practice, 2nd edition. Kendall Hunt Publishing Company, Dubuque, USA, 2004.

13. Potter JD. Epidemiology informing clinical practice: from bills of mortality to population laboratories. Nat Clin Pract Oncol 2005;2:625-34.

14. Clemmesen J. Statistical Studies in the Aetiology of Malignant Neoplasms. Acta Pathologica et Microbiologica Scandinavica 1965; Supplementum 174:I

15. Wagner G. Cancer Registration: Historical Aspects. In: Parkin DM, Wagner G, Muir C. The Role of the Registry in Cancer Controled [volume 66]. The International Agency for Research on Cancer, Lyon, France, 1985.

16. Wagner G. History of cancer registration. In: Jensen OM, Parkin, D.M., MacLennan, R., Mui, C.S., Skeet, R.G. Cancer Registration: Principles and Methods [volume 95]. International Agency for Research on Cancer, Lyon, France, 1991.

17. Doll R, Payne P, Waterhouse J, Editors. Cancer Incidence in Five Continents. International Union Against Cancer, Geneva, Switzerland, 1966.

18. Curado MP, Edwards B, Shin HR, Storm H, Ferlay J, Heanue M, Boyle P. Cancer Incidence in Five Continents, Volume IX [volume 160]. International Agency for Research on Cancer, Lyon, France, 2007.

19. Parkin DM. The role of cancer registries in cancer control. Int J Clin Oncol 2008;13:102-11.

20. Asplund K. Kriterier som ska utmärka ett regionalt cancercentrum (RCC) (In English: Criteria that must distinguish a Regional Cancer Center (RCC)) [memorandum 31 January 2011]. Government Offices of Sweden, Ministry of Health and Social Affairs, Stockholm, Sweden, 2011.

21. Rosén M. Översyn av de nationella kvalitetsregistren. Guldgruvan i hälso och sjukvården, förslag till gemensam satsning 2011–2015 (In English: Inquiry of the national quality registers. The gold mine in health care. Proposals for a national investment 2011–2015). Government Offices of Sweden, Ministry of Health and Social Affairs, Stockholm, Sweden, 2010.

22. Boström M. Plan för fortsatt arbete i uppdraget utveckling av kvalitetsregistren (In English: Plan for further work in the development of the quality registries) [memorandum 8 May 2011]. Government Offices of Sweden, Ministry of Health and Social Affairs, and Swedish Association of Local Authorities and Regions, 2011.

23. International Statistical Classification of Diseases and Related Health Problems, Tenth Revision (ICD-10). World Health Organization, Geneva, Switzerland 2009.

24. Berrino F, Verdecchia A, Lutz JM, Lombardo C, Micheli A, Capocaccia R, Group EW. Comparative cancer survival information in Europe. Eur J Cancer 2009;45:901-8.

25. Woods LM, Rachet B, Coleman MP. Origins of socio-economic inequalities in cancer survival: a review. Ann Oncol 2006;17:5-19.

26. Berrino F, Micheli A, Sant M, Capocaccia R. Interpreting survival differences and trends. Tumori 1997;83:9-16.

27. Enstrom JE, Austin DF. Interpreting cancer survival rates. Science 1977;195:847-51.

28. Welch HG, Schwartz LM, Woloshin S. Are increasing 5-year survival rates evidence of success against cancer? JAMA 2000;283:2975-8.

29. Bailar JC, Gornik HL. Cancer undefeated. N Engl J Med 1997;336:1569-74.

30. Beral V, Peto R. UK cancer survival statistics. BMJ 2010;341:c4112.

31. International Statistical Classification of Diseases, Injuries, and Causes of Death. Sixth Revision (ICD-6). World Health Organization, Geneva, Switzerland, 1948.

32. Personal communication with Lars-Age Johansson on cause of death registration in general and for Sweden in particular. Johansson LA, Unit for Healthcare and Social Welfare Statistics, Department of Statistics, Monitoring and Evaluation, National Board of Health and Welfare, Stockholm, Sweden, 2011.

