Comparing survival from cancer using population-based cancer registry data – methods and applications

Xue Qin Yu

Doctor of Philosophy Thesis

The University of Sydney

June 2007

Certificate of Originality

This thesis is submitted to the University of Sydney in fulfilment of the requirement for the Degree of Doctor of Philosophy.

The work presented in this thesis is, to the best of my knowledge and belief, original except as acknowledged in the text. I hereby declare that I have not submitted this material, either in full or in part, for a degree at this or any other institution.

Signature:… *Date: …15 June 2007……….*

Acknowledgements

I completed this thesis while employed as a Statistician at the Cancer Epidemiology Research Unit (CERU) at The Cancer Council New South Wales (NSW). This thesis would not have been completed without the help and support of many people. Firstly, I would like to thank my primary supervisor, Professor Bruce Armstrong, for his support and for patiently guiding me through the challenging 5 years of my PhD. I greatly value his crucial input at every stage of my thesis, without which it would have remained unfinished. In addition, I wish to thank my associate supervisors, Associate Professors Robert Gibberd and Dianne O'Connell, for their statistical expertise, guidance and encouragement throughout the whole 5 years. I am extremely grateful to all three of my supervisors for their time, kindness and insight.

I would also like to thank additional co-authors of the four papers included in this thesis – Dr Paul Dickman for technical support in applying the new statistical approach he developed for relative survival modelling, David Smith for familiarising me with the data I have used and supporting me in its early use and for being a good friend and colleague for so many years, Professor Alan Coates for providing constructive criticisms and clinical insights into the paper in the fourth chapter, and Professor Michal Abrahamowicz for providing comprehensive comments on aspects of analysis and interpreting the results in the paper in the fifth chapter.

In addition, I am indebted to the reviewers and editors of the journals that published the 3 papers in this thesis. Permissions to reprint either in print or electronic form have been granted by publishers of these articles, Springer Science + Business Media, Elsevier Ltd, John Wiley & Sons Inc respectively. My thanks go to Dr Elisabeth Heseltine, a renowned writer & teacher of scientific writing, for proofreading the preface and Chapters 1 and 7 of this thesis.

My sincere thanks also go to the CERU team at The Cancer Council NSW for their understanding of my cultural differences. Thanks also go to the NSW Central Cancer Registry for providing the data for my thesis, and the NSW Colorectal Cancer Care Survey for providing the survey data for the analysis in Chapter 6.

Most of all I wish to thank my wife, my daughter and my son who put up with me for more years than they care to remember. Their love and support has been unwavering, and for that I thank them from the bottom of my heart.

Structure of the thesis

Three published manuscripts and one accepted manuscript subject to minor revision form the body of this thesis. Two chapters precede them. The first is an introductory chapter, which contains brief statements describing the research problems investigated, overall objectives and specific aims of the thesis. It includes a brief account of the usefulness of survival analysis using population-based data in cancer control and in evaluating performance of cancer services both temporally and spatially.

The second contains a review of the literature on geographical and temporal variation in survival using population-based cancer registry data, which sets the scene for the following chapters. I also critically evaluated the limitations and strengths of the studies in this field in relation to the following chapters.

The third chapter is our paper "Yu XQ, O'Connell DL, Gibberd RW, Smith DP, Dickman PW, Armstrong BK. Estimating regional variation in cancer survival: a tool for improving cancer care. *Cancer Causes Control* 2004; 15:611-8". The chapter mainly describes the methods for estimating regional variation in cancer survival and use of these estimates to identify cancers with the greatest potential for improving care outcomes with prioritisation of actions targeted against such cancers.

The fourth chapter is our paper "Yu XQ, O'Connell DL, Gibberd RW, Armstrong BK. A population-based study from New South Wales, Australia 1996-2001: area variation in survival from colorectal cancer. *Eur J Cancer* 2005; 41:2715-21". The paper describes an application of the methods described in the third chapter to colorectal cancer, gives a more complete analysis of regional variation in colorectal cancer survival and explores the possible reasons for the survival difference among regions in New South Wales, Australia.

The fifth chapter is our paper "Yu XQ, O'Connell DL, Gibberd RW, Coates AS, Armstrong BK. Trends in survival and excess risk of death after a diagnosis of cancer in 1980 to 1996 in New South Wales Australia. *Int J Cancer* 2006; 119:894-900". In this paper, I examined time trends in survival for patients diagnosed with any of 28 cancers between 1980 and 1996 in NSW, with adjustment for disease spread at diagnosis and cancer histology.

The sixth chapter is our paper "Yu XQ, O'Connell DL, Gibberd RW, Abrahamowicz M, Armstrong BK. Misclassification of colorectal cancer stage and area variation in survival accepted by *Int J Cancer* subject to minor revision". In this paper, I examined the accuracy of spread of cancer at diagnosis recorded in the registry database by comparing it with the staging information for the same patients collected from a patterns of care survey, and then assessed the impact of measurement error on estimates of regional variation in stage-specific survival, an extension of Chapter 4.

The last chapter provides a broad discussion and conclusions drawn from the work described in the thesis. I broadly discuss the usefulness of survival analysis based on cancer registry data, before and after these publications, and bring all the findings together, including identifying cancer types or geographical areas with extremely poor performance requiring prompt action to improve outcomes of patient care.

In addition, there are sections with acknowledgements, an abstract, the author's contribution and an appendix that lists publications related to the thesis and other publications arising from thesis-related work.

Abstract

Over the past decade, population-based cancer registry data have been used increasingly worldwide to evaluate and improve the quality of cancer care. The utility of the conclusions from such studies relies heavily on the data quality and the methods used to analyse the data.

Interpretation of comparative survival from such data, examining either temporal trends or geographical differences, is generally not easy. The observed differences could be due to methodological and statistical approaches or to real effects. For example, geographical differences in cancer survival could be due to a number of real factors, including access to primary health care, the availability of diagnostic and treatment facilities and the treatment actually given, or to artefact, such as lead-time bias, stage migration, sampling error or measurement error. Likewise, a temporal increase in survival could be the result of earlier diagnosis and improved treatment of cancer; it could also be due to artefact after the introduction of screening programs (adding lead time), changes in the definition of cancer, stage migration or several of these factors, producing both real and artefactual trends. In this thesis, I report methods that I modified and applied, some technical issues in the use of such data, and an analysis of data from the State of New South Wales (NSW), Australia, illustrating their use in evaluating and potentially improving the quality of cancer care, showing how data quality might affect the conclusions of such analyses.

This thesis describes studies of comparative survival based on population-based cancer registry data, with three published papers and one accepted manuscript (subject to minor revision).

In the first paper, I describe a modified method for estimating spatial variation in cancer survival using empirical Bayes methods (which was published in *Cancer Causes and Control* 2004). I demonstrate in this paper that the empirical Bayes method is preferable to standard approaches and show how it can be used to identify cancer types where a focus on reducing area differentials in survival might lead to important gains in survival.

In the second paper (published in the *European Journal of Cancer* 2005), I apply this method to a more complete analysis of spatial variation in survival from colorectal cancer in NSW and show that estimates of spatial variation in colorectal cancer can help to identify subgroups of patients for whom better application of treatment guidelines could improve outcome. I also

show how estimates of the numbers of lives that could be extended might assist in setting priorities for treatment improvement.

In the third paper, I examine time trends in survival from 28 cancers in NSW between 1980 and 1996 (published in the *International Journal of Cancer* 2006) and conclude that for many cancers, falls in excess deaths in NSW from 1980 to 1996 are unlikely to be attributable to earlier diagnosis or stage migration; thus, advances in cancer treatment have probably contributed to them.

In the accepted manuscript, I described an extension of the work reported in the second paper, investigating the accuracy of staging information recorded in the registry database and assessing the impact of error in its measurement on estimates of spatial variation in survival from colorectal cancer. The results indicate that misclassified registry stage can have an important impact on estimates of spatial variation in stage-specific survival from colorectal cancer. Thus, if cancer registry data are to be used effectively in evaluating and improving cancer care, the quality of stage data might have to be improved.

Taken together, the four papers show that creative, informed use of population-based cancer registry data, with appropriate statistical methods and acknowledgement of the limitations of the data, can be a valuable tool for evaluating and possibly improving cancer care. Use of these findings to stimulate evaluation of the quality of cancer care should enhance the value of the investment in cancer registries. They should also stimulate improvement in the quality of cancer registry data, particularly that on stage at diagnosis. The methods developed in this thesis may also be used to improve estimation of geographical variation in other count-based health measures when the available data are sparse.

