The Epidemiology of Leukaemia in the UK

Fatima Bhayat

MBChB, MSc Public Health (MPH)

Thesis submitted to the

University of Nottingham for the degree of

Doctor of Philosophy, October 2009

ACKNOWLEDGEMENTS

Richard Hubbard, who supervised my PhD with tireless enthusiasm and a great sense of humour, and for whom the glass is always half-full.

Emma Das-Gupta, whose co-supervision was always encouraging.

Chris Smith who provided the data.

Laila Tata and **Tim Card** who shared their impressive knowledge of both general practice data, as well as STATA with me.

Yasmin Eyre and **Alan Smyth,** who so generously shared their office space.

Tim and **Joe C**, whose support and understanding over the past few years made this possible.

SUMMARY

INTRODUCTION

Approximately 9% of all new malignant diagnoses in the UK are due to haematological malignancies. The acute and chronic leukaemias constitute 2.5 % of all cancers and leukaemia is the 12th most common cancer registered in the UK. Approximately 7 000 people are diagnosed with the disease and more than 4 300 people die from leukaemia in the UK each year. As such, they have an important impact on the health of the public and represent a significant cost to the health care budget.

AIMS AND OBJECTIVES

The research presented in this thesis firstly aimed to quantify the incidence of and mortality from the acute and chronic leukaemias in the UK, and to define their associations with gender, age, socioeconomic class, calendar time, and geographic region of residence. A further aim was to determine whether the use of non-steroidal anti-inflammatory drugs (NSAIDs) had a protective effect on the incidence of and mortality from these leukaemias, as has been shown to be the case for a number of other cancers. Finally, the impact of alcohol consumption on leukaemia incidence and mortality was investigated. A surprising result from the incidence and mortality studies was that survival in AML, but not other leukaemias, was worse with increasing socioeconomic

deprivation. This generated an additional hypothesis surrounding potential class bias in bone marrow transplantation in these patients, a new area that was also investigated, in addition to the original aims and objectives of the research.

METHODS

Both general practice and hospital data were used to conduct these population-based studies. 'The Health Improvement Network' (THIN) general practice dataset was used to conduct the cohort studies of incidence and mortality, as well the case-control studies investigating non-steroidal anti-inflammatory drug use and alcohol consumption, as potential risk factors for leukaemia. Hospital Episode Statistics (HES) data were used to investigate the additional hypothesis generated by results of the incidence and mortality studies, which showed that mortality in AML patients worsens with increasing socioeconomic deprivation.

RESULTS

A total of 4162 cases of leukaemia were identified, 2314 (56%) of whom were male. The overall incidence of leukaemia is 11.25 per 100 000 person-years and is independent of socioeconomic class. Median survival from leukaemia is 6.58 years and mortality increases with increasing age at diagnosis. The prognosis in AML is dismal and worsens with increasing socioeconomic deprivation, a phenomenon not seen in other leukaemias.

Bone marrow transplantation declines with increasing socioeconomic deprivation (p for trend <0.01). Patients with AML in the most deprived socioeconomic quintile are 40% less likely to have a bone marrow transplantation than those in the most advantaged socioeconomic class (OR 0.60, p<0.01, 95% C.I. 0.49 - 0.73), even after adjusting for gender, age at diagnosis, year of bone marrow transplantation and co-morbidity.

The risk of leukaemia overall appears to increase marginally with increased use of NSAIDs prior to diagnosis. This is not seen when individual leukaemia subtypes are examined, however, except perhaps in CLL where patients who had received 2-5 prescriptions/year were 29% more likely to be diagnosed with CLL than those who had not had any NSAID prescriptions (O.R. 1.29, p=0.05, 95% C.I. 1.00 - 1.67). There is no statistically significant association between exposure to NSAIDs prior to leukaemia diagnosis, and survival.

There is no statistically significant association between alcohol consumption and risk of developing leukaemia overall, nor with any of the leukaemia subtypes studied here. Alcohol consumption is associated with a lower risk of death in leukaemia overall (HR 0.83, p=0.04, 95% C.I. 0.69 - 0.99), as well as in ALL (HR 0.14, p<0.01, 95% C.I. 0.04 - 0.44) and CLL (HR 0.71, p=0.02, 95% C.I. 0.53 - 0.96), when compared to those who had not consumed any alcohol.

CONCLUSIONS

The age and gender patterns of leukaemia incidence and mortality in the subtypes studied here are consistent with the published literature. Time trends in incidence and mortality must be interpreted with caution due to changes in case ascertainment and classification of leukaemia subtypes over time. The increase in incidence of ALL, CLL and AML over the past 20 years has not been accompanied by improvements in mortality with time. Whether the increase in incidence over time is real or due to better recording, or both, remains unclear.

Similar mortality across socioeconomic gradients in the leukaemias studied here suggests equal access to and uptake of services. The exception to this is in AML, where poorer survival among AML patients from lower socioeconomic classes is seen. AML patients from lower socioeconomic classes are less likely to undergo bone marrow transplantation than their better off counter-parts and this phenomenon is independent of co-existing illness.

The use of NSAIDs does not reduce the risk of developing leukaemia, nor do they improve survival. This research found no statistically significant association between alcohol consumption and incident leukaemia, although alcohol consumption was found to be associated with a reduced risk of death from leukaemia overall, ALL and CLL. Better recording of alcohol

consumption in THIN will enhance its value in investigating associations of alcohol consumption and disease or other outcomes.

TABLE OF CONTENTS

LIST OF TABLES

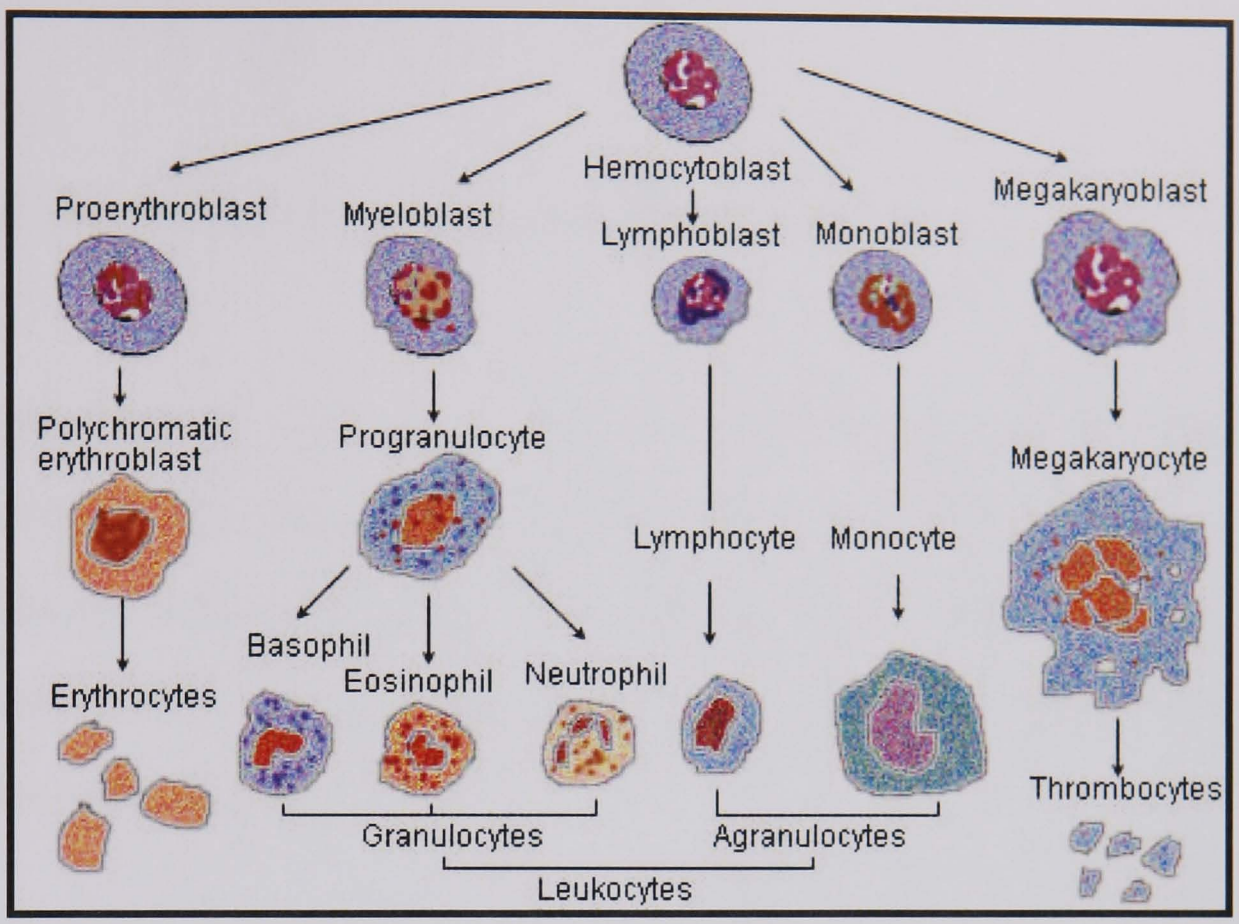
LIST OF FIGURES

1 INTRODUCTION

Approximately 9% of all new malignant diagnoses in the UK are due to haematological malignancies[1]. The acute and chronic leukaemias constitute 2.5 % of all cancers and leukaemia is the 12th most common cancer registered in the UK[1]. Approximately 7 000 people are diagnosed with the disease and more than 4 300 people die from leukaemia in the UK each year[1]. As such, they have an important impact on the health of the public and represent a significant cost to the health care budget.

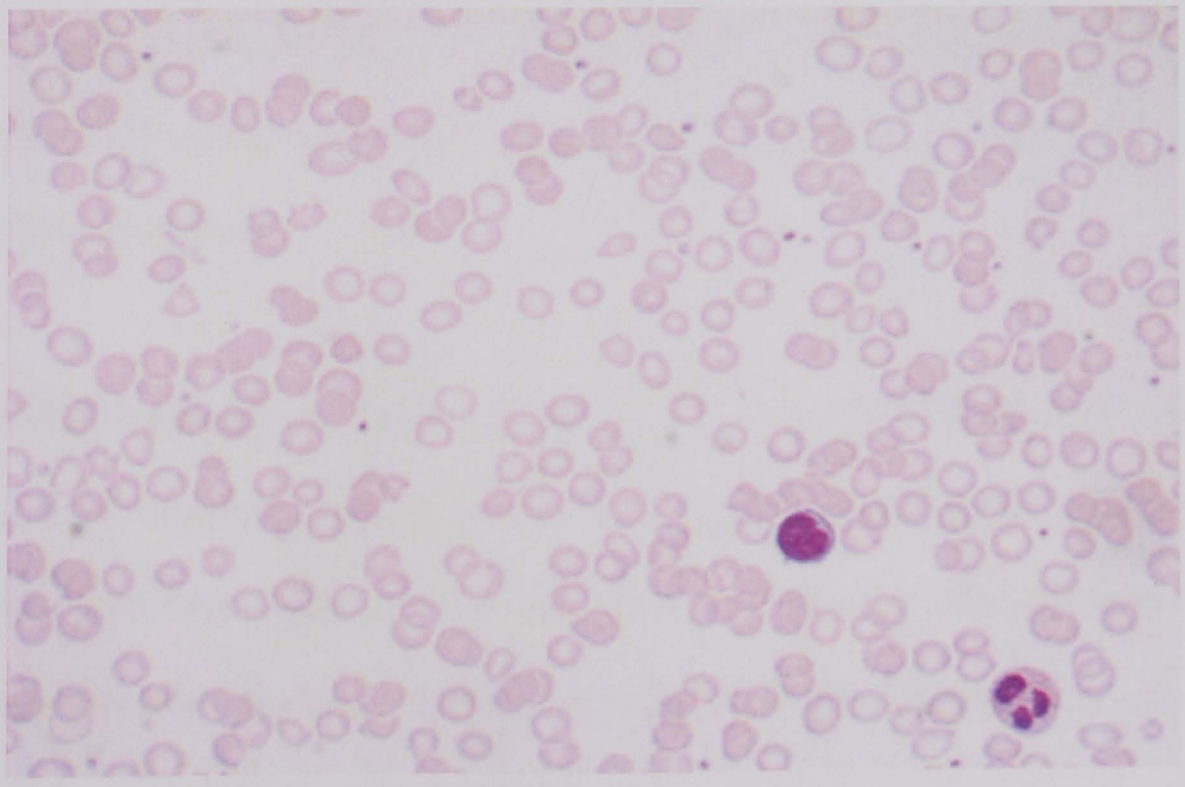
This introductory chapter firstly summarizes the clinical features of the acute and chronic leukaemias. A literature review then follows discussing what is already known about the epidemiology of leukaemia as it pertains to gender, age, geographical region, socioeconomic class and trends over time. The associations of these factors with both the incidence of and mortality from leukaemia are discussed. The available literature on aetiological factors implicated in leukaemia incidence and mortality is then reviewed. The aims and rationale of this project are then presented, and the chapter concludes with an outline of the contents of subsequent chapters.

Figure 1.1: Differentiation of Haematology Cell Lines



© Wikipedia

Figure 1.2: Normal Peripheral Blood Film



© www.imperial.ac.uk

1.1 CLINICAL FEATURES

1.1.1 Acute Lymphoblastic Leukaemia (ALL)

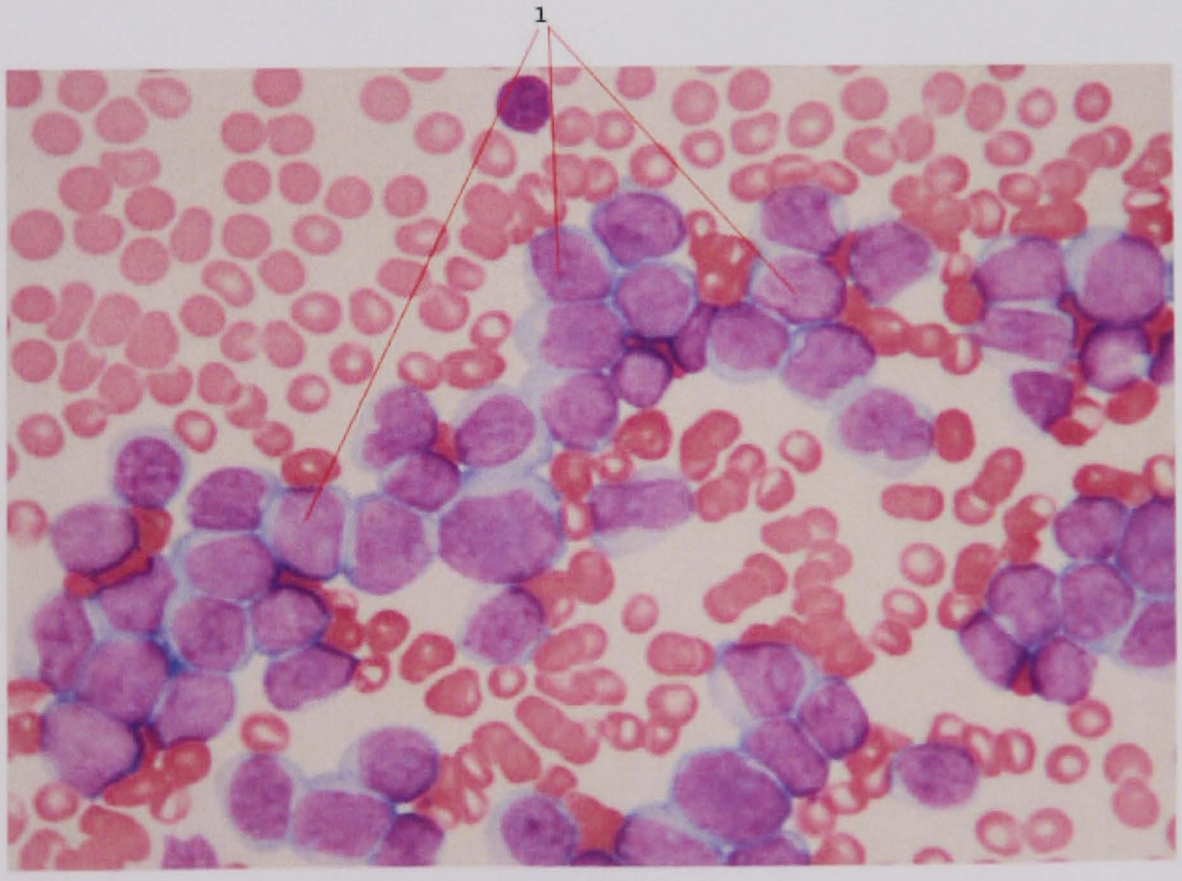
Pathogenesis. ALL results from a clonal proliferation of B or T blast cells, in the earliest stages of lymphoid maturation. Blast cells infiltrate the bone marrow and are also present in the peripheral blood[2]. Normal cell differentiation is shown in Figure 1.1 and a normal peripheral blood film is shown in Figure 1.2.

Clinical findings. Clinical manifestations are mainly those of bone marrow failure secondary to bone marrow infiltration. The symptoms include: tiredness; bruising and bleeding; bacterial infections; bone pain; enlarged lymph nodes; headache and vomiting. Physical findings include: pallor; purpura and bruising; lymphadenopathy; hepatosplenomagaly; bone tenderness and fever.

Laboratory features. Anaemia; leucopaenia; thrombocytopaenia; circulating blast cells on blood film (marked with '1' in Figure 1.3) and bone marrow infiltration with blast cells are common laboratory findings. Cytogenetic analyses are also conducted as this provides important prognostic information upon which treatment regimes are based.

Treatment. ALL treatment regimes used in adults and children depend on risk stratification, which incorporates an assessment of several prognostic factors in each case. Chemotherapy is the mainstay of treatment and is carried out in 4 phases: (1) induction, which clears the bone marrow of blast cells and replaces them with normal cells; (2) consolidation, which further reduces the leukaemia cell burden; (3) Central Nervous Systems (CNS) prophylaxis, which aims to prevent CNS involvement; and (4) maintenance therapy, which is given over 2-3 years. Patients who are deemed to be in poor risk categories at presentation, as well as those who relapse, are considered for allogeneic bone marrow transplantation[2].

