Socioeconomic inequalities in cancer survival in the Netherlands and Great Britain

small-area based studies using cancer registry data

Carola Schrijvers

Financial support for the printing of this thesis was obtained from the Department of Public Health, Erasmus University Rotterdam and the Comprehensive Cancer Centre South, Eindhoven.

CIP-GEGEVENS KONINLIJKE BIBLIOTHEEK, DEN HAAG

Schrijvers, Cornelia Theodora Maria

Socioeconomic inequalities in cancer survival in the Netherlands and Great Britain/Cornelia Theodora Maria Schrijvers.
Thesis Rotterdam - With ref. - With summary in Dutch
ISBN 90-9009028-0
Subject headings: survival / cancer / socioeconomic / prognostic factors / The Netherlands / Great Britain

Cover: Anne-Lore Kuryszczuk
Printer: Elinkwijk B.V., Utrecht

Socioeconomic inequalities in cancer survival in the Netherlands and Great Britain

small-area based studies using cancer registry data

Sociaaleconomische verschillen in kanker overleving in
Nederland en Groot-Brittannië
studies gebaseerd op kankerregistratie gegevens

Proefschrift

Ter verkrijging van de graad van doctor
aan de Erasmus Universiteit Rotterdam
op gezag van de Rector Magnificus,
Prof.dr. P.W.C. Akkermans, M.A.
en volgens het besluit van het college voor promoties.

De openbare verdediging zal plaatsvinden op
woensdag 17 januari 1996 om 13.45 uur

door

Cornelia Theodora Maria Schrijvers

geboren te Oss

Promotiecommissie

Promotor:	Prof. dr. J.P. Mackenbach
Co-promotor:	Dr. J.W.W. Coebergh
Overige leden:	Prof. dr. A.F. Casparie
	Prof. dr. M.P. Coleman
	Prof. dr. J.A. van Dongen

Contents

Chapters 2-6 are based on the following papers and manuscripts:

2 Schrijvers CTM, Mackenbach JP. Cancer patient survival by socioeconomic status in seven countries: a review for six common cancer sites.
J Epidemiol Community Health 1994;48:441-446[1]

4 Schrijvers CTM, Stronks K, Mheen H van de, Coebergh JWW, Mackenbach JP. Validation of cancer prevalence data from a postal survey by comparison with cancer registry records.
Am J Epidemiol. 1994;139:408-414[2]

4 Schrijvers CTM, Coebergh JWW, Heijden LH van der, Mackenbach JP. Socio-Economic Status and breast cancer survival in the Southeastern Netherlands, 1980-1989
Eur J Cancer 1995;31A:1660-1664[3]

4 Schrijvers CTM, Coebergh JWW, Heijden LH van der, Mackenbach JP. Socioeconomic variation in cancer survival in the Southeastern Netherlands, 1980-1989.
Cancer 1995;75:2946-2953[4]

5 Schrijvers CTM, Mackenbach JP, Lutz J-M, Quinn MJ, Coleman MP. Deprivation and survival from breast cancer.
Br J Cancer 1995;72:738-743[5]

5 Schrijvers CTM, Mackenbach JP, Lutz J-M, Quinn MJ, Coleman MP. Deprivation, stage at diagnosis and cancer survival.
In press Int J Cancer[6]

4 Socioeconomic status and co-morbidity among incident cancer patients in the Southeastern Netherlands, 1993. (manuscript)

6 Socioeconomic variation in cancer survival in the Southeastern Netherlands and the South Thames area: a comparison. (manuscript)

Chapter 1. Introduction

Cancer is the second most important cause of death in the Netherlands, as it is in many developed countries. In 1990 about 30% of all deaths in the Netherlands could be ascribed to cancer.[1] In general, survival from cancer is rather poor but it varies by such characteristics as the organ of origin of the tumour and age of the patient. Socioeconomic status is another factor which has been found to be of prognostic importance for cancer patients.[2] In general, cancer patients from lower socioeconomic groups have a lower survival rate than patients from higher socioeconomic groups.

The subject of this thesis is the association between socioeconomic status and cancer survival in the southeastern Netherlands and the area covered by the South Thames Regional Health Authority (RHA) in South East England. Both a description of and explanations for variation in survival by socioeconomic status in these two areas are given. The studies reported in this thesis, can be placed within a broader framework of research on socioeconomic inequalities in health (paragraph 1.1) as well as within a narrower framework of research on socioeconomic inequalities in cancer (paragraph 1.2). In the final paragraph of this introduction, the aims of the studies reported in this thesis are presented (paragraph 1.3).

1.1 Socioeconomic inequalities in health: a brief introduction

Our society is characterised by a system of social stratification, which is caused by an unequal distribution of material and other resources among the inhabitants. People hold a relative position on the social hierarchy, which is summarized by the term socioeconomic status, in indicators of socioeconomic status such as education, income, and occupation. Each refers to a different aspect of social stratification. Education determines the access to information and the ability to process this information, income is important with regard to access to material goods, while occupation refers to the prestige, privileges and power associated with holding specific jobs.[3]

1.1.1 A description of socioeconomic inequalities in health

Socioeconomic inequalities in health can be defined as systematic differences in the prevalence or incidence of health problems between people of higher and lower socioeconomic status.[4] Research into socioeconomic inequalities in health was initiated in the 19th century by medical doctors who were organised in for example the sanitary movement, which gathered information on important public health issues. They showed among others that the poor segments of society had higher mortality rates than the rich segments.[5] In the 20th century, the establishment of the welfare state was thought to have reduced socioeconomic inequalities in health substantially, because an important characteristic of the welfare state is equal access to the health care system for everybody, regardless of socioeconomic status. It became evident however, that socioeconomic inequalities in health are still present in European countries.[6]

Great Britain has a long tradition of research into socioeconomic inequalities in health. For example the association between social class, based on occupation, and mortality has been registered and described since the 19th century.[7] An important landmark, both at the national and international level, was the publication of the Black report in 1980, which showed that "from birth to old age, those at the bottom of the social scale have much poorer health and quality of life than those at the top".[8] This conclusion from the Black report applies to many other countries in Europe and to a variety of health indicators.[6]

The Netherlands have a much shorter history of research into socioeconomic inequalities in health than Great Britain and therefore less evidence on such differences has become available. Recently, a growing number of research projects has shown that also in the Netherlands, a lower socioeconomic status is associated with a higher frequency of many health problems, such as the prevalence of health complaints, prevalence of many chronic conditions and adult mortality.[9,10] The number of reported health complaints, measured with a list of symptoms and sensations, is on average higher in lower socioeconomic groups. The Netherlands Health Interview Survey reports for the period 1981-1985 an average of 8.4 complaints in respondents with primary school only, whereas the average number of complaints is 5.2 in respondents with a university education. Findings for income as socioeconomic indicator are similar.[11] The number of chronic conditions per 100 persons, as reported in the Netherlands Health Interview Survey for the period 1981-1985, was found to be about 50% higher among people with the lowest educational level than among those belonging to the highest educational category. The difference was smaller when income was used as indicator of socioeconomic status.[11] There are also results from condition-specific analyses: one study showed that the relative risk for the prevalence in the lowest versus the highest socioeconomic category of most specific chronic conditions, such as lung diseases, diabetes, and back complaints, lies between 1.10 and 1.30 when 3 occupational status categories were distinguished. The largest relative risk was found for chronic bronchitis (1.68).[12] Finally, four longitudinal studies on socioeconomic inequalities in mortality among men aged 35-64 years, covering the period from the 1950s onwards, have revealed that the relative risk of dying for the lowest versus the highest socioeconomic group varies between about 1.20 and about 2.00. This variation in study results can probably be ascribed to differences in the study population and design, but the results all show the same pattern: higher mortality in the lower socioeconomic groups.[13-16]

1.1.2 Explanations of socioeconomic inequalities in health

Four categories of possible explanations of socioeconomic inequalities in health are distinguished in the Black report, which are: (1) artefact explanations, (2) social selection, (3) materialist and structuralist explanations, and (4) behavioural and cultural explanations.[8]

1. Artefact explanations
Artefact explanations suggest that the association between socioeconomic status and health is due to errors in the process of measurement. For example, socioeconomic differences in self-reported morbidity could be caused by differential misreporting of morbidity by socioeconomic status. Actually, socioeconomic inequalities in the prevalence of specific chronic conditions (chronic non-specific lung disease, heart disease and diabetes mellitus) were found to be underestimated as a result of differential misreporting.[17] In general, the artefact explanation is thought to be of little importance in explaining socioeconomic inequalities in health.[18]

2. Social selection
Social selection means that health has an effect on socioeconomic status, rather than the other way around. Two types of health related selection can be distinguished; the first one is health related intragenerational social mobility. This implies that adults with a bad health status move downwards in the social hierarchy more often and move upwards less often as compared to persons in good health. This results in a relatively large number of people with ill-health at the bottom of the social hierarchy. This type of explanation has probably only a modest effect on socioeconomic inequalities in health.[19,20] The second type of health related selection is intergenerational social mobility. This implies that people with a bad health status during childhood/early adulthood move less often to a higher socioeconomic position and more often to a lower position than their parents' position, as compared to people with a good health during childhood. This explanation is probably more important, but the available evidence on this type of selection is still very sparse.[18]

The other two explanations of socioeconomic inequalities in health, which will be discussed below, are part of the causation theory. This theory assumes that socioeconomic status has a causal but indirect effect on health through a differential distribution of determinants of health across socioeconomic groups.

3. Materialist/structuralist explanations
The third explanation states that material deprivation has an effect on health and refers to exposure to hazards which are unequally distributed across socioeconomic groups. Examples of such hazards are exposure to health-damaging chemicals in certain occupations and poor-quality housing, which are more common in the lower socioeconomic groups.[18]

4. Behavioural/cultural explanations
Finally, behavioural and cultural explanations emphasize the role of the differential distribution across socioeconomic groups of adverse health-related behaviours in causing socioeconomic inequalities in health. Examples of factors with a higher prevalence in lower socioeconomic groups are smoking, adverse dietary habits, and a lack of physical exercise.[18]

Other possible explanations

The distinction of four categories of explanations of socioeconomic inequalities in health in the Black report has given rise to a debate on other possible explanations which are part of the causation mechanism. An example is the unequal distribution of psychosocial stress across socioeconomic groups.[21] This is a plausible explanation of socioeconomic inequalities in health, as these have been observed for so many different health problems. A higher general susceptibility to disease among those with a lower socioeconomic status might be related to such psychosocial stressors as adverse life events and continuous psychosocial burdens.

Another factor which may be responsible for socioeconomic inequalities in health is unequal access to the health care system for people from different socioeconomic groups, which may even exist in countries such as the Netherlands and the United Kingdom in which this access is assumed to be equal for everybody. Socioeconomic inequalities in medical consumption in the Netherlands have been observed in a study conducted in the early 1990s in the Southeastern Netherlands.[22] This study showed that, after adjustment for socioeconomic differences in health status, persons with a lower educational level visit their general practitioner more often than persons with a higher educational level. For prescribed drugs there are no socioeconomic differences in utilization, while the specialist and physiotherapist are visited less frequently by people from lower educational groups as compared to people from higher educational groups. People with a lower educational level use drugs without prescription less often than those with a higher educational level. The results for hospital admissions revealed no clear picture.

Results from studies on inequalities in the provision and use of health care services in the United Kingdom have suggested that the National Health Service does not guarantee equal access for everybody to the health care system. As in none of the studies reported here adjustment was made for health status, the results should be interpreted with caution however. A study using data from the British General Household Survey for the years 1983-1987 showed that residents of socially deprived areas have higher than average general practitioner consultation rates.[23] A recent report showed that for some conditions (hernia, cholecystectomy, hip operations) general practitioner consultations increase with social deprivation, while operation rates do not appear to show the same pattern.[24] Two studies on access to services for the management of ischaemic disease by deprivation have shown less access for residents of poorer areas[25,26] while another study has shown no such differences.[27]

Policy measures

The accumulated evidence on the existence and causes of socioeconomic inequalities in health asks for policy measures to reduce these inequalities. In our society, health is very important and therefore the tendency exists to consider all socioeconomic inequalities in health as unjust. On the other hand, freedom of choice for every citizen is a central principle and therefore it is more appropriate to

consider socioeconomic inequalities in health as unjust only in so far as these arise from an unequal opportunity for everyone to achieve health. The determinants of health leading to unjust socioeconomic inequalities in health should be subject to policy. These determinants are living conditions beyond the control of the individual (physical and social environment and health care) and conditions of choice (e.g. the knowledge of an individual about the health risks of a certain behaviour). Apart from unjust inequalities, we may distinguish unavoidable and acceptable inequalities in health. If the causes of socioeconomic inequalities in health are determined by nature, as is the case with the distribution of genetic factors, this inequality may be unfair but is also unavoidable, as the distribution of these factors cannot be changed. Other inequalities are acceptable, as these are the results of free individual choices. This approach implies that not all socioeconomic inequalities in health are necessarily unjust, so that not all determinants of such inequalities have to be a target for policy measures.[28]

1.2 Socioeconomic inequalities in cancer

1.2.1 A description of socioeconomic inequalities in cancer

The association between socioeconomic status and cancer mortality is similar to that for many other diseases: higher mortality rates have been observed among socioeconomically disadvantaged people for all cancers combined, as well as for a large number of specific cancers.[29,30] Socioeconomic inequalities in cancer mortality are the end result of socioeconomic inequalities in the incidence of and survival from cancer. For most cancers, incidence is unequally distributed across socioeconomic groups,[31] which might (partly) be explained by the differential distribution of cancer risk factors across socioeconomic groups. As with cancer mortality,[32] the strength and direction of the association between socioeconomic status and cancer incidence differs per cancer. For example, most studies have found a higher incidence of breast cancer among women at the upper end of the social scale.[31] This may be explained by a higher prevalence of risk factors such as nulliparity and late age at first birth[33] among women of high socioeconomic status. Lung cancer incidence is, on the other hand, higher in the lower socioeconomic groups,[31] which might be explained by a higher prevalence of smoking in these groups.[34-36]

Studies on the association between socioeconomic status and cancer survival have revealed rather consistent findings. For most cancers it was found that people with a high socioeconomic status live longer after a cancer diagnosis than those with a low socioeconomic status. Kogevinas and co-authors found, for patients diagnosed between 1971 and 1981 in England and Wales, that owner occupiers (high socioeconomic status) had better survival than council tenants (low socioeconomic status) for 11 out of 13 cancers in males and 12 out of 15 cancers in females.[37] A Swedish national study found better survival for white collar workers

than for blue collar workers for 10 out of 13 cancers in both males and females among patients diagnosed between 1961 and 1979.[38] An older study, on patients diagnosed between 1940 and 1969 in Iowa, United States, found that for each of 39 cancers, survival of indigent patients was poorer than survival of non-indigent patients.[39]

In the Netherlands, only one study investigated the association between socioeconomic status and overall cancer mortality, which was higher in men with a low socioeconomic status as compared to men with a high socioeconomic status.[13] The association between socioeconomic status and the incidence of cancers of the lung[40], breast[41], and colon[42] was the subject of another Dutch study. Lung cancer incidence was higher in men with a low education, while men with a high education had a higher risk to develop colon cancer than men with a low education. For women no association was found between socioeconomic status and the incidence of breast[41] and colon cancer[42], while lung cancer incidence was not investigated in women. The association between socioeconomic status and cancer survival has never before been investigated in the Netherlands.

1.2.2 Explanations of socioeconomic inequalities in cancer survival

Several possible explanations of socioeconomic inequalities in cancer survival have been studied and hypothesized; these can be grouped in four main categories.[2]

1. Differences in tumour biology

The histological type of a tumour is both an important biological feature of a tumour and an important prognostic factor. Its possible effect on socioeconomic inequalities in survival can be illustrated with the example of lung cancer. Lung cancer patients diagnosed with small-cell tumours experience lower survival than patients diagnosed with other histological types.[43,44] Small-cell lung tumours have been found to be very closely linked with tobacco smoking,[45] which is more common in lower socioeconomic groups.[34-36] Part of the lower lung cancer survival of patients with a low socioeconomic status as found in some studies[38,44] might be explained by a higher frequency of small-cell tumours in this group of lung cancer patients. The principle of a differential distribution of histological types across socioeconomic groups may also apply to other cancers.

Another biological feature of a tumour is the part of an organ (subsite) in which it originated, which may be important in colorectal and stomach cancer. For example, survival rates differ for subsites in colon cancer[46] and the distribution of subsites may vary across socioeconomic groups. Part of the socioeconomic variation in survival may therefore be explained by the distribution of subsites across socioeconomic groups.

2. Differences in delay in diagnosis

Delay in diagnosis of cancer can be defined as the time interval between the onset of symptoms and the diagnosis of cancer. This total delay can be subdivided in several periods: for example the time between the onset of symptoms of cancer and the first contact with the health care system (patient delay) and the time between this first contact and definitive diagnosis and/or start of the treatment (diagnostic delay).[47] A short delay or an earlier diagnosis and subsequent treatment may positively affect the natural history of the disease and therefore postpone death and result in a true survival advantage. However, the effect of a short delay, through an earlier diagnosis, may also result in advancing the time of diagnosis without postponing a patients' death. The time which is added erroneously in this way to a persons' survival time is called lead time and the bias resulting from this time (lead time bias) should always be considered as a possible artefact explanation of any gradient in survival by socioeconomic status.

Reliable data on delay are seldom available from medical records and therefore from cancer registries, which constitute the data source for many studies on socioeconomic variation in cancer survival. The evidence on the association between socioeconomic status and delay is mainly based on studies that used data from clinical records, which are based on interviews with patients at the time of hospital admission. Some of these studies have found a longer delay in the lower socioeconomic groups,[48-50] while others have found no association between socioeconomic status and delay.[51,52]

The impact of delay on socioeconomic inequalities in cancer survival can be studied, indirectly, through stage of disease at diagnosis, because delay is related to stage of disease at diagnosis. In studies on breast cancer it was shown that a shorter period of delay results in less advanced stages,[50,53-57] and a less advanced stage in general results in a better survival. The stage distribution of cancer patients with a low socioeconomic status was found to be less favourable (more advanced stages) than the stage distribution of patients with a high socioeconomic status for cancers of the breast,[48,58-61] colon,[46,51,62] and cervix.[63]

3. Differences in treatment

It has often been suggested that socioeconomic inequalities in cancer survival may also be caused by differences in the type of treatment received by patients from different socioeconomic groups.[2] However, only a few studies have taken such differences in treatment into account. Moreover, most data on treatment come from cancer registry records and concern the broad type of primary treatment, because more detailed information on treatment (on factors such as compliance, doses and frequency of chemo- and radiotherapy) is not available from registry records. In two American studies it was found that within several treatment groups, socioeconomic inequalities in survival were still apparent. These findings suggest that major differences in treatment are not responsible for these inequalities,[64,65] although differences in treatment may well exist in each broad treatment category.

A higher frequency of mastectomy as opposed to lumpectomy or partial mastectomy was found in less educated breast cancer patients in the USA, after adjustment for tumour size and co-morbidity.[66] A study on the treatment of non-small-cell lung cancer patients in the USA found that those who had private medical insurance were more likely to be treated with surgery than those with another or no medical insurance. Among patients who did not have surgery, those with private insurance were more likely to receive another form of therapy (radiation or chemotherapy).[67]

4. Differences in host resistance

A striking feature of the results from studies on socioeconomic inequalities in cancer survival is the fact that for most cancers the same association between socioeconomic status and cancer survival was found. A plausible explanation for this finding is a lower host resistance among the socioeconomically disadvantaged, leading to a more rapid tumour growth and spread, and resulting in more advanced stages. Host resistance could be lower in patients of low socioeconomic status because of poor nutrition, more co-morbidity, and adverse psychosocial factors such as stressful life events, a low ability to cope with a cancer diagnosis and a lack of social support. Stressful life events,[68,69] and a lack of social support[68,70-72] have been found to be more common among people with a low socioeconomic status in general, but the evidence on the role of these factors in cancer patients is conflicting. One study found that being able to express emotion is an important positive prognostic factor for patients with metastatic breast cancer.[73] A study based on the experience of a small cohort of breast cancer patients provides limited evidence that social stress decreases and social involvement increases survival time.[74] An experimental group of breast cancer patients receiving psychotherapy survived longer than a control group of patients which did not receive such therapy.[75] On the other hand, for breast cancer patients with metastatic disease, disease-related variables probably outweigh the influence of psychosocial factors in determining length of survival.[76]

The possible explanations of socioeconomic inequalities in cancer survival can also be grouped according to the scheme applied in the Black Report (paragraph 1.1.1). They all fit within the causation theory, although the artefact explanation and selection may also play a role. Those differences in tumour biology that are caused by life style characteristics, can be regarded as behavioural and cultural explanations. The same is true for delay in diagnosis, caused by either differences in knowledge about health or attitude towards health care. Part of the variation in delay may be caused by socioeconomic differences in access to health care, which can be placed under the heading of structuralist explanations. Differences in treatment, after adjustment for biological features of a tumour, and differences in host resistance relate both to behavioural/cultural and to materialist/structuralist explanations.

1.3 This thesis

This thesis reports the results of a study on the association between an area-based measure of socioeconomic status and survival from the most common cancers in the area covered by the population based Eindhoven cancer registry (Southeastern Netherlands). This association has been quantified and furthermore, possible explanations of the association between socioeconomic status and cancer survival have been studied. In order to place the results from the Dutch study in a broader perspective, another study was undertaken on the association between an area-based measure of socioeconomic status and cancer survival in part of the area covered by the population based Thames Cancer Registry (Southeast England). The aims of this study were also to quantify the association between socioeconomic status and cancer survival and to study the impact of possible explanatory factors on this association.

Study aims

The specific aims of the studies reported in this thesis are:

1. To describe variation in cancer survival by socioeconomic group for patients diagnosed with common cancers between 1980 and 1989 in two areas: the Southeastern Netherlands and the area covered by the South Thames Regional Health Authority (RHA).
   In both areas, survival from cancers of the lung, breast, colorectum, prostate, and stomach was investigated. The number of patients from the South Thames area was much larger than in the Southeastern Netherlands, and therefore also less frequent cancers could be studied, which are cancers of the bladder, pancreas, ovary, uterus, and cervix.
2. To investigate the impact of a number of prognostic factors on the association between socioeconomic status and cancer survival in both areas. The prognostic factors were:
   (1) histological type and subsite of the tumour, as indicators of tumour biology; subsite was only thought to be of prognostic importance for stomach and colorectal cancer;
   (2) stage of disease at diagnosis, as indicator of delay in diagnosis;
   (3) type of treatment;
   (4) number of life events and number of co-morbid conditions, as indicators of host status; these factors were only studied in the Southeastern Netherlands.
3. To compare results from the studies in the Southeastern Netherlands and the South Thames area.

Contents of this thesis

Chapter 2 contains a review on studies conducted since the 1950s on cancer survival by socioeconomic status in seven countries for six common cancers. The major methodological characteristics of the studies in this thesis are described in chapter 3. The results of studies on socioeconomic variation in cancer survival and the impact of prognostic factors on the association between socioeconomic status and survival are discussed in chapters 4 (Southeastern Netherlands) and 5 (South Thames), followed by a comparison of the results from both areas in chapter 6. Chapter 7, the discussion, evaluates the gained insights from this thesis, while taking several methodological issues into account.

References

1. Centraal Bureau voor de Statistiek. Sterfte naar belangrijke doodsoorzaken, 1970-1990.'s-Gravenhage: sdu/uitgeverij/cbs-publikaties, 1992.
2. Leon D, Wilkinson RG. Inequalities in prognosis: socio-economic differences in cancer and heart disease survival. In: Fox J (ed). Health Inequalities in European Countries. Aldershot: Gower, 1989.
3. Giddens E. Sociology (second edition). Cambridge: Polity Press, 1993.
4. Kunst AE, Mackenbach JP. Measuring socioeconomic inequalities in health. Copenhagen: WHO regional office for Europe, 1994.
5. Houwaart ES. De hygiënisten: artsen, staat en volksgezondheid in Nederland 1840-1900. Groningen: Historische Uitgevery Groningen, 1991.
6. Fox J (ed). Health Inequalities in European Countries. Aldershot: Gower, 1989.
7. Registrar General. Occupational mortality, Decennial Supplement England & Wales, 1851. London: HMSO, 1851.
8. Townsend P, Davidson N, Whitehead M (eds). Inequalities in health (The Black Report & The Health Divide). London: Penguin Books, 1988.
9. Mackenbach JP. Ongezonde verschillen. Over stratificatie en gezondheid in Nederland. Assen: Van Gorcum, 1994.
10. Mackenbach JP. Socioeconomic inequalities in health in the Netherlands: impact of a five year research programme. Br Med J 1994;309:1487-1491.
11. Centraal Bureau voor de Statistiek. Netherlands Health Interview Survey 1981-1985. Den Haag: Staatsuitgevery, 1988.
12. Bos GAM van den. Zorgen van en voor chronisch zieken (thesis). Amsterdam: Bohn, Scheltema & Holkema, 1989.
13. Appels A, Otten F, Mendes de Leon C, Sturmans F, Mulder P, Schuurman P. De KRIS follow-up studie VII. Sociaal-economische status en gezondheid. Tijdschr Soc Gezondheidsz 1990;68:298-305.
14. Mares NEHM, Aben DJM, Schouten EG, Kok FJ, Heide-Wessel C van der, Heide RM van der. Inkomen en sterfte: resultaten van 25 jaar vervolgonderzoek bij mannelijke Amsterdams ambtenaren. Ned Tijdschr Geneeskd 1988;132:1109-1113.
15. Doornbos G, Kromhout D. Educational level and mortality in a 32-year follow-up study of 18-year-old men in the Netherlands. Int J Epidemiol 1990;19:374-379.
16. Duijkers TJ, Kromhout D, Spruit IP, Doornbos G. Inter-mediating risk factors in the relation between socio-economic status and 25-year mortality (the Zutphen Study). Int J Epidemiol 1989;18:658-662.
17. Mackenbach JP, Looman CWN, Meer JBW van der. Misreporting of chronic conditions in a health interview survey: differentials by level of education (submitted).
18. Davey Smith G, Blane D, Bartley M. Explanations for socio-economic differentials in mortality. Eur J Public Health 1994;4:131-144.
19. Fox AJ, Goldblatt PO, Adelstein AM. Selection and mortality differentials. J Epidemiol Community Health 1982;36:69-79.
20. Fox AJ, Goldblatt PO, Jones DR. Social class mortality differentials: artefact, selection or life circumstances? J Epidemiol Community Health 1985;39:1-8.
21. Marmot MG, Shipley MJ, Rose G, Inequalities in death-specific explanations of a general pattern? Lancet 1984;1:1003-1006.
22. Meer JBW van der, Looman CWN, Mackenbach JP. Sociaal-economische verschillen in

medische consumptie. Rotterdam: Instituut Maatschappelijke Gezondheidszorg, 1993.
23. Balarajan R, Yuen P, Machin D. Deprivation and general practitioner workload. BMJ 1992;304:529-534.
24. Chaturvedi N, Ben-Schlomo Y. From the surgery to the surgeon: does deprivation influence consultation and operation rates? British Journal of General Practice 1995;45:127-131.
25. Findlay IN, Dargie HJ, Dyke T. Coronary angiography in Glasgow: relation to coronary heart disease and social class. Br Heart J 1991;66:70.
26. Ben-Schlomo Y, Chaturvedi N. Assessing equity in access to health care provision in the UK: does where you live affect your chance of getting coronary artery bypass graft? J Epidemiol Community Health 1995;49:200-204.
27. Kee F, Gaffney S, Currie S, O'Reilly D. Access to coronary catherisation: fair shares for all? BMJ 1993;307:1305-1307.
28. Stronks K, Gunning-Schepers LJ. Should equity in health be target number 1? Eur J Public Health 1993;3:104-111.
29. Logan WPD. Cancer mortality by occupation and social class 1851-1971. London: HMSO, 1982.
30. Valkonen T, Martelin T, Rimpelä A, Notkola V, Savela S. Socio-economic mortality differences in Finland 1981-90. Helsinki: Statistics Finland, 1993.
31. Loon AJM van, Brug J, Goldbohm RA, Brandt PA van den. Differences in cancer incidence and mortality among socio-economic groups. Scan J Soc Med 1995;23:110-120.
32. Davey Smith G, Leon D, Shipley MJ, Rose G. Socioeconomic differentials in cancer among men. Int J Epidemiol 1991;20:339-345.
33. Kelsey JL, Gammon MD, John EM. Reproductive factors and breast cancer. Epidemiol Rev 1993;15:36-47.
34. Marmot MG, Kogevinas M, Elston MA. Social/economic status and disease. Ann Rev Public Health. 1987;8:111-135.
35. Pocock SJ, Cook DG, Shaper AG, Phillips AN, Walker M. Social class differences in ischaemic heart disease in British men. Lancet 1987;2:197-201.
36. Wagenknecht LE, Perkins LL, Cutter GR et al. Cigarette smoking behaviour is strongly related to educational status: the CARDIA study. Preventive Medicine 1990;19:158-169.
37. Kogevinas M, Marmot MG, Fox AJ, Goldblatt PO. Socioeconomic differences in cancer survival. J Epidemiol Community Health 1991;45:216-219.
38. Vågerö D, Persson G. Cancer survival and social class in Sweden. J Epidemiol Community Health 1987;41:204-209.
39. Berg JW, Ross R, Latourette HB. Economic status and survival of cancer patients. Cancer 1977;39:467-477.
40. Loon AJM van, Goldbohm RA, Brandt PA van den. Lung cancer: is there an association with socioeconomic status in The Netherlands? J Epidemiol Community Health 1994;49:65-69.
41. Loon AJM van, Goldbohm RA, Brandt PA van den. Socioeconomic status and breast cancer incidence: a prospective cohort study. Int J Epidemiol 1994;23:899-905.
42. Loon AJM van, Goldbohm RA, Brandt PA van den. Socioeconomic status and colon cancer incidence: a prospective cohort study. Br J Cancer 1995;71:882-887.
43. Bonett A, Roder D, Esterman A. Monitoring case survival by histological type for lung cancer, breast cancer and non-Hodgkin's lymphomas: one function of a population-based registry. Commun Health Stud 1982;6:223-227.
44. Bonett A, Roder D, Esterman A. Determinants of case survival for cancers of the lung, colon, breast and cervix in South Australia. Med J Aust 1984;141:705-709.
45. Kreyberg L. Lung cancer and tobacco smoking in Norway. Br J Cancer 1955;9:495-510.
46. Auvinen A. Social class and colon cancer survival in Finland. Cancer 1992;70:402-409.
47. Funch DP. Diagnostic delay in symptomatic colorectal cancer. Cancer 1985;56:2120-2124.
48. Richardson JL, Langholz G, Bernstein L, Burciaga C, Danley K, Ross RK. Stage and delay in breast cancer diagnosis by race, socioeconomic status, age and year. Br J Cancer 1992;65:922-926.
49. Elwood JM, Moorehead WP. Delay in diagnosis and long-term survival in breast cancer. BMJ 1980;31 May:1291-1294.
50. Gould Martin K, Paganini-Hill A, Casagrande C, Mack T, Ross RK. Behavioral and biological determinants of surgical stage of breast cancer. Preventive Medicine 1982;11:429-440.
51. Vineis P, Fornero G, Magnino A, Giacometti, R, Ciccone G. Diagnostic delay, clinical stage, and social class: a hospital based study. J Epidemiol Community Health 1993;47:229-231.
52. Samet JM, Hunt WC, Lerchen ML, Goodwin JS. Delay in seeking care for cancer symp-

toms: a population-based study of elderly New Mexicans. J Nat Cancer Inst 1988;80:432-438.
53. Wilkinson GS, Edgerton F, Wallace HJ, Reese P, Patterson J, Priore R. Delay, stage of disease and survival from breast cancer. J Chron Disease 1979;32:365-373.
54. Hainsworth PJ, Henderson MA, Bennett RC. Delayed presentation in breast cancer: relationship to tumour stage and survival. The Breast 1993;2:37-41.
55. Robinson E, Mohilever J, Zidan J, Sapir D. Delay in diagnosis of cancer. Possible effects on the stage of disease and survival. Cancer 1984;54:1454-1460.
56. Porta M, Gallén M, Malats N, Planas J. Influence of diagnostic delay upon cancer survival: an analysis of five tumour sites. J Epidemiol Community Health 1991:45:225-230.
57. Maguire A, Porta M, Malats N, Gallén M, Pinol JL, Fernandez E. Cancer survival and the duration of symptoms. An analysis of possible forms of the risk function. Eur J Cancer 1994;30A:785-792.
58. Farley TA, Flannery JT. Late-stage diagnosis of breast cancer in women of lower socioeconomic status: public health implications. Am J Public Health 1989;79:1508-1512.
59. Wells BL, Horm JW. Stage at diagnosis in breast cancer: race and socioeconomic factors. Am J Public Health 1992;82:1383-1385.
60. Mandelblatt J, Andrews H, Kerner J, Zauber A, Burnett W. Determinants of late stage diagnosis of breast and cervical cancer: the impact of age, race, social class, and hospital type. Am J Public Health 1991;81:646-649.
61. Karjalainen S, Pukkala E. Social class as a prognostic factor in breast cancer survival. Cancer 1990;66:819-826.
62. Brenner H, Mielck A, Klein R, Ziegler H. The role of socioeconomic factors in the survival of patients with colorectal cancer in Saarland/Germany. J Clin Epidemiol 1991;44:807-815.
63. Lamont DW, Symonds RP, Brodie MM, Nwabineli NJ, Gillis CR. Age, socio-economic status and survival from cancer of cervix in the West of Scotland 1980-1987. Br J Cancer 1993;67:351-357.
64. Lipworth L, Abelin T, Connelly RR. Socio-economic factors in the prognosis of cancer patients. J Chron Dis 1970;23:105-116.
65. Lipworth L, Bennett B, Parker P. Prognosis of nonprivate cancer patients. J Nat Cancer Inst 1972;48:11-16.
66. Muss HB, Hunter CP, Wesley M, et al. Treatment plans for black and white women with stage II node-positive breast cancer. The National Cancer Institute black/white cancer survival study experience. Cancer 1992;70:2460-2467.
67. Greenberg ER, Chute CG, Stukel T, et al. Social and economic factors in the choice of lung cancer treatment. A population-based study in two rural states. N Engl J Med 1988;318:612-617.
68. Murrell SA, Norris FH. Differential social support and life change as contributors to the social class-distress relationship in older adults. Psychol and Aging 1991;6:223-231.
69. Myers JK, Lindenthal JJ, Pepper MP. Social class, life events and psychiatric symptoms: a longitudinal study. In: Dohrenwend BS, Dohrenwend, BP (ed). Stressful life events: their nature and effects. New York: Wiley, 1974: 191-206.
70. Raats G, Sanderman R, Ormel H. Ongelijke kansen op geestelijke gezondheid(szorg). Groningen: Universiteitsdrukkerij, 1987.
71. Ranchor AV, Bouma J, Sanderman R. Vulnerability and social class: differential patterns of personality and social support over the social classes. In: Ranchor AV, Bouma J, Sanderman R. Sociaal-economische status en het myocardinfartct. Groningen: Vakgroep Gezondheidswetenschappen/Noordelijk Centrum voor Gezondheidsvraagstukken (Verslag ten behoeve van de programmacommissie Sociaal-Economische Gezondheidsverschillen), 1993.
72. Matthews KA, Kelsey SF, Meilahn EN, Kuller LH, Wing RR. Educational attainment and behavioral and biologic risk factors for coronory heart disease in middle-aged women. Am J Epidemiol 1989;129:1132-1144.
73. Derogatis LR, Abelhoff MD, Melisaratos N. Psychological coping mechanism and survival time in metastatic breast cancer. J Am Med Assoc 1979;242:1504-1508.
74. Marshall JR, Funch DP. Social environment and breast cancer. A cohort analysis of patient survival. Cancer 1983;52:1546-1550.
75. Spiegel D, Kraemer HC, Bloom JR, Gottheil E. Effects of psychosocial treatment on survival of patients with metastatic breast cancer. Lancet 1989;October 14:888-891.
76. Jamison FN, Burish TG, Wallston KA. Psychogenic factors in predicting survival of breast cancer patients. J Clin Oncol 1987;5:768-772.

Chapter 2. Cancer patient survival by socioeconomic status: a review for six common cancer sites

2.1 Introduction*

Socioeconomic differences in mortality have been reported for a variety of causes of death including cancer.[1,2] Cancer mortality is generally higher in people of low socioeconomic status (SES) compared with people of a high SES. This mortality disadvantage may be the result of socioeconomic differences in cancer incidence or cancer survival.

Socioeconomic differences in cancer incidence and cancer survival do not call for the same health policy measures. Differences in cancer incidence ask for interventions in the area of primary prevention, whereas socioeconomic differences in cancer survival ask for policy measures in the area of secondary prevention or treatment.

We have tried to establish the size and consistency of socioeconomic differences in cancer survival, on the basis of a systematic review of the available published studies on the subject. This review deals with socioeconomic differences in cancer patient survival for a number of common cancer sites: colon, rectum, lung, prostate, breast, and cervix.

2.2 Methods

The study material was selected through Medline and the references of papers and books, which resulted in 40 papers on socioeconomic differences in cancer survival. To enable a useful comparison of the results of the reviewed studies, some exclusion criteria were developed.

Studies on patients diagnosed in the 1950s or earlier were excluded.

Hospital based studies were excluded because cancer patients treated in specific hospitals may not be representative of cancer patients in the general population. In particular, socioeconomic contrast may be larger in the general population than in a hospital population.

Studies covering fewer than five years of follow-up were excluded, because for many cancers survival differences may not yet be apparent shortly after diagnosis.

Three measures of SES were considered to be unfit for our purpose. Studies using race as a measure were excluded, because it is difficult to separate the impact of SES and other race related factors on survival. Studies that used hospital type or insurance status as a socioeconomic measure were also excluded, as we consider both variables to be intermediate in the SES-survival association.

Studies that reported on fewer than 200 cancer deaths were excluded from this review. This number of events is the minimum needed to indicate a relative risk (RR) of dying of 1.5 when two socioeconomic groups with equal numbers are compared (with $\alpha = 0.05$ and $ß = 0.20$).[3]

---

* Schrijvers CTM, Mackenbach JP. J Epidemiol Community Health 1994;48:441-446

Cancer sites for which fewer than three papers on SES and survival were available were not considered in this review.

Finally, 14 studies remained for inclusion in the review. Table 1 presents the most important characteristics of the selected papers, which are ordered by country of origin of the study population.[4-19]

The country of origin of the study population may be a determinant of the strength of socioeconomic differences in cancer survival. In general, these differences are expected to be smaller in countries like Sweden, with good access to health care services for the entire population.

The measures of SES are divided into two broad categories: measures on the individual level such as education[8], occupation[14,16-18], or housing tenure[13] and ecological measures in which the place of residence of cancer patients is used to assign a socioeconomic score. These measures are either based on census tract[4], block group[5], postcode[6,7,11,15], electoral ward[12], or community of residence.[19]

Table 1 shows that most studies cover the 1970s and early 80s with the exception of three studies which cover an incidence period starting in the 60s.[4,8,16]

From table 1 it can be seen that different measures of survival were used. If the survival of cancer patients is studied, deaths due to causes other than the cancer(s) of interest must be excluded. In a number of studies the exact cause of death was known, and therefore patients dying from causes other than the specific cancer could be treated as censored in the survival analysis. The resulting measure is called the corrected survival rate.[4,11,15,17-19] The relative survival rate, which is the ratio of the observed and expected survival rate,[16,17] is usually calculated when reliable information on the exact cause of death is not available. The expected survival rate is based on life tables of the general population.

A few studies did not report on the exclusion of deaths from other causes.[5,8,12,14] In two other studies the distributions of deaths related and not related to cancer were similar in the different socioeconomic categories and the authors did not therefore correct for deaths from other causes.[6,7] Finally, the standardised case fatality ratio was employed in one study[13], in which the case fatality rates of the entire study population for the cancer in question were used as a standard.

For most studies an RR of dying for the lowest compared with the highest SES category was taken directly from the paper.[5-7,11,15,17-19] For two studies[8,12], we calculated an RR of dying with 95 % confidence intervals (95% CI).[20] For one study, the ratios of standardised case fatality rates were calculated; these are presented for men and women separately.[13] For two studies we present a survival ratio,[4,16] because an RR of dying could not be calculated. A survival ratio is the ratio of the survival rate of the lowest to the highest SES group and indicates worse survival for the lowest SES group if it is below 1.00.

Table 1. Study-population, measure of socioeconomic status (SES) and measure of survival for 14 published reports on socioeconomic differences in cancer survival

Ref no	Population	Cancer site	No of patients	SES measure	Year of diagnosis	Measure of survival
4	Hawaii, USA	Colon Rectum	1446 881	Ecological: weighted score based on: average years of education and average income per census tract; 3 categories	1960-74	Corrected survival rate
5	Northwestern Washington State, USA	Breast	1506	Ecological: social class, several indicators per block group of residence; 2 categories	1973-83	Survival rate*
6	USA	Prostate	2513	Ecological: education, % of high school graduates, ≥ 25 years, per postcode of residence; 4 categories	1977-81	Survival rate†
7	USA	Rectum Colon	1528 3617	Ecological: education, % of high school graduates, ≥ 25 years, per postcode of residence; 3 categories	1977-82	Survival rate†
8,9 10	Boston, USA Tokyo, Japan	Breast Breast	563 814	Individual: education, years of schooling; 2 categories	1965-66 1965-67	Survival rate*
11	South Australia	Lung Colon Breast	2934 2227 2676	Ecological: income, median male income per postcode of residence; 3 categories	1977-82	Corrected survival rate
12	Sheffield, UK	Cervix	548	Ecological: occupation, % of semiskilled/unskilled workers per electoral ward; 5 categories	1971-84	Survival rate*
13	England & Wales	Breast Lung Colon Rectum Prostate	Total 17844	Individual: housing tenure; 2 categories	1971-81	Standardised case fatality ratio
14	South Thames RHA, UK	Cervix	1728	Individual: social class (occupation); 5 categories	1977-81	Survival rate*
15	West of Scotland, UK	Cervix	1588	Ecological: unweighed average of 4 census variables per postcode of residence; 7 categories	1980-87	Corrected survival rate
16	Sweden	Colon Rectum Prostate Lung Breast Cervix	5774 3707 4752 7540 11531 4087	Individual: occupation; 2 categories	1961-79	Relative survival rate
17	Finland	Breast	10181	Individual: social class (occupation); 4 categories	1971-80	Relative survival rate, Corrected survival rate
18	Finland	Colon	2969	Individual: social class (occupation); 4 categories	1979-82	Corrected survival rate
19	Saarland, Germany	Colon Rectum	1465 1162	Ecological: occupation: % of blue collar workers aged 15-65y per community of residence; education: % with no more than 9 years schooling per community of residence; 3 categories	1974-83	Corrected survival rate

* Whether a correction for causes of death other than the cancer was made is unknown
† No correction for other causes of death was made because the distributions of deaths related and not related to cancer were similar in the various SES categories

For one study, only graphs were presented in the paper.[16] We therefore obtained the original life tables from which five year relative survival rates had been abstracted and calculated 95% CIs from these.[21]

For most studies we used the number of SES categories originally distinguished by the authors. For one study,[8] we reduced the original number of four categories to two, to provide a sufficient number of patients per category.