33. Folkhälsorapport 2009 (In English: Public Health Report 2009) [art. nr. 2009-126-71]. National Board of Health and Welfare, Stockholm, Sweden, 2009.

34. Folkbokföring (In English: Population registration) [Internet]. Swedish Tax Agency, Solna, Sweden, 2011. [cited 15 June 2011]. Available from: http://www.skatteverket.se/privat/folkbokforing.4.18e1b10334ebe8bc800039.html

35. Folkbokföringen igår och idag (In English: The population records yesterday and today) [Internet]. Swedish Tax Agency, Solna, Sweden, 2011. [cited 15 June 2011]. Available from: http://www.skatteverket.se/privat/folkbokforing/omfolkbokforing/folkbokforingigaridag.4.76a43be412206334b89800023140.html

36. Johansson L-A, Dödsorsaksstatistik - Historik, produktionsmetoder och tillförlitlighet (In English: Cause of death statistics - History, production and reliability) [art. nr. 2010-4-33]. National Board of Health and Welfare, Stockholm, Sweden, 2010.

37. Lag (2001:99) om den officiella statistiken (In English: The Official Statistics Act). Government Offices of Sweden, Ministry of Finance, Stockholm, Sweden, 2001.

38. Förordning (2001:100) om den officiella statistiken (In English: Official Statistics Ordinance). Government Offices of Sweden, Ministry of Finance, Stockholm, Sweden, 2001.

39. Sveriges officiella statistic (In English: Official Statistics of Sweden) [Internet]. Statistics Sweden, Stockholm, Sweden, 2011. [cited 10 September 2011]. Available from: http://www.scb.se/Pages/List____139368.aspx.

40. Lag (1998:543) om hälsodataregister (In English: The Health Data Registers Act). Government Offices of Sweden, Ministry of Health and Social Affairs, Stockholm, Sweden, 1998.

41. Förordning (2001:709) om cancerregister hos Socialstyrelsen. (In English: Ordinance for cancer register at the National Board of Health and Welfare). Government Offices of Sweden, Ministry of Health and Social Affairs, Stockholm, Sweden, 2001.

42. Ändring i föreskrifterna och allmänna råden (SOSFS 2003:13) om uppgiftsskyldighet till cancerregistret vid Socialstyrelsen. SOSFS 2006:15 (M). National Board of Health and Welfare, Stockholm, Sweden, 2006.

43. Planering av onkologisk sjukvård. Råd och anvisningar från Socialstyrelsen (In English: Planning of oncological care. Recommendations and guidelines from the National Board of Health and Welfare) [art. nr 32]. National Board of Health and Welfare, Stockholm, Sweden, 1974.

44. Möller TR. Cancer Care Programmes: The Swedish Experience. In: Parkin DM, Wagner G, Muir C. The Role of the Registry in Cancer Control [volume 66]. The International Agency for Research on Cancer, Lyon, France, 1985.

45. Mattsson B, Wallgren A. Completeness of the Swedish Cancer Register. Non-notified cancer cases recorded on death certificates in 1978. Acta Radiol Oncol 1984;23:305-13.

46. Barlow L, Westergren K, Holmberg L, Talbäck M. The completeness of the Swedish Cancer Register: a sample survey for year 1998. Acta Oncol 2009;48:27-33.

47. This information is partly based on personal experience and knowledge as I have been working at the National Board of Health and Welfare as the opposite party to the regional cancer registries on data deliveries between 1998 and 2006, and partly on personal communication with three of the six regional cancer registries that had the courtesy to answer my emails about what they do to ensure completeness and quality of their cancer register.

48. Dödsorsaker 2004 (In English: Causes of Death 2004) [art. nr. 2007-42-1]. National Board of Health and Welfare, Stockholm, Sweden, 2007.

49. Johansson LA, Björkenstam C, Westerling R. Unexplained differences between hospital and mortality data indicated mistakes in death certification: an investigation of 1,094 deaths in Sweden during 1995. J Clin Epidemiol 2009;62:1202-9.