Figure 1.3: ALL Peripheral Blood Film



1.1.2 Chronic Lymphocytic Leukaemia (CLL)

Pathogenesis. CLL is characterised by the accumulation of small mature B lymphocytes in the peripheral blood, bone marrow and secondary lymphoid organs[3].

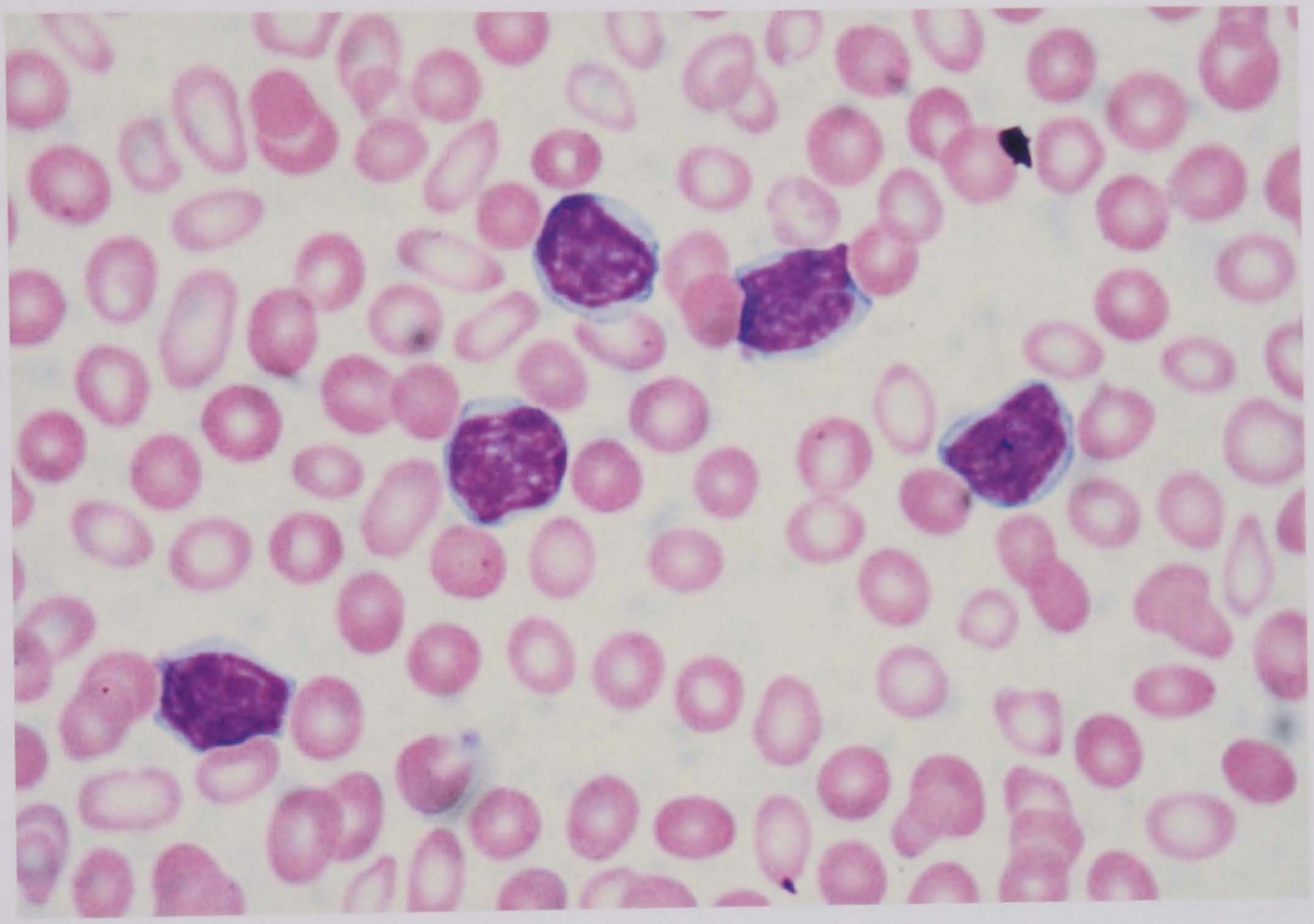
Clinical features. While most patients are asymptomatic at diagnosis, those with symptoms may present with: tiredness; bleeding and bruising; bacterial infections; night sweats; and/or fever and weight loss. Lymphadenopathy and splenomegaly are the main physical findings. Patients may have associated autoimmune conditions such as autoimmune haemolytic anaemia (AIHA) and autoimmune thrombocytopenic purpura (ATP)[2].

Laboratory features. These include monoclonal lymphocytosis; anaemia; thrombocytopaenia; bone marrow infiltration and hypogammaglobulinaemia. Peripheral blood film shows small mature lymphocytes with fragile membranes to the extent that they are damaged during the preparation of slides, leading to their appearance as 'smudge cells', as seen in Figure 1.4[3]. Cytogenetic investigations are also undertaken and provide some prognostic information[2].

Treatment. While asymptomatic patients do not require treatment, progressive disease is treated with chemotherapy. Antibody therapy is

sometimes combined with chemotherapy, and bone marrow transplantation is sometimes undertaken in younger patients[2].

Figure 1.4: CLL Peripheral Blood Film



© commons.wikimedia.org

1.1.3 Acute Myeloid Leukaemia (AML)

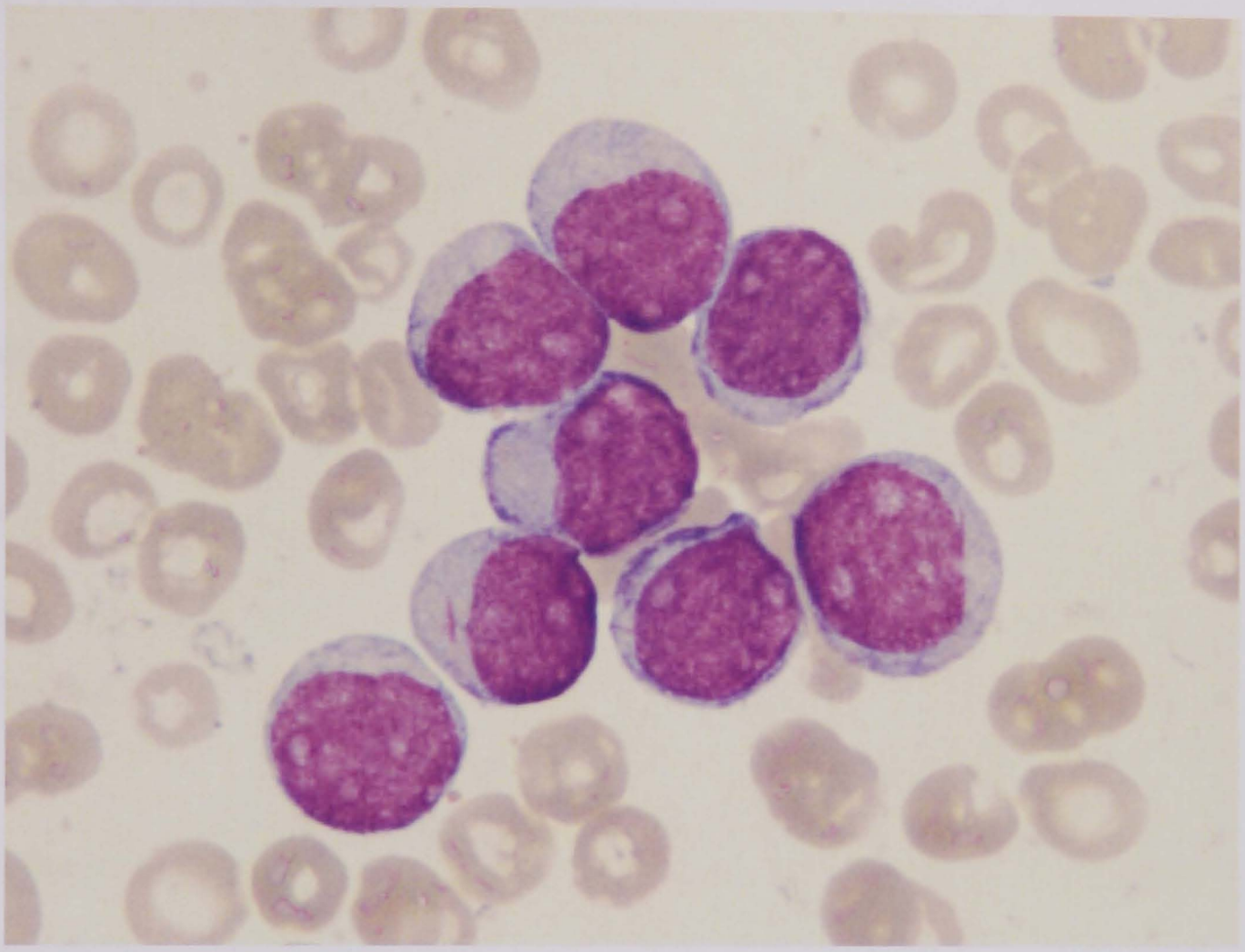
Pathogenesis. AML is a clonal disorder of myeloid precursor cells that leads to infiltration of the bone marrow with immature cells. This, in turn, results in impaired neutrophil, platelet and red cell production. Blast cells also appear in the peripheral blood[2].

Clinical features. Again, symptoms and signs of bone marrow failure, as described above, occur. Blast cell infiltration of other organs can also occur, causing gum hypertrophy, skin lesions, lymphadenopathy and splenomegaly.

Laboratory features. Anaemia, thrombocytopaenia, disseminated intravascular coagulopathy (DIC), leucocytosis, leucopaenia and blast cells in the bone marrow (large, darkly stained cells seen in the centre of Figure 1.5) are among the laboratory features. Cytogenetic features may again provide prognostic information.

Treatment. There are 3 main components to treatment: (1) intensive chemotherapy to induce remission; (2) supportive care, which is inevitably required to manage the neutropaenia and mucosal inflammation that follows chemotherapy; and (3) bone marrow transplantation for patients in poor risk disease groups and for those who have relapsed[2].

Figure 1.5: AML Bone Marrow



1.1.4 Chronic Myeloid Leukaemia (CML)

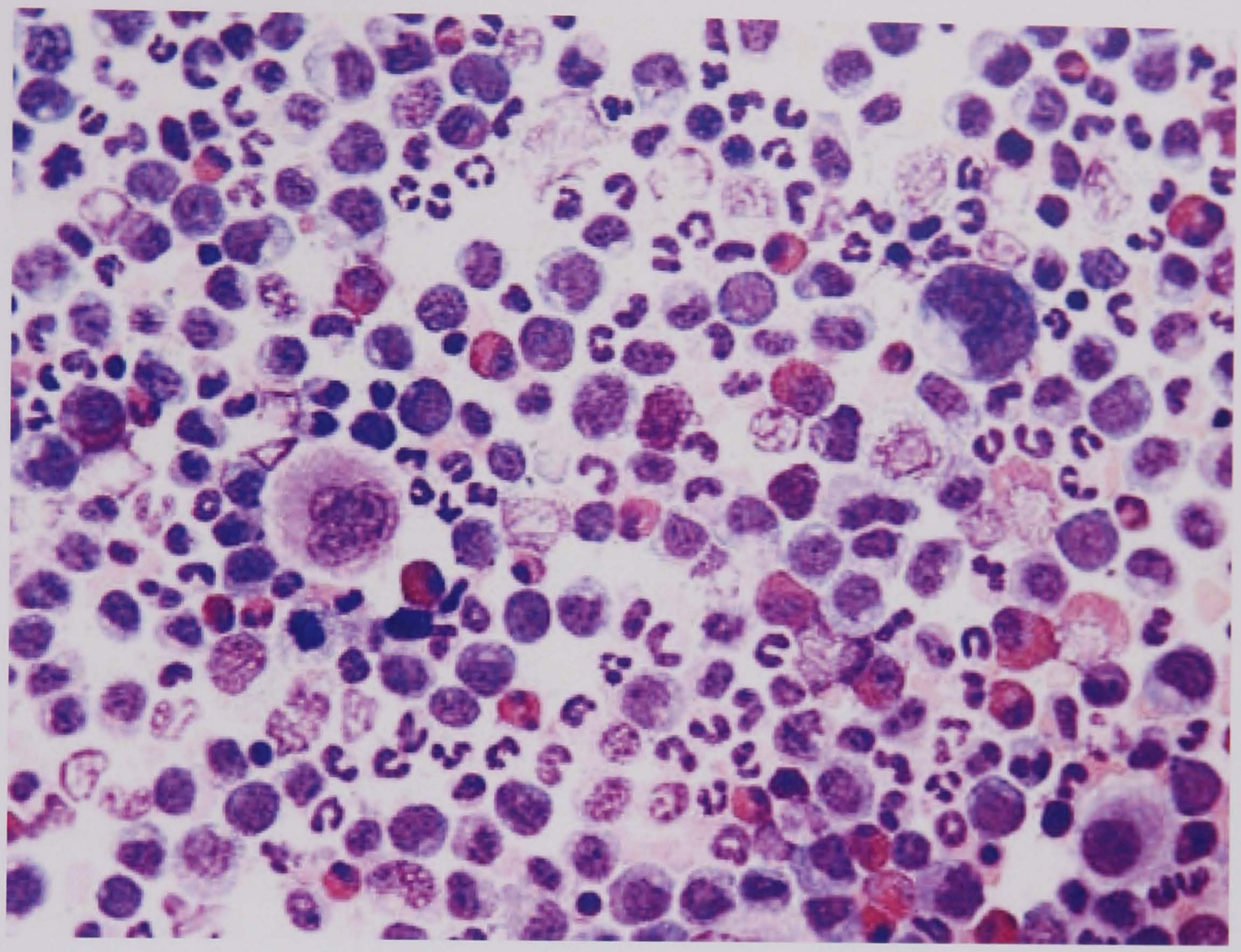
Pathogenesis. CML is a clonal disease in which a genetically altered stem cell proliferates, generating a population of differentiated cells that gradually replaces normal haemopoesis and expands the total myeloid mass[4].

Clinical features. Symptoms include: fatigue; weight loss; sweats and anorexia, while pallor and splenomegaly are the most common physical findings.

Laboratory features. Laboratory findings include: anaemia; leucocytosis; abnormal platelet counts; myelocytes, metamyelocytes and basophils in the peripheral blood (see Figure 1.6); and hypercellular bone marrow with increased white blood cell production.

Treatment. Treatment entails chemotherapy, which reduces the white cell count to normal; α-interferon which also controls the blood count; and bone marrow transplantation, which is the only known curative therapy for CML[2].

Figure 1.6: CML Blood Film



© emedicine.medscape.com

1.2 LITERATURE REVIEW

1.2.1 Search Strategy

A search of the available peer-reviewed literature was conducted at the start of this project to identify publications examining the epidemiology of leukaemia. The following three databases were searched as far back as 1950: CINAHL, EMBASE and MEDLINE. The electronic resource 'eLibrary

Gateway' of the University of Nottingham Library was used to access publications. The search terms '(leuk$ AND epidemiology) were used, producing 361 results. Publications were then limited to those in English, and those applicable to humans, leaving 242 publications. After removing duplicates, 164 remained for review. An additional search using the terms '(poverty AND leuk$)' was also conducted, and limited in the same way. This produced an additional 95 publications. Finally a third search, of the terms '((socioeconomic AND deprivation) AND leuk$)' was conducted. This search was also limited to English language publications and applicability to humans, and produced 4 additional results once duplicates were removed. A total of 263 publications were therefore reviewed and the relevant ones are summarised below.

Publications reporting the associations of gender, age, socioeconomic class, trends over time and geographic region, with incidence and mortality are discussed first, and these are presented by leukaemia sub-type. A summary of the literature as it pertains to a range of environmental factors suspected of being of aetiological importance in leukaemia then follows.

1.2.2 Results

1.2.2.1 ALL

INCIDENCE. ALL is the commonest childhood cancer, but it may present at any age[2]. Incident ALL is believed to show a slight male preponderance and

there is wide regional variation in its incidence across the world[5 6]. Although a registry-based study specifically examining sex ratios and risks of haematological malignancies, found no gender differences in children under 10 yrs, the overall incidence of ALL in the UK was greater in men than women[7]. Another later UK study of children aged 0-14 years also found a male excess in ALL incidence, although the statistical significance of the gender difference was not tested[8]. A recent study of ALL incidence among teenagers and young adults in England found neither an association with socioeconomic class, nor significant regional variation in incidence within the UK[9]. This was consistent with findings from a case-control study of 0 to 14 year olds in the UK, which also showed no association with socioeconomic class[10]. Reported trends over time suggest that annual increases until about two decades ago were followed by a subsequent plateau in incidence in children[6 11].

MORTALITY. The prognosis of ALL is better in girls than boys, and better in children than adults, with around 70% of children and 40% of adults surviving long-term[12]. A review of children treated in UK ALL trials from 1972 to 1990, found that although results of treatment have improved over time, survival was significantly better in girls than boys, independent of age and white cell count at presentation[13]. A similar survival advantage was seen over time, and in girls, in a European study, although this study did not adjust for white cell count at presentation[14]. Studies also demonstrated significant inter-country variation in survival within Europe[14 15] with better survival in Nordic countries than in eastern Europe, but there was no statistically significant regional

variation in 1- or 5-year survival between regions within the UK in the period 1986 to 1990[16]. A UK study examining survival in childhood ALL diagnosed between 1971 and 1990, using national cancer registry data, found no statistically significant socioeconomic class gradient in survival[16].