2.3 Results

Table 2 shows the results of the selected papers ordered by cancer site.

Colon cancer

For colon cancer seven studies were included in the review. Two studies showed no association between SES and survival.[7,16] The other five studies all indicated a small survival advantage for colon cancer patients from the higher socioeconomic group.[4,11,13,18,19] In one of these studies the survival difference was not statistically significant at the 5% level,[4] while in another only the raised RR for men was statistically significant.[13] Finally, in one study there was no information on statistical significance.[18]

Rectal cancer

For cancer of the rectum, five studies are presented in table 2. Differences in survival were apparent in one study,[19] in which the RR of dying in the lowest compared with the highest SES group was statistically significantly ($p < 0.05$) larger than 1.00. Three other studies also showed worse survival for the lowest SES group,[4,7,16] although in two studies this was not a statistically significant difference,[4,7] and in the third study this was only the case for men.[16] Finally, one study showed (not statistically significant) opposite results for men (RR = 1.18) and women (RR = 0.82).[13]

Lung cancer

In the case of lung cancer, two studies presented a small, not statistically significant survival advantage for the highest SES group.[13,16] In one other study it was only mentioned that no survival difference was found.[11]

Prostatic cancer

For cancer of the prostate one study found a rather high RR of dying for the lowest SES category ($p=0.03$).[6] The results of the two other studies showed either a slight, not statistically significant, survival advantage for the lowest SES category[13] or for the highest SES category.[16]

Table 2. Results expressed as relative risk or survival ratio for the lowest relative to the highest socioeconomic status group

Study ref	Relative risk of dying (95% CI or p value)	Survival ratio (95% CI or p value)	Adjusted for:
Colon			
4		0.82 (0.66-1.02)	Age, sex, race, stage
7	0.97 ($p > 0.05$)		Age, sex, race, stage
11	1.26 (1.04-1.52)		Age, place of residence
13	M: 1.44 ($p < 0.05$) F: 1.11 ($p > 0.05$)		Age, period of follow-up
16		M,F: 1.00 ($p > 0.05$)*	Age, sex, follow-up year
18	1.15†		
19	1.22 (1.01-1.47)		Age, sex, stage, year of diagnosis, region, district
Rectum			
4		0.79 (0.60-1.05)	Age, sex, stage, race
7	1.09 ($p > 0.05$)		Age, sex, stage, race
13	M: 1.18 ($p > 0.05$) F: 0.82 ($p > 0.05$)		Age, period of follow-up
16		M: 0.83 ($p < 0.05$) F: 0.91 ($p > 0.05$)	
19	1.32 (1.09-1.60)		Age, sex, stage, year of diagnosis, region, district
Lung			
11	No difference†		Age, histology, birth place
13	M: 1.08 ($p > 0.05$) F: 1.13 ($p > 0.05$)		Age, period of follow-up
16		M: 0.93 ($p > 0.05$) F: 0.90 ($p > 0.05$)*	
Prostate			
6	1.86 ($p=0.03$)		Age, race
13	0.91 ($p > 0.05$)		Age, period of follow-up
16		0.94 ($p > 0.05$)*	
Breast			
5	1.52 (1.28-1.88)		Age, race, stage, histology
8	Boston:1.32 (1.08-1.61) Tokyo:1.30 (0.91-1.86)		Age
11	1.35 (1.04-1.74)		Age, histology
13	0.98 ($p > 0.05$)		Age, period of follow-up
16		0.91 ($p < 0.05$)*	
17	1.28 ($p < 0.05$)		Age, period of diagnosis, follow-up year
Cervix			
12	1.1 (0.99-1.23)		
14	No difference ($p>0.05$)		Age, stage
15	1.11 (0.64-1.92)		Age, stage, histology, tumour grade, health board, year of treatment
16		0.91 ($p < 0.05$)*	

M = male, F= female
* p value for this study < 0.05 when the 95% CIs for 5 year relative survival rate for the two SES groups do not overlap; † p value not reported

Breast cancer

Data on socioeconomic differences in breast cancer survival come from six studies in this review. Except for one study,[13] they all showed a raised RR of dying for patients with the lowest SES.[5,8,11,16,17] However, the results for Japan in one study were not statistically significant.[8]

Cervical cancer
Finally, for cancer of the cervix only one study showed a statistically significant higher survival rate for the highest SES group.[16] Two studies showed a slight survival advantage for the highest SES group, which was not statistically significant,[12,15] while in the fourth study no difference in survival between SES groups was found.[14]

2.4 Discussion

We have reviewed results from 14 studies on socioeconomic differences in survival for six cancer sites. As can be seen from table 2, survival differences are generally rather small. Furthermore, results differ in relation to the cancer site. With regard to the results, we distinguished between three types of studies: (1) those showing a statistically significant difference in survival; (2) studies showing survival differences, which are not statistically significant; and (3) studies showing no survival difference according to SES.

For cancers of the breast, colon, rectum, and cervix, most studies showed better survival for patients from higher socioeconomic groups. For these cancers all the statistically significant differences suggest a survival advantage for those of higher SES, and most of the non-significant differences agreed with this.

The results are unclear both for lung cancer and cancer of the prostate. For lung cancer, only small, non-significant survival differences were found in two studies,[13,16] with the higher SES groups showing an advantage, while no difference was found in the third study.[11] For cancer of the prostate, the results of one study which showed significantly better survival for the higher SES group,[6] were contradicted by one[13] of two studies that showed non-significant results.

In general, socioeconomic survival differences are thought to be larger in cancers of relatively good prognosis,[22] as earlier detection and treatment can be of greater influence on the survival for these cancers. This is more or less confirmed by our review, although the picture is less clear than expected.

For breast cancer, overall survival is rather good[23] and survival differences are relatively large. For lung cancer, which has the lowest overall survival probability of the six cancers studied,[23] very small survival differences were found. The remaining four cancer sites have an intermediate level of survival.[23] For cancers of the colon, rectum, and cervix survival differences according to SES were found, which is in concordance with their overall level of survival. For cancer of the prostate, which has a better overall survival than colon cancer, the results are less clear.

The general pattern of socioeconomic survival differences described above seems to be quite coherent. However, the results of the separate studies may have been influenced by their study design (for example, study population, measure of SES, period of diagnosis) and data analysis (for example, number of other factors for which adjustment was made in the survival analysis, the reporting of confidence

intervals). We will briefly mention some of the differences in study design and data analysis.

As we have already stated, study results might depend on the country of origin of the study population. We did not observe a systematic difference, however, in study results per country. Another important feature of study design concerns the measure of SES which is used in a study. In general, ecological measures are more prone to misclassification than measures based on individual characteristics. This misclassification is probably not related to the outcome and therefore results in a bias towards the null hypothesis. For example, in one study the measure of SES was based on the median male income per postcode of residence.[11] This measure was also applied to female survival data, therefore causing even more misclassification.

Some individual measures of SES, such as housing tenure,[13] are only rough indicators. This could account for the inconsistency of the results from this study with those from other studies, for example, for breast cancer.

Overall, studies using an ecological measure[4-7,11,12,15,19] did not differ substantially in their results from those using an individual measure of SES.[8,13,14,16-18]

The measure of outcome employed in a study on SES and cancer survival is another characteristic which may influence the study results. In studies using the relative survival rate as outcome, the expected survival rate is based on life tables of the general population. However, life expectancy of people from lower socioeconomic groups is lower than life expectancy of the general population. Therefore, their relative survival rate is underestimated, while for higher socioeconomic groups it is overestimated. Karjalainen and Pukkala[17] compared socioeconomic differences in relative and corrected survival rates and showed that by using the relative survival rate the absolute difference in rates between the highest and lowest social class was larger. The ratio of survival rates of the highest and lowest social class was very similar using either the relative or corrected survival rate however. A small overestimation of socioeconomic differences in cancer patient survival can result from using the relative survival rate, as in the Swedish study.[16] For cancer of the cervix only, however, this study[16] does show a larger difference in survival according to SES than the other studies.[12,14,15]

Although it was not clear whether correction for deaths from causes other than cancer was made in four studies,[5,8,12,14] results of these studies did not differ substantially from those of other studies.

The number and type of variables for which adjustment in the survival analysis was made also varied across studies, which made a comparison of results rather difficult. For cervical cancer, however, the results from three UK studies are consistent, although in the analysis of one study adjustment was made for many variables,[15] while in two other studies this was not the case.[12,14]

It is important to know, as we noted in the introduction, whether socioeconomic differences in cancer mortality are mainly caused by incidence or survival differentials. We therefore compared our findings on survival with published data

on cancer mortality according to SES for the six cancer sites which were studied. The selected studies concern patients diagnosed between the second half of the 1960s and the beginning of the 1980s in Finland,[1] Australia,[24,25] New Zealand,[26] Switzerland,[27] the UK[28-30] and the USA.[31]

For rectal cancer no association exists between SES and mortality.[25,26,30] Mortality is higher in lower socioeconomic groups for cancers of the lung[1,24-30] and cervix.[1,25,27,29] For cancers of the colon, prostate, and breast, either no mortality differences were found (colon,[25,30] prostate,[26,29,31] breast[27,29]) or there was a higher mortality in higher SES groups (colon,[24,26] prostate,[25,27,30] breast[1,25]).

If we compare our findings on cancer survival with the published data on socioeconomic differences in cancer mortality, we come to the following conclusions.

For lung cancer, the higher mortality in the lower socioeconomic groups cannot be ascribed to socioeconomic differences in survival, which seemed to be rather small and insignificant. Mortality differences must therefore be the result of differences in incidence. This is confirmed by findings from studies on SES and lung cancer incidence, which showed a higher incidence for the socially disadvantaged.[25,29,32-34]

For cancer of the cervix, higher mortality was found for the lower socioeconomic groups, while small survival differences were found in the reviewed papers for this cancer. These mortality differences must therefore be the result of the socioeconomic differences in cancer incidence which have been reported in several studies and which indicate a higher incidence in the lower SES groups.[25,29,33,35]

For cancers of the breast and colon, mortality was higher in higher SES groups in some, but not all, studies, while survival seemed to be better in these groups. Thus, for these cancers, the better survival for patients from higher SES groups could somewhat weaken the positive association between SES and mortality, or make it totally disappear in some situations. The incidence for these cancers is higher in the higher SES groups (breast,[25,33-36] colon[25,33,34,36]) which confirms that mortality differences for these cancers are also mainly caused by differences in incidence.

For rectal cancer, no socioeconomic gradient in mortality was found, but survival differences do exist. With regard to incidence too, no socioeconomic gradient was found,[25,33] which makes the evidence on the impact of incidence and survival differences according to SES on mortality differences rather inconclusive.

Finally for cancer of the prostate mortality was higher in higher SES groups in some studies, while results on survival were inconsistent. The mortality differences according to SES for this cancer seem to be caused by socioeconomic differences in incidence. This is confirmed by the finding in several studies that the incidence of this cancer is higher in men from high socioeconomic groups.[25,33,34,36,37]

We conclude that overall the impact of socioeconomic differences in cancer survival on differences in cancer mortality is low. Socioeconomic differences in

cancer mortality are mainly caused by differences in incidence. Health policy measures in the field of primary prevention aimed at known cancer risk factors should therefore be taken to reduce socioeconomic differences in cancer mortality.

References

1. Valkonen T, Martelin T, Rimpelä A. Socio-economic mortality differences in Finland 1971-1985. Helsinki: Central Statistical Office of Finland, 1990.
2. Fox AJ, Goldblatt PO. Longitudinal Study, socio-demographic mortality differentials 1971-1975. London: Office of Population Censuses and Surveys, Series LS no 1. London: HMSO, 1982.
3. Freedman LS. Tables of the number of patients required in clinical trials using the logrank test. Stat Med 1982;1:121-129.
4. Wegner EL, Kolonel LN, Nomura AMY, Lee J. Racial and socioeconomic status differences in survival of colorectal cancer patients in Hawaii. Cancer 1982;49:2208-2216.
5. Bassett MT, Krieger N. Social class and black-white differences in breast cancer survival. Am J Public Health 1986;76:1400-1403.
6. Dayal HH, Polissar L, Dahlberg S. Race, socioeconomic status, and other prognostic factors for survival from prostate cancer. J Natl Cancer Inst 1985;74:1001-1006.
7. Dayal H, Polissar L, Yang CY, Dahlberg S. Race, socioeconomic status, and other prognostic factors for survival from colo-rectal cancer. J Chron Dis 1987;40:857-864.
8. Morrison AS, Lowe CR, MacMahon B, Ravnihar B, Yuasa S. Incidence risk factors and survival in breast cancer: report on five years of follow-up observation. Eur J Cancer 1977;13:209-214.
9. Salber EJ, Trichopoulos D, MacMahon B. Lactation and reproductive histories of breast cancer patients in Boston, 1965-1966. J Natl Cancer Inst 1969;43:1013-1024.
10. Yuasa S, MacMahon B. Lactation and reproductive histories of breast cancer patients in Tokyo, Japan. Bull World Health Organ 1970;42:195-204.
11. Bonett A, Roder D, Esterman A. Determinants of case survival for cancers of the lung, colon, breast, and cervix in South Australia. Med J Aust 1984;141:705-709.
12. Milner PC, Watts M. Effect of socioeconomic status on survival from cervical cancer in Sheffield. J Epidemiol Community Health 1987;41:200-203.
13. Kogevinas M, Marmot MG, Fox AJ, Goldblatt PO. Socioeconomic differences in cancer survival. J Epidemiol Community Health 1991;45:216-219.
14. Murphy M, Goldblatt P, Thornton-Jones H, Silcocks P. Survival among women with cancer of the uterine cervix: influence of marital status and social class. J Epidemiol Community Health 1990;44:293-296.
15. Lamont DW, Symonds RP, Brodie MM, Nwabineli NJ, Gillis CR. Age, socio-economic status and survival from cancer of cervix in the West of Scotland 1980-1987. Br J Cancer 1993;67:351-357.
16. Vågerö D, Persson G. Cancer survival and social class in Sweden. J Epidemiol Community Health 1987;41:204-209.
17. Karjalainen S, Pukkala E. Social class as a prognostic factor in breast cancer survival. Cancer 1990;66:819-826.
18. Auvinen A. Social class and colon cancer survival in Finland. Cancer 1992;70;402-409.
19. Brenner H, Mielck A, Klein R, Ziegler H. The role of socioeconomic factors in the survival of patients with colorectal cancer in Saarland/Germany. J Clin Epidemiol 1991;44:807-815.
20. Vandenbroucke JP, Hofman A, Stiphout WAHJ van. Grondslagen der epidemiologie. Utrecht: Bunge, 1991:157-158.
21. Parkin DM, Hakulinen T. Analysis of survival. In: Jensen OM, Parkin DM, MacLennan R, Muir CS, Skeet RG (eds). Cancer registration: principles and methods. IARC scientific publications no. 95. Lyon: IARC, 1991.
22. Berg JW, Ross R, Latourette HB. Economic status and survival of cancer patients. Cancer 1977;39:467-477.
23. Hanai A, Fujimoto I. Survival rate as an index in evaluating cancer control. In: Parkin DM, Wagner G, Muir CS (eds). The role of the registry in cancer control. IARC scientific publications no. 66. Lyon: IARC, 1985.
24. McMichael AJ. Social class (as estimated by occupational prestige) and mortality in Australian males in the 1970s. Community Health Studies 1985;9:220-230.
25. Williams J, Clifford C, Hopper J, Giles G. Socioeconomic status and cancer mortality and incidence in melbourne. Eur J Cancer 1991;27:917-921.

26. Pearce NE, Howard JK. Occupation, social class and male cancer mortality in New Zealand, 1974-1978. Int J Epidemiol 1986;15:456-462.
27. Levi F, Negri E, La Vecchia C, Te VC. Socioeconomic groups and cancer risk at death in the Swiss canton of Vaud. Int J Epidemiol 1988;17:711-717.
28. Balarajan R, McDowall ME. Regional socioeconomic differences in mortality among men in Great Britain today. Public Health 1988;102:33-43.
29. Kogevinas E. Longitudinal Study. Socio-demographic differences in cancer survival. Office of population censuses and surveys, Series LS no. 5, London: HMSO, 1990.
30. Davey-Smith G, Leon D, Shipley MJ, Rose G. Socioeconomic differentials in cancer among men. Int J Epidemiol 1991;20:339-345.
31. Ernster VL, Selvin S, Sacks ST, Austin DF, Brown SM, Winkelstein W. Prostatic cancer: mortality and incidence rates by race and social class. Am J Epidemiol 1978;107:311-320.
32. Devesa SS, Diamond EL. Socioeconomic and racial differences in lung cancer incidence. Am J Epidemiol 1983;118:818-831.
33. Vågerö D, Persson G. Occurrence of cancer in socioeconomic groups in Sweden. Scan J Soc Med 1986;14:151-160.
34. Sharp L, Black RJ, Harliness EF, Finlayson AR, Muir CS. Cancer registration statistics Scotland 1981-1990. Edinburgh: Scottish Cancer Intelligence Unit, 1993.
35. Devesa SS, Diamond EL. Association of breast cancer and cervical cancer incidences with income and education among whites and blacks. J Natl Cancer Inst 1980;65:515-528.
36. Rimpelä AH, Pukkala EI. Cancers of affluence: positive social class gradient and rising incidence trend in some cancer forms. Soc Sci Med 1987;24:601-606.
37. Ross RK, McCurtis JW, Henderson BE, Menck HR, Mack TM, Martin SP. Descriptive epidemiology of testicular and prostatic cancer in Los Angeles. Br J Cancer 1979;39:284-292.

Chapter 3. Materials and methods

3.1 Data sources

The studies reported in this thesis are based on different data sources. The majority of the data came from two population based cancer registries, which will be discussed in part 1 of this chapter. Data for the study on socioeconomic variation in cancer survival in the Southeastern Netherlands came from the Eindhoven cancer registry (paragraph 3.1.1), while the association between deprivation and survival in the South Thames area was studied with data from Thames Cancer Registry (paragraph 3.1.2). Some basic issues involving the quality of cancer registration are also discussed (paragraph 3.1.3), as is the Longitudinal Study on Socio-Economic Health Differences (LS-SEHD). This study provided data to study the association between SES and a number of prognostic factors in the Southeastern Netherlands (paragraph 3.1.4).

3.1.1 Eindhoven Cancer Registry, The Netherlands

This regional cancer registry is population based and started operating in 1955. It is the oldest regional cancer registry in the Netherlands. In 1985 (midyear of the study period) the registry covered an area of about 2500 km^2, with almost 1 million inhabitants (7% of the Dutch population) in the Southeastern part of the Netherlands. Since 1989, the mid-western part of the province Brabant is also covered by the registry, resulting in a total population of about 2.2 million inhabitants. In this study we cover the period 1980-1989 and therefore we only report on patients living in the area of about 1 million inhabitants as mentioned above.

Registration is based on notifications of newly diagnosed cases from the departments of pathology, surgery and other hospital departments, as well as from the regional radiotherapy institute and from medical records departments. Data are collected from the medical records of the newly diagnosed patients during regular visits to these institutions, generally within 6 months after diagnosis. Incidence for the 1980's has been reported.[1-3]

The (active) follow-up of deaths consists of systematic checks of the vital status of patients, both through hospitals and in municipal population registers. Less than 1% of the patients diagnosed in the period 1975 to 1985 proved to be lost to follow-up.[4] In the survival study reported in this thesis, follow-up of patients ends at July 1, 1991.

3.1.2 Thames Cancer Registry, Great Britain

This population based cancer registry has been recording cancer in the population of South East England since 1960. Until 1984, it covered the territory of the South East and South West Thames Regional Health Authority (RHA) and in 1985 coverage was extended to North East and North West Thames RHAs. For the survival study reported in this thesis, only data from South East and South West Thames were used, as the study concerned the period of diagnosis between 1980

and 1989. In the remainder of this thesis the total of both areas will be referred to as South Thames. The registry covers an area which contains about a quarter of the population of England and Wales (14 million people).

Data are collected actively from hospitals and other health care facilities which include pathology, haematology and cytology laboratories, wards and outpatient units, and departments of radiotherapy. Furthermore, death certificates are an important source of information, as will be described in the next section of this chapter. Incidence for the 1980s has been reported.[5-8]

The follow-up of deaths of cancer patients is passive, which means that all deaths (both cancer and non-cancer deaths) are notified to the Registry, cancer deaths by the Office of Population Censuses and Surveys, and deaths due to other causes of people already registered with cancer by the National Health Service Central Register[7,8]. Up to 4% of cancer registrations remain untraced at the latter register.[9] In the survival study reported in this thesis, the follow-up of patients ends at December 31, 1992.

3.1.3 Quality of cancer registry data

The quality of cancer registry data concerns both the validity of the recorded information and the completeness of registration. In this paragraph we will discuss three indicators of data quality: two indicators of validity and one indicator of completeness. These indicators were used by the editors of Cancer Incidence in Five Continents (volume VI)[10], to judge on the suitability of registry data to be included in this monograph. Both the Eindhoven and Thames Cancer Registry contribute data to this monograph.

Histological verification

Validity of cancer registration can be defined as the proportion of cases recorded with a given characteristic (e.g. sex, age, cancer site) which truly have the attribute. One commonly used indicator of the validity of diagnostic information is the percentage of cancer registrations confirmed by histology (HV%).[10] Histological verification of suspected tissue by a histopathologist is usually taken as the gold standard of diagnostic evidence. Cases registered without histological confirmation of diagnosis may often have advanced disease, be older or receive palliative care and they may therefore have a lower survival than histologically confirmed cases. On the other hand, some of these cases may not have cancer at all. The HV% is assessed per cancer site, thus taking into account the possibility that reliable alternative diagnostic methods are available.[10] In Table 1, the HV% contains both cases diagnosed by histology and cytology and it is clearly higher in the Southeastern Netherlands as compared to the South Thames area, both for all sites combined as well as for the most common cancers separately.[10]

Death certificate only cases

A high percentage of cases registered on the basis of a death certificate only is generally considered to be a negative indicator of validity. This indicator shows for how many registrations no other information than a death certificate mentioning cancer can be obtained. In countries where the death certificate is a public document, cancer registries obtain information about persons dying with cancer in the registry's territory; cancer can be the underlying or contributing cause of death. This procedure is followed by Thames Cancer Registry, but not by the Eindhoven Cancer Registry, as the death certificate is not a public document in the Netherlands.

If a patient, notified through a death certificate, is not already known to the Thames Cancer Registry, data on the clinical diagnosis, date of diagnosis and treatment is searched for.[11,12] For about one third of these patients, clinical details could not be found. These cases are the real death-certificate-only (DCO) registrations who made up almost 20 percent of all registrations in the period 1983-1987 (table 1). The percentage of DCO cases is higher in cancers with a low survival rate (lung and stomach) as compared to cancers with an overall better survival (colorectum, prostate and breast). Furthermore, access to specialised care may also be an important determinant of the proportion of DCO-cases.

The date of diagnosis of DCO cases is unknown and they can therefore not be used in survival calculations. If most DCO cases visited a physician in the terminal stage of their life and therefore no treatment was initiated, the survival rate without these DCO cases would be an overestimation of the true survival rate in the population, as the DCO cases have a lower survival rate.

Mortality/incidence ratio

Completeness of cancer registration is the proportion of all incident cancers in the target population which are included in the data base of a cancer registry. Incompleteness can be minimized by using multiple data sources from a wide variety of sectors of the health care system where cancer patients are diagnosed and treated.

One indirect method of measuring completeness is to compare the number of cancer registrations with the number of cancer deaths in the same population and time period, which results in the mortality/incidence (M/I) ratio. If this ratio exceeds 1, it is usually a signal of incompleteness. The M/I ratio will be equal to (1-survival probability) in a steady state of constant incidence and survival and if reporting of cause of death was accurate. Site specific evaluation of the M/I ratio is necessary, as for cancers with a poor survival the ratio will be close to 1, while for cancers with a good survival the M/I ratio will be lower. A direct comparison of the M/I ratio in both areas is not possible; e.g. because overall survival is higher for most cancer sites in the Southeastern Netherlands than in the South Thames area.[13] The M/I ratios in table 1 are indeed mostly higher for the South Thames data as compared to the Southeastern Netherlands.

Table 1. Indices of data quality, six most common cancers and all sites*, Southeastern Netherlands and South Thames, 1983-1987[10]

	Southeastern Netherlands		South Thames	
	Males	Females	Males	Females
Lung				
HV%	89	86	54	50
DCO	-	-	22	24
M/I	98	95	93	91
Breast				
HV%	na	97	na	73
DCO	na	-	na	12
M/I	na	39	na	56
Colon				
HV%	93	94	67	64
DCO	-	-	19	22
M/I	63	69	72	71
Rectum				
HV%	97	97	77	73
DCO	-	-	13	16
M/I	47	44	59	60
Prostate				
HV%	95	na	69	na
DCO	-	na	16	na
M/I	53	na	63	na
Stomach				
HV%	94	91	59	49
DCO	-	-	24	29
M/I	82	90	89	88
All sites*				
HV%	88	90	63	65
DCO	-	-	19	18
M/I	73	58	75	68

HV%: % with histological verification; DCO: death certificate only; M/I: mortality/incidence ratio; na: not applicable * All sites but nonmelanoma skin cancer

An independent case ascertainment method to estimate completeness is to be preferred, as this involves a comparison of cancer registry data with an independent source of information.[14,15] No such direct measure of completeness for the 1980s is available for either of the registries. Recently, a comparison was made between the 1992 data of the Eindhoven Cancer Registry and data of the National Hospital Discharge Registry which registers diagnoses of all hospitalized people in the Netherlands. This comparison showed some incompleteness for pancreas cancer,

and for lung cancer in the elderly, while overall incompleteness was 2% (Coebergh JWW, personal communication). Thames Cancer Registry has recently carried out a research project to estimate the completeness of registration, using both routinely recorded information from the registry's data base and death certificates. This has shown that, five years after diagnosis, overall completeness was approximately 92% (Bullard J, personal communication).

We conclude that the HV% is relatively low for data from Thames Cancer Registry as compared with the Eindhoven Cancer Registry. Furthermore, the DCO% is rather high for the Thames data, but unfortunately this indicator of validity cannot be calculated for the Eindhoven Cancer Registry data, as the death certificate is not a public document in the Netherlands. Both the M/I ratio as indicator of completeness and more recent study results show that incompleteness is probably not very large in both areas.

3.1.4 The Longitudinal Study on Socio-Economic Health Differences

The Longitudinal Study on Socio-Economic Health Differences (LS-SEHD) is a prospective cohort study which started in 1991. For this study, an aselect sample (stratified by age, degree of urbanization and socioeconomic status) of approximately 27000 persons was drawn from the population registers in an area in the Southeastern part of the Netherlands, which is completely covered by the Eindhoven Cancer Registry. The persons in this sample received a postal questionnaire, resulting in a response rate of 70.1% (n=18973). There were small differences in response according to some background characteristics. Response was lower in the largest city Eindhoven (69%) as compared to the smallest municipalities (73%). The two lowest socioeconomic groups had a response rate of 68%, while it was 73% in the highest socioeconomic category (socioeconomic status was based on the postcode of residence). Women had a higher response rate (72.4%) than men (67.8%), while the response rate increased with age: 15-34 years (67.2%), 35-54 years (69.2%), 55-74 years (73.1%).

The LS-SEHD aims at assessing the contribution of different mechanisms and factors to the explanation of socioeconomic inequalities in health in the Netherlands. The postal survey contained questions on the highest level of education attained, and the occupational level of the respondent and occupation of the main breadwinner in the respondents' household. The indicators of health measured through the postal survey were: perceived general health, subjective health complaints and chronic conditions. Finally, a number of explanatory factors of socioeconomic inequalities in health have been measured: health-related life style factors, structural/environmental factors, psychosocial stress-related factors, childhood environment, cultural factors, psychological factors, and health in childhood.

Follow-up information of the participants in this study will be collected from different sources. Information on changes of address, marital status, and vital status

will be obtained from the population registers of the municipalities in the study area. Furthermore, the medical cause of death will be retrieved by linkage to the national cause-of-death register. The national hospital admission register will be used to measure the incidence of specific chronic conditions, by diagnosis at discharge and counting first admissions for each condition only. Finally, the Eindhoven Cancer Registry will be used to measure the incidence of cancer in the study population.[16]

3.2 Measures of Socioeconomic Status

3.2.1 Introduction

In most studies on socioeconomic variation in cancer survival, data from population based cancer registries have been used and in these, the socioeconomic status of individuals has rarely been measured directly. An alternative for individual measures of socioeconomic status are area-based measures, which have frequently been applied in the United States and the United Kingdom. In most studies on socioeconomic variation in cancer survival in these countries, census data have been used to determine the average socioeconomic level of each small area.[17-23]

In the Netherlands, the regional cancer registries do not contain data on the socioeconomic status (such as occupation and education) of individual cancer patients. Furthermore, recent census data are not available in the Netherlands, as the last census was held in 1971. We therefore used a measure of socioeconomic status which has been developed for marketing purposes, which is based on the place of residence at time of diagnosis of each individual cancer patient (paragraph 3.2.2).

We have also used an area-based measure of deprivation in our study on socioeconomic variation in cancer survival in the area covered by the South Thames RHA. The data base of Thames Cancer Registry does contain information on the occupation of cancer patients, but this is incomplete or missing for a large proportion of patients. Area-based measures of deprivation are much more integrated in British research as compared to the Netherlands, not in the least due to the availability of data from the ten-yearly census, which has been used to develop single and combined area-based measures of deprivation. One of these measures is the Carstairs Index,[24] a well-known measure of material deprivation which has been used in the British study (paragraph 3.2.3).

3.2.2 The Dutch Study

The measure of socioeconomic status developed for this study is area-based, as mentioned before. Through the postcode of residence at time of diagnosis, each patient was first assigned to one of 45 categories of a sociodemographic classification which was then collapsed into 3 or 5 categories. Several steps were taken to

derive the measure of socioeconomic status as used in this study, using information at different levels of aggregation which is described in the next paragraph. Furthermore, the results of studies which aimed at validating the area-based measure will be discussed.

Development of the area-based measure of socioeconomic status

Table 2 shows which steps were taken to develop the area-based measure of socioeconomic status and the information that was used at different levels of aggregation. We acquired data at level 3 from CCN marketing systems; the steps from level 1 to 2 and level 2 to 3 were implemented by CCN, while the step from level 3 to 4 was constructed by us.

Level 1 refers to the original data gathered by various agencies on a large number of socioeconomic and demographic characteristics of individual people. Examples of these socioeconomic variables are: occupation, education, and type of health insurance, while examples of demographic variables are: age, sex, and marital status. The majority of the data collected at level 1, came from face-to-face interviews in which questions were asked about all the members of a respondents' household. These interviews contained a question on the highest educational level attained by the main breadwinner in the household in which three categories were distinguished (low: primary school or lower vocational; intermediate: lower general or intermediate vocational; and high: intermediate/higher general, higher vocational, university).

These individual data from the interviews have been used by CCN marketing systems to estimate the average level and distribution of a number of socioeconomic and demographic characteristics at the postcode level (level 2). In this way, data are available on socioeconomic and demographic variables for each postcode area in the Netherlands (on average containing 16 households). Examples are: occupation (% of main breadwinners per postcode area in each of 5 categories), education (% of main breadwinners per postcode area in each of 3 categories), while examples of demographic variables are: the age-distribution and the average number of persons per household in each postcode area.

Table 2. Data used at each level of aggregation to derive the area-based measure of socioeconomic status in the Dutch study

Level	Individual	Postcode	45 sociodemographic categories	5 socioeconomic categories
	(1)	(2)	(3)	(4)
Data	socioeconomic and demographic data collected in interviews	average values and distribution of socioeconomic and demographic data	average values and distribution of socioeconomic and demographic data	average number of years of education

The information on approximately 20 variables and their separate categories at the postcode level (level 2) was used by the marketing agency to assign each postcode-area to one of 45 categories of a sociodemographic classification (level 3), using a non-hierarchical cluster analysis.[25] The resulting classification is a nominal typology of 45 categories and examples of descriptions given to some of these categories are: "rural with a high socioeconomic status", "higher income with older children", and "young with a high income".

The registration area of the Eindhoven Cancer Registry consists of 22,853 postcode areas, which, on the basis of a cluster analysis, have each been assigned to one of the 45 categories of the classification. The 45 categories were finally collapsed by us into 5 hierarchical socioeconomic categories (level 4). We calculated the average number of years of education at the national level for each of these 45 categories, ordered them according to this number, and divided the distribution into quintiles based on the percentage of persons in the Netherlands living in postcode sectors belonging to each of the 45 categories.

The average number of years of education for each of the 45 categories was calculated by multiplying the percentage of main breadwinners in each of 3 educational categories by a corresponding number of years of education and taking the sum of the resulting three figures, using the following formula:

(7.5 x % with low educ.)+(10 x % with intermediate educ.)+(15 x % with high educ.)/total %

The 3 educational categories in this formula refer to the highest attained level of education (low: primary school or lower vocational; intermediate: lower general or intermediate vocational; and high: intermediate/higher general, higher vocational or university), while the corresponding number of years of education in the 3 educational categories was 7.5 in the lowest, 10 in the intermediate and 15 in the highest educational category.

Table 1 (appendix) shows to which of the 5 socioeconomic categories each of the 45 categories of the original classification has been assigned. These 5 socioeconomic categories were used in the survival analyses for cancers of the lung, breast, and colorectum. As the total number of patients for cancers of the prostate and stomach was relatively small, the 45 categories were also divided into 3 socioeconomic categories, based on tertiles of the underlying population (table A, appendix).

Results of the validation of the area-based measure of socioeconomic status

We have conducted different types of studies to validate the area-based measure of socioeconomic status, using data from the postal survey which is part of the baseline data collection of the Longitudinal Study on Socio-Economic Health Differences (LS-SEHD) (paragraph 3.1.4). For respondents to the LS-SEHD postal survey (n=18973), data were available on education, occupation, occupation of the main breadwinner, the score on the area-based measure of socioeconomic status and a number of health indicators.

Validation studies will be discussed with respect to individuals (level 1), postcodes (level 2) and 45 sociodemographic categories (level 3). Furthermore, the association between several indicators of socioeconomic status (individual education, occupation and the area-based measure in 3 or 5 categories) and several health indicators will be discussed.

1. Validation at the individual level

Firstly, we validated the area-based measure at the individual level by crosstabulating the area-based measure in 5 socioeconomic categories (level 4) with individual education in 4 categories (level 1), using data of respondents (n=18227) to the LSSEHD survey. The results of this comparison are shown in table 3, from which we observe a higher percentage of respondents with a low education in the lower categories of the area-based measure and a higher percentage of respondents with a high education in the higher categories of the area-based measure. Overall, correspondence between the measures is moderate, the Pearson correlation coefficient between the area-based measure and education at the individual level was 0.25. The results were very similar for all possible combinations of the area-based measure (3 or 5 categories) and education at the individual level (3, 4 or 7 categories). Similar values have been reported from another study[28] in which the Pearson correlations between individual-level and census tract-level socioeconomic variables ranged between 0.2 and 0.4, using several types of socioeconomic variables. These levels of correlations imply an underestimation of the association between individual socioeconomic status and cancer survival if area-based measures of socioeconomic status are used, under the assumption of nondifferential misclassification of socioeconomic status.[26]

Table 3. Association between the area-based measure of socioeconomic status (5 categories) and individual level education (4 categories)

	Education*				
Area-based	Low (1)	2	3	High (4)	Total
Low (1)	34.2	42.4	16.7	6.7	100
2	25.4	43.7	20.4	10.5	100
3	22.0	40.4	22.1	15.5	100
4	19.1	37.5	24.8	18.6	100
High (5)	14.8	34.5	25.5	25.2	100
Total	23.3	39.4	21.8	15.5	100

* Measured as: (1) primary school only; (2) lower vocational and lower general; (3) intermediate vocational and intermediate/higher general; (4) higher vocational and university

2. Validation at the postcode level

Information on education at the postcode level was used by the marketing agency to assign each postcode to one of 45 sociodemographic categories. We selected the postcode areas for which at least 6 respondents were found in the LS-SEHD survey (381 out of a total of 2615 postcodes) and we calculated the average number of years of education per postcode area with the formula as described on page 30. For each of the 381 postcodes, the average number of years of education as derived from the original classification could be compared with the average number of years of education as calculated with the LS-SEHD survey data on education. Both average numbers were subsequently used to assign a postcode to one of the 5 socioeconomic categories, applying the classification as given in table 1 (appendix).

The distribution of postcodes across 5 socioeconomic categories based on either of these sets of average numbers of years of education was compared. Table 4 shows that 52% of the postcodes that had been assigned to the lowest category of the area-based measure, was also assigned to the lowest category if LS-SEHD survey data were used to calculate the average number of years of education. The percentage of postcodes which was assigned to the same socioeconomic category was much lower for the categories 2 to 4, while it was about 61% for the highest category.

The Pearson correlation coefficient between the two variables was 0.51. We may conclude from this exercise that although misclassification is fairly substantial, validity at the postcode level is satisfactory, given that the assignment of postcodes to one of 45 categories by the marketing agency was based on a large number of socioeconomic and demographic variables, and that education was only one of them.

Table 4. Association between the measure of socioeconomic status in 5 categories based on individual data from the LS-SEHD postal survey and from the classification of 45 sociodemographic categories

	Categories based on LS-SEHD data					
Original classification	Low (1)	2	3	4	High (5)	Total
Low (1)	52.3	16.6	13.5	7.3	10.3	100
(2)	20.8	33.2	20.8	6.3	18.9	100
(3)	19.5	9.8	24.3	9.8	36.6	100
(4)	16.7	20.0	20.0	20.0	23.3	100
High (5)	1.4	11.6	13.2	13.0	60.8	100
Total	32.8	17.3	16.0	9.4	24.5	100

3. Validation at the level of 45 sociodemographic categories

We compared the average number of years of education according to marketing agency data for each of the 45 categories (level 3) with the average number of years of education according to LS-SEHD survey data. The LS-SEHD survey data directly apply to the inhabitants of the postcode areas included in the analysis. We included only 32 of the 45 sociodemographic categories in this analysis, as for these at least 100 respondents were found in the LS-SEHD survey. In both series of analyses, the calculation was done with the same formula (page 30) and using the same 3 categories of education.

Table 5. Average number of years of education per sociodemographic category based on marketing agency data or LS-SEHD postal survey data

Socio-demo. cat.*	Resp. (N)	Average education agency†	Average education LS-SEHD‡	Socio-demo. cat.*	Resp. (N)	Average education agency†	Average education LS-SEHD‡
22	721	7.8	8.6	8	468	10.1	10.2
21	181	8.2	9.0	31	301	10.2	10.2
39	2056	8.2	9.0	6	155	10.3	11.5
26	112	8.2	8.9	11	163	10.3	11.7
18	256	8.5	9.0	12	822	10.6	10.5
20	1449	8.6	9.0	27	251	10.6	10.7
25	320	8.8	9.9	17	240	10.8	11.1
36	182	8.9	8.5	32	192	10.9	10.7
23	177	8.9	9.9	5	1148	10.9	10.4
38	928	8.9	9.6	2	610	11.0	10.4
24	162	9.2	9.9	28	176	11.0	10.3
44	231	9.4	9.2	30	123	11.0	11.0
7	665	9.5	9.3	4	1358	11.2	10.6
19	1333	9.7	9.6	1	1239	12.8	11.9
10	473	9.7	10.7	29	159	13.4	11.6
9	305	9.8	10.0	3	881	13.8	11.7

* The 32 categories were ordered by the average number of years of education based on marketing agency data
† average years of education as calculated with data from the marketing agency
‡ average years of education as calculated with data from the LS-SEHD postal survey

Table 5 shows that the variation in average number of years of education between the sociodemographic categories is larger when marketing agency data are used as compared to data from the LS-SEHD survey. This could mean that socioeconomic contrast is smaller in the study area of the LS-SEHD as compared to the

Netherlands at the national level, to which the marketing data apply. The Spearman rankcorrelation coefficient between these two series was 0.87, which implies that the area-based measure is a good indicator of the average level of education for each of the 45 sociodemographic categories.

4. Study of the association between socioeconomic status and health, using both individual and area-based measures of socioeconomic status

The validation studies carried out at different levels of aggregation suggest an underestimation of the association between individual level socioeconomic status and several health measures, such as cancer survival. We determined the association between the area-based measure of socioeconomic status and health and compared this with the effect of a number of individual measures of socioeconomic status on health, using data from the LS-SEHD postal survey. The measures of socioeconomic status in this analysis were: the area-based measure both in 5 and in 3 categories, the education of the respondent in 7 and 3 categories, the occupation of the respondent in 6 categories, and the occupation of the main breadwinner in 6 categories. Occupation was classified according to the Erikson, Goldthorpe and Portocarero (EGP) scheme, which consists originally of ten levels.[27] The 6 categories distinguished in this analysis were: (1) unskilled manual workers and low skilled manual workers, (2) high skilled manual workers, (3) self-employed, (4) routine non-manual employees, (5) lower grade professionals, (6) higher grade professionals. The measures of outcome were: perceived general health, subjective health complaints and chronic conditions. Perceived general health was measured by the answer on the question "How do you rate your health in general?". The answer was dichotomized into (very) good versus less than good (fairly good; sometimes good, sometimes bad; bad). Subjective health complaints were measured by means of a checklist, containing 13 questions on complaints about the heart, stomach, etc. This variable was dichotomized into 3 or less versus 4 or more complaints. Finally, the number of chronic conditions was measured through a checklist of 23 chronic conditions (e.g. diabetes, low back pain, cancer, heart disease etc), and the variable was dichotomized into none versus at least one chronic condition. The following confounders were taken into account: sex, age (5 year groups), marital status (4 categories), religious affiliation (5 categories), and degree of urbanization (5 categories). The association between socioeconomic status and each of these health measures was expressed in an odds ratio and 95% confidence interval, resulting from logistic regression analyses, after adjustment for confounding variables.