Author’s contribution

The author conducted the work presented in this thesis under the supervision of three supervisors: Professor Bruce Armstrong and Associate Professors Robert Gibberd and Dianne O’Connell. I was the major contributor to all aspects of the study: planning the research, literature review, data management, analysis and statistical modelling, interpretation of results, writing of manuscripts for peer-reviewed journals and writing the thesis.

Table of contents

Chapter 1 Introduction

About this chapter

This chapter provides the background for my thesis. First, I describe the aims of this thesis. In the following 'Background' section, I provide a brief overview of the use of population-based cancer registry data in evaluating the quality of cancer care in other countries to provide the context for the following chapters. In this section, I discuss the strengths and limitations of population-based cancer registry data and survival analysis and their usefulness in evaluating and potentially improving the quality of cancer care. In addition, I describe the research problems investigated in the following chapters and methods for overcoming issues in the use of population-based data in evaluating the quality of cancer care and the justification for doing this research.

Thesis aims

My primary aim in conducting the research contained in this thesis is to examine the usefulness of population-based cancer registry data in evaluating and potentially improving the quality of cancer care. In addition, this thesis has two secondary aims. The first is to explore and apply appropriate statistical methods to overcome the limitations of such data in measuring health service performance, such as sampling errors, confounding factors, lead-time bias, and measurement errors. The second is to assess the impact of registry data quality on estimates of geographical or temporal differences in survival.

Background – Survival is an important measure of patient care

Although the efficacy of cancer therapy is best evaluated by studying patient outcomes in randomised clinical trials, the patients in most clinical trials have generally not been followed long enough for long-term survival to be evaluated. Studies using population-based cancer registries allow for much longer follow-up. Moreover, in contrast to a clinical trial, in which patients might be excluded on the basis of disease stage at diagnosis, age or co-morbid illness, studies using population-based registry data represent the experience of a general population of people with cancer. They thus represent the full spectrum of care that patients receive or fail to receive during and after diagnosis, not care from particular hospitals or cancer therapy centres. More importantly, the results of clinical trials might differ from actual treatment outcomes in everyday practice.[1] Population-based data can provide insights into the practical effectiveness of treatments that cannot be obtained from trials. However, there are some disadvantages in using population-based cancer registry data for measuring health service performance since cancer registry data are primarily used to estimate cancer incidence rates. Most cancer registries do not collect treatment and co-morbidity information for the patients, and cancer stage at diagnosis is also often not collected. Even for registries that do collect this information, the data quality is not as good as in clinical trials: some data may not be accurately recorded or may be missing altogether. Despite these limitations, I argue that studies using population-based registry data are worth doing because of their representativeness and generalisability.

During the past decade, population-based cancer registry data have been used worldwide to assess progress in cancer management and to compare cancer outcomes between population subgroups and geographical areas. Such studies generally focus on broad patterns of cancer care, with a special interest in quality and outcome of care and geographical, ethnic, social and temporal variation. There is ongoing interest in understanding the causes of such variation and in developing effective interventions to reduce the frequency of inappropriate or poor-quality care and, in some instances, to reduce inequality in care or outcome between population sub-groups.

Cancer registries are recognised as being among the best of all public health surveillance systems with regard to completeness, accuracy, data availability and timeliness of reporting. The data are a ready source for addressing some questions related to the quality of cancer care, which is currently a high priority for cancer-related health services research. Because of their legal mandate, cancer registries are uniquely situated to identify a population-based sample of incident cancer patients and are the best source for measuring the quality of cancer care in whole populations.[2] In the past, much of the information collected was used to study cancer epidemiology and the causes of cancer; but was underutilised to measure the performance of cancer care.

Survival can be used to quantify the effectiveness of early detection and treatment at the population level and is thus an important component in monitoring cancer control. Registry data often include the stage, histopathology, vital status and some patient descriptors. Patient survival, therefore, can be used as the basis for quality measures particularly if there is the capacity for risk adjustment. The aims of such studies include: monitoring the global effect of diagnostic and treatment improvements, establishing priorities for healthcare investment and research, estimating the potential for further improvement, planning clinical trials, and carrying out studies on the equity of the health service.

Relative survival is the most widely used method of survival analysis based on population-based cancer registry data. It is usually defined as a measure of excess mortality due to a diagnosis of cancer; that is, it is "corrected" for background mortality.[3] It compares the survival of cancer patients to that of the general population in which they belong. The advantage of relative survival is that no information on cause of death is required; this information is either unavailable or inaccurate for many cancer registries.[4;5] Even if cancer registries routinely collect cause of death information, it is difficult to separate causes "entirely due to cancer" from causes "completely unrelated to cancer" and the "shades of grey" between them. Thus, relative survival provides a more objective and possibly more accurate means of removing the effect of mortality from other causes.[6]

Although there has been increased use of population-based cancer registry data for cancer-related health services research, since the work for this thesis began, only a few studies from Europe have been published that systematically address methods for overcoming the limitations of such data, and none have been published in Australia. To the best of my knowledge, there was no published study in the literature that both quantified the measurement error in disease stage at diagnosis, a critical factor in survival, and examined its impact on temporal or geographical area variation in survival. There was, therefore, a need to address these issues using data from population-based cancer registries, especially paying attention to the quality of the data and to reduce the measurement error on estimated geographical variation or temporal trends.

In this thesis, I report on the methods modified to address the above issues and explore the technical issues involved. I illustrate these methods for evaluating and potentially improving the quality of cancer care, and also show how data quality can affect the conclusions drawn from such analyses. Chapter 3 is primarily about dealing with sampling errors by using the empirical Bayes method, and the use of stage adjustment is aimed at addressing a confounding factor and lead-time bias. In Chapter 4, I provide more detailed analyses to address these issues and interpretation of the NSW Cancer Registry data. Chapter 5 provides comprehensive analyses of trends in survival by taking lead-time bias and stage migration into account. Stage migration is dealt with by careful interpretation of the results of survival over time, trends in incidence and changes in stage distribution over time. In Chapter 6, I address an aspect of measurement error of disease stage at diagnosis recorded on the registry database as well as other issues dealt with in Chapters 3 and 4. In Chapter 7, I bring all the findings from this thesis together and discuss what the thesis as a whole contributes to new knowledge about the usefulness of survival analysis based on cancer registry data.

Chapter 2 Literature Review

About this chapter

In this chapter I have provided a detailed review of previous studies of survival from cancer using population-based cancer registry data. The studies included in this review are studies of geographical variation in survival and trends in survival published since 1996. I chose to limit the review to the most recent studies because they are more relevant to this thesis.

Survival variation between geographical areas

Studies of geographical variation in cancer survival since 1996 have examined numerous sites, including breast,[7-20] colon and rectum,[7-12;21-28] lung,[7-12;16;22;29-31] prostate,[7-12;16;22;32] melanoma,[8;12;16;22;23] stomach,[7;10;12;16;22;23;33] cervix,[8;12;16;22] ovary,[7;10;12;16] and non-Hodgkin lymphoma (NHL),[8;12;16] and in many countries. Some such studies focused on a single cancer type with an in-depth analysis while others included selected, frequently occurring cancers, providing an overview of cancer survival across different cancer types. The relevant studies included in this review are summarised in Table 1.

Among those studies including multiple types of cancers, EUROCARE is the first large-scale international project to compare cancer survival across a number of different countries, using data from population-based cancer registries.[34-37] The EUROCARE database included 4 million cancer cases from 65 cancer registries in 20 European countries. The main results of its analyses revealed wide international differences in survival among cancer patients diagnosed in 12 European countries between 1978 and 1985,[37] and among patients diagnosed in 17 European countries between 1985 and 1989.[38] Similar kinds of studies comparing survival from multiple cancer types between several geographical areas have been conducted in Australia,[8] Canada,[9] Nordic countries,[22;23] Scotland[7] and the USA.[10] Generally these studies included the more frequently occurring cancers, such as breast, colon and rectum, lung and prostate cancers and melanoma.

For those studies focusing on a single cancer type, the most studied cancer types were breast and colorectal cancer because these two cancers affect large numbers of people, and have an important impact on the population. Moreover, due to the availability of methods for early diagnosis and effective treatments if detected early, geographical variation in survival is likely to exist. The data from the EUROCARE study highlighted wide differences in survival from breast cancer between European countries.[14;15] These studies are based on large numbers of cases (119,139 and 145,000 respectively) across 12-17 countries, where the cancer registries supplied uniform data, thus the results are representative of large geographical areas in Europe. Similar results for breast cancer were also found in England, Scotland and the USA. Using the English National Cancer Registry data, Mullee et al[39] reported that for 78,904 women diagnosed with breast cancer between 1992 and 1994, large and significant variation was found in 5-year relative survival (from 66% to 85%) among 99 health authorities ($p<0.001$). Twelves et al[20] analysed data on 1617 breast cancer patients diagnosed in 1987 in Scotland and found that Health Board was an important determinant of patient survival at both 5 years ($p=0.04$) and 10 years ($p=0.004$) after adjusting for age, clinical stage, oestrogen receptor status, pathological node status and pathological tumour size. Using the Surveillance, Epidemiology, and End Results (SEER) registry data, Goodwin et al[13] found significant variation in survival across the 66 health service areas ($p<0.0001$) in the older (aged 65 and older) US population of women diagnosed with breast cancer in 1985-1991.

