Master's Thesis in Economics
Stockholm School of Economics
Fall 2007

Does cancer drug vintage affect cancer survival?

MATTIAS BERNOW

19040@student.hhs.se

Abstract

The effects of modern medicine on longevity is presently debated in the field of Health Economics. Until recently the consensus among health economists and epidemiologists has been that interventions by modern medicine have had an almost negligible effect on survival. During the last few years more and more material has been published indicating that there actually is a dependence between disease survival and intervention of modern medicine. This thesis investigates whether mean cancer drug vintage, i.e. the innovative degree of the utilized drugs, affects disease outcome in terms of cancer survival. Using a more advanced model and higher quality data than other similar studies, this thesis finds that increased mean cancer drug vintage positively affect cancer survival. This finding is in line with the most recent flank of previous research in the field.

Tutor: Magnus Johannesson
Examiner: Niklas Zethraeus
Discussants: Emma Hemmingson and Maria Lundén
Presentation: December 14, 2007

Acknowledgments:
This thesis is the result of a collaboration of efforts by many people. To start, I would like to thank my tutor, Magnus Johannesson, for support and inspiration. I would like to thank Dr Nils Wilking for medical advice, invaluable contacts and a never failing enthusiasm. I am indebted to Dr Hermann Brenner, Adam Gondos and Alex Penfold for generously sharing input data. I would like to extend my appreciation to the entire staff at i3 - European Health Economics for a fantastic summer. Finally, I want to express my most humble gratitude to my brother, Mårten Bernow, for invaluable support all throughout the work of this thesis.

1 Introduction

Cancer disease, annually accounting for 7 million deaths worldwide, is a major cause of premature death second only to cardiovascular diseases (Jönsson and Wilking 2007, piii8). As research expands the treatment options for cancer, the costs of new treatments tend to rise. Up-to-date technology and drugs are in general more costly than older ones, for comparison see FASS (2007). As health care resources are scarce, optimally allocating recourses is of great importance. In order to realize that ambition, extensive economic evaluation is needed. An important part of the evaluation is to examine if new treatments are more effective than older ones as this gives policymakers direction whether to go for the new, more expensive technology and drugs or not. This part of the resource allocation problem is exceedingly interesting in the ever expanding field of new drugs.

From a greater perspective it can be questioned to what extent research and development in medicine affect health and longevity. It is common knowledge that the major increases in public health during the last centuries are due to improvements in public health, environment, and nutrition (Folland et al 2001). Until recent years, the consensus among health economist was that the contribution by modern medicine to global longevity has been quite modest (Lichtenberg 2005a). Several textbooks on health economics support this view; see Santerre and Neun (2000), Henderson (1999) and Folland et al (2001). Recent research, though, has shown that this consensus might not be entirely correct. Cutler and McClellan (2001) concludes, in their case studies of five conditions (heart attack, low birth weight infants, depression, cataracts and breast cancer), that technological innovations in medicine have had important positive impacts on health.

Evidence that mean cancer drug vintage, i.e. the innovative degree of the utilized drugs, has a positive effect on cancer survival was first presented in "*A pan-European comparison regarding patient access to cancer drugs*" (Wilking and Jönsson 2005), a 90 page report from Karolinska Institute. Almost two years later an updated and improved version "*A global comparison regarding patient access to cancer drugs*" (Jönsson and Wilking 2007) was published in Annals of Oncology. The findings of the reports have been challenged by Coleman (2006 and 2007). Coleman argues among other things that the model used is too simplistic in two matters: 1) it applies to blunt estimates for survival and drug use, 2) the

model does not extend the analysis to include other medical technology advances[1] such as improvements in surgery or screening.

The purpose of this thesis is to explore if cancer drug vintage has an effect on cancer survival. The hypothesis is that mean cancer drug vintage has a positive effect on cancer survival. This thesis adds to the previous work by (Jönsson and Wilking 2007) and alleviates Coleman's first point of critique by using actual survival rates and drug usage instead of proxies for the survival and drug usage variables. Further, the aspect of technological advances within medicine, i.e. Coleman's second point of critique, will be mitigated by giving the used model flexibility to take into account also these type of effects[2]. The thesis will be based on available, most recent, survival data (1996-2004) matched with actual drug sales for six cancer sites and nine countries.

This introduction is followed by a background on the research field associated to this thesis, the thesis' contribution and its hypothesis, in section 2. The methodology, where the utilized variables are defined and the regression models are introduced is presented in section 3, while section 4 describes the data set and comments on the data gathering procedure. Section 5 contains both a general presentation and a discussion of the results. Finally, in section 6, conclusions are presented and followed by suggestions for further research. For the interested layman there is also a brief overview of the field of cancer disease and treatment to be found in appendix A.

[1] Please refer to section 3.4.2 for more information.
[2] Please refer to section 3.4.2 for more information.

2 Research field background, contribution and hypothesis

This section starts of with a brief introduction to health economics. Next, an overview of the research field of economic efficiency of new health care technology spending is presented. The following section zooms in on a separate part of the above research field, concerning drug vintage and survival and provides a starting point for the investigation of this thesis. This is followed by how this thesis contributes to the previous research and finally the hypothesis is presented.

2.1 Health economics

Healthcare spending accounts for a substantial part of the gross domestic product in all developed countries. During 1970-2005 healthcare expenditures in Sweden increased from 6.7 percent of GDP to 9.1 percent (OECD 2007). The increasing costs are not a local but a global phenomena (OECD 2007). Increasing healthcare costs and finite health care resources necessitated the development of tools for economic evaluation and so the field of health economics was developed (Kobelt 2002). In the textbook *Health Economics: An Introduction to Health Economics* health economics is defined as "*the application of the theories, tools and concepts of economics as a discipline to the topics of health and health care*" (Kobelt 2002, p. 10). In short, the field deals with issues relating to the allocation of scarce resources in order to improve health. In large parts of the developed world the health care sector is financed and provided for in such ways that the price mechanism is largely set aside. Not being able to rely on the price mechanism, economic evaluation, i. e. health economics, has come to play a vital role as decision makers and policy setters decide how to allocate resources.

2.2 Background on the economic efficiency of new health care technology spending

That new technology comes at a higher price than older technology or at least more spending is supported by Cutler and McClellan (2001) as well as the fact that patents keep prices of new technology high until they expire.

Until recent years health economists have questioned the role of modern medicine in increasing global longevity and consequently to economic growth. The quotations below also published by Lichtenberg (2005c) are taken from four textbooks and show examples of these attitudes toward the contribution of modern medicine:

"*the empirical evidence indicates [that] the overall contribution of medical care to health is rather modest at the margin. . . education, lifestyle, the environment, and income [are] the major contributing factors*" (Santerre and Neun 2000, p. 69).

"*increase in life expectancy [has] been much more influenced by economic development than improvements in medical care /.../ the most important medical advances are being brought about by improvements in information technology, not pills and scalpels*" (Getzen 1997, p.330).

"*Research on the relationship between health status and medical care frequently has found that the marginal contribution of medical care to health status is rather small /.../ any significant improvements in health status are more likely to originate from factors other than medical care /.../ Factors that determine the level of health include income and education, environmental and life-style factors, and genetics*" (Henderson 1999, p. 142).

"The historical declines in population mortality rates were not due to medical interventions because effective medical interventions became available to populations largely after the mortality had declined. Instead, public health, improved environment, and improved nutrition probably played substantial roles" (Folland et al 2001, p. 118).

As true as it might be that improvements in public health, environment, nutrition, education and lifestyle can explain most of the increased longevity for the last 100 years, that still does not exclude the possibility that the effect of modern medicine on longevity is positive. For a long time, it was questioned whether the positive effects of increased survival in terms of production and economic growth outweighed the costs of developing and producing the health care provided. However, in 2001 Cutler and McClellan presented evidence to the contrary. In their article (Cutler and McClellan 2001) on case studies of five conditions, namely heart attack, low birth weight infants, depression, cataracts and breast cancer, they find that technological innovations in medicine have had important positive economic impacts on health, i.e. that the economic benefits of survival have outweighed the costs of developing and providing the treatment. In four of the five conditions technological innovations in medicine are on net positive. For the fifth (breast cancer) there is no clear result. They find that "*technology often leads to more spending, but outcomes improve by even more*" (p. 23) and that "*medical spending as a whole is worth the increased cost of care*"(p. 11) which, if correct, has great implications for public policy makers. Looking at the increase in US healthcare costs from 1980 to 2000, Luce et al (2006) find that every extra dollar spent on overall healthcare services generated health gains valued at $1.15 to $1.94. In line with these findings Lichtenberg (2005a), having researched 47 major chronic conditions, concludes that the gains from new drugs in terms of increased production due to increased ability to work exceeds the costs by 150 percent.

2.3 Background on cancer drug vintage and cancer survival

The research presented in section 2.2 conclude that it is economically viable to develop new drugs and treatments for some conditions or at least that it has been so in the past. A prerequisite thereof is that new drugs results in better health or increased survival than already existing drugs. Whether this is the case is debated in a separate part of the research field, presented below.

Looking at the relationship between introduction of new drugs and increased longevity Lichtenberg (2005b-c & 2006) finds evidence that the introduction and use of new drugs have improved people's health and increased longevity. The first work on trying to estimate the effect of new cancer drugs on cancer survival was published in "A pan-European comparison regarding patient access to cancer drugs" (Wilking and Jönsson 2005). In a comment to the report (pp. 86-90), Frank Lichtenberg performs two econometric studies where he examines the relationship between access to cancer drugs and cancer survival. The first study examines the contribution of the introduction of new cancer drugs to cancer survival in the USA during 1975-1995. The second study looks at the effect of access to cancer drugs to survival in several countries during 2000. Both studies are based on availability of cancer drugs, i.e. number of registered cancer drugs for different indications. The results indicate that access to new cancer drugs affect cancer survival positively. The findings of the report was later challenged by Coleman (2006). He argues that the use of "available cancer drugs" as a proxy for "utilized cancer drugs" is not scientifically correct. Whether this is true can be debated, but it is commonly known that market uptake for new drugs are sometimes slow, meaning that availability and usage do not necessarily have to correspond.

Two years later "*A global comparison regarding patient access to cancer drugs*" (Jönsson and Wilking 2007) was published. In the report, the last article "*The effect of drug vintage on cancer survival and mortality*" examines the effect of cancer drug innovation on cancer survival. The hypothesis is that mean cancer drug vintage has a positive effect on cancer survival rates, defining drug vintage as the first year in which the drug was launched somewhere in the world. In this report, the authors conduct the studies based on actual usage of cancer drugs instead of cancer drug availability. Due to lack of three-dimensional data, i.e. data points referable to country, year and cancer site at the same time, the analysis of the effect of cancer drug vintage on cancer survival is carried out in three two-dimensional regression models, i.e. by cancer site and year but not by country, by cancer site and country

but not by year, and finally by country and year but not by cancer site.