From table 6 it can be seen that for perceived general health the gradient in odds ratios is the same for each measure of socioeconomic status. The odds of having less than good perceived general health is consistently higher in the lower socioeconomic categories. The strength of the association between socioeconomic status and perceived health is greater for the individual measures than for the area-based measures of socioeconomic status, as indicated by the odds ratios, which are

higher for occupation and education than for the area-based measures. The results for subjective health complaints and chronic conditions are similar to those for perceived health. Overall, the gradient for the different socioeconomic measures is in the same direction but much weaker when the area-based measure is used. For chronic conditions, the odds ratios are much smaller in general than for the other two health indicators and the gradient is also less consistent, especially when occupation is used as socioeconomic indicator.

We conclude that the association between individual level socioeconomic status and health (perceived general health, subjective health complaints, chronic conditions) is underestimated when area-based measures of socioeconomic status are used. This finding should be carefully extrapolated to the survival analyses of the Dutch study reported in this thesis, as survival is an objective health measure and the health measures as reported in the postal survey are subjective. We may assume however, that the association between individual level socioeconomic status and cancer survival will also be underestimated which is also indicated by the results of the other validation studies described in this chapter. We do not know the size of the underestimation however.

Table 6. Association between socioeconomic status and perceived general health, subjective health complaints, and chronic conditions: odds ratio and 95% confidence interval adjusted for age, sex, religion, marital status and degree of urbanization

SES measure	% in each category	Health measure: Perceived general health (less than good)	Subjective health complaints (at least 4)	Chronic conditions (at least 1)
Area-based				
5 categories*				
1 (low)	26.6	1.92 (1.74-2.12)	1.49 (1.36-1.64)	1.16 (1.06-1.26)
2	14.9	1.57 (1.39-1.76)	1.35 (1.22-1.51)	1.13 (1.02-1.25)
3	17.1	1.37 (1.22-1.53)	1.14 (1.02-1.26)	1.12 (1.01-1.23)
4	14.6	1.33 (1.18-1.50)	1.16 (1.04-1.29)	1.11 (1.00-1.23)
5 (high)	26.8	1.00	1.00	1.00
3 categories*				
Low	35.4	1.71 (1.57-1.86)	1.43 (1.32-1.55)	1.12 (1.04-1.20)
Medium	28.8	1.30 (1.19-1.42)	1.13 (1.04-1.23)	1.08 (1.00-1.16)
High	35.8	1.00	1.00	1.00
Education				
7 categories*				
1 (low)	20.5	4.93 (3.93-6.20)	3.27 (2.70-3.96)	1.39 (1.18-1.63)
2	23.5	3.33 (2.65-4.17)	2.27 (1.88-2.74)	1.30 (1.11-1.52)
3	14.7	2.25 (1.78-2.85)	1.92 (1.58-2.33)	1.30 (1.10-1.53)
4	14.2	2.36 (1.87-2.99)	1.90 (1.57-2.31)	1.48 (1.26-1.74)
5	8.3	1.63 (1.25-2.11)	1.57 (1.27-1.94)	1.18 (0.99-1.42)
6	13.5	1.60 (1.26-2.04)	1.50 (1.23-1.82)	1.18 (1.00-1.39)
7 (high)	5.3	1.00	1.00	1.00
3 categories*				
Low	44.0	2.67 (2.42-2.95)	1.87 (1.71-2.04)	1.17 (1.08-1.27)
Medium	28.9	1.54 (1.38-1.71)	1.34 (1.22-1.47)	1.21 (1.11-1.32)
High	27.1	1.00	1.00	1.00
Occupation				
Respondent†				
1 (low)	21.4	3.34 (2.70-4.13)	2.23 (1.83-2.71)	1.24 (1.05-1.46)
2	19.5	2.58 (2.09-3.20)	1.99 (1.64-2.42)	1.25 (1.05-1.48)
3	3.8	2.63 (1.98-3.48)	2.03 (1.55-2.65)	1.07 (0.84-1.37)
4	23.0	1.62 (1.30-2.02)	1.42 (1.17-1.73)	1.34 (1.13-1.59)
5	25.0	1.27 (1.02-1.57)	1.20 (0.99-1.45)	1.44 (1.22-1.71)
6 (high)	7.3	1.00	1.00	1.00
Breadwinner†				
1 (low)	18.2	3.11 (2.55-3.78)	2.22 (1.85-2.66)	1.15 (0.99-1.34)
2	23.2	2.44 (2.02-2.96)	1.91 (1.61-2.28)	1.23 (1.05-1.45)
3	4.1	2.29 (1.75-2.98)	1.85 (1.44-2.38)	0.95 (0.75-1.20)
4	19.4	1.55 (1.27-1.90)	1.51 (1.26-1.81)	1.30 (1.11-1.51)
5	26.2	1.19 (1.02-1.45)	1.19 (0.99-1.41)	1.41 (1.20-1.66)
6 (high)	8.9	1.00	1.00	1.00

* n=17255 † n=10777

3.2.3 The British study

In this study, an area-based measure of deprivation was used. Through the postcode of residence at diagnosis, each patient was assigned to a census enumeration district and subsequently to 1 of 5 deprivation categories. In this paragraph, this process will be explained. Furthermore, the association between social class based

on individual occupation and the area-based measure will be discussed.

Calculation of scores on the Carstairs Index
Theoretically, 4 levels of aggregation can be distinguished in the process to derive the Carstairs Index, which is shown in table 7. Data on individuals were obtained from the 1981 census (level 1) on many variables. The smallest level of aggregation of census data in the British study is the level of census enumeration districts (level 3) (average 400 households), as census data at the postcode level was not directly available to us (level 2).

Table 7. Level of aggregation and data used at each level to derive the area-based measure of material deprivation (Carstairs Index)

Level	Individual	Postcode	Census Enumeration District	5 categories of material deprivation
	(1)	(2)	(3)	(4)
Data	socioeconomic and demographic data collected in interviews	socioeconomic and demographic data	average values for 4 indicators of material deprivation	average values and distribution of 4 indicators of material deprivation

Data from the 1981 census were used on 4 indicators of material deprivation[24] at the census enumeration district level. These indicators are overcrowding (persons in private households living at a density of more than 1 person per room, as a proportion of all persons living in private households), male unemployment (proportion of economically active males seeking work), low social class (proportion of all persons in private households with head of household in social class IV or V), and car ownership (proportion of all persons in private households without a car). Information on these four variables was combined into a single score: for each enumeration district, the average for Great Britain on the four variables was subtracted from the actual value and then divided by the population (Great Britain) standard deviation (s.d.). Then the sum of the scores on the four variables was calculated.
The Carstairs scores for all enumeration districts in Great Britain were ranked from low ('affluent') to high ('deprived') and the distribution was divided into quintiles, which resulted in 5 categories of material deprivation (level 4). Each patient could be assigned to one of these 5 categories through the postcode of residence (level 2) at time of diagnosis, which was linked to the corresponding census enumeration district (level 3).

Validation of the Carstairs index at the individual level
As mentioned before, data on social class based on occupation has limitations, e.g. the variable is incomplete or missing for a large proportion of patients. Despite

such limitations, a comparison was made between two different measures of socioeconomic status. The first is social class based on occupation of individual cancer patients (5 categories: I. Professional; II. Intermediate; III. Skilled and unskilled manual; IV. Partly skilled; V. Unskilled)[28] and the second is the Carstairs Index in 5 categories. All patients for whom both measures of socioeconomic status were known were selected for this comparison, which resulted in 36% of all patients that had already been selected for survival analysis (n=55501). The percentage of patients for which both measures were known was low for the most common female cancers: breast (16%), uterus (19%), cervix (22%) and ovary (23%), but higher in the other cancers: colorectum (33%), bladder (39%) pancreas (42%), stomach (43%), prostate (44%), and lung (50%).

As the distribution of social class across deprivation categories did not differ substantially for men and women and between cancer sites (results not shown), we combined data on both sexes and the 10 most common cancers. Table 8 shows a higher percentage of patients with social class IV or V in the lower categories of the Carstairs index. For social class III we found little contrast, while the percentage of patients in both social class I and II was higher in the higher categories of the Carstairs Index. The correlation coefficient between social class and the Carstairs Index was 0.22, which is in agreement with the findings from the Dutch study on the correlation between individual and area-based measures of SES. This finding is not surprising, as also in the British study we used an area-based measure of socioeconomic status in which several variables were used to calculate a summary score, and social class was only one of these variables. Again, under the assumption that misclassification is nondifferential, one may expect an underestimation of the association between deprivation and survival at the individual level if the area-based measure of deprivation is used. However, we should carefully interpret these findings, as they only apply to 36 percent of the study population and this selection could have resulted in a bias of the results.

Table 8. Distribution of social class within deprivation categories, 10 most common cancer sites*, South Thames, 1980-1989

	Social class					
Carstairs	V (low)	IV	III	II	I (high)	Total
1 (low)	16.9	20.9	49.1	11.3	1.8	8.2
2	10.7	20.1	52.8	14.2	2.2	18.3
3	7.5	18.0	51.9	19.5	3.1	21.8
4	5.1	16.4	50.0	23.7	4.8	24.1
5 (high)	3.7	13.1	44.8	30.4	8.0	27.6
Total	7.2	16.9	49.4	21.9	4.6	100

* 36 percent of total number of patients diagnosed between 1980 and 1989

Socioeconomic status and survival, using both social class and the area-based measure of socioeconomic status.

We compared the socioeconomic gradient in survival for patients for whom both the Carstairs index and social class were known (see table 9). It is important to note that the comparison was only made for a selection of patients, as only for about a third of all patients we did have information on both measures of socioeconomic status. Furthermore, the overall survival for the 3 most common cancers differs for this selection of patients from the survival for all patients (see also chapter 5).

For cancer of the lung, the gradient in survival is more consistent and steeper when the Carstairs Index is used as compared to social class. For breast cancer, the gradient is steeper when individual social class is the measure of socioeconomic status. This is also true for colorectal cancer but to a larger degree; the gradient is absent when the Carstairs Index is used, while it is rather steep using social class (table 10).

From this comparison we may draw the tentative conclusion that the association between individual social class and survival is underestimated for cancer of the breast and colorectum, and overestimated for lung cancer when the Carstairs Index is used.

Table 9. Number and percentage of patients by deprivation and social class, cancers of the lung, breast, colorectum, South Thames, 1980-1989

	Lung				Breast				Colorectum			
SES	Carstairs		Social Class		Carstairs		Social Class		Carstairs		Social Class	
High	4949	24.5	636	3.1	1507	31.5	285	6.0	2769	29.8	548	5.9
2	4700	23.3	3757	18.6	1100	23.0	1362	28.5	2328	25.1	2207	23.8
3	4522	22.4	10182	50.5	1029	21.5	2317	48.5	2025	21.8	4528	48.8
4	4013	19.9	3805	18.9	777	16.3	580	12.1	1582	17.0	1463	15.8
Low	1979	9.8	1783	8.8	365	7.6	234	4.9	580	6.2	538	5.8
Total	20163	100	20163	100	4778	100	4778	100	9284	100	9284	100

Table 10. 5 year RSR by deprivation and social class, cancers of the lung, breast, colorectum, South Thames

	Lung		Breast		Colorectum	
SES:	Carstairs	Social Class	Carstairs	Social Class	Carstairs	Social Class
High	7.0 (6.2-7.8)	6.6 (4.3-8.9)	62 (58-66)	61 (54-68)	29 (27-31)	32 (27-37)
(2)	6.5 (5.6-7.4)	7.0 (6.0-8.0)	58 (54-62)	63 (60-66)	32 (30-34)	34 (31-37)
(3)	6.6 (5.8-7.4)	6.2 (5.6-6.8)	56 (53-59)	58 (55-61)	30 (27-33)	29 (27-31)
(4)	5.2 (4.4-6.0)	6.1 (5.2-7.0)	56 (52-60)	51 (45-57)	26 (23-29)	26 (23-29)
Low	5.6 (4.4-6.8)	6.3 (4.9-7.7)	55 (49-61)	51 (42-60)	30 (25-35)	23 (18-28)
Total	6.3 (5.9-6.7)	6.3 (5.9-6.7)	58 (56-60)	58 (56-60)	30 (28-32)	30 (28-32)
Ratio High/Low	1.25	1.05	1.13	1.20	0.97	1.39

3.3 Survival analyses

In the last paragraph of this chapter we discuss the most important measures of outcome in the studies reported in this thesis. In univariate survival analyses (paragraph 3.3.1) the measure of outcome was the relative survival rate and in the multivariate analyses (paragraph 3.3.2) it was the hazard ratio.

3.3.1 Univariate survival analyses

Cancer survival is a measure of outcome to quantify the effect of cancer detection and treatment on the natural history of the disease, and can simply be expressed as the percentage of patients alive at a certain point in time after diagnosis. However, two components of total mortality need to be considered. The first one is mortality from the specific cancer under study and the second is mortality from other causes. Survival rates should be calculated with adjustment for mortality from causes other than the specific cancer under study, to indicate the mortality which can be attributed to cancer.

One way to adjust for mortality from other causes, is to obtain information on the cause of death of each cancer patient, followed by a survival analysis in which patients who died from other causes than the specific cancer under study are censored. This type of analysis results in the corrected survival rate and relies on information from death certificates. However, most cancer registries, such as the Eindhoven Cancer Registry, do not have access to death certificates.

Another approach to adjust for mortality from causes other than cancer, is to calculate the relative survival rate, which does not rely on information from death certificates. The relative survival rate (%) is the ratio of the observed survival rate

in a group of cancer patients to the survival rate expected in a group similar to the patients in such characteristics as age and sex, but free of the specific disease under study.[29] The expected survival rate is based on a life-table from the general population. In the studies reported in this thesis, we used the method of Hakulinen[30] to calculate expected survival, with Chiang's approximation.[31] The analyses were conducted with a computer program package from the Finnish Cancer Registry.[32]

An example: a 5 year relative survival rate of 60 percent for a group of breast cancer patients means that after 5 years of follow-up, 60 percent of the patients is still alive, considering that breast cancer is the only cause of death. If a relative survival rate is below 100 percent it means that, during the specified time interval, survival was lower and mortality was higher in the patient group than in a similar group of persons from the general population (who were free of the cancer under study). A relative survival rate may become 100 percent, which implies equal mortality in the patient group and the comparison group from the general population, or even exceed 100 percent for some cancer sites after a long period of follow-up, which implies that the life expectancy of the patients under study exceeds that of the general population.

In order to calculate relative survival for different socioeconomic groups, expected survival should be based on social class specific life tables. Such life tables are not available for either of the two study areas, and therefore we used one single life table to calculate expected survival for different socioeconomic groups: either the life-table of the regional population (Southeastern Netherlands for the registration area of the Eindhoven Cancer Registration) or the national population (England & Wales for the South Thames area). Through this procedure, expected survival is probably underestimated for the higher socioeconomic groups which results in an overestimation of the relative survival rate in these groups. For lower socioeconomic groups, the expected survival could be overestimated and relative survival underestimated. The gradient in survival by socioeconomic status may thus be overestimated by using a single life table to calculate expected survival for each socioeconomic group. A Finnish study on breast cancer, in which both the corrected and relative survival by social class were calculated, showed that the overestimation is rather small for all patients combined. The ratio of corrected survival rates for the highest and lowest social class was 1.12 using the relative survival rate and 1.10 using the corrected survival rate.[33]

3.3.2 Multivariate survival analyses

In the studies described in this thesis, the model of Hakulinen and Tenkanen[34] was used to study the impact of several prognostic factors simultaneously on the annual excess mortality due to the cancer under study. The measure of outcome in the multivariate analyses is the hazard ratio, which can be defined as the probability of dying from the specific cancer under study (so adjusted for mortality from other

causes) for each category of the covariates included in the model.

The model is a multiple regression model for relative survival, in which total mortality is a combination of additive and multiplicative risks. The addition applies to expected mortality for demographically similar individuals in the general population and of the disease related mortality hazard. The latter is a multiplicative function of risk corresponding to the proportional hazards model: it depends on a covariate and a function constant by time.

The estimation of the parameters is based on the maximum likelihood method.[35] The statistical significance of each of the parameters corresponding to each covariate in the model is tested by the change in deviance which has a chi-square distribution with the number of degrees of freedom depending on the number of levels of the studied covariate. The multivariate analyses were carried out with GLIM.[36]

References

1. Bakker D, Coebergh JWW, Crommelin MA, Verhagen-Teulings MTh. Netherlands, Eindhoven. In: Muir CS, Waterhouse JAH, Mack T, Powel J, Whelan SL (eds). Cancer incidence in five continents, volume V. IARC scientific publications no. 88. Lyon: IARC, 1987: 574-579.
2. Coebergh JWW, Verhagen-Teulings MTh, Crommelin MA, Masseling E, Heijden LH van der. Netherlands, Eindhoven. In: Parkin DM, Muir CS, Whelan SL, Gao Y-T, Ferlay J, Powell J, eds. Cancer incidence in five continents, volume VI. IARC scientific publications no. 120. Lyon: IARC, 1992: 666-669.
3. Coebergh JWW, Heijden LH van der (eds). Cancer incidence and survival 1975-1987. Eindhoven: Comprehensive Cancer Centre South, Eindhoven Cancer Registry publications no. 2, 1991.
4. Coebergh JWW, Crommelin MA, Heijden LH van der, Hop WCJ, Verhagen-Teulings, MTh. Survival of cancer patients in southeastern Netherlands. Ned Tijdschr Geneeskd 1991;135:9-38-943.
5. Skeet RG, Thornton-Jones H, Murrells TJ. UK, England, South Thames Region, 1978-1982. In: Muir CS, Waterhouse JAH, Mack T, Powell J, Whelan SL, eds. Cancer incidence in five continents, volume V. IARC scientific publications no. 88. Lyon: IARC, 1987: 664-667.
6. Chamberlain J, Bourne HM, Thornton-Jones H. UK, England, South Thames Region, 1983-1987. In: Parkin DM, Muir CS, Whelan SL, Gao Y-T, Ferlay J, Powell J (eds). Cancer incidence in five continents, volume VI. IARC scientific publications no. 120. Lyon: IARC, 1992: 790-793.
7. Thames Cancer Registry. Cancer in South East Thames, 1987-1989: cancer incidence, prevalence and survival in residents of the District Health Authorities in South East Thames. Sutton: Thames Cancer Registry, 1992a.
8. Thames Cancer Registry. Cancer in South West Thames 1987-1989: cancer incidence, prevalence and survival in residents of the District Health Authorities in South West Thames. Sutton: Thames Cancer Registry, 1992b.
9. Swerdlow AJ. Cancer registration in England and Wales: some aspects relevant to interpretations of the data. J Roy Stat Soc 1986;149:146-160.
10. Parkin DM, Muir CS. Comparability and quality of data. In: Parkin DM, Muir CS, Whelan SL, Gao Y-T, Ferlay J, Powell J (eds). Cancer incidence in five continents, volume VI. IARC scientific publications no. 120. Lyon: IARC, 1992: 45-55.
11. Skeet RG. The Thames Cancer Registry. In: Jensen OM, Parkin DM, MacLennan R, Muir CS, Skeet RG, eds. Cancer registration: principles and methods. IARC scientific publications no. 95. Lyon: IARC, 1991: 237-245.
12. Thames Cancer Registry. Cancer in South East England, 1991: cancer incidence, prevalence and survival in residents of the District Health Authorities in South East England. Sutton: Thames Cancer Registry, 1994.
13. Berrino F, Sant M, Verdecchia A, Capocaccia R, Hakulinen T, Estève J (eds). Survival of cancer patients in Europe. The Eurocare study. IARC scientific publications no. 132. Lyon: IARC, 1995.
14. Parkin DM, Chen VW, Ferlay J, Galceran J, Storm HH, Whelan SL. Comparability and

quality control in cancer registration. IARC technical report no. 19. Lyon: IARC, 1994: 32-37.
15. Schouten LJ, Höppener P, Brandt PA van den, Knottnerus JA, Jager JJ. Completeness of cancer registration in Limburg, the Netherlands. Int J Epidemiol 1993;22:369-376.
16. Mackenbach JP, Mheen H van de, Stronks K. A prospective cohort study investigating the explanation of socio-economic inequalities in health in the Netherlands. Soc Sci Med 1994;38:299-308.
17. Wegner EL, Kolonel LN, Nomura AMY, Lee J. Racial and socioeconomic status differences in survival of colorectal cancer patients in Hawaii. Cancer 1982;49:2208-2216.
18. Bassett MT, Krieger N. Social class and black-white differences in breast cancer survival. Am J Public Health 1986;76:1400-1403.
19. Dayal HH, Power RN, Chiu C. Race and socio-economic status in survival from breast cancer. J Chronic Des 1982;35:675-683.
20. Dayal HH, Polissar L, Dahlberg S. Race, socioeconomic status, and other prognostic factors for survival from prostate cancer. JNCI 1985;74:1001-1006.
21. Dayal H, Polissar L, Yang CY, Dahlberg S. Race, socioeconomic status, and other prognostic factors for survival from colo-rectal cancer. J Chron Dis 1987;40:857-864.
22. Lamont DW, Symonds RP, Brodie MM, Nwabineli NJ, Gillis CR. Age, socio-economic status and survival from cancer of cervix in the West of Scotland 1980-1987. Br J Cancer 1993;67:351-357.
23. Carnon AG et al. Relation between socioeconomic deprivation and pathological prognostic factors in women with breast cancer. BMJ 1994;309:1054-1057.
24. Carstairs V, Morris R. Deprivation and Health in Scotland. Aberdeen: Aberdeen University Press, 1991.
25. Nijdam B, Buuren H van. Statistiek voor de sociale wetenschappen. Alphen a/d Rijn: Samsom, 1983.
26. Greenwald HP, Polissar NL, Borgatta EF, McCorkle R. Detecting survival effects of socioeconomic status: problems in the use of aggregate measures. J Clin Epidemiol 1994;47:-903-909.
27. Erikson R, Goldthorpe JH, Portocarero L. Intergenerational class mobility in three western European societies: England, France and Sweden. Br J Sociology 1979;30:415-441.
28. OPCS. Classification of occupations 1990. London: HMSO, 1990.
29. Ederer F, Axtell LM, Cutler SJ. The relative survival rate: a statistical methodology. National Cancer Institute Monographs 1961;6:101-121.
30. Hakulinen T. Cancer survival corrected for heterogeneity in patient withdrawal. Biometrics 1982;38:933.
31. Chiang CL. Introduction to stochastic processes in Biostatistics. New York: Wiley, 1968:272.
32. Hakulinen T, Abeywickrama KH. A computer program package for relative survival analysis. Comp Prog Biomed 1985;19:197-207.
33. Karjalainen S, Pukkala E. Social class as a prognostic factor in breast cancer survival. Cancer 1990;66:819-826.
34. Hakulinen T, Tenkanen L. Regression analysis of relative survival rates. Appl Stat 1987;36:-309-317.
35. Kleinbaum DG, Kupper KL, Muller E. Applied regression analysis and other multivariate methods. Boston: PWS-Kent, 1988: 483-519.
36. Payne CD. The Glim System Release 3.77. Generalized linear interactive modeling manual. Oxford: Numerical Algorithms Groups, 1985.

Appendix

Table 1. Percentage of main breadwinners in 3 educational categories, the average number of years of education, the percentage of inhabitants and categories of the area-based measure, for each of 45 sociodemographic categories in the Dutch study

Sociodemographic category	% low education*	% interm education*	% high education*	education: average years	% inhab, Dutch population	area-based measure (5 cat)	area-based measure (3 cat)
22	90.5	7.7	1.6	7.8	2.03	1	1
21	77.1	16.0	3.3	8.2	1.38	1	1
26	76.9	18.8	3.6	8.2	1.47	1	1
39	78.1	17.6	3.7	8.2	3.61	1	1
41	76.3	18.5	3.9	8.3	3.78	1	1
42	71.1	24.9	3.6	8.4	0.81	1	1
43	72.7	16.8	5.8	8.4	0.86	1	1
18	69.6	23.5	4.8	8.5	1.24	1	1
20	67.6	25.6	6.3	8.6	2.82	1	1
40	64.6	25.0	7.5	8.7	3.15	1	1
25	61.5	30.5	7.4	8.8	4.39	2	1
23	61.1	29.0	8.3	8.9	2.61	2	1
36	62.3	28.3	8.6	8.9	1.28	2	1
38	58.4	27.6	8.8	8.9	3.22	2	1
24	58.4	27.9	13.1	9.2	2.80	2	2
45	53.3	31.8	10.6	9.2	0.80	2	2
35	52.3	33.7	12.7	9.3	0.26	2	2
44	46.8	34.9	12.7	9.4	1.26	2	2
7	40.8	46.3	10.9	9.5	4.63	2	2
37	44.7	39.7	13.9	9.6	0.92	3	2
10	42.5	38.0	16.1	9.7	2.27	3	2
19	41.7	40.3	14.7	9.7	4.82	3	2
9	40.4	39.6	16.9	9.8	1.67	3	2
33	41.8	35.4	19.9	9.9	0.32	3	2
8	33.9	43.9	19.4	10.1	3.56	3	2
31	33.4	45.0	20.9	10.2	5.12	3	2
6	34.4	39.6	23.1	10.3	1.05	3	2
11	35.5	37.6	23.9	10.3	0.38	4	2
14	35.8	35.3	28.1	10.5	0.88	4	2

12	23.3	52.4	23.0	10.6	4.27	4	2
27	24.3	49.9	24.2	10.6	2.33	4	3
17	19.3	54.2	26.0	10.8	3.45	4	3
5	19.3	51.6	26.9	10.9	4.52	4	3
13	32.1	33.5	33.0	10.9	2.62	4	3
32	23.5	45.7	29.0	10.9	1.65	5	3
2	24.3	41.6	31.2	11.0	1.86	5	3
28	19.4	49.7	29.6	11.0	2.60	5	3
30	25.3	41.0	33.1	11.0	0.85	5	3
4	13.4	54.4	31.6	11.2	4.06	5	3
34	15.1	36.8	47.6	12.0	0.89	5	3
1	8.8	31.0	59.4	12.8	2.08	5	3
16	7.5	24.5	67.8	13.2	0.52	5	3
29	8.0	19.0	72.7	13.4	1.64	5	3
15	5.4	20.9	73.6	13.5	1.25	5	3
3	4.0	18.3	77.3	13.8	2.15	5	3

* The percentages for low, medium and high education refer to the entire Dutch population and do not add up to 100, as a category of students and a category 'unknown' is excluded from the calculations

Chapter 4. Socioeconomic status and cancer survival in the Southeastern Netherlands

4.1 Socioeconomic status and breast cancer survival in the Southeastern Netherlands, 1980-1989*

4.1.1 Introduction

Breast cancer is the most common cancer among females in the Netherlands[1] as in many developed countries. Dutch women experience one of the highest incidence rates in the world.[2] The 5-year relative survival rate of breast cancer patients in the period 1975-1985 in the Southeastern Netherlands was 69%.[3]

Socioeconomic differences in breast cancer survival have been reported in studies from the United States,[4] Finland,[5] Sweden,[6] Australia,[7] Scotland[8] and England & Wales.[9] Except for one,[9] these studies on patients diagnosed in the 1960s or later, showed that breast cancer patients of low socioeconomic status (SES) have a higher chance of dying from their disease than breast cancer patients of high SES.

This paper is the first report on the impact of SES on breast cancer survival in the Netherlands, a country that is characterized by a relative lack of geographical and financial barriers to primary and specialized care. A description of the association between an area-based measure of SES and breast cancer survival in the 1980s is given and possible explanations of this association were studied. Regarding the latter, it was tested whether the difference in survival from breast cancer by SES can be explained by the distribution of a number of prognostic factors: stage at diagnosis, morphology, and treatment.

4.1.2 Patients and methods

Patients

Data for this study were derived from the population based Eindhoven Cancer Registry, which serves an area of about one million inhabitants (about 7% of the Dutch population) in the Southeastern part of the Netherlands.[2] The registry identifies newly diagnosed cases of cancer through routine reports from departments of pathology and radiotherapy, through inpatient records from all eight community hospitals in the region, as well as through data from specialized departments and hospitals outside of the region.[2,10] In this region the distance to a hospital is always less than 30 kilometres and that to a radiotherapy department is always less than 50 kilometres. All hospitals use the same criteria for the clinical assessment and treatment of breast cancer patients as they adhere to the guidelines developed by the regional Breast Cancer Study Group.[11]

The records of all women diagnosed with an invasive tumour of the breast

---

* Schrijvers CTM, Coebergh JWW, Heijden LH van der, Mackenbach JP. Eur J Cancer 1995;31A:1660-1664

between 1980 and 1989 (n=3959) were checked. Patients with an unknown basis of diagnosis (n=3), diagnosis based on autopsy (n=2), or unknown address at diagnosis (n=21) were excluded from the basic material. The remaining 3933 patients were followed-up until July 1, 1991, through the virtually complete municipal registries in the area, to determine their vital status. This was unknown for 5 patients, thus finally 3928 patients were included in the study.

Both patients with (96%) and those without (4%) a histologically confirmed breast tumour were included in the study, as there was no systematic difference in the proportion of patients with a histologically confirmed breast tumour according to SES group.

SES

Because no data on the SES of individual patients was directly available from the cancer registry, a proxy measure of SES was used, based on the place of residence at time of diagnosis of each patient. Data to develop the proxy measure were obtained from a commercial marketing agency, which has assigned each postcode (average of 16 households) in our study area to one of 45 sociodemographic categories, using a wide range of socioeconomic and sociodemographic survey data at the postcode level. The central variable in our analysis is education; the agency provided us with information on the percentage of main breadwinners in 3 educational groups (low, medium, high) for each of the 45 sociodemographic categories. These 3 educational groups encompassed several types of schooling, and we assigned an average number of years of education to each of them: 7.5 years to the lowest educational category (years of education between 6 and 9 years), 10 years to the medium educational category (years of education either 10 or 11) and 15 years to the highest educational category (years of education between 12 and 18).

The information on the percentage of main breadwinners in each of these 3 educational groups was then used to calculate a summary measure of the average number of years of education for each of the 45 sociodemographic categories. The 45 categories were then ranked from low (7.8 years) to high (13.8 years) according to their summary score on education and 5 socioeconomic categories were constructed, based on quintiles of the underlying population. So the highest SES category (1) contains about 20% of the population living in areas with the highest educational level, while the lowest SES category (5) contains about 20% of the population living in areas with the lowest educational level. Finally, each woman was assigned to one of the 5 categories of SES, through her postcode of residence at time of diagnosis.

We validated the proxy measure of SES in a subsample of respondents to a postal survey, which had been carried out in a part of the registration area of the Eindhoven cancer registry.[12] The subsample consisted of respondents living in one of 381 postcode areas for which at least 6 respondents were found in the survey, as the postcode area was the unit of measurement in this analysis. Each postcode could be assigned to one of the 5 socioeconomic categories of the proxy measure.

For respondents to the survey, data on the educational level was known and for each of the 381 postcodes, we calculated the average number of years of education with the survey data and then assigned each postcode to one of the 5 socioeconomic groups of the proxy measure (using the same procedure as with the marketing agency data). For each postcode we thus had two scores: (1) a score from 1 to 5 based on data from the original classification of the marketing agency and (2) a score from 1 to 5 based on data from respondents to the survey. The Pearson correlation coefficient between the two variables was 0.51, which is rather high for this type of comparison. We may conclude from this exercise that validity at the postcode level is satisfactory, given that the assignment of postcodes to one of 45 categories by the marketing agency was based on a large number of socioeconomic and sociodemographic variables, of which education was only one variable.

Prognostic factors

We studied the impact of a number of potential confounders and intermediate variables, which were treated as categorical in the analysis. As potential confounders of the SES-survival association we studied: age at diagnosis (3 categories: younger than 50, 50 to 64, and 65 years or older), period of diagnosis (2 categories: 1980 to 1984 and 1985 to 1989) and degree of urbanization of the place of residence at diagnosis (3 categories: smallest, intermediate, and largest municipalities). The following potential intermediate variables in the association between SES and survival were studied: stage at diagnosis (4 categories: localized (only local involvement of a tumour), regional (tumour growth confined to the breast and regional lymph nodes), distant (spread to other organs), and unknown), morphology (3 categories: ductal carcinoma, lobular carcinoma or other)[13] and treatment (5 categories: surgery only, surgery plus radiotherapy, surgery plus endocrine therapy, surgery plus chemotherapy, and no surgery).

Univariate analyses

The survival time of patients was calculated as the number of days between the date of diagnosis and either the date of death or the end of follow-up (July 1, 1991), whatever occurred first. As no information on the exact cause of death was available, the Relative Survival Rate (RSR) was used to correct for deaths due to causes other than breast cancer. The RSR[14] is the ratio of the observed survival rate of a group of cancer patients to the expected survival rate in a group similar to the patient group with respect to age, sex, and calendar period of observation. In this study the expected survival rate is based on life tables of the population of the registration area of the Eindhoven Cancer Registry, which were obtained from the Netherlands Central Bureau of Statistics. These life-tables each applied to a 2-year calendar period and were age- and sex- specific. RSR's and 95% Confidence Intervals (CI) were calculated with the computer program for cancer survival studies from the Finnish Cancer Registry.[15]

Multivariate analyses

The multivariate analyses were conducted with a regression model adapted to the RSR[16] using GLIM.[17] The measure of effect in the multivariate analyses was the hazard ratio (HR), which expresses the probability of death from breast cancer for a specific category of patients relative to a reference category (with a HR of 1.00).

The entire period of follow-up was divided into two periods: up to 5 and 6 to 12. Because the probability to die from breast cancer was not equal for these two periods it was necessary to correct for this difference in HRs by including this variable in the model. At each step in the multivariate analysis an extra variable was added to a model which contained follow-up period and SES. First, possible confounders were added to the model and then possible intermediate variables. For a variable to be included in the final model, it had to cause a change in HRs of the SES variable after addition to the model. Furthermore, the reduction in deviance due to a variable, with a corresponding difference in degrees of freedom, using the chi-square distribution, had to be statistically significant ($p < 0.05$).

At each step in the analysis, a test for trend with the SES variable was also conducted by including it as a continuous variable in the model. The reduction in deviance due to the continuous SES variable was then evaluated, using the chi-square distribution with 1 degree of freedom.

4.1.3 Results

Table 1 contains the 5- and 10-year RSR for the five SES categories, uncorrected for other factors. Both the 5- and 10-year RSR appeared to be higher for the higher SES categories, although a clear gradient was not apparent and 95% CI's overlapped.

Table 1. Five and ten year relative survival rate (%) according to socioeconomic status, breast cancer, Southeastern Netherlands, 1980-1989

SES	N (%)	5 year RSR%	10 year RSR%
1 (high)	795 (20.2)	77 (73 - 81)*	64 (58 - 70)
2	430 (10.9)	74 (69 - 79)	64 (55 - 73)
3	814 (20.7)	75 (71 - 79)	65 (58 - 72)
4	987 (25.1)	72 (68 - 76)	61 (55 - 67)
5 (low)	902 (23.0)	73 (70 - 76)	57 (50 - 64)
Total	3928 (100)	74 (72 - 76)	62 (59 - 65)

SES: socioeconomic status * 95% confidence interval between brackets

The distribution of age ($p < 0.001$) and degree of urbanization ($p < 0.001$) differed statistically significantly per SES category (table 2), while for the other variables this was not the case: period of diagnosis ($p=0.61$), stage ($p=0.08$), morphology ($p=0.11$), and treatment ($p=0.93$). For stage however, we found a higher percentage of patients diagnosed with a distant stage in the lower SES categories.

Table 2. Distribution of possible confounders and intermediate variables according to socioeconomic status, breast cancer, Southeastern Netherlands, 1980-1989

	SES						
	high 100 %	2 100 %	3 100 %	4 100 %	low 100 %	Total 100 %	X^2 test
Age							
- 49	34.1	29.8	25.1	30.2	24.7	28.6	$p < .001$
50 - 64	34.2	33.0	34.6	33.8	38.5	35.1	
65 -	31.7	37.2	40.3	36.0	36.8	36.3	
Period of diagnosis							
80-84	44.2	42.8	43.2	43.0	46.2	44.0	
85-89	55.8	57.2	56.8	57.0	53.8	56.0	$p=0.61$
Degree of urbanization							
Smallest	2.9	3.5	10.3	15.9	8.3	9.0	
Intermediate	52.6	33.7	40.4	57.3	38.5	46.0	$p < .001$
Largest	44.5	62.8	49.3	26.8	53.2	45.0	
Stage							
Local	49.6	49.3	46.7	46.9	48.4	48.0	
Regional	35.8	31.9	33.2	33.6	31.8	33.4	
Distant	5.4	6.5	6.3	6.8	8.6	6.8	$p=0.08$
Unknown	9.2	12.3	13.8	12.7	11.2	11.8	
Morphology							
Ductal	82.4	77.4	82.2	78.1	79.0	80.0	
Lobular	10.7	13.1	9.3	13.5	12.6	11.8	$p=0.11$
Other	6.9	9.5	8.5	8.4	8.4	8.2	
Treatment							
Su	20.1	22.1	19.9	22.0	22.3	21.3	
Su + Ra	56.2	52.1	55.4	55.1	56.4	55.4	
Su + En	7.4	8.4	8.6	7.9	6.5	7.7	$p=0.93$
Su + Ch	7.9	7.9	7.9	7.1	6.8	7.4	
No Su	8.4	9.5	8.2	7.9	8.0	8.2	

SES: socioeconomic status; Su: Surgery; Ra: Radiotherapy; En: Endocrine therapy; Ch: Chemotherapy

Table 3 contains the results of the multivariate analyses, showing the HRs for the five SES categories for the different models, with the highest SES category as a reference category. Period of diagnosis and degree of urbanization were added to a model with follow-up period and SES, and appeared to be no confounders of the SES-survival association and are therefore not presented in table 3. In model 1 which included follow-up period and SES, the gradient in HRs was clear and the lower SES categories showed higher HRs. The p-value for the test for trend was 0.037.

When age was included in the model (model 2) the HRs for SES were reduced substantially, while the reduction in deviance was also statistically significant. The CIs around HRs for the five SES categories overlapped, but a gradient was apparent with higher HRs for the lower SES categories (test for trend, p=0.073).

After a correction for stage (model 3), differences in HRs became much smaller and the gradient disappeared (p=0.841). The reduction in deviance due to stage was also statistically significant. Morphology (model 4) and treatment (model 5) changed HRs only moderately but because the reduction in deviance due to these variables was statistically significant, they were kept in the final model.

Table 3. Hazard ratio and 95% confidence interval by socioeconomic status, breast cancer, Southeastern Netherlands, 1980-1989[28]

	SES					
Model	high*	2	3	4	low	Test for trend
Model 1: Follow-up period, and SES						
Hazard ratio	1.00	1.09	1.09	1.17	1.24	p=.037
95% CI		(0.86-1.38)	(0.90-1.33)	(0.97-1.41)	(1.03-1.49)	
Model 2: Follow-up period, SES, and age						
Hazard ratio	1.00	1.06	1.04	1.15	1.18	p=.073
95% CI		(0.84-1.33)	(0.86-1.26)	(0.96-1.38)	(0.99-1.42)	
Model 3: Follow-up period, SES, age, and stage						
Hazard ratio	1.00	1.09	1.04	1.06	1.03	p=.841
95% CI		(0.88-1.34)	(0.87-1.25)	(0.90-1.26)	(0.87-1.22)	
Model 4: Follow-up period, SES, age, stage, and morphology						
Hazard ratio	1.00	1.07	1.04	1.06	1.03	p=.802
95% CI		(0.87-1.33)	(0.87-1.24)	(0.90-1.26)	(0.87-1.22)	
Model 5: Follow-up period, SES, age, stage, morphology, and treatment						
Hazard ratio	1.00	1.04	1.03	1.04	1.03	p=0.792
95% CI		(0.84-1.29)	(0.87-1.23)	(0.88-1.23)	(0.87-1.22)	

SES: socioeconomic status; CI: confidence interval * reference category

4.1.4 Discussion

Our results suggest that socioeconomic differences in breast cancer survival exist in the Netherlands: after a correction for age, mortality due to breast cancer was 18 percent higher in the lowest SES category than in the highest SES category. Although CIs for the different SES categories overlapped, a gradient in HRs for different SES categories was apparent ($p=0.073$). Socioeconomic differences in breast cancer survival could mainly be ascribed to differences in the stage-distribution between the SES categories, particularly to differences in the percentage of patients diagnosed with a metastasis, which was 8.6 for the lowest and 5.4 for the highest SES category.

Before we continue with the interpretation of our findings, some methodological issues concerning the proxy measure of SES have to be considered. The measure of SES is ecological and based on the average number of years of education per postcode area of residence, and therefore misclassification, resulting in an underestimation of the SES-survival gradient, cannot be ruled out. The results from our validation study showed however, that our measure of SES is a very reasonable indicator of SES at the postcode level.

The postcode of residence at the time of diagnosis was used to assign each patient to a socioeconomic category. The area of residence of a patient and therefore her SES score could have changed during the follow-up period. It seems very unlikely however that migration after the diagnosis of cancer was differential according to SES.

Due to the use of one single life-table to correct for other causes of death than breast cancer, we may have overestimated the gradient in survival by SES. Expected survival might be overestimated for lower SES groups and therefore relative survival is underestimated and the HR is overestimated. For higher SES groups, expected survival might be underestimated, and therefore the relative survival overestimated and the HR underestimated. In a Finnish study, it was shown that this overestimation of the SES-survival gradient is probably not very large.[5] In this study, the socioeconomic gradient in both corrected survival (censoring of cases dying from other causes than breast cancer) and relative survival were calculated. The ratio of survival rates of the highest and lowest social class was somewhat higher when the RSR was used (1.12) as compared to the corrected survival rate (1.10). This overestimation of the SES-survival gradient is probably smaller in the Netherlands than in Finland, as socioeconomic variation in general mortality is smaller in the Netherlands than in Finland.[18]

A direct comparison of our findings with those from other studies[4-9] is rather difficult, as studies differ in design and data analysis. In most studies a better survival for higher SES groups was found. However, in a study on English breast cancer patients diagnosed between 1971 and 1981, a non-significant better survival was found for council tenants (low SES) than for owner-occupiers (high SES).[9] In a study on Swedish breast cancer patients (period of diagnosis 1961-1979) the RSR

of white collar workers (high SES) was about seven percent higher than that of blue collar workers (low SES), without a correction for other prognostic factors.[6] The relative risk of case fatality in low SES women from South-Australia (1977-1982) was 1.35 (95% CI: 1.04-1.74) after a correction for age and histology.[7] Five year survival was 66% in the highest SES group and 55% in the lowest SES group in patients diagnosed in west of Scotland in the period 1980-1987, using an area-based measure of SES.[8] Even in studies which adjusted for differences in stage-distribution across SES groups, a statistically significant higher risk of dying for the lowest SES group was found,[4,5] which is not the case in our study. For Finnish breast cancer patients (1971-1980) from the highest social class, the relative risk of dying after a correction for age, period of diagnosis, and stage was 0.78 (95% CI: 0.68-0.90).[5] Women from the United States (1979-1983) living in areas with at least 35% working class, experienced a relative risk of mortality of 1.52 (95% CI: 1.28-1.88), compared to women living in areas with less than 35% working class, adjusted for race, age, stage, and histology.[4] Our results are thus in the same direction as those from most studies conducted in other countries. The strength of the association seems to be relatively weak however in the Netherlands, with an age corrected HR for the lowest SES category of 1.18.