1.2.2.2 CLL

INCIDENCE. CLL is the most common leukaemia in the Western world and mainly affects people aged over 50 years[2]. CLL is believed to be twice as common in men as in women at any given age[17], and its incidence is known to increase with age[3]. In the UK, a recent registry-based study confirmed a male excess specific to people in the 5th and 7th decade[7]. There is also large regional variation in incidence, being rare in China and Korea, and virtually absent in Japan[3]. Regional variation in incidence in the UK has also been noted in a study conducted in the mid 1980s[17]. The incidence appears to be increasing as measured by registration rates[6], although an earlier review found incidence rates increased between 1945 and 1957 and then showed a declining trend since 1968[11]. The association of socioeconomic class with CLL has not been investigated in the UK.

MORTALITY. While the natural history is very variable, the majority of patients survive without symptoms or progression, such that more than half die of unrelated causes[17]. A recent review estimates survival ranges from months to decades, with a median survival of 7.5 years[3]. Age per se is not a

prognostic factor, since the clinical course of the disease is identical in younger and older patients. The causes of death are CLL-related in the younger patients, and related to second malignancies or are unrelated in the elderly[3]. Better rates of remission have been achieved over recent years, but it is not clear whether overall survival has improved[3]. Two reviews found no trend in survival over time[6 11]. Associations between survival in CLL and socioeconomic class, and UK region have not been investigated.

1.2.2.3 AML

INCIDENCE. While AML can occur at any age, it becomes increasingly common with increasing age[18] and is the most common form of leukaemia in adults[2]. A male excess in AML incidence has been noted in the UK, particularly after the age of 55[7 18], as well as in an earlier Finnish study[19]. A registry based UK study found no association between socioeconomic class and AML incidence[7], consistent with the findings of another UK study of teenagers and young adults which also found no socioeconomic class trend in AML incidence[9]. There was, however, regional variation within the UK after adjusting for socio-economic class, with the highest incidence noted in the south-east and the lowest incidence in the east of England. Regional variation in AML incidence was also found in Finland in an earlier study[19]. The incidence of AML in the UK is believed to have increased between 1961 and 1978, but not between 1984 and 1988[6].

MORTALITY. Survival declines with increasing age[2 18], and survival beyond 1 year is rare in those aged over 70[2]. A UK study of adults with AML found no significant gender differences in 1- and 3-year survival[18]. In a study examining the effect of gender and age on survival in European children with AML, girls had better survival than boys overall, but there was no gender difference in survival under the age of 5 years[20]. These findings with regard to gender were similar across most of the European countries studied. The same study also found that survival varied from one country to another, being worse in Eastern and Southern Europe, and better in Western Europe, and that survival improved considerably over the time-period of the study (1978-1989). A subsequent study examining survival in children diagnosed between 1990 and 1994 showed that survival had continued to improve across Europe, but that the earlier pattern of inter-country variation remained[15].

1.2.2.4 CML

INCIDENCE. The incidence of CML increases with increasing age and is rare in children[2]. A greater incidence in males than females is a consistent finding within the UK, and worldwide[4 7 18]. In a study of teenagers and young adults, the incidence of CML was found to increase with increasing socioeconomic deprivation[9]. This study also found regional variation in incidence with higher rates in London and the southwest than elsewhere in the UK. The incidence of CML is believed to have remained stable between 1943 and 1977 in Denmark[6]. In the UK, however, registrations have declined in the period 1984 to 1998[6].

MORTALITY. The median survival is around 4 to 5 years and depends on the response to initial treatment[4]. In a study of adults with CML in the south east of England, survival was better in those under the age 65 than in older patients, with no survival difference between men and women[18]. A registry-based study investigating survival among European children with CML found survival in girls was significantly better than in boys[20]. This study also found that regional variation in survival exists across Europe, and that survival has improved over the period of the study (1978-1989). Mortality has also been shown to have declined in both men and women in a review covering the period between 1970 and 2000[11].

1.2.2.5 AETIOLOGICAL FACTORS

A number of environmental and lifestyle factors have been implicated in the aetiology of leukaemia, but results of research have been inconclusive for many of them.

IONISING RADIATION. A review of the associations between environmental factors and a number of leukaemias confirmed that exposure to ionising radiation is a well documented risk factor for acute non-lymphocytic leukaemia (ANLL)[21]. The same review also concluded that while exposure to high doses of ionising radiation may precede the development of CML, no such association exists with CLL. Another recent review similarly concluded

that exposure to ionising radiation is strongly associated with AML[22] . ALL in children has been associated with in-utero exposure to ionising radiation[21].

ELECTROMAGNETIC FIELDS. Exposure to strong electro-magnetic fields has also been implicated in the development of ANLL, although the difficulty in quantifying such exposure makes any firm conclusions impossible[21 22].

ORGANIC SOLVENTS. Occupational handling of benzene and other organic solvents has been implicated in the aetiology of leukaemia. A strong association has been noted between occupational exposure to benzene and AML[22] and ANLL[23], for example. The biological mechanisms are believed to surround chromosomal aberrations seen in leukaemic cells of patients who have been exposed to these organic solvents. These aberrations are believed to play a role in activation of oncogenes, with consequent malignancy. A latency period of 10 years or more appears to exist between the start of exposure to these compounds and a diagnosis of ANLL[23]. In contrast, in a study examining the occupations of AML patients, organic solvents were not found to be associated with AML[19]. This study utilised Finnish cancer registry data, and compared the exposure of AML cases with that of registry patients with other cancers. An important weakness of this study in exploring associations with exposure to solvents and other chemicals was that occupation at the time of AML diagnosis was used. Given the latency period between exposure and leukaemia diagnosis found by others, it is likely that positive associations may have been missed as a result of using this approach.

While studies have shown an unusually high incidence of CLL among farmers in the USA, this was not seen in Scandinavian farmers[21]. Reasons for this difference may include differences in the specific chemicals farmers are exposed to, difficulty in adequately measuring, and then comparing the levels of exposure within and between studies, as well as differences in study design and methodology. There is no strong evidence to suggest CML is related to exposure to organic solvents, although they may increase the risk of other haematological malignancies such as Non-Hodgkin's Lymphoma, Hodgkin's disease and myeloma. These associations will not be explored further as they are not relevant to the research presented here.

ALCOHOL. The relationship between alcohol consumption and leukaemia has been examined before, but results have been inconsistent.

In a case-control study of 578 white men with leukaemia, alcohol consumption was positively associated with ALL, but not leukaemia overall or CML[24]. Odds ratios for ALL did not reach statistical significance, however, nor did investigators find a dose response gradient with the quantity of alcohol consumed, a pattern which may have suggested causality, had it been found.

In another case-control study of 164 case-control pairs, Pogoda et al. found that alcohol consumption was associated with a decreased risk of AML in adults, but results did not reach statistical significance[25]. The authors

acknowledged that alcohol consumption in their study population was related to higher socioeconomic class, and that their controls were also of higher socioeconomic class than their cases. This represents an important potential source of bias in their results. As only education was controlled for in the analysis, this combination of factors may have resulted in residual confounding with respect to socioeconomic class and consequently produced spuriously low odds ratios.

In another case-control study of 765 incident cases of acute de novo leukaemia in adults and 618 controls, regular drinkers had a reduced relative risk of leukaemia compared to non-drinkers[26]. When consumption patterns and different types of alcohol were examined, light and moderate beer intake was associated with a reduced risk of leukaemia that was statistically significant, while moderate or heavy wine intake was associated with an increased relative risk of leukaemia, although results for the latter did not reach statistical significance. As in the study by Pagoda et al.[25], alcohol consumption was more prevalent among controls than cases and education was the only marker of socioeconomic class adjusted for. Again this may have resulted in downwardly biased odds ratios if the controls were of higher socioeconomic class than the cases in this study. Unfortunately, the similarity or otherwise of cases and controls with respect to socioeconomic class is not reported in the paper. Selection bias related to social class may also explain the inverse association found with beer if, for example, beer intake is inversely associated with socioeconomic class and controls were of lower socioeconomic class than cases.

A more recent multi-centre case-control study exploring the associations of alcohol intake with a number of leukaemia subtypes was also inconclusive[27]. In this study of 649 cases and 1771 controls, any alcohol intake was found to be associated with a reduced risk of leukaemia overall, ALL and CLL, compared to those who never drank alcohol. The contrary was true for AML and CML. None of these odds ratios reached statistical significance, however. This study also examined the type and quantity of alcohol consumed and found an inverse association with leukaemia overall and moderate intake of all alcohol, wine, beer and spirits. Similar inverse associations were seen in ALL, CLL and AML with all alcohol. A positive association with leukaemia overall was found for high-level consumption of all alcohol, wine and beer. Again, results did not reach statistical significance. CML was positively associated with alcohol consumption at all levels, but odds ratios again did not reach statistical significance.

A review examining the evidence for an association between alcohol consumption and leukaemia, as well as other cancers, concluded that there were insufficient studies of appropriate size to draw any firm conclusions with regard to leukaemia[28].

SMOKING. A number of studies have shown an association between smoking and AML and ANLL, but not for all leukaemias. Results are not consistent, however, and have often been confounded by social class[29-33].

DIET. Animal studies and ecological studies of leukaemia in humans have suggested that diet may influence the risk of leukaemia, but this has not been confirmed in studies directly examining this association. A large study compared leukaemia incidence as recorded in cancer registry data from 24 countries with international food supply data[34]. A positive correlation, which was stronger in men than women, was found between total calorie-intake and lymphoid as well as total leukaemia. No such correlation was seen for myeloid leukaemias.

PERINATAL AND REPRODUCTIVE FACTORS. As part of a large childhood cancer study in the UK, a number of perinatal factors and their associations with childhood leukaemia were examined[35]. Hyperemesis was associated with all leukaemias, and polyhydramnios and anaemia with AML. Although babies who developed leukaemia were heavier at birth, these results were of borderline statistical significance. Babies who developed common B-cell precursor ALL were more likely to have been born to mothers who had had a previous molar pregnancy. Down's Syndrome was also shown to be strongly associated with ALL and AML, consistent with earlier reports.

Maternal medication use during pregnancy has also been examined as a potential aetiological factor in infant leukaemia. A case-control study examined medications used by 243 mothers of infants diagnosed with leukaemia[36]. Controls were selected from mothers whose children did not have leukaemia. Their use of 27 specific drugs (including several antibiotics, clomiphene and levothyroxine) as recorded in their medical records was then

compared. None of the associations seen reached statistical significance. Possible reasons for this include the small numbers of cases with each leukaemia sub-type and/or the large number of drugs examined.

PRESCRIPTION MEDICATION. The use of non-steroidal anti-inflammatory drugs (NSAIDs) has been investigated in both epidemiological studies and clinical trials as potentially reducing the risk of several malignancies. Few studies have explored the association of NSAID use and risk of leukaemia. A case-control study of 412 AML cases found NSAIDs had a protective effect on the risk of one AML subtype (M2), but results did not reach statistical significance[37]. Another study of 81 post-menopausal women with leukaemia (including 35 cases of CLL, 28 of AML and 5 of CML) found a statistically significant protective effect with NSAID use in leukaemia overall, as well as with aspirin use in leukaemia overall, AML and CLL[38]. However, both these studies relied on self-reported drug histories (in the former, subjects were required to recall their drug intake up to 10 years prior to their diagnosis) subjecting results to recall bias.

While chemotherapy is an established risk factor for AML, other prescription drugs as a potential aetiological factor in the development of leukaemia have not been investigated to any great extent. Long-term benzodiazepine use has been implicated in the incidence of AML of the M4 subtype, but results did not reach statistical significance[37]. Studies have also suggested that the use of hormone replacement therapy (HRT) may be associated with certain cancers as a result of oestrogen receptor stimulation. Given that

haemopoetic cells also express oestrogen receptors, HRT use has also been investigated as a risk factor for leukaemia. Only 2 studies were identified in the peer-reviewed literature. A US study, linking data from a cohort of women in Iowa to population-based registry data identified 63 cases of CLL over a 13-year follow-up period[39]. No association with CLL risk was found. A later study, using the same study population, but followed up for a longer time period, identified 201 cases of leukaemia, including 87 CLL cases and 74 cases of AML[40]. Again, no association was seen between current or previous HRT use and leukaemia overall, AML, or CLL.

1.3 PROJECT AIMS AND RATIONALE

Less is known about the acute and chronic leukaemias compared to other cancers, such as lung or breast cancer, for example. This is particularly the case with regard to the UK burden of disease and aetiology. As illustrated by the review above, specific aspects of the epidemiology of leukaemias in the UK, such as associations with socioeconomic class, have not been investigated before for many of the leukaemia subtypes. This gap in our knowledge, with respect to both incidence and mortality, can be addressed using general practice data. Hence the quantification of the incidence of and mortality with the acute and chronic leukaemias in the UK, producing up-to-date incidence and mortality figures for these diseases, forms the basis of this research project.

By using computerised general practice data a number of risk factors for these malignancies are also examined, including gender, age, socio-economic class, calendar time and geographical location. These initial incidence and mortality studies generated hypotheses surrounding access to treatment in certain groups, and these are examined in greater detail using hospital data.

The aetiological factors investigated here were chosen to reflect both the gaps in knowledge that currently exist, as well those research questions that could be answered using the data sources available to me. The role of alcohol consumption and smoking, for example, where published studies are limited in number, small and/or have produced inconclusive results, was chosen because it was possible to answer these questions using these data. Furthermore, general practice data have not been used to examine these factors before.

Factors such as exposure to ionising radiation, proximity to electromagnetic fields, occupational exposure to organic solvents, and diet, on the other hand, cannot be investigated using these data, and hence this was not undertaken.

Although it is possible to investigate the association between leukaemia and the use of many drugs, NSAIDs were chosen as they have been demonstrated to be protective in other cancers, and published studies investigating their role in the leukaemias have thus far been small,

underpowered and inconclusive. This research project therefore represented an opportunity to definitively answer this question.

While the over-arching aim of this research was to assess the usefulness of 'THIN' data for haematology epidemiology research, a number of questions surrounding the natural history, aetiology, and uptake of certain health services have been answered.

Results from these studies will not only to inform the planning of health care services by providing up-to-date stratified incidence and mortality figures, but also provide new aetiological insights into these conditions and identify areas for further research.

1.4 OUTLINE OF SUBSEQUENT CHAPTERS

The next chapter discusses several aspects of the methodology employed in the studies presented here. Chapter 3 then reports the incidence of and mortality from leukaemias overall in the UK, stratified by gender, age at diagnosis, Townsend Score of socioeconomic deprivation, year of diagnosis and geographical region. Chapter 4 reports the incidence of and mortality from several leukaemia subtypes in a similar way. In the light of findings from this study, bone marrow transplantation across socioeconomic class strata was then investigated using hospital data, and results of this study are reported in Chapter 5. Chapter 6 describes the findings of a case-control

study examining the use of non-steroidal anti-inflammatory drugs and their association with leukaemia incidence and mortality. This is followed by another case-control study presented in Chapter 7, which investigates the association of alcohol consumption and leukaemia incidence and mortality. The existing literature will be discussed in greater depth within the context of the individual studies reported in each chapter.

2 METHODOLOGY

This chapter discusses key aspects of the methodology specific to the research presented in this thesis. These include: available general practice datasets and more specifically,'The Health Improvement Network' (THIN); Hospital Episode Statistics (HES), the Townsend Score of material deprivation, and the Charlson Co-morbidity Index (CCI).

2.1 GENERAL PRACTICE DATA

Commercial companies have long recognised the potential of collating patient data held in UK general practices into large databases for the purpose of medical research[41]. In the 1980s companies encouraged the use of computers in general practice by publicising practice management software as well as by offering financial incentives to adopt computerisation. In the 1990s government facilitated the funding of primary care computers, and also established service targets that were financially rewarded. The evidence practices were required to produce in order to demonstrate that targets had been met was most easily obtained by interrogating practice computer systems. These factors accelerated computerisation and a concomitant increase in data collection in general practice. Databases containing anonymised patient records obtained from general practice systems now represent a hugely important medical research resource, and

have been widely used in epidemiological research[42]. More than 97% of people in the UK are registered with a general practitioner, which makes general practice databases an excellent source of data for this type of research.

2.1.1 THIN data

The dataset used for this research is 'The Health Improvement Network' (THIN) dataset which comprises data from over 330 general practices and includes 5.7 million patients, 2.5 million of whom are actively contributing data and can be prospectively followed. Prospective data collection started in September 2002, but data from practices that have held electronic records from as early as 1987 have also been included[43]. The total number of usable patients in the dataset is 5 395 612, with 2 592 133 actively contributing data on 1st July 2007 when data for these studies were extracted. Data recorded from 1987 to July 2007 have been used in these studies.

Over 35 million person-years of data are contained in the dataset and data include patient demographics, such as date of birth, gender, household size and period in database. All diagnoses made, along with referrals to hospitals and emergency visits are included, as are details of all drugs prescribed in primary care. Additional health data such as height, weight, alcohol consumption, smoking habits and blood pressure are also held. The Townsend score, a measure of socioeconomic deprivation, is also held for each patient, as are markers of pollution to which they may be exposed.

The number of patients in the dataset at mid-year compared to the UK population are shown below for previous years (1998 to 2006)[43].

Table 2.1: THIN Population

Year	UK Population	THIN Population	% Coverage
1988	57 165 777	1 307 398	2.3
1989	57 364 985	1 435 486	2.5
1990	57 567 259	1 533 168	2.7
1991	57 438 700	1 653 748	2.9
1992	57 563 100	1 740 605	3.0
1993	57 672 500	1 848 102	3.2
1994	57 797 400	1 979 507	3.4
1995	57 928 000	2 076 790	3.6
1996	58 043 000	2 163 589	3.7
1997	58 167 200	2 231 958	3.8
1998	58 305 300	2 286 916	3.9
1999	58 481 100	2 319 774	4.0
2000	58 643 200	2 348 228	4.0
2001	58 789 194	2 374 474	4.0
2002	59 207 000	2 404 968	4.1
2003	59 554 000	2 437 720	4.1
2004	59 834 300	2 435 264	4.1
2005	60 209 500	2 446 248	4.1
2006	60 533 000	2 426 358	4.0

THIN data therefore currently represent around 4% of the UK population. Since only general practices which use particular IT systems contribute data to THIN, and greater use of these particular IT systems occurs in the south-east of England, contributing practices are potentially more likely to serve patients from higher social classes than those practices that do not contribute to THIN. This has the potential to introduce a social class bias into studies that use these data. Matching for general practice in the design of case-control studies and/or adjusting for social class in analyses can minimise this potential bias. Both of these approaches have been used in this project.