The first model regresses survival and drug vintage by primary cancer site and year for the USA. This model uses data on several cancer sites during the period 1992-2003. The survival data is drawn from SEER 9 registries[3] and vintage data from the MEDSTAT Marketscan database[4].

The second model analyzes survival and drug vintage by primary cancer site and country for patients diagnosed 1990-1994. Data is collected for several cancer sites and five European countries. Survival data is taken from the GLOBOCAN 2002 database[5], which provides data on incidence[6] and prevalence[7]. Survival rates are then estimated by dividing prevalence by incidence. Drug utilization is derived from IMS Oncology Analyzer[8]. Unfortunately drug utilization was not available for 1990-1994 so the authors hypothesize that the drug use for 1990-1994 was serially correlated with that of 2002-2006.

The third model looks at mortality and drug vintage by country and year for all cancer sites combined. Mortality data is collected for 20 countries and the years 1995-2003. In this model the authors use mortality data instead of survival data, which they believe to be superior, due to the fact that survival rates were not available for the time period for which they had data on drug utilization.

In all of the three analysis's a positive relation between increases in mean drug vintage and increase in survival is found. Nevertheless, once again, the results were contested by Coleman (2007). This time Coleman's main critique was the way in which the survival estimates and drug utilization of the second model were calculated. He argues that dividing prevalence by incidence is a too unsophisticated estimate of survival rates and backs his argument up by pointing at differences between the estimated survival rates and previously published ditto. As for the assumption that the drug use of 1990-1994 is serially correlated with that of 2002-2006, Coleman dismisses it due to large shifts in the drug consumption that took place in between the periods.

---

[3] SEER 9 registries include Atlanta, Connecticut, Detroit, Hawaii, Iowa, New Mexico, San Francisco-Oakland and Seattle-Puget Sound, a combined population that in many ways equal the entire population of the USA.
[4] MEDSTAT contains information on prescriptions for hundreds of thousands of individuals in the USA.
[5] Based on data from WHO.
[6] For explanation, please refer to Appendix A, section A2.
[7] For explanation, please refer to Appendix A, section A2.
[8] Based on 70 000 patients compiled by more than 1 600 clinicians.

Even though Coleman (2005 & 2007) question the findings of Wilking and Jönsson (2005) and Jönsson and Wilking (2007), no macro-level empirical evidence (with aggregated data for several cancer sites, countries and years) contradicting Jönsson and Wilking has, as of today, been presented. However, on the micro level, i.e. clinical trials comparing newer drugs to older ones, some studies has been published reporting no increase in survival due to newer cancer drugs, for example see von der Maase et al (2005).

2.4 Contribution of this study

This thesis will take the second model of "*A global comparison regarding patient access to cancer drugs*" (Jönsson and Wilking 2007, pp. iii72-74), as starting point. The contribution of this thesis is that it will use actual survival rates and drug use instead of proxies. Further, the aspect of technological advances within medicine will be taken into account by extending the model from two to three dimensions, i.e. from cancer site and country to cancer site, country and year. Adding a dummy variable for the dimension year to the model, will allow us to correct for the yearly increase in survival due to technological advances in medicine, i.e. the yearly changes in survival not attributable to changes in drug vintage. This is possible as the collected data sample can be sorted in three dimension: by cancer site, country and year respectively. It will be the first time this kind of analysis is carried out and is believed by Jönsson and Wilking (2007) to a superior way to test the hypothesis. The analysis will be based on available, most recent, survival data (1996-2004) matched with actual drug sales for six cancer sites and nine countries.

2.5 Hypothesis

In this thesis the research of Jönsson and Wilking (2007), who concluded that mean cancer drug vintage has positive effect on cancer survival, will be built upon with higher quality data and an improved model. To fulfill the purpose of this thesis, i.e. to investigate if cancer drug vintage has an effect on cancer survival, the following hypothesis will be examined:

Hypothesis: Mean cancer drug vintage has a positive effect on cancer survival.

3 Methodology

The hypothesis will be tested in two models, described in section 3.4. In the regression models the effect of cancer drug vintage, i.e. the innovative degree of the utilized drugs, on cancer survival is tested. The dependent variable, survival, is described in section 3.1. The explanatory variable, drug vintage, is defined and calculated in two different ways in accordance with Jönsson and Wilking (2007). This will be described in section 3.2. The difference between the two models lie in which control variables are used. The control variables for the two models are all dummy variables in different dimensions, namely country, cancer site and year. The control variables will be described in section 3.3.

The reason for using two models are:

1. One model enables us to compare the result, when using better quality data, with previous research.

2. Building a second model enables us to include one more control variable, i.e. to increase the model from two to three dimensions (from cancer site and country to cancer site, country and year). The new variable, i.e. a dummy variable for the dimension year, clears the survival data of increases due to other technology advances within medicine than that of drugs. This refinement adds to the theoretical consistency of the model and thus increases the validity of the result.

3.1 Dependent variable

It can be discussed what measure: survival or mortality is the better to quantify disease outcome as both measures are affected by incidence[9]. The risk of choosing survival is that an improvement in screening would lead to an increase in diagnosing both malign[10] and benign[11] tumors, which in turn may lead to over-diagnosing cancerous tumors, as is the case for PSA-testing for prostate cancer (Carroll 2005). Over-diagnosing results in increased incidence statistics. As the over-diagnosed (benign) tumors do not have a fatal outcome this generate a

[9] For explanation, please refer to Appendix A, section A2.
[10] A malign tumor is cancerous.
[11] A benign tumor is non-cancerous.

misleading increase in survival. Cancer mortality[12], on the other hand, is dependent on how many, in absolute numbers, that die from the disease. If there is a sudden drop in cancer mortality this does not necessarily mean that treatment has improved. Instead it could be an effect of decreased incidence due to improved prevention.
In sum, using mortality as outcome measure is risky as it does not take changes in incidence in to account. Using survival, incidence is taken into account but, except for actual changes, the measure is affected by developments in healthcare technology. As the model used in this thesis does account for technological changes from year to year, survival is the preferred outcome measure of cancer. This is also in line with previous research, see Jönsson and Wilking (2007).

When it comes to measuring outcome of a terminal disease clinicians often refer to the outcome in terms of survival rates, i.e. the percentage of the patient population that survives for a certain period of time from the time of diagnosis. Common measures of survival are 1-year, 5-year and 10-year survival. There are several methods to calculate survival rates. The survival rates in this thesis are calculated according to the most up to date method called model-based period analysis[13]. This method has so far only been developed for calculating 5-year relative survival. Due to this reason the 5-year relative survival will be used to measure cancer survival in this thesis.

When comparing survival rates it is important to remember that the population diagnosed with a disease may not correspond to the average population of the country, in terms of age and sex. Thus, the cancer patient population has to be standardized so that it in terms of survival can be compared to a healthy population of the same gender and age. The survival rates used in this thesis have been age adjusted according to the standard cancer patient population described by Corazziari et al (2004).

In sum, the dependent variable used to quantify disease outcome in this thesis is *5-year relative survival*, measuring the percentage of the patients diagnosed with cancer that survives at least five years from the time of diagnosis.

[12] For explanation, please refer to Appendix A, section A2.
[13] For explanation, please refer to Appendix A, section A2.

3.2 Explanatory variable

The explanatory variable used to measure the degree of innovation of utilized drugs is *mean cancer drug vintage*. The preferred way to measure drug vintage is as a weighted average year of initial launch of the drugs used to treat a certain disease.

Mean cancer drug vintage, $Vintage_{ijk}$, for each cancer site i, in country j, in year k is calculated:

$$Vintage_{ijk} = \Sigma_d [N_{dijk} * YEAR_d] / \Sigma_d N_{dijk}$$

Where

N_{dijk} = the sales of cancer drug d used to treat patients with cancer site i, in country j, in year k.

$YEAR_d$ = the year of initial launch of cancer drug d somewhere in the world.

This method results in values of mean cancer drug vintage, $Vintage_{ijk}$, taking the form of a year and a decimal value, for example 1995.6, which equals the mean year of initial launch of the drugs utilized to treat cancer item i, in country j, in year k.

An alternative way to estimate mean cancer drug vintage is to divide the utilized cancer drugs into old drugs and new drugs. This is done by choosing a certain year as limit such that drugs launched after this year are defined as new and drugs launched during or previous to this year are defined as old. In this form, mean cancer drug vintage is calculated:

$$Vintage_{ijk} = \Sigma_d [N_{dijk} * f(YEAR_d)] / \Sigma_d N_{dijk}$$

where

N_{dijk} = the sales of cancer drug d used to treat patients with cancer site i, in country, j in year k.

$f(YEAR_d)$ = 1 if $YEAR_d >$ limit year

= 0 if $YEAR_d \leq$ limit year

$YEAR_d$ = the year of initial launch of cancer drug d somewhere in the world.

By using f($YEAR_d$), instead of just the simple $YEAR_d$, mean cancer drug vintage, $Vintage_{ijk}$, becomes a fraction from 0 (none) to 1 (all) which equals the fraction of patients treated with drugs launched after the defined limit year. In order to make it easier to understand the outcome of the regressions when using this version of the mean cancer drug vintage variable, the fractions calculated are multiplied by 100. By doing this both the dependent and explanatory variable are expressed in percentage units.
It is important to acknowledge that calculated in this way, dividing the drugs into two groups, the vintage variable becomes a less precise instrument of measuring the drug vintage. Moreover, it will be more sensitive to small inconsistencies in data, which is observed in section 4.3.

In "*A global comparison regarding patient access to cancer drugs*" (Jönsson & Wilking 2007) the authors used both these methods of defining and estimating mean cancer drug vintage. They believed the previous method to be a better estimate, but had to turn to the latter as the result from using the former was not significant.

3.3 Control variables

In the models used to test the hypothesis three types of control variables exist, disease dummies, country dummies and year dummies.

3.3.1 Disease dummies

The collected data sample consists of six different cancer diseases. It is common knowledge that there are differences in survival for different cancer sites. An example from collected data is that in Finland, in 1996 the 5-year relative survival for breast cancer was 79.2 percent while the corresponding value for lung cancer was only 10.6 percent. To allow for these differences disease dummies, α_i, are introduced. The dummies get the value of 1 if the pair of survival and drug vintage values are referred to a certain disease, otherwise it takes the value of 0. Five disease dummies, i = 1 to 5, are introduced. Breast cancer, being the first disease in alphabetical order in this thesis, is used as base case.