The most important explanatory factor of socioeconomic differences in breast cancer survival in our study was stage of disease at diagnosis. In several studies, it was found that women from lower SES groups are diagnosed with more advanced stages of breast cancer than women from higher SES groups.[5,19-22] Such differences in stage distribution may be related to the length of delay between the occurrence of the first symptoms and the time of diagnosis, which might be shorter in more educated and better informed women. In some studies, delay was found to be longer for women of lower SES,[22-24] and a longer delay was found to be related to more advanced stages of breast cancer,[24-26] while it is related to lower survival.[26]

In our study, stage was only moderately associated with SES, as only a distant stage was more common among lower SES women. Credit to this moderate association may be good access to primary and specialized care in the Southeastern Netherlands, as a result of relatively short distances to a hospital, good supply of health services, and a health insurance system without major financial obstacles: in the study-period only 0.4% of the Dutch population was uncovered by health insurance.[27]

Less attention has been given to socioeconomic differences in treatment as an explanation for survival differences. It could be argued that the choice of treatment, given the extent of disease at diagnosis, might be related to the SES of breast cancer patients. Although we found no differences in treatment according to SES after adjustment for stage (results not shown), socioeconomic differences in the quality of treatment of breast cancer patients may exist. Such differences could not be evaluated however through the rather crude indicator of treatment used in this study. In any case, such differences cannot be responsible for large differences in survival in the Netherlands.

Our findings on the influence of stage on socioeconomic differences in breast cancer survival indicate that, with regard to secondary prevention of breast cancer, special attention should be given to women of lower SES. During the study period, a breast cancer screening program at the population level was absent, and it is now being implemented in the Netherlands. Through health education programs, women from lower SES groups should be extra stimulated to participate in such a screening program as well as to practice breast self examination. Such programs, together with keeping up good general access to health care services for the entire population, may lead to a further reduction of socioeconomic differences in breast cancer survival in the Netherlands.

References

1. Winter GA de, Coebergh JWW, Leeuwen FE van, Schouten LJ (eds). Incidence of cancer in the Netherlands, 1989. Utrecht: Netherlands Cancer Registry, 1992.
2. Bakker D, Coebergh JWW, Crommelin MA, Verhagen-Teulings MTh. Netherlands. Eindhoven Cancer Registry. In: Muir CS, Waterhouse J, Mack T, Powell J, Whelan S (eds). Cancer incidence in five continents, volume V. IARC scientific publications no 88. Lyon: IARC, 1987: 574-579.
3. Coebergh JWW, Heijden LH van der (eds). Cancer incidence and survival 1975-1987 in Southeastern Netherlands. Eindhoven: Eindhoven Cancer Registry, 1991.
4. Bassett MT, Krieger N. Social class and black-white differences in breast cancer survival. Am J Public Health 1986;76:1400-1403.
5. Karjalainen S, Pukkala E. Social class as a prognostic factor in breast cancer survival. Cancer 1990;66:819-826.
6. Vågerö D, Persson G. Cancer survival and social class in Sweden. J Epidemiol Community Health 1987;41:204-209.
7. Bonett A, Roder D, Esterman A. Determinants of case survival for cancers of the lung, colon, breast and cervix in South Australia. Med J Aust 1984;141:705-709.
8. Carnon AG, Ssemwogerere A, Lamont DW, Hole DJ, Mallon EA, George WD, Gillis CR. Relation between socioeconomic deprivation and pathological prognostic factors in women with breast cancer. Br Med J 1994;309:1054-1057.
9. Kogevinas M, Marmot MG, Fox AJ, Goldblatt PO. Socioeconomic differences in cancer survival. J Epidemiol Community Health 1991;45:216-219.
10. MacLennan R, Muir C, Steinitz R, Winkler A (eds). Cancer registration and its techniques. IARC scientific publications no 21. Lyon: IARC, 1978.
11. Voogd AC, Beek MWPM van, Crommelin MA, Kluck HM, Repelaer van Driel OJ, Coebergh JWW. Management of early breast cancer in southeast Netherlands since 1984. Acta Oncologica 1994;33:753-757.
12. Mackenbach JP, Mheen H van de, Stronks K. A prospective cohort study investigating the explanation of socio-economic inequalities in health in the Netherlands. Soc Sci Med 1994;38:299-308.
13. International Classification of Diseases for Oncology. First edition 1976. Geneva: World Health Organisation, 1976.
14. Ederer F, Axtell LM, Cutler SJ. The relative survival rate: a statistical methodology. Bethesda, U.S.A.,: National Cancer Institute, monograph no. 6, 1961: 101-121.
15. Hakulinen T, Abeywickrama KH. A computer program package for relative survival analysis. Comp Prog in Biomed 1985;19:197-207.
16. Hakulinen T, Tenkanen L. Regression analysis of the relative survival rates. Appl Stat 1987;36:309-317.
17. Payne CD. The GLIM system Release 3.77. Generalized linear interactive modelling manual. Oxford: Numerical Algorithms Group, 1985.
18. Kunst AE, Mackenbach JP. An international comparison of socio-economic inequalities in mortality. Rotterdam: Instituut Maatschappelijke Gezondheidszorg, 1992.
19. Farley TA, Flannery JT. Late-stage diagnosis of breast cancer in women of lower socioeconomic status: public health implications. Am J Public Health 1989;79:1508-1512.
20. Wells B, Horm JW. Stage at diagnosis in breast cancer: race and socioeconomic factors. Am J Public Health 1992;82:1383-1385.

21. Mandelblatt J, Andrews H, Kerner J, Zauber A, Burnett W. Determinants of late stage diagnosis of breast and cervical cancer: the impact of age, race, social class and hospital type. Am J Public Health 1991;81;646-649.
22. Richardson JL, Langholz B, Bernstein L, Burciaga C, Danley K, Ross RK. Stage and delay in breast cancer diagnosis by race, socioeconomic status, age and year. Br J Cancer 1992;65-:922-926.
23. Elwood JM, Moorehead WP. Delay in diagnosis and long-term survival in breast cancer. BMJ 1980;31 May:1291-1294.
24. Gould-Martin K, Paganini-Hill A, Casagrande C, Mack T, Ross RK. Behavioral and biological determinants of surgical stage of breast cancer. Preventive Medicine 1982;11:429-440.
25. Hainsworth PJ, Henderson MA, Bennett RC. Delayed presentation in breast cancer: relationship to tumour stage and survival. The Breast 1993;2:37-41.
26. Wilkinson GS, Edgerton F, Wallace HJ, Reese P, Patterson J, Priore R. Delay, stage of disease and survival from breast cancer. J Chron Dis 1979;32:365-373.
27. Maandbericht gezondheidsstatistiek, oktober 1991, Den Haag: CBS, 1991: 4-6.
28. Schrijvers CTM, Coebergh JWW, Heijden LH van der, Mackenbach JP. Socioeconomic variation in cancer survival in the Southeastern Netherlands, 1980-1989. Cancer 1995;75:-2946-2953.

4.2 Socioeconomic variation in cancer survival in the Southeastern Netherlands, 1980-1989*

4.2.1 Introduction

Socioeconomic variation in the survival of patients with cancer has been reported for a variety of cancer sites in studies from several countries. In general, patients with a low socioeconomic status (SES) live for a shorter period after a cancer diagnosis than patients with a high SES, although this is not true to the same extent for each cancer site.[1]

Several prognostic factors for cancer have been mentioned as possible explanatory factors for socioeconomic variation in cancer survival. The most important is stage of disease at diagnosis, with more advanced disease among patients in lower SES groups. Other possible explanations are differences in treatment, host resistance, and tumour characteristics.[2]

This is the first combined report on socioeconomic variation in survival from the most common cancer sites in the Netherlands, a country characterized by good general access to the health care system at the level of general practitioner and specialized care. This equal access is reflected in a good supply of health services and a health insurance system without major financial obstacles. Less than 1% of the Dutch population was not receiving health insurance coverage in the study period.[3] The basic coverage of health services by health insurance was sufficient generally and did not vary clearly by type of health insurance. Furthermore, in the study region, the distance to a hospital is always less than 30 kilometres, and the distance to the only radiotherapy department is always less than 50 kilometres.

In this paper, we describe socioeconomic variation in survival from cancer of the lung, breast, colorectum, prostate, and stomach, and we tried to explain such variation by studying the distribution of a number of prognostic factors (stage at diagnosis, histological type, and treatment) across socioeconomic groups.

4.2.2 Patients and methods

Data source

Data for this study came from the Eindhoven Cancer Registry, a population based cancer registry, which included data for an area of about one million inhabitants (approximately 7% of the Dutch population) in the Southeastern Netherlands.

The registry identified newly diagnosed cases of cancer through routine reports from departments of pathology and radiotherapy, through inpatient records from all eight community hospitals in the region, and through data from specialized departments and hospitals outside of the region.[4-6]

---

* Schrijvers CTM, Coebergh JWW, Heijden LH van der, Mackenbach JP. Cancer 1995;75:2946-2953

Patients
All patients diagnosed from 1980 to 1989 with an invasive tumour of the lung (n=4591), breast (n=3928), colorectum (n=3558), prostate (n=1484), or stomach (n=1455) were selected for study.

Socioeconomic status
Data on the socioeconomic status of individual patients were not available in the cancer registry. Therefore, we used a measure of SES based on the postcode of residence at the time of diagnosis for each patient. In the Netherlands, a marketing agency assigned each postcode sector, with an average of 16 households, to 1 of 45 categories of a sociodemographic classification based on various data sources. The original classification of 45 categories was divided into 5 broader categories, constructed by ranking the 45 categories by the average number of years of education of the main breadwinners in the corresponding areas from low (7.8 years) to high (13.8 years), and by dividing this distribution into quintiles. Because the total number of patients with cancer of the prostate and stomach was relatively small, the 45 categories also were grouped into 3 categories, based on tertiles of the underlying population.

We validated the proxy measure of socioeconomic status with data from the baseline measurement of a prospective cohort study that began in 1991. For this study, an aselect sample (stratified by age, degree of urbanization, and socioeconomic status) of approximately 27000 persons was drawn from the population registers in an area in the Southeastern part of the Netherlands, an area which is completely included in the Eindhoven Cancer Registry. The persons in this sample received a postal questionnaire; the response rate was 70.1% (n=18973)[7]. The postal survey provided information on the education of the respondents who also could be assigned through their postcode of residence to 1 of the 45 categories of our sociodemographic classification. The educational status and proxy measure were known for 18227 respondents, and this was the number of respondents that we used for two types of validation studies. Firstly, validity at the level of the 45 sociodemographic categories was assessed. For each of the 45 categories of the classification, we calculated the average number of years of education a second time, using the postal survey data. The 45 categories then were ranked a second time according to the average number of years of education based on these postal survey data. The rank-correlation coefficient between the two series ([1] average number of years of education based on the original classification given by the marketing agency [2] average number of years of education based on the postal survey data) was 0.87. Secondly, validity at the individual level was assessed by calculating the Pearson correlation coefficient between individual education in 7 categories (survey data) and the area-based measure in 5 categories, which was 0.30. The proxy measure was, thus, a reasonable indicator of the socioeconomic status at the aggregate level, whereas it was a less valid measure of the socioeconomic status at the individual level. This suggests that we may underesti-

mate the association between SES and survival at the individual level.[8]

Prognostic factors

We studied the impact of available prognostic factors on the association between socioeconomic status and survival, as possible confounders (age, sex, period of diagnosis, and urbanization) and intermediate factors (stage at diagnosis, histological type, and treatment). Furthermore, subsites were distinguished for cancer of the colorectum ([1] rectum, [2] sigmoid, [3] ascending colon, [4] transverse and descending colon, and [5] other subsites)) and stomach ([1] cardia, [2] pylorus and [3] stomach excluding cardia and pylorus). Age was studied in several categories, taking the varying age distribution of cases per cancer site into account: lung (-59, 60-69, and 70+ years), breast (-49, 50-64, and 65+ years), colorectum (-59, 60-69, 70-79, and 80+ years), prostate (-69, 70-79, and 80+ years) and stomach (-59, 60-69, and 70+ years). The period of diagnosis (1980-1989) was divided in two 5-year periods because the survival was better in the second 5-year period for most cancer sites. Degree of urbanization of the place of residence at diagnosis was studied in 3 categories: smallest, intermediate, and largest municipalities. Tumour staging initially was based on the International Union Against Cancer classification of malignant tumours[9]. Each patient was assigned to one of four stage categories that we distinguished as localized (tumour confined to the organ of origin), regional (spread to regional lymph nodes), distant (spread to other organs), and unknown. For cancer of the prostate, localized and rarely classified regional tumours were combined. The histological classification for cancer of the lung was small cell and non small cell; for the breast it was ductal, lobular, and other; and for the colorectum, prostate, and stomach it was adenocarcinoma and other. The classification of treatment for cancer of the lung was chemotherapy, surgery, no treatment, other; for the breast it was surgery, surgery plus radiotherapy, surgery plus endocrine therapy, surgery plus chemotherapy, and no surgery; for the colorectum it was surgery, surgery plus radiotherapy, other, and no treatment; for the prostate it was surgery, surgery plus other treatment, endocrine therapy plus other treatment, other, no treatment; and for the stomach it was surgery, other, and no treatment.

Survival analyses

The patients were followed actively through municipal population registries to determine their vital status as of July 1, 1991. These registries have a virtually complete coverage of the population and are maintained continuously with respect to deaths and changes of address. Patients who moved from the study area were traced through the municipal registry of their new place of residence. Less than 1% of all patients proved to be lost to follow-up[10].

The survival time of each patient was calculated as the number of days between the date of diagnosis and either the date of death or the end of follow-up (July 1, 1991), whichever occurred first. The survival time then was divided by 365.25 to calculate the survival time in years.

Because no information on the exact cause of death was available, the Relative Survival Rate (RSR) expressed as a percentage was used to correct for deaths due to other causes than the cancer under study. The RSR is the ratio of the observed survival of a group of patients with cancer to the expected survival, which is the survival they would experience if they were subject to the same overall mortality as a group from the general population similar to the patient group with respect to age, sex, and calendar period of observation.[11] The expected survival rate is calculated with life tables for the regional population, which were obtained from the Netherlands Central Bureau of Statistics. Survival analyses were conducted with the computer program for cancer survival studies from the Finnish Cancer Registry.[12]

The multivariate analyses were conducted with a regression model adapted to the RSR[13] using Generalized Linear Interactive Modelling (GLIM).[14] The measure of effect in these analyses was the hazard ratio (HR), which gives the probability of death from the type of cancer under study for a specific group of patients relative to a reference category, which has a hazard ratio of 1.00.

In this study, we present HRs that apply to the entire period of follow-up, which was divided into two periods (up to 5 and 6-12 years). Because the probability of death from cancer was not equal for these two periods, it was necessary to correct for this difference in hazards by including this variable in the model. We started with a model that contained duration of follow-up in two periods (up to 5 and 6-12 years) and SES. We first added the possible confounders (age, sex, period of diagnosis, urbanization), followed by subsite for colorectum and stomach cancer, and then the intermediate factors for all sites (stage, histological type, and treatment). The addition of each factor was evaluated by testing for statistical significance ($P < 0.05$) the reduction in deviance caused by that factor, with a corresponding difference in degrees of freedom, as compared with the model without the factor. At each step in the analysis, a test for trend with the SES variable also was conducted by including it as a continuous variable in the model. The reduction in deviance due to the continuous SES variable then was evaluated using the chi-square distribution with 1 degree of freedom.

4.2.3 Results

Univariate Analyses

Table 1 shows the 5-year relative survival rate by site and SES. For lung cancer the 5-year RSR was higher in the higher SES groups, although the highest SES group had a lower 5-year RSR than the second highest SES group. For breast cancer, we observed the same pattern, with the highest 5-year RSR in the higher SES groups, but the gradient was not consistent. For colorectal cancer, the 5-year RSR was also higher in the higher SES groups, with a clear gradient in survival according to SES. The 5-year RSR for cancer of the prostate was slightly lower in the lower SES groups, whereas for stomach cancer, the highest 5-year RSR was

observed in the lowest SES group, with a clear gradient.

Table 1. Five year relative survival rate, 95% confidence interval, and number of cases, by cancer site and socioeconomic status, Southeastern Netherlands, 1980-1989

		SES					
Cancer site		High	(2)	(3)	(4)	Low	Total
Lung	RSR%	15	17	14	12	11	13
	95% CI	12-18	13-21	11-17	9-15	9-13	12-14
	No. of cases	717	420	944	1172	1338	4591
Breast	RSR%	77	74	75	72	73	74
	95% CI	73-81	69-79	71-79	68-76	70-76	72-76
	No. of cases	795	430	814	987	902	3928
Colo-rectum	RSR%	55	54	50	48	49	51
	95% CI	50-60	47-61	45-55	44-52	45-53	48-54
	No. of cases	688	378	706	865	921	3558
		High	Intermediate		Low		Total
Prostate	RSR%	61	60		59		60
	95% CI	53-69	52-68		52-66		55-65
	No. of cases	427	482		575		1484
Stomach	RSR%	18	20		25		22
	95% CI	13-23	15-25		21-29		19-25
	No. of cases	310	493		652		1455

SES: socioeconomic status; RSR: relative survival rate: CI: confidence interval

Prognostic Factors

Regarding the distribution of prognostic factors across SES groups, we only found a few clear patterns. For breast cancer, the percentage of women diagnosed with a metastasis was higher in the lower SES groups: low (1): 8.6, (2): 6.8, (3): 6.3, (4): 6.5, high (5): 5.4. For stomach cancer, the percentage of patients diagnosed with a metastasis was 32 in the highest SES group, 29 in the intermediate category, and 27 in the lowest category.

Multivariate Analyses

The results of the multivariate analyses are shown in tables 2-6. Prognostic factors that did not significantly improve the fit of the preceding model are neither presented in these tables nor mentioned in the text.

For lung cancer (table 2), higher hazards were found in the lower SES groups in a model with follow-up period, SES and age (model 1). Adjustment for stage (model 2), histological type (model 3), and treatment (model 4), did not change the

hazard ratios for SES substantially. In the final model (follow-up period, SES, age, stage, histological type, and treatment), the gradient in survival by SES was still apparent, with an excess hazard of death of 16% in the lowest SES group. At each step in the analysis, the hazard ratio from the test for trend on the SES variable was approximately 1.05.

Table 2. Hazard ratio and 95% confidence interval by socioeconomic status, lung cancer, Southeastern Netherlands, 1980-1989

	SES					
	High*	2	3	4	Low	Test for trend
Model 1: Follow-up period, SES, and age						
Hazard Ratio	1.00	0.97	1.05	1.14	1.18	1.05
95% CI		0.80-1.18	0.90-1.24	0.98-1.33	1.02-1.36	1.01-1.08
Model 2: Follow-up period, SES, age, and stage						
Hazard Ratio	1.00	0.97	1.06	1.15	1.17	1.05
95 % CI		0.81-1.15	0.92-1.22	1.01-1.32	1.03-1.34	1.02-1.08
Model 3: Follow-up period, SES, age, stage, and histology						
Hazard Ratio	1.00	0.98	1.06	1.15	1.18	1.05
95% CI		0.82-1.16	0.93-1.22	1.01-1.32	1.04-1.34	1.02-1.08
Model 4: Follow-up period, SES, age, stage, histology, and treatment						
Hazard Ratio	1.00	1.02	1.07	1.15	1.16	1.04
95% CI		0.87-1.19	0.94-1.21	1.02-1.30	1.03-1.31	1.01-1.07

SES: socioeconomic status; CI: confidence interval
* Reference category

In a model with follow-up period, SES, and age for breast cancer (model 1, table 3), a clear gradient in hazards by SES was found (HR, test for trend: 1.04), with an 18% excess hazard of death in the lowest SES category. After a correction for stage (model 2), differences in HRs between SES groups became much smaller, and the gradient disappeared (HR, test for trend: 1.00). Histological type (model 3) and treatment (model 4) changed the HRs for SES only moderately, and from the final model (model 4), which contains follow-up period, SES, age, stage, histological type, and treatment, no gradient in HRs by SES emerged (HR, test for trend: 1.00).

Table 3. Hazard ratio and 95% confidence interval by socioeconomic status, breast cancer, Southeastern Netherlands, 1980-1989

	SES					
	High*	2	3	4	Low	Test for trend
Model 1: Follow-up period, SES, and age						
Hazard Ratio	1.00	1.06	1.04	1.15	1.18	1.04
95% CI		0.84-1.33	0.86-1.26	0.96-1.38	0.99-1.42	1.00-1.09
Model 2: Follow-up period, SES, age, and stage						
Hazard Ratio	1.00	1.09	1.04	1.06	1.03	1.00
95% CI		0.88-1.34	0.87-1.25	0.90-1.26	0.87-1.22	0.97-1.04
Model 3: Follow-up period, SES, age, stage, and histology						
Hazard Ratio	1.00	1.07	1.04	1.06	1.03	1.00
95% CI		0.87-1.33	0.87-1.24	0.90-1.26	0.87-1.22	0.97-1.04
Model 4: Follow-up period, SES, age, stage, histology, and treatment						
Hazard Ratio	1.00	1.04	1.03	1.04	1.03	1.00
95% CI		0.84-1.29	0.87-1.23	0.88-1.23	0.87-1.22	0.97-1.04

SES: socioeconomic status; CI: confidence interval
* Reference category

Table 4. Hazard ratio and 95% confidence interval by socioeconomic status, colorectal cancer, Southeastern Netherlands, 1980-1989

	SES					
	High*	2	3	4	Low	Test for trend
Model 1: Follow-up period, SES, age, and subsite						
Hazard Ratio	1.00	1.00	1.06	1.15	1.17	1.05
95% CI		0.79-1.28	0.87-1.30	0.95-1.40	0.97-1.41	1.00-1.09
Model 2: Follow-up period, SES, age, subsite, and stage						
Hazard Ratio	1.00	1.06	1.11	1.27	1.20	1.05
95% CI		0.86-1.32	0.93-1.33	1.07-1.50	1.02-1.41	1.01-1.09
Model 3: Follow-up period, SES, age, subsite, stage, and histology						
Hazard Ratio	1.00	1.08	1.14	1.27	1.18	1.05
95% CI		0.87-1.33	0.95-1.35	1.07-1.50	1.01-1.39	1.01-1.09
Model 4: Follow-up period, SES, age, subsite, stage, histology, and treatment						
Hazard Ratio	1.00	1.06	1.14	1.27	1.14	1.04
95% CI		0.86-1.31	0.95-1.35	1.08-1.50	0.97-1.34	1.00-1.08

SES: socioeconomic status; CI: confidence interval.
* Reference category

For colorectal cancer, a clear gradient (HR, test for trend: 1.05) in HRs by SES was found in a model that included follow-up period, SES, age, and subsite (table 4). Adjustment for stage caused an increase of the HRs for SES, especially in the second lowest category (model 2), whereas adjustment for histological type (model 3) and treatment (model 4) caused only minor changes in the HRs. In the final model (follow-up period, SES, age, subsite, stage, histological type, treatment), the SES-survival gradient still existed (HR, test for trend: 1.04) but was interrupted in the second lowest category.

For cancer of the prostate (table 5) the intermediate and low SES group showed a higher HR than the high SES group in model 1 (SES and age). After adjustment for stage (model 2) and treatment (model 3), the HRs in the final model (SES, age, stage, and treatment) remained unchanged as compared with the first model. The association between SES and survival was weak, however, as the confidence intervals around the HR from the tests for trend were broad.

The hazards in the lowest and intermediate SES groups for stomach cancer (table 6) were lower than the hazard for the highest SES group (model 1, follow-up period, SES, age, and subsite). The gradient in survival by SES from this model is not statistically significant, however (HR test for trend: 0.95). After adjustment for stage (model 2), socioeconomic differences in HRs disappeared (HR, test for trend: 0.99), whereas adjustment for treatment (model 3) caused a minor change in HRs. In the final model (follow-up period, SES, age, subsite, stage, and treatment), the gradient in hazard by SES had disappeared (HR, test for trend: 0.99).

Table 5. Hazard ratio and 95% confidence interval by socioeconomic status, prostate cancer, Southeastern Netherlands, 1980-1989

	SES			
	High*	Intermediate	Low	Test for trend
Model 1: SES and age				
Hazard Ratio	1.00	1.05	1.20	1.10
95% CI		0.77-1.43	0.95-1.59	0.95-1.26
Model 2: SES, age, and stage				
Hazard Ratio	1.00	1.13	1.20	1.09
95% CI		0.85-1.50	0.92-1.56	0.96-1.24
Model 3: SES, age, stage, and treatment				
Hazard Ratio	1.00	1.07	1.20	1.10
95% CI		0.81-1.42	0.93-1.56	0.96-1.25

SES: socioeconomic status; CI: confidence interval.
* Reference category

Table 6. Hazard ratio and 95% confidence interval by socioeconomic status, stomach cancer, Southeastern Netherlands, 1980-1989

	SES			
	High*	Intermediate	Low	Test for trend
Model 1: Follow-up period, SES, age, and subsite				
Hazard Ratio	1.00	0.92	0.89	0.95
95% CI		0.71-1.20	0.69-1.15	0.84-1.07
Model 2: Follow-up period, SES, age, subsite, and stage				
Hazard Ratio	1.00	1.04	0.99	0.99
95% CI		0.84-1.29	0.81-1.21	0.89-1.09
Model 3: Follow-up period, SES, age, subsite, stage, and treatment				
Hazard Ratio	1.00	1.06	1.00	0.99
95% CI		0.86-1.31	0.82-1.21	0.90-1.09

SES: socioeconomic status; CI: confidence interval
* Reference category

4.2.4 Discussion

Our results indicate that socioeconomic variation in survival of patients with cancer of common sites exists in the Netherlands. After adjustment for confounding variables, up to 20% higher hazards of death were found for patients from low SES areas diagnosed with cancer of the lung, breast, colorectum, and prostate as compared with patients from high SES areas. For stomach cancer, a higher hazard of death was found for patients from high SES areas. Overall, socioeconomic variation in survival was small, and for cancers of the lung, colorectum and prostate, it could not be explained by the distribution of the prognostic factors stage, histological type, and treatment. For breast and stomach cancer, socioeconomic differences in survival could be ascribed mainly to differences between SES categories in the percentage of patients diagnosed with a metastasis.

Some methodological issues could influence the interpretation of our results. First, like many other previous investigators in this area, we used an ecological measure of SES, which was assigned to each individual patient according to the postcode of residence at the time of diagnosis. The results of our validation study on the socioeconomic score indicate that this measure was an appropriate indicator of socioeconomic status at the aggregate level, whereas its use may lead to an underestimation of socioeconomic variation at the individual level.

Secondly, to correct for other causes of death than the specific cancer under study, we could not use SES-specific life tables, because these are not available in the Netherlands. The expected survival of patients from lower SES groups may be

overestimated using a life table of the total regional population, which results in an underestimation of the RSR and overestimation of the HR for the lower SES groups. For patients in the higher SES groups, the expected survival rate probably was underestimated, resulting in an overestimation of the RSR and underestimation of the HR. Thus, using one single life table may have resulted in an overestimation of the socioeconomic gradient in relative survival, except for stomach cancer for which we found a reverse association between SES and survival. Karjalainen and Pukkala[15] determined breast cancer survival by social class with two different measures of survival. First, they used the corrected survival rate in which cases dying from other causes than breast cancer were censored and no life tables from the general population were needed. Secondly, they used the relative survival rate, with the expected survival based on the life table for the entire Finnish population. The overestimation of the social class-survival gradient using the RSR as measure of outcome was small; the ratio of survival rates of the highest and lowest social class was slightly higher when the RSR was used (1.12) as compared with the corrected survival rate (1.10). We have no reason to assume that the overestimation of the SES-survival gradient was larger in the Netherlands than in Finland.

The higher risk of dying from cancer for patients from lower SES areas also was found in other studies on cancer of the lung[2], breast[15-20], colon[2,16,17,21-23], rectum[16,21,23], colorectum[24], and prostate.[16,25] A higher survival rate from stomach cancer in the lower SES groups was found in one study in females but not in males.[2] A few studies showed no association between SES and survival from cancers of the lung[16,17] or stomach[16] or a better survival in lower SES groups for cancers of the prostate[2], breast[2], and rectum in females but not in males.[2] Socioeconomic differences in survival found in other studies generally were of similar magnitude as in our study, with the exception of breast cancer, for which substantially higher relative risks of dying were found in other studies.[1] Therefore, in the Netherlands, a country with fairly equal access to health care services, we found similar gradients of survival by socioeconomic status as in other countries.

In our study, stage of disease at diagnosis explained most of the socioeconomic variation in breast and stomach cancer survival. In other studies on breast cancer, the socioeconomic gradient in survival could not be entirely explained by the distribution of stage across socioeconomic groups.[15,18] The association between SES and breast cancer stage in our study may be caused by socioeconomic differences in the length of delay between the occurrence of the first symptoms and the time of diagnosis. Women with a high SES women may have a greater tendency to seek medical advice for cancer symptoms than women with a low SES. We found that stage could not explain the socioeconomic variation in survival for cancer of the colorectum, lung, and prostate, which concurred with the results from others on colorectal[21,24] and prostate cancer.[25] Treatment was studied in broad categories and was associated with stage. After adjustment for stage, no substantial influence of treatment on the socioeconomic survival gradient was found.

Our findings show that fairly equal access to health care services, a situation

which has been achieved in the Netherlands, does not guarantee equal survival chances for all cancer patients. Because we mainly focused our research on stage at diagnosis (assumed to be related to access to the health care system), the next step would be to study other possible determinants of socioeconomic variation in cancer survival. Such determinants include host factors such as nutritional status, psychological well being, social support, immune response, tumour aggressiveness, and co-morbidity, and factors other than access, which determine quality of care, such as hospital size and type and adherence to guidelines.

References

1. Schrijvers CTM, Mackenbach JP. Cancer patient survival by socioeconomic status in seven countries: a review for six common cancer sites. J Epidemiol Community Health 1994;48:-441-446.
2. Kogevinas M, Marmot MG, Fox AJ, Goldblatt PO. Socioeconomic differences in cancer survival. J Epidemiol Community Health 1991;45:216-219.
3. Maandbericht gezondheidsstatistiek, oktober 1991, Den Haag: CBS, 1991: 4-6.
4. MacLennan R, Muir C, Steinitz R, Winkler A (eds). Cancer registration and its techniques. IARC scientific publications no. 21. Lyon: IARC, 1978.
5. Bakker D, Coebergh JWW, Crommelin MA, Verhagen-Teulings MTh. Netherlands, Eindhoven. In: Muir C, Waterhouse J, Mack T, Powel J, Whelan S (eds). Cancer incidence in five continents, Vol V. IARC scientific publications no. 88. Lyon: IARC, 1987: 574-579.
6. Coebergh JWW, Verhagen-Teulings MTh, Crommelin MA, Masseling E, Heijden LH van der. In: Parkin DM, Muir CS, Whelan SL, Gao YT, Ferlay J, Powell J (eds). Cancer incidence in five continents, Vol VI. IARC scientific publications no. 120. Lyon: IARC, 1992: 666-669.
7. Mackenbach JP, Mheen H van de, Stronks K. A prospective cohort study investigating the explanation of socio-economic inequalities in health in the Netherlands. Soc Sci Med 1994;38-:299-308.
8. Greenwald HP, Polissar NL, Borgatta EF, McCorkle R. Detecting survival effects of socioeconomic status: problems in the use of aggregate measures. J Clin Epidemiol 1994;47-:903-909.
9. Hermanek P, Sobin LH (eds). UICC TNM Classification of malignant tumours. 4th ed. UICC. Berlin: Springer-Verlag, 1987.
10. Coebergh JWW, Crommelin MA, Heijden LH van der, Hop WCJ, Verhagen-Teulings MTh. Survival of cancer patients in southeastern Netherlands. Ned Tijdschr Geneeskd 1991;-135:938-943.
11. Cutler SJ, Ederer F. Maximum utilisation of the life table method in analyzing survival. J Chronic Dis 1958;8:699-712.
12. Hakulinen T, Abeywickrama KH. A computer program package for relative survival analysis. Comput Methods Programs Biomed 1985;19;197-207.
13. Hakulinen T, Tenkanen L. Regression analysis of relative survival rates. Applied Statistics 1987;36:309-317.
14. Payne CD. The GLIM system release 3.77. Generalized linear interactive modelling manual. Oxford: Numerical Algorithms Group, 1985.
15. Karjalainen S, Pukkala E. Social class as a prognostic factor in breast cancer survival. Cancer 1990;66:819-826.
16. Vågerö D, Persson G. Cancer survival and social class in Sweden. J Epidemiol Community Health 1987;41:204-209.
17. Bonett A, Roder D, Esterman A. Determinants of case survival for cancers of the lung, colon, breast and cervix in South Australia. Med J Aust 1984;141:705-709.
18. Bassett MT, Krieger N. Social class and black-white differences in breast cancer survival. Am J Public Health 1986;76:1400-1403.
19. Morrison AS, Lowe CR, MacMahon B, Ravnihar B, Yasa S. Incidence risk factors and survival in breast cancer: report on five years of follow-up observation. Eur J Cancer 1977;13:209-214.
20. Carnon AG, Ssemwogerere A, Lamont DW, Hole DJ, Mallon EA, Gillis CR. Relation between socioeconomic deprivation and pathological prognostic factors in women with breast

cancer. Br Med J 1994;309:1054-1057.
21. Wegner EL, Kolonel LN, Nomura AMY, Lee J. Racial and socioeconomic status differences in survival of colorectal cancer patients in Hawaii. Cancer 1982;49:2208-2216.
22. Auvinen A. Social class and colon cancer survival in Finland. Cancer 1992;70:402-407.
23. Dayal H, Polissar L, Yang CY, Dahlberg S. Race, socioeconomic status, and other prognostic factors for survival from colo-rectal cancer. J Chronic Dis 1987;40:857-864.
24. Brenner H, Mielck A, Klein R, Ziegler H. The role of socioeconomic factors in the survival of patients with colorectal cancer in Saarland/Germany. J Clin Epidemiol 1991;44:807-815.
25. Dayal HH, Polissar L, Dahlberg S. Race, socioeconomic status, and other prognostic factors for survival from prostate cancer. J Natl Cancer Inst 1985;74:1001-1006.

4.3 Validation of cancer prevalence data from a postal survey by comparison with cancer registry records*

4.3.1 Introduction

In measuring the prevalence of chronic health conditions, many epidemiologic investigations rely on self-reports of respondents to a health interview survey containing a checklist of chronic conditions. Only a few studies have reported on the validity of these data, on the basis of comparisons with either clinical examinations or medical records.[1-11] In general, the results are disappointing at the individual level, but it has remained unclear to what extent prevalence estimates, both for the total population and for subpopulations, are really biased.

This paper reports the results of a study that aimed to validate data on self-reported cancer prevalence obtained through a postal survey conducted in the Southeastern Netherlands in 1991. We had the rather unique opportunity to compare self-reported cancer prevalence at the individual level with data on cancer prevalence from a population based cancer registry in the same area, which was the gold standard in this study. This enabled us to assess the effect of misclassification on the prevalence estimate for the whole population, and to assess the effect of differential misclassification on estimates of prevalence ratios by age, sex, education, and urbanization.

4.3.2 Materials and methods

In 1991, a postal survey was conducted among 27070 noninstitutionalized inhabitants (15-74 years) of the Southeastern part of the Netherlands. The municipal registries were used as a sampling frame. These registries have a practically complete coverage of the population and are kept up-to-date continuously with respect to births, deaths, changes in marital status, and changes of address. The response rate of the survey was 70.1 percent, which resulted in a study population of 18973 respondents. This survey is part of the baseline measurement of a prospective cohort study.[12]

One of the questions in the survey was: "Do you currently suffer from any of the following chronic conditions or did you suffer from any of these during the last five years?" A list of 23 chronic conditions was then presented, one of which was "malignant disease or cancer".

The answer to this question (yes or no) was linked with records from the Eindhoven Cancer Registry.[13] This population based cancer registry serves an area of about one million inhabitants, containing the area where the survey took place. The registry identifies newly diagnosed cases of cancer through routine reports

---

* Schrijvers CTM, Stronks K, Mheen DH van de, Coebergh JWW, Mackenbach JP. Am J Epidemiol 1994;139:408-414

from departments of pathology and radiotherapy and through inpatient records from all community hospitals in the region, as well as through data from specialized departments and hospitals outside of the region.[13] Incident cases of primary nonmelanoma skin cancer are also registered[13,14], for which outpatient dermatology records are also checked.

Survey data and registry records were linked in a two-step procedure. First, a combination of the respondent's date of birth and the first two characters of his or her last name at birth was used as a linking key. For respondents to the survey, these data came from the municipal registries and were missing for 468 respondents. In a second step, these and other data (such as sex, initials, surname, address, and postal code) that were present in both the cancer registry and the survey were used for a visual inspection, checking the correctness of matches which appeared in the first step. Finally, the registry records were searched intensively for the names of respondents with a false positive report of cancer.

A total of 565 respondents did not answer the survey question on cancer. For these respondents, no record linkage took place. The survey question asked about cancer prevalence in the previous five years (1986-1991). However, prevalent cases could have been diagnosed before 1986; therefore, all computerized records, starting in 1971, were used for record linkage. The period 1981-1991 was taken as the reference period in our analysis. The results from other periods of registry review will also be discussed.

Record linkage resulted in a data file for 17940 respondents (18973 - 468 - 565) containing age, sex, education, and urbanization from the survey and year of diagnosis and primary cancer site from the cancer registry. Age was divided into four categories: 15-44, 45-54, 55-64, and 65-74 years. Educational level was classified according to number of years of education: low (6-9 years), intermediate (10 years) and high (11-18 years). The degree of urbanization of the respondents' place of residence was classified in four categories, from high for the large cities to low for the smallest rural communities. For specific subpopulations, directly age-standardized prevalences were calculated with 95% Confidence Intervals (CI), using the entire study population as a standard.

4.3.3 Results

Table 1 shows the results of the record linkage for the period 1981-1991. A total of 355 respondents (1.98 percent) reported that they had cancer, while 17585 respondents reported no cancer. The number of respondents who reported cancer but were not found in the registry (false positives) was 94 (26.5 percent), while the number of respondents who did not report cancer but were found in the registry (false negatives) was 212 (1.21 percent).

The sensitivity of survey data was 0.552 (95 percent confidence interval (CI) 0.507-0.597), and the specificity was 0.995 (95 percent CI 0.994-0.996) (table 2).

Table 1. Results of record linkage between cancer prevalence data from a postal survey and cancer registry records (1981-1991), Southeastern Netherlands, 1991

	Respondent found in cancer registry		
Postal survey	Yes	No	Total
Yes	261 (233)*	94 (122)	355
No	212 (115)	17373 (17470)	17585
Total	473 (348)	17467 (17592)	17940

* Numbers in parentheses refer to record linkage after the exclusion of nonmelanoma skin cancer from the cancer registry records.

Table 2. Sensitivity, specificity, and predictive values of positive and negative answers in record linkage between cancer prevalence data from a postal survey and cancer registry records (1981-1991), Southeastern Netherlands, 1991

			Predictive value	
	Sensitivity	Specificity	Positive answer	Negative answer
Including nonmelanoma skin cancer	0.552 (0.507-0.597)	0.995 (0.994-0.996)	0.735 (0.689-0.781)	0.988 (0.986-0.990)
Excluding nonmelanoma skin cancer	0.670 (0.621-0.719)	0.993 (0.992-0.994)	0.656 (0.607-0.705)	0.993 (0.992-0.994)

* Numbers in parentheses, 95% confidence interval

The percentage of respondents who were found in the cancer registry but who did not report that they had cancer was 45 (212 out of 473) for the total group. The percentage of these negative self-reports varied from 100 for cervical cancer to 16 for breast cancer (table 3).

A substantial number of false negatives (n=97 or 46 percent) were registered with nonmelanoma skin cancer. Therefore, we also determined the validity of survey data after the exclusion of nonmelanoma skin cancer from the cancer registry. When the 125 prevalent cases of nonmelanoma skin cancer were excluded from the cancer registry records, the number of false positives became 122 (34.4 percent), and the number of false negatives became 115 (0.65 percent) (table 1). Consequently, sensitivity improved to 0.670 (95 percent CI 0.621-0.719), and specificity became 0.993 (95 percent CI 0.992-0.994) (table 2).

Table 3. Percentage of patients found in a cancer registry who did not report cancer in a postal survey (negative self-reports), for cancer sites with ≥ 10 prevalent cases found in the cancer registry: Southeastern Netherlands, 1991

Cancer site	No. of patients found in the cancer registry	% of negative self-reports
Cervix	10	100
Skin (nonmelanoma)	125	78
Bladder	30	53
Skin (melanoma)	20	50
Uterus	19	47
Prostate	22	32
Colon	36	28
Rectum	20	25
Lung	32	22
Breast	85	16
All sites combined	473	45

Including nonmelanoma skin cancer, the prevalence of cancer was 1.98 percent (95 percent CI 1.78-2.18) based on survey data, while the prevalence based on registry records was 2.64 percent (95 percent CI 2.41-2.87). Thus, cancer prevalence was underestimated by 25 percent ((2.64-1.98)/2.64 x 100 percent) using survey data.

After the exclusion of nonmelanoma skin cancer from registry records, cancer prevalence was overestimated by a negligible 2 percent ((1.94-1.98)/1.94 x 100 percent).

Table 4 shows the age-standardized prevalence estimates for subgroups of the study population, according to both the survey and the cancer registry. The underestimation of cancer prevalence in the survey was larger for men (ratio of prevalences=0.64) than for women (ratio of prevalences=0.85). This resulted in a prevalence ratio of women to men of 1.37 using survey data, as opposed to 1.02 using registry data. For respondents below the age of 45 years, the prevalence of cancer was overestimated using survey data (ratio of prevalences = 1.46). For the other age categories, the prevalence was underestimated, especially for respondents aged 65 years or more. The ratio of prevalence in old respondents to young respondents was much lower according to the survey data (9.47) than according to the registry (21.7). In the highest educational category, prevalence was well estimated (ratio of prevalences = 1.01), while for the low and intermediate categories, prevalence was underestimated. The ratio of prevalences for the lowest educational category relative to the highest was underestimated using survey data (1.01) as compared with registry data (1.31). The prevalence of cancer was underestimated by the survey in every category of urbanization, particularly in rural

communities. The urban:rural prevalence ratio was overestimated using survey data (2.03) as compared with registry data (1.72).