THIN data are organised into 6 files: patient; medical; therapy; additional health data (AHD) data file, which contains information on preventative healthcare, tests and immunisations; postcode variable indicators (PVI); and dosage. A patient identifier links the patient data files to each other[43].

STRENGTHS AND LIMITATIONS:

Data are collected in an automated way during the routine activities of a general practice, thus not interfering in the delivery of care to patients. This also means that the data recorded reflect 'real life' practice. Since information is continually up-dated, investigations can be conduced into newly marketed drugs or recently adopted diagnostic tests or other health technologies used in general practice, for example. Access to these established data also allows for the rapid conduct of a range of study designs. Furthermore, the

selection of population controls at the same time as cases in the conduct of case-control studies is also facilitated.

A complete computerised record of a patient's healthcare is built up over the time they are registered with their GP. If a general practice becomes computerised after a patient has registered, events considered medically important by the GP are entered onto the electronic system from the paper notes. Similarly, if a patient transfers from a practice that does not contribute to THIN to one that does, only medically important events are likely to be recorded in the electronic record at the new practice. As such, a patient will typically only have a complete electronic record in THIN for a part of their life. Data are recorded for the purposes of patient and practice management rather than for research, and hence will reflect those data that are considered relevant to the patient's care.

THIN data also include the dates a patient registers and leaves a practice, which allows follow-up period to be taken into account in study designs. In terms of demographic information, THIN data are anonymised to the extent that names, addresses, NHS numbers and exact dates of birth are excluded. Gender and age are, however, included. A unique household identifier links patients who live at the same address, or who are members of the same family, provided they are registered at the same general practice.

Medical conditions and symptoms reported to the GP are recorded electronically during the consultation. A hierarchical system of codes (Read

Clinical Classification version 2) that can be cross-referenced to the International Classification of Disease (ICD) is used for this purpose. Referrals to secondary care are also recorded. Information from secondary care and other information received by the practice is transcribed and entered retrospectively. Where information is entered retrospectively, only those aspects considered significant and relevant are likely to be entered, which represents a potential source of weakness with regard to completeness of the patient's electronic record. Symptoms are more likely to be recorded if a prescription is issued as a result of the consultation. Furthermore, illnesses such as common colds and headaches, which are not routinely consulted for will not be recorded in THIN (or other general practice databases) and hence prevalence estimates of these conditions based on such data will be spuriously low.

The diagnostic validity of general practice data has been demonstrated for a range of malignant diseases, including breast cancer and non-melanoma skin cancer[44-46]. The validity of non-malignant diagnoses, including gastrointestinal disease, liver disease, autism, venous thromboembolism, as well as chronic obstructive pulmonary disease (COPD), has also been demonstrated[44 47-50]. The quality of prescription data within these databases has also been the subject of investigation, and have been shown to be accurate[51 52]. These validation studies used direct observation of medical records as well as questionnaires sent to GPs to assess validity.

Two popular methods of testing validity, namely direct observation of records and 'face validity', have both been used to assess the validity of THIN data. The former method was used in a study that specifically examined the validity of non-melanoma skin cancer (NMSC) diagnoses in THIN[46]. Patients with this diagnosis were firstly identified in THIN. Questionnaires were then sent to the GPs of a random sample of 40 of these patients asking how their diagnoses were confirmed. 37/40 (93%) had had their diagnoses confirmed by a letter from the hospital or a pathology report, and 2 had left the practice before the diagnosis could be confirmed. Gribbon et.al. have used the 'face validity' method in an incidence and mortality study of idiopathic pulmonary fibrosis and sarcoidosis in the UK[53]. They demonstrated that incidence and mortality rates in THIN were consistent with what previous studies, using alternative data sources, had found.

While the other published studies of diagnostic validity in general practice databases tended to use the GPRD, the similarities between these datasets in terms of the data collection methods and quality control standards employed, can be expected to result in similar levels of data accuracy. This has in fact been tested in a series of case-control studies examining several disease-disease and disease-drug associations in both THIN and the GPRD. This examination of 'face validity' concluded that THIN data were as valid as GPRD data[52].

Prescriptions issued in general practice are generated electronically and a copy is issued to patients who have them filled by a pharmacist. As a result,

prescriptions are well recorded in THIN. Prescriptions for controlled drugs, immunisations and those issued on home visits, however, are not done electronically, and so need to be recorded onto computer systems retrospectively. This has the potential to result in under-recording of prescriptions for drugs issued under these circumstances.

Similarly, drugs prescribed in secondary care will not be recorded in THIN unless further prescriptions need to be issued by the general practitioner, or they are of sufficient relevance to the GP's management of the patient so as to be recorded retrospectively. Psychiatry patients may obtain their medication from community mental health teams, while contraception may be obtained from family-planning clinics, increasing the likelihood that these medications are also under-recorded in THIN. While these are potential weaknesses of all general practice databases, they are not likely to have impacted on the studies presented here, as they do not involve such prescriptions.

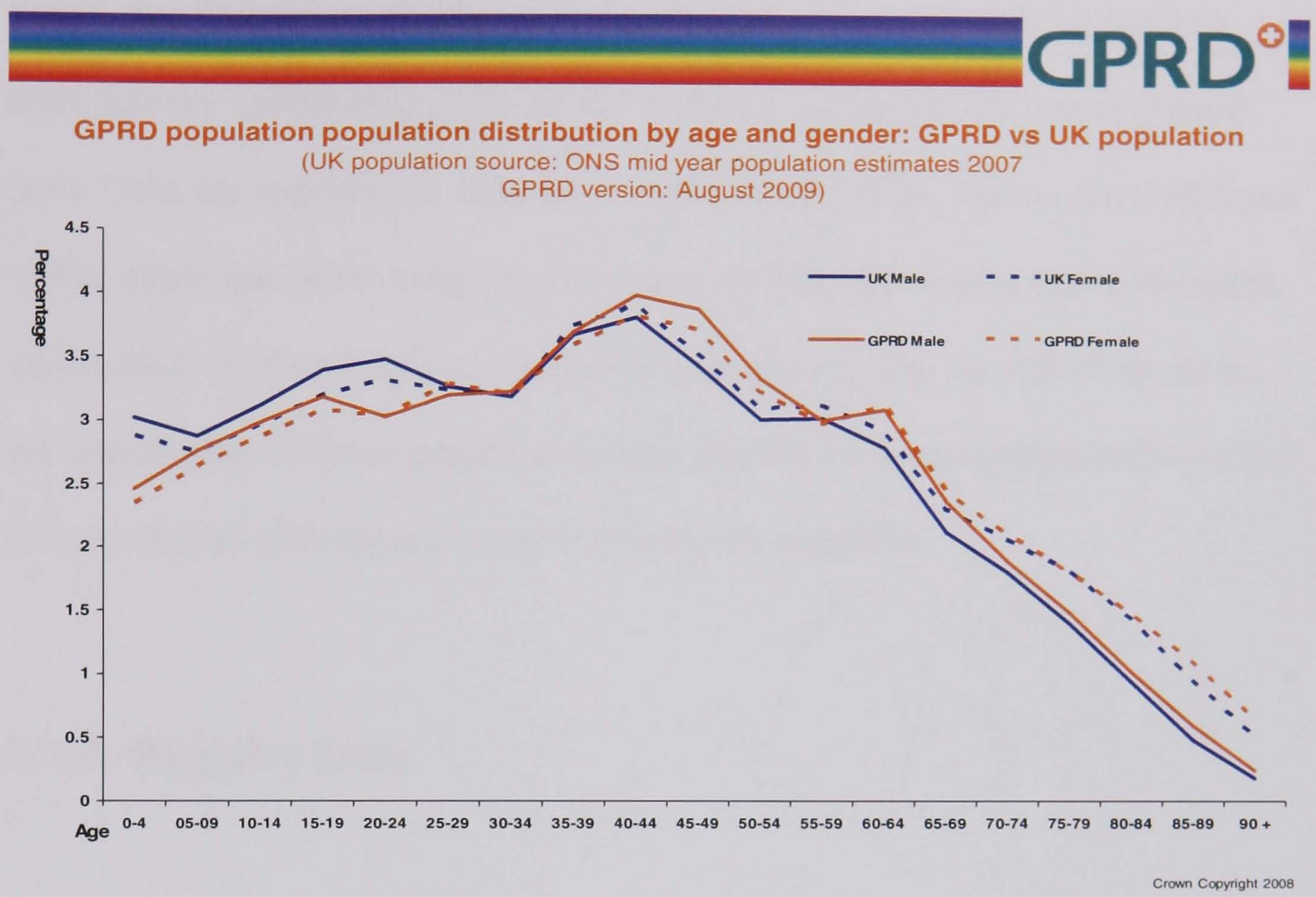
Non-prescription drugs (i.e. those obtained 'over the counter') are also not recorded in GP prescription databases. Children, people over the age of 65, and certain subgroups of patients such as pregnant women and diabetics, are entitled to free prescriptions, however, and so there exists a financial incentive to obtain drugs that can be bought 'over the counter', on prescription instead. Results of studies of these patient groups are therefore less likely to be biased by this weakness. This potential under-recording is relevant to one of the studies presented here involving non-steroidal anti-

inflammatory drug (NSAID) use, and will be discussed in greater detail in the relevant chapter. Prescription data held in THIN reflects prescriptions issued, not those filled, nor whether the prescribed medication has actually been taken by patients or not. This weakness is common to all general practice databases.

2.1.2 GPRD

The GPRD is a similar dataset to THIN in origin, content and format. The GPRD contains data from around 500 practices and represents about 5.5% of the UK population[54]. There are over 4 million patients currently actively contributing to the dataset, which contains 63 million person-years of data. This dataset is also reported to be broadly representative of the UK population. Published data in terms of gender, at least, supports this assertion (see Figure 2.1 below), although the distribution by geographic region and social class compared to that of the UK population, for example, is less clear.

Figure 2.1: GPRD Population



Although linkage of the GPRD to socioeconomic class (as well as to Office of National Statistics data, cancer registry data, and Hospital Episode Statistics) is currently being undertaken, less than half of contributing practices are being linked in this way [personal communication]. This represents a significant disadvantage of this dataset compared to THIN, in the conduct of research such as that presented here where the focus is on socioeconomic class.

2.1.3 Qresearch

Over 600 UK practices that use the EMIS clinical computer system contribute data to the Qresearch database[55]. As with the above datasets, data from both actively contributing patients are included along with historical patient data. Data are reported to date back to the early 1990s. Again, the limitations of this database are in keeping with those of other general practice datasets. Information made publicly available by Qresearch gives no indication as to the distribution of these practices across the UK, or how representative of the UK population with regard to demography the data are.

2.1.4 Registry Data

While registry data make a valuable contribution to our knowledge, there are areas where these data are incomplete. Although post-code data are included in registry data, the socioeconomic class of patients is not held. While this limitation can be overcome by mapping post-codes to socioeconomic class, other limitations make registry data unsuitable for the range of studies presented here. Information on co-morbid illness is not held, for example, making adjustment for co-morbidity difficult without linking registry data to other datasets. Registry data also do not incorporate other health data, such as smoking and alcohol consumption, which limit their use in aetiological studies that require these lifestyle factors to be taken into account. The absence of prescription data, for example, limits its value in the pharmaco-epidemiological aspects of this research. While linkage of various UK data sources are currently underway, these are far from complete. The

GPRD is currently being linked to HES data, ONS data as well as registry data, but fewer than 50% of the English practices that contribute data to GPRD, have thus far consented to having their data linked to these other data sources [personal communication]. Once this linkage is complete, the resulting dataset would be a hugely valuable resource to epidemiological research. It is not at all clear when this project will be completed though, or what the likely extent of the eventual coverage achieved will be.

There is therefore a need to utilize the best currently available datasets that *do* include this additional information to conduct epidemiological research quickly, cheaply, and efficiently in the meantime.

2.2 HOSPITAL EPISODE STATISTICS

Since bone marrow transplantation is undertaken in hospitals, hospital data were deemed more appropriate than general practice data for the study of bone marrow transplantation in AML, which is reported in Chapter 5.

Hospital Episode Statistics (HES) were introduced in 1986 and measure all hospital inpatient and day surgery activity[56]. HES data cover over 50 million people and all NHS hospitals, encapsulating 90 to 95% of all in-patient care. These are record-level data administered by The NHS Information Centre for Health and Social Care, on behalf of the Secretary of State for Health. Data are extracted from routine data flows between healthcare providers and commissioners and used to populate the 3 main HES datasets (admitted patient care, outpatient and Accident and Emergency datasets).

The admitted patient dataset, which includes inpatient and day-case records, was used in anonymous form for this research. Cumulative data are extracted quarterly for this dataset and it is updated annually, in addition. There are approximately 16 million records in this dataset for each financial year. Data from the 1989/90 financial-year are available, although the records of individual patients have only been linked since 1997/98 onwards. Data held include patients' demographic information such as date of birth, gender and region of residence, details of diagnoses and treatments received, as well as administrative details such as admission and discharge

dates, along with the place they were treated (NHS Trust or independent sector hospital, for example).

HES data are provided in tables, each of which represent one financial year, and contain all hospital episodes for that year. The contents of the tables provided are determined by the specific fields that have been requested (and paid for) by researchers. For the purposes of this research only those patients and fields that were required to address the specific research question were purchased. Data for all patients with a diagnosis of AML (as defined by specific ICD-9 and 10 codes provided to the data supplier) recorded in HES were requested. The data fields requested were age, month and year of birth, a patient identification number specific to HES (HESID), start and end dates of each hospital episode, all recorded diagnoses and procedures for each of these episodes, the health authority of residence of the patient, and the Lower Super Output Area (a measure of socio-economic deprivation) of the patient.

STRENGTHS AND LIMITATIONS:

HES data are anonymised and contain no patient names. Furthermore, access to GP and consultant codes is restricted to preserve anonymity where the small size of sub-sets of data, for example, may result in the identification of people. This limitation had no bearing on the results presented here.

As with all healthcare datasets, changes in geographical boundaries may contribute to fluctuations in data. While this may have affected data used in this research, the impact will have been random and is therefore unlikely to have biased results.

Unlike general practice data, HES data do not include a denominator population. This presented no problems since those aspects of this research using HES data did not require a general denominator population.

Inaccuracies with respect to coding are a further potential problem with all databases. While this is likely to have affected this research, errors were again likely to be random, occurring across all socioeconomic classes and so are not likely to have biased these results.

HES data provided to researchers are reported to have a lag time of between 9 and 12 months, representing the time taken for data to be extracted from hospital systems, cleaned by The NHS Information Centre, and then provided to researchers.

Non-return of data is a further potential limitation of HES data. While the NHS Information Centre liases closely with the more than 300 NHS organisations and treatment centres to encourage them to submit complete and valid data, submitted data are then also subject to 'cleaning' processes to minimise the effect of missing and invalid data. Although data have improved over time, some inaccuracies are bound to remain.

The diagnostic validity of HES data has not been investigated to nearly the extent that general practice datasets have been. There is, however, some evidence that the accuracy of three-digit ICD-9 codes are between 86 and 91% for well-recognised acute conditions, and that the accuracy of procedure codes (first 2 digits of OPCS-4) is also very good[57]. Diagnostic codes were also found to be improving over time as the incentives for collecting accurate data change and hospital information systems become more sophisticated[57]. Another study found that procedures recorded in HES were consistent with those recorded in clinical datasets for several cardiovascular and colorectal cancer surgical procedures[58].

2.3 TOWNSEND SCORE

Townsend Score of material deprivation is the measure of socioeconomic status recorded in THIN, and hence used in these studies. The score is based on a combination of four variables namely: unemployment; car ownership; home ownership and overcrowding, which produce a ranking of a particular small geographic area (of about 150 homes, called a Lower Super Output Area) relative to others[59]. The data used to calculate the Townsend Score are derived from the 1991 census. Unemployment is determined by the percentage of economically active residents aged 16-59/64 who is unemployed. Car ownership and home ownership are defined as the percentage of private households who do not possess a car, and are not owner-occupied, respectively. Overcrowding is defined by the percentage of

private households with more than one person per room. The Townsend Score is the summation of the standardized scores (z scores) for each variable. The greater the Townsend Score the greater the degree of deprivation in the area. Although the Townsend Score assigned to an individual in the dataset is not an assessment of that individual's own socio-economic circumstances, it represents the small homogenous socio-geographic area, comprising about 150 homes, in which they live.

2.4 CHARLSON CO-MORBIDITY INDEX

Co-morbid illnesses not only influence the diagnosis and management of cancers, they also predict survival in several malignancies[60]. Studies presented here have utilised the Charlson Co-morbidity Index (CCI) when adjusting for patients' co-existing medical conditions.

The Charlson Index was developed about 20 years ago and was based on the 1-year mortality of medical patients admitted to a North American hospital. The scoring takes into account the presence of 19 different medical disease groups, each of which carries a weight ranging from 1 to 6, depending on the relative risk of death within 12 months associated with the presence of the particular disease group. The disease groups utilised and their weights are shown in Table 2.2.