3.3.2 Country dummies

The collected data sample also consists of nine countries. It also is common knowledge that there are differences in cancer survival for different countries. An example from collected data is that in Finland, in 1996 the 5-year relative survival for breast cancer was 79.2 percent

while the corresponding value for Slovenia was 67.7 percent. To allow for these differences country dummies, δ_j, are introduced. The dummies get the value of 1 if the pair of survival and drug vintage values are referred to a certain country, otherwise it takes the value of 0. Eight country dummies, j = 1 to 8, are introduced. Finland, being the first country in alphabetical order in the data sample, is used as base case.

3.3.3 Year dummies

Finally, the collected data is a time period 1996-2004, consisting of nine years. It is common knowledge that as healthcare technology develops there is a positive trend in survival from year to year. An example from collected data is that in Finland, in 1996 the 5-year relative survival for breast cancer was 79.2 percent while the corresponding value for 2004 had improved to 86.9 percent. To allow for these differences year dummies, φ_k, are introduced. The dummies get the value of 1 if the pair of survival and drug vintage values are referred to a certain year, otherwise it takes the value of 0. Eight year dummies, k = 1 to 8, are introduced. The year of 1996, being the first year in chronological order in the data sample, is used as base case.

3.4 Modelling

Two sets of models will be used to analyze the collected panel data. The first model, the one used in "*A global comparison regarding patient access to cancer drugs*" (Jönsson and Wilking 2007, pp.iii72-74) is included in order to be able to compare the results. The second model is in essence constructed like the first model but is extended to make use of the fact that the data sample is three-dimensional. As the sample is multi-dimensional, (two-dimensional in model 1 and three-dimensional in model 2) difference-in-difference models estimated with fixed effects can be used. This kind of estimation method allows for a much better control of potentially confounding factors than cross-sectional or time-series models would (Stock and Watson 2003, pp. 385-388 and 419-420). The difference between model 1 and model 2 is that they are carried out in two and three dimensions respectively.

3.4.1 Model 1 - two dimensions: cancer site and country

This model was used in "*A global comparison regarding patient access to cancer drugs*" (Jönsson and Wilking 2007, pp. iii72-74). It analyses the effect of mean cancer drug vintage on cancer survival by primary cancer site and by country. The regressions in model 1 are run

with fixed effects in order to be able to compare the results with the results found by Jönsson and Wilking (2007). However, in this thesis actual survival rates are used instead of proxies. The model has the following form:

$$Survival_{ij} = c + \beta Vintage_{ij} + \alpha_i + \delta_j + \varepsilon_{ij}$$

where:

$Survival_{ij}$ = the 5-year relative survival rate in percent for cancer type I, in country j.

$Vintage_{ij}$ = a measure of the vintage of drugs used to treat cancer site i, in country j.

c = a constant.

α_i = disease dummies 1 to 5, a fixed effect for cancer site i.

δ_j = country dummies 1 to 8, a fixed effect for country j.

ε_{ij} = a disturbance.

3.4.2 Model 2 - three dimensions: cancer site, country and year

This model is equal to model 1 but it does also take the dimension year into account. The year dimension does correct for the yearly improvements in technology, besides those explained by the drug vintage variable. For cancer survival, the most important factors being accounted for by the technology factor ought to be increased density and improvements of radiology equipment, new surgical techniques and increased cancer screening and prevention programs. The model takes the following form:

$$Survival_{ijk} = c + \beta Vintage_{ijk} + \alpha_i + \delta_j + \varphi_k + \varepsilon_{ijk}$$

where:

$Survival_{ijk}$ = the 5-year relative survival rate in percent for cancer type i, in country j, in year k.

$Vintage_{ijk}$ = a measure of the vintage of drugs used to treat cancer site i, in country j, in year k.

c = a constant

α_i = disease dummies 1 to 5, a fixed effect for cancer site i

δ_j = country dummies 1 to 8, a fixed effect for country j

φ_k = year dummies 1 to 8, a fixed effect for year k

ε_{ijk} = a disturbance

This is actually the model Jönsson and Wilking (2007) wanted to use but left for further research as the data needed for such a model was not available at the time. For the first time, the three dimensional data needed for this model has been collected for this thesis.

4 Sample and Data

4.1 Sample selection and data gathering

To be able to run the above models in order to test the hypothesis, data on survival and drug sales was required. A sample of six cancer sites in eleven countries for the period 1995-2004 was set out to be collected. The cancer sites included were colorectal, lung, breast, ovarian, non-Hodgkin lymphoma and leukaemia. The reason for choosing those sites is that they all have been viewed upon as breakthrough fields in oncology at some point during the last 40 years. The countries selected were Estonia, Finland, Germany, Italy, the Netherlands, Norway, Slovenia, Switzerland, UK, Poland and USA. The reason for choosing those countries was to include countries with dissimilar uptake of new drugs, a fact that might strengthen the analysis. The period 1995-2004 was chosen due to the fact that it is the time period closets to present day that it was possible to find data for[14]. Further, choosing a recent time period does make the results more interesting from a policy perspective as they better describe today's reality for patients, clinicians and researchers, than would data from 20 years ago. The selection of and the reasoning behind the data sample was made in collaboration with Dr Nils Wilking[15].

Survival data for each cancer site, country and year has been provided and processed by Dr Herman Brenner and Adam Gondos, both at the Division of Clinical Epidemiology and Aging Research at the German Cancer Research Centre in Heidelberg, Germany. Raw data for their computations was collected from national cancer registries[16], which are commonly used for health statistics.

What drugs to include in the calculation of the mean drug vintage variable was limited to drugs used for treating the six cancer sites in the sample. Key drugs, i.e. molecules, for each cancer site were identified by Dr Nils Wilking. All in all 21 molecules were selected.[17] Data on drug sales for those molecules was provided by IMS Health[18], a world leading company

[14] Please refer to section Appendix A, section A2, for information on how it is possible in year 2007 to calculate 5-year relative survival of patients diagnosed in 2004.
[15] Docent Nils Wilking at Karolinska Institute, specialist in Oncology at Karolinska Hospital.
[16] For information on what registries were used and their population coverage please refer to Appendix B.
[17] For information on which molecules please refer to Appendix B.
[18] http://www.imshealth.com/

when it comes to supplying data regarding the pharmaceutical industry. The drug sales data was presented in dollars and specified for each country and year.
To compute mean drug vintage in three dimensions, the sales data of the 21 molecules had to be distributed on the six cancer sites. For ten of the 21 molecules only one indication was registered during the sample period. For nine of the remaining eleven molecules real numbers on utilization for different indications (sites) were available from IMS Oncology Analyzer, a product of IMS Health. These data cover the period 2002-2006 for five European countries. Using this as an estimate for all of the countries, for the whole period is not optimal, but remains the best estimate available. This is a weakness in data which might affect outcome of the models, but does most probably not. For the last two drugs, Topotecan and Cisplatin estimates of usage for different cancer sites were made based on personal communication with GlaxoSmithKlein, a pharmaceutical company selling Topotecan, and Dr Nils Wilking. Also those two estimations could deviate from the true value, which in the end might affect the results achieved. Most probably the differences between the estimates and reality are rather small. In the end the main findings from this thesis are built on total sample data for which drug utilization still is referred to the correct country and year. Due to this fact, eventual errors in the estimates of drug utilization between cancer sites will cancel out and ought not to have any negative effects on the main findings.

4.2 Missing data

When gathering drug sales data for 1995 to 2004, sales data for 1995 was not provided. Due to this 1995 was dropped from the data sample. Further, the two Italian cancer registries turned out to have considerable differences for some of the survival estimates. This might be due to random errors and perhaps a weighted average of the estimates from the two registries would have made a good estimate. Though, in order not to corrupt data Italy was dropped from the data sample.

When gathering data on drug sales it became evident that data on Estonia was corrupt. Due to this Estonia was dropped from the sample.

In the entire data sample there are no missing values for survival. For mean drug vintage the following data points are missing: breast cancer diagnosed in 1996 for Norway, Poland, Slovenia and Switzerland.

4.3 Data description

Descriptive statistics for sample data is presented in Table 4.3.1 below. As can be seen, data consist of little more than 480 data points for each of the survival and vintage variables. When reflecting over the numbers in this table it is important to remember that when looking at a certain part of the sample, for example a year, it is affected by variations in the other two dimensions. For example, the largest differences in survival lie between the different diseases and not between certain years or countries. Due to this the standard deviations for survival are much larger for a certain year or country than for a certain disease. Still some conclusions can be drawn from the data.

Looking at the years it can be observed that there exists a positive trend in both survival and drug vintage. This poses no problem as the model is a difference-in-difference model (Stock and Watson 2003, pp. 385-388 and 419-420).

When comparing survival for the different cancer sites it can be seen that lung cancer has the worst survival, while breast cancer under optimal conditions have a 92.3 percent 5-year relative survival rate. Looking at the spread (min-max) of survival, i.e. comparing the worst country the worst year with the best county the best year, dramatic differences exist for each cancer site. Thus the positive trend observed when looking at the years hold for each of the diseases.

For the individual countries it can only be concluded that there exist differences in both survival and mean drug vintage between countries.

It is difficult to say anything about the values of the vintage >1995 variable due to the way it is calculated. As stated in section 3.2, this method is more sensible to small inconsistencies in data. That is the reason for the aberrant max value for 1997. The value is due to a vintage data point (colorectal cancer, Norway, 1997) which deviates from the general trend for Norway. The vintage (Year)-variable which is less sensitive to deviating data points is not affected in the same way.