After the exclusion of nonmelanoma skin cancer, cancer prevalence was still underestimated in men but not in women (table 4). The overestimation of prevalence in the youngest age group became larger; underestimation in the other age groups became much smaller (65-74 years) or disappeared (45-64 years). For the lowest and highest educational categories, an overestimation of cancer prevalence was now found, while an underestimation remained apparent in the intermediate category. For every degree of urbanization, underestimation of cancer prevalence changed into an overestimation.

The difference in female:male prevalence ratios between survey data and registry data became smaller after the exclusion of nonmelanoma skin cancer. For education, the difference in prevalence ratios (low:high) also became much smaller, while for age and urbanization, the differences in prevalence ratios based on survey data or registry data remained of similar magnitude. Thus, after the exclusion of nonmelanoma skin cancer, some prevalence ratios became smaller, but the patterns of prevalence ratios essentially remained the same.

Table 4. Age-standardized prevalences (%) of cancer based on data from a postal survey and from cancer registry records (1981-1991), according to sex, age, educational level, and urbanization: Southeastern Netherlands, 1991

	Prevalence				
	Survey (a)	Registry (b), including nonmelanoma skin cancer	Registry (c), excluding nonmelanoma skin cancer	Ratio (a:b), including nonmelanoma skin cancer	Ratio (a:c), excluding nonmelanoma skin cancer
Sex					
Women	2.28 (1.97-2.58)*	2.67 (2.34-3.00)	2.11 (1.82-2.40)	0.85	1.08
Men	1.67 (1.40-1.93)	2.61 (2.28-2.94)	1.76 (1.49-2.04)	0.64	0.95
Ratio (women:men)	1.37	1.02	1.20		
Age (years)					
15-44	0.51 (0.34-0.68)	0.35 (0.21-0.49)	0.28 (0.15-0.41)	1.46	1.82
45-54	1.46 (1.11-1.82)	1.97 (1.56-2.38)	1.44 (1.09-1.79)	0.74	1.01
55-64	3.12 (2.58-3.66)	3.90 (3.30-4.50)	3.08 (2.54-3.62)	0.80	1.01
65-74	4.83 (4.03-5.64)	7.61 (6.62-8.60)	5.38 (4.53-6.23)	0.63	0.90
Ratio (old:young)	9.47	21.7	19.2		
Educational level					
Low	2.01 (1.70-2.33)	2.59 (2.24-2.94)	1.88 (1.58-2.18)	0.78	1.07
Intermediate	1.94 (1.56-2.31)	2.97 (2.50-3.44)	2.18 (1.78-2.58)	0.65	0.89
High	1.99 (1.59-2.39)	1.97 (1.57-2.37)	1.77 (1.39-2.14)	1.01	1.12
Ratio (low:high)	1.01	1.31	1.06		
Urbanization					
1 (rural)	1.13 (0.67-1.59)	1.73 (1.16-2.30)	1.06 (0.61-1.51)	0.65	1.07
2	1.56 (1.13-2.00)	2.41 (1.87-2.95)	1.55 (1.11-1.99)	0.65	1.01
3	1.83 (1.29-2.37)	2.26 (1.66-2.86)	1.52 (1.03-2.02)	0.81	1.20
4 (urban)	2.29 (2.01-2.58)	2.98 (2.65-3.31)	1.90 (1.64-2.16)	0.77	1.21
Ratio (urban:rural)	2.03	1.72	1.79		

* Numbers in parentheses, 95% confidence interval

4.3.4 Discussion

Our study shows that the prevalence of cancer was underestimated by the postal survey we conducted in a region in the Southeastern Netherlands. In other studies, it was also found that, for cancer, data as reported in a survey do not correspond satisfactorily to those obtained from clinical examinations[2] or medical records.[4,5,8] Furthermore, the degree of underestimation differed by sex, age, education, and degree of urbanization, which leads to a biased picture of variation in cancer prevalence between subgroups of the population if one uses data from a postal survey.

The period of registry review chosen in this study was 10 years, because we assumed that most cases that were prevalent in 1986-1991 were incident in 1981-1991. The influence of choosing a broader time window was determined by considering all computerized records of the Eindhoven Cancer Registry (1971-1991). Sensitivity declined from 0.552 to 0.484, and specificity barely changed from 0.992 to 0.996. The percentage of false positives decreased from 26.5 to 20.8, while the percentage of false negatives increased from 1.21 to 1.71. The prevalence based on the registry consequently became higher, and underestimation of the prevalence by the survey increased from 25 percent to 39 percent. This higher underestimation may partly have been due to the linkage of respondents who were diagnosed many years previously and were cured before 1986, and therefore correctly did not report in the survey that they had cancer. If the period of registry review is shortened to 1986-1991, the percentage of false positives increases dramatically to 40.8, but it is clear that many of these cases may have been incident before 1986. We consider the results reported in this paper based on the time-window 1981-1991 to be best estimates.

The underestimation of cancer prevalence by the survey is due to a substantial number of false negative cases, 46 percent of which (n=97) were registered as nonmelanoma skin cancer in the Eindhoven Cancer Registry, 87 of which were basal cell carcinoma. It is likely that nonmelanoma skin cancer is not always considered a "malignant disease" by patients and thus may not be reported in surveys. However, 22 percent of cases of nonmelanoma skin cancer found in the cancer registry were reported in the survey. This is logical, since no distinction was made between cancer sites in our questionnaire as in many other health interview surveys. Perhaps the validity of survey questionnaires on cancer can be improved by explicitly including or excluding nonmelanoma skin cancer. However, although this would improve the validity of overall prevalence estimates, differences between subgroups would still be biased.

The percentage of respondents who appeared to have cancer without reporting it also seemed to be higher for other cancer sites with a relatively high frequency of less severe histologic types, such as cervical cancer (table 3). This finding is in agreement with the finding by Chambers et al.[5] that reporting is less accurate for less severe cancers. Not only could underreporting be due to respondents' misun-

derstanding the term "malignant disease or cancer," but it may also be a taboo subject among certain segments of society, e.g., older and less educated people. In addition, there may be differences in whether doctors inform their patients that they have cancer, e.g., according to educational level.

Incompleteness of the cancer registry, which was the gold standard in our study, cannot be excluded as an explanation for some of the 94 false positive cases. Some of the false positive respondents could have moved into the registration area around Eindhoven with a cancer diagnosis that was made elsewhere. The year of the respondent's arrival at the current address was available for 47 of the 94 false positive cases; 16 persons arrived before 1970, 14 arrived during the period 1971-1980, and 17 arrived after 1980. If (some of) the cancers in the latter category were diagnosed outside of the registration area and these 47 respondents were representative of the 94 false positive cases, up to about one third of these cases (17/47 or 36 percent) may in reality not be false positive cases.

Finally, for both false negative and false positive cases, general errors in filling in the questionnaire could have occurred and could explain part of the discrepancy between survey data and cancer registry data.

We do not know whether our results on cancer can be generalized to other chronic conditions measured by health interview surveys. The existence of taboos may have contributed to less accurate prevalence estimates for cancer, but this may be of less importance or no importance for other chronic conditions. However, the overall validity of survey data concerning cancer prevalence was reported to be higher than that for data on other chronic conditions.[1,2,6] We therefore expect the situation with respect to other chronic conditions to be worse, not better, than that with cancer.

References

1. Belloc NB. Validation of morbidity survey data by comparison with hospital records. J Am Stat Assoc 1954;49:819-846.
2. Krueger DE. Measurement of prevalence of chronic disease by household interviews and clinical evaluations. Am J Public Health 1957;47:953-960.
3. Meltzer JW, Hochstim JR. Reliability and validity of survey data on physical health. Public Health Rep 1970;85:1075-1086.
4. Madow WG. Net differences in interview data on chronic conditions and information derived from medical records. Vital and health statistics, series 2, no. 57. Hyattsville, MD: National Center for Health Statistics, 1973.
5. Chambers LW, Spitzer WO, Hill GB, Helliwell BE. Underreporting of cancer in medical surveys: a source of systematic error in cancer research. Am J Epidemiol 1976;104:141-145.
6. Colditz GA, Martin P, Stampfer MJ, et al. Validation of questionnaire information on risk factors and disease outcomes in a prospective cohort study of women. Am J Epidemiol 1986;123:894-900.
7. Ollson L, Svärdsudd K, Nilsson G, Ringqvist I, Tibblin G. Validity of a postal questionnaire with regard to the prevalence of myocardial infarction in a general population sample. Eur Heart J 1989;10:1011-1016.
8. Harlow SD, Linet MS. Agreement between questionnaire data and medical records: the evidence for accuracy of recall. Am J Epidemiol 1989;129:233-248.
9. Linet MS, Harlow SD, McLaughlin JK, McCaffrey LD. A comparison of interview data and medical records for previous medical conditions and surgery. J Clin Epidemiol 1989;42:1207-1213.
10. Midthjell K, Holmen J, Bjørndal A, Lund-Larsen PG. Is questionnaire information valid in

the study of a chronic disease such as diabetes? The Nord-Trøndelag diabetes study. J Epidemiol Community Health 1992;46:537-542.
11. Heliövaara M, Aromaa A, Klaukka T, Knekt P, Joukamaa H, Impivaara O. Reliability and validity of interview data on chronic diseases: The Mini-Finland Health Survey. J Clin Epidemiol 1993;46:181-191.
12. Mackenbach JP, Mheen DH van de, Stronks K. A prospective cohort study investigating the explanation of socio-economic inequalities in health in the Netherlands. Soc Sci Med 1994;38:299-308.
13. Coebergh JWW, Verhagen-Teulings MTh, Crommelin MA, Masseling E, Heijden LH van der. In: Parkin DM, Muir CS, Whelan SL, Gayo Y-T, Ferlay J, Powell J (eds). Cancer incidence in five continents, volume VI. IARC scientific publications no. 120. Lyon: IARC, 1992: 666-669.
14. Coebergh JWW, Neumann HAM, Vrints LW, Heijden LH van der, Meijer WJ, Verhagen-Teulings MTh. Trends in the incidence of non-melanoma skin cancer in the SE Netherlands 1975-1988: a registry-based study. Br J Dermatol 1991;125:353-359.

4.4 Socioeconomic status and prognostic factors among prevalent cancer cases

4.4.1 Introduction

Cancer registries do often not contain information on factors which have been hypothesized repeatedly to be important determinants of socioeconomic variation in cancer survival, such as co-morbidity, social support and psychological wellbeing[1]. We therefore used information on this type of prognostic factors from prevalent cases of cancer that had been identified through a postal survey (paragraph 4.3). We studied the association between socioeconomic status of prevalent cases of cancer and (a) the number of chronic conditions they reported in a postal survey and (b) the number of life events they experienced during the past twelve months, as reported in the survey.

4.4.2 Patients and methods

Data were derived from a postal survey, conducted in 1991, which is the baseline measurement of the Longitudinal Study on Socio-Economic Health Differences (SEHD). The design and objectives of this study have been described elsewhere,[2] and in chapter 3. One of the questions in the survey was "Do you currently suffer from any of the following conditions or did you suffer from any of these during the last 5 years?". A list of 23 chronic conditions was presented, one of which was "malignant disease or cancer". Other chronic conditions on this list were for example diabetes, low back pain, and heart disease. Prevalent cases of cancer were first selected according to their answer (yes or no) to this question and then their answers were validated against cancer registry records as described in 4.3. This procedure resulted in 261 prevalent cases of cancer. We then excluded patients who had been diagnosed before 1986 (n=73) to restrict the time passed between a cancer diagnosis and the actual time of the survey. We also excluded patients for which the educational level was unknown (n=7). This resulted in 181 prevalent cases of cancer with various cancer sites, of which the most common sites were breast (n=54), colon (n=20) and lung (n=17). Due to these small numbers, it was not possible to conduct analyses for specific cancer sites.

The number of other chronic conditions per patient was counted and then the variable was dichotomized into none (n=49 or 27.1%) versus at least one other chronic condition (n=132 or 72.9%). The survey also contained a question on specific life events during the year preceding the survey, such as unemployment of the respondent or his/her partner, serious illness of the partner, a divorce or a worsened financial situation. The number of events was counted and a dichotomized variable was created with categories none (n=89 or 49.7%) versus at least one event (n=90 or 50.3%). The highest attained educational level of the respondent was classified in 3 categories: (1) primary school or lower vocational (n=96), (2) lower general or intermediate vocational (n=50), (3) intermediate/higher general,

higher vocational, university (n=35). Furthermore, we used the area-based measure of socioeconomic status in 3 categories: low (n=74), intermediate (n=51) or high (n=56) and in 5 categories: low (n=64), 2 (n=20), 3 (n=38), 4 (n=18), high (n=41) (see 3.2).

Logistic regression analyses were conducted to calculate an odds ratio with 95% confidence intervals, adjusted for sex and age (5-year categories) with the highest socioeconomic group as reference category.

4.4.3 Results

The association between socioeconomic status and chronic conditions (none versus at least one) and life events (none versus at least one) is given in tables 1 and 2 for each of the 3 measures of socioeconomic status. Both the results from univariate analyses and logistic regression analyses (adjusted for age and sex) are presented.

Table 1 shows that the percentage of patients with at least 1 chronic condition did not vary systematically between educational categories. This is also reflected in the odds ratios for both the highest and intermediate category, which were both greater than 1, but the odds ratio for the intermediate category was higher than that for the lowest educational category. If the area-based measure in either 3 or 5 categories was used, we observed no clear pattern in the number of chronic conditions by socioeconomic status. We may therefore conclude that there seems to be no clear association between the socioeconomic status of prevalent cancer patients and the number of chronic conditions they reported in a postal survey.

Table 2 shows that patients from the low educational group have more often experienced at least 1 life event during the 12 months preceding the survey than patients in the higher educational groups. When the area-based measure of socioeconomic status in 3 categories was used, we observed the same pattern: the percentage of patients with at least 1 life event was higher in the lowest and intermediate socioeconomic group. When we used the area-based measure in 5 categories, we observed the same type of pattern, but the gradient was inconsistent. We may conclude that there seems to be an association between the socioeconomic status (for 2 out of 3 measures) of prevalent cancer cases and the number of life events they experienced.

Table 1. Number of chronic conditions, odds ratio with 95% confidence interval, by socioeconomic status, prevalent cancer cases, Southeastern Netherlands, 1991

	Number of chronic conditions		
	0	at least 1	Odds ratio
Education			
Low	28.1	71.9	1.27 (0.50-3.20)*
Intermediate	22.0	78.0	1.48 (0.52-4.24)
High	31.4	68.6	1.00
Area-based			
Low	25.7	74.3	1.01 (0.46-2.36)
Intermediate	31.4	68.6	0.71 (0.29-1.74)
High	25.0	75.0	1.00
Area-based			
Low	23.4	76.6	1.51 (0.58-3.91)
2	30.0	70.0	1.06 (0.30-3.74)
3	34.2	65.8	0.70 (0.35-1.93)
4	16.7	83.3	2.76 (0.62-1.22)
High	29.3	70.7	1.00
Total	27.1 (n=49)	72.9 (n=132)	

* 95% confidence interval

Table 2. Number of life events, odds ratio with 95% confidence interval, by socioeconomic status, prevalent cancer cases, Southeastern Netherlands, 1991

	Number of life events		
	0	at least 1	Odds ratio
Education			
Low	43.2	56.8	2.11 (0.90-4.91)*
Intermediate	57.1	42.9	1.06 (0.42-2.65)
High	57.1	42.9	1.00
Area-based			
Low	46.6	53.4	1.61 (0.77-3.35)
Intermediate	48.0	52.0	1.41 (0.64-3.08)
High	55.4	44.6	1.00
Area-based			
Low	42.9	57.1	1.69 (0.74-3.85)
2	45.0	55.0	1.38 (0.45-4.25)
3	56.8	43.2	0.91 (0.36-2.28)
4	55.6	44.4	0.84 (0.27-2.65)
High	53.7	46.3	1.00
Total	49.7 (n=89)	50.3 (n=90)	

* 95% confidence interval

4.4.4 Discussion

The association between socioeconomic status and two possible prognostic factors was studied among prevalent cancer cases. No association was found between the socioeconomic status of these cancer patients and the number of chronic conditions, while patients with a low socioeconomic status reported more often at least one adverse life event during the past twelve months.

The results as reported here apply to a group of prevalent cancer cases diagnosed during the period 1986-1991 with a variety of cancers. However, if an association between socioeconomic status and the number of life events is found in such a heterogeneous group it points at possible associations for diagnostic subgroups of patients.

A lack of an association between socioeconomic status and the number of chronic conditions might (partly) be caused by differential misreporting. In another study, socioeconomic inequalities in the prevalence of specific chronic conditions (chronic non-specific lung disease, heart disease and diabetes mellitus) were found to be underestimated as a result of differential misreporting.[3] This could explain why we did not find an association between socioeconomic status and the number of chronic conditions. Other data sources might be used to obtain more objective information on chronic conditions in cancer patients and this was done in another study which investigated socioeconomic status and co-morbidity, using information abstracted from clinical records (see 4.5). Misreporting of life events might have been a less serious problem, as this involved events which can be easily remembered and it does not require specific medical knowledge on diagnoses, such as with specific chronic conditions.

The association between socioeconomic status and survival and the role of the studied prognostic factors on this association could not be studied with this material. Earlier studies in which psychosocial factors such as stressful life events were related to cancer survival showed conflicting results[4-6]. A carefully designed prospective study on socioeconomic variation in cancer survival would have to combine both disease related factors (such as stage of disease at diagnosis and histological type of the tumour) and host factors such as psychosocial factors and co-morbidity. Our analysis shows that such a study might reveal that in the Netherlands, host factors do play a role in causing socioeconomic inequalities in cancer survival.

References

1. Leon D, Wilkinson RG. Inequalities in prognosis: socio-economic differences in cancer and heart disease survival. In: Fox J (ed). Health Inequalities in European Countries. Aldershot: Gower, 1989.
2. Mackenbach JP, Mheen DH van de, Stronks K. A prospective cohort study investigating the explanation of socio-economic inequalities in health in the Netherlands. Soc Sci Med 1994;38-:299-308.
3. Mackenbach JP, Looman CWN, Meer JBW van der. Misreporting of chronic conditions in a health interview survey: differentials by level of education (submitted).
4. Derogatis LR, Abelhoff MD, Melisaratos N. Psychological coping mechanism and survival

time in metastatic breast cancer. J Am Med Assoc 1979;242:1504-1508.
5. Marshall JR, Funch DP. Social environment and breast cancer. A cohort analysis of cancer patient survival. Cancer 1983;52:1546-1550.
6. Jamison FN, Burish TG, Wallston KA. Psychogenic factors in predicting survival of breast cancer patients. J Clin Oncol 1987;5:768-772.

4.5 Socioeconomic status and co-morbidity among incident cancer patients in the Southeastern Netherlands, 1993

4.5.1 Introduction

Most studies on the association between socioeconomic status and cancer survival found better survival for patients from higher socioeconomic groups.[1-3] Several possible explanations for these findings have been studied such as socioeconomic variation in the stage-distribution at diagnosis, biological features of a tumour and treatment.[4] However, studies in which adjustment was made for an important prognostic factor such as stage of disease at diagnosis, still showed a higher risk of dying from cancer in patients with a low socioeconomic status.[5-9] Apart from residual confounding by stage, other possible explanations of the association between socioeconomic status and cancer survival have been hypothesized. Concomitant diseases could be more frequent among cancer patients of lower socioeconomic status, and a larger number of co-morbid conditions and/or more severe conditions will in general lead to a poorer prognosis.[10]

In this chapter, we report the results of a preliminary study on socioeconomic variation in the number and type of serious co-morbid conditions among cancer patients diagnosed in 1993 in the registration area of the Eindhoven Cancer Registry. The following two hypotheses were tested: (1) co-morbidity is more common among cancer patients with a low socioeconomic status than among cancer patients with a high socioeconomic status (2) cancer patients with a low socioeconomic status experience more severe types of co-morbid conditions than cancer patients with a high socioeconomic status.

4.5.2 Patients and methods

Data for this study came from the Eindhoven Cancer Registry, a population based cancer registry covering an area of about one million inhabitants in the Southeastern part of the Netherlands. The registry identifies newly diagnosed cases of cancer through routine reports from departments of pathology and radiotherapy, through inpatient records from all eight community hospitals in the region, as well as through data from specialized departments and hospitals outside of the region.

In 1993, the Eindhoven Cancer Registry started to register serious co-morbidity at the time of diagnosis for all new patients in the registration area, according to a 'pre-fixed' list of conditions developed by Charlson[11]. Information on co-morbid conditions was abstracted from medical records during routine registration practice by experienced clerks, who had received supplementary training before starting this task. The abstracted information concerned the type of co-morbid condition and each condition was registered separately.

For this analysis, all patients diagnosed in 1993 with one of the five most common cancers (lung, breast, colorectum, prostate, and stomach) were selected

for study (n=1750). Socioeconomic status was divided in three categories (low, intermediate and high), based on the postcode of residence at time of diagnosis of each patient (paragraph 3.2). The number of co-morbid conditions per patient was calculated (range 0 to 5) and a new variable indicating this number was constructed with 4 categories (0, 1, 2 to 5, unknown). Patients with missing information on co-morbidity were excluded (n=110 or 6.3%), leaving 1640 patients for the analysis.

Apart from descriptive analyses, logistic regression analyses were performed to quantify the association between socioeconomic status and the number of co-morbid conditions (none versus at least one condition), while adjusting for age (5-year categories) and sex. The measure of effect is the odds ratio, and corresponding 95% confidence intervals were calculated.

4.5.3 Results

The distribution of patients across socioeconomic groups and cancers is given in table 1.

Table 1. Number and percentage of patients by socioeconomic status, five most common cancers, Southeastern Netherlands, 1993

	Breast		Lung		Colorectum		Prostate		Stomach		Total	
SES	N	%	N	%	N	%	N	%	N	%	N	%
Low	160	35.0	189	42.8	153	39.8	81	33.7	44	37.3	627	38.2
Intermediate	172	37.6	149	33.7	140	36.5	76	31.7	45	38.1	582	35.5
High	125	27.4	104	23.5	91	23.7	83	34.6	29	24.6	432	26.3
Total	457	100	442	100	384	100	240	100	118	100	1641	100

The association between socioeconomic status and the number of co-morbid conditions for the five cancer sites combined is given in table 2: the percentage of patients without a serious co-morbid condition was relatively high in the high socioeconomic group. The percentage of patients with one or at least two conditions was higher in the low and intermediate socioeconomic group as compared with the high socioeconomic group. For breast cancer we observed the same pattern as for all sites combined, with a higher percentage of patients without a co-morbid condition in the higher socioeconomic groups. Furthermore, the percentage of patients with one condition was higher in the low socioeconomic group, while the percentage with at least two conditions was lower in the high socioeconomic group. The results for lung cancer are less clear. Although more patients were registered without a co-morbid condition in the high socioeconomic group there was no gradient, neither for this category nor for the other two categories (one or at least two conditions). The results for colorectal cancer were very similar to those for lung cancer: no clear gradient for the categories without or with one condition and

a lower percentage with a least two co-morbid conditions in the high socioeconomic category. For cancer of the prostate, the percentage of patients without a co-morbid condition did not differ between socioeconomic groups. The percentage with one condition was higher in the low socioeconomic group, while the percentage with at least two conditions was lower in this patient group. For stomach cancer there was no clear pattern in the distribution of number of co-morbid conditions across socioeconomic groups.

Table 3 shows the percentage of patients in categories of co-morbidity for each of four age groups. The percentage of patients without a co-morbid condition was higher in each younger age group, while the percentage with one, or at least two conditions was higher in each older age group.

Table 2. Percentage of patients by socioeconomic status and number of co-morbid conditions, five most common cancers, Southeastern Netherlands, 1993

	All sites				**Breast**				**Lung**			
	Number of conditions				Number of conditions				Number of conditions			
SES	0	1	2-5	Total	0	1	2-5	Total	0	1	2-5	Total
Low	46.6	35.7	17.7	100	54.4	34.4	11.2	100	39.2	36.0	24.8	100
Intermediate	48.6	33.3	18.1	100	68.0	18.6	13.4	100	32.2	41.6	26.2	100
High	57.4	30.6	12.0	100	79.2	17.6	3.2	100	50.0	33.7	16.3	100
Total	50.1	33.5	16.4	100	66.3	23.9	9.8	100	39.4	37.3	23.3	100
	Colorectum				**Prostate**				**Stomach**			
	Number of conditions				Number of conditions				Number of conditions			
SES	0	1	2-5	Total	0	1	2-5	Total	0	1	2-5	Total
Low	48.4	30.7	20.9	100	48.1	42.0	9.9	100	40.9	45.5	13.6	100
Intermediate	45.0	38.6	16.4	100	48.7	34.2	17.1	100	37.8	42.2	20.0	100
High	51.6	36.3	12.1	100	47.0	34.9	18.1	100	37.9	44.8	17.3	100
Total	47.9	34.9	17.2	100	47.9	37.1	15.0	100	39.0	44.1	16.9	100

Table 3. Percentage of patients by age and number of co-morbid conditions, five most common cancers, Southeastern Netherlands, 1993

	Number of co-morbid conditions per patient			
Age	0	1	2-5	Total
- 44	86.7	12.2	1.1	100
45-59	67.4	26.5	6.1	100
60-74	43.0	38.1	18.9	100
75 +	31.9	42.5	25.6	100
Total	48.1	35.1	16.8	100

Table 4 shows the results from logistic regression analyses on the association between socioeconomic status and the number of co-morbid conditions (none versus at least one), both for all sites combined and for each site separately. The results are given in table 4 as odds ratios both unadjusted and adjusted for sex and age, and cancer site for all sites combined.

Table 4. Odds ratio and 95% confidence interval, at least one co-morbid condition at time of diagnosis by cancer site and socioeconomic status, Southeastern Netherlands, 1993

	Cancer site					
	All sites		Breast		Lung	
SES	unadj.	+sex,age site	unadj.	+age	unadj.	+sex,age
Low	1.55 1.21-1.98	1.56 1.20-2.05	3.19 1.88-5.44	3.35 1.85-6.06	1.55 0.96-2.52	1.69 1.01-2.83
Intermediate	1.43 1.12-1.84	1.54 1.17-2.02	1.79 1.04-3.06	2.19 1.20-3.99	2.10 1.26-3.52	2.39 1.37-4.15
High*	1.00	1.00	1.00	1.00	1.00	1.00
	Colorectum		Prostate		Stomach	
SES	unadj.	+sex,age	unadj.	+age	unadj.	+sex,age
Low	1.14 0.69-1.92	1.16 0.66-2.03	0.95 0.52-1.76	0.96 0.51-1.80	0.88 0.34-2.31	1.00 0.35-2.88
Intermediate	1.31 0.77-2.22	1.25 0.70-2.21	0.93 0.50-1.74	0.96 0.50-1.85	1.01 0.38-2.63	0.86 0.30-2.42
High*	1.00	1.00	1.00	1.00	1.00	1.00

unadj = unadjusted; * reference category

For all sites combined, we found that both the low and intermediate socioeconomic group have a 55% greater probability on being diagnosed with at least one co-morbid condition than patients from the highest socioeconomic group (adjusted for sex, age and site). Breast cancer patients in the low and intermediate socioeconomic group also have a higher chance on being diagnosed with a least one other serious condition than patients in the highest socioeconomic group. The same was found for lung cancer, but the probability on being diagnosed with a least one co-morbid condition was higher for patients in the intermediate socioeconomic group than for patients in the low socioeconomic group. The pattern for colorectal cancer was similar to that for lung cancer, but the odds ratios were smaller and the confidence intervals were wider and therefore no firm conclusions can be drawn for this cancer. For cancers of the prostate and stomach we found no clear association between socioeconomic status and the odds ratio for at least 1 other condition at diagnosis.

The association between socioeconomic status and the prevalence of major co-

morbid conditions is described in tables 5-8. Table 5 shows that for the five cancers combined, COPD was less common in the high socioeconomic group. The differences in prevalence between the socioeconomic groups for the other conditions were smaller: more hypertension, a history of myocardial infarction and other conditions in the lower socioeconomic groups.

Table 5. Prevalence of major co-morbid conditions by socioeconomic status, 5 most common cancers, Southeastern Netherlands, 1993

	SES							
	Low		Intermediate		High		Total	
Condition	N	%	N	%	N	%	N	%
COPD	101	16.1	89	15.3	38	8.8	228	13.9
Hypertension	89	14.2	79	13.6	49	11.3	217	13.2
History of MI	80	12.8	63	10.8	40	9.2	183	11.1
Other cancers*	55	8.8	63	10.8	44	10.2	162	9.9
Diabetes	34	5.4	34	5.8	18	4.2	86	5.2
Other conditions	124	19.8	84	14.4	62	14.3	305	18.6

COPD: Chronic Obstructive Pulmonary Disease; MI: Myocardial Infarction
* Excluding basal cell carcinoma of the skin

Table 6 shows that among breast cancer patients hypertension, diabetes, a history of myocardial infarction and COPD were more common among patients in the low and intermediate socioeconomic group than among patients in the high socioeconomic group.

Table 6. Prevalence of major co-morbid conditions by socioeconomic status, breast cancer, Southeastern Netherlands, 1993

	SES							
	Low		Intermediate		High		Total	
Condition	N	%	N	%	N	%	N	%
Hypertension	38	23.8	31	18.0	10	8.0	79	17.3
Diabetes	15	9.4	13	7.6	3	2.4	31	6.8
History of MI	12	7.5	9	5.2	3	2.4	24	5.2
COPD	12	7.5	9	5.2	2	1.6	23	5.0
Other cancers*	6	3.8	5	2.9	5	4.0	16	3.5
Other conditions	12	7.5	18	10.5	7	5.6	37	8.1

MI: Myocardial Infarction; COPD: Chronic Obstructive Pulmonary Disease
* Excluding basal cell carcinoma of the skin

For lung cancer patients COPD, a history of MI and other conditions were more common in the low and intermediate socioeconomic groups (table 7).

Table 7. Prevalence of major co-morbid conditions by socioeconomic status, lung cancer, Southeastern Netherlands, 1993

	SES							
	Low		Intermediate		High		Total	
Condition	N	%	N	%	N	%	N	%
COPD	50	26.5	42	28.2	17	16.3	109	24.7
Other cancers*	23	12.2	24	16.1	13	12.5	60	13.6
History of MI	28	14.8	22	14.8	9	8.7	59	13.3
Hypertension	22	11.6	13	8.7	13	12.5	48	10.9
Diabetes	15	7.9	16	10.7	8	7.7	39	8.8
Other conditions	41	21.7	37	24.8	12	11.5	90	20.4

COPD: Chronic Obstructive Pulmonary Disease; MI: Myocardial Infarction
* Excluding basal cell carcinoma of the skin

For colorectal cancer other cancers were more common in the highest socioeconomic group, while hypertension, COPD, and other conditions were less common in this patient group as compared to the intermediate and low socioeconomic group. There was no clear pattern in the prevalence of specific other cancers across socioeconomic groups.

Table 8. Prevalence of major co-morbid conditions by socioeconomic status, colorectal cancer, Southeastern Netherlands, 1993

	SES							
	Low		Intermediate		High		Total	
Condition	N	%	N	%	N	%	N	%
Other cancers*	18	11.8	18	12.9	14	15.4	50	13.0
Hypertension	20	13.1	20	14.3	9	9.9	49	12.8
COPD	19	12.4	23	16.4	4	4.4	46	12.0
History of MI	18	11.8	10	7.1	14	15.4	42	10.9
Other conditions	45	29.4	34	24.3	18	19.8	97	25.3

COPD: Chronic Obstructive Pulmonary Disease; MI: Myocardial Infarction
* Excluding basal cell carcinoma of the skin

Similar analyses were done for cancers of the prostate and stomach, which showed no clear variation in the prevalence of specific co-morbid conditions between socioeconomic groups (results not shown).

4.5.4 Discussion

We studied the association between socioeconomic status and co-morbidity at time of diagnosis of cancer patients in the Southeastern Netherlands in 1993. Overall, confidence intervals are wide and therefore a careful interpretation of the results is necessary. For the five most common cancers combined, patients with a low and intermediate socioeconomic status more often had at least one other chronic condition than patients with a high socioeconomic status. Site specific analyses showed that only for breast cancer a clear gradient in the number of co-morbid conditions by socioeconomic status existed, while the pattern for cancers of the lung and colorectum showed elevated odds ratios for the lower socioeconomic groups, but no clear gradient. Furthermore the results for colorectal cancer were not statistically significant. For cancers of the prostate and stomach, we found no clear association between socioeconomic status and the number of co-morbid conditions.

The prevalence of specific chronic conditions showed variation by socioeconomic status for all sites combined and for the individual cancer sites. For example COPD was more common in the lower socioeconomic groups for all sites combined and for lung cancer, which is in agreement with the higher prevalence of smoking in the lower socioeconomic groups.[12] In general, the differences in the prevalence of specific conditions for individual sites were small however.

Our data have some limitations which should be kept in mind while interpreting them. Firstly, socioeconomic variation in co-morbidity was investigated by studying the distribution of some broad categories of number of conditions, without considering the severity and duration of these conditions, as such information was not available from the clinical records. In general, we may assume however, that cancer survival is lower with an increasing number of co-morbid conditions, as was found in other studies,[10,11] independent of the severity of individual conditions.

Secondly, there may well be incompleteness of registration of co-morbidity by clinicians. For example, in patients with a metastasis one might expect a systematic underreporting of co-morbidity, as this may have little consequences for their treatment. This was not confirmed by our study however. For all sites combined, the association between socioeconomic status and the number of co-morbid conditions per category of stage showed a similar pattern for patients with metastasis as for patients with higher stages. Furthermore, exclusion of patients with a metastasis from the site-specific analysis did not change the results.

For all sites combined and for some specific sites, we found that patients with a low and intermediate socioeconomic status had more co-morbid conditions at the time of their cancer diagnosis. Other studies have found that the presence of co-

morbid conditions at time of diagnosis adversely affects cancer survival.[10,11] This effect could be a direct one: co-morbid conditions may affect the course of cancer. The reverse may also be true: (treatment for) cancer may accelerate the course of other pathological conditions, and this may result in a greater risk of death from other conditions, which constitute an important part of mortality in cancer patients.[13]

Overall, socioeconomic differences in the prevalence of specific conditions were not very large, so that these differences may have small implications for socioeconomic variation in survival. With respect to specific conditions, the largest difference was found in breast cancer patients for hypertension, which was more common in the low and intermediate socioeconomic group, and which could result in a higher risk of dying of cardiovascular disease in these groups. Whereas differences for other conditions were much smaller, a combination of specific conditions may result in a survival advantage for patients with a high socioeconomic status.

We conclude, from this preliminary analysis, that there seems to be an association between socioeconomic status and the number of co-morbid conditions in cancer patients. There was variation in the prevalence of specific co-morbid conditions for some sites, while a lower survival rate for lower socioeconomic groups has been found in many cancer sites. Therefore, co-morbidity might play a role in explaining socioeconomic variation in cancer survival but probably not to the same extent for all sites. As this cross sectional study included only 1 year of registration, future (longitudinal) studies will have to elaborate on the possible role of co-morbidity as explanatory factor of socioeconomic variation in cancer patient survival.

References

1. Berg JW, Ross R, Latourette HB. Economic status and survival of cancer patients. Cancer 1977;39:467-477.
2. Vågerö D, Persson G. Cancer survival and social class in Sweden. J Epidemiol Community Health 1987;41:204-209.
3. Kogevinas M, Marmot MG, Fox AJ, Goldblatt PO. Socioeconomic differences in cancer survival. J Epidemiol Community Health 1991;45:216-219.
4. Leon D, Wilkinson RG. Inequalities in prognosis: socio-economic differences in cancer and heart disease survival. In: Fox J (ed). Health Inequalities in European Countries. Aldershot: Gower, 1989.
5. Wegner EL, Kolonel LN, Nomura AMY, Lee J. Racial and socioeconomic status differences in survival of colorectal cancer patients in Hawaii. Cancer 1982;49:2208-2216.
6. Dayal HH, Polissar L, Dahlberg S. Race, socioeconomic status, and other prognostic factors for survival from prostate cancer. J Natl Cancer Inst 1985;74:1001-1006.
7. Bassett MT, Krieger N. Social class and black-white differences in breast cancer survival. Am J Public Health 1986;76:1400-1403.
8. Karjalainen S, Pukkala E. Social class as a prognostic factor in breast cancer survival. Cancer 1990;66:819-826.
9. Brenner H, Mielck A, Klein R, Ziegler H. The role of socioeconomic factors in the survival of patients with colorectal cancer in Saarland/Germany. J Clin Epidemiol 1991;44:807-815.
10. Satariano WA, Ragland DR. The effect of comorbidity on 3-year survival of women with primary breast cancer. Ann Intern Med 1994;120:104-110.
11. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis 1987;40:373-383.
12. Wagenknecht LE, Perkins LL, Cutter GR et al. Cigarette smoking behaviour is stonrly

related to educational status: the CARDIA study. Preventive Medicine 1990;19:158-169.
13. Brown BW, Brauner C, Minnotte MC. Non-cancer deaths in white adult cancer patients. J Nat Cancer Inst 1993;85:979-987.

Chapter 5. Deprivation and cancer survival in the South Thames area

5.1 Deprivation and survival from breast cancer*

5.1.1 Introduction

A 25% reduction in breast cancer mortality by the year 2000 among women invited for screening was set as a target for the Health of the Nation strategy in England.[1] This reduction is unlikely to be reached by a reduction in the incidence of breast cancer, because incidence at ages 45-74 is still rising,[2] and the major risk factors so far identified for breast cancer, such as nulliparity, late age at first birth and late age at menopause,[3] are not amenable to intervention. Improvement in survival is a more promising approach to the reduction of breast cancer mortality: this is the focus of the National Breast Screening Programme.[4] Considerations of equity would require different socioeconomic groups of patients to have equal chances of survival from breast cancer.[5] It is therefore important to monitor any socioeconomic variation in breast cancer survival and if possible to determine its causes.

Socioeconomic variation in breast cancer survival has been reported from Finland, Sweden, England & Wales, Scotland, the USA and Australia, using either individual,[6-8] or area-based measures[9-13] of socioeconomic status. These studies have shown that breast cancer patients from higher socioeconomic groups have higher survival rates, except for the English study which found a weak reverse gradient.[8]

We studied variation in breast cancer survival between categories of deprivation in the area covered by the South Thames Regional Health Authority (RHA), which includes London south of the River Thames and the counties of Kent, Surrey and Sussex, with a population around 6.5 million. We examined the influence of several prognostic factors on this variation, and evaluated the potential effect on mortality of eliminating any gradient in survival by category of deprivation.

5.1.2 Patients and methods

Data source and patients

Data for this study came from the Thames Cancer Registry, a population based cancer registry covering a population of 14.1 million people in Southeast England. The Registry has been operating continuously since 1960, covering the territory of what is now South Thames RHA until 1984. Coverage was extended to the territory of North Thames RHA in 1985, but because we analysed survival for women diagnosed from 1980, only women resident in South Thames RHA were included. The methods and data quality indices of the Registry have been described[14,15] and incidence for the 1980s reported.[16-19]

---

* Schrijvers CTM, Mackenbach JP, Lutz J-M, Quinn MJ, Coleman MP
Br J Cancer 1995;72:738-743

All 35,000 female residents of South Thames RHA who were diagnosed with a malignant breast tumour in the decade 1980-89 were eligible for study. The mean age at diagnosis was 63 years (range 30 to 99 years). The 2,822 (8.1%) women for whom the date of death was known but the date of diagnosis unknown (death certificate only cases, DCO), were excluded from analysis because their survival time could not be calculated. A further 2,502 (7.1%) cases with an incomplete or unknown postcode were also excluded, since their census enumeration district could not be reliably determined (see below). A total of 29,676 women (84.8% of those eligible) were included in survival analyses. No distinction was made between cases for which histological evidence of malignancy was (77.3%) or was not available to the Registry, because this percentage did not differ systematically between deprivation categories.

Deprivation score

The measure of deprivation for each woman was based on her usual residence at diagnosis, by linking the full postcode of residence to the corresponding census enumeration district (ED). Nationally, each ED contains on average 400 households. For each of the 14,386 EDs in South Thames, data from the 1981 census were obtained on four variables: overcrowding (proportion of persons in private households living at a density of more than one person per room as a proportion of all persons in private households), male unemployment (proportion of economically active males who are seeking work), low social class (proportion of all persons in private households with head of household in social class 4 or 5) and car ownership (proportion of all persons in private households with no car).

The Carstairs Index combines these four variables for a given small geographic area into a single score, considered to represent material deprivation.[20] The value of each variable for each ED is first standardised by subtracting the mean value for Great Britain as a whole, and dividing the result by the population standard deviation. The sum of the four standardised scores for each ED provides its Carstairs Index.

Each ED in South Thames was then assigned to one of five deprivation categories, constructed by ranking the Carstairs scores for all EDs in Great Britain from low ('affluent') to high ('deprived') and dividing this distribution into quintiles.

Prognostic factors

Age was initially studied in three categories: 30-49, 50-64 and 65-99 years, but survival patterns across deprivation categories were very similar for the two youngest age groups, and they were combined for analysis. Period of diagnosis was studied in two quinquennia, 1980-84 and 1985-89, since overall survival from breast cancer was higher in the later period. Stage at diagnosis (clinical or pathological) was explicitly stated in the medical records for less than 20% of breast cancer patients.[21] A simplified stage is routinely constructed by Registry staff for all cases, however, using pathology reports, operation notes and other information:

it is available for some 80% of cases. Stage was categorised in three groups for this study: local (tumour confined to the breast), regional (involvement of lymph nodes) and metastasis (spread to other organs). Patients for whom the stage at diagnosis was unknown were included in the analysis as a fourth category. Morphology was studied in three categories: ductal, other specific morphology and unknown morphology. Finally, type of treatment was studied in seven broad categories: surgery; surgery plus radiotherapy; surgery plus chemotherapy; surgery plus radiotherapy plus chemotherapy; radiotherapy plus chemotherapy; no treatment, and treatment unknown.

Survival analysis

The survival time in years for each woman was calculated as the number of days between the date of diagnosis and the date of death or December 31, 1992 (whichever occurred first) divided by 365.25. Potential follow-up time ranged from 3 to 13 years.