All these studies on breast cancer survival, among different countries with large study populations, suggest that the effect of place of residence is probably real. What are the causes of such variation? There are a variety of possible explanations but two are most likely: stage at diagnosis and efficacy of treatment. To separate these two and to analyse their effect on survival between geographical areas, reliable data on disease stage and treatment are required. However, the variation in survival in the Scottish study remained significant after taking account of differences in stage between health boards, which led them to believe that the most likely explanation was differences in local or adjuvant systemic treatment between Health Boards.[20] The EUROCARE high resolution study is an extension of the EUROCARE study with additional collection of information on disease stage, staging procedures, and treatments for a representative sample of cases from 17 European cancer registries. Using these data Sant et al[17] concluded that stage at diagnosis is a key explanation of differences in breast cancer survival across Europe. In this study, the authors used multiple regression models to assess the impact of disease stage on 5-year relative survival for breast cancer in 1990-92 between 17 cancer registries in 6 European countries. The first model, only adjusting for age, did not substantially change the relative survival pattern. After additional adjustment for stage and surgery, the relative excess risk (RER) of death in most regions moved towards 1, indicating that stage was a key determinant of regional variation in survival. A further movement of the RER towards

1 occurred after adjustment for the number of lymph nodes examined. However, the fact that the RER remained higher in 3 regions suggests that the management of breast cancer patients is not optimal in those areas.[17] Applying the known effective treatment to all sub-groups in the population would, therefore, reduce the geographical variation in breast survival, as suggested by Goodwin et al.[13]

Geographical variation in survival for colorectal cancer is also well studied. Studying variation in survival among 5147 colorectal cancer cases diagnosed in 1991-95 in the Wessex region, south England, Kim et al[26] concluded that district of treatment showed a highly significant relationship to survival after adjusting for stage and surgery type. Gatta et al[24] reported that 5-year relative survival from colorectal cancer differed significantly across European countries, ranging from 23% to 59% for colon cancer and 22% to 53% for rectal cancer. Using the data from the EUROCARE high resolution study, Gatta et al[25] found that survival from colorectal cancer varied markedly between European countries and there was a twofold range in the risk of death from this cancer even after adjustment for surgery and disease stage at diagnosis. Differences in survival from cancers of the colon and rectum are especially marked in the first 6 months after treatment, suggesting that there are effects from stage at diagnosis and/or access to optimal care.[25] However, Prior et al[27] and Woodman et al[40] disagreed with the view and concluded that the observed survival differentials may not be due to differences in the quality of care but may reflect misdiagnosis or a failure of some European registries to register all patients with advanced disease. In a study comparing survival between European and US patients diagnosed with colorectal cancer in 1990-91 using data from 10 European and 9 SEER registries, Ciccolallo et al[21] concluded that US-Europe survival differences in colorectal cancer were large but seem to be mostly attributable to differences in stage at diagnosis. After adjusting for age, sex, site, stage, resection and number of lymph nodes examined, the relative excess risks of death due to colorectal cancer were similar in 5 of the 10 European registries when compared to that of the US, cases in one registry had a lower relative excess risk than US patients.[21]

Other studies of geographical variation in survival for single cancers were for lung, head and neck, prostate and stomach cancers. This group of cancers are studied less often because they have very poor prognosis (lung and stomach) and are less likely to show any geographical differences in survival or results relating to them are difficult to interpret because of lead-time bias and length bias (prostate). Berrino et al[41] reported significant differences in relative survival for head and neck cancer patients diagnosed in 1985-89 between Eastern European countries and the rest of Europe, and the differences persisted after correcting for the distribution of cancer subsite. Using data from 45 European cancer registries, Post et al[32] reported that 5-year relative survival from prostate cancer varied markedly from 40% to 72% among European countries. For lung cancer, the survival was very poor but age-standardised relative survival varied significantly within Europe, with 5-year relative survival for patients diagnosed in 1985-89 being between 5% and 12% for males and 7% and 18% for females.[30] Similar results were also found in Denmark[31] and England.[29] Using Danish national cancer registry data, Madsen et al[31] observed that survival from lung cancer was dependent on the place of residence and after adjusting for age and stage, 5-month survival differed significantly between regions. Cartman et al[29] also found significant variation in one-year survival from lung cancer between health authority districts in England and this variation remained after adjusting for age at diagnosis.

Studies comparing survival between geographical areas involving multiple cancer types were conducted in Australia,[8] Canada,[9] Italy,[11] Nordic countries,[22;23] Scotland[7] and the USA,[10] and Europe,[34;35;37;38] including cancers of the colon and rectum,[7-12;16;22;23] breast,[7-12;16;22;23] prostate,[7-12;16;22;23] cervix,[8;12;16;22] stomach,[7;10;12;16;22;23] ovary,[7;10;12;16;22] and melanoma,[8;12;16;22;23] and other cancer sites.[8;10;12;16;22;23] In the report "Cancer survival in Australia 1992-1997", men living in rural and remote areas were found to have poorer 5-year relative survival for all cancers, lung and prostate cancer and melanoma; women in those areas had poorer survival for lung and cervical cancers, compared with the metropolitan areas and large rural centres.[8] Similarly in Scotland, Campbell et al[7] found increasing distance from a cancer centre was associated with poorer survival and patients living in more remote areas were less likely to be diagnosed before they died, especially for stomach, breast and colorectal cancers. Part of the reason for this may be due to differences in stage at diagnosis between rural and urban areas, as increasing distance from a cancer centre was found to be associated with a high chance of disseminated disease at diagnosis (p=0.04).[42]

In Canada, Ellison et al[9] found that the survival rates of Canadians with cancer depended on in which region of the country they lived. They reported that on average, Canadian women with breast cancer have an 82 percent chance of surviving five years after diagnosis. This rose to 85 percent for women living in British Columbia, but dropped to 76 percent in Newfoundland and Labrador. The same pattern was repeated for men with prostate cancer.[9]

Gatta et al[11] analysed population-based data on survival of cancer and compared survival between geographical areas in Italy. The data were from three cancer registries in Italy with 90,431 cases. They found that there were significant differences in survival between different areas of the country, particularly for cancers that respond well to treatment (eg. cancers of the breast, prostate and large bowel).[11] They thought that the unequal provision of care might explain the differences in survival between northern-central Italy and the south. Regional variation in cancer survival has been reported in the USA,[10] including cancers of the stomach, colon, rectum, lung, breast, uterus, and prostate. Significant variation (p=0.01) in survival among SEER areas was observed for all individual cancers studied except for cancer of the ovary (p=0.04). After adjusting for stage, significant variation remained except for cancers of the ovary and bladder.[10]

Intercountry differences in survival were also reported for many European countries, and between Europe and the USA. Using data from cancer registries in the Nordic countries, Engeland et al[23] reported that Danish patients had a markedly lower relative survival than patients in other Nordic countries for cancers of the stomach, colon and rectum, breast and prostate. Sant et al[16] calculated 5-year relative survival for 1,836,287 patients diagnosed with one of 13 cancers between 1978 and 1989, using data from 20 cancer registries in 13 European countries. They found large disparities in survival between countries for most solid tumours (large bowel, breast, kidney, cervix, ovary, prostate and stomach, and melanoma of the skin). Generally survival was highest in Northern Europe and lowest in Eastern Europe, and was also low in the UK and Denmark. Less marked regional variation was found for the lymphomas and regional variation was not observed for cancers of the lung or brain.[16] Analysing data from 738,075 European and 282,398 US patients for each of the 12 major cancers diagnosed in 1985-1989, Gatta et al[12] compared survival in Europe and the United States and found that Europeans had significantly lower survival than US patients for most cancers; greater differences were found for cancers of the prostate, colon, rectum, breast, uterus and melanoma. The survival differences were small for lung cancer and lymphomas, and no significant differences in survival were observed for stomach cancer.[12]