Table 4.3.1: Descriptive statistics

	Survival		Vintage (Year)		Vintage (>1995)	
	Min / Max	Mean / Median	Min / Max	Mean / Median	Min / Max	Mean / Median
Data sample	(Observations)	(StDev)	(Observations)	(StDev)	(Observations)	(StDev)
All	9.1 / 92.3	47.0	1971.0 / 2000.5	1990.8 / 1991.9	0.0 / 99.5	22.8 / 6.1
	(486)	(20.6)	(482)	(6.0)	(482)	(29.9)
1996	9.1 / 80.2	44.1 / 43.8	1971.0 / 1995.4	1985.1 / 1986.6	0.0 / 7.1	0.2 / 0.0
	(54)	(19.4)	(50)	(7.1)	(50)	(1.2)
1997	9.3 / 82.4	44.8 / 44.4	1971.6 / 1996.5	1987.1 / 1988.8	0.0 / 99.5	4.9 / 0.0
	(54)	(19.8)	(54)	(7.4)	(54)	15.9
1998	9.6 / 84.3	45.6 / 45.0	1972.0 / 1996.9	1988.9 / 1990.1	0.0 / 80.2	10.1 / 1.8
	(54)	(20.1)	(54)	(5.8)	(54)	(17.2)
1999	9.8 / 86.0	46.4 / 45.5	1971.5 / 1995.5	1989.8 / 1991.1	0.0 / 91.31	12.9 / 4.1
	(54)	(20.4)	(54)	(5.3)	(54)	(19.3)
2000	10.0 / 87.6	47.1 / 46.1	1971.6 / 1996.6	1990.5 / 1991.4	0.0 / 95.3	16.4 / 5.8
	(54)	(20.7)	(54)	(5.1)	(54)	(23.2)
2001	10.2 / 89.0	47.8 / 46.6	1975.8 / 1997.3	1992.0 / 1993.0	0.0 /97.7	25.7 / 14.6
	(54)	(21.0)	(54)	(3.9)	(54)	(27.4)
2002	10.4 / 90.2	48.5 / 47.3	1982.7 / 1999.0	1993.7 / 1994.3	1.0 / 98.2	39.8 / 38.2
	(54)	(21.3)	(54)	(3.3)	(54)	31.0
2003	10.5 / 91.3	49.2 /47.9	1986.0 / 1999.9	1994.3 / 1994.8	1.0 / 98.9	44.7 / 43.6
	(54)	(21.6)	(54)	(3.4)	(54)	(33.2)
2004	10.5 / 92.3	49.9 / 48.5	1989.8 / 2000.5	1995.0 / 1995.1	0.4 / 99.0	49.1 /48.3
	(54)	(21.8)	(54)	(3.0)	(54)	(34.8)
Breast	67.5 / 92.3	79.5 / 80.3	1983.7 / 1996.2	1992.5 / 1993.1	0.0 / 53.4	20.9 / 17.1
	(81)	(6.0)	(81)	(3.0)	(81)	19.0
Colorectal	30.1 / 67.8	54.54 / 56.6	1994.3 / 1998.1	1995.3 / 1995.4	0.0 / 99.5	29.1 / 25.9
	(81)	(8.6)	(77)	(0.8)	(77)	(27.1)
Leukemia	22.4 / 55.8	54.5 / 44.1	1980.6 / 2000.5	1991.1 / 1989.0	0.0 / 91.5	29.7 / 0.7
	(81)	(7.7)	(81)	(6.3)	(81)	(38.1)
Lung	9.1 / 20.4	14.1 / 14.0	1984.0 / 1993.4	1990.5 / 1990.9	0.0 / 9.8	2.9 / 2.7
	(81)	(3.1)	(81)	(1.8)	(81)	2.4
Non-Hodgkin lymphoma	34.6 / 65.9	52.4 / 52.8	1971.0 / 1997.6	1984.4 / 1985.3	0.0 / 99.0	47.4 / 52.7
	(81)	(7.0)	(81)	(0.1)	(81)	38.2
Ovarian	32.5 / 46.0	39.2 / 38.8	1984.0 / 1993.5	1990.9 / 1991.4	0.0 / 35.3	7.5 / 5.5
	(81)	(3.0)	(81)	(1.9)	(81)	(7.5)
Finland	10.5 / 86.9	47.5 / 47.0	1972.4 / 1999.6	1991.4 / 1992.0	0.0 / 94.2	20.2 / 4.9
	(54)	(22.2)	(54)	(5.2)	(54)	29.9
Germany	13.8 / 83.6	50.3 / 49.7	1972.8 / 2000.5	1992.0 / 1993.0	0.0 / 93.6	28.6 / 14.7
	(54)	(19.8)	(54)	(5.2)	(54)	29.8
Netherlands	14.4 / 84.3	47.3 / 47.6	1971.6 / 1999.5	1989.5 / 1991.3	0.0 / 86.2	17.4 / 4.2
	(54)	20.6	(54)	(6.6)	(54)	(25.3)
Norway	12.5 / 85.6	49.3 / 50.4	(1971.5 / 2000.1	1992.1 / 1992.9	0.0 / 99.5	26.1 / 4.0
	(54)	(20.9)	(53)	(5.3)	(53)	(33.1)
Poland	12.4 / 73.2	26.5 / 35.55	1971.5 / 2000.1	1988.3 / 1988.8	0.0 / 91.3	17.5 / 0.1
	(54)	(17.7)	(53)	(7.0)	(53)	(25.7)
Slovenia	10.0 / 76.8	43.3 / 40.5	1972.8 / 1999.8	1989.7 / 1990.0	0.0 / 89.0	15.3 / 0.0
	(54)	(18.5)	(53)	(6.5)	(53)	29.0
Switzerland	18.7 / 88.9	52.5 / 53.4	1972.2 / 1999.2	1991.6 / 1991.7	0.0 / 89.3	32.0 / 18.3
	(54)	(19.9)	(53)	(5.3)	(53)	31.5
United Kingdom	9.1 / 80.3	45.2 / 48.1	1971.5 / 1999.7	1989.6 / 1990.8	0.0 / 88.3	19.9 / 3.8
	(54)	(20.2)	(54)	(6.3)	(54)	(27.6
USA	15.8 / 92.3	51.3 / 48.4	1971.0 / 1999.1	1992.6 / 1992.9	0.0 / 99.0	28.6 / 11.1
	(54)	(22.1)	(54)	(5.1)	(54)	(33.0)

5 Results and discussion

5.1 Model 1 - two dimensions: cancer site and country

Table 5.1.1: Results from model 1

Dependent variable	Survival	Survival
Intercept	-456.551	78.754
	(<0.001)***	(<0.001)***
Vintage	0.269	0.068
	(<0.001)***	(<0.001)***
Colorectal	-25.380	-25.17
	(<0.001)***	(<0.001)***
Leukemia	-36.791	-37.765
	(<0.001)***	(<0.001)***
Lung	-64.934	-64.243
	(<0.001)***	(<0.001)***
Non-Hodgkin Lymphoma	-24.984	-28.945
	(<0.001)***	(<0.001)***
Ovarian	-39.882	-39.401
	(<0.001)***	(<0.001)***
Germany	2.659	2.243
	(0.001)***	(0.020)**
Netherlands	0.239	-0.076
	(0.775)	(0.918)
Norway	1.639	1.444
	(0.033)**	(<0.050)**
Poland	-9.966	-10.614
	(<0.001)***	(<0.001)***
Slovenia	-3.609	-3.725
	(<0.001)***	(<0.001)***
Switzerland	4.984	4.242
	(<0.001)***	(<0.001)***
United Kingdom	-1.801	-2.268
	(0.019)**	(0.020)**
USA	3.457	3.200
	(<0.001)***	(<0.001)***
Cancer site	All	All
Country	All	All
Years	All	All
Vintage	Year	>1995
Adjusted R^2	0.963	0.966

*** significance at the one percent level

** significance at the five percent level

** significance at the ten percent level

Table 5.1.1 shows the results of model 1 when the regression is run with the whole data sample. Two different approaches with regard to the vintage variable are presented. In the first column the preferred vintage variable measured as weighted year, denoted Year, is found. The second column presents the result when using the vintage variable measured as the fraction of patients treated with drugs launched after 1995.

For the preferred vintage variable model 1 produces a significant result at the one percent level. The coefficient for the vintage variable is positive in line with the hypothesis. When the model was used in "*A global comparison regarding patient access to cancer drugs*" (Jönsson and Wilking 2007, pp.iii72-74) the authors found a positive, but not significant result for this variable. Looking at the adjusted R^2 it seems that the explanatory power of the model is very good. However, this is partly an effect due to the fact that there are several dummies in the model that correct for differences in the sample.

Turning to the included dummies for the whole sample they are all significant at the one percent level except for the United Kingdom which is significant at the five percent level and the Netherlands which is not significant at all. Not much can be learned from their values, which reflect that survival differ for different diseases and countries when compared to the base case: breast cancer in Finland. When doing such comparisons it is better to look at the descriptive statistics or preferably directly on input data.

Limiting the sample to the individual diseases give the results presented below in Table 5.1.2. From the table it can be seen that the preferred vintage variable is positive and significant at the one percent level for all diseases except for ovarian cancer. The values of the vintage coefficients vary between the different diseases indicating that mean drug vintage have different impacts. To draw any further conclusions regarding the coefficient values or to compare them should be done with an outmost wariness. Though it ought not to be faulty to assume that the effect of mean drug vintage is larger for colorectal cancer than for leukemia.

Limiting the data sample to the individual countries generate the results presented in Table 5.1.3 below. It shows that the mean drug vintage variable measured as a weighted average year is positive and significant at the one percent level in all countries, except for Slovenia which is significant at the five percent level. There are small variations in the vintage coefficient between the different countries. The variations are however too small to say anything about how the countries compare to each other.

Table 5.1.2: Results from model 1 when limiting sample data to individual diseases

Dependent variable	Survival	Survival	Survival	Survival	Survival	Survival	Survival
Intercept	-456.551	-1511.679	-3716.230	-329.484	-891.020	-518.004	-440.529
	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(0.244)
Vintage	0.269	0.286	1.892	0.186	0.453	0.287	0.242
	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(0.203)
Germany	2.659	-4.331	0.557	5.281	5.289	6.139	1.118
	(0.001)***	(<0.001)***	(0.680)	(<0.001)***	(<0.001)***	(<0.001)***	(0.194)
Netherlands	0.239	-0.733	-2.544	2.460	4.227	0.851	-3.292
	(0.775)	(0.417)	(0.055)*	(<0.001)***	(<0.001)***	(0.365)	(<0.001)***
Norway	1.639	-2.387	-0.656	6.480	2.426	2.198	0.068
	(0.033)**	(0.010)***	(0.633)	(<0.001)***	(<0.001)***	(<0.018)**	(0.937)
Poland	-9.966	-12.850	-19.016	-15.835	4.765	-13.196	-4.262
	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***
Slovenia	-3.609	-11.005	-14.119	0.323	1.820	1.724	-1.776
	(<0.001)***	(<0.001)***	(<0.001)***	(0.517)	(<0.001)***	(<0.065)*	(<0.084)*
Switzerland	4.984	-0.487	2.677	11.925	9.305	3.590	0.643
	(<0.001)***	(0.590)	(0.070)*	(<0.001)***	(<0.001)***	(<0.001)***	(0.451)
United Kingdom	-1.801	-8.047	-6.822	8.545	0.470	-0.324	-5.653
	(0.019)**	(<0.001)***	(<0.001)***	(<0.001)***	(0.136)	(0.726)	(<0.001)***
USA	3.457	1.376	3.897	3.124	6.679	4.856	-1.251
	(<0.001)***	(0.142)	(0.004)***	(<0.001)***	(<0.001)***	(<0.001)***	(0.152)
Cancer site	All	Breast	Colorectal	Leukemia	Lung	Non-Hodgkin	Ovarian
Country	All	All	All	All	All	All	All
Years	All	All	All	All	All	All	All
Vintage	Year	Year	Year	Year	Year	Year	Year
Adjusted R^2	0.963	0.899	0.888	0.982	0.961	0.923	0.655