To adjust for mortality from other causes than breast cancer, we used the relative survival rate as measure of outcome in the univariate analyses. The relative survival rate, expressed as a percentage (RSR%), is the ratio of the survival observed in the group of cancer patients and the survival that would be expected if they were subject to the same overall mortality rates by age and calendar period as the general population.[22] Expected survival was computed from the England and Wales life table for 1981. The computer program from the Finnish Cancer Registry was used to calculate the RSR and its 95% confidence interval (CI).[23]

Multivariate analysis was conducted with a proportional hazards model adapted to the RSR[24] using GLIM.[25] The measure of outcome was the hazard ratio, which expresses the probability of death for a specific category of patients relative to a referent category with probability of death defined as unity.

The basic model included the duration of follow-up (up to 5 and 6-13 years) and deprivation category: prognostic factors were added as categorical variables in a fixed order; first, period of diagnosis, then factors considered to be intermediate in any association between deprivation and survival, namely stage at diagnosis, morphology and type of treatment. The improvement in fit of the model obtained from each additional prognostic factor was tested for statistical significance at the 5% level using the chi-square distribution for the reduction in deviance from the preceding model with the corresponding difference in degrees of freedom. The statistical significance of the trend in the hazard ratio across deprivation categories was tested by examining the effect of adding deprivation category to the model as a continuous variable.

Mortality reduction

We estimated the reduction in mortality 5 years after breast cancer diagnosis which might be achieved if any socioeconomic gradient in survival were eliminated. In order to obtain the number of deaths that would have been expected if all women

had experienced the survival of the most affluent group, cumulative (crude) death rates at five years were calculated for each 5 year age group in the most affluent patient category and applied to the numbers of women in the corresponding age group in the other deprivation categories. The potential reduction in mortality was calculated for the age groups 30-64 and 65-99 and for each deprivation category, as both the absolute and the percentage difference between observed and expected deaths. A similar calculation was done for the age group 50-69 years, which will be monitored for breast cancer mortality in relation to the national Breast Screening Programme.[26]

5.1.3 Results

A third (34%) of the women with breast cancer lived in the 32.9% of areas categorised to the most affluent quintile of the Carstairs Index, while only 6% lived in the 8.9% of areas categorised as the most deprived (table 1). These distributions reflect both the relative affluence of South Thames within Great Britain and the higher incidence of breast cancer in more affluent women.

Table 1. Number (%) of enumeration districts, number (%) of cases, and relative survival rates at 5 and 10 years by deprivation category, breast cancer, South Thames, 1980-1989

Deprivation category	Number of EDs	% of EDs	Number of cases	% of cases	5 year RSR (95% CI)	10 year RSR (95% CI)
Affluent	4739	32.9	10097	34.0	71 (69-73)	59 (57-61)
(2)	3251	22.6	7147	24.1	67 (65-69)	54 (52-56)
(3)	2763	19.2	6107	20.6	63 (62-64)	51 (49-53)
(4)	2359	16.4	4536	15.3	64 (62-66)	50 (47-53)
Deprived	1274	8.9	1789	6.0	60 (57-63)	48 (44-52)
Total	14386	100	29676	100	67 (66-68)	54 (53-55)

ED: enumeration district; RSR: relative survival rate; CI: confidence interval

Survival at both 5 and 10 years was higher in the more affluent patient groups. The difference in survival between the most affluent and most deprived category increased slightly with time since diagnosis (figure 1). The absolute difference in survival between these two groups was more than 10%, and the survival gradient across deprivation categories was clear, although women in the third and fourth categories had similar survival rates.

The survival gradient across deprivation categories was steeper for older women than for younger women (table 2).

The distribution of prognostic factors by deprivation category was therefore studied separately for these two age groups; an example is shown in table 3 for stage at diagnosis. For women aged 30-64 years, there was no consistent pattern in stage by deprivation category. Among women aged 65-99 years, the distribution of stage at diagnosis was more advanced in the most deprived group, of whom 17% presented with metastases.

Differences in stage distribution by age and deprivation category were generally small, however, and the patterns of survival by stage were very similar for the age groups 30-64 and 65-99 years. Stage-specific survival rates are therefore presented in table 4 for all ages combined. In every category of stage, survival at five years was higher for women from more affluent areas, with a clear gradient.

Figure 1. Breast cancer, South Thames, 1980-1989:
Relative survival (%) in women from the most affluent and most deprived enumeration districts, by time since diagnosis

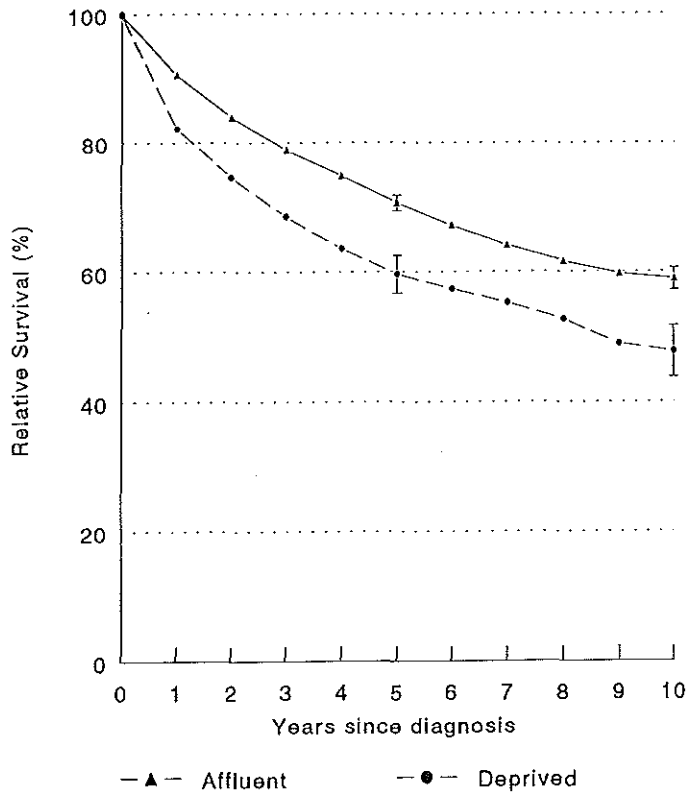


Table 2 Five year relative survival by deprivation category and age group, breast cancer, South Thames, 1980-1989

		Deprivation category				
Age group	Affluent	(2)	(3)	(4)	Deprived	Total
30-64 years						
5-year RSR	73	70	66	65	64	69
95% CI	71-75	68-72	64-68	63-67	61-67	68-70
No. of cases	5609	3495	2912	2234	910	15160
65-99 years						
5-year RSR	67	63	60	62	53	63
95% CI	65-69	61-65	58-62	59-65	49-57	62-64
No. of cases	4488	3652	3195	2302	879	14516

RSR: relative survival rate; CI: confidence interval

Table 3. Stage at diagnosis (%) by age group and deprivation category, breast cancer, South Thames,1980-1989

		Deprivation category				
Stage	Affluent	(2)	(3)	(4)	Deprived	Total
30-64 years						
Local	47.8	48.0	50.5	48.6	48.3	48.5
Regional	23.1	25.4	24.6	25.1	27.7	24.5
Metastasis	7.5	8.0	7.7	9.0	7.5	7.9
Unknown	21.6	18.7	17.2	17.3	16.6	19.1
	100.0	100.0	100.0	100.0	100.0	100.0
65-99 years						
Local	49.2	49.7	51.1	47.6	41.3	49.0
Regional	18.2	18.3	17.8	18.5	18.5	18.2
Metastasis	9.9	10.7	11.0	12.0	17.3	11.1
Unknown	22.7	21.3	20.1	21.9	22.9	21.7
	100.0	100.0	100.0	100.0	100.0	100.0

Table 4. Five year relative survival rate by deprivation category and stage at diagnosis, breast cancer, South Thames, 1980-1989

Stage		Deprivation category: Affluent	(2)	(3)	(4)	Deprived	Total
Local	RSR	84	82	78	80	77	81
	95% CI	83-85	80-84	76-80	78-82	73-81	80-82
	No. of cases	4892	3488	3103	2181	802	14466
Regional	RSR	64	61	58	57	56	60
	95% CI	61-67	58-64	55-61	53-61	51-61	56-64
	No. of cases	2111	1555	1285	986	415	6352
Metastasis	RSR	26	23	21	23	16	23
	95% CI	22-30	19-27	17-25	18-28	10-22	21-25
	No. of cases	865	671	578	477	220	2811
Unknown	RSR	65	57	50	52	49	57
	95% CI	63-67	53-61	46-54	48-56	43-55	56-58
	No. of cases	2229	1433	1141	892	352	6047
Total	RSR	71	67	63	64	60	67
	95% CI	69-73	65-69	62-64	62-66	57-63	66-68
	No. of cases	10097	7147	6107	4536	1789	29676

RSR: relative survival rate; CI: confidence interval

Multivariate analysis was conducted separately for the two age-groups (table 5). Within these broad age categories, analysis of finer sub-divisions of age did not alter the relationship between deprivation and survival. For women aged 30-64 years, there was a clear gradient in the probability of death across deprivation categories, with higher hazard ratios for the more deprived groups (model 1). Addition of period of diagnosis did not change the hazard ratios (model 2). Adjustment for stage at diagnosis altered the hazard ratios for individual deprivation categories only slightly (model 3), while neither morphology nor type of treatment had any substantial influence on the hazard ratios (models 4 and 5). In the final model, including duration of follow-up, period of diagnosis, stage, morphology and type of treatment, the gradient in survival across deprivation categories was still apparent, with a 36% excess hazard of death in the most deprived category.

For women aged 65 years and over, the gradient of hazard ratio by deprivation category was more marked, especially for the most deprived category (hazard ratio 1.69; model 1). Adjustment for stage at diagnosis reduced the gradient (model 3), while adjustment for morphology (model 4) had little effect. Adjusting for the type of treatment (model 5) mainly reduced the hazard in the most deprived group; in this model, including the same variables as for younger women, the socioeconomic

gradient in survival was also still apparent, with a similar 34% excess hazard of death in the most deprived category.

For both age groups and in each model, addition of each prognostic factor significantly improved the fit over that of the preceding model, and the trend in hazard ratio across deprivation categories was statistically significant (2-sided p-value <0.00001 in each case). Finer sub-division of period of diagnosis and follow-up time did not alter the results in either of the age-groups.

Of the 12,911 deaths that occurred within 5 years of breast cancer diagnosis, 960 (7.4%) might have been avoided if all women had experienced the survival of the most affluent category (table 6). There was a higher percentage of potentially avoidable deaths in the more deprived categories: 6.5%, 12.3%, 11.8% and 17.8% in categories 2-5, respectively. The potential reduction in mortality was larger in women aged 30-64 years (506 deaths, 10% of all deaths) than in women aged 65-99 years (454 deaths, 5.8%). Finally, in the age group 50-69 years, the overall potential reduction in mortality at five years was just over 10% (507) of all deaths, reaching 22% (74 deaths) in the most deprived category.

Table 5. Hazard ratios and 95% confidence intervals (CI) by age and deprivation category; adjustment for prognostic factors, breast cancer, South Thames, 1980-1989

	30-64 years				65-99 years			
Deprivation category	Hazard ratio	95% CI	Difference[a] in: deviance	d.f.	Hazard ratio	95% CI	Difference[a] in: deviance	d.f.
Model 1: Deprivation, follow-up period (0-5 and 6-13 years)								
Affluent	1.00		259	5	1.00		295	5
(2)	1.15	1.05-1.27			1.17	1.02-1.33		
(3)	1.30	1.18-1.43			1.24	1.08-1.42		
(4)	1.31	1.18-1.46			1.23	1.06-1.43		
Deprived	1.35	1.17-1.57			1.69	1.41-2.03		
Model 2: Deprivation, follow-up period and period of diagnosis (1980-1984 and 1985-1989)								
Affluent	1.00		5	1	1.00		11	1
(2)	1.15	1.05-1.27			1.16	1.02-1.33		
(3)	1.30	1.18-1.44			1.24	1.09-1.42		
(4)	1.31	1.18-1.46			1.23	1.06-1.43		
Deprived	1.35	1.16-1.57			1.68	1.39-2.03		
Model 3: Deprivation, follow-up period, period of diagnosis and stage at diagnosis								
Affluent	1.00		1874	3	1.00		1995	3
(2)	1.16	1.06-1.26			1.15	1.02-1.29		
(3)	1.34	1.23-1.46			1.23	1.09-1.38		
(4)	1.30	1.18-1.43			1.16	1.01-1.32		
Deprived	1.39	1.22-1.59			1.47	1.25-1.74		
Model 4: Deprivation, follow-up period, period of diagnosis, stage, morphology								
Affluent	1.00		83	2	1.00		101	2
(2)	1.16	1.07-1.27			1.15	1.03-1.29		
(3)	1.35	1.24-1.48			1.23	1.09-1.38		
(4)	1.31	1.19-1.43			1.18	1.03-1.34		
Deprived	1.41	1.24-1.61			1.46	1.24-1.72		
Model 5: Deprivation, follow-up period, period of diagnosis, stage, morphology and treatment								
Affluent	1.00		1071	6	1.00		1073	6
(2)	1.12	1.03-1.21			1.17	1.06-1.28		
(3)	1.32	1.22-1.44			1.23	1.11-1.35		
(4)	1.30	1.19-1.42			1.16	1.04-1.29		
Deprived	1.36	1.21-1.54			1.34	1.17-1.54		

[a] Difference from preceding model. For model 1, difference from model including only constant term; 5 d.f. refer to 4 d.f. for deprivation and 1 d.f. for follow-up period

5.1.4 Discussion

Our results show a gradient in survival for women diagnosed with breast cancer in the South Thames region between 1980 and 1989 according to a measure of material deprivation in the small area of their residence at diagnosis. Survival among women from deprived areas was lower than for women from affluent areas during the entire 13-year follow-up period and at all ages, but the gradient in survival across deprivation categories was steeper for older women (65-99 years). The hazard ratio for the most deprived category was 1.35 for younger women and 1.69 for older women, but after adjustment for calendar period of diagnosis, stage at diagnosis, morphology and type of treatment, the excess hazard was still about 35% for both age groups.

Four methodological issues affect the interpretation of these results. First, the area-based measure of deprivation used here (Carstairs Index) is a proxy measure for the deprivation of individual breast cancer patients at the time of diagnosis, and therefore the gradient in survival by deprivation might be underestimated. However, this measure has been shown to have a stronger association with mortality than social class based on occupation, while there are many problems with measuring social class based on occupation, especially for women.[27]

We used information from the 1981 census to assign a deprivation score to women diagnosed between 1980 and 1989. This could have resulted in misclassification if the socioeconomic characteristics of some enumeration districts changed substantially between 1981 and the time of breast cancer diagnosis for residents of such districts. Such changes cannot be ruled out, but are unlikely to have occurred differentially according to deprivation category, and would be expected to cause under-estimation of any differences in breast cancer survival by deprivation category.

A second potential bias arises from the use of national rather than regional life tables to adjust for expected mortality. All-cause mortality was higher in England and Wales as a whole than in South Thames,[28] so expected survival will be lower (and relative survival higher) than if regional life tables had been used. It seems unlikely, however, that differences between the various deprivation categories in life expectancy calculated nationally or regionally would be so great as to produce substantial bias in the relative survival gradient for breast cancer.

Similarly, use of a single life table for all women may also be criticised, since all-cause mortality varies with social class: this might exaggerate any underlying gradient in relative survival from breast cancer. Separate life tables for social classes or deprivation categories are unavailable, however. There is some evidence that the gradient in relative survival from breast cancer is robust to differences between socio-economic groups in mortality from other causes. The ratio of breast cancer survival in Finland between the highest and lowest social classes was 1.10 with corrected survival rates (censoring deaths from other causes) and 1.12 with relative survival rates.[7]

Table 6. Observed, expected[a] and avoidable[b] deaths at 5 years, by age and deprivation category, breast cancer, South Thames, 1980-1989

Deprivation category	Age group	No. of deaths observed	expected	Avoidable deaths %	No.
Affluent	30-64	1764	1674	-	-
	65-99	2282	2282	-	-
	Total	3956	3956	-	-
	50-69	1486	1486	-	-
(2)	30-64	1150	1058	8.0	92
	65-99	1965	1854	5.6	111
	Total	3115	2912	6.5	203
	50-69	1154	1049	9.1	105
(3)	30-64	1070	884	17.4	186
	65-99	1783	1617	9.3	166
	Total	2853	2501	12.3	352
	50-69	1076	896	16.7	180
(4)	30-64	836	678	18.9	158
	65-99	1263	1174	7.0	89
	Total	2099	1852	11.8	247
	50-69	831	683	17.8	148
Deprived	30-64	345	275	20.3	70
	65-99	543	455	16.2	88
	Total	888	730	17.8	158
	50-69	333	259	22.2	74
Total	30-64	5075	4569	10.0	506
	65-99	7836	7382	5.8	454
	Total	12911	11951	7.4	960
	50-69	4880	4373	10.4	507

[a] from elimination of survival gradient across deprivation categories
[b] difference between observed and expected deaths (see text).

A third methodological issue concerns the exclusion from analysis of DCO cases, for which survival time is unknown. In this study the percentage of such cases was similar (8-9%) in all deprivation categories. We were able to estimate the effect of excluding DCO cases on observed survival (Bullard J, personal communication). As a ratio of the observed (unadjusted) survival at 5 years in the most affluent group, observed survival at 5 years in groups 2 to 5 respectively was 0.92, 0.87, 0.88 and 0.82, respectively. These ratios became 0.91, 0.86, 0.86 and 0.81, after correction for the exclusion of DCO cases, and their exclusion could thus have had very little effect on the gradient in survival reported here.

Fourth, the stage at diagnosis used in these analyses is not identical to the TNM stage. The key advantages are that, unlike TNM stage, it is available for most cases; it is simple; it has been assigned by Registry staff with a standard definition over many years, and, for cases where both stage codes are available, it has almost identical prognostic significance (Lutz J-M, personal communication). It has been argued that the most important explanatory factor for socioeconomic variation in breast cancer survival is a difference in the stage distribution between deprivation categories, and in some studies deprived women have been shown to present at a more advanced stage than affluent women.[7,29,30] No such pattern was observed in Scotland,[13] or for younger women in this study. For older women, differences in the stage distribution did explain part of the variation in survival, the hazard ratio for the most deprived group falling from 1.68 to 1.47 after adjustment for stage. Our results are similar to those from other studies in which survival differences between socioeconomic groups persisted after correction for stage at diagnosis.[7,9,11] Part of the gradient in survival by deprivation could be due to residual confounding by stage. If women from deprived areas were diagnosed less accurately than women from affluent areas, they would be understaged more often, leading to greater misclassification of stage at diagnosis in women from deprived areas. This assumption could not be tested with cancer registry data however.

Our findings suggest that special attention to early detection and rapid referral of breast cancer should be given to women aged 65 or more living in deprived areas. A strength of area-based analyses is that such women could be identified through their area of residence, perhaps for special health education programmes. The other prognostic factors that we studied had little impact (type of treatment) or no impact (morphology) on the gradient in survival by deprivation category.

We conclude that a gradient in breast cancer survival according to deprivation still existed after adjustment for stage at diagnosis, morphology and broad category of treatment. Other factors might be responsible for the observed gradient in breast cancer survival by deprivation category, such as a poorer host resistance among the deprived patients, which could be related to more co-morbidity, an adverse nutritional status, less social support, and negative psychological factors such as less ability to cope with a cancer diagnosis. Aspects of the health care system which might be related to the lower survival of the lower socioeconomic groups are, apart from the type of treatment, adverse hospital referral patterns, the lower quality or appropriateness of treatment and worse compliance with treatment in these groups of patients. For most of these factors, however, information is not available from cancer registry records, and other approaches will be required to study their impact. Preliminary results from a study in our territory of breast cancer patients aged less than 50 suggest that survival was significantly affected by the use of adjuvant therapy.[31] Hospital referral patterns are being examined.

The Health of the Nation target for breast cancer envisages a 25% reduction in breast cancer mortality among women aged 50-69 by the year 2000. Improving the survival of breast cancer patients living in less affluent areas would make a

substantial contribution to this target. In women aged 50-69 years the overall reduction would have been over 10% five years after diagnosis. Our results suggest that one way of achieving this improvement would be to focus on socioeconomic differences in stage at presentation in older women. In younger women, other factors, so far unidentified, are responsible for the socioeconomic gradient in breast cancer survival.

Acknowledgement

CTM Schrijvers was supported by a grant from the Research and Development programme of South West Thames Regional Health Authority

References

1. Department of Health. The health of the nation: a strategy for health in England. London: HMSO, 1992.
2. Coleman MP, Estève J, Damiecki P, Arslan A, Renard H. Trends in cancer incidence and mortality. IARC scientific publications no. 121. Lyon: IARC, 1993.
3. Kelsey JL, Gammon MD, John EM. Reproductive factors and breast cancer. Epidemiol Rev 1993;15:36-47.
4. Chamberlain J, Moss SM, Kirkpatrick AE, Michell M, Johns L. National Health Service breast screening programme results for 1991-2. Br Med J 1993a;307:353-356.
5. Expert advisory group on cancer to the chief medical officers of England and Wales. Consultative document: a policy framework for commissioning cancer services. London: Department of Health, 1994.
6. Vågerö D, Persson G. Cancer survival and social class in Sweden. J Epidemiol Community Health 1987;41:204-209.
7. Karjalainen S, Pukkala E. Social class as a prognostic factor in breast cancer survival. Cancer 1990;66:819-826.
8. Kogevinas M, Marmot MG, Fox AJ, Goldblatt PO. Socioeconomic differences in cancer survival. J Epidemiol Community Health 1991;45:216-219.
9. Dayal HH, Power RN, Chiu C. Race and socio-economic status in survival from breast cancer. J Chronic Dis 1982;35:675-683.
10. Bonett A, Roder D, Esterman A. Determinants of case survival for cancers of the lung, colon, breast and cervix in South Australia. Med J Aust 1984;141:705-709.
11. Bassett MT, Krieger N. Social class and black-white differences in breast cancer survival. Am J Public Health 1986;76:1400-1403.
12. Ansell D, Whitman W, Lipton R, Cooper R. Race, income, and survival from breast cancer at two public hospitals. Cancer 1993;72:2974-2978.
13. Carnon AG, Ssemwogerere A, Lamont DW, Hole DJ, Mallon EA, George WD, Gillis CR. Relation between socioeconomic deprivation and pathological prognostic factors in women with breast cancer. Br Med J 1994;309:1054-1057.
14. Skeet RG. The Thames Cancer Registry. In: Jensen OM, Parkin DM, MacLennan R, Muir CS, Skeet RG (eds). Cancer registration: principles and methods. IARC scientific publications no. 95. Lyon: IARC, 1991: 237-245.
15. Thames Cancer Registry. Cancer in South East England, 1991: cancer incidence, prevalence and survival in residents of the District Health Authorities in South East England. Sutton: Thames Cancer Registry, 1994.
16. Skeet RG, Thornton-Jones H, Murrells TJ. UK, England, South Thames Region, 1978-1982. In: Muir CS, Waterhouse JAH, Mack T, Powell J, Whelan SL (eds). Cancer incidence in five continents, volume V. IARC scientific publications no. 88. Lyon: IARC, 1987: 664-667.
17. Chamberlain J, Bourne HM, Thornton-Jones H. UK, England, South Thames Region, 1983-1987. In: Parkin DM, Muir CS, Whelan SL, Gao Y-T, Ferlay J, Powell J (eds). Cancer incidence in five continents, volume VI. IARC scientific publications no. 120. Lyon: IARC, 1993b: 790-793.
18. Thames Cancer Registry. Cancer in South East Thames, 1987-1989: cancer incidence, prevalence and survival in residents of the District Health Authorities in South East Thames. Sutton: Thames Cancer Registry, 1992a.
19. Thames Cancer Registry. Cancer in South West Thames 1987-1989: cancer incidence,

prevalence and survival in residents of the District Health Authorities in South West Thames. Sutton: Thames Cancer Registry, 1992b.
20. Carstairs V, Morris R. Deprivation and health in Scotland. Aberdeen: Aberdeen University Press, 1991.
21. Chouillet AM, Bell CMJ, Hiscox JG. Management of breast cancer in southeast England. Br Med J 1994;308:168-171.
22. Cutler SJ, Ederer F. Maximum utilisation of the life table method in analysing survival. J Chronic Dis 1958;8:699-712.
23. Hakulinen T, Abeywickrama KH. A computer program package for relative survival analysis. Comp Prog Biomed 1985;19:197-207.
24. Hakulinen T, Tenkanen L. Regression analysis of relative survival rates. Appl Stat 1987;36:-309-317.
25. Payne CD. The GLIM system release 3.77. Generalized linear interactive modelling manual. Oxford: Numerical Algorithms Group, 1985.
26. Department of Health. The health of the nation: specification of national indicators. London: HMSO, 1993.
27. Carstairs V, Morrris R. Deprivation and mortality: an alternative to social class? Comm Med 1989;11:210-219.
28. OPCS. Mortality statistics: area, 1985. Series DH5 no.12. London: HMSO, 1987.
29. Farley TA, Flannery JT. Late-stage diagnosis of breast cancer in women of lower socioeconomic status: public health implications. Am J Public Health 1989;79:1508-1512.
30. Wells BL, Horm JW. Stage at diagnosis in breast cancer: race and socioeconomic factors. Am J Public Health 1992;82:1383-1385.
31. Wolfe CDA, Barton J, Bourne HM, Richards MA. Variation in the incidence and management of primary breast cancer in women under 50 years of age. J Epidemiol Community Health 1993;47:400 (abstract).

5.2 Deprivation, stage at diagnosis and cancer survival*

5.2.1 Introduction

The association between socioeconomic status and cancer mortality has been studied for many years.[1] Attempts to explain the variation in cancer mortality rates between socioeconomic groups have mainly focused on studying socioeconomic variation in cancer incidence and the distribution of cancer risk factors across groups. The variation in cancer survival and its possible determinants, for example, socioeconomic differences in the timing of cancer detection and treatment, have been studied less intensively.[2,3] Nevertheless, the potential for reduction of cancer mortality among the socioeconomically disadvantaged, by improving such factors as early detection and adequate treatment, seems promising, and socioeconomic variation in cancer survival should be monitored systematically. We therefore examined the association between an area-based measure of deprivation and cancer survival in the ten most common cancers in the area covered by the South Thames Regional Health Authority (RHA).

The most recent studies that have dealt with the association between deprivation and survival in a large number of common cancers have not studied the impact of stage of disease at diagnosis on this association.[4,5] We were able to study the impact of stage of disease at diagnosis, an important prognostic factor, which may point to socioeconomic variation in the early detection of cancer.

5.2.2 Patients and methods

Patients

Data for this study came from the records of the Thames Cancer Registry, a population-based cancer registry covering a population of about 14 million people in south-east England. From 1960 to 1984 the registry covered the territory of the South Thames RHA; in 1985 coverage was extended to include the territory of the North Thames RHA. In this analysis, patients diagnosed from 1980 were studied and therefore only patients resident in South Thames RHA were included for study. The methods and data quality indices of the Registry have been described[6] and incidence for the 1980s reported.[7-10]

The records of all patients (men and women) diagnosed between 1980 and 1989 with a malignant tumour in one of the ten most common cancer sites and aged 30 to 99 years at diagnosis were checked (n=192,082). The cancers included were lung, breast, colorectum, bladder, prostate, stomach, pancreas, ovary, uterus and cervix. Two categories of patients were excluded from the analyses: patients with an incomplete or unknown postcode (n=11,495 or 6%), since their census enume-

* Schrijvers CTM, Mackenbach JP, Lutz J-M, Quinn MJ, Coleman MP
In press Int J Cancer

ration district could not be reliably determined (see below). The survival of these patients did not differ substantially from the survival of patients that were included in the analysis. Patients for whom the date of death was known but not the date of diagnosis (death certificate only cases, DCO) were also excluded from analyses, as their survival time could not be calculated (n=24,905 or 13%). A total of 155,682 (81% of the original data set) were included in the survival analyses: 73,444 men and 82,238 women (table 1).

Deprivation score

The measure of deprivation for each patient was based on the address at time of diagnosis, by linking the full postcode of residence to the corresponding census enumeration district. Data from the 1981 census on four indicators of material deprivation were obtained for each enumeration district (average of 400 households in Great Britain): overcrowding (proportion of persons in private households living at a density of more than one person per room), male unemployment (proportion of economically active males seeking work), low social class (proportion of all persons in private households with head of household in social class IV or V) and car ownership (proportion of all persons in private households without a car).

The Carstairs Index combines standardised scores on these four variables into a single score for each census enumeration district,[11] using the mean value and standard deviation for Great Britain as a standard.

Each of the 14,386 enumeration districts in South Thames RHA was assigned to one of five deprivation categories, which were constructed by ranking the Carstairs scores for all enumeration districts in Great Britain from low ("affluent") to high ("deprived"), and by forming quintiles based on the underlying population distribution.

Table 1. Number (percentage) of patients by cancer and deprivation category, 10 most common cancers, men and women, South Thames RHA, 1980-1989

	Deprivation category					
Cancer	Affluent (1)	(2)	(3)	(4)	Deprived (5)	Total
Lung	10088 (25.0)	9160 (22.7)	8989 (22.3)	8223 (20.4)	3819 (9.5)	40279
Breast	10097 (34.0)	7147 (24.1)	6107 (20.6)	4536 (15.3)	1789 (6.0)	29676
Colorectum	8530 (30.7)	6901 (24.8)	6002 (21.6)	4662 (16.8)	1701 (6.1)	27796
Bladder	3896 (30.1)	3065 (23.7)	2809 (21.7)	2282 (17.6)	905 (7.0)	12957
Prostate	4155 (33.2)	3278 (26.2)	2524 (20.1)	1852 (14.8)	723 (5.8)	12532
Stomach	2828 (25.9)	2560 (23.4)	2414 (22.1)	2200 (20.1)	929 (8.5)	10931
Pancreas	1979 (29.2)	1666 (24.6)	1438 (21.2)	1203 (17.8)	486 (7.2)	6772
Ovary	1994 (33.2)	1491 (24.8)	1216 (20.3)	916 (15.3)	382 (6.4)	5999
Uterus	1537 (30.9)	1261 (25.4)	1064 (21.4)	775 (15.6)	331 (6.7)	4968
Cervix	935 (24.8)	805 (21.3)	856 (22.7)	763 (20.2)	413 (10.9)	3772
Total	46039 (29.6)	37334 (24.0)	33419 (21.5)	27412 (17.6)	11478 (7.4)	155682

Prognostic factors

Age was studied in two or three categories, depending on the age-distribution of cases for each cancer: lung, bladder and stomach (30-64, 65-74 and 75-99), breast, ovary, uterus and pancreas (30-64 and 65-99), colorectum and prostate (30-74 and 75-99), and cervix (30-44, 45-64 and 65-99). The results of analyses in which age was studied in much smaller categories did not differ from those presented in this paper. Data for men and women were combined for cancers of the lung, colorectum, stomach and pancreas, as both overall survival and the gradient in survival by deprivation were very similar. Gender was included as a possible confounder for these cancers. Survival from bladder cancer was clearly higher for men than for women, and therefore we will also discuss the results for this cancer for men and women separately. Period of diagnosis was included in the analysis in two five year periods: 1980 to 1984 and 1985 to 1989. Stage at diagnosis (clinical or pathological) was explicitly stated in 20% of the medical records for all cancers combined. A simplified stage is routinely constructed by Registry staff for all cases, however, using pathology reports, operation notes and other information. Stage was originally categorised in three groups: local (tumour confined to the organ of origin); regional (involvement of regional lymph nodes) and metastasis (spread to distant organs). Patients for whom the stage at diagnosis was unknown were also included in the analysis as a fourth category. For most of the cancers, the percentage of patients diagnosed with a regional disease was rather low, and therefore the four categories were distinguished only for cancers of the lung, breast, colorectum and stomach. For the other cancers, the distinction was local, non-local (regional and metastasis combined), and unknown. Data for cancers of the colon and rectum were combined and adjustment was made for subsite in five categories: (1) rectum, (2) sigmoid, (3) ascending colon, (4) transverse and descending colon, and (5) other subsites. Furthermore, subsites were distinguished for stomach cancer: (1) cardia, (2) pylorus, (3) stomach excluding cardia and pylorus.

Survival analysis

Cases diagnosed between 1980 and 1989 were followed up until the date of death or 31 December 1992, whichever occurred first.

To adjust for mortality from causes other than the cancer under study, we used the relative survival rate as measure of outcome in the univariate analyses. The relative survival rate, expressed as a percentage (RSR%) is the ratio of observed survival in a group of cancer patients to the survival that would be expected if they were subject to the same overall mortality rates by age, gender, and calendar period as the general population.[12] The England and Wales life table for 1981 was used to calculate expected survival. The RSR and its 95% confidence interval (CI) were calculated by a computer program from the Finnish Cancer Registry.[13]

In the multivariate analyses, the measure of outcome was the hazard ratio which expresses the probability of death from the cancer under study for a specific category of patients relative to a reference category (which has a hazard ratio of

unity). These analyses were conducted with a proportional hazards model adapted to the RSR[14] using GLIM.[15]
We started with a basic model which included duration of follow-up in two categories (up to 5 years and 6 to 13 years) and deprivation category, and then added the other variables (age, gender, period of diagnosis, stage at diagnosis and subsite for colorectal and stomach cancer). The improvement in fit due to each variable was tested for statistical significance at the 5% level with the Chi-square test. The statistical significance of the trend in the hazard ratio across deprivation categories was tested by examining the effect of adding deprivation category to the model as a continuous variable (one degree of freedom).

5.2.3 Results

About 30% of all patients lived in areas categorised as the most affluent quintile of the Carstairs Index, while only 7.4% lived in areas categorised as the most deprived quintile. This distribution reflects the relative affluence of South Thames within Great Britain. For a few cancers, the percentage of patients in the two most affluent groups is higher than the percentage for all cancers combined: these are breast, prostate and ovary and to a lesser extent colorectum and uterus. For cancers of the lung, stomach and cervix, the percentage of patients in the two most deprived categories is higher than the percentage for all cancers combined. For cancers of the bladder and pancreas, there was a similar distribution across deprivation categories as for all cancers combined. This variation in the distribution of patients across deprivation categories per cancer as compared with all cancers combined probably reflects variation in incidence by deprivation category (table 1).

Relative survival five years after diagnosis was better for patients from affluent areas than for patients from deprived areas for cancers of the lung, breast, colorectum, bladder, prostate, uterus and cervix. For these cancers we observed a gradient in survival by deprivation, which was interrupted in the second lowest category of deprivation for cancers of the lung, breast, prostate, uterus and cervix. The gradient in survival by deprivation for bladder cancer was present only in men (5-year RSR%: affluent (1) 69, (2) 67, (3) 66, (4) 63, deprived (5) 62), but not in women. For cancers of the stomach, pancreas and ovary, no clear difference in 5-year RSR% by deprivation category was observed (table 2).

Table 2. Five year RSR and 95% confidence interval by deprivation category and cancer, 10 most common cancers, South Thames RHA, 1980-1989

	Deprivation category				
Cancer	Affluent	2	3	4	Deprived
Lung	8.0 (7.3-8.7)*	7.3 (6.6-8.0)	7.2 (6.5-7.9)	6.3 (5.7-6.9)	6.5 (5.6-7.4)
Breast	71 (69-73)	67 (65-69)	63 (62-64)	64 (62-66)	60 (57-63)
Colorectum	40 (39-41)	41 (39-43)	39 (37-41)	36 (34-38)	36 (33-39)
Bladder	66 (64-68)	65 (63-67)	63 (61-65)	62 (59-65)	61 (56-66)
Prostate	52 (49-55)	48 (45-51)	45 (42-48)	40 (37-43)	42 (36-48)
Stomach	11 (10-12)	12 (11-13)	10 (9-11)	12 (10-14)	12 (9-15)
Pancreas	3.8 (2.7-4.9)	3.3 (2.2-4.4)	4.0 (2.8-5.2)	3.6 (2.3-4.9)	4.5 (2.2-6.8)
Ovary	30 (28-32)	27 (24-30)	27 (24-30)	30 (26-34)	27 (22-32)
Uterus	76 (73-79)	71 (68-74)	71 (67-75)	66 (61-71)	67 (60-74)
Cervix	62 (59-62)	57 (53-61)	55 (51-59)	58 (54-62)	54 (49-59)

* 95% confidence interval

For each of the cancers we saw a similar gradient in 10-year RSR% by deprivation category as for the 5-year RSR% (results not shown). For cancers with an overall 5-year RSR% below 20% (pancreas, lung, stomach), we also examined the survival gradient by deprivation, 1 and 2 years after diagnosis. For cancer of the pancreas, we observed a higher 1-year RSR% in affluent patients than in deprived patients, but two years after diagnosis the gradient had disappeared. For lung and stomach cancer, the results for survival one and two years after diagnosis were similar to those for survival five and ten years after diagnosis (results not shown).

The stage distribution by cancer and deprivation category showed no systematic pattern for most cancers. Only for breast cancer, and to a lesser degree for cancer of the prostate, we found a higher percentage of patients with non-local disease in the more deprived patient groups (table 3).

Table 3. Percentage of patients with non-local stage at diagnosis by cancer and deprivation category, 10 most common cancers, South Thames RHA, 1980-1989

	Deprivation category					
Cancer	Affluent	2	3	4	Deprived	Total
Lung	30.3	30.9	29.1	28.9	29.3	29.8
Breast	29.5	31.2	30.5	32.2	35.5	30.9
Colorectum	32.6	33.2	32.6	31.9	32.2	32.6
Bladder	4.6	4.7	5.1	5.1	4.8	4.8
Prostate	26.8	26.3	28.3	27.8	31.4	27.4
Stomach	37.3	36.7	36.4	36.5	35.3	36.6
Pancreas	31.4	33.6	34.9	30.2	29.8	32.3
Ovary	36.7	39.4	43.2	39.1	37.4	39.1
Uterus	9.6	10.5	9.1	11.4	6.6	9.8
Cervix	10.1	11.6	12.4	10.2	13.6	11.3

Table 4 shows the results from the multivariate analyses, by cancer and deprivation category. The hazard ratios presented in this table were adjusted for follow-up period, age, gender, period of diagnosis and subsite in colorectal and stomach cancer. These hazard ratios combine results for the entire period of follow-up, so for example the hazard ratio of 1.13 for the most deprived category of lung cancer patients means that during the entire period of follow-up, the annual excess probability of dying was 13% in the most deprived category as compared with the most affluent category.

The results from the multivariate analyses are in agreement with the results from the univariate analyses: the hazard ratios were higher in the more deprived categories for cancers of the lung, breast, colorectum, bladder, prostate, uterus and cervix. The trend in hazard ratios by deprivation was statistically significant for these cancers. For cancers of the stomach, pancreas and ovary, no clear gradient in hazard ratios by deprivation category was observed, and the trend in hazard ratios by deprivation was not statistically significant.

We tested the improvement in fit of the preceding model resulting from the addition of stage, which was statistically significant for each of the cancers. However, the addition of stage caused no large changes in the hazards for the five deprivation categories in most cancers (table 5). The changes in hazard ratios for deprivation were largest for cancers of the uterus and cervix, especially in the most deprived patient group. For cervical cancer, the hazard ratio for the most deprived patient group changed from 1.35 to 1.27 and for cancer of the uterus from 1.46 to 1.59.

Table 4. Hazard Ratio (HR) and 95% confidence interval (CI) by cancer and deprivation category, 10 most common cancers, South Thames RHA, 1980-1989[1]

	Deprivation category					
Cancer	Affluent	2	3	4	Deprived	Slope
	HR	HR (95% CI)	HR (95% CI)	HR (95% CI)	HR (95% CI)	HR (95% CI)
Lung	1.00	1.04 (0.96-1.14)	1.09 (1.00-1.19)	1.12 (1.03-1.22)	1.13 (1.01-1.26)	1.03 (1.01-1.06)
Breast	1.00	1.15 (1.02-1.30)	1.27 (1.12-1.44)	1.27 (1.11-1.46)	1.47 (1.22-1.76)	1.09 (1.06-1.13)
Colorectum[2]	1.00	0.99 (0.92-1.08)	1.04 (0.96-1.13)	1.14 (1.05-1.25)	1.14 (1.01-1.29)	1.04 (1.02-1.06)
Bladder	1.00	1.02 (0.86-1.21)	1.10 (0.93-1.31)	1.24 (1.04-1.48)	1.24 (0.97-1.58)	1.07 (1.02-1.12)
Prostate	1.00	1.10 (0.94-1.30)	1.18 (1.00-1.40)	1.34 (1.12-1.60)	1.37 (1.07-1.76)	1.09 (1.04-1.14)
Stomach[2]	1.00	0.98 (0.87-1.11)	1.05 (0.93-1.19)	1.03 (0.91-1.17)	1.11 (0.93-1.33)	1.02 (0.99-1.06)
Pancreas	1.00	1.01 (0.88-1.15)	0.96 (0.83-1.11)	1.07 (0.92-1.25)	1.04 (0.84-1.28)	1.01 (0.97-1.05)
Ovary	1.00	1.10 (0.91-1.34)	1.13 (0.93-1.39)	1.05 (0.84-1.31)	1.10 (0.81-1.50)	1.02 (0.97-1.08)
Uterus	1.00	1.21 (0.90-1.61)	1.18 (0.87-1.60)	1.48 (1.09-2.02)	1.46 (0.97-2.20)	1.11 (1.02-1.20)
Cervix	1.00	1.15 (0.94-1.40)	1.29 (1.07-1.57)	1.17 (0.96-1.44)	1.35 (1.07-1.71)	1.06 (1.01-1.12)

[1] Results from models with follow-up period, deprivation category, age, (sex), and period of diagnosis [2] Also adjusted for subsite

5.2.4 Discussion

Our results show that patients from deprived areas had worse survival than those from affluent areas for cancers of the lung, breast, colorectum, bladder, prostate, uterus and cervix, but not for cancers of the stomach, pancreas or ovary. The excess hazard of death for patients from the most deprived category ranged from 11% for colorectal and lung cancer to 59% for cancer of the uterus (adjusted for age, gender, period of diagnosis and stage at diagnosis). This shows an enormous potential for a reduction of cancer mortality by improving the survival rates of patients from deprived areas. Stage of disease at diagnosis was not an important explanatory factor of the association between deprivation and survival.

We considered a number of methodological aspects that might have influenced our results. First, we used an area-based measure of deprivation (Carstairs Index),

which is a proxy measure of the deprivation of individual cancer patients. We did not use social class based on occupation as measure of deprivation, since this was incomplete or missing for a large proportion of patients, especially for women. Furthermore, deprivation was found to be more strongly associated with mortality than social class based on occupation (Carstairs & Morris, 1989). On the other hand, using an area-based measure could result in misclassification and therefore in underestimation of the gradient in survival by deprivation. For most cancers our results were similar to those from another English study in which an individual measure of deprivation was used.[5] Better survival for patients with a low socioeconomic status was found in this study for cancers of the prostate, breast, and stomach and rectum in females. If the association between deprivation and survival for these cancers would have been underestimated in our study as a result of misclassification, the findings from both studies on these cancers would be even more divergent.