50. Karlsson T. Befolkningsstatistik 2009 (In English: Population Statistics 2009) [art. nr. BE0101]. Population and Welfare Department, Statistics Sweden, Örebro, Sweden, 2010.

51. Edler B, Johansson N. Kartläggning av folkbokföringsfelet (In English: Identifying the error in the population registration) [art. nr. 2006:7]. Swedish Tax Agency, Solna, Sweden, 2006.

52. Övertäckningen i Registret över totalbefolkningen (In English: The over-coverage in the Total Population Register) [art. nr. BE96BR1005]. Population and Welfare Department, Statistics Sweden, Örebro, Sweden, 2010.

53. Ederer F, Axtell LM, Cutler SJ. The Relative Survival Rate: A Statistical Methodology. National Cancer Institute Monographs 1961;6:101-21.

54. Björkenstam C, Edberg A, Ayoubi S, Rosén M. Are cancer patients at higher suicide risk than the general population? Scand J Public Health 2005;33:208-14.

55. Estève J, Benhamou E, Raymond L. Descriptive Epidemiology, Statistical Methods in Cancer Research [volume IV]. International Agency for Research on Cancer, Lyon, France, 1994.

56. Howlader N, Ries LA, Mariotto AB, Reichman ME, Ruhl J, Cronin KA. Improved estimates of cancer-specific survival rates from population-based data. J Natl Cancer Inst 2010;102:1584-98.

57. Dödsorsaker 2010 – Causes of Death 2010 [art. nr. 2011-7-6]. National Board of Health and Welfare, Stockholm, Sweden, 2011.

58. Ederer F, Heise H. Instructions to IBM 650 programmers in processing survival computations. Methodological note No. 10. End Results Evaluation Section, National Cancer Institute, Bethesda (MD), USA, 1959.

59. Hakulinen T. Cancer survival corrected for heterogeneity in patient withdrawal. Biometrics 1982;38:933-42.

60. Perme MP, Stare J, Estève J. On Estimation in Relative Survival. Biometrics 2011; doi:10.1111/j.1541-0420.2011.01640.x.

61. Hakulinen T, Seppä K, Lambert PC. Choosing the relative survival method for cancer survival estimation. Eur J Cancer 2011; doi:10.1016/j.ejca.2011.03.011.

62. Global Health Observatory [Internet]. World Health Organization, Geneva, Switzerland, 2011. [cited 15 June 2011]. Available from: http://www.who.int/gho/en/

63. Brenner H, Gefeller O. An alternative approach to monitoring cancer patient survival. Cancer 1996;78:2004-10.

64. Braine T. Race against time to develop new antibiotics. Bulletin of the World Health Organization [volume 89]. World Health Organization, Geneva, Switzerland, 2011.

65. Talbäck M, Stenbeck M, Rosén M. Up-to-date long-term survival of cancer patients: an evaluation of period analysis on Swedish Cancer Registry data. Eur J Cancer 2004;40:1361-72.

66. Brenner H, Gefeller O, Hakulinen T. Period analysis for 'up-to-date' cancer survival data: theory, empirical evaluation, computational realisation and applications. Eur J Cancer 2004;40:326-35.

67. Brenner H, Hakulinen T. Up-to-date cancer survival: period analysis and beyond. Int J Cancer 2009;124:1384-90.

68. Oksanen H. Modelling the survival of prostate cancer patients [PhD thesis]. Tampere school of public health, University of Tampere, Finland, 1998.

69. Talbäck M, Dickman PW. Estimating expected survival probabilities for relative survival analysis – Exploring the impact of including cancer patient mortality in the calculations. Eur J Cancer 2011; doi:10.1016/j.ejca.2011.08.010.

70. Stenbeck M, Rosén M, Holm LE. Cancer survival in Sweden during three decades, 1961-1991. Acta Oncol 1995;34:881-91.

71. Talbäck M, Rosén M, Stenbeck M, Dickman PW. Cancer patient survival in Sweden at the beginning of the third millennium - predictions using period analysis. Cancer Causes Control 2004;15:967-76.