*** significance at the one percent level

** significance at the five percent level

** significance at the ten percent level

Table 5.1.3: Results from model 1 when limiting sample data to individual countries

Dependent variable	Survival	Survival	Survival	Survival	Survival	Survival	Survival	Survival	Survival	Survival
Intercept	-456.551	-486.713	-706.108	-304.404	-562.531	-264.469	-132.208	-906.650	-528.614	-780.657
	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(0.040)**	(0.139)	(<0.001)***	(<0.001)***	(<0.001)***
Vintage	0.269	0.286	0.394	0.194	0.323	0.168	0.103	0.497	0.303	0.435
	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(0.010)***	(0.024)**	(<0.001)***	(<0.001)***	(<0.001)***
Colorectal	-25.380	-26.314	-20.794	-26.852	-24.074	-21.911	-28.156	-22.190	-24.646	-24.848
	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***
Leukemia	-36.791	-43.207	-34.097	-39.774	-35.040	-46.239	-32.142	-30966	-26.635	-42.868
	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***
Lung	-64.934	-72.568	-63.174	-67.525	-68.235	-55.650	-60.769	-62.909	-64.732	-68.308
	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***
Non-Hodgkin Lymphoma	-24.984	-29.399	-18.658	-39.774	-25.221	-31.380	-18.561	-24.736	-21.863	-26.875
	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***
Ovarian	-39.882	-42.208	-37.229	-44.540	-40.407	-33.890	-33.533	-41.234	-40.058	-45.770
	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***
Cancer site	All	All	All	All	All	All	All	All	All	All
Country	All	Finland	Germany	Netherlands	Norway	Poland	Slovenia	Switzerland	UK	USA
Years	All	All	All	All	All	All	All	All	All	All
Vintage	Year	Year	Year	Year	Year	Year	Year	Year	Year	Year
Adjusted R^2	0.963	0.994	0.994	0.999	0.993	0.986	0.992	0.991	0.991	0.987

*** significance at the one percent level
** significance at the five percent level
** significance at the ten percent level

Unlike the results of model 1 above, Jönsson and Wilking (2007) do not attain a significant result with the preferred vintage variable. Thus, they turn to the alternative drug vintage variable measuring the fraction of patients treated with drugs launched after a defined limit year. Running their model with several different limit years, Jönsson and Wilking (2007) find a relationship, significant at the one percent level, between 5-year survival and mean drug vintage only for the limit year >1985. It yields a positive vintage coefficient of 0.383.
To be able to compare results, the regression of model 1 was run with the alternative vintage variable for some different limit years. The years tested were >1980, >1985, >1990, >1995 and >2000. Only >1995 was significant at the one percent level. As Table 5.1.1 shows it was positive with the coefficient 0.068. In Table 5.1.1 only the regression using >1995 is presented.

Considering the difference in the limit years, >1985 and >1995, yielding a significant vintage variable in Jönsson and Wilking (2007) and this thesis respectively, it seems reasonable since the year of diagnosis for the patient cohorts regarded also differ by approximately ten years, 1990-1994 and 1996-2004. Regarding the value of the vintage coefficient found by Jönsson and Wilking (2007), it is almost six times larger than the one found in this thesis. It is difficult to say if anything can be concluded from the difference in the magnitude of the coefficients. The fact that the models include different cancer sites might explain some, if not most, of the difference. In the end, both studies find vintage coefficients that are significant and positive in line with the hypothesis of this thesis.

To conclude the result of model 1, the hypothesis can not be rejected. The result of model 1 is in line with previous research, supporting the hypothesis that there exists a positive and significant relationship between cancer survival and mean cancer drug vintage.

5.2 Model 2 - three dimensions: cancer site, country and year

Table 5.2.1: Results from model 2

Dependent variable	Survival	Survival
Intercept	-35.529	77.514
	(0.701)	(<0.001)***
Vintage	0.057	0.038
	(0.223)	(<0.001)***
1997	0.311	0.244
	(0.677)	(0.739)
1998	0.973	0.813
	(0.202	(0.270)
1999	1.660	1.442
	(0.032)**	(0.052)*
2000	2.346	2.038
	(0.003)***	(0.006)***
2001	2.976	2.400
	(<0.001)***	(0.002)***
2002	3.576	2.568
	(<0.001)***	(0.002)***
2003	4.232	3.070
	(<0.001)***	(<0.001)***
2004	4.860	3.576
	(<0.001)***	(<0.001)***
Colorectal	-24.899	-25.020
	(<0.001)***	(<0.001)***
Leukemia	-37.090	-37.503
	(<0.001)***	(<0.001)***
Lung	-65.344	-64.776
	(<0.001)***	(<0.001)***
Non-Hodgkin Lymphoma	-26.695	-28.156
	(<0.001)***	(<0.001)***
Ovarian	-40.216	-39.799
	(<0.001)***	(<0.001)***
Germany	2.781	2.495
	(0.001)***	(0.001)***
Netherlands	-0.159	-0.159
	(0.829)	(0.824)
Norway	1.763	1.589
	(0.017)**	(0.028)**
Poland	-10.660	-10.726
	(<0.001)***	(<0.001)***
Slovenia	-4.001	-3.902
	(<0.001)***	(<0.001)***
Switzerland	4.990	4.563
	(<0.001)***	(<0.001)***
United Kingdom	-2.183	-2.276
	(0.003)***	(0.002)***
USA	3.704	3.451
	(<0.001)***	(<0.001)***
Cancer site	All	All
Country	All	All
Years	All	All
Vintage	Year	>1995
Adjusted R^2	0.967	0.968

*** significance at the one percent level

** significance at the five percent level

** significance at the ten percent level

Table 5.2.1 shows the results of model 2 when the regression is run with the full data sample. Two different approaches in regard of the vintage variable are presented. In the first column the preferred vintage variable is measured as weighted year, denoted Year. The second column presents the result when using the vintage variable measured as the fraction of patients treated with drugs launched after 1995, denoted >1995.

When the regression of model 2 is run with the preferred vintage variable, the result is not significant. The coefficient for the vintage variable is positive in line with the hypothesis but no conclusions can be drawn as it is not significant. The regression is re-run with the vintage variable measured as the fraction of patients treated with drugs launched after some different limit years. The years tested were >1980, >1985, >1990, >1995 and >2000. Only >1995 was significant at the one percent level. As Table 5.2.1 shows it was positive with the coefficient 0.038. In Table 5.2.1 only the regression using >1995 is presented. For the same reason as stated in previous section, using >1995 is in line with Jönsson and Wilking (2007). Looking at the adjusted R^2 it seems that the explanatory power of the model is very good. However, as for model 1, this is partly an effect due to the fact that there are several dummies in the model that correct for differences in the sample.

The dummies only tell how each factor represented by the dummy relates to dimension in the base case, breast cancer for disease dummies, Finland for country dummies and 1996 for year dummies. As stated in previous section, the best way to find out how they relate to each other is to look at the descriptive statistics or preferably sample data.

Limiting the sample to the individual diseases give the results presented below in Table 5.2.2. From the table it can be seen that the mean drug vintage variable measured as a fraction is only significant for lung cancer. For the other diseases it is insignificant at the ten percent level and in some cases show a negative sign.

Limiting the data sample to the individual countries generate the results presented in Table 5.2.3 below. It shows that the mean drug vintage variable measured as a fraction is positive and significant at the one percent level for the Netherlands, Switzerland and USA. For the other countries the vintage variable is insignificant and in two cases negative.

Limiting the sample to nine periods of one year give the results presented in Table 5.2.4 below. From the table it can be seen that the mean drug vintage variable measured as a fraction is positive and significant at the five percent level for the years 2001-2003 and

positive and significant at the ten percent level for 2004. For the other years it is insignificant and positive except for 1996 which is negative

Table 5.2.2: Results from model 2 when limiting sample data to individual diseases

Dependent variable	Survival	Survival	Survival	Survival	Survival	Survival	Survival
Intercept	77.514	78.678	53.704	38.077	9.474	48.118	39.589
	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***
Vintage	0.038	0.0001	-0.006	-0.006	0.110	0.010	-0.048
	(<0.001)***	(0.994)	(0.565)	(0.735)	(0.005)***	(0.509)	(0.142)
1997	0.244	1.3	0.733	0.478	-0.063	1.096	0.618
	(0.739)	(0.003)***	(0.329)	(0.288)	(0.816)	(0.184)	(0.416)
1998	0.813	2.521	1.849	0.957	0.214	1.935	1.034
	(0.270)	(<0.001)***	(0.015)**	(0.036)**	(0.430)	(0.037)**	(0.180)
1999	1.442	3.665	2.939	1.436	0.458	2.907	1.403
	(0.052)*	(<0.001)***	(<0.001)***	(0.002)***	(0.096)*	(0.005)***	(0.071)*
2000	2.038	4.764	4.096	1.926	0.715	3.868	1.767
	(0.006)***	(<0.001)***	(<0.001)***	(<0.001)***	(0.010)***	(0.001)***	(0.024)**
2001	2.400	5.808	5.254	2.521	1.049	4.725	2.083
	(0.002)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(0.008)***
2002	2.568	6.774	6.392	3.294	1.301	5.647	2.625
	(0.002)***	(<0.001)***	(<0.001)***	(0.014)**	(<0.001)***	(<0.001)***	(0.002)***
2003	3.070	7.684	7.511	3.836	1.583	6.606	3.070
	(<0.001)***	(<0.001)***	(<0.001)***	(0.010)***	(<0.001)***	(<0.001)***	(<0.001)***
2004	3.576	8.562	8.618	4.322	1.830	5.523	3.461
	(<0.001)***	(<0.001)***	(<0.001)***	(0.006)***	(<0.001)***	(<0.001)***	(<0.001)***
Germany	2.495	-3.769	2.420	5.454	5.488	5.950	1.583
	(0.001)***	(<0.001)***	(0.001)	(<0.001)***	(<0.001)***	(<0.001)***	(0.041)**
Netherlands	-0.159	-1.044	-1.593	2.100	4.037	-1.256	-3.462
	(0.824)	(0.016)**	(0.014)**	(<0.001)***	(<0.001)***	(0.179)	(<0.001)***
Norway	1.589	-1.422	0.535	6.667	2.800	2.262	0.129
	(0.028)**	(0.001)***	(0.480)	(<0.001)***	(<0.001)***	(0.007)***	(0.863)
Poland	-10.726	-12.868	-19.320	-16.319	3.803	-15.477	-5.057
	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***
Slovenia	-3.902	-10.799	-13.645	0.183	1.063	0.615	-2.711
	(<0.001)***	(<0.001)***	(<0.001)***	(0.682)	(<0.001)***	(0.480)	(0.001)
Switzerland	4.563	0.198	4.939	11.913	8.731	3.774	0.764
	(<0.001)***	(0.688)	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(0.316)
United Kingdom	-2.276	-7.477	-5.672	8.034	-0.496	-1.623	-6.126
	(0.002)***	(<0.001)***	(<0.001)***	(<0.001)***	(0.049)**	(0.063)	(<0.001)***
USA	3.451	3.788	4.623	3.376	6.135	5.792	-1.080
	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(0.175)
Cancer site	All	Breast	Colorectal	Leukemia	Lung	Non-Hodgkin	Ovarian
Country	All	All	All	All	All	All	All
Years	All	All	All	All	All	All	All
Vintage	>1995	>1995	>1995	>1995	>1995	>1995	>1995
Adjusted R^2	0.968	0.978	0.976	0.946	0.972	0.939	0.737