Table 5. Hazard Ratio (HR) and 95% confidence interval (CI) by cancer and deprivation category, 10 most common cancers, South Thames RHA, 1980-1989[1]

	Deprivation category					
Cancer	Affluent	2	3	4	Deprived	Slope
	HR	HR (95% CI)	HR (95% CI)	HR (95% CI)	HR (95% CI)	HR (95% CI)
Lung	1.00	1.04 (0.96-1.12)	1.09 (1.01-1.18)	1.13 (1.04-1.22)	1.11 (1.00-1.23)	1.03 (1.01-1.06)
Breast	1.00	1.15 (1.06-1.25)	1.29 (1.19-1.41)	1.24 (1.13-1.36)	1.43 (1.27-1.62)	1.09 (1.06-1.11)
Colorectum[2]	1.00	1.02 (0.95-1.09)	1.03 (0.96-1.11)	1.16 (1.07-1.25)	1.11 (1.00-1.24)	1.04 (1.02-1.06)
Bladder	1.00	1.05 (0.91-1.22)	1.07 (0.93-1.25)	1.22 (1.05-1.42)	1.22 (0.99-1.51)	1.06 (1.02-1.10)
Prostate	1.00	1.11 (0.99-1.25)	1.21 (1.07-1.36)	1.37 (1.20-1.56)	1.34 (1.12-1.60)	1.09 (1.06-1.13)
Stomach[2]	1.00	0.98 (0.87-1.10)	1.04 (0.93-1.18)	1.02 (0.90-1.15)	1.11 (0.93-1.32)	1.02 (0.99-1.05)
Pancreas	1.00	1.02 (0.89-1.16)	0.95 (0.83-1.09)	1.08 (0.93-1.26)	1.04 (0.85-1.28)	1.01 (0.97-1.05)
Ovary	1.00	1.12 (0.97-1.30)	1.09 (0.94-1.28)	1.08 (0.91-1.29)	1.12 (0.88-1.43)	1.02 (0.98-1.07)
Uterus	1.00	1.26 (0.99-1.61)	1.24 (0.96-1.59)	1.45 (1.11-1.88)	1.59 (1.12-2.24)	1.11 (1.04-1.19)
Cervix	1.00	1.12 (0.93-1.34)	1.26 (1.05-1.51)	1.14 (0.94-1.37)	1.27 (1.02-1.58)	1.05 (1.00-1.10)

[1] Results from models with follow-up period, deprivation category, age, (sex), and period of diagnosis [2] Also adjusted for subsite

Second, DCO cases were excluded from the analyses. However, the percentage of such cases varied systematically only across deprivation categories for lung cancer, with a higher DCO% in more deprived categories (table 6). We were able to estimate the effect of DCO cases on observed lung cancer survival, as the date of diagnosis of as many DCO cases as possible diagnosed during 1986-1987 had been obtained through the Family Health Service Authorities. The survival of these cases could therefore be calculated and the ratio of survival reduction (%) to DCO (%) was used to estimate the impact of DCO bias on our results (Bullard J, personal communication). The ratio of the observed 5-year survival rate in the most affluent group and the observed 5-year survival rate in groups 2 to 5 of lung cancer patients respectively was 0.91, 0.90, 0.80 and 0.85. These ratios became 0.91, 0.89, 0.78 and 0.81, after correction for the exclusion of DCO cases, suggesting a small under-estimation of the deprivation by survival gradient in lung cancer after the exclusion of DCO cases.

Another possible source of bias is the use of a single life table, due to the absence of social- class-specific life tables, to calculate expected survival in order to adjust for mortality due to causes other than the cancer under study. This bias could have resulted in overestimation of the gradient in relative survival by deprivation. In general, overall mortality is higher in deprived areas, and therefore the expected survival of patients from more deprived groups may be overestimated if a life table from the general population is used. This may have resulted in underestimation of the relative survival for the more deprived groups. For the affluent patient groups, expected survival may be underestimated and relative survival overestimated. There is some evidence, from a Finnish study on social class and breast cancer survival, that this overestimation of the deprivation-survival gradient is rather small. The ratio of survival between the highest and lowest social class was 1.10 with corrected survival rates (censoring deaths from other causes) and 1.12 with relative survival rates.[17]

Furthermore, we considered the possibility that bias of the results arose from the use of a national rather than a regional life table to adjust for expected mortality. Overall mortality was lower in South Thames than in England and Wales as a whole,[18] so expected survival was lower and relative survival higher than if a regional life table had been used. It is unlikely, however, that differences in life expectancy calculated nationally or regionally would be systematic by deprivation category and therefore result in a bias of the relative survival gradient by deprivation.

Finally, the use of a life table based on a single year (1981) might have caused a bias of the results, but this concerns overall survival. It seems unlikely that the association between deprivation and survival was biased, since the distribution of cases across years of diagnosis did not vary between deprivation categories.

Table 6. DCO cases (%) by cancer and deprivation category, 10 most common cancers, South Thames RHA, 1980-1989

	Deprivation category					
Cancer	Affluent	2	3	4	Deprived	Total
Lung	15.8	16.2	17.3	18.3	19.9	17.2
Breast	7.9	8.9	9.0	9.4	9.4	9.4
Colorectum	14.1	13.3	13.3	12.9	15.9	13.6
Bladder	6.8	6.5	5.7	7.8	7.4	6.7
Prostate	12.2	12.7	13.5	12.5	13.5	12.7
Stomach	19.9	17.6	17.8	18.9	21.8	18.9
Pancreas	22.5	20.1	20.8	21.1	20.3	21.1
Ovary	11.4	11.9	13.4	13.2	12.8	12.3
Uterus	5.9	6.2	5.6	5.7	8.3	6.0
Cervix	5.9	5.7	5.5	7.6	4.8	6.0

The stage variable which we used in this analysis is not identical to the TNM stage. However, the resulting reduction in deviance was substantial in our analyses, suggesting reasonable validity of the stage variable. This was confirmed when the prognostic significance of this stage variable and the TNM stage were found to be very similar in breast cancer cases for which both stage codes were available (Lutz J-M, personal communication).

We further considered the possibility that lead-time bias might explain part of the survival with deprivation gradient in this study. The analyses in this paper included an adjustment for stage of disease at diagnosis. As the period between origin of the tumour and diagnosis is associated with stage, adjustment was made, indirectly, for lead-time. Of course, residual confounding due to stage, and therefore lead-time, may still explain part of the association between deprivation and survival. However, if lead-time bias were the explanation, we would expect a larger difference in survival by deprivation in the first period of follow-up than in the later period, and this was clearly not the case in our study. The results from survival analyses of patients with metastatic disease showed better survival for patients from the most affluent group as compared to the most deprived group for six out of the seven cancers for which we had observed better survival in the affluent patients; the only exception was lung cancer. However, the gradient in survival by deprivation was not consistent for any of these cancers and 95% CIs were rather broad for most categories.

Only for cancers of the cervix and uterus did we find a substantial change in the hazard ratios for the five deprivation categories resulting from the addition of stage. Although stage is a very important prognostic factor in general, it cannot explain the gradient in survival by deprivation. This was also found in other

studies, in which socioeconomic variation in survival from either breast cancer[17,19,20], colorectal cancer,[21,22] colon cancer[23] or cancer of the prostate[24] persisted after adjustment for stage. Similar results were found in older studies for many cancer sites.[25-27]

Another possible explanation for the gradient in survival by deprivation is a difference in the management of cancer patients from various deprivation categories. It could be argued that this factor may be more important in cancers with an overall medium or good prognosis, and that for such cancers socioeconomic variation in survival would be larger. We calculated the ratios of survival rates of the most affluent and the most deprived patient group for each cancer and ranked these according to the overall survival of the cancers. From these results we observed no clear gradient, which is in agreement with other results.[2] We did find, however, that the variation in survival by deprivation was absent or rather small in cancers with relatively poor or moderate overall survival (pancreas, lung, stomach, ovary, colorectum), while variation was larger in cancers with a fairly high overall survival (prostate, bladder, cervix, breast, uterus).

Our results show that the gap in survival between cancer patients from affluent and deprived areas is large, both in absolute and relative terms. We found no evidence for an explanation of the survival gradient by deprivation in terms of large variation in stage of disease at diagnosis. Other determinants of socioeconomic variation in early detection of cancer by deprivation should be studied, such as the biological aggressiveness of a tumour, and host factors which may interact with the tumour. Examples of such factors are co-morbidity, psychological factors, nutritional status, social support, and immune response. In addition, determinants of treatment should be studied, such as adherence to guidelines and type of hospital.

Acknowledgement

CTM Schrijvers was supported by a grant from the Research and Development programme of South West Thames Regional Health Authority

References

1. Logan WPD. Cancer mortality by occupation and social class 1851-1971. IARC scientific publications no. 36, Studies on medical and population subjects no. 44. London: HMSO, Lyon: IARC, 1982.
2. Kogevinas, E. Longitudinal Study. Socio-demographic differences in cancer survival. OPCS series LS no 5. London: HMSO, 1990.
3. Schrijvers CTM, Mackenbach JP. Cancer patient survival by socioeconomic status in seven countries: a review for six common cancer sites. J Epidemiol Community Health 1994;48-:441-446.
4. Vågerö D, Persson G. Cancer survival and social class in Sweden. J Epidemiol Community Health 1987;41:204-209.
5. Kogevinas M, Marmot MG, Fox AJ, Goldblatt PO. Socio-economic differences in cancer survival. J Epidemiol Community Health 1991;45:216-219.
6. Skeet RG. The Thames Cancer Registry. In: Jensen OM, Parkin DM, MacLennan R, Muir CS, Skeet RG (eds). Cancer registration: principles and methods. IARC scientific publications

no. 95. Lyon: IARC, 1991: 237-245.
7. Skeet RG, Thornton-Jones H, Murrelss, TJ. UK, England, South Thames Regions, 1978-1982. In: Muir CS, Waterhouse JAH, Mack T, Powell J, Whelan SL (eds). Cancer incidence in five continents, volume V. IARC scientific publications no 88. Lyon:IARC, 1987: 664-667.
8. Chamberlain J, Bourne HM, Thornton-Jones H. UK, England, South Thames Regions, 1983-1987. In: Parkin DM, Muir CS, Whelan SL, Gao Y-T, Powell J (eds). Cancer incidence in five continents, volume VI. IARC scientific publications no 120. Lyon: IARC, 1993: 790-793.
9. Thames cancer registry. Cancer in South East Thames, 1987-1989: cancer incidence, prevalence and survival in residents of the District Health Authorities in South East Thames. Sutton: Thames Cancer Registry, 1992a.
10. Thames cancer registry. Cancer in South West Thames, 1987-1989: cancer incidence, prevalence and survival in residents of the District Health Authorities in South East Thames. Sutton: Thames Cancer Registry, 1992b.
11. Carstairs V, Morris R. Deprivation and health in Scotland. Aberdeen: Aberdeen University Press, 1991.
12. Cutler SJ, Ederer F. Maximum utilization of the life table method in analyzing survival. J Chronic Dis 1958;8:699-712.
13. Hakulinen T, Abeywickrama KH. A computer program package for relative survival analysis. Comp Prog in Biomed 1985;19:197-207.
14. Hakulinen T, Tenkanen L. Regression analysis of relative survival rates. Appl Stat 1987;36:-309-317.
15. Payne CD. The GLIM system release 3.77. Generalized linear interactive modelling manual. Oxford: Numerical Algorithms Group, 1985.
16. Carstairs V, Morris R. Deprivation and mortality: an alternative to social class? Comm Med 1989;11: 210-219.
17. Karjalainen S, Pukkala E. Social class as a prognostic factor in breast cancer survival. Cancer 1990;66:819-826.
18. OPCS. Morality statistics: area, 1985. Series DH5 no. 12. London: HMSO, 1987.
19. Bassett MT, Krieger N. Social class and black-white differences in breast cancer survival. Am. J Public. Health 1986;76:1400-1403.
20. Gordon NH, Crowe JP, Brumberg DJ, Berger NA. Socioeconomic factors and race in breast cancer recurence and survival. Am J Epidemiol 1992;135:609-618.
21. Wegner EL, Kolonel LN, Nomura AMY, Lee J. Racial and socioeconomic status differences in survival of colorectal cancer patients in Hawaii. Cancer 1982;49:2208-2216.
22. Brenner H, Mielck A, Klein R, Ziegler H. The role of socioeconomic factors in the survival of patients with colorectal cancer in Saarland/Germany. J Clin Epidemiol 1991;44:807-815.
23. Auvinen A. Social class and colon cancer survival in Finland. Cancer 1992;70:402-409.
24. Dayal HH, Polissar L, Dahlberg S. Race, socioeconomic status, and other prognostic factors for survival from prostate cancer. J Natl Cancer Inst 1985;74:1001-1006.
25. Linden G. The influence of social class in the survival of cancer patients. Am J Public Health 1969;59:267-274
26. Lipworth L, Abelin T, Connelly RR. Socioeconomic factors in the prognosis of cancer patients. J Chron Dis 1970;23:105-116.
27. Berg JW, Ross R, Latourette HB. Economic status and survival of cancer patients.Cancer 1977;39:467-477.

Chapter 6. Socioeconomic variation in cancer survival in the Southeastern Netherlands and the South Thames area: a comparison

6.1 Introduction

The results of studies on socioeconomic variation in cancer survival were described for the Southeastern Netherlands in chapter 4 and for the area covered by the South Thames Regional Health Authority (RHA) in chapter 5. A comparison of socioeconomic variation in cancer survival for the two areas is now presented, evaluating both the pattern of socioeconomic variation in survival and the steepness of the socioeconomic gradient in survival. Furthermore, the impact of stage of disease at diagnosis on the association between socioeconomic status and survival in the two areas was compared. The histological type of the tumour and treatment were not included in this analysis as it was shown in earlier chapters that their distribution did not vary systematically between socioeconomic groups.

6.2 Patients and methods

The data concerned patients diagnosed between 1980 and 1989 in either the Southeastern Netherlands or the South Thames area with a malignant tumour of the lung, breast, or colorectum, the 3 most common cancers in both areas with a large enough number of patients to reliably determine socioeconomic variation in survival for patients from the Southeastern Netherlands. Only patients aged 30 to 99 years were included. Therefore the number of patients for the Southeastern Netherlands is lower in this analysis than in chapter 4, in which no lower limit for age was set. The end of follow-up was set at July 1, 1991 while it was December 31, 1992 in chapter 5 for the South Thames area. The results presented in this chapter may therefore differ slightly from those in chapter 5.

The measures of socioeconomic status were the same as in the earlier chapters (see also chapter 3). For the Southeastern Netherlands, the measure of socioeconomic status was based on the postcode of residence at time of diagnosis, through which each patient was assigned to one of 45 categories of a sociodemographic classification. The 45 categories were collapsed into 5 socioeconomic categories, based on quintiles of the underlying population and using education as main indicator of socioeconomic status. Table 1 contains the distribution of both the entire Dutch population and the population of the Southeastern Netherlands across the 5 socioeconomic categories. The distribution of cases across socioeconomic groups for different cancers probably reflects socioeconomic variation in cancer incidence: a higher lung cancer incidence in the lower socioeconomic groups and a higher incidence of cancers of the breast and colorectum in the higher socioeconomic groups.

Table 1. Distribution of population and cancer cases across socioeconomic status (SES) groups, Southeastern Netherlands, cancers of the lung, breast and colorectum, 1980-1989

	Population			Cancer cases					
SES	Nether-lands	South-eastern Nether-lands	3 most common cancers	Lung		Breast		Colorectum	
High	19.5%	20.1%	18.2%	716	15.6%	790	20.2%	679	19.3%
(2)	18.5%	14.4%	10.2%	418	9.1%	426	10.9%	371	10.6%
(3)	19.7%	21.4%	20.4%	943	20.6%	811	20.8%	703	20.0%
(4)	21.2%	22.4%	25.0%	1172	25.6%	981	25.1%	854	24.3%
Low	21.1%	21.7%	26.2%	1336	29.1%	899	29.1%	906	25.8%
Total	100%	100%	100%	4585	100%	3907	100%	3513	100%

Table 2. Distribution of population and cancer cases across deprivation categories, South Thames, cancers of the lung, breast and colorectum, 1980-1989

	Population			Cancer cases					
Depri-vation category	Great Britain	South Thames	3 most common cancers	Lung		Breast		Colorectum	
Affluent	20%	32.9%	29.5%	10088	25.0%	10097	34.0%	8530	30.7%
(2)	20%	22.6%	24.0%	9160	22.7%	7147	24.1%	6901	24.8%
(3)	20%	19.2%	21.5%	8989	22.3%	6107	20.6%	6002	21.6%
(4)	20%	16.4%	17.6%	8223	20.4%	4536	15.3%	4662	16.8%
Deprived	20%	8.9%	7.4%	3819	9.5%	1789	6.0%	1701	6.1%
Total	100%	100%	100%	40279	100%	29676	100%	27796	100%

For the South Thames area, we used the Carstairs Index as measure of deprivation, which was assigned to each patient through the census enumeration district of residence at the time of diagnosis. The five categories of the area-based measure were originally constructed using the population distribution of Great Britain, resulting in 5 equally sized deprivation categories. When all the enumeration districts in the South Thames area were assigned to one of these 5 categories, the distribution across deprivation categories became more skewed (table 2). This is due to the fact that the South Thames area is a relatively affluent area within Great Britain, and therefore the number of enumeration districts assigned to the affluent

categories was much higher than the number assigned to the deprived categories. The distribution of cancer cases across deprivation categories again probably reflects variation in incidence per deprivation category: a higher lung cancer incidence in the deprived categories and a higher incidence of breast cancer and colorectal cancer in the affluent categories.

The following prognostic factors were studied: age (breast cancer: 30-64 and 65-99; lung cancer: 30-64, 65-74 and 75-99; colorectal cancer: 30-74 and 75-99), sex, period of diagnosis (1980-1984 and 1985-1989), and stage of disease at diagnosis in the same categories as in earlier chapters (local, regional, distant and unknown). Subsite was also included in the analysis of colorectal cancer, in 5 categories: rectum, sigmoid, ascending colon, transverse and descending colon, and other subsites.

In data analyses, the same techniques were used as in earlier chapters, resulting in the same measures of outcome as in the studies reported in chapters 4 and 5. The univariate analyses were conducted with the computer program for cancer survival studies from the Finnish Cancer Registry[1]. The measure of outcome in the univariate analyses was the relative survival rate (RSR)[2] and in the multivariate analyses, it was the hazard ratio.[3] In the multivariate analyses we started with a basic model including duration of follow-up in two periods (up to 5 and 6-12 years) and socioeconomic status (5 categories), to which we added the possible confounders age (2 or 3 categories), sex, and period of diagnosis (2 categories) (model 1). We then added stage of disease at diagnosis to these models (model 2). Addition of each factor was evaluated by testing the change in deviance for statistical significance, in relation to the corresponding change in degrees of freedom for this factor.

6.3 Results

6.3.1 Lung cancer

Table 3 shows the 5-year RSR by socioeconomic group and area of residence for lung cancer. Overall 5 year survival was higher in the Southeastern Netherlands than in the South Thames area. The gradient in survival by socioeconomic status was similar for both areas, but the ratio of 5-year RSR for the highest and lowest socioeconomic group was higher in the Southeastern Netherlands (1.36) than in the South Thames area (1.26).

The distribution of stage of disease at diagnosis across socioeconomic groups in both areas is given in table 4. For all patients combined, so regardless of SES, we observed a much higher percentage of patients registered with local disease in the South Thames area as opposed to the Southeastern Netherlands. This seemed to be at odds at first sight with the lower overall lung cancer survival in the South Thames area. Consequently, the percentage of lung cancer patients registered with regional or metastatic disease was higher in the Southeastern Netherlands.

Table 3. Five year relative survival rate by socioeconomic status, lung cancer, Southeastern Netherlands and South Thames, 1980-1989

Socioeconomic Status	Southeastern Netherlands	South Thames
High	15 (12-18)*	7.8 (7.2-8.4)
(2)	17 (12-21)	7.3 (6.7-7.9)
(3)	14 (11-17)	7.1 (6.4-7.8)
(4)	12 (9-15)	6.3 (5.6-7.0)
Low	11 (9-13)	6.2 (5.3-7.1)
Total	13 (12-14)	7.1 (6.8-7.4)
Ratio High/Low	1.36	1.26

* 95% confidence interval between brackets

The stage distribution did not vary systematically between socioeconomic groups in either of the areas. The observed differences in stage distribution between the two areas could either reflect a true difference or might be caused by a difference in the classification of stage. We will come back to this issue in the discussion. The patterns of stage distribution across socioeconomic groups did not change substantially after the exclusion of patients for whom no information on stage was available.

Table 4. Stage of disease (%) at diagnosis by socioeconomic status (SES), lung cancer, Southeastern Netherlands and South Thames, 1980-1989

	Southeastern Netherlands						South Thames					
Stage	High	(2)	(3)	(4)	Low	Tot	High	(2)	(3)	(4)	Low	Tot
Local	20.0	21.3	19.3	19.9	19.5	19.8	38.9	38.4	39.2	38.4	37.6	38.6
Regional	18.3	18.7	20.3	20.2	18.4	19.3	7.0	7.3	6.4	6.3	5.8	6.7
Distant	27.2	24.6	26.5	26.3	28.7	27.0	23.3	23.7	22.6	22.6	23.5	23.1
Unknown	34.5	35.4	33.9	33.6	33.5	33.9	30.8	30.7	31.7	32.7	33.1	31.6
Total	100	100	100	100	100	100	100	100	100	100	100	100

Table 5 shows stage specific lung cancer survival for both areas. In two of the four distinguished stage categories, 5-year survival was higher in the Southeastern Netherlands than in South Thames. The difference in survival was especially large for patients with local disease.

The ratios of 5-year RSR for patients with regional disease, metastatic disease or

or unknown stage of disease and the 5-year RSR of patients with local disease (reference category) are also given in table 5. The contrast in survival between categories of stage was much more marked in the Southeastern Netherlands than in the South Thames area. This is also reflected in the 5-year RSR for South Thames patients with local and regional disease, which were similar. This implies again that the classification of stage in the South Thames area is probably not correct as was also shown by the relatively low percentage of lung cancer patients with regional disease in this area (table 4).

Table 5. Five year relative survival rate by stage, lung cancer, Southeastern Netherlands and South Thames, 1980-1989

	Southeastern Netherlands			South Thames		
Stage	5 year RSR	95% CI	Survival ratio (ref=local)	5 year RSR	95% CI	Survival ratio (ref=local)
Local	35	31-39	1.00	10.5	9.8-11.2	1.00
Regional	12	9-15	0.34	11.7	10.2-13.2	1.11
Distant	1	0.4-1.6	0.03	2.3	2.0-2.6	0.22
Unknown	10.3	8.4-12.2	0.29	5.5	5.0-6.0	0.52

RSR: Relative Survival Rate; CI: confidence interval; ref: reference category

Table 6 shows the results from the multivariate analyses on lung cancer for both areas. Period of diagnosis (both areas) and sex (South Thames) are not shown in this table as they proved not to confound the association between socioeconomic status and survival. The gradient in hazard ratios across socioeconomic groups after adjustment for confounders, was slightly steeper in the Southeastern Netherlands than in the South Thames area, but the corresponding HR was of borderline statistical significance. The addition of stage to the model caused negligible changes in the hazard ratios for both areas. These findings are in agreement with those from the univariate analyses in which we found no clear association between socioeconomic status and stage in either of the two areas.

Table 6. Hazard ratio by socioeconomic status, lung cancer, Southeastern Netherlands and South Thames, 1980-1989

	Southeastern Netherlands		South Thames	
Model:	Fu,SES, age,sex	+Stage	Fu,SES, age	+Stage
	HR (95% CI)	HR (95% CI)	HR (95% CI)	HR (95% CI)
Affluent	1.00	1.00	1.00	1.00
(2)	0.96 (0.75-1.23)	0.94 (0.77-1.15)	1.04 (0.95-1.14)	1.04 (0.95-1.13)
(3)	1.06 (0.87-1.30)	1.06 (0.90-1.24)	1.09 (0.99-1.20)	1.11 (1.01-1.21)
(4)	1.13 (0.94-1.38)	1.13 (0.96-1.32)	1.12 (1.02-1.24)	1.13 (1.03-1.24)
Deprived	1.17 (0.98-1.41)	1.17 (1.00-1.36)	1.14 (1.00-1.29)	1.13 (1.01-1.27)
Slope	1.05 (1.00-1.09)	1.05 (1.02-1.08)	1.04 (1.01-1.06)	1.03 (1.01-1.06)

FU: follow-up; SES: socioeconomic status; HR: hazard ratio; CI: confidence interval

6.3.2 Breast Cancer

For breast cancer, we observed a higher 5 year survival for patients in the Southeastern Netherlands than in the South Thames area. The gradient in breast cancer survival by socioeconomic status was more consistent and steeper in the South Thames patients and the ratio of the 5 year RSR for the highest and lowest socioeconomic group was 1.17 in the South Thames area and 1.05 in the Southeastern Netherlands (table 7).

Table 7. Five year relative survival rate by socioeconomic status, breast cancer, Southeastern Netherlands and South Thames, 1980-1989

Socioeconomic Status	Southeastern Netherlands	South Thames
High	77 (73-81)*	70 (69-71)
(2)	74 (69-79)	66 (65-67)
(3)	75 (71-79)	63 (61-65)
(4)	72 (68-76)	63 (61-65)
Low	73 (69-77)	60 (57-63)
Total	74 (72-76)	66 (65-67)
Ratio High/Low	1.05	1.17

* 95% confidence interval between brackets

The variation in stage distribution between the areas for breast cancer was less marked than for lung cancer. The percentage of patients registered with regional disease was much higher in the Southeastern Netherlands, while the percentage with either a metastatic disease or stage unknown was higher in the South Thames area. For breast cancer, we found a higher percentage of patients with metastatic disease in the lower socioeconomic groups for both areas, although the differences between the highest and lowest socioeconomic group were not very large. In the Southeastern Netherlands the percentage of patients with a metastatic disease ranged from 5.4 in the highest to 8.6 in the lowest socioeconomic group, while it ranged from 8.6 in the highest to 12.3 in the lowest socioeconomic group in the South Thames area (table 8).

Table 8. Stage of disease (%) at diagnosis by socioeconomic status (SES), breast cancer, Southeastern Netherlands and South Thames, 1980-1989

	Southeastern Netherlands						South Thames					
SES:	High	(2)	(3)	(4)	Low	Tot	High	(2)	(3)	(4)	Low	Tot
Local	49.5	49.5	46.7	46.8	48.5	48.0	48.4	48.8	50.8	48.1	44.8	48.7
Regional	35.9	31.5	33.0	33.7	31.8	33.4	20.9	21.8	21.0	21.7	23.2	21.4
Distant	5.4	6.6	6.3	6.8	8.6	6.8	8.6	9.4	9.5	10.5	12.3	9.5
Unknown	9.2	12.4	13.9	12.6	11.1	11.8	22.1	20.0	18.7	19.7	19.7	20.4
Total	100	100	100	100	100	100	100	100	100	100	100	100

For breast cancer patients 5-year survival was higher in 3 out of 4 stage categories in the Southeastern Netherlands than in the South Thames area. Only for patients with metastatic disease the 5-year survival was higher in the South Thames area (table 9). For breast cancer patients, we found that the ratio of survival rates for patients with metastatic disease as compared to those with local disease (reference category), was much smaller (larger contrast) in the Southeastern Netherlands, while for the other stage categories the ratios were of similar magnitude in both areas.

Table 9. Five year relative survival rate by stage, breast cancer, Southeastern Netherlands and South Thames, 1980-1989

	Southeastern Netherlands		South Thames	
	5 year RSR (95% CI)	survival ratio (ref=local)	5 year RSR (95% CI)	survival ratio (ref=local)
Local	91 (89-93)	1.00	81 (80-82)	1.00
Regional	67 (64-70)	0.74	60 (58-62)	0.74
Distant	12 (7-17)	0.13	24 (22-26)	0.30
Unknown	62 (56-68)	0.68	57 (55-59)	0.70

RSR: Relative Survival Rate; CI: confidence interval; ref: reference category

Table 10 shows the results from the multivariate analyses of breast cancer survival in both areas. Period of diagnosis did not confound the association between SES and survival in the Southeastern Netherlands and is therefore not mentioned in table 9.
The gradient in hazard ratios by socioeconomic status was much steeper in the South Thames area, which was also found in the univariate analyses. Adjustment for stage caused only small changes in the hazard ratios in the South Thames area, while in the Southeastern Netherlands, differences in hazards by socioeconomic status became much smaller after adjustment for stage.

Table 10. Hazard ratio by socioeconomic status, breast cancer, Southeastern Netherlands and South Thames, 1980-1989

	Southeastern Netherlands		South Thames	
Model:	Fu,SES, Age	+Stage	FU,SES, Age,Per	+Stage
	HR (95% CI)	HR (95% CI)	HR (95% CI)	HR (95% CI)
Affluent	1.00	1.00	1.00	1.00
(2)	1.01 (0.71-1.44)	1.04 (0.79-1.37)	1.15 (1.02-1.30)	1.15 (1.06-1.25)
(3)	1.02 (0.76-1.36)	1.02 (0.82-1.28)	1.27 (1.12-1.44)	1.29 (1.19-1.41)
(4)	1.14 (0.87-1.50)	1.05 (0.85-1.30)	1.27 (1.11-1.46)	1.24 (1.13-1.36)
Deprived	1.20 (0.92-1.57)	1.03 (0.83-1.27)	1.46 (1.22-1.76)	1.43 (1.26-1.62)
Slope	1.05 (0.99-1.12)	1.01 (0.96-1.06)	1.09 (1.06-1.13)	1.09 (1.06-1.11)

FU: Follow-up; SES: socioeconomic status; HR: hazard ratio; CI: confidence interval

6.3.3 Colorectal cancer

The results in table 11 show that both also for colorectal cancer, 5 year survival is higher for patients from the Southeastern Netherlands than for patients from the South Thames area. Furthermore, both the pattern and gradient in survival by socioeconomic status were very similar for both areas.

Table 11. Five year relative survival rate by socioeconomic status, colorectal cancer, Southeastern Netherlands and South Thames, 1980-1989

Socioeconomic Status	Southeastern Netherlands	South Thames
High	54 (49-59)*	41 (39-43)
(2)	53 (46-60)	41 (40-42)
(3)	50 (45-55)	39 (37-41)
(4)	48 (43-53)	36 (34-38)
Low	48 (44-52)	37 (34-40)
Total	50 (48-52)	39 (38-40)
Ratio High/Low	1.13	1.11

* 95% confidence interval between brackets

For colorectal cancer, the differences in stage distribution between the areas were much smaller than for the other cancers. We observed a slightly higher percentage with local and regional disease in the Southeastern Netherlands and a lower percentage with unknown stage in this area. There was no clear association between SES and stage in either of the areas. After the exclusion of patients without information on stage, the differences in stage distribution between the areas remained similar (table 12).

Table 12. Stage of disease (%) at diagnosis by socioeconomic status (SES), colorectal cancer, Southeastern Netherlands and South Thames, 1980-1989

	Southeastern Netherlands						South Thames					
SES:	High	(2)	(3)	(4)	Low	Tot	High	(2)	(3)	(4)	Low	Tot
Local	38.3	45.9	45.9	45.0	42.3	43.2	41.8	42.1	40.7	40.8	40.1	41.4
Regional	19.2	18.4	16.9	19.1	17.8	18.3	14.3	15.1	14.7	13.1	14.2	14.4
Distant	21.5	18.0	15.5	14.3	18.4	18.1	18.3	18.1	17.9	18.8	18.0	18.2
Unknown	21.0	17.7	21.6	21.6	21.5	20.5	25.5	24.6	26.7	27.3	27.7	26.0
Total	100	100	100	100	100	100	100	100	100	100	100	100

For colorectal cancer, patients in the Southeastern Netherlands with local disease and unknown stage showed a much higher 5-year survival than patients with the same stage in the South Thames area. Relative survival for patients from the two areas with regional disease was similar, while (as for lung and breast cancer) for patients with metastatic disease, those living in the Southeastern Netherlands had a lower 5-year RSR than those living in South Thames. Again, we observed a larger contrast in relative survival by stage in the Southeastern Netherlands as compared to South Thames (table 13).

Table 14 shows the results from the multivariate analyses of colorectal cancer survival. Sex and period of diagnosis did not confound the association between socioeconomic status and survival and are therefore not presented in the models.
The gradient in hazard ratios by socioeconomic status from a model with SES and confounders seemed to be somewhat steeper for colorectal cancer patients from the Southeastern Netherlands, with borderline statistical significance. Adjustment for stage had hardly any effect in the South Thames area, while the effect of stage was larger in the Southeastern Netherlands.

Table 13. Five year RSR by stage, colorectal cancer, Southeastern Netherlands and South Thames, 1980-1989

	Southeastern Netherlands		South Thames	
	5 year RSR (95% CI)	survival ratio (ref=local)	5 year RSR (95% CI)	survival ratio (ref=local)
Local	76 (72-80)	1.00	60 (58-62)	1.00
Regional	39 (33-45)	0.51	37 (35-39)	0.62
Distant	3 (1-5)	0.04	9 (8-10)	0.15
Unknown	49 (45-53)	0.64	31 (29-33)	0.52

RSR: Relative Survival Rate; CI: confidence interval; ref: reference category

Table 14. Hazard ratio by deprivation, colorectal cancer, Southeastern Netherlands and South Thames, 1980-1989

	Southeastern Netherlands		South Thames	
Model:	Fu,SES, Age, subsite,	+Stage	FU,SES, Age, subsite	+Stage
	HR (95% CI)	HR (95% CI)	HR (95% CI)	HR (95% CI)
Affluent	1.00	1.00	1.00	1.00
(2)	0.99 (0.75-1.31)	1.08 (0.86-1.35)	0.98 (0.87-1.08)	1.02 (0.94-1.11)
(3)	1.05 (0.84-1.32)	1.11 (0.92-1.33)	1.04 (0.94-1.16)	1.05 (0.96-1.14)
(4)	1.16 (0.93-1.44)	1.30 (1.09-1.54)	1.14 (1.02-1.27)	1.14 (1.04-1.25)
Deprived	1.19 (0.96-1.47)	1.21 (1.02-1.43)	1.14 (0.97-1.34)	1.12 (0.98-1.28)
Slope	1.05 (1.00-1.10)	1.06 (1.02-1.10)	1.04 (1.01-1.07)	1.04 (1.01-1.06)

FU: follow-up; SES: socioeconomic status; HR: hazard ratio; CI: confidence interval

6.4 Discussion

The aim of the analyses described in this chapter was to find out whether socioeconomic variation in survival from the 3 most common cancers differs between two areas within western Europe: the Southeastern Netherlands and the area covered by the South Thames Regional Health Authority. Furthermore, the impact of stage on the association between socioeconomic status and survival was studied for both areas.

The results of the comparison suggest that socioeconomic variation in survival from lung and colorectal cancer is similar in both areas, with a better survival in the higher socioeconomic groups. For breast cancer, we also observed better survival for the higher socioeconomic groups in both areas, with clearly larger socioeconomic variation in survival in the South Thames area than in the Southeastern Netherlands.

Overall, the impact of stage of disease at diagnosis on the association between socioeconomic status and survival appeared to be small. Univariate analyses showed that only in breast cancer patients, metastatic disease was more common in the lower socioeconomic groups in both areas. In the multivariate analyses, adjustment for stage only had an effect on the hazard ratios for different socioeconomic groups of breast cancer patients from the Southeastern Netherlands. In this area, the socioeconomic gradient in breast cancer survival disappeared after adjustment for stage, while it remained unchanged for breast cancer patients from the South Thames area. For cancers of the lung and colorectum we observed no substantial change in hazard ratios for socioeconomic categories after adjustment for stage in either of the study-areas.

We considered some methodological issues which might have influenced the

study results. The type of data used and the method of data analysis were similar for the two study areas. We have used the same inclusion criteria, study period, follow-up period and analytical methods to study socioeconomic variation in cancer survival in both areas. For the South Thames area however, a substantial number of patients could not be included in the analysis, either because their postcode was unknown or because their registration was based on a death certificate only. The survival of cases with unknown postcode did not differ substantially from the survival of the other cases however (see 5.2). Furthermore, the effect of excluding DCO cases on the gradient in survival by deprivation proved to be not very large (see 5.2).

The measure of socioeconomic status used in both countries differs with respect to content and level of measurement. The area-based measure developed for the Dutch analyses was based on about 20 socioeconomic and sociodemographic variables. Each postcode had been assigned to one of 45 sociodemographic categories, according to these 20 variables, while we mainly focused on education as the variable of interest in our analysis. In the British study, four indicators of material deprivation were used to calculate a score for each census enumeration district, from which we derived the Carstairs Index. Each small area was assigned to one of the 5 deprivation categories, defined by quintiles of the national population distribution, according to the combined score on these 4 variables. This difference in construction of the socioeconomic variables in the two areas may have caused more misclassification of the educational level of postcodes in the Southeastern Netherlands. On the other hand, information at the smallest level concerned only 16 households per postcode on average in the Netherlands and about 400 households per census enumeration district on average in the South Thames area, which would imply less misclassification of the socioeconomic score for the Southeastern Netherlands than for the South Thames area.

Another methodological issue concerns a possible difference in the quality of the data on stage of disease at diagnosis between the cancer registries. This is indicated by a number of findings: firstly, the percentage of lung cancer patients diagnosed with local disease was much higher in the South Thames area than in the Southeastern Netherlands, while the percentage of patients with regional disease was exceptionally low. This is not what we would have expected, given the low overall lung cancer survival in the South Thames area. Furthermore, we observed a larger contrast in relative survival rates between categories of stage in the Southeastern Netherlands as compared to the South Thames area. These findings imply substantial understaging of lung cancer patients from the South Thames area, but we do not know whether this is differential by deprivation category. An indicator of the accuracy of staging in the various deprivation categories could be the percentage of patients without information on stage, which is only slightly higher in the more deprived patient groups however. Taking all these findings together we might conclude that the staging of lung cancer patients or recording of stage by the registry in the South Thames area is not nearly as good as in the Southeastern

Netherlands.
For breast cancer we observed smaller differences in the distribution of stage categories between the areas than for lung cancer, although the percentage of patients classified with regional disease was clearly lower for breast cancer in the South Thames area. However, also for breast cancer survival data point to possible differences in the quality of information on stage. Contrast in survival for patients with metastatic disease as compared to those with local disease was larger in the Southeastern Netherlands than in the South Thames area. This is further illustrated by the effect of adjustment for stage on the socioeconomic gradient in survival. In both areas we had observed a higher percentage of patients in the lower socioeconomic groups diagnosed with metastatic disease. However, adjustment for stage explained most of the survival gradient in the Southeastern Netherlands, while for South Thames the gradient remained similar to the unadjusted gradient.
Finally the results of the analyses for colorectal cancer also imply misclassification of stage in the South Thames area. Again we found clearly more contrast in survival by stage category in data from the Southeastern Netherlands as compared to South Thames. However, understaging in these patients seemed less severe than in patients with lung cancer.
In summary, these findings all point at a less valid staging system or recording of stage in the South Thames area than in the Southeastern Netherlands. This could be due to understaging by clinicians, which is reflected in a higher percentage of patients with local disease, especially in lung cancer and to a lesser extent for the other cancers. The understaging of cancer patients in the South Thames area could be caused by a lack of access to specialised care for at least part of the patients. This access could be different for patients from different socioeconomic groups, but this was not suggested by our findings on the stage distribution across socioeconomic groups.

In earlier chapters we discussed the possibility of overestimating the socioeconomic gradient in survival as a result of using only one life table to adjust for mortality from causes other than the cancer under study (expected survival). In a study on male mortality according to level of education in nine different countries in the 1970s (age range 35-54), socioeconomic variation in general mortality was larger in England and Wales than in the Netherlands.[5] In both countries mortality was higher in the lower socioeconomic groups. If the socioeconomic gradient in overall mortality is larger in the South Thames area than in the Southeastern Netherlands, the overestimation of the gradient in survival by socioeconomic status may be larger in the South Thames area as well. This above referenced comparative study on overall mortality only concerns men in a specific age range, and therefore its results can only be an indicator of possible effects of socioeconomic variation in overall mortality on the socioeconomic gradient in cancer survival.

We found a clear difference in the impact of socioeconomic status on breast cancer survival for both areas, with a much steeper gradient in breast cancer survival in the South Thames area, while for lung and colorectal cancer this was

not the case. One possible explanation for this finding is that early detection, resulting in a better prognosis, is more common in breast cancer than in lung and colorectal cancer. Socioeconomic variation in the early detection of breast tumours (less advanced disease in the higher socioeconomic groups) may well be larger in the South Thames area than in the Southeastern Netherlands. This was not reflected in a comparison of the stage distribution across socioeconomic groups in both areas however, but this might be explained by the understaging of patients in the South Thames area.

Overall survival is higher in breast cancer than in lung and colorectal cancer, and it has been hypothesized that in cancers with a good overall prognosis the impact of treatment on socioeconomic variation in survival may be larger.[4] So if socioeconomic variation in treatment were part of the explanation of the differences in survival gradient between both areas, these differences could be expected to be maximal for breast cancer. For the Southeastern Netherlands, this explanation is not appropriate because any variation in survival by socioeconomic status disappeared in the Southeastern Netherlands after adjustment for stage. For the South Thames area, it could well be a plausible explanation for (part of) the socioeconomic variation in cancer survival. However, in our previous study of the association between socioeconomic status and broad treatment categories in both areas, no systematic differences in treatment between socioeconomic groups appeared. If the quality of and compliance with treatment would differ between socioeconomic groups, treatment differences by socioeconomic status could explain (part of) the variation in survival by deprivation in the South Thames area.

From our study of the socioeconomic gradient in survival in two areas within western Europe, we can draw some tentative conclusions. Socioeconomic variation in survival from lung and colorectal cancer was similar in the two study areas, whereas socioeconomic variation in breast cancer survival was much larger in the South Thames area. Stage of disease at diagnosis did not seem to be a major determinant of this variation in either of the two areas, except for breast cancer in the Southeastern Netherlands. We need to be careful in drawing firm conclusions based on the stage data from Thames Cancer Registry however. A next step would be to conduct a similar study, while trying to avoid differences in design between both studies, especially differences in the measure of socioeconomic status and differences in the quality of stage information.