Table 1. Summary of studies examining geographical variation in cancer survival

First author (year)	Cancer type(s)	Population & setting	Analysis	Description of results
Australian Institute of Health and Welfare (2003)[8]	All cancers combined, colorectal, lung, melanoma, breast, cervix, prostate and NHL	National cancer statistics clearing house data: Australia	Relative survival	Men living outside metropolitan areas and large rural centres had poorer 5-year survival for all cancers, lung and prostate cancer and melanoma; women in those areas had poorer survival for lung and cervical cancer
Berrino (1998)[41]	Head and neck	EUROCARE II	Relative survival	Significant differences in survival for head and neck cancer patients diagnosed in 1985-1989 were observed between Eastern countries and the rest of Europe and the differences persisted after correcting for differences in the distribution of sub-site
Blomqvist (1997)[43]	Colon	Swedish National Cancer Registry	Relative survival for six 3-year periods from 1973 to 1990 and multivariate analysis	Relative survival from colon cancer improved substantially from 1973 to 1990 and differences in survival between areas decreased during the study period; the authors think that the convergence of survival mainly can be attributed to improvements in care provided in the primary catchment areas of local and county hospitals
Campbell (2000)[7]	Lung, colorectal, breast, prostate, stomach, ovary	Scottish Cancer Registry data	Cox regression	Increasing distance from a cancer centre was associated with poorer survival for colorectal, breast and stomach cancers after adjusting for age, sex and settlement size
Cartman (2002)[29]	Lung	Northern and Yorkshire Cancer Registry data: UK	Kaplan-Meier survival and Cox regression	One-year survival varied significantly from 23% to 19% among 4 groups of health authority districts and this difference was still present after adjusting for age at diagnosis
Ciccolallo (2005)[21]	Colon and rectum	10 EUROCARE registries and 9 SEER registries in USA	Relative survival modelling	3-year relative survival was 69% for USA, 57% for European patients; after adjusting for age, sex and site, relative excess risk was significantly higher than the USA in all EUROCARE registries from 1.07 to 2.22
Dickman (1997)[22]	12 common cancers	National Cancer Registry data: Denmark, Finland, Norway, Sweden	Relative survival modelling	Significant regional variation in survival was found for 9 of the 12 cancers studied, including cancers of the breast, lung, colon, rectum and prostate, and melanoma (detailed results for regional variation were presented for Sweden only)
Eaker (2005)[44]	Breast	One clinical breast cancer register in Sweden	Relative survival and Poisson regression	7-year relative survival was lower in county A compared with several other counties; this difference decreased after adjusting for diagnostic activity and after county A began to strictly adhere to the regional breast cancer guidelines these differences disappeared
Ellison (2001)[9]	Breast, colorectal, lung and prostate	Canadian Cancer Registry data	Age-standardised relative survival	Age-standardised relative survival differed significantly among provinces for prostate, breast and male lung cancer, but not for colorectal cancer

First author (year)	Cancer type(s)	Population & setting	Analysis	Description of results
Engeland (1998)[23]	Stomach, colon, rectum, breast, uterus, prostate, melanoma	National Cancer Registry data: Denmark, Finland, Norway, Sweden, Iceland	Relative survival	Danish patients had a markedly lower relative survival than the patients in other Nordic countries for cancers of the stomach, colon, rectum, breast, and prostate
Farrow (1996)[10]	Stomach, colon, rectum, lung, breast, uterus, ovary, prostate, bladder	9 SEER registries USA	Relative survival, Cox regression model	Significant variation in survival among SEER areas was observed for all cancers studied (p<0.01), except ovary (p=0.04). After adjusting for stage, variation remained significant except for cancers of the ovary and bladder
Gatta (1998)[24]	Colorectal cancer	EUROCARE data from 17 European countries	Relative survival	5-year relative survival from colorectal cancer differed significantly across European countries, ranging from 23% to 59% for colon cancer and 22% to 53% for rectal cancer
Gatta (2000a)[25]	Colorectal cancer	11 EUROCARE cancer registries	3-year observed survival, Cox model	3-year survival ranged from 25% to 59% across 11 population registries. After correction for stage, significantly different survival still existed between populations, with relative risk from 0.76 to 1.81
Gatta (2000b)[12]	12 major cancers	41 EUROCARE cancer registries and 9 SEER registries in USA	Relative survival	Europeans had significantly lower survival than US patients for most cancers; greater differences were for cancers of the prostate, colon, rectum, breast, uterus and melanoma
Goodwin (2002)[13]	Breast	SEER data: USA	Kaplan-Meier survival	Kaplan-Meier survival curves indicated significantly different survival among the 66 health service areas (p<0.0001)
Janssen-Heijnen (1998)[30]	Lung	44 EUROCARE cancer registries	Age-standardised relative survival	Age-standardised relative survival varied greatly within Europe, with 5-year relative survival from 5% to 12% in males and from 7% to 18% in females
Kim (2000)[26]	Colon and rectum	South West Cancer Intelligence Unit: England	Observed survival and Cox regression model	Variation in survival from colorectal cancer across districts in southern England persisted after adjusting survival rates for age, stage and surgery type
Madsen (2002)[31]	Lung	National Cancer Registry data: Denmark	Cox regression model	Survival was dependent on place of residence. After adjusting for age & stage 5-month survival differed significantly between areas
Mullee (2004)[39]	Breast	National Cancer Registry data: England and Wales	Relative survival, regression model	Large and significant variation was found in 5-year relative survival (from 66% to 85%) among 99 health authorities (p<0.001)
Post (1998)[32]	Prostate	Data from 45 EUROCARE registries	Relative survival	5-year relative survival varied markedly from 40% to 72% among European countries
Prior (1998)[27]	Colon	North Western Regional Cancer Registry: England	Relative survival	Relative survival for the North West cohort, based on the exclusion of cases with only a clinical diagnosis, was very close to that of the European cohort, suggesting the disadvantage in survival in the UK found in the EUROCARE study may be due to failure to record all advanced cases in some European registries

First author (year)	Cancer type(s)	Population & setting	Analysis	Description of results
Quinn (1998)[14]	Breast	Data from 42 EUROCARE registries	Relative survival	5-year relative survival in 1985-1989 differed widely among countries, the lowest being 58% in Slovakia and highest being 81% in Sweden
Sant (1998)[15]	Breast	Data from 25 EUROCARE registries	Relative survival	5-year relative survival varied from 61% to 78% across European countries
Sant (2001)[16]	13 common cancers	Data from 20 EUROCARE registries	Relative survival	Survival was highest in Northern Europe and lowest in Eastern Europe, and also low in the UK and Denmark. The differences between 4 European regions decreased over time for cancers of the large bowel, breast, and melanoma and Hodgkin's disease
Sant (2003)[17]	Breast	Data from 17 EUROCARE registries	Regression model for relative survival	5-year relative survival varied significantly from 66% to 86% across 9 European countries. After adjusting for stage and number of lymph nodes examined, survival differences between countries were greatly reduced
Sant (2004)[18]	Breast	Data from SEER registry and 17 EUROCARE registries	Regression model for relative survival	5-year relative survival was 89% in USA and 79% in Europe. Significant difference was found after adjusting for age at diagnosis and surgery, but disappeared after full adjustment for stage
Spilsbury (2005)[19]	Breast cancer patients with surgery	Western Australia Cancer Registry data linked with hospital records	Relative survival (RSR) and Cox regression	Living in regional areas was associated with poorer survival ($p<0.001$), but residential location was not associated with survival after adjusting for treatment and health-related factors
Twelves (2001)[20]	Breast	Scottish Cancer Registry data	Kaplan-Meier survival, Cox regression model	Significant variation in the risk of death across health boards at 5 years ($p=0.04$) and at 10 years ($p=0.004$)
Verdecchia (2004)[33]	Gastric cancer	Data from 47 EUROCARE registries	Regression model	Significant differences in survival persisted among European countries after adjusting for age, subsite and histology

There tends to be greater geographical variation in survival when the comparison is between countries with large differences in political system and/or level of economic development (Table 2). For example, large and significant intercountry variation was found in the EUROCARE project for most cancers studied, with eastern European countries having lower survival generally compared with the average European rates.[16] When comparing countries with similar political and economic development, the variation seems to be smaller. For example, for most cancers, the variation in survival between the Nordic countries was smaller than the differences found in the EUROCARE comparisons.[23] Furthermore, variation in survival is less likely when the comparison is between areas in the same country. For example, the authors did not find significant geographical variation in colorectal cancer survival in either Australia or Canada,[8;9] while wide and significant variation was found within Europe.[16;24;25] On the other hand, as found in the EUROCARE studies,[34] geographical variation in survival tends to be smaller[16;25] or even disappears[8] for lymphomas, probably because of the relative lack of available effective therapies, compared with many other cancers.