Table 2.2: Charlson Co-morbidity Scoring

CONDITION	WEIGHT
Myocardial infarct	1
Congestive cardiac failure	1
Peripheral vascular disease	1
Cerebrovascular disease	1
Dementia	1
Chronic pulmonary disease	1
Connective tissue disease	1
Ulcer disease	1
Mild liver disease	1
Diabetes	1
Hemiplegia	2
Moderate or severe renal disease	2
Diabetes with end-organ damage	2
Any tumour	2
Leukaemia	2
Lymphoma	2
Moderate or severe liver disease	3
Metastatic solid tumour	6
AIDS	6

While the Charlson Co-morbidity Index was initially validated in breast cancer, it has since been validated in a range of malignancies, and for

predicting mortality over periods of time longer than just 12 months[60]. The Charlson Co-morbidity Index has been found to have good reliability and to correlate well with mortality and progression-free survival outcomes. It can also be easily modified to account for age, for example. Limitations of the index include: co-morbid conditions taken into account are limited to 19 medical disease groups; non-malignant haematological conditions, such as anaemia, for example, are excluded; and its ability to predict outcomes within 6 months is reduced.

For these studies, all the diagnoses within each of the 19 groups utilised by the CCI were compared to the medical histories of the patients recorded in THIN. A co-morbidity score was then calculated for each patient and for each of the controls.

3 INCIDENCE AND MORTALITY OVERALL

3.1 ABSTRACT

Introduction: Although leukaemia constitutes 2.5% of all newly diagnosed malignancies in the UK, there is a paucity of accurate up-to-date data on the variation of its incidence with gender, age, socioeconomic status, calendar time and geographical region. General practice data were used to study this as it offered several advantages over cancer registry data. **Aims and Objectives:** The quantification of the incidence of and mortality from leukaemia in the UK and its variation with gender, age, socioeconomic status, year of diagnosis, and geographical region. **Methods:** General practice data from 'The Health Improvement Network' (THIN) dataset was used. A list of READ codes was used to identify cases of leukaemia in the dataset and steps were taken to ensure only incident, rather than prevalent cases were included in the analysis. Denominator data were also derived from THIN. Variables of interest were age at diagnosis, gender, Townsend Score, year of diagnosis and health authority. Crude incidence rates, as well as incidence rate ratios stratified by the variables of interest were calculated. The median survival and hazard ratios for death with leukaemia were calculated. All analyses were conducted using STATAv9. **Results:** 4 162 cases of leukaemia were identified, 2 314 (56%) of whom were male. 267 (6.4%) cases were diagnosed in children aged under 20, of whom 147 (55%)

were male. The overall incidence of leukaemia is 11.25 per 100 000 person-years, and 3.22 per 100 000 person-years in children aged under 20yrs. The age at which leukaemia is diagnosed shows a bimodal distribution with a peak in incidence in children aged 5 and under, and a further peak in the 8th decade of life. The relative risk (RR) of leukaemia in women is 0.66 ($p<0.01$, 95% C.I. 0.62-0.70) compared to men, when adjusted for age-category, Townsend score, health authority and year of diagnosis, but there is no gender difference in those under the age of 20yrs.The RR of leukaemia increases with increasing age (p for trend<0.01), but is independent of Townsend Score. The overall incidence of leukaemia has increased over time (p for trend<0.01), when mutually adjusted for all other variables of interest, but has decreased in the sub-group aged under 20yrs (p for trend<0.01). The median survival from leukaemia is 6.58 years. The risk of death is lower in women than men when adjusted for age at diagnosis, Townsend score, year of diagnosis and health authority (HR 0.88, p=0.01, 95% C.I. 0.79 - 0.97), and greater with increasing age at diagnosis (p for trend<0.01). The increased risk of death with lower socioeconomic class is of borderline statistical significance (p for trend=0.05). Mortality has remained stable over time, while some geographical variation in both incidence and mortality were found. Mortality in the sub-group of children is independent of all the variables of interest. **Discussion:** Greater mortality in lower socio-economic classes may reflect a greater incidence of those leukaemia subtypes with a poorer prognosis in these groups; more co-morbidity in these groups; and/or a class bias in access to and up-take of treatment. These factors will be investigated in greater depth in subsequent chapters.

3.2 INTRODUCTION

Leukaemia constitutes 2.5 % of all cancers and is the 12th most common cancer registered in the UK[1]. Approximately 7000 people are diagnosed with the disease and more than 4 300 people die from leukaemia in the UK each year[1]. Although statistics available through the Office of National Statistics, cancer charities and cancer registries in the UK shed some light on the variation in incidence of leukaemia by gender and age, as well as trends over time, there is a paucity of data on its association with patients' socioeconomic circumstances.

3.3 AIMS AND OBJECTIVES

This study set out to quantify the incidence of and mortality from leukaemia overall in the UK, as well as its variation with gender, age, socioeconomic status, calendar time and geographical region. Children, as a sub-group of the study population were also of interest.

3.4 METHODS

Data source: 'The Health Improvement Network' (THIN) dataset was used. The total number of usable patients in the dataset was 5 395 612, with 2 592 133 actively contributing data on 1st July 2007 when data for this study was

extracted. Data recorded from 1987 to July 2007 have been used in this study.

Case ascertainment: A list of READ codes was used to identify all cases with a diagnosis of leukaemia in the dataset (see Appendix B). I compiled this list by collating the descriptions of all ICD-10 diagnostic codes that apply to leukaemia and identifying the applicable READ codes from the READ code dictionary accompanying THIN data. Many retrospective diagnoses are entered into patient records at the time someone first joins a practice, or when a practice first starts to use diagnostic software. For this reason, cases were only included in these studies if their first ever recording of a diagnosis of leukaemia occurred at least 12 months after their general practice records were computerised. This ensured that only incident, rather than prevalent cases were identified for inclusion in the study.

Data management for cases: As mentioned in Chapter 2,THIN data are organised into 6 files: patient; medical; therapy; additional health data (AHD) data file, which contains information on preventative healthcare, tests and immunisations; postcode variable indicators (PVI) and dosage. A patient identifier links the patient data files to each other. Data for the leukaemia cases used in these studies were provided to me in the form of 6 files as mentioned above, and all subsequent data management and statistical analyses were performed by me. The files were linked to each other using the unique patient identifier included in THIN data.

The specific variables of interest required for these analyses were then created as follows: Age at diagnosis was calculated as the age at which a leukaemia diagnosis was first recorded in a given patient's record in the dataset. This was then grouped into 19, 5-year age categories with category 1 being those aged 0-5yrs and category 19 representing those aged over 90. To further enhance statistical power, age at diagnosis was further grouped into 20-year age-bands, and year of diagnosis into 5-year bands for the purposes of calculating adjusted incidence and mortality rates. Age and year-band were then included in the analysis as categorical variables, as were gender, Townsend Score and health authority. Male gender was used as the baseline in gender analyses. The 13 health authorities are those 10 defined for England by the Department of Health in July 2006[61], in addition to Scotland, Wales and Northern Ireland.

Denominator data: In order to calculate incidence and mortality rates the THIN mid-year population (1st July 2007) were used as denominator data. These were provided to me as aggregate data stratified by gender, age, Townsend score and health authority, and contained over 37 million person-years of data. This method was used in order to reduce the volume of data that would have to be handled. Processing individual person-time data would have exceeded the capacity of available computing resources.

Analysis: Crude incidence rates for leukaemia, stratified by gender, age at diagnosis, Townsend score, year of diagnosis and health authority, were calculated. Poisson Regression was then used to determine incidence rate

ratios independently for gender, age, Townsend Score, year of diagnosis and health authority. Incident rate ratios mutually adjusted for all the variables of interest were then calculated. Incidence rate ratios were then calculated in the same way for children under the age of 20 years at the time of diagnosis. The median survival from leukaemia was determined, and as well as the risk of death (Hazard Ratios) stratified by age, gender, Townsend Score, year of diagnosis and health authority. All analyses were conducted using STATAv9.

Ethics: Ethical approval for study was obtained from the Nottingham Research Ethics Committee.

3.5 RESULTS

3.5.1 Incidence

A total of 4 162 cases of leukaemia were identified, of whom over half were male. Over 6% of all cases were diagnosed in children aged less than 20 years, of whom over half were male (Table 3.1).

Table 3.1: Gender Distribution of Leukaemia Cases

	MALES		FEMALES		TOTAL
	N	%	N	%	
All Cases	2 314	55.60	1 848	44.40	4 162
Under 20yrs	147	55.06	120	44.94	267

The overall incidence of leukaemia in this study population is 11.25 per 100 000 person-years, and 3.22 per 100 000 person-years in children aged less than 20 years. Crude incidence rates stratified by gender, age category, Townsend Score, year of diagnosis and health authority are shown in Table 3.2 to Table 3.8.

Table 3.2: Crude Incidence Rates by Gender

	Cases	Person-years	I. R.*	95% Conf. Interval
Overall	4 162	36 982 494	11.25	10.91 - 11.60
Male	2 314	18 278 041	12.66	12.15 - 13.18
Female	1 848	18 704 453	9.88	9.44 - 10.34

*Incidence Rates per hundred thousand person-years

Table 3.3: Crude Incidence Rates by Age at Diagnosis

Age	Cases	Person-years	I.R.*	95% Conf. Interval
<6	131	2 282 229	5.74	4.84 - 6.81
6-10	66	2 129 032	3.10	2.43 - 3.94
11-15	38	2 010 582	1.89	1.37 - 2.59
16-20	40	1 851 851	2.16	1.58 - 2.93
21-25	34	2 537 313	1.34	0.96 - 1.87
26-30	32	2 500 000	1.28	0.90 - 1.81
31-35	53	2 746 113	1.93	1.47 - 2.52
36-40	66	2 784 810	2.37	1.86 - 3.02
41-45	103	3 179 012	3.24	2.67 - 3.93
46-50	142	2 431 506	5.84	4.95 - 6.88
51-55	199	2 271 689	8.76	7.62 - 10.06
56-60	311	2 114 208	14.71	13.17 - 16.44
61-65	391	2 195 395	17.81	16.13 - 19.67
66-70	506	1 630 679	31.03	28.44 - 33.85
71-75	590	1 444 308	40.85	37.69 - 44.29
76-80	593	1 192 679	49.72	45.87 - 53.89
81-85	460	996 708	46.15	42.12 - 50.57
86-90	279	433 901	64.30	57.18 - 72.31
>90	128	237 499	53.90	45.32 - 64.09

*Incidence Rates per hundred thousand person-years

Table 3.4: Crude Incidence Rates by 20-year Age Bands

Age	Cases	Person-years	I .R.*	95% Conf. Interval
<20	267	8 291 925	3.22	2.86 - 3.63
20-39	193	10 604 395	1.82	1.58 - 2.10
40-59	755	10 000 000	7.55	7.03 - 8.11
60-79	2 080	6 463 642	32.18	30.83 - 33.60
80^{+}	867	1 667 949	51.98	48.63 - 55.55

*Incidence Rates per hundred thousand person-years

Table 3.5: Crude Incidence Rates by Townsend Score

Townsend Score	Cases	Person-years	I.R.*	95% Conf. Interval
1	946	8 423 864	11.23	10.54 - 11.97
2	922	7 180 685	12.84	12.04 - 13.70
3	784	6 950 354	11.28	10.51 - 12.09
4	695	6 222 023	11.17	10.37 - 12.04
5	446	4 486 921	9.94	9.06 - 10.90
No Record	369	3 731 041	9.89	8.93 - 10.96

*Incidence Rates per hundred thousand person-years

Table 3.6: Crude Incidence Rates by Year of Diagnosis

Year	Cases	Person-years	I.R.*	95% Conf. Interval
1987	1	41 778	2.39	0.34 - 16.99
1988	2	286 755	0.70	0.17 - 2.79
1989	20	503 119	3.98	2.57 - 6.16
1990	74	1 112 781	6.65	5.29 - 8.35
1991	116	1 379 310	8.41	7.01 - 10.09
1992	126	1 494 661	8.43	7.08 - 10.04
1993	147	1 644 295	8.94	7.60 - 10.51
1994	188	1 787 072	10.52	9.11 - 12.13
1995	197	1 921 951	10.25	8.91 - 11.79
1996	185	2 095 130	8.83	7.64 - 10.20
1997	198	2 229 729	8.88	7.72 - 10.20
1998	243	2 316 491	10.49	9.25 - 11.90
1999	236	2 381 432	9.91	8.72 - 11.26
2000	260	2 418 604	10.75	9.52 - 12.14
2001	307	2 467 845	12.44	11.12 - 13.91
2002	362	2 520 891	14.36	12.96 - 15.92
2003	317	2 568 881	12.34	11.05 - 13.77
2004	356	2 581 580	13.79	12.43 - 15.30
2005	305	2 618 025	11.65	10.42 - 13.04
2006	350	2 617 801	13.37	12.04 - 14.84

*Incidence Rates per hundred thousand person-years

Table 3.7: Crude Incidence Rates by Year-band of Diagnosis

Year	Cases	Person-years	I.R.*	95% Conf. Interval
1987-1991	213	3 322 932	6.41	5.60 - 7.33
1992-1996	843	8 949 044	9.42	8.81 - 10.08
1997-2001	1 244	11 813 865	10.53	9.96 - 11.13
2002-2006	1 690	12 910 618	13.09	12.48 - 13.73

*Incidence Rates per hundred thousand person-years

Table 3.8: Crude Incidence Rates by Health Authority

Health Authority	Cases	Person-years	I.R.*	95% Conf. Interval
London	400	4 415 011	9.06	8.21 - 9.99
East of England	339	3 228 571	10.50	9.44 - 11.68
East Midlands	213	1 797 468	11.85	10.36 - 13.55
North East	156	1 258 064	12.40	10.60 - 14.51
North West	466	4 000 000	11.65	10.63 - 12.75
Northern Ireland	82	1 156 558	7.09	5.71 - 8.80
Scotland	261	2 242 268	11.64	10.31 - 13.14
South Central	476	4 103 448	11.60	10.60 - 12.69
South East Coast	424	3 424 878	12.38	11.26 - 13.62
South West	485	3 733 641	12.99	11.89 - 14.20
Wales	234	1 908 646	12.26	10.78 - 13.93
West Midlands	358	3 520 157	10.17	9.17 - 11.28
Yorkshire & Humber	268	2 200 328	12.18	10.80 - 13.73

*Incidence Rates per hundred thousand person-years

The results of Poisson regression showing incidence rate ratios independently for gender, age category, Townsend Score, year of diagnosis and health authority are shown in Table 3.9 to Table 3.15. Incidence rate ratios mutually adjusted for all the variables of interest are shown in Table 3.16 and Table 3.17.

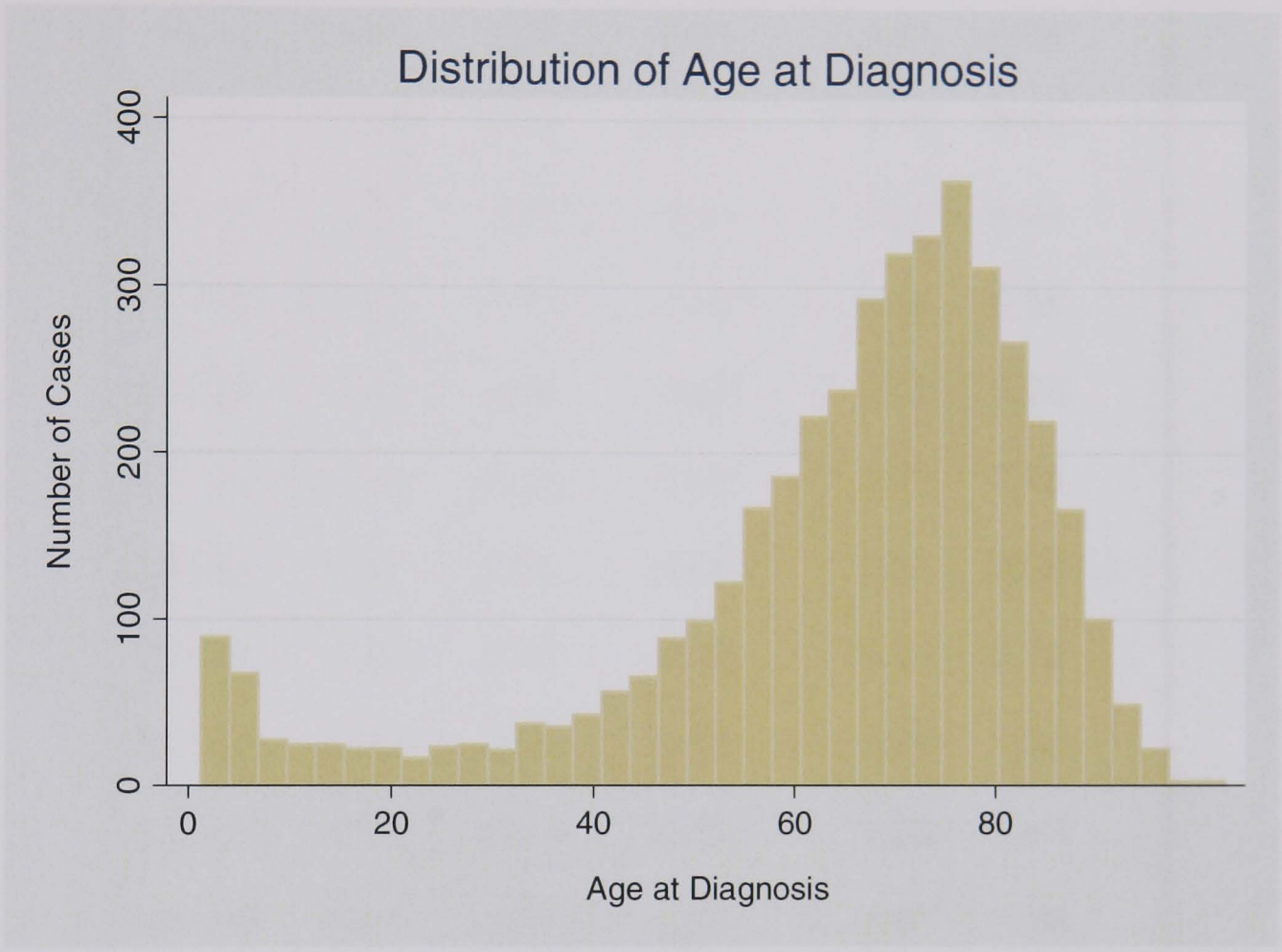
With regard to gender, the unadjusted risk of leukaemia in women is 11% lower than in men (IRR 0.79, p<0.01, 95% C.I. 0.74-0.84). When adjusted for age-category at diagnosis, Townsend Score, year of diagnosis and health authority, however, the risk is 34% lower in women than men (IRR 0.66, p<0.01, 95% C.I. 0.62-0.70).