*** significance at the one percent level

** significance at the five percent level

** significance at the ten percent level

Table 5.2.3: Results from model 2 when limiting sample data to individual countries

Dependent variable	Survival	Survival	Survival	Survival	Survival	Survival	Survival	Survival	Survival	Survival
Intercept	77.514	80.515	75.963	80.793	78.491	69.228	71.023	79.666	72.832	82.678
	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***
Vintage	0.038	0.006	0.010	0.022	0.011	-0.017	-0.006	0.074	0.015	0.046
	(<0.001)***	(0.597)	(0.392)	(<0.001)***	(0.103)	(0.358)	(0.594)	(<0.001)***	-0.255	(0.001)***
1997	0.244	0.687	0.822	0.331	0.772	-0.590	-0.079	0.519	0.764	1.149
	(0.739)	(0.397)	(0.313)	(0.317)	(9.186)	(0.595)	(0.927)	(0.615)	-0.343	(0.220)
1998	0.813	1.377	1.709	0.674	1.726	0.090	0.405	-0.128	1.521	1.567
	(0.270)	(0.095)*	(0.043)**	(0.046)**	(0.004)***	(0.936)	(0.637)	(0.906)	(0.063)*	(0.107)
1999	1.442	2.029	2.614	0.963	2.526	0.749	0.971	0.829	2.220	2.375
	(0.052)*	(0.017)**	(0.003)***	(0.006)***	(<0.001)***	(0.504)	(0313)	(0.444)	(0.009)***	(0.019)**
2000	2.038	2.677	3.420	1.230	3.281	1.432	1.356	1.385	2.913	3.429
	(0.006)***	(0.002)***	(<0.001)***	(0.001)***	(<0.001)***	(0.208)	(0.120)	(0.212)	(0.001)***	(0.001)***
2001	2.400	3.294	4.251	1.455	3.955	2.069	1.897	1.362	3.501	4.058
	(0.002)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(0.073)*	(0.035)**	(0.244)	(<0.001)***	(<0.001)***
2002	2.568	3.865	4.950	1.381	4.667	3.099	2.468	1.874	4.034	4.686
	(0.002)***	(<0.001)***	(<0.001)***	(0.001)***	(<0.001)***	(0.020)**	(0.012)**	(0.121)	(<0.001)***	(<0.001)***
2003	3.070	4.501	5.773	1.585	5.419	3.821	2.908	2.381	4.657	5.311
	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(0.007)***	(0.004)***	(0.056)*	(<0.001)***	(<0.001)***
2004	3.576	5.151	6.560	1.804	6.152	4.612	3.399	3.086	5.310	5.916
	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(0.003)***	(0.001)***	(0.015)**	(<0.001)***	(<0.001)***
Colorectal	-25.020	-25.392	-19.770	-26.318	-24.049	-31.986	-27.952	-21.807	-24.071	-24.505
	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***
Leukemia	-37.503	-43.313	-34.163	-40.321	-35.352	-46.784	-32.142	-31.501	-28.031	-43.881
	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***
Lung	-64.776	-72.606	-63.087	-67.318	-68.432	-56.629	-61.112	-61.752	-65.631	-69.089
	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***
Non-Hodgkin	-28.156	-31.007	-21.437	-31.499	-27.563	-33.892	-19.378	-29.130	-25.485	-30.809
	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***
Ovarian	-39.799	-42.070	-37.020	-44.301	-40.309	-34.934	-33.885	-40.148	-40.675	-46.905
	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***
Cancer site	All	Finland	Germany	Netherlands	Norway	Poland	Slovenia	Switzerland	United Kingdom	USA
Country	All	All	All	All	All	All	All	All	All	All
Years	All	All	All	All	All	All	All	All	All	All
Vintage	>1995	>1995	>1995	>1995	>1995	>1995	>1995	>1995	>1995	>1995
Adjusted R^2	0.968	0.996	0.995	0.999	0.998	0.989	0.994	0.993	0.995	0.995

*** significance at the one percent level

** significance at the five percent level

** significance at the ten percent level

Table 5.2.4: Results from model 2 when limiting sample data to individual years

Dependent variable	Survival	Survival	Survival	Survival	Survival	Survival	Survival	Survival	Survival	Survival
Intercept	77.514	75.157	76.941	78.006	78.834	79.793	79.160	78.506	78.840	79.498
	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***
Vintage	0.038	-0.325	0.022	0.022	0.027	0.024	0.084	0.107	0.110	0.121
	(<0.001)***	(0.604)	(0.626)	(0.659)	(0.575)	(0.543)	(0.029)**	(0.032)**	(0.038)**	(0.087)*
Colorectal	-25.020	-22.121	-25.404	-25.302	-25.154	-25.195	-25.316	-25.409	-25.942	-26.956
	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***
Leukemia	-37.503	-34.524	-35.312	-35.931	-36.390	-36.914	-37.430	-42.522	-42.632	-44.588
	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***
Lung	-64.776	-61.793	-63.041	-63.846	-64.536	-65.289	-64.372	-62.625	-63.729	-63.641
	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***
Non-Hodgkin	-28.156	-26.858	-27.032	-27.629	-27.895	-27.976	-30.591	-31.662	-31.711	-32.601
	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***
Ovarian	-39.799	-36.947	-38.251	-38.983	-39.593	-40.200	-39.394	-39.012	-39.182	-39.115
	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***
Germany	2.495	1.933	1.942	2.146	2.464	2.576	2.143	2.244	2.652	2.368
	(0.001)***	(0.407	(0.451)	(0.395)	(0.309)	(0.285)	(0.346)	(0.329)	(0.258)	(0.347)
Netherlands	-0.159	1.050	0.729	0.449	0.196	-0.155	0.053	-0.495	-1.041	-1.980
	(0.824)	(0.651)	(0.773)	(0.855)	(0.935)	(0.948)	(0.981)	(0.826)	(0.651)	(0.415)
Norway	1.589	1.027	1.029	1.413	1.597	1.694	1.128	1.848	2.066	1.632
	(0.028)**	(0.676)	(0.693)	(0.566)	(0.506)	(0.475)	(0.617)	(0.412)	(0.371)	(0.505)
Poland	-10.726	-8.353	-10.835	-10.887	-10.912	-10.960	-10.123	-10.793	-10.980	-11.996
	(<0.001)***	(0.002)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***	(<0.001)***
Slovenia	-3.902	-1.413	-3.568	-3.707	-3.757	-3.903	-3.894	-4.515	-4.314	-5.307
	(<0.001)***	(0.566)	(0.163)	(0.138)	(0.131)	(0.113)	(0.087)*	(0.049)**	(0.068)*	(0.032)**
Switzerland	4.563	4.027	4.334	4.169	4.462	4.680	3.566	4.458	4.695	4.637
	(<0.001)***	(0.107)	(0.092)*	(0.123)	(0.080)*	(0.061)*	(0.134)	(0.056)*	(0.050)**	(0.066)*
United Kingdom	-2.276	-2.597	-2.499	-2.377	-2.279	-2.215	-2.294	-2.072	-2.169	-2.649
	(0.002)***	(0.267)	(0.325)	(0.335)	(0.344)	(0.350)	(0.305)	(0.357)	(0.348)	(0.279)
USA	3.451	2.731	2.607	2.780	3.101	3.615	3.241	4.112	4.186	3.563
	(<0.001)***	(0.298)	(0.305)	(0.275)	(0.215)	(0.138)	(0.158)	(0.073)*	(0.077)*	(0.169)
Cancer site	All	1996	1997	1998	1999	2000	2001	2002	2003	2004
Country	All	All	All	All	All	All	All	All	All	All
Years	All	All	All	All	All	All	All	All	All	All
Vintage	>1995	>1995	>1995	>1995	>1995	>1995	>1995	>1995	>1995	>1995
Adjusted R^2	0.968	0.960	0.952	0.956	0.959	0.962	0.967	0.967	0.967	0.964

*** significance at the one percent level

** significance at the five percent level

** significance at the ten percent level

Dividing the data sample in the three dimensions only one disease, three countries and four years generate significant results, all of them positive. It can be questioned if these positive and significant factors drive the result of the regression for the full sample. In order to test this model 2 was re-run, still with the vintage variable measured as a fraction of patients treated with drugs launched after 1995, with the potentially driving factors excluded, i.e. excluding lung cancer, the Netherlands, Switzerland, USA and 2001-2004. Thus, the sample consists of only 150 survival estimates and 147 vintage estimates. The result is presented below in table 5.2.5.

From Table 5.2.5 it can be seen that the survival variable is positively dependent on the vintage variable at the ten percent significance level. Knowing this, the hypothesis that the excluded factors were driving the result for the full sample can be rejected. The fact that the vintage variable was insignificant for most of the diseases, countries and years might have been due to the small samples: 81, 54 and 54 respectively. Thus, the result presented in table 5.2.1 holds.