For most cancers, survival differences between regions are probably due to variation in earliness of detection or diagnosis, or to regional differences in access to care and quality of care delivered. Numerous studies have provided scientific evidence of the relationships between treatment and outcomes, such as cancer specific survival or cancer recurrence. Supportive evidence for these explanations comes from recent detailed studies of breast and colorectal cancer, which showed strong evidence that the application of effective treatments reduced the differences in survival between geographical areas (colon cancer in Sweden)[43] and that regional differences in survival disappeared after adjusting for differences in treatment for breast cancer in Australia and Sweden.[19;44] In a population-based study comparing survival differences in breast cancer between counties in different time periods in Sweden, Eaker and colleagues[44] found that 7-year relative survival for breast cancer was significantly lower in one county in 1992-1993, but that this survival difference had disappeared in 2000-2002 after this county began, in 1996-97, to adhere strictly to the regional breast cancer guidelines, including recruiting new staff members to the multidisciplinary breast team, active quality assurance of mammography screening and stricter adherence to guidelines for treating breast cancer patients. In an Australian study comparing survival after breast cancer surgery, Spilsbury et al[19] found that women living in regional areas had poorer 5-year relative survival ($p<0.001$) than their counterparts in metropolitan areas, however, residential location was no longer associated with survival after adjusting for treatment, and health-related factors.

Table 2. Summary of results of studies examining geographical variation in survival by cancer type

Cancer type	First author (year)	Extent of geographical variation in survival observed
Breast	Australian Institute of Health and Welfare (2003)[8]	Small but significant variation: 5-year relative survival (RSR): 79.9% in other rural areas (lowest) vs 82.8% national average
	Campbell (2000)[7]	Strong evidence of variation observed: after adjusting for age, sex and settlement size, increasing distance from a cancer centre was associated with poorer survival - odds ratio (OR) ranging from 1.0 in ≤5km to 2.9 ≥38 km (p for trend <0.001)
	Dickman (1997)[22]	Significant variation observed (p<0.01): with high number of deaths that could be prevented (284) (for Sweden only)
	Ellison (2001)[9]	Significant variation observed with 5-year RSR ranging from 76% (lowest) to 85% (highest) by province
	Engeland (1998)[23]	5-year RSR was markedly lower (69%) in Denmark than other Nordic countries (74% and 79% for Norway and Finland respectively)
	Farrow (1996)[10]	Significant regional variation observed: relative risk of deaths (RR) ranged from 0.77 to 1.19 across SEER areas (p<0.01)
	Gatta (2000b)[12]	Large differences in survival observed: Europeans had a higher (1.75) relative excess risk of death (RER) than people in the USA
	Goodwin (2002)[13]	Considerable variation was observed among 66 health service areas (p<0.0001)
	Mullee (2004)[39]	Large and significant variation observed: 5-year RSR ranged from 66% to 85% among 99 health authorities (p<0.001)
	Quinn (1998)[14]	Wide variation observed: 5-year RSR ranged from 56% to 83% across 13 European countries for patients diagnosed in 1987-89
	Sant (1998)[15]	Significant variation observed: 5-year RSR ranged from 61% to 78% across 11 European countries for women diagnosed in 1983-85
	Sant (2001)[16]	Wide and significant variation observed: 5-year RSR ranged from 55.8% in Eastern Europe to 80.5% in Northern Europe for patients diagnosed in 1987-89
	Sant (2003)[17]	Wide and significant variation observed: 5-year RSR varied from 66% to 85% across 9 European countries
	Sant (2004)[18]	Significant variation observed: 5-year RSR was 89% in the USA and 79% in Europe for women diagnosed in 1990-92
	Spilsbury (2005)[19]	Significant variation observed: living in regional areas was associated with poorer survival (p<0.001)
	Twelves (2001)[20]	Significant variation in the risk of death across health boards at 5 years (p=0.04) and at 10 years (p=0.004)
Colon and rectum	Australian Institute of Health and Welfare (2003)[8]	Variation was not significant between geographical areas for both sexes
	Campbell (2000)[7]	Evidence of variation observed: after adjusting for age, sex and settlement size, increasing distance from a cancer centre was associated with poorer survival - OR ranging from 1.0 in ≤5km to 1.8 ≥38 km (p for trend = 0.02)
	Ciccolallo (2005)[21]	Significant variation observed between Europe and the USA: 3-year RSR was 69% for USA and 57% for Europeans
	Dickman (1997)[22]	Significant variation observed (p<0.01): with high number of deaths that could be prevented (382) (for Sweden only)
	Ellison (2001)[9]	Little inter-provincial variation was observed for colorectal cancer

Cancer type	First author (year)	Extent of geographical variation in survival observed
Colon and rectum **(continued)**	Engeland (1998)[23]	Denmark had lowest 5-year RSR among the Nordic countries since the mid 1970s
	Farrow (1996)[10]	Significant regional variation observed: RR ranged from 0.89 to 1.11 for colon cancer and 0.92 to 1.12 for rectal cancer across SEER areas (p<0.01)
	Gatta (1998)[24]	Large and significant variation observed within Europe: 5-year RSR ranged from 23% to 59% for colon cancer and 22% to 53% for rectal cancer
	Gatta (2000a)[25]	Wide and significant variation observed within Europe: 3-year observed survival ranged from 25% to 59% across 11 populations covered by registries
	Gatta (2000b)[12]	Large differences in survival observed: Europeans had a higher (1.5) RER than patients in the USA diagnosed in 1985-89
	Kim (2000)[26]	Highly significant variation between districts observed: district of treatment was a significant determinant of survival (p<0.0001)
	Sant (2001)[16]	Wide and significant variation observed: 5-year RSR ranged from 23% in Eastern Europe to 52% in Northern Europe for patients diagnosed in 1987-89
Lung	Australian Institute of Health and Welfare (2003)[8]	Significant variation in 5-year RSR observed: men living in small rural centres (8.0%) and other rural areas (8.5%) had significantly lower survival than the national average (10.5%)
	Campbell (2000)[7]	Evidence of variation observed: after adjusting for age, sex, deprivation and settlement size, increasing distance from a cancer centre was associated with poorer survival – hazard ratio (HR) ranged from 1.0 in ≤5km to 1.1 ≥38 km (p for trend = 0.02)
	Cartman (2002)[29]	Wide variation in survival observed: one-year survival was significantly better in the districts with highest rates of active treatment (23%) compared with 19% for those with lowest treatment rates
	Dickman (1997)[22]	Significant variation observed (p<0.01): with 249 deaths that could be prevented (for Sweden only)
	Ellison (2001)[9]	Significant variation observed with 5-year RSR for men ranging from 8% (lowest) to 15% (highest) by province; for women the variation was not significant
	Farrow (1996)[10]	Significant regional variation observed: RR ranged from 0.95 to 1.11 for lung cancer across SEER areas (p<0.01)
	Gatta (2000b)[12]	Significant differences in survival observed: Europeans had a higher RER (1.2) than patients in the USA diagnosed in 1985-89
	Janssen-Heijnen (1998)[30]	Wide variation observed: 5-year RSR varied greatly within Europe from 5% to 12% in males and 7% to 18% in females
	Madsen (2002)[31]	Significant variation observed: after adjusting for age and stage, 5-month survival varied significantly between regions
Prostate	Australian Institute of Health and Welfare (2003)[8]	Significant variation observed: men living in 'capital cities' and 'metropolitan areas' had significantly better survival, and men in 'rural and remote areas' had significantly poorer survival compared with the national average
	Campbell (2000)[7]	Evidence of variation observed: after adjusting for age, sex, deprivation and settlement size, increasing distance from a cancer centre was associated with poorer survival – HR ranged from 1.0 in ≤5km to 1.2 ≥38 km (p for trend = 0.04)
	Dickman (1997)[22]	Large and significant variation observed (p<0.01): with 390 deaths that could be prevented (for Sweden only)