Table 3.9: Incidence Rate Ratios by Gender

Gender	IRR	Std. Err.	P>z	95% Conf. Interval
Males	1	-	-	-
Females	0.79	0.03	<0.01	0.74 - 0.84

The bimodal distribution of age at time of diagnosis is shown in Figure 3.1.

Figure 3.1: Distribution of Age at Diagnosis



The risk of leukaemia in those aged between 20 and 40 is only 56% that of those aged less than 20 years (IRR 0.56, $p<0.01$, 95% C.I. 0.47-0.68), when adjusted for gender, Townsend Score, year of diagnosis and health authority. After the age of 40 the RR of leukaemia rises, reaching as much as 17 times that of the under 20's in those aged 80 and over (IRR 17.03, $p<0.01$, 95% C.I. 14.81-19.57), when similarly adjusted (p for trend<0.01).

Table 3.10: Incidence Rate Ratios by Age at Diagnosis

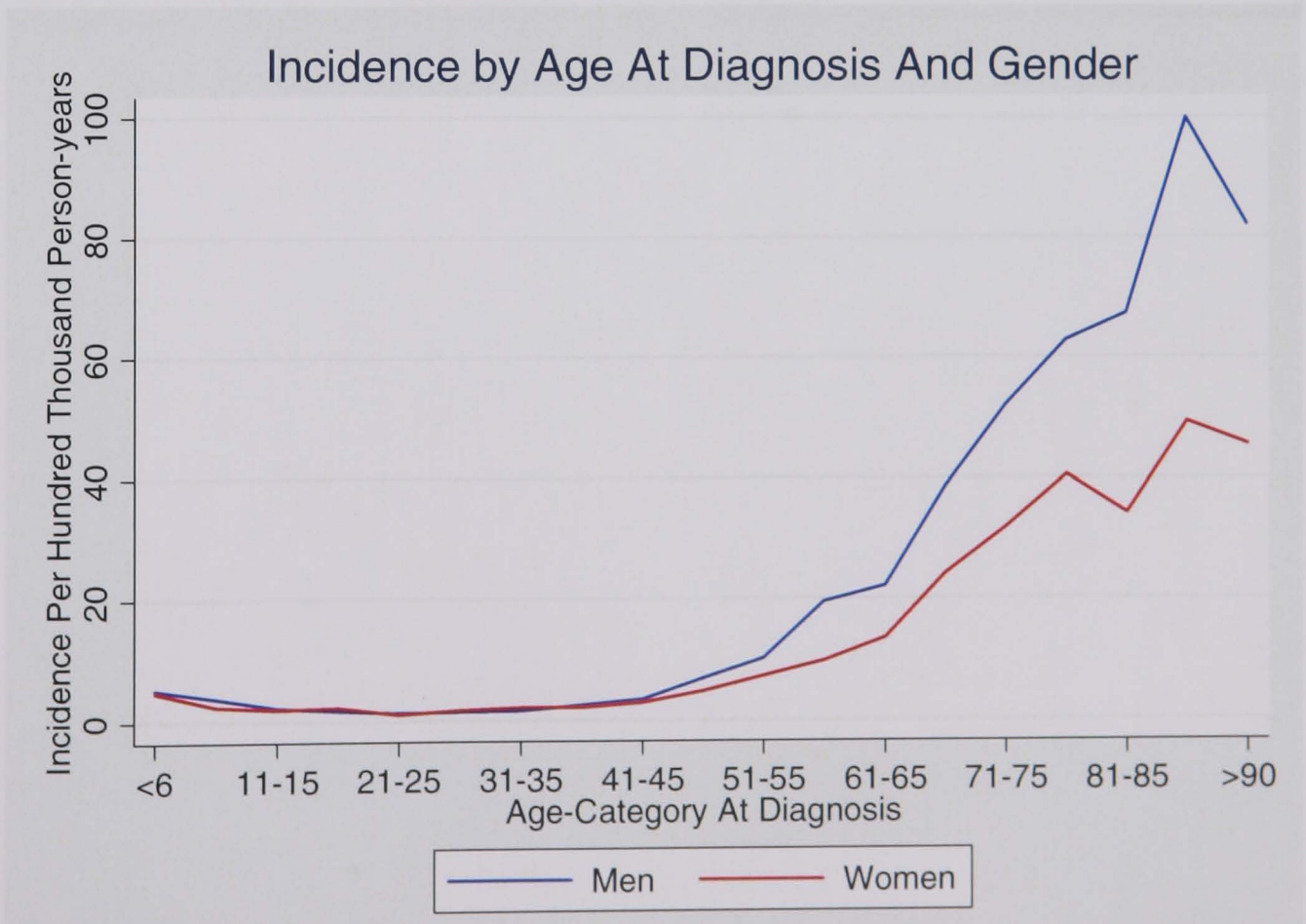
Age	IRR	Std. Err.	P>z	95% Conf. Interval
<6	1	-	-	-
6 - 10	0.55	0.08	<0.01	0.41 - 0.74
11 - 15	0.34	0.06	<0.01	0.23 - 0.48
16 - 20	0.37	0.07	<0.01	0.26 - 0.54
21 - 25	0.24	0.05	<0.01	0.16 - 0.35
26 - 30	0.20	0.04	<0.01	0.13 - 0.23
31 - 35	0.32	0.05	<0.01	0.23 - 0.45
36 - 40	0.42	0.06	<0.01	0.31 - 0.56
41 - 45	0.56	0.07	<0.01	0.43 - 0.72
46 - 50	0.99	0.12	0.93	0.78 - 1.26
51 - 55	1.52	0.17	<0.01	1.22 - 1.90
56 - 60	2.51	0.27	<0.01	2.04 - 3.09
61 - 65	3.01	0.31	<0.01	2.47 - 3.68
66 - 70	5.25	0.52	<0.01	4.32 - 6.38
71 - 75	7.01	0.69	<0.01	5.79 - 8.50
76 - 80	8.36	0.82	<0.01	6.90 - 10.13
81 - 85	7.94	0.80	<0.01	6.52 - 9.67
86 - 90	11.10	1.19	<0.01	8.99 - 13.69
> 90	9.31	1.17	<0.01	7.27 - 11.92

Table 3.11: Incidence Rate Ratios by 20-year Age-bands

Age	IRR	Std. Err.	P>z	95% Conf. Interval
<20	1	-	-	-
20-39	0.55	0.05	<0.01	0.46 - 0.67
40-59	2.29	0.16	<0.01	1.99 - 2.64
60-79	9.65	0.63	<0.01	8.48 - 10.97
80+	15.83	1.12	<0.01	13.78 - 18.18

The incidence of leukaemia increases more sharply in men than women after the age of 40 years (Figure 3.2).

Figure 3.2: Incidence by Age at Diagnosis and Gender



In terms of socioeconomic deprivation, the relative risk of leukaemia tends to decrease with increasing Townsend Score (i.e. increasing deprivation), although the test for trend did not reach statistical significance (p for trend = 0.08 across Townsend Scores 1 to 5) in multivariate analyses.

Table 3.12: Incidence Rate Ratios by Townsend Score

Townsend Score	IRR	Std. Err.	P>z	95% Conf. Interval
1	1	-	-	-
2	1.13	0.05	0.01	1.03 - 1.24
3	0.99	0.05	0.82	0.90 - 1.09
4	0.99	0.05	0.81	0.89 - 1.09
5	0.87	0.51	0.02	0.78 - 0.98
No Record	0.89	0.06	0.05	0.78 - 1.00

Table 3.13: Incidence Rate Ratio by Year of Diagnosis

Year	IRR	Std. Err	P>z	95% Conf. Interval
1987	1	-	-	-
1988	0.29	0.36	0.31	0.03 - 3.21
1989	1.66	1.70	0.62	0.22 - 12.37
1990	2.78	2.80	0.31	0.39 - 19.98
1991	3.51	3.53	0.21	0.49 - 25.15
1992	3.52	3.53	0.21	0.49 - 25.19
1993	3.73	3.75	0.19	0.52 - 26.68
1994	4.39	4.40	0.14	0.62 - 31.34
1995	4.28	4.29	0.15	0.60 - 30.55
1996	3.69	3.70	0.19	0.52 - 26.32
1997	3.71	3.72	0.19	0.52 - 26.45
1998	4.38	4.39	0.14	0.61 - 31.24
1999	4.14	4.15	0.16	0.58 - 29.50
2000	4.49	4.50	0.13	0.63 - 31.98
2001	5.20	5.20	0.10	0.73 - 37.00
2002	6.00	6.01	0.07	0.84 - 42.70
2003	5.15	5.16	0.10	0.72 - 36.70
2004	5.76	5.77	0.08	0.81 - 41.01
2005	4.87	4.88	0.11	0.68 - 34.67
2006	5.58	5.59	0.09	0.78 - 39.74

Table 3.14: Incidence Rate Ratio by Year-band of Diagnosis

Year- band	IRR	Std. Err.	P>z	95% Conf. Interval
1987-1991	1	-	-	-
1992-1996	1.47	0.11	<0.01	1.27 - 1.71
1997-2001	1.64	0.12	<0.01	1.42 - 1.90
2002-2006	2.04	0.15	<0.01	1.77 - 2.36

The incidence of leukaemia has increased over calendar time. The relative risk of leukaemia in 2002-2006 was nearly twice that of the period 1987-1991, when adjusted for gender, age-category, Townsend Score and health authority (p for trend<0.01). When only data for the past decade is analysed, the incidence has remained stable, however (p for trend= 0.09).

The relative risk of leukaemia does not vary significantly by health authority compared to London, when adjusted for gender, age-category, Townsend Score and year of diagnosis. The exception to this is in Northern Ireland where the risk is 23% lower than London (IRR 0.77, p=0.04, 95% C.I. 0.61-0.98), when adjusted for gender, age-category, Townsend Score and year of diagnosis.

Table 3.15: Incidence Rate Ratio by Health Authority

Health Authority	IRR	Std. Err.	P>z	95% Conf. Interval
London	1	-	-	-
East of England	1.18	0.09	0.02	1.02 - 1.37
East Midlands	1.27	0.11	0.01	1.07 - 1.51
North East	1.36	0.13	<0.01	1.13 - 1.65
North West	1.26	0.09	<0.01	1.10 - 1.45
Northern Ireland	0.79	0.10	0.06	0.62 - 1.01
Scotland	1.28	0.10	<0.01	1.09 - 1.50
South Central	1.26	0.09	<0.01	1.10 - 1.44
South East Coast	1.36	0.10	<0.01	1.19 - 1.57
South West	1.42	0.10	<0.01	1.24 - 1.62
Wales	1.33	0.11	<0.01	1.13 - 1.57
West Midlands	1.14	0.08	0.07	0.99 -1.32
Yorkshire & Humber	1.32	0.11	<0.01	1.12 - 1.54

Table 3.16: Mutually Adjusted Incident Rate Ratios

	IRR**	Std. Err.	P>z	95% Conf. Interval
Gender				
Males	1	-	-	-
Females	0.66	0.02	<0.01	0.62 - 0.70
Age at Diagnosis				
<20	1	-	-	-
20-39	0.56	0.05	<0.01	0.47 - 0.68
40-59	2.28	0.16	<0.01	1.98 - 2.63
60-79	9.85	0.65	<0.01	8.66 - 11.20
80+	17.03	1.21	<0.01*p<0.01	14.81 - 19.57
Townsend Score				
1	1	-	-	-
2	1.05	0.05	0.30	0.96 - 1.15
3	0.97	0.05	0.55	0.88 - 1.07
4	0.98	0.05	0.76	0.89 - 1.09
5	0.90	0.05	0.09 *p=0.08	0.80 - 1.02
No record	0.93	0.06	0.26	0.82 - 1.06
Year of Diagnosis				
1987-1991	1	-	-	-
1992-1996	1.44	0.11	<0.01	1.24 - 1.67
1997-2001	1.59	0.12	<0.01	1.38 - 1.84
2002-2006	1.95	0.14	<0.01*p<0.01	1.69 - 2.25

Table 3.17: Mutually Adjusted Incident Rate Ratios (Cont.)

Health Authority				
London	1	-	-	-
East of England	1.01	0.08	0.91	0.87 - 1.17
East Midlands	1.12	0.10	0.20	0.94 - 1.33
North East	1.19	0.12	0.07	0.99 - 1.44
North West	1.12	0.08	0.10	0.98 - 1.29
Northern Ireland	0.77	0.10	0.04	0.61 - 0.98
Scotland	1.17	0.09	0.06	0.99 - 1.37
South Central	1.00	0.07	0.97	0.87 - 1.15
South East Coast	1.12	0.08	0.12	0.97 - 1.29
South West	1.09	0.08	0.21	0.95 - 1.25
Wales	1.12	0.10	0.19	0.95 - 1.32
West Midlands	1.03	0.08	0.74	0.89 - 1.19
Yorkshire & Humber	1.09	0.09	0.29	0.93 - 1.28

**All IRRs are adjusted for all other variables in the table.

*p = test for trend across ordered categories.

Cases for whom there is no record of Townsend Score were excluded from the trend analysis of Townsend Score.

Children aged less than 20yrs: In the analysis of this sub-group of 267 children, the overall incidence of leukaemia is 3.22 per hundred thousand person-years. The risk of leukaemia is similar in males and females, when adjusted for gender, Townsend Score, year of diagnosis and health authority. Furthermore, multivariate analysis reveals an increasing trend in incidence with increasing Townsend Score (i.e. increasing deprivation), although this does not reach statistical significance (p for trend=0.07). A statistically significant decreasing trend over calendar time is also seen in multivariate analysis (p for trend<0.01). There was some variation in incidence by geographic region, with higher incidence rates in the Midlands, the North East, Northern Ireland and Wales, compared to London. (See Table 3.18 and Table 3.19)

Table 3.18: Mutually Adjusted Incident Rate Ratios (< 20 years)

	IRR**	Std. Err.	P>z	95% Conf. Interval
Gender				
Males	1	-	-	-
Females	0.97	0.13	0.84	0.76 - 1.25
Townsend Score				
1	1	-	-	-
2	1.39	0.28	0.10	0.94 - 2.05
3	1.31	0.26	0.17	0.89 - 1.93
4	1.50	0.31	0.05	0.99 - 2.25
5	1.53	0.38	0.09 *p=0.07	0.94 - 2.47
No record	1.43	0.43	0.24	0.79 - 2.59
Year of Diagnosis				
1987-1991	1	-	-	-
1992-1996	0.45	0.12	<0.01	0.27 - 0.78
1997-2001	0.35	0.10	<0.01	0.20 - 0.59
2002-2006	0.35	0.09	<0.01 *<0.01	0.21 - 0.58

Table 3.19: Mutually Adjusted Incident Rate Ratios (< 20 years) (Cont.)

Health Authority				
London	1	-	-	-
East of England	1.23	0.37	0.48	0.68 - 2.23
East Midlands	2.94	1.12	0.01	1.39 - 6.22
North East	2.37	1.01	0.04	1.03 - 5.45
North West	1.15	0.28	0.58	0.71 - 1.85
Northern Ireland	2.44	1.03	0.04	1.06 - 5.58
Scotland	1.73	0.52	0.07	0.95 - 3.12
South Central	1.05	0.30	0.87	0.60 - 1.82
South East Coast	0.93	0.27	0.80	0.53 - 1.64
South West	1.15	0.33	0.62	0.66 - 2.03
Wales	1.88	0.55	0.03	1.06 - 3.34
West Midlands	1.15	0.32	0.61	0.67 - 1.99
Yorkshire & Humber	1.63	0.51	0.12	0.88 - 3.00

**All IRRs are adjusted for all other variables in the table.

*p = test for trend across ordered categories.

Cases for whom there is no record of Townsend Score were excluded from the trend analysis of Townsend Score.

3.5.2 Mortality

The median survival from leukaemia is 6.58 years, and is more than a year longer in women than men. See Table 3.20.

Table 3.20: Median Survival by Gender

Gender	Cases	Median Survival (years)	Std. Err.	95% C. I.
Men	2 293	6.08	0.25	5.42 - 6.86
Women	1 829	7.12	0.35	6.14 - 8.62
Overall	4 122	6.58	0.32	6.00 - 7.05

Mutually adjusted hazard ratios for the risk of death from leukaemia are presented in Table 3.21 and Table 3.22. The risk of death from leukaemia is 11% lower in women than men when adjusted for age, Townsend Score, year of diagnosis and health authority (HR 0.89, p=0.02, 95% C.I. 0.81-0.98), and increases with increasing age at diagnosis (p for trend <0.01). The risk of death increases with increasing Townsend Score (p for trend=0.05), but there has been no significant change in mortality over time. The risk of death from leukaemia is greater in the East Midlands, North East, North West, South West, West Midlands and Yorkshire & Humber, compared to London as baseline, when adjusted for gender, age at diagnosis, Townsend Score and year of diagnosis.