References

1. Hakulinen T, Abeywickrama KH. A computer program package for relative survival analysis. ComP Prog Biomed 1985;19:197-207.
2. Cutler SJ, Ederer F. Maximum utilisation of the life table method in analyzing survival. J Chronic Dis 1958;8:699-712.
3. Hakulinen T, Tenkanen L. Regression analysis of relative survival rates. Appl Stat 1987;36:-309-317.
4. Kogevinas E. Longitudinal Study. Sociodemographic differences in cancer survival. OPCS series LS no. 5. London: HMSO, 1990.
5. Kunst AE, Mackenbach JP. The size of mortality differences associated with educational status in nine industrialized countries. Am J Public Health 1994;84:932-937.

Chapter 7. Discussion

7.1 Summary of the results

The first aim of the studies reported in this thesis was to describe socioeconomic variation in cancer survival for two areas within western Europe: the Southeastern Netherlands and the area covered by the South Thames Regional Health Authority (RHA) in Southeast England. In the Southeastern Netherlands, we found, after adjustment for confounding variables, an up to 20% higher probability of death for patients from low socioeconomic status (SES) areas diagnosed with cancer of the lung, breast, colorectum, and prostate, as compared to patients from high SES areas. For stomach cancer, a higher probability of death was found for patients from high SES areas. In the South Thames area, we found a higher probability of death for deprived patient groups as compared to affluent patient groups for 7 out of the 10 most common cancers (lung, breast, colorectum, bladder, prostate, uterus, and cervix). This excess mortality for deprived patients ranged from 13% for lung cancer to 48% for breast cancer, after adjustment for confounding variables. For the remaining 3 cancers (stomach, pancreas, and ovary), no statistically significant variation in survival by deprivation was found, although the results were indicative of better survival in the affluent patient groups.

The second aim of this thesis was to investigate the impact of a number of prognostic factors, which were grouped in four main categories, on the association between SES and cancer survival. Firstly, socioeconomic differences in some biological features of a tumour (subsite in colorectal and stomach cancer, histological type of a tumour) were studied. For both study areas, we found no clear association between SES and either subsite in colorectal or stomach cancer or the histological type of a tumour in any of the cancers.

We further studied the association between SES and stage of disease at diagnosis, as an indicator of delay in diagnosis. For most cancers we found no association between SES and stage of disease at diagnosis in both areas. In the Southeastern Netherlands, we only found a higher proportion of patients with a metastatic disease in the lower socioeconomic groups for breast cancer and in the higher socioeconomic groups for stomach cancer. In the South Thames area, only for breast and prostate cancer we found a higher percentage of patients diagnosed with an advanced disease in the lower socioeconomic groups as compared with the higher socioeconomic groups. Stage has also been studied in conjunction with survival and it proved to explain most of the variation in survival by SES for breast and stomach cancer in the Southeastern Netherlands. In the South Thames area stage explained part of the association between deprivation and survival for older women (65-99 years) with breast cancer, but had no effect for younger women (30-64 years) with breast cancer or for patients with cancer of the prostate. It seems likely that there was misclassification of stage in the South Thames area in the 1980s, which might have been caused by understaging.

The third possible explanation of socioeconomic variation in cancer survival which was studied is treatment. We investigated whether differences in treatment by SES exist, even in countries like the Netherlands and England, where access to

and use of health care services is assumed to be equal for every citizen. In both areas, we observed no clear difference in type of treatment by SES in any of the cancers studied. However, treatment was studied in broad categories, and a more detailed analysis of for example the quality of cancer treatment might have shown differences by SES (see also 7.2.3).

Finally, the association between SES and some host factors (co-morbidity and life events) was studied in the Southeastern Netherlands. A larger number of co-morbid conditions at diagnosis for patients from lower socioeconomic groups was found as compared to those from higher socioeconomic groups, and furthermore low SES patients reported a larger number of life events. Unfortunately, co-morbidity and life events could not be studied in conjunction with survival and therefore the results on these factors are only indicative of their possible impact on socioeconomic variation in cancer survival.

In summary, these studies suggest that both stage and factors related to host resistance (co-morbidity and life events) explain part of the socioeconomic variation in cancer survival. However, stage was only an explanatory factor of socioeconomic variation in survival from breast and stomach cancer in the South-eastern Netherlands. In the South Thames area, stage of disease at diagnosis did not explain the survival differences by deprivation category. Furthermore, the effect of co-morbidity and life events on survival could not be studied.

The third aim of this thesis was to compare the socioeconomic gradient in survival for the 3 most common cancers (lung, breast, colorectum) in the two study areas: the Southeastern Netherlands and the South Thames area. Socioeconomic variation in survival from lung and colorectal cancer was similar in both areas, whereas for breast cancer, variation in survival by SES was larger in the South Thames area than in the Southeastern Netherlands. The impact of stage of disease at diagnosis on socioeconomic variation in survival was also studied for the 3 cancers in both areas. In multivariate analyses, adjustment for stage only diminished the socioeconomic gradient in breast cancer survival in the Southeastern Netherlands, while stage had no effect in either breast cancer in the South Thames area or in lung and colorectal cancer in both areas. The absence of an effect of stage in the South Thames area could be caused by misclassification of stage.

7.2 Methodological issues

7.2.1 Data sources

Population based cancer registries are the major data sources for the studies described in this thesis. Criteria of quality of cancer registry data, such as validity and completeness, need to be met before these data can be used in a population based study. As discussed in chapter 3, the percentage of cancer registrations confirmed by histology (HV%) is a clear indicator of validity and this percentage is substantially lower for data from the Thames Cancer Registry as compared to the

Eindhoven Cancer Registry, for each of the studied cancers. This finding has implications for the overall survival rates calculated with data from Thames Cancer Registry. In general, survival of cases for which histological evidence is not available is lower than survival of cases for which it is available (results now shown). It seems however, that our study-results on survival by level of deprivation have not been biased, as the HV% did not vary systematically between deprivation categories.

The percentage of cases registered on the basis of a death certificate only (DCO-cases), is another indicator of validity, which was evaluated for the Thames Cancer Registry. In general this percentage was rather high for this registry, but again the comparison of survival between deprivation categories was not biased. Although survival for DCO-cases is lower than survival for other cases, the percentage of DCO-cases did not differ systematically by deprivation category, except for cancers of the lung and breast. The effect of variation in the DCO% by deprivation on the association between deprivation and survival for these cancers was found to be very small (see 5.2).

In the Netherlands, due to strict interpretation of privacy regulations by the Central Bureau of Statistics, death certificates cannot be used as an additional source of notification of cancer cases and therefore the DCO% cannot be calculated. We have no reason to assume however, that for this one indicator of quality of registry data the Eindhoven Cancer Registry would show bad results, while for other indicators of validity and completeness the results are good, such as for the HV% and the mortality/incidence ratio. This ratio is an indirect measure of completeness of cancer registration (see chapter 3) and our evaluation of completeness of registration in chapter 3 was reassuring, especially as recent estimates of completeness were high in both study areas.

We conclude, that both the relatively low HV% and the relatively high DCO% for data from the Thames Cancer Registry are an issue of concern. However, as discussed above, it is very unlikely that our main results have been biased, as we did not observe an unequal distribution of these indicators of validity across deprivation categories.

7.2.2 Measures of socioeconomic status

In an ecological study the unit of analysis is a group of individuals. The studies reported in this thesis are not strictly ecologic but rather mixed-ecological, which means that the measure of SES is area-based, while information on outcome (cancer survival) and prognostic factors is based on individual data.[1] Information on the SES of individual patients was not available from the Eindhoven Cancer Registry. Occupation and social class of cancer patients were known from the records of Thames Cancer Registry, but only for about a third of all patients, with even a lower percentage for women. We therefore did not use these measures in our survival analyses.

Theoretically, the area-based measures of SES employed in these studies may be looked upon in two ways. Firstly, they can be seen as proxy measures of the SES of individual cancer patients, who are assigned to a socioeconomic category according to their postcode of residence at time of diagnosis.
An important problem related to this type of analysis is known as the 'ecological fallacy' and results from making a causal inference about individual phenomena on the basis of observations of groups.[2] The results of our validation studies in chapter 3 imply that the association between SES at the individual level and cancer survival may be somewhat underestimated if an area-based measure of SES is used. Secondly, the area-based measures of SES can be seen as characteristics of SES at the aggregate level, without any intention to interpret them as individual level socioeconomic indicators. Segregation of individuals into small areas which vary in socioeconomic level, is an important feature of social stratification. Various studies have shown that the socioeconomic level of a small area has health effects over and above those of the SES of an individual.[3-5]
Both the individual level effect and the area level effect give insight into the type of policy measures that could be taken to reduce socioeconomic inequalities both in general health and in cancer survival. Except for policy measures aimed at groups of people identifiable through a common factor, such as a low educational level, policy measures may aim at a reduction of socioeconomic variation in cancer survival at the small-area level. From this perspective, an important strength of area-based analyses is that people can also be identified and approached for e.g. health education at the neighbourhood level. Furthermore, extra resources (e.g. health services with good access) may be provided for the inhabitants of specific small areas.

7.2.3 Prognostic factors

Stage of disease at diagnosis
With a few exceptions, we found no differences between socioeconomic groups in the distribution of stage of disease at diagnosis. Other studies have reported a clearer association between SES and stage, with more advanced stages in the lower socioeconomic groups in for example breast cancer[6,7], colorectal cancer[8,9] and cancer of the prostate.[10]
We considered the possibility that the information on stage from the Eindhoven Cancer Registry was (partly) invalid. Data on stage from this registry were registered according to the TNM classification[11] and are based on pathological evidence supplemented with clinical data for a small percentage of patients. The TNM data were used to construct somewhat broader categories of stage than those in the TNM system (local, regional, distant, unknown), as this lowered the number of patients in the category "stage unknown." We also studied the distribution of stage categories according to the TNM classification across socioeconomic groups and this yielded similar results.

Another, more plausible explanation for our findings on small socioeconomic variation in stage distribution might be that the health care system is reasonably accessible for everyone in the Southeastern Netherlands (see also chapter 4). As we reported in the introduction, only small socioeconomic inequalities in the utilization of health care services seem to exist in the Netherlands.[12]
In the South Thames area, the main staging system is not based on the TNM classification. The simplified staging system, based on pathology reports, operation notes and other information, distinghuished the same stage categories as in the Southeastern Netherlands (local, regional, distant, unknown). The prognostic significance of this variable has been found to be similar to stage according to the TNM classification in breast cancer (Lutz J-M, personal communication). It seems likely however, that there is understaging in this area, which might even be differential by level of deprivation. This would be the case if different diagnostic techniques would have been applied to the various socioeconomic groups. Unfortunately, this assumption could not be tested with cancer registry data. The evidence on socioeconomic inequalities in the access to the health care system is even more limited in England, but there might well be differences in access between socioeconomic groups as is also suggested by the large inequalities in cancer survival that we found in the South Thames area (chapter 5).

Treatment
We could only study broad categories of type of treatment to test the hypothesis that patients from different socioeconomic groups are treated unequally, given the biological features of their disease. We found no effect of treatment on the association between SES and survival, which is in agreement with the results from other studies.[13,14]
We considered the possibility that an analysis of more detailed treatment information might have yielded clearer results on the association between SES and treatment. An example of a more detailed analysis is our study of the association between SES and the type of surgery (radical mastectomy versus breast conserving) given to breast cancer patients in the Southeastern Netherlands. We found no clear variation in type of breast surgery by SES (results not shown), which is not in agreement with the results from others who found a higher frequency of mastectomy in less educated women.[15]

Other factors
The results of our studies suggest that other factors than those included in our analyses, explain (part of) the socioeconomic variation in cancer survival. For most cancers in both areas, adjustment for stage, histological type of the tumour, and treatment did not substantially reduce the socioeconomic gradient in survival.
Other possible explanatory factors of socioeconomic variation in cancer survival are not routinely registered and therefore they could not be included in the study. Examples are host factors such as nutritional status, psychological well-being,

social support, immune response, tumour-aggressiveness, and co-morbidity. Some of the most important prognostic factors could be studied however, such as stage of disease at diagnosis, while others such as co-morbidity and adverse life events were not directly available to include in the survival analyses but these could be studied in separate analyses.

7.2.4 Outcome

Survival time, calculated as the time between the date of diagnosis and either the date of death or the end of follow-up, is a very solid measure of outcome. However, the possibility of lead-time bias should also be considered in this study. Lead-time is the time which is added to the survival time of a patient, due to an earlier diagnosis, without improvement of the prognosis. Lead-time might be longer in patients from higher socioeconomic groups, through a shorter delay in these patients, than for patients from lower socioeconomic groups. These differences in delay may result in a longer survival in the higher socioeconomic groups which is an artefact as it does not reflect a true survival advantage.
By adjustment for stage, some of the variation in lead-time might be taken into account, as a shorter delay in general results in a less advanced stage at diagnosis. However, residual confounding by stage, and therefore lead-time, may still explain part of the variation in survival by SES, because stage was classified in rather broad categories. For most cancers, we did not find a substantially larger difference in survival by SES in the first period of follow-up than in the second period (results not shown). This suggests that lead-time is probably not a substantial problem in our studies, as the effect of lead-time is largest in the first period of follow-up, in which the mortality is generally higher than in the second period of follow-up.

7.3 Policy measures

Socioeconomic inequalities in cancer mortality are the end result of two independent forces: socioeconomic variation in cancer incidence and cancer survival. From our own results and from the review of earlier studies on socioeconomic variation in cancer survival (chapter 2) we conclude that socioeconomic variation in cancer mortality seems to be mainly caused by socioeconomic variation in cancer incidence. This has important implications for the type of health policy measures that should be taken to reduce socioeconomic inequalities in cancer mortality. Such policy measures should primarily focus on risk factors and aim at primary prevention of cancer. However, our knowledge on cancer risk factors amenable to successful intervention with lasting effects is limited. One exception is smoking as the major determinant of lung cancer which is also more common in the lower socioeconomic groups.[16] However, measures taken so far to try to spread the message on the adverse health effects of smoking only had limited success, and in

the age group 15-19 in the Netherlands, the prevalence of smoking seems to have increased since 1991.[17]
Another example is breast cancer, in which little can be done to change the distribution of known risk factors such as nulliparity, late age at first birth and late age at menopause[18] across socioeconomic groups.

Although socioeconomic variation in cancer survival is smaller than socioeconomic variation in cancer incidence, it is an important study target as it shows potential for reducing socioeconomic inequalities in cancer mortality. Apart from a description of the association between SES and cancer survival, studies of the impact of possible explanatory factors on this association could result in concrete targets for policy measures.

The prognostic factors that we found to be of importance in explaining socioeconomic variation in cancer survival are stage of disease at diagnosis and, indirectly, co-morbidity and life events. In order to change the distribution of these factors between socioeconomic groups it is important to identify the underlying causes of socioeconomic variation in these prognostic factors which might be amenable to intervention.
We found more advanced stages for some of the cancers in the lower socioeconomic groups of patients. Such socioeconomic variation in stage of disease at diagnosis is most likely caused by variation in patients' delay which can be the result of socioeconomic differences in knowledge about health in general or about cancer symptoms. Health policy should be aimed at more and better health education for people with a low socioeconomic background, both in general, as well as regarding cancer symptoms more specifically. This might result in a shorter delay and therefore in earlier detection of cancer in these patient groups. A tendency to delay may also be caused by a difference in attitude towards health and health care between socioeconomic groups.[19,20] This shows the importance of keeping up the good general access to the health care system, both at the level of general practitioner and specialized care. Furthermore, doctors' delay could be a possible determinant of socioeconomic variation in stage distribution, which implies that more rapid referral of cancer patients by general practitioners to specialists should become common practice, given a good general access to diagnostic services.

We found more co-morbidity and adverse life events among cancer cases with a low SES than among those with a high SES. These prognostic factors could be indicators of a persons' host resistance and are related to factors and circumstances embedded in everyday life, which are in return strongly determined by a relatively low income. In fact, this implies that policy measures should not be restricted to one particular area but that they should involve improvement of housing conditions, neighbourhoods, social networks, income etc.

In the introduction, determinants of health were subdivided into those that are leading to either unjust, unavoidable, or acceptable inequalities in health. The prognostic factors which were identified in this thesis to be of importance in

causing socioeconomic variation in cancer survival may at least partly be considered as unjust. Especially socioeconomic variation in stage distribution can be seen as unacceptable, if this is caused by differences in knowledge between different socioeconomic groups. The other factors which were identified to be of possible prognostic importance, co-morbidity and life events, are also largely determined by factors which are beyond the control of the individual and by conditions of choice, and these should also be the subject of health policy measures.[21]

7.4 Recommendations for future studies

Population based cancer registries have a potential advantage as compared with other data sources for studies on SES and cancer. Data from such registries may be assumed to give a complete picture of the occurrence of and survival from cancer in a specific geographic area. Unfortunately, until now, only a few cancer registries have gathered data on the SES of cancer patients. Routine registration of SES should become common practice in hospitals, so that cancer registries could add this item to their minimum data set. In terms of research, adoption of information on education of cancer patients in the database of cancer registries, would enable monitoring trends in the incidence of and survival from cancer by SES. This would be an important tool to register effects of, and study implications for policy measures in the area of primary or secondary prevention of cancer. Population data on the number of people by age, sex and socioeconomic category are needed however to calculate both socioeconomic variation in incidence and mortality. In the Netherlands, such population data could be extrapolated from surveys, and in Great Britain, they are available through the census.

We have identified a number of prognostic factors which are of importance in causing socioeconomic variation in cancer survival. We recommend that the distribution of such factors as stage of disease at diagnosis and its possible determinants across socioeconomic groups are studied in more detail. One of these determinants is delay, which has been found to be longer in cancer patients from lower socioeconomic groups as compared to patients from higher socioeconomic groups,[22-24] but little is known about the factors causing a longer delay in low SES patients. Especially insights into the underlying determinants of socioeconomic variation in delay are necessary to accomplish progress in the study of socioeconomic variation in cancer patient survival.

The studies reported in this thesis have shown that socioeconomic variation in cancer survival cannot completely be explained by the prognostic factors which are relatively easily available from cancer registry records. Future studies should also focus on other possible determinants of socioeconomic variation in cancer patient survival, e.g. tumour agressiveness, social support and nutritional status (see 7.2.3). Other study designs are necessary to investigate socioeconomic variation in survival and its possible determinants, such as prospective cohort studies in which

extra data are gathered on prognostic factors, besides using data from cancer registry records. Furthermore, studies on cancer (survival) that are not directly designed to study socioeconomic variation should also routinely gather information on the SES of cancer patients. This could yield new insights into the impact of prognostic factors on socioeconomic variation in cancer survival.

References

1. Susser M. The logic in ecological I. The logic of analysis. Am J Public Health 1994;84:825-829.
2. Morgenstern H. Uses of ecological analysis in epidemiologic research. Am J Public Health 1982;72:1336-1344.
3. Kaplan GA, Wilson TW, Cohen RD, Kauhanen J, Wu M, Salonen JT. Social functioning and overall mortality: prospective evidence from the kuopio ischemic heart disease risk factor study. Epidemiology 1994;5:495-500
4. Carstairs V, Morris R. Deprivation and mortality: an alternative to social class? Comm Med 1989;11:210-219.
5. Carstairs V, Morris R. Deprivation: explaining differences in mortality between Scotland and England and Wales. BMJ 1989;299:881-889.
6. Karjalainen S, Pukkala E. Social class as a prognostic factor in breast cancer survival. Cancer 1990;66:819-826
7. Bassett MT, Krieger N. Social class and black-white differences in breast cancer survival. Am J Public Health 1986;76:1400-1403
8. Wegner EL, Kolonel LN, Nomura AMY, Lee J. Racial and socioeconomic status differences in survival of colorectal cancer patients in Hawaii. Cancer 1982;49:2208-2216
9. Brenner H, Mielck A, Klein R, Ziegler H. The role of socioeconomic factors in the survival of patients with colorectal cancer in Saarland/Germany. J Clin Epidemiol 1991;44:807-815
10. Dayal HH, Polissar L, Dahlberg S. Race, socioeconomic status, and other prognostic factors for survival from prostate cancer. J Natl Cancer Inst 1985;74:1001-1006
11. Hermanek P, Sobin LH (eds). UICC TNM Classification of malignant tumours. 4th ed. UICC. Berlin: Springer-Verlag, 1987.
12. Meer JBW van der, Looman CWN, Mackenbach JP. Sociaal-economische verschillen in medische consumptie. Rotterdam: Instituut Maatschappelijke Gezondheidszorg, 1993.
13. Lipworth L, Abelin T, Connelly RR. Socio-economic factors in the prognosis of cancer patients. J Chron Dis 1970;23:105-116
14. Lipworth L, Bennett B, Parker P. Prognosis of nonprivate cancer patients. J Nat Cancer Inst 1972;48:11-14.
15. Muss HB, Hunter CP, Wesley M, et al. Treatment plans for black and white women with stage II node-positive breast cancer. The National Cancer Institute black/white cancer survival study experience. Cancer 1992;70:2460-2467.
16. Wagenknecht LE, Perkins LL, Cutter GR et al. Cigarette smoking behaviour is strongly related to educational status: the CARDIA study. Preventive Medicine 1990;19:158-169.
17. StiVoRo. Jaarverslag 1992. Den Haag: Stichting Volksgezondheid en Roken, 1993.
18. Kelsey JL, Gammon MD, John EM. Reproductive factors and breast cancer. Epidemiol Rev 1993;15:36-47.
19. Sharp K, Ross CE, Cockerham WC. Symptoms, beliefs, and the use of physician services among the disadvantaged. J Health Soc Behav 1983;24:255-263.
20. Blaxter M. Self-definition of health status and consulting rates in primary care. Q J Soc Affairs 1985;1:131-171.
21. Stronks K, Gunning-Schepers LJ. Should equity in health be target number 1? Eur J Public Health 1993;3:104-111.
22. Richardson JL, Langholz G, Bernstein L, Burciaga C, Danley K, Ross RK. Stage and delay in breast cancer diagnosis by race, socioeconomic status, age and year. Br J Cancer 1992;65:922-926.
23. Elwood JM, Moorehead WP. Delay in diagnosis and long-term survival in breast cancer. BMJ 1980;31 May,1291-1294.
24. Gould Martin K, Paganini-Hill A, Casagrande C, Mack T, Ross RK. Behavioral and biological determinants of surgical stage of breast cancer. Preventive Medicine 1982;11:429-440.

Summary

Socioeconomic inequalities in health are systematic differences in the prevalence or incidence of health problems between people of higher and lower socioeconomic status (SES). The most important indicators of SES are occupational status, educational level and income. In general, people with a low socioeconomic status have more health problems than people with a high socioeconomic status. One of the health measures for which socioeconomic inequalities have been described is cancer survival.

In this thesis, results of studies on the association between socioeconomic status and survival from the most common cancers in the Southeastern Netherlands and the area covered by the South Thames Regional Health Authority (RHA) in Great Britain was studied. The aim of both studies was to describe and explain socioeconomic variation in survival. With respect to the latter, the distribution of a number of prognostic factors across socioeconomic groups was studied as well as the impact of such factors on socioeconomic variation in cancer survival. These prognostic factors are: stage of disease at diagnosis, histological type of a tumour and treatment for both areas, and co-morbidity and life events for the Southeastern Netherlands only.

Chapter two contains a review of earlier studies on cancer survival by socioeconomic status in seven countries and for six common cancer sites, which are: colon, rectum, lung, prostate, breast, and cervix. The results of these studies showed that cancer patients with a relatively high socioeconomic status had better survival than patients with a relatively low socioeconomic status for cancers of the colon, rectum, breast, and cervix, while for cancers of the lung and prostate, results were unclear. Furthermore, these studies showed that in general, socioeconomic differences in cancer survival are small, and therefore their contribution to socioeconomic differences in cancer mortality is probably small too.

In chapter 3, the patients and methods employed in the cancer survival studies are described. Socioeconomic variation in cancer survival in both areas was studied with data from two population based cancer registries: the Eindhoven Cancer Registry in the Southeastern Netherlands and Thames Cancer Registry in Southeast England.

In both studies, patients diagnosed in the decade 1980-1989, with either a tumour of the lung, breast, colorectum, prostate or stomach, were included in the study. The number of patients from the South Thames area was much larger and therefore also less frequent cancers could be studied, which are: bladder, pancreas, ovary, uterus, and cervix. In the Southeastern Netherlands, patients were classified by socioeconomic status based on their postcode of residence at time of diagnosis (3 or 5 categories). In the South Thames area, a score on the Carstairs Index, belonging to the corresponding census enumeration district of residence, was used to assign each patient to one of five deprivation categories. The measures of outcome were the same in both studies: to correct for deaths due to causes other than the cancer under study, the relative survival rate was used as the measure of outcome in the univariate analyses, and the hazard ratio in the multivariate analyses.

In chapter 4 the results of the studies on socioeconomic variation in cancer

survival in the Southeastern Netherlands are reported. A lower survival was found among patients with a low socioeconomic status for cancers of the lung, breast, colorectum and prostate, while for stomach cancer, survival was lower for patients with a high socioeconomic status. The results for cancers of the prostate and stomach need to be interpreted with caution, as these were based on relatively small numbers of patients.
For cancers of the lung, colorectum, and prostate, socioeconomic variation in survival could not be explained by the distribution of the prognostic factors stage, histology, and treatment. For breast and stomach cancer, socioeconomic variation in survival could mainly be ascribed to socioeconomic differences in the number of patients diagnosed with a metastatic disease. An advanced disease was more common among lower socioeconomic groups of breast cancer patients and higher socioeconomic groups of stomach cancer patients.

The association between socioeconomic status and co-morbidity and life events was studied for prevalent cancer cases who had reported cancer in a postal survey. In a separate study, the respondent's answer to the survey question on cancer had been validated against records from the Eindhoven Cancer Registry. This validation study showed that the survey underestimated cancer prevalence in the population by 25%. However, after the exclusion of nonmelanoma skin cancer, cancer prevalence was overestimated by a negligible 2%. Misclassification of cancer by the postal survey was differential according to age, sex, education, and urbanization, and this did not change after the exclusion of nonmelanoma skin cancer.
Among prevalent cases of cancer, identified through the postal survey, those with a low socioeconomic status reported more often at least one adverse life event during the year preceding the survey, while no association was found between socioeconomic status and co-morbidity. Socioeconomic variation in co-morbidity was also studied among patients diagnosed in 1993 with one the five most common cancers and registered by the Eindhoven Cancer Registry. We found that for the five sites combined, patients from lower socioeconomic groups were more often diagnosed with a least one other chronic condition than patients with a high socioeconomic status. This pattern was also found among patients with breast cancer, and to a lesser degree in cancers of the lung and colorectum.

The association between deprivation and cancer survival in the South Thames area is described in chapter 5. For breast cancer, we found a clear gradient in survival by deprivation, with better survival for women from more affluent areas. In younger women (30-64 years) the survival gradient by deprivation could not be explained by the prognostic factors stage, morphology and treatment, while for older women (65-99 years), part of the variation in survival by deprivation was due to a higher percentage of deprived women with an advanced disease.
For most other cancers a lower survival was found among deprived patients: lung, colorectum, prostate, bladder, uterus, and cervix. For these cancers, stage of disease at diagnosis did not explain the survival differences by deprivation category. Furthermore, the histological type of a tumour and treatment could not explain

variation in survival by deprivation category.
For cancers of the stomach, pancreas and ovary, no statistically significant variation in survival by deprivation category was found.

In chapter 6, the association between socioeconomic status and survival from the three most common cancers (lung, breast, colorectum) was compared for both areas, using the same methods as in the earlier chapters. Both univariate and multivariate analyses showed that the gradient in survival by socioeconomic status was similar in patients from both areas for cancers of the lung and colorectum, while the socioeconomic gradient in breast cancer survival was much steeper in the South Thames area than in the Southeastern Netherlands. We found no clear association between socioeconomic status and stage of disease at diagnosis, except for breast cancer. In both areas, breast cancer patients with a low socioeconomic status were more often diagnosed with an advanced disease than patients with a high socioeconomic status. However, in a multivariate analysis, stage did not explain socioeconomic variation in survival for South Thames patients, while it did for patients from the Southeastern Netherlands. The absence of an effect of stage in the South Thames area could be caused by misclassification.

Socioeconomic inequalities in cancer mortality seem to be mainly caused by socioeconomic inequalities in cancer incidence. However, the potential of reducing socioeconomic inequalities in cancer incidence by targeting risk factors is limited. Reducing socioeconomic inequalities in cancer survival shows potential for reducing socioeconomic inequalities in cancer mortality.
We identified a number of prognostic factors which explain part of the socioeconomic gradient in cancer survival: stage of disease at diagnosis, co-morbidity, and life events. Future studies should focus on the underlying determinants of socioeconomic variation in these prognostic factors, such as delay as a determinant of late stage at diagnosis. Furthermore, other possible determinants of socioeconomic variation in cancer survival should be studied, such as tumour aggressiveness, social support and nutritional status.

Samenvatting

Sociaaleconomische verschillen in gezondheid zijn systematische verschillen in de prevalentie en incidentie van gezondheidsproblemen tussen mensen met een relatief hoge en mensen met een relatief lage sociaaleconomische status (SES). De belangrijkste indicatoren van SES zijn beroep, opleiding en inkomen. Mensen met een lage SES hebben over het algemeen een slechtere gezondheid dan mensen met een hoge SES. Eén van de gezondheidsmaten waarvoor sociaaleconomische verschillen zijn beschreven is kankeroverleving.

Dit proefschrift bevat de resultaten van een studie naar het verband tussen SES en kankeroverleving voor de meest voorkomende vormen van kanker in Zuidoost Nederland. Daarnaast is het verband tussen SES en kankeroverleving voor de meest voorkomende kankers bestudeerd voor patiënten uit het gebied ten Zuiden van de rivier de Thames in Zuidoost Engeland. Voor beide gebieden is het verband tussen SES en overleving beschreven en zijn verklaringen voor overlevingsverschillen naar SES bestudeerd. Hiertoe is allereerst de verdeling van een aantal prognostische factoren over sociaaleconomische groepen bestudeerd en daarna is de invloed van deze factoren op sociaaleconomische verschillen in overleving bepaald. De volgende prognostische factoren werden in beide gebieden bestudeerd: stadium bij diagnose, histologisch type van de tumor en behandeling. Verder werden co-morbiditeit en life events ook gerelateerd aan de SES van kankerpatiënten in Zuidoost Nederland.

Hoofdstuk 2 bevat een overzicht van de resultaten van eerdere studies naar het verband tussen SES en kankeroverleving in zeven landen en voor zes veel voorkomende kankers, namelijk colon, rectum, long, prostaat, borst en cervix. Deze studies laten zien dat kankerpatiënten met een relatieve hoge SES een betere prognose hebben dan patiënten met een relatief lage SES. Dit geldt voor colon-, rectum-, borst- en cervix- kanker, terwijl de resultaten voor long- en prostaatkanker onduidelijk waren. Uit deze studies bleek dat sociaaleconomische verschillen in kankeroverleving over het algemeen klein zijn en dat hun bijdrage aan sociaaleconomische verschillen in kankersterfte waarschijnlijk ook klein is.

Hoofdstuk 3 bevat een beschrijving van de patiënten en methoden die in de overlevingsstudies zijn gebruikt. Voor de studie van sociaaleconomische verschillen in kankeroverleving werden in beide gebieden data gebruikt van een zogenaamde population based kankerregistratie: de IKZ kankerregistratie in Eindhoven, Zuidoost Nederland en de Thames Cancer Registry in Londen, Zuidoost Engeland.
In beide studies werden patiënten opgenomen die tussen 1980 en 1989 werden gediagnostiseerd met één van de volgende kankers: long, borst, colorectum, prostaat en maag. Het aantal patiënten in de Thames registratie was veel hoger en daardoor konden ook de minder frequente kankers in de Engelse studie worden opgenomen, namelijk blaas, pancreas, ovarium, uterus en cervix.
In Zuidoost Nederland werden patiënten ingedeeld in een aantal SES groepen (3 of 5 groepen) op basis van de postcode van de woonplaats ten tijde van diagnose. In de Engelse studie vormde een score op de Carstairs Index, die materiële deprivatie meet, uitgangspunt om patiënten in te delen in 5 groepen. Dit gebeurde op basis van de woonplaats van patiënten ten tijde van de census van 1980.
De uitkomstmaten waren hetzelfde in beide studies: om te corrigeren voor sterfte

aan andere doodsoorzaken dan de betreffende kanker werden de relatieve overleving respectievelijk de hazard ratio gebruikt als uitkomstmaat in de univariate respectievelijk multivariate analyses.

In hoofdstuk 4 worden de resultaten beschreven van de studies naar sociaaleconomische verschillen in kankeroverleving in Zuidoost Nederland. Voor een aantal kankers werd een lagere overleving gevonden voor patiënten met een lage SES, namelijk long, borst, colorectum en prostaat. Voor maagkanker werd een betere overleving gevonden bij patiënten met een lage SES. De resultaten voor prostaat en maagkanker dienen met de nodige voorzichtigheid te worden geïnterpreteerd, omdat ze gebaseerd zijn op relatief kleine aantallen patiënten.
Voor een aantal kankers (long, colorectum, prostaat) konden sociaaleconomische verschillen in overleving niet worden verklaard door verschillen in de verdeling van de prognostische factoren stadium, histologie en behandeling over SES groepen. Voor borst- en maagkanker konden sociaaleconomische verschillen in overleving met name worden toegeschreven aan sociaaleconomische verschillen in het percentage patiënten dat werd gediagnostiseerd met een metastase op afstand. Een metastase op afstand kwam vaker voor bij borstkanker patiënten met een lage SES en bij maagkanker patiënten met een hoge SES.

Het verband tussen SES en co-morbiditeit en life events werd bestudeerd bij prevalente kankerpatiënten die kanker hadden gerapporteerd in een postenquête. Het antwoord op de enquêtevraag naar het vóórkomen van kanker werd in een afzonderlijke studie gevalideerd. Hiertoe werd gebruik gemaakt van gegevens van de IKZ kankerregistratie. Uit deze valideringsstudie bleek dat de prevalentie van kanker met 25% werd onderschat middels de postenquête. Na uitsluiting van patiënten met nonmelanoma huidkanker werd de prevalentie van kanker met slechts 2% overschat middels de postenquête. De misclassificatie van kanker middels de postenquête was differentieel naar leeftijd, geslacht, opleiding en urbanisatiegraad en deze patronen veranderden niet na uitsluiting van nonmelanoma huidkanker.
Life events werden gerapporteerd door prevalente kankerpatiënten, die eerder waren geïdentificeerd middels een postenquête. Hieruit bleek dat patiënten met een lage SES vaker minimaal één negatieve life event gedurende het jaar voorafgaand aan de enquête rapporteren dan patiënten met een hoge SES. In deze patiënten groep werd geen verband gevonden tussen SES en co-morbiditeit.
Co-morbiditeit werd ook bestudeerd bij patiënten die in 1993 werden gediagnostiseerd met één van de vijf meest voorkomende kankers in Zuidoost Nederland. Bij patiënten met een lage SES werd vaker minimaal één andere chronische aandoening gevonden ten tijde van de kanker diagnose dan bij patiënten met een hoge SES. Dit gold zowel voor alle kankers samen als voor borstkanker en in mindere mate ook voor long- en dikke darm- kanker.

In hoofstuk 5 wordt het verband tussen deprivatie en kankeroverleving in het Zuid Thames gebied beschreven. Dit verband was duidelijk aanwezig voor borstkanker, met een betere overleving voor vrouwen met een hoge SES. Het verband tussen deprivatie en overleving kon niet worden verklaard door de prognostische factoren

stadium, morfologie en behandeling bij vrouwen tussen de 30 en 64 jaar. Bij vrouwen tussen de 65 en 99 jaar kon een deel van de variatie in overleving naar deprivatie worden verklaard uit het hogere percentage vrouwen met een metastase op afstand onder vrouwen met een relatief lage SES.
Voor de meeste andere kankers werd ook een lagere overleving gevonden onder patiënten uit de lage SES groepen: long, colorectum, prostaat, blaas, uterus en cervix. Noch stadium bij diagnose, noch de prognostische factoren behandeling en histologie, droegen veel bij aan de verklaring van de gevonden overlevingsverschillen.
Voor een aantal kankers werd geen statistisch significant verschil in overleving naar deprivatie gevonden en wel voor maag-, pancreas- en ovariumkanker.

In hoofstuk 6 is een vergelijking gemaakt van het verband tussen SES en de overleving voor de drie meest voorkomende kankers (long, borst, colorectum) in beide onderzoeksgebieden. Hierbij werden dezelfde methoden gebruikt als in de hoofdstukken 4 en 5. Zowel de univariate als de multivariate analyses lieten zien dat de sociaaleconomische gradiënt in overleving voor long- en colorectale- kanker vergelijkbaar was in beide gebieden. De gradiënt in borstkanker overleving naar SES was veel sterker in het Zuid Thames gebied dan in Zuidoost Nederland. Er werd geen duidelijk verband gevonden tussen SES en stadium bij diagnose, behalve voor borstkanker. In beide gebieden werd vaker een metastase op afstand gevonden bij patiënten met een lage SES dan bij patiënten met een hoge SES. In de multivariate analyses droeg stadium alleen bij aan de verklaring van sociaaleconomische verschillen in overleving voor borstkankerpatiënten uit Zuidoost Nederland. Een gebrek aan effect van stadium in het Zuid Thames gebied werd mogelijk veroorzaakt door misclassificatie van stadium.

Sociaaleconomische verschillen in kankersterfte lijken vooral veroorzaakt te worden door sociaaleconomische verschillen in kankerincidentie. Echter, de mogelijkheden om sociaaleconomische verschillen in kankerincidentie te verkleinen middels de beïnvloeding van risicofactoren zijn beperkt. Het verkleinen van sociaaleconomische verschillen in kankeroverleving biedt mogelijkheden met het oog op het verkleinen van sociaaleconomische verschillen in kankersterfte. Uit de studies in dit proefschrift kwamen een aantal prognostische factoren naar voren die een deel van de sociaaleconomische gradiënt in kankeroverleving verklaren: stadium bij diagnose, co-morbiditeit en life events. Toekomstige studies zouden zich moeten richten op de determinanten van sociaaleconomische verschillen in deze prognostische factoren, zoals b.v. delay als determinant van een vergevorderd stadium bij diagnose. Verder is het belangrijk om andere potentiële determinanten van sociaaleconomische verschillen in kankeroverleving te bestuderen, zoals de agressiviteit van een tumor, sociale steun en de voedingstoestand van kankerpatiënten.

Dankwoord

Op deze plaats wil ik een aantal mensen bedanken die een bijdrage hebben geleverd aan het tot stand komen van dit proefschrift. Allereerst wil ik mijn promotor en dagelijks begeleider, Johan Mackenbach bedanken voor het in goede banen leiden van mijn onderzoek en voor zijn belangrijke inhoudelijke bijdrage aan dit proefschrift. Jan-Willem Coebergh zorgde voor een goede samenwerking met de IKZ kankerregistratie. I want to thank Michel Coleman for giving me the opportunity to work at the Thames Cancer Registry and for his support during my stay in Great Britain.

Het grootste deel van mijn tijd heb ik doorgebracht op de 'werkvloer' van het Instituut Maatschappelijke Gezondheidszorg, waarvan ik mij de goede contacten met collega's zal blijven herinneren. In het bijzonder wil ik Karien Stronks bedanken voor haar voortdurende steun en het 'meedenken' in alle fasen van mijn onderzoek. Ook Dike van de Mheen heeft mij steeds gesteund en was als kamergenoot vaak eerste aanspreekpunt voor uiteenlopende zaken. Caspar Looman wil ik bedanken voor zijn belangrijke rol bij het leren werken met de overlevingsanalyses.

Dit onderzoek was niet mogelijk geweest zonder de gegevens van de IKZ kankerregistratie en de mensen die deze gegevens hebben verzameld en verwerkt. Speciale dank aan Louis van der Heijden die werkelijk alle problemen met 'de data' aankon. De overige medewerkers van het IKZ wil ik bedanken voor hun gastvrijheid en het vertrouwen dat ze in mij hebben gesteld.

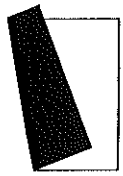
Many thanks to all those at the Thames Cancer Registry who contributed to the studies in this thesis by collecting, storing or interpreting the data that I used for my research.

Een speciaal woord van dank voor mijn paranimfen, Xandra en Saskia, met wie ik hierna nog vele gezellige uren hoop door te brengen.

Tot slot dank ik mijn ouders omdat ze mij altijd onvoorwaardelijk hebben gesteund en natuurlijk Léon die als geen ander weet hoe hij dit proefschrift moet beoordelen.

Curriculum vitae

Carola Schrijvers werd geboren op 1 oktober 1966 in Oss. Ze doorliep het Atheneum A aan het Maasland College te Oss. In 1985 begon ze met de studie Huishoudwetenschappen aan de Landbouwuniversiteit te Wageningen. Het doctoraalprogramma bevatte twee afstudeervakken gezondheidsleer, één afstudeervak voorlichtingskunde en een stage op het International Agency for Research on Cancer in Lyon, Frankrijk.
Na het behalen van het doctoraalexamen in 1991 startte ze met haar promotieonderzoek aan het Instituut Maatschappelijke Gezondheidszorg van de Erasmus Universiteit Rotterdam. Een gedeelte van het in dit proefschrift beschreven onderzoek werd uitgevoerd bij de Thames Cancer Registry in Londen. Sinds april 1995 werkt ze als onderzoeker bij het Instituut voor Verslavings Onderzoek in Rotterdam.



IKZ
INTEGRAAL KANKERCENTRUM ZUID
Samenwerkingsverband voor de
kankerbestrijding in Zuid-Nederland

Voor dit onderzoek werd gebruik gemaakt van gegevens van het Integraal Kankercentrum Zuid (IKZ) in Eindhoven. Het IKZ werkt aan integrale zorg voor mensen met kanker en ondersteunt hulpverleners die daarbij betrokken zijn in ziekenhuizen en in de thuiszorg in het gebied Noord-Brabant en Noord-Limburg.

Eén van de activiteiten van het IKZ is de kankerregistratie, waarbij het gaat om het verzamelen en bewerken van gegevens over alle vormen van kanker. Deze registratie is in 1955 gestart in het oostelijk deel van de regio. Zij heeft inmiddels veel informatie over kanker in Nederland voortgebracht, met name door onderzoek in samenwerking met de Erasmus Universiteit Rotterdam, zoals het onderzoek beschreven in dit proefschrift.

De gegevens die door de registratie-medewerkers worden verzameld, betreffen onder andere demografische kenmerken van patiënten, gebruikte diagnostiek, toegepaste behandeling en follow-up. Clinici en wetenschappers gebruiken deze informatie voor velerlei onderzoek. Hierdoor wordt bijvoorbeeld inzicht verkregen in het voorkomen van kanker, de effecten van preventieve maatregelen en de benodigde toekomstige voorzieningen.