Cancer type	First author (year)	Extent of geographical variation in survival observed
Prostate (continued)	Ellison (2001)[9]	Large and significant variation observed with 5-year RSR for men ranging from 67% (lowest) to 91% (highest) by province
	Engeland (1998)[23]	Large and significant variation observed: with poorer 5-year RSR (41%) in Denmark compared with 56% in Norway and 59% in Finland
	Farrow (1996)[10]	Significant regional variation observed: RR ranged from 0.83 to 1.18 for prostate cancer across SEER areas (p<0.01)
	Gatta (2000b)[12]	Large and significant differences in survival observed: Europeans had a higher RER (2.8) than patients diagnosed in the USA in 1985-89
	Post (1998)[32]	Large and significant variation observed: 5-year RSR varied markedly from 40% to 72% among European countries
	Sant (2001)[16]	Wide and significant variation observed within Europe: 5-year observed survival ranged from 43% to 66% across 11 populations covered by registries
Melanoma	Australian Institute of Health and Welfare (2003)[8]	Significant variation observed: men living in 'other rural areas' had poorer 5-year RSR than the national average, and variation was not significant for women
	Dickman (1997)[22]	Large and significant variation observed (p<0.01): with 408 deaths that could be prevented (for Sweden only)
	Engeland (1998)[23]	No marked differences were observed between the Nordic countries
	Gatta (2000b)[12]	Large and significant differences in survival observed: Europeans had a higher RER (1.8) than patients in the USA diagnosed in 1985-89
	Sant (2001)[16]	Wide and significant variation observed within Europe: 5-year observed survival ranged from 62% to 88% across 11 populations covered by registries
Stomach	Campbell (2000)[7]	Strong evidence of variation observed: after adjusting for age, sex and settlement size, increasing distance from a cancer centre was associated with poorer survival - OR ranging from 1.0 in ≤5km to 3.9 ≥38 km (p for trend <0.001)
	Dickman (1997)[22]	Small but significant variation observed (p<0.01): with 74 deaths that could be prevented (for Sweden only)
	Engeland (1998)[23]	Moderate variation observed: with poorer 5-year RSR (10%) for males in Denmark compared with 16% in Norway and 18% in Finland; 5-year RSR was similar for females
	Gatta (2000b)[12]	Little difference in survival observed: Europeans had a higher (21%) 5-year RSR than patients in the USA diagnosed (19%) in 1985-89
	Sant (2001)[16]	Wide and significant variation observed within Europe: 5-year observed survival ranged from 12% to 24% across 11 populations covered by registries
	Verdecchia (2004)[33]	Significant variation observed: survival varied significantly across European countries after adjusting for age, subsite, histology and disease stage at diagnosis
Cervix	Australian Institute of Health and Welfare (2003)[8]	No significant differences in 5-year RSR were observed across geographical areas
	Dickman (1997)[22]	No significant regional variation in 5-year RSR was observed (p=0.60) (for Sweden only)
	Gatta (2000b)[12]	Moderate but significant differences in survival observed: Europeans had a lower (61.8%) 5-year RSR than patients in the USA (66.1%) diagnosed in 1985-89
	Sant (2001)[16]	Significant variation observed within Europe: 5-year observed survival ranged from 49% to 65% across 11 populations covered by registries

Cancer type	First author (year)	Extent of geographical variation in survival observed
Ovary	Campbell (2000)[7]	No significant differences observed: distance from a cancer centre was not associated with survival (p for trend = 0.26)
	Farrow (1996)[10]	No significant regional variation observed: RR ranged from 0.94 to 1.07 for ovarian cancer across SEER areas after adjusting for age ($p>0.05$)
	Gatta (2000b)[12]	Moderate but significant differences in survival observed: Europeans had a lower (32.9%) 5-year RSR than patients in the USA (39.5%) diagnosed in 1985-89
	Sant (2001)[16]	Significant variation observed within Europe: 5-year observed survival ranged from 26% to 42% across 11 populations covered by registries
NHL	Australian Institute of Health and Welfare (2003)[8]	No significant differences in 5-year RSR were observed across geographical areas
	Gatta (2000b)[12]	Small but significant differences in survival observed: Europeans had a lower (46.7%) 5-year RSR than patients in the USA (50.3%) diagnosed in 1985-89
	Sant (2001)[16]	Regional variation was less marked for the lymphomas: 5-year observed survival ranged from 45% to 51% across 11 populations covered by registries
Hodgkin's disease	Gatta (2000b)[12]	Small but significant differences in survival observed: Europeans had a lower (71.7%) 5-year RSR than patients in the USA (74.9%) diagnosed in 1985-89
	Sant (2001)[16]	Regional variation was less marked for the lymphomas: 5-year observed survival ranged from 66% to 72% across 11 populations covered by registries

RSR – relative survival **RR – relative risk** **RER – relative excess risk of death** **HR – hazard ratio** **NHL – non-Hodgkin lymphoma**

Trends in survival over time

Monitoring trends in patient survival is a useful approach for assessing performance in diagnosing and treating cancer patients over time. Cancer registries that record all new cases in a well-defined population represent the only way to assess real changes in the management of cancer at the population level. Patient survival rates provide useful information for doctors, patients and policy-makers, although estimates must be interpreted with care. The literature shows (Table 3) that survival from most cancers has improved since the 1980s in Canada,[45] England,[46] Singapore,[47] Sweden[48] and the USA.[49] EUROCARE is a large study designed to reliably compare population-based cancer survival across different countries in Europe. Comparing EUROCARE 2 and 3, the European average survival was found to have increased from the mid 1980s to 1992-94 for most major cancers.[28] The increases in 5-year relative survival were marked for prostate cancer (55% to 68%), and substantial for breast cancer (70% to 77%), colorectal cancer (44% to 52%) and melanoma (68% to 78% for men and 82% to 88% for women), but only moderate for lung cancer (7.5% to 9.2% for men and 8.1% to 9.8% for women).[28] Across the Atlantic in the USA, improvement in 5-year relative survival was also seen for most cancers between 1975-79 and 1995-2000.[49] The cancers with large gains in survival were cancers of the prostate, colon and rectum, breast and kidney, and NHL, while improvement was limited for cancers of the lung, pancreas, liver and oesophagus.[49] Similar results were also found in Canada,[45] England,[46] Singapore,[47] Sweden[48] and Switzerland.[50] These studies did not, however, consider the effect of stage differences over time in their comparisons of survival over time. Therefore, there has been debate over the extent to which improved treatment has contributed to the trend in survival.[51-53]

Although there are a variety of possible explanations for such an improvement, two of the most likely are earlier detection and improved treatment. Increasing survival over time may reflect improvements in early detection by screening programs, improved treatment, or a combination of the two. Generally, one can separate the effect of earlier diagnosis from the effect of more effective treatment by comparing stage-specific survival or overall survival controlling for stage. If stage-specific survival stays the same over time, then better survival can be interpreted as a result of a more favourable stage distribution rather than more effective treatment.

A few studies have taken account of changes in stage when assessing survival trends over time. A systematic analysis of patient survival over 40 years was conducted using Finnish Cancer Registry data.[54] Also there have been studies of individual cancers of the breast[55-58]and colon and rectum.[59-62] All these studies found that increasing survival was partially independent of trends in stage over time. Thus authors interpreted the overall survival trends as being due to both earlier diagnosis and improvement in the management of cancer.

There are some good examples of substantial improvements in survival that have probably been achieved by advances in treatment. In a study comparing survival from rectal cancer from 1960 to 1989 in Sweden, Dahlberg et al[63] found that 5-year relative survival rates were very similar across regions in Sweden between 1960 and 1984; however significantly better survival was seen in one county compared to the rest of Sweden between 1985 and 1989. They believed that this improvement was a result of the introduction of combined preoperative radiotherapy and total mesorectal excision (TME) in 1985 in this county while not in the rest of Sweden.[63] A more recent study supported this: as the improved surgical technique and preoperative radiotherapy was gradually adopted in the rest of Sweden since the early 1990s, the survival differences seen in 1985-89 disappeared.[64] In an Australian study based on registry data linked with hospital records, authors observed a 15% increase in relative survival from ovarian cancer between the periods of 1982-87 (38.8%) and 1994-98 (53.5%), and they thought that the improved outcome was a result of a dramatic shift to the more aggressive, cytoreductive surgery in Western Australia over the past 20 years.[65] In an English study, Reeves and colleagues[66] studied survival trends from 1972 to 1984 in people aged less than 35 years diagnosed with leukaemia, Hodgkin's disease, and testicular cancer using data from the National Cancer Registry in England. They found that 5-year relative survival for childhood leukaemia and testicular cancer improved dramatically from the mid 1970s and to a lesser extent for Hodgkin's disease. They argued that these improvements were mainly due to the introduction of new therapeutic modalities, cisplatin-based chemotherapy for testicular cancer and chemotherapeutic protocols together with increased centralisation of care for childhood leukaemia.[66]