Table 3.21: Mutually Adjusted Hazard Ratios

	HR**	Std. Err.	P>z	95% C. I.
Gender				
Males	1	-	-	-
Females	0.89	0.04	0.02	0.81 - 0.98
Age at Diagnosis				
<20	1	-	-	-
20-39	1.94	0.40	<0.01	1.29 - 2.91
40-59	2.16	0.36	<0.01	1.56 - 2.99
60-79	3.97	0.62	<0.01	2.93 - 5.38
80+	7.32	1.16	<0.01*p<0.01	5.36 - 9.99
Townsend Score				
1	1	-	-	-
2	1.08	0.08	0.31	0.93 - 1.25
3	1.31	0.10	<0.01	1.13 - 1.52
4	1.17	0.09	0.04	1.00 - 1.37
5	1.13	0.11	0.20 *p=0.05	0.94 - 1.36
No record	1.44	0.14	<0.01	1.19 - 1.75
Year of Diagnosis				
1987-1991	1	-	-	-
1992-1996	0.91	0.09	0.32	0.74 - 1.10
1997-2001	0.93	0.09	0.47	0.76 - 1.13
2002-2006	0.92	0.09	0.39 *p=0.98	0.75 - 1.12

Table 3.22: Mutually Adjusted Hazard Ratios (Cont.)

Health Authority				
London	1	-	-	-
East of England	1.26	0.15	0.06	0.99 - 1.59
East Midlands	1.32	0.18	0.04	1.01 - 1.72
North East	1.61	0.22	<0.01	1.23 - 2.11
North West	1.55	0.17	<0.01	1.25 - 1.92
Northern Ireland	0.94	0.21	0.77	0.60 - 1.46
Scotland	0.87	0.12	0.34	0.66 - 1.15
South Central	1.22	0.14	0.08	0.98 - 1.53
South East Coast	1.19	0.14	0.14	0.94 - 1.50
South West	1.38	0.15	<0.01	1.11 - 1.72
Wales	1.25	0.17	0.09	0.96 -1.63
West Midlands	1.36	0.16	0.01	1.07 - 1.72
Yorkshire & Humber	1.30	0.17	0.04	1.01 - 1.67

**All HRs are adjusted for all other variables in the table.

*p = test for trend across ordered categories.

Cases for whom there is no record of Townsend Score were excluded from trend analysis of Townsend Score.

Children aged less than 20yrs. The risk of death from leukaemia in this group is independent of gender, Townsend Score and year of diagnosis.

3.6 DISCUSSION

The overall incidence of leukaemia in this study is 11.25 per 100 000 person years. Men are more commonly affected than women overall, but the risk of leukaemia under the age of 40 years is similar in men and women. The age at which leukaemia is diagnosed shows a bimodal distribution with a peak in incidence in children aged 5 years and under, and a further peak in the 8th decade of life. The incidence of leukaemia is independent of socioeconomic deprivation. There is an increase in the incidence of leukaemia over calendar time, and some variation in incidence with geographical region.

The median survival from leukaemia is 6 years and 7 months in this study. Mortality is greater in men than in women, and increases with increasing age at diagnosis. Increasing mortality (of borderline significance) is seen with increasing deprivation. Mortality has remained stable over time, but varies with geographic region.

In the subgroup of children, however, leukaemia incidence is independent of gender and socioeconomic class, but has decreased over calendar time. Again there is some variation by geographical region. Leukaemia mortality in children is independent of gender, Townsend Score and year of diagnosis.

The overall incidence of leukaemia found in this study is consistent with the crude incidence rate of 11.7 per 100 000 population published by Cancer

Research UK for 2004[1]. The overall protective effect of the female gender on the relative risk of leukaemia, despite a similar risk until the age of 40 years, is explained by a sharper increase in incidence in men compared to women after the age of 40 (Figure 3.2). This pattern of divergence in incidence has been noted in other developed countries[6]. An infectious aetiology for leukaemia is supported by this difference in that there are known gender differences in susceptibility to infections such as polio and herpes simplex, for example. The role of hormonal influences such as the use of HRT in women is also an area for further exploration. Two previous studies that investigated HRT as a risk factor for leukaemia were inconclusive [39 40]. Both studies contained only small numbers of leukaemia cases, however. The finding here of a poorer prognosis in men than women with leukaemia is also consistent with other published data[1 62].

Studies of leukaemia and social class prior to the 1980's have mainly found higher incidences of leukaemia in higher social classes[6 63 64], in both adults and children. Since the 1980's, however, studies have consistently reported inverse associations with socioeconomic class[65]. This apparent change in direction of the association may be due to differences in study design and measures of socioeconomic deprivation that have been used over time, however. Most studies conducted prior to the 1980's were ecological and these tended to show a positive association of leukaemia with higher social class. In contrast, after 1980 most studies were case-control studies and used individual-level measures of income and education, rather than ecological-level indicators of socio-economic status[65]. This study has now

shown a lack of association of leukaemia incidence with socioeconomic class, but an impact on mortality when small-area measures of socio-economic status, such as Townsend Score, are used.

Except for Northern Ireland, this study showed no significant differences in incidence among health authorities compared to London when adjusted for gender, age-category, deprivation and year of diagnosis. Significant regional differences in mortality consistent with a 'north-south divide' are evident, however, even when adjusted for socioeconomic deprivation. A study examining nearly 4 000 cases of leukaemia in teenagers and young adults between 1979-2001, found statistically significant variation in leukaemia incidence by English region[9]. In contrast to this study though, the highest incidences of leukaemia were found in London and the south of England when adjusted for Townsend score, but results were not adjusted for gender or year of diagnosis. This difference may reflect a real change in the pattern of incident leukaemia since that study was done, or differences in study design. Regional mortality rates published by the Office of National Statistics (ONS) also show wide regional variation, but because the denominator population used by the ONS was based on 1996 UK population estimates, figures between that study and this one cannot be directly compared[62].

The initial increase in incidence over calendar time followed by a stable incidence over the last decade found in this study is consistent with those trends published by Cancer Research UK, which shows an increase in incidence until the mid 1990s followed a subsequent plateau[1]. While the

initial increasing trend may represent a true increase in leukaemia incidence, it is more likely to represent a combination of the latter with demographical changes in the population over that time, improvements in diagnostic accuracy, and improvements in cancer data recording and collation. Given that the diagnosis of leukaemia increases with age, the expansion in the elderly population over the time period in question would have contributed to an increase in the number of cases over time. Improvements in specialist medical care for the elderly would also have resulted in greater case ascertainment in this group as more asymptomatic and/or milder disease was detected by the increased use of routine blood testing. Furthermore, once blood tests have shown abnormalities, more invasive diagnostic investigations are more likely to have been carried out in this age group in recent years than in the past. With the advent of cancer registries and the development of regional specialist cancer services, there has also been better recording, collection and collation of data over time. In terms of mortality, an analysis of nearly 57 000 patients with leukaemia in the UK found that 5-year survival increased for men diagnosed over the period 1986-1999, but not for women[66]. A study reporting age-specific mortality rates for leukaemia covering the period 1910 to 1997, showed that in children under 5, mortality increased until the 1950's and then declined steadily since[11]. For those aged over 65 years, however, an increase in mortality rates was again seen until the 1950's, but the subsequent decline only continued until the 1980's. This was then followed by plateau in mortality rates.

The main strength of this study is that the large size of the study population has allowed precise estimates of the incidence and mortality of leukaemia in the UK to be calculated. By using THIN data, access to a number of variables of interest by which to stratify results has been possible, which would not have been the case had registry data been utilised. A further strength is that the denominator population was obtained from the same dataset as the cases, which has enabled accurate stratified incidence rate and hazard ratios to be determined. The methodology employed here to minimise the inclusion of prevalent cases means that incidence rates are not spuriously elevated. The risk of masking any trends over time has also been minimised by excluding prevalent cases in this way.

A further strength of this study is its design. The vast majority of previous studies have been case-control studies, with most requiring the involvement of patients and/or carers in completing interviews and/or questionnaires[65]. Case-control studies like these, while efficient for studying rare diseases, are prone to bias, including socioeconomic class bias, in the selection of cases and controls. This potentially results in disproportionately more people from higher social classes participating in studies. Using general practice data has minimised this type of selection bias, while addressing the difficulty of studying a rare disease by using appropriate statistical techniques, i.e. Poisson regression.

A potential weakness of this study is the issue of diagnostic validity in the dataset. This is unlikely to have impacted on these results given that

leukaemia is a secondary care diagnosis made only after confirmatory pathological tests. General practitioners are therefore unlikely to record this diagnosis in patients' records in error. The similarity of the disease incidence, gender distribution and trends over time presented here, with those published from registry data, provides some evidence of the 'face validity' of leukaemia in THIN. Direct examination of patients' medical records to confirm leukaemia diagnoses is possible, but this is time-consuming and expensive. Due to the limited resources available, it was not possible to examine individual patient records in this way for this project.

In summary, the incidence of leukaemia in the UK is independent of socioeconomic class, but mortality appears to be worse with increasing deprivation. One explanation for this may be that those leukaemia subtypes with worse prognoses are more common among people from lower socioeconomic classes. This issue is addressed in the following chapter, which examines the incidence of and mortality from a number of leukaemia subtypes. While different subtypes are known to have different peak ages of onset and different mortality patterns, the next study will determine whether these patterns are consistent across the spectrum of socioeconomic class.

Greater co-morbidity in patients from lower socioeconomic classes, and/or a class bias in access to treatments such bone marrow transplantation are other possible reasons for finding greater mortality rates among leukaemia patients from lower socioeconomic classes. This will also be examined in subsequent studies.

Although the incidence of leukaemia has remained relatively stable over the past decade, the number of people affected will continue to rise as the population ages. This has significant implications for the planning of future health care services.

This study has demonstrated the value of general practice data in leukaemia research, adding to the growing evidence of the value of such data in epidemiological research.

MAIN FINDINGS

- **THIN data are a valuable resource for leukaemia research.**
- **Mortality in leukaemia patients from lower socioeconomic classes is worse than among the better off.**

QUESTIONS ARISING

- **Are leukaemia subtypes with poorer prognoses more common in people from lower socioeconomic classes?**
- **Does a socioeconomic class gradient in mortality exist for all leukaemia subtypes?**

4 INCIDENCE AND MORTALITY BY LEUKAEMIA SUBTYPE

4.1 ABSTRACT

Introduction: The previous study showed that while the incidence of leukaemia overall was independent of social class, mortality was worse with increasing socioeconomic deprivation. This study will determine whether these trends hold true for all leukaemia subtypes. **Aims and Objectives:** The quantification of the incidence of and mortality from leukaemia subtypes in the UK and their variation with gender, age, socioeconomic status, geographical region, and year of diagnosis. **Methods:** General practice data from 'The Health Improvement Network' (THIN) dataset were used as before. Incident cases of 6 leukaemia subtypes of interest were identified from READ codes and denominator data were again derived from THIN. Crude incidence rates, incidence rate ratios, as well as the median survival and hazard ratios for risk of death from these leukaemia subtypes were calculated as before. **Results:** 3 226 (78%) of the 4 162 leukaemia cases were identified as falling into one of the 6 subtypes of interest. CLL is the most common and ALL the least common subtype. CLL, AML and unspecified Lymphocytic Leukaemia are more common in men. The incidence of all leukaemia subtypes increases with increasing age at diagnosis, and the incidence of all subtypes studied

are independent of socioeconomic class. All subtypes, except unspecified Lymphocytic Leukaemia and CML, show an increasing trend in incidence over time. Median survival is best in ALL and worst in AML. The prognosis for women with CLL and unspecified Lymphocytic Leukaemia is better than for men. The risk of death increases with increasing age at diagnosis (p for trend<0.01) for all subtypes. Mortality in AML increases with increasing socioeconomic deprivation (p for trend=0.01), but is independent of deprivation in all other subtypes. Mortality is independent of the year of diagnosis. There is some regional variation in both incidence and mortality.

Discussion: Similar mortality from most leukaemia subtypes across socioeconomic gradients and geographical locations suggests that on the whole, access to and uptake of diagnostic services is equal across these strata. However, poorer survival in AML patients from lower socioeconomic classes, despite no socioeconomic class gradient in incidence warrants further exploration.

4.2 INTRODUCTION

The previous study showed that while the incidence of leukaemia overall was independent of socioeconomic class, survival appears to worsen with increasing socioeconomic deprivation. Furthermore, despite little regional variation in incidence, there is wide regional variation in survival even when socioeconomic deprivation is adjusted for. This study will examine whether these findings hold true for all leukaemia subtypes.

4.3 AIMS AND OBJECTIVES

To quantify the incidence of and mortality from leukaemia subtypes in the UK, and their variation with gender, age, socioeconomic status, calendar time and geographic region.

4.4 METHODS

The data source and case ascertainment methods used were the same as in the previous study. Again, I did all data management and statistical analyses myself.

Data management of cases: The following leukaemia subtypes were identified from the same list of READ codes described before (See Appendix

B): Acute Lymphocytic Leukaemia (ALL), Chronic Lymphocytic Leukaemia (CLL), unspecified Lymphocytic or Lymphoid Leukaemia, Acute Myelocytic Leukaemia (AML), Chronic Myelocytic Leukaemia (CML), and unspecified Myeloid Leukaemia. Cases that were described as lymphoid or myeloid, but which could not be identified as either acute or chronic were categorised as 'unspecified' lymphoid or myeloid leukaemia, respectively. These are not, therefore, discrete disease entities. Each case was assigned the most specific diagnosis in their THIN record. This ensured that even if the first recoding of a leukaemia diagnosis for a given patient was non-specific, but subsequent records were more specific because bone marrow biopsy results had became available in the interim, for example, then the more specific diagnosis was captured.

Denominator data: These were derived and used as before.

Variables of interest: Age at diagnosis was calculated as before, and was grouped into 5, 20-year age categories with category 1 being those aged <20yrs and category 5 representing those aged 80yrs and above. Year of diagnosis was grouped into 4, 5-year bands with 1987-1991 being the first year-band and 2002-2006 the most recent. As before, age and year of diagnosis were then included in the analysis as categorical variables, along with gender, Townsend Score and health authority. Male gender was again used as the baseline in gender analyses.

Analysis: The 6 different subtypes described above were analysed separately. Crude incidence rates for leukaemia subtypes, stratified by gender, age at diagnosis, Townsend Score, year of diagnosis and health authority, were calculated. Poisson Regression was then used to determine incidence rate ratios, both independently and mutually adjusted for gender, age, Townsend Score, year of diagnosis and health authority. The median survival for each of the leukaemia subtypes was then calculated. Hazard ratios for the risk of death, adjusted for gender, age-category, Townsend Score, year of diagnosis and health authority, was then determined by Cox Regression. All analyses were conducted using STATAv9.

Ethics: Ethical approval for study was obtained from the Nottingham Research Ethics Committee.

4.5 RESULTS

3 226 (78%) of the 4 162 leukaemia cases were identified as falling into one of the 6 subtypes of interest. The 936 cases that did not have a sufficiently specific diagnosis to be further categorised were broadly representative of the overall study population.

With respect to gender, for example, 56% of them were male and 44% were female. The age at diagnosis of the uncategorized cases showed a bimodal distribution (Figure 4.1) consistent with that of the overall population, as seen

in Figure 4.2. In terms of socioeconomic class, the number of uncategorized cases decline with increasing Townsend Score (i.e. with increasing socioeconomic deprivation) (see Figure 4.3 below) as is seen with the overall leukaemia population studied (see Figure 4.4 below and Table 3.5). The increase in the number of cases recorded over time in this group between 1987-1991 and 1997-2001 follows a similar pattern to the overall leukaemia population as is shown in Table 4.1. The number of uncategorized cases then fell in the period 2002-2006. The uncategorized cases as a proportion of all leukaemia diagnoses fell from 1992-1996 through 2002-2006, which suggests that the recording of leukaemia diagnoses in general practice data has improved over time. This is probably the result of a combination of more specific diagnoses being made, and then reported to GPs from secondary care, and/or more specific diagnoses being recorded in the general practice computer systems. The health authorities in which these patients reside also reflect those of the leukaemia cases overall, with more cases from the South Central and fewest cases from Northern Ireland (see Figure 4.5 and Figure 4.6). Excluding these 936 cases in subsequent analyses is therefore unlikely to have introduced important bias into the results that follow.