72. Talbäck M, Stenbeck M, Rosén M, Barlow L, Glimelius B. Cancer survival in Sweden 1960-1998 - developments across four decades. Acta Oncol 2003;42:637-59.

73. Berrino F, De Angelis R, Sant M, Rosso S, Bielska-Lasota M, Lasota M, Coebergh J, Santaquilani M. Survival for eight major cancers and all cancers combined for European adults diagnosed in 1995-99: results of the EUROCARE-4 study. Lancet Oncol 2007;8:773-83.

74. Verdecchia A, Francisci S, Brenner H, Gatta G, Micheli A, Mangone L, Kunkler I. Recent cancer survival in Europe: a 2000-02 period analysis of EUROCARE-4 data. Lancet Oncol 2007;8:784-96.

75. Coleman MP, Forman D, Bryant H, Butler J, Rachet B, Maringe C, Nur U, Tracey E, Coory M, Hatcher J, McGahan CE, Turner D, et al. Cancer survival in Australia, Canada, Denmark, Norway, Sweden, and the UK, 1995-2007 (the International Cancer Benchmarking Partnership): an analysis of population-based cancer registry data. Lancet 2011;377:127-38.

76. Coleman M, Quaresma M, Berrino F, Lutz J, De Angelis R, Capocaccia R, Baili P, Rachet B, Gatta G, Hakulinen T, Micheli A, Sant M, et al. Cancer survival in five continents: a worldwide population-based study (CONCORD). Lancet Oncol 2008;9:730-56.

77. Cancer i Sverige - Insjuknande och överlevnad utifrån regional och socioekonomisk indelning (In English: Cancer in Sweden - Morbidity and survival based on regional and socio-economic subdivision) [art. nr. 2011-3-26]. National Board of Health and Welfare, Stockholm, Sweden, 2011.

78. Vågerö D, Persson G. Cancer survival and social class in Sweden. J Epidemiol Community Health 1987;41:204-9.

79. Sant M, Allemani C, Capocaccia R, Hakulinen T, Aareleid T, Coebergh JW, Coleman MP, Grosclaude P, Martinez C, Bell J, Youngson J, Berrino F, et al. Stage at diagnosis is a key explanation of differences in breast cancer survival across Europe. Int J Cancer 2003;106:416-22.

80. Feinstein AR, Sosin DM, Wells CK. The Will Rogers phenomenon. Stage migration and new diagnostic techniques as a source of misleading statistics for survival in cancer. N Engl J Med 1985;312:1604-8.

81. Larsson S. Completeness and reliability of lung cancer registration in the Swedish Cancer Registry. Acta Pathol Microbiol Scand A 1971;79:389-98.

82. Flam F, Rutqvist LE. Under-registration of gestational trophoblastic disease in the Swedish Cancer Registry. Eur J Epidemiol 1992;8:683-6.

83. Helgesson O, Bengtsson C, Lapidus L, Merck C, Sparén P. Malignant disease observed in a cohort of women. A validation of Swedish Cancer Registry data. Scand J Soc Med 1994;22:46-9.

84. Garne JP, Aspegren K, Möller T. Validity of breast cancer registration from one hospital into the Swedish National Cancer Registry 1971-1991. Acta Oncol 1995;34:153-6.

85. Frödin JE, Ericsson J, Barlow L. Multiple primary malignant tumors in a national cancer registry - reliability of reporting. Acta Oncol 1997;36:465-9.

86. Ekström AM, Signorello LB, Hansson LE, Bergström R, Lindgren A, Nyrén O. Evaluating gastric cancer misclassification: a potential explanation for the rise in cardia cancer incidence. J Natl Cancer Inst 1999;91:786-90.

87. Aström M, Bodin L, Tidefelt U. Adjustment of incidence rates after an estimate of completeness and accuracy in registration of acute leukemias in a Swedish population. Leuk Lymphoma 2001;41:559-70.

88. Lindblad M, Ye W, Lindgren A, Lagergren J. Disparities in the classification of esophageal and cardia adenocarcinomas and their influence on reported incidence rates. Ann Surg 2006;243:479-85.