However, a few studies had reported that the improved survival in the most recent periods was due mainly to a more favourable stage distribution. Using data from a specialised colorectal cancer registry in Italy from 1984 to 1997, Ponz de Leon and colleagues[67] found that 5-year relative survival was significantly more favourable in 1990-91 than it was in 1984-89; but this upward trend was associated with a sharp increase in newly detected localised tumours in the more recent period. Thus, they concluded that the improved survival from colorectal cancer may be attributed to several concomitant factors, such as wider use of colonoscopy, increased education of patients and more attention given to symptoms.[67] Another study in Northern Ireland examined changes in survival from melanoma of the skin in two 5-year periods: 1984-88 and 1994-98.[68] McMullen and co-workers found that patients diagnosed in the 2nd period had a one-third lower risk of dying than those in the 1st period. However, after adjustment for Breslow thickness and ulceration, the favourable survival in the 2nd period became non-significant (relative hazard=0.88, 95% confidence intervals (CI):

0.69-1.13). They concluded that the improved survival in the 2nd period was mainly due to detecting thinner melanoma and melanoma with less ulceration rather than advances in treatment.[68]

Table 3. Summary of studies examining trends in cancer survival over time

First author (year)	Cancer type(s)	Population & setting	Analysis	Description of results
Angell-Andersen (2004)[59]	Colon and rectum	National cancer registry: Norway	Relative survival	50,993 cases diagnosed with colorectal cancer between 1958 and 1997 were included in the analysis. 5-yr relative survival increased 3% per 5-yr diagnostic period, from 40% in 1958-62 to 56% and 60% in 1993-97 for males and females respectively.
Birgisson (2005)[64]	Colon and rectum	Swedish cancer registry	RSR over 8 5-yr periods from 1960 to 1999	5-yr RSR improved significantly from 39.6% in 1960-64 to 57.2% in 1995-99 for colon cancer and from 36.1% to 57.6% for rectal cancer respectively; with the largest improvement of survival being in later periods studied
Chia (2001)[47]	17 common cancers	National cancer registry: Singapore	RSR over 5 5-yr periods from 1968 to 1992	Significant 5-yr RSR increases were seen for 9 cancers, including cancer of the colon-rectum, breast, cervix and NHL; no changes in survival over 25 years for cancers of the liver and pancreas
Coleman (2004)[46]	20 common cancers	National cancer registry: England	RSR over 3 5-yr periods from 1986 to 1999	5-yr RSR increased for 11 of the 16 cancers for men, including colon, rectum, prostate and melanoma, not for cancers of the lung, pancreas and brain; for women 9 of the 17 cancers studied were increased significantly from 1986 to 1999 including breast, colon, rectum, ovary, but not for lung, pancreas, cervix and brain
Dahlberg (1998)[63]	Rectum	Swedish cancer registry	RSR and multivariate analyses	5-yr RSR were very similar across regions in Sweden between 1960 and 1984; however, significantly better survival was seen in the county of Uppsala than in the rest of Sweden between 1985 and 1989 as a result of the introduction of combined preoperative radiotherapy and total mesorectal excision (TME) in 1985 in Uppsala which was not widely performed in the rest of Sweden.
Dickman (1999)[54]	Over 30 selected cancer types including colon, rectum, lung, breast, prostate and melanoma	Finnish cancer registry	Relative survival over four 10-year periods from 1955 to 1994	A gradual increase in 5-, and 10-yr relative survival from 1955-64 to 1985-94 was observed for almost all individual cancers with the exception of cervical cancer. Compared with the 1975-84 period 5-yr relative survival in 1985-94 improved for patients with major cancers, except for cancers of the pancreas, liver, lung & cervix
Du (2002)[60]	Colon and rectum	National cancer registry: Singapore	Stage-specific survival by 5 5-yr periods from 1968 to 1992 and Cox regression	5-yr RSR improved greatly over the study period for both colon and rectal cancer. After adjusting for age, sex, ethnic groups and clinical stage, period of diagnosis remained a significant factor for survival; the hazard ratio was 0.47 and 0.40 for the most recent period for colon and rectal cancer respectively

First author (year)	Cancer type(s)	Population & setting	Analysis	Description of results
Ellison (2004)[45]	Breast, lung, colorectal and prostate	Canadian Cancer Registry	Change in 5-year relative survival over time	Between 1985-87 and 1992-94, increases in 5-year relative survival were dramatic for prostate cancer, large for breast cancer, and smaller for colorectal cancer, but little change for lung cancer
Engel (2002)[69]	Ovary	Munich cancer registry: Germany	Kaplan-Meier (KM) survival curve	Overall, the 5-year and 10 year stage-specific survival for those diagnosed in 1988-1997 improved only slightly compared with those diagnosed in 1978-1987
Faivre-Finn (2002)[61]	Colon	Cote-d'Or cancer registry: France	Logistic regression	Survival improved over the study period with 5-yr relative survival being 33.0% (1976-79), 40.9% (1980-83), 52.5% (1984-87), 58.9% (1988-91) and 55.3% (1992-95), after adjusting for age, stage, subsite and emergency surgery or not, period of diagnosis was strongly related to survival (p<0.0001)
Jensena (2003)[55]	Breast	National cancer registry: Denmark	Kaplan-Meier survival curves	An overall improvement in survival was seen in two of the 3 counties studied with the 3rd one experiencing no change in survival.
Laurivick (2003)[65]	Ovary	Western Australian cancer registry linked with hospital records	Relative survival and logistic regression	A 15% increase in relative survival was observed between 1982-87 (38.8%) and 1994-98 (53.5%) with a significant increase in surgical procedures over the same period
Levi (2000)[50]	29 common cancers and all cancers combined	Vaud cancer registry: Switzerland	RSR over 4 periods of diagnosis	5-yr RSR for all cancers increased from 51% in 1974-78 to 64% in 1989-93; most cancer sites showed an increase in 5-yr RSR, including colon, rectum, lung, breast, prostate and melanoma
Martijn (2003)[62]	Rectum (diagnosed from 1980 to 2000)	Eindhoven cancer registry: the Netherlands	Relative survival and Cox regression for 3 periods	5-yr relative survival increased from 49% in 1980-89 to 55% in 1990-94, and 61% in 1995-2000; after adjusting for age, gender, subsite and stage, significant improvements in the overall prognosis were found for patients aged less than 60 years and those 60-74 years old, but not for those aged 75 or over
McMullen (2004)[68]	Melanoma	Northern Ireland cancer registry	KM survival curves and Cox regression	5-yr survival increased from 71.0% in 1984-88 to 77.4% in 1994-98; after allowing for age and sex, patients diagnosed in the 2nd period had a 30% lower risk of dying than those diagnosed in the 1st period (RH=0.71). However, after additional adjustment for Breslow depth and ulceration, the favorable survival in the recent period became non-significant (RH=0.88 95% CI: 0.69-1.13)
Ponz de Leon (2000)[67]	Colon and rectum	Modena colorectal cancer registry: Italy	KM survival curves	Survival was significantly higher in patients diagnosed in 1990-91 for both colon (p=0.004) and rectal cancer (p=0.02) than in 1984-89

First author	Cancer type(s)	Population &	Analysis	Description of results
Reeves (1999)[66]	Breast, testicular cancer, Hodgkin's disease, acute lymphatic leukaemia, acute myeloid leukaemia	National cancer registry data: England (only a subset of those 15-34 years old but for leukemia those aged 0-14 years also included)	Change in relative survival over 4 periods between 1972 and 1984	5-year RSR for childhood leukaemia improved from 9% in 1972-75 to 34% in 1982-84 for AML, whereas for ALL it increased from 45% to 75% over the same period; for those aged 15-34 years, 5-year RSR for AML increased from 6% in 1972-75 to 31% in 1982-84, figures for ALL showed an increase from 22% to 44% over the same period; 5-year RSR for testicular cancer also showed a gradual and significant increase; while the increase for Hodgkin's disease was moderate. There was little change for breast cancer
Sant (2003)[28]	12 groups of common cancers	EUROCARE-3 data	RSR comparison between 1983-85 and 1992-94	Overall European 5-yr RSR increased from 1983-85 to 1992-94 for the major cancers: lung (7.5% - 9.2% for men; 8.1% - 9.8% for women), colorectum (44% - 52%), melanoma (68% - 78%), breast (70% - 77%), prostate (55% - 68%)
Spilsbury (2005)[19]	Breast cancer patients with surgery	Western Australian cancer registry data linked with hospital records	RSR over 4 4-yr periods from 1982 to 1997	5-yr RSR from breast cancer after surgery increased significantly ($p<0.001$) from 77% in 1982-85 to 86% in 1994-97. After adjusting for age, hospital type, comorbidity and treatment type, period of diagnosis remained significant
Talback (2003)[48]	All cancers combined, oesophagus, colon, rectum, lung, breast, cervix, uterus, ovary, prostate, melanoma etc 40 cancer types	Swedish cancer registry	5, 10, 15, 20 year relative survival from 1960 to 1998	This study included 1,021,421 patients diagnosed with one of the 40 cancers in Sweden in 1960-1998. During the 1990s substantial survival improvements were observed not only for uncommon cancers, such as testicular cancer, Hodgkin's lymphoma, but also for cancer of the rectum, kidney and melanoma. Survival for breast and cervical cancer also improved during the 1990's, but not for pancreatic, liver and lung cancers
Taylor (2003)[56]	Breast	NSW state cancer registry: Australia	RSR over 5 5-yr periods	5-yr RSR improved over the 5 periods of diagnosis from 1972 to 1996, especially from the late 1980s onwards
Thomson (2004)[57]	Breast: diagnosed in 1987 and 1993	Scottish cancer registry	Kaplan-Meier survival curve and Cox regression for both cohorts	Significant improvements in survival was seen for those diagnosed in 1993, after adjustment for all the clinical/pathological factors the risk of death was still lower in 1993 ($p=0.03$); with additional adjustment for all healthcare delivery the cohort effect remained significant ($p=0.03$); screening appeared to make the biggest contribution among all the healthcare & demographic factors
Webb (2004)[58]	Breast	Case-control study: Queensland, Australia	Relative survival	5-yr RSR was better among women diagnosed in 1990-94 (84%) than those diagnosed in 1981-84 (74%), after adjusting for tumour size and nodal status, this difference reduced but still persisted