Figure 4.1: Age Distribution of Unclassified Cases

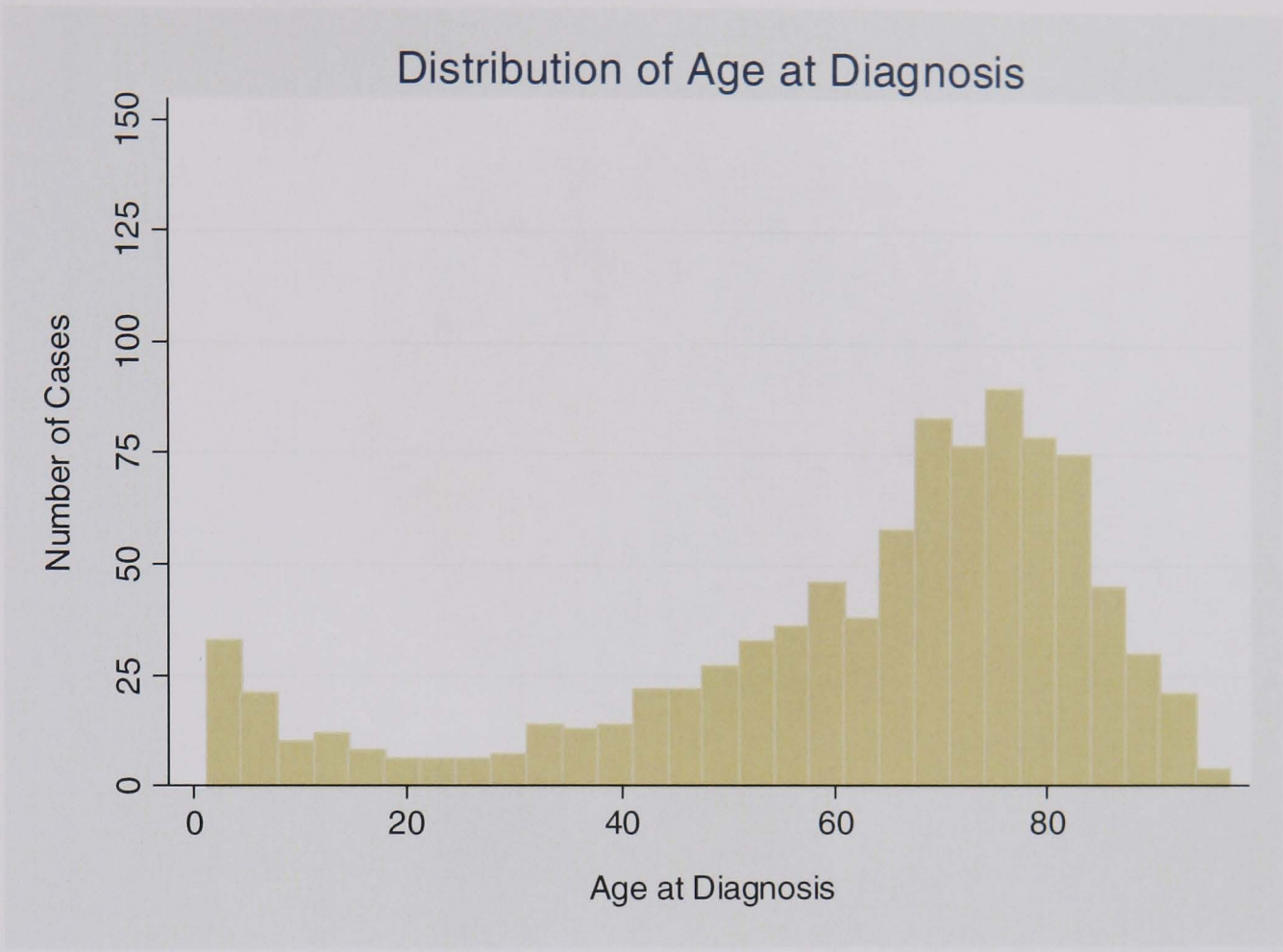


Figure 4.2: Age Distribution of All Leukaemia Cases

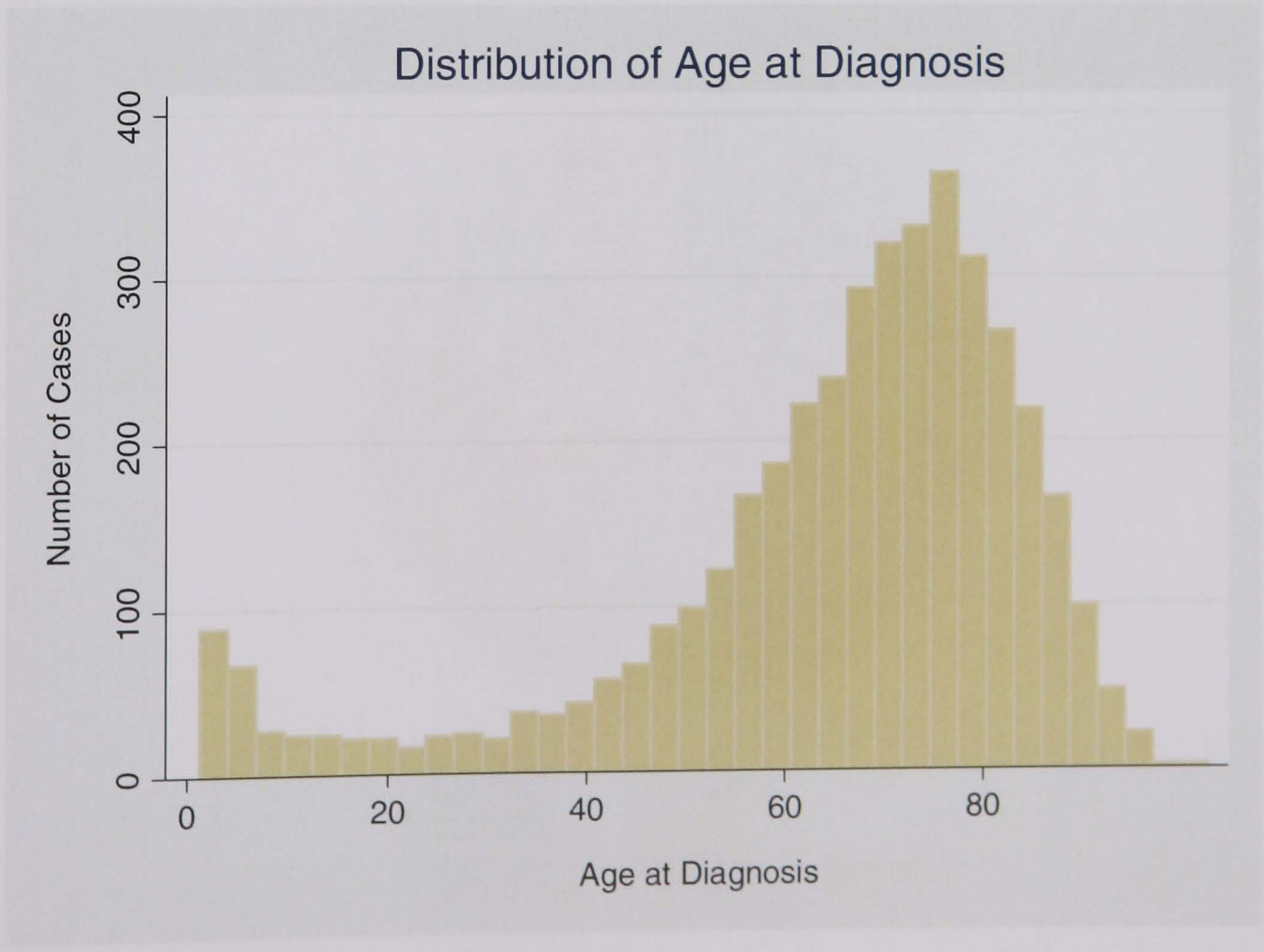


Figure 4.3: Townsend Score Distribution of Unclassified Cases

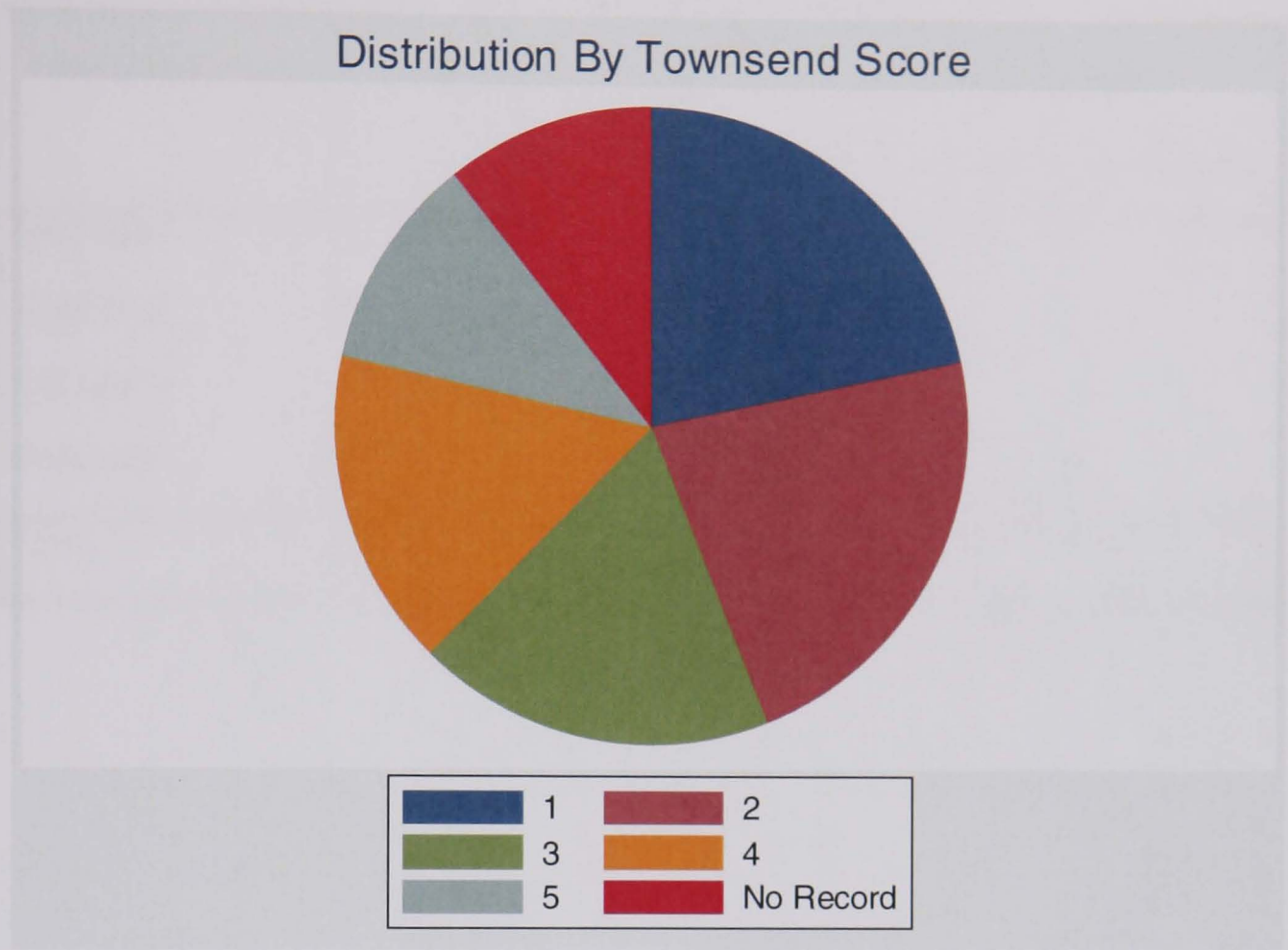


Figure 4.4: Townsend Score Distribution of All Leukaemia Cases

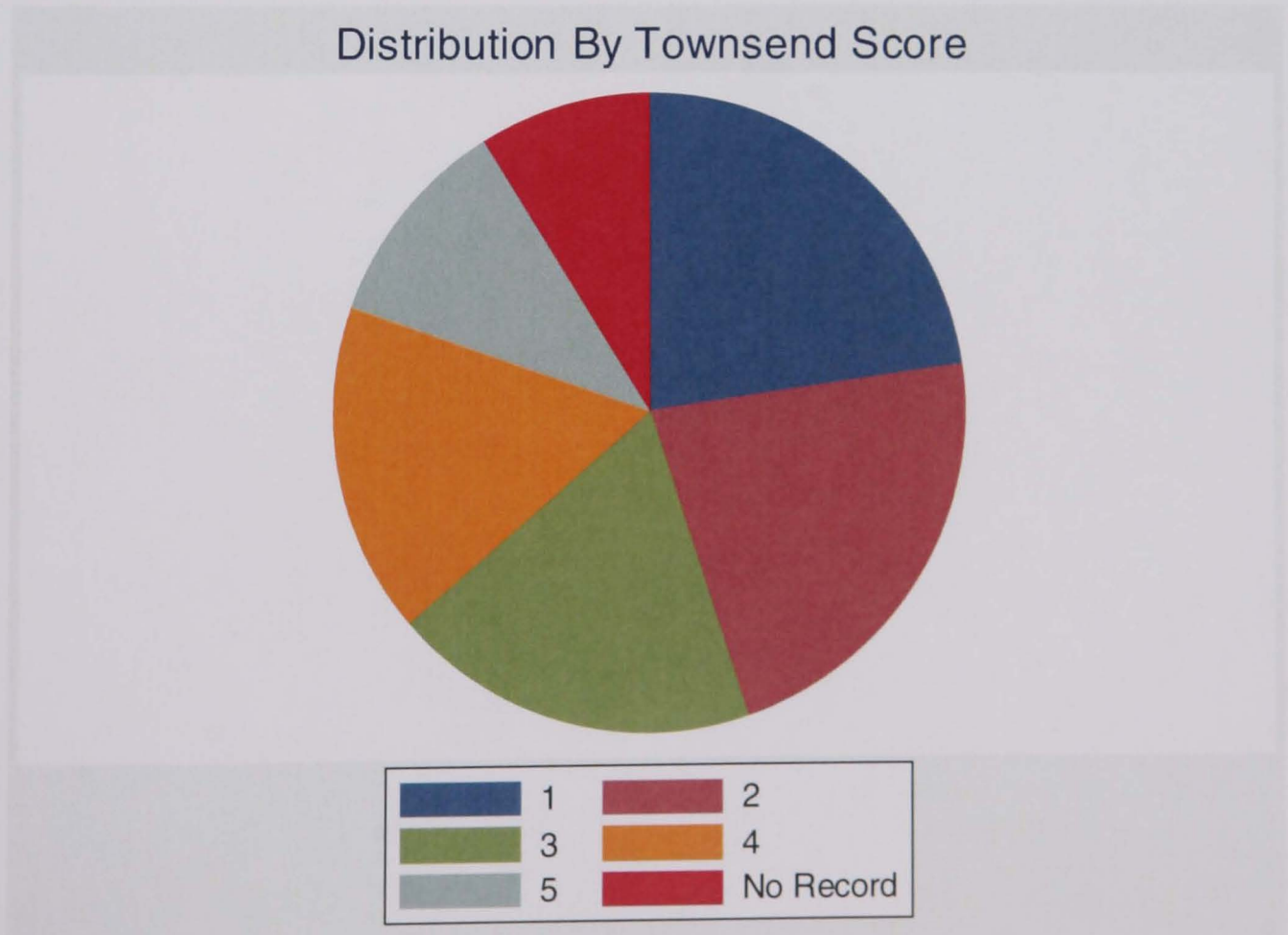


Table 4.1: Uncategorized vs All Leukaemia Cases by Year of Diagnosis

Year of Diagnosis	All Leukaemia Cases	Uncategorized Cases	
	Number	Number	% Of all leukaemia cases
1987-1991	213	62	29
1992-1996	843	283	34
1997-2001	1 244	310	25
2002-2006	1 690	257	15
Total	4 162	936	22

Figure 4.5: Health Authority Distribution of Unclassified Cases

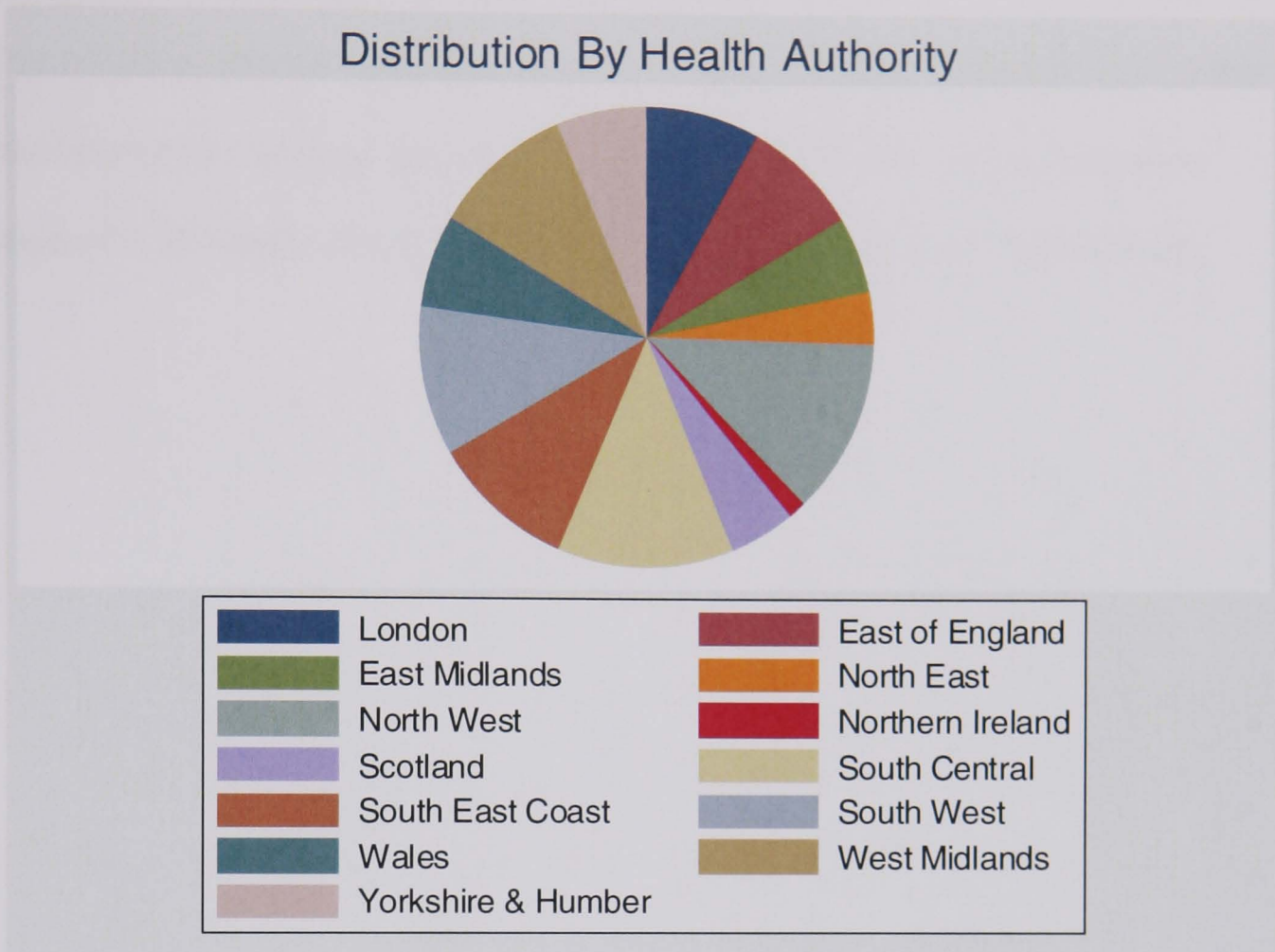