89. Turesson I, Linet MS, Björkholm M, Kristinsson SY, Goldin LR, Caporaso NE, Landgren O. Ascertainment and diagnostic accuracy for hematopoietic lymphoproliferative malignancies in Sweden 1964-2003. Int J Cancer 2007;121:2260-6.

90. Holmäng S, Amsler-Nordin S, Carlson K, Holmberg E, Johansson SL. Completeness and correctness of registration of renal pelvic and ureteral cancer in the Swedish Cancer Registry. Scand J Urol Nephrol 2008;42:12-7.

91. Måsbäck A, Andersson G, Olsson H. Problems in assessing multiple cutaneous melanoma. A review on the accuracy of a population based cancer registry. Cancer Epidemiol 2010;34:262-6.

92. Lambe M, Eloranta S, Wigertz A, Blomqvist P. Pancreatic cancer; reporting and long-term survival in Sweden. Acta Oncol 2011; doi:10.3109/0284186X.2011.599338.

93. Heuman S. Patientdata och läkemedel m.m. (In English: Patient information and medicines etc.) [art. nr. 2007:48]. Government Offices of Sweden, Ministry of Health and Social Affairs, Stockholm, Sweden, 2007.

94. Lag om ändring i sekretesslagen (1980:100) SFS 2008:356 (In English: Act to amend of the Official Secrecy Act). Government Offices of Sweden, Ministry of Justice, Stockholm, Sweden, 2008.

95. Offentlighets- och sekretesslag (2009:400), (In English: Public Access to Information and Secrecy Act). Government Offices of Sweden, Ministry of Justice, Stockholm, Sweden, 2009.

96. Skeet RG. Quality and quality control. In: Jensen OM, Parkin, D.M., MacLennan, R., Mui, C.S., Skeet, R.G. Cancer Registration: Principles and Methods [volume 95]. International Agency for Research on Cancer, Lyon, France, 1991.

97. Tillräcklig kvalitet och kriterier för officiell statistik. Riktlinjer från Rådet för den officiella statistiken. (In English: Adequate quality and criteria for official statistics. Guidelines from the Council for Official Statistics.) [art. nr X42OP0601]. The Council for Official Statistics, Statistics Sweden, Örebro, Sweden, 2006.

98. Parkin DM, Bray F. Evaluation of data quality in the cancer registry: principles and methods Part II. Completeness. Eur J Cancer 2009;45:756-64.

99. Talbäck M, Dickman PW. Predicting the survival of cancer patients recently diagnosed in Sweden and an evaluation of predictions published in 2004. Acta Oncol 2011; doi:10.3109/0284186X.2011.626444.

100. Brenner H, Hakulinen T. Up-to-date long-term survival curves of patients with cancer by period analysis. J Clin Oncol 2002;20:826-32.

101. Brenner H, Söderman B, Hakulinen T. Use of period analysis for providing more up-to-date estimates of long-term survival rates: empirical evaluation among 370,000 cancer patients in Finland. Int J Epidemiol 2002;31:456-62.

102. Brenner H, Hakulinen T. Advanced detection of time trends in long-term cancer patient survival: experience from 50 years of cancer registration in Finland. Am J Epidemiol 2002;156:566-77.

103. Ellison LF. An empirical evaluation of period survival analysis using data from the Canadian Cancer Registry. Ann Epidemiol 2006;16:191-6.

104. Hinchliffe SR, Dickman PW, Lambert PC. Adjusting for the proportion of cancer deaths in the general population when using relative survival: A sensitivity analysis. Cancer Epidemiol 2011; doi:10.1016/j.canep.2011.09.007.

105. The Health and Medical Service Act (1982:763, with amendments up to and including SFS 2002:163). Government Offices of Sweden, Ministry of Health and Social Affairs, Stockholm, Sweden, 2003.

106. Detta är Socialstyrelsen (In English: This is The National Board of Health and Welfare) [art. nr. 2011-5-6]. National Board of Health and Welfare, Stockholm, Sweden, 2011.