Methodological issues for interpreting temporal or geographical variation

Temporal and geographical variation in survival can be due to several factors, including the quality of treatment received. It may also be due to other factors, such as variation in the extent of disease at diagnosis and in the histology and grade of tumours, or artefacts in the data, such as stage migration. Therefore, interpretation of comparative survival, either time trends or geographical differences based on population-based cancer registry data, is generally not easy. For example, temporal or geographical differences in cancer survival could be attributed to a number of real factors, including access to primary health care, the availability of diagnostic and treatment facilities and the treatment actually given, or it might simply be due to statistical artefacts (lead-time bias, confounding factors, stage migration, sampling error or measurement error). Alternatively a combination of several of these factors may produce both real and artefactual differences.

Lead-time and length bias

If a cancer is detected by screening, the survival time is increased by an amount known as lead time,[70] without necessarily postponing the time of death. For example, in a comparison of survival between Europe and the USA using EUROCARE and SEER registry data, Gatta et al[12] found that higher survival in the USA patients for breast, colon and rectum, and prostate cancers coincided with a higher incidence of these cancers, thus they thought this was likely due to the inclusion of more early stage small tumours identified by screening in the USA cohort.

Length-bias occurs when slow-growing, less aggressive cancers with good prognosis are preferentially detected by screening. An extreme form of length-bias is overdiagnosis; that is the detection of tumours that would never have caused harm to the patient if they had not been screen-detected. The large uptake of prostate-specific antigen (PSA) testing, starting around

1990, in the USA led to escalating incidence of prostate cancer[71] and corresponding survival increases. This coincidence suggests that the apparent increase in survival since the early 1990s was mainly due to length bias (and probably some lead-time bias), because the patients diagnosed with PSA are substantially 'diluted' with many nonfatal cases. The same was found in Europe when studying survival trends for prostate cancer.[28;72] Thus, differences in screening rates across geographical areas or over time could affect comparisons of survival rates between regions or time trends.

Stage-adjusted or stage-specific estimates of survival may be helpful in understanding temporal trends in survival, but they should be interpreted with caution. Because some cancers, such as breast, colorectal and prostate, are commonly diagnosed through screening programs, survival for these cancers is susceptible to the effects of lead-time bias and length bias. These effects may lengthen apparent survival time without providing benefit to the patient. Theoretically, stage adjustment should reduce the magnitude of lead-time bias but does not eliminate it because there is a within stage shift.[73] Such adjustment, however, has little or no effect on length bias[73] because through screening, a totally new class of cancers is diagnosed with a different behaviour (ie no capacity to kill) than tumours detected otherwise in the same stage categories.

Stage migration

Stage migration[74] can bias geographical or temporal comparisons of survival. In studies of trends in survival, it can produce an apparent improvement in stage-specific or stage-adjusted survival as a result of the introduction of more sensitive staging tools. That is, patients with otherwise clinically silent metastatic lesions are reclassified to more advanced stages because these tools find at least some of these silent metastases. This shift from earlier to more advanced stages appears to improve survival in both earlier and advanced stages but results in no overall improvement for the entire cohort. For example, in a German study examining trends in survival for ovarian cancer from 1978 to 1997, Engel et al[69] found that survival improved about 10% for early stages but survival for the total sample hardly changed. They thought that stage migration may have contributed to the apparent improvements in survival from early stage tumours since classification was more accurate in the recent period.[69]

By the same token, in studies of geographical differences in survival, stage migration might affect the comparability of stage-specific or stage-adjusted survival estimates because tumours classified as “localized” in one area are possibly “more localized” than tumours assigned the same stage in other areas. In a study aimed at improving the interpretation of survival differences between Europe and US women with breast cancer, Sant and colleagues[18] used multiple regression models to assess the impact of stage and staging procedure on survival estimates for women diagnosed in 1990-92. They found that the differences in relative excess risk gradually reduced from model 1, only adjusting for age at diagnosis, model 2 with additional adjustment for surgery, to model 3 with further adjustment for stage at diagnosis, but the difference was still significant (RER for EUROCARE = 1.12, 95% CI: 1.03-1.22). They used the adjustment for number of lymph nodes examined pathologically to partially address the problem of stage migration, and found that after this adjustment in the final model the differences in RER reduced further (RER for EUROCARE=1.07, 95% CI: 0.98-1.17) and became not statistically significant.[18] Thus, stage migration must be considered in interpreting survival differences temporally or geographically.

Sampling error

When survival rates are calculated for areas with small populations, the estimates can be affected by sampling error, reducing the precision of estimates for individual areas. Sampling error observed in areas with small populations can also mask true regional variation in cancer survival. For example, in a study of fire and burn-related mortality rates by counties in the USA, 77 (25%) counties with high rates had five or fewer observed deaths over the entire 9 year study period and the authors believed that the high rates in many of these counties were likely due to sampling error rather than true elevation in underlying risk.[75] Similarly when comparing survival for non-common cancers the number of cases would be small in sparsely populated areas, and addition or omission of one or two cases would cause drastic changes in their estimates. Consequently, use of such estimates can introduce errors in decision-making for health service planning.

Measurement error

Disease stage at diagnosis is considered the most important prognostic factor in most cancers. Therefore, any comparison of outcome between populations or at different times must be adjusted for stage. To disentangle the effect of early diagnosis with the effect of treatment in explaining geographical or temporal differences in survival, it is essential to have reliable data on cancer stage. Measurement error in stage would be expected to reduce our ability to control for the effects of earlier detection and thus, may lead to incorrect inferences about the effects of treatment. In a recent study examining variability in axillary lymph nodes dissection for breast cancer among 19 hospitals in the Netherlands, Schaapveld et al[76] found that a more extensive surgical dissection or pathological examination of the specimen resulted in a higher number of positive nodes, and false negative axillary staging, due to low level axillary dissection or lymph nodes examined in some hospitals, would influence the risk of recurrence and hence survival of individual breast cancer patients.[77]

Other kinds of measurement error may also be important in some cancer registries. They may include, incomplete ascertainment of incident cases with, for example, a bias towards ascertainment of more advanced disease (if, for example, many cancers are first detected from death notifications);[40] inaccurate diagnosis of the type of cancer;[27] and failure to ascertain all deaths that have occurred in the cohort of cancer patients[78] (eg, through passive follow-up by linkage to death registries, which will miss deaths in people who have migrated from the area covered by the deaths data base and deaths that fail to link because of incompleteness or error in the identifiers in either database). In a more recent study assessing the impact on survival estimates of incomplete ascertainment of cancer cases and cases identified through death certificate only (DCO), Robinson et al[79] found that there were large effects of incompleteness and DCO cases on estimated survival figures. They recommended that future studies should take variation in the proportion of DCO and completeness into consideration when comparing survival between different populations.[79]

In summary, this literature review demonstrated that survival analysis using population-based cancer registry data can be used in monitoring and evaluating the diagnosis and treatment of cancer at the population level. There was wide variation in survival from many cancers among European countries, SEER areas in the USA, and areas in Canada and Australia. Patient survival has improved since the 1980s for most cancers in most countries included in this review. However, there are many limitations of such studies and researchers using such data should be aware of them when assessing them, analysing the data and interpreting the results.