Table 5.2.5: Control of model 2

Dependent variable	Survival
Intercept	76.867
	(<0.001)***
Vintage	0.043
	(0.066)*
1997	0.141
	(0.894)
1998	0.891
	(0.401)
1999	1.541
	(0.153)
2000	2.139
	(0.054)*
Colorectal	-25.754
	(<0.001)***
Leukemia	-34.882
	(<0.001)***
Non-Hodgkin Lymphoma	-26.554
	(<0.001)***
Ovarian	-37.040
	(<0.001)***
Germany	1.605
	(0.157)
Norway	0.938
	(0.410)
Poland	-13.233
	(<0.001)***
Slovenia	-3.992
	(0.001)***
United Kingdom	-2.635
	(0.019)**
Cancer site	All
Country	All
Years	All
Vintage	>1995
Adjusted R^2	0.931

*** significance at the one percent level

** significance at the five percent level

** significance at the ten percent level

The value of the vintage coefficient found in the "*A global comparison regarding patient access to cancer drugs*" (Jönsson and Wilking 2007) is ten times larger than the one found with this model. What value is the correct one might be discussed but is in essence not very interesting. As for model 1, the fact that the models include different cancer sites might explain some, if not most, of the difference. What is interesting is the fact that the new model, model 2, confirms the findings of "*A pan-European comparison regarding patient access to cancer drugs*" (Wilking and Jönsson 2005), "*A global comparison regarding patient access to cancer drugs*" (Jönsson and Wilking 2007) and model 1. This new model, never before tested on sample data, created to be superior to previous models by correcting for the development

of technology that occurs over time, supports the hypothesis that cancer survival is positively dependent on the mean cancer drug vintage. To conclude the findings of model 2, it does, just like model 1, support the hypothesis that there exist a positive and significant relationship between cancer survival and mean cancer drug vintage.

Most probably the finding of this thesis that mean cancer drug vintage has a positive effect on cancer survival, in line with previous research, can be explained to a large extent by the breakthroughs in endocrine therapy[19] and biotherapy[20]. Since the late nineties a number of drugs from both research fields have been launched globally. The introduction of anti-hormones, aromatase inhibitors and monoclonal antibodies[21] have improved cancer treatment and dramatically reduced the burden of highly toxic agents, resulting in fewer side effects and increased cancer patient survival. Considering the slow market uptake of new oncology drugs (Wilking and Jönsson 2005) there is reason to believe that there is still room for improvement in cancer survival due to drugs already available to the market. If this is the case we will continue to see an increase in the vintage variable and a corresponding increase in cancer survival for the years to come.

[19] For more information, please refer to Appendix A, section A5.
[20] For more information, please refer to Appendix A, section A5.
[21] For more information, please refer to Appendix A, section A5.

6 Conclusions and final remarks

Starting with results of model 1, the sample data for this thesis supports the hypothesis and confirms the results found by Jönsson and Wilking (2007) in their corresponding model, showing that mean cancer drug vintage has a positive effect on cancer survival.

Heading on to model 2, this model was suggested in the "*A global comparison regarding patient access to cancer drugs*" (Jönsson and Wilking 2007) but was never developed nor tested due to the lack of the three dimensional sample data needed. In this thesis the model has seen the light of day for the first time. When tested on collected sample data it does support the hypothesis that cancer survival is dependent on mean drug vintage, that the relationship is positive and that it is highly significant with a p-value less than 0.001.

Having corrected cancer survival rates for differences between countries, differences between different cancer sites and differences over time, such as advances in medicine technology, the result of this thesis indicate a that mean cancer drug vintage have a positive effect on cancer survival. Further testing by removal of the countries, cancer sites and years, identified as strong contributors to the outcome, do not alter the result. However, it must be kept in mind that the model used in this thesis does not take important confounding variables like improvements of radiology equipment, surgery or screening into account. The correction of the survival rates described above reduce their influence in the model, but does not fully correct for them. Thus, the finding of this thesis must be handled with great carefulness.

Finally, it is important to remember that no finding is stronger than the sample data it is derived from. The sample data for this thesis consist a weakness in the matter of how cancer drug sales were estimated among the six diseases. However, as the main finding is derived from the full sample data, consisting all of the six cancer sites, it should make no big difference as the drug sales are still attributed to the right year and the right country.

In conclusion the findings of this thesis indicate that cancer survival is positively affected by mean cancer drug vintage.

6.1 Further research

It can be questioned whether the vintage effect is equal for different cancer sites. Reasonably the effect of mean cancer drug vintage should be larger for cancer sites for which there has

been recent breakthroughs in the cancer drug development leading to large increases in survival. Due to this it would be interesting to test in future research if the vintage factor affects different cancer sites similarly or if there is a difference in magnitude of the effect.

7 References

- Brenner, H. and Gefeller, O. (1996) "An alternative approach to monitoring cancer patient survival", *Cancer*, 78, 2004–2010.

- Brenner, H. and Gefeller, O. (1997) "Deriving more up-to-date estimates of long-term patient survival", *Journal of Clinical Epidemiology*, 50, 211–216.

- Brenner, H. and Hakulinen, T. (2006) "Up-to-date and Precise Estimates of Cancer Patient Survival: Model-based Period Analysis", *American Journal of Epidemiology*, 164, 689-696.

- Boyle, P., d'Onofrio, A., Maisonneuve, P. et al. (2005) "Measuring progress against cancer in Europe: has the 15% decline targeted for 2000 come about?", *Annals of Oncology*, 14, 1312–1325.

- Carroll, P. (2005) "Early stage prostate cancer – Do we have a problem with over-detection, over treatment or both?", *The Journal of Urology*, 173, 1061-1062.

- Coleman, M. (2006) "New drugs and survival: does the Karolinska report make sense?", *Cancerworld*, 14, 26-29.

- Coleman, M. (2007) "Editorial", *Annals of Oncology*, 18, 1433–1435.

- Corazziari, I., Quinn, M. and Capocaccia, R. (2004) "Standard cancer patient population for age standardizing survival ratios", *European Journal of Cancer*, 40, 2307–2316.

- Cutler, D. and McClellan, M. (2001) "Is Technological Change in Medicine Worth It?", *Health Affairs* 20(5), 11–29.

- FASS, (2007) *FASS (Pharmaceutical Specialities in Sweden)*, Läkemedelsinformation AB. Also available [online]: http://www.fass.se [2007-11-30]

- Folland, S., Goodman A. and Stano, M. (2001) *The Economics of Health and Health Care*, third edition. Upper Saddle River, NJ: Prentice Hall.

- Getzen, T. (1997) *Health Economics: Fundamentals and Flow of Funds*. New York, NY: John Wiley and Sons.
- Hanahan, D. and Weinberg, RA. (2000) "The hallmarks of cancer", *Cell*, 100, 57–70.
- Henderson, J. (1999) *Health Economics and Policy*. Cincinnati, OH: South-Western Publishing Co.
- Jönsson, B. and Wilking, N. (2007) "A global comparison regarding patient access to cancer drugs", *Annals of Oncology*, 18, Supplement 3.
- Kobelt, G. (2002) *Health Economics: An Introduction to Economic Evaluation*. London: Office of Health Economics.
- Lichtenberg, F. (2005a) "Availability of new drugs and Americans' ability to work", *Journal of Occupational and Environmental Medicine*, 47, 373–380.
- Lichtenberg, F. (2005b) "Pharmaceutical knowledge-capital accumulation and longevity", In Corrado, C., Haltiwanger, J., Sichel, D. (eds): *Measuring Capital in the New Economy*, University of Chicago Press.
- Lichtenberg, F. (2005c) "The Impact of New Drug Launches on Longevity: Evidence from Longitudinal, Disease-Level Data from 52 Countries, 1982–2001", *International Journal of Health Care Finance and Economics*, 5, 47–73.
- Lichtenberg, F. (2006) "The impact of increased utilization of HIV drugs on longevity and medical expenditure: an assessment based on aggregate U.S. time-series data", *Expert Review Pharmacoeconomics and Outcomes Research*, 6, 425–436.
- Luce, B., Mauskopf, J., Sloan, F. et al. (2006) "The Return on Investment in Health Care: From 1980 to 2000", *Value in Health*, 9, 146-156.
- OECD (2007) *OECD Health Data 2007*. Available [online]: www.oecd.org/health/healthdata [2007-11-30].
- Ringborg, U., Henriksson, R. and Friberg, S. (1998) *ONKOLOGI*. Borås: Liber.

- Santerre, R. and Neun, S. (2000) *Health Economics: Theories, Insights, and Industry Studies*, revised edition. Orlando, FL: Dryden Press.

- Stock, J. and Watson, M. (2003) *Introduction to Econometrics*. Boston, MA: Addison Wesley.

- Talbäck, M., Stenbeck, M. and Rosén, M. (2004) "Up-to-date long-term survival of cancer patients: an evaluation of period analysis on Swedish Cancer Registry data", *European Journal of Cancer*, 40, 1361–1372.

- Villa, L., Costa, R., Petta, C. et al. (2005) "Prophylactic quadrivalent human papillomavirus (types 6, 11, 16, and 18) L1 virus-like particle vaccine in young women: a randomised double-blind placebo-controlled multicentre phase II efficacy trial", *Lancet Oncology*, 6, 271–278.

- von der Maase, H., Sengelov, L., Roberts, J.T. et al. (2005), "Long-term survival results of a randomized trial comparing gemcitabine plus cisplatin, with methotrexate, vinblastine, doxorubicin, plus cisplatin in patients with bladder cancer", *Journal of Clinical Oncology*, 23, 4602-4608.

- Wilking, N. and Jönsson, B. (2005) "A pan-European comparison regarding patient access to cancer drugs". Available [online]: http://ki.se/content/1/c4/33/52/Cancer_Report.pdf [2007-11-30]

- World Health Organization (2004) "Priority medicines for Europe and the world.". Available [online]: http://whqlibdoc.who.int/hq/2004/WHO_EDM_PAR_2004.7.pdf. [2007-11-30]

Appendix A - Cancer

This section will give the layman readers, i.e. most economist and others not having studied medicine, a brief insight in the mysteries of cancer. The pathological mechanism behind the disease will be described in layman terms. Next, the effects in terms of mortality and costs will be discussed. Prevention of the disease will be touched up on and finally treatment in general and cancer drugs in certain will be described. Where noting else is indicated the text is built on ONKOLOGI (Ringborg et al 1998), a textbook for medical school students taking the compulsory course in oncology.

A1 The Disease

Cancer is a disease resulting from one or several maladies in the cell cycle. The cell cycle, i.e. the lifecycle of a cell, is responsible for its growth, replication and apoptosis (programmed cell death). In order for cancer to appear, not one but several events have to occur. In a very simple, schematic way the complicated series of events leading to cancer can be discussed in terms of *ignition* and *promotion*. The first step, ignition, is an irreversible mutation of the DNA in a somatic cell[22]. Some kinds of chemical substances, ultraviolet or ionised radiation and genetic materials (such as viruses) are known factors that can cause such mutations to the DNA. Though, ignition is in itself not enough to turn a cell into a cancer cell. There has to be some other factor (promotion), direct or indirect, stimulating cell growth.

Damages in the DNA and other external factors alter the production of the cancer cell leading to an over expression of signal transmitters and receptors as well as other products of the cell machinery. This creates a cell with an extraordinary possibility of endless replication, removes the limitations of growth and provides the ability to avoid being terminated, resulting in what is commonly known as a tumour. Figure A1 was first published in the article “The hallmarks of cancer” by Hanahan and Weinberg (2000) and has since been cited by many. It identifies six important characteristics of cancer cells which all possible targets for the development of new cancer drugs.

[22] Somatic cells are all cells of the human body except for stem cells and gametes (=spermatozoa and ova).

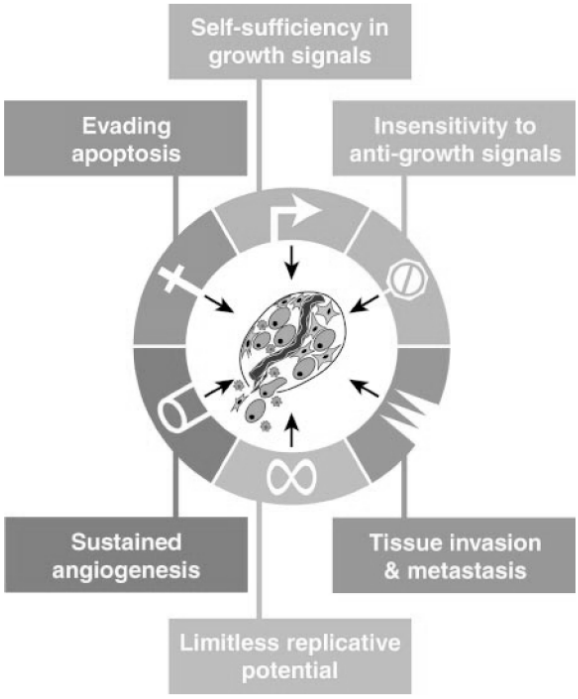

Figure A1 *The six hallmarks of cancer. Apted from Hanahan and Weinberger (2000).*

A2 Survival, incidence, prevalence and mortality

The burdens of cancer can be measured in many ways. Statistics on patient survival is important as reference material. It provides a basis for evaluation and decision making for patients, clinicians and scientists and should for that reason be as up-to-date as possible. Traditional methods for analyzing survival have important limitations with regard to up-to-dateness of long-term cumulative survival estimates. This is also true for survival statistics on cancer patients. To reduce these limitations Herman Brenner and Olaf Gefeller in 1996 presented a new approach on estimating cancer survival called “Period analysis” (Brenner, Gefeller 1996 and 1997). Period analysis reduces the time lag of the cancer survival estimates by some 5–10 years (Talbäck et al 2004). In 2006 an advanced and even more up-to-date method of computing survival statistics at minimal loss of precision called “Model-based period analysis” was introduced (Brenner, Hakulinen 2006). The method provides 5-year relative survival estimates deduced from the aggregated survival data available at the end of each year, i.e. the 5-year relative survival for patients diagnosed during 2004 can be calculated with very good precision already at the end of the year. Consequently 5-year relative survival enables survival rates to be as up-to-date as patients, doctors and scientists can possibly wish for.

Except for survival, terms like incidence, prevalence and mortality are often discussed. Incidence is the rate at which a disease occurs, i.e. the number of new cases of the disease occurring during a certain period in a population at risk. Prevalence is the number of cases of a specific disease present in a given population at a certain time, which is the same as the aggregated incidence. Mortality is the death rate, often referred as the ratio of actual deaths to expected deaths of those not having cancer. It is estimated that in 2002 there was 11 million cases of cancer reported world wide resulting in 7 million deaths. Mortality from cancer is second only to cardiovascular diseases. The cancer cases were distributed over the world according to Figure A1. The most common malignancies were lung, breast, colorectal and stomach cancer followed by prostate and liver cancer. The most commonly used measure of the burden of cancer, developed by WHO and the World Bank, is 'disability adjusted life years' (DALYs). This is a sum of the years (in full health) lost due to mortality and disability. Measured in this way, close to 10 million DALYs were lost to cancer in EU25 in 2002 (Jönsson and Wilking 2007, pp. iii8-9).

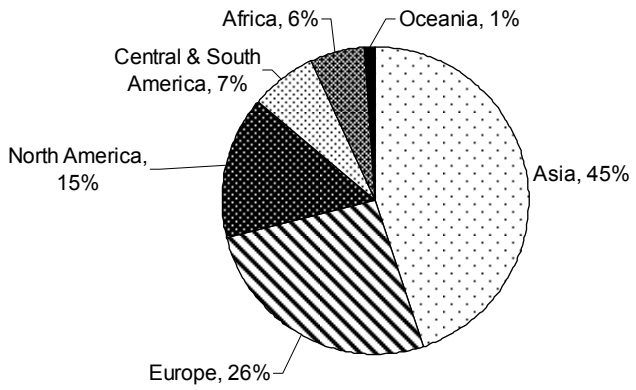

Figure A2 *Geographic distribution of worldwide cancer incidence in 2002 of the estimated total of 11 million cases (Jönsson and Wilking 2007, p. iii8).*

A3 Cost of cancer

The cost of cancer to society can be separated into direct and indirect costs. The total cost, i.e. the sum of direct and indirect costs is vast. Direct costs are the health care costs such as prevention and treatment. Those are changing over time but several studies estimate them somewhere between 4.1 to 9.0 % of total health care budgets for industrialized countries (Jönsson and Wilking 2007, piii10). Europe alone spent more than € 50 billion in 2004 on direct cancer costs (Jönsson and Wilking 2007, piii10). Indirect costs, mainly due to

productivity losses due to inability to work and mortality before 65 years of age (Jönsson and Wilking 2007, piii9), are estimated to account for 70-85 % of total costs (WHO 2004).

A4 Prevention.

The best way to decrease the suffering due to cancer is of course to discover it at an early stage of the disease or even better to avoid it in the first place. Hopefully this will be accomplished one day through the ever ongoing research on screening techniques and on agents (factors) causing cancer. One of the latest breakthroughs is the introduction of vaccines against human papillomaviruses (HPV), causative agent (ignition) of cervical cancer in women (Villa et al 2005).

Someday in the future we might fully understand and learn what causes cancer. Still it would be naïve to believe that equipped with this knowledge we would be able to eradicate the disease. Even though it has been known for over 50 years that smoking cause cancer people still start smoking. In Sweden for example, an increasing percentage of female smoking has resulted in a doubled incidence of female lung cancer since 1980. On a European level preventions are estimated to be able to lower mortality about nine percent over the next 15 years (Boyle et el 2003). Understanding that prevention will never eradicate the problem of cancer, means that the treatment of cancer will always remain an important cornerstone in fighting the disease.

A5 The treatment

In general terms treatment of cancer can be divided in to 3 broad fields; surgery, radiation therapy and drug therapy.

Surgery was for several centuries the only way of treating cancer. Developments of cancer surgery have followed in the footsteps of general surgery with the introduction of ether anaesthesia (1846) and antiseptic technique (1867). One of the first properly described surgery techniques for a certain cancer site was the gastric resection by Billroth (1881). Surgery remains to this date one of the most important tools in diagnosing, treating, giving symptom relief as well as curing cancer.

The field of radiotherapy was laid open by Wilhelm Röntgen's discovery in 1895. Diagnostic x-ray was developed and spread over the western world within three months. Treatment wise x-ray was mainly used against benign skin conditions in the early years. In Sweden,

experiences of x-ray cancer treatment were reported in 1899 and those were later internationally published in 1910. While x-ray remains an important screening technique, radiotherapy today is a highly developed clinical speciality using ionised radiation for treatment and curing of cancer patients.

While surgery and radiotherapy remain good quality treatments to relief large parts of a patients tumour burden and cure approximately 50 percent of the patients, the hope of actually being able to cure the remaining 50 percent is now days to a great extent set to cancer drug treatment. Cancer drug therapy as we know it today actually has its roots in the weapon industry development of mustard gas after World War I. As the properties of the cell toxic gas were understood, scientists realised its potential and developed it further to 'nitrogen mustard', the first actual chemotherapy drug. Since then, great progress has been made and cancer drug therapy today can be divided into chemotherapy, endocrine[23] therapy and biotherapy. Chemotherapeutic drugs are cell toxic and are given in cocktails of several drugs effecting different mechanisms in the cancer cell. Chemotherapeutic drugs can further be divided after mechanism and group into alkylating agents and platinum substances, topoisomerase inhibitors, anti-metabolites, mitosis inhibitors and others.

Endocrine therapy was first described during late 19th century. Important progress has been made during the last decades through the discovery of new effective drugs such as Tamoxifen, LHRH-antagonists and aromatase inhibitors. The introduction of endocrine therapy represents a step away from highly toxic agents to well-defined molecular targets. For some cancer sites, such as prostate-, thyroid- and corpus cancer, endocrine therapy has greater effect than chemotherapy; for breast and prostate cancer equal; and for leukaemia it serves as an important complement to chemotherapy.

Biotherapy is often referred to as the fourth patway and acts through the cancer patients own immune system. Monoclonal antibodies, i.e. antibodies with the identical binding sites to certain targets, can either affect cells by them selves or by a substance being conjugated[24] to the antibody. The latter can be viewed up on as a target guided weapon, where an already known cell toxic substance is guided to its target by the antibody. bo This way has some side

[23] endocrine = hormonal
[24] conjugated = attached to

effects, all attributed to the carried substance. The unconjugated monoclonal antibody acts by binding to certain receptors on the cancer cell activating its apoptosis or by inducing a response in the patients own immune system, activating it towards the cancer cells. The upside of unconjugated monoclonal antibodies is that they have very few side effects.

Appendix B – Sample data

Table B1: Cancer registry data

Selected countries	Cancer registries	Approximate population coverage (%)
Estonia	National	100
Finland	National	100
Germany	Saarland	1.3
Italy	Torino, Tuscany	4.0*
Netherlands	Eindhoven	6.6
Norway	National	100
Poland	Cracow	1.3
Slovenia	National	100
Switzerland	Geneva	5.3
UK	Scotland	8.5
USA	SEER	30

*2 registries together

Table B2: Key molecules

Molecule	**Initial launch date**
Anastrozole	Sep 1995
Bevacizumab	Feb 2004
Capecitabine	May 1998
Carboplatin	Dec 1985
Cetuximab	Dec 2003
Cisplatin	Dec 1978
Docetaxel	Apr 1995
Doxorubicin	Jan 1971
Epirubicin	Apr 1984
Erlotinib	Nov 2004
Exemestane	Nov 1999
Gemcitabine	Jun 1995
Imatinib	May 2001
Irinotecan	Apr 1994
Letrozole	Nov 1996
Oxaliplatin	Jul 1996
Paclitaxel	Dec 1992
Rituximab	Nov 1997
Topotecan	Jul 1996
Trastuzumab	Oct 1998
Vinorelbine	Jun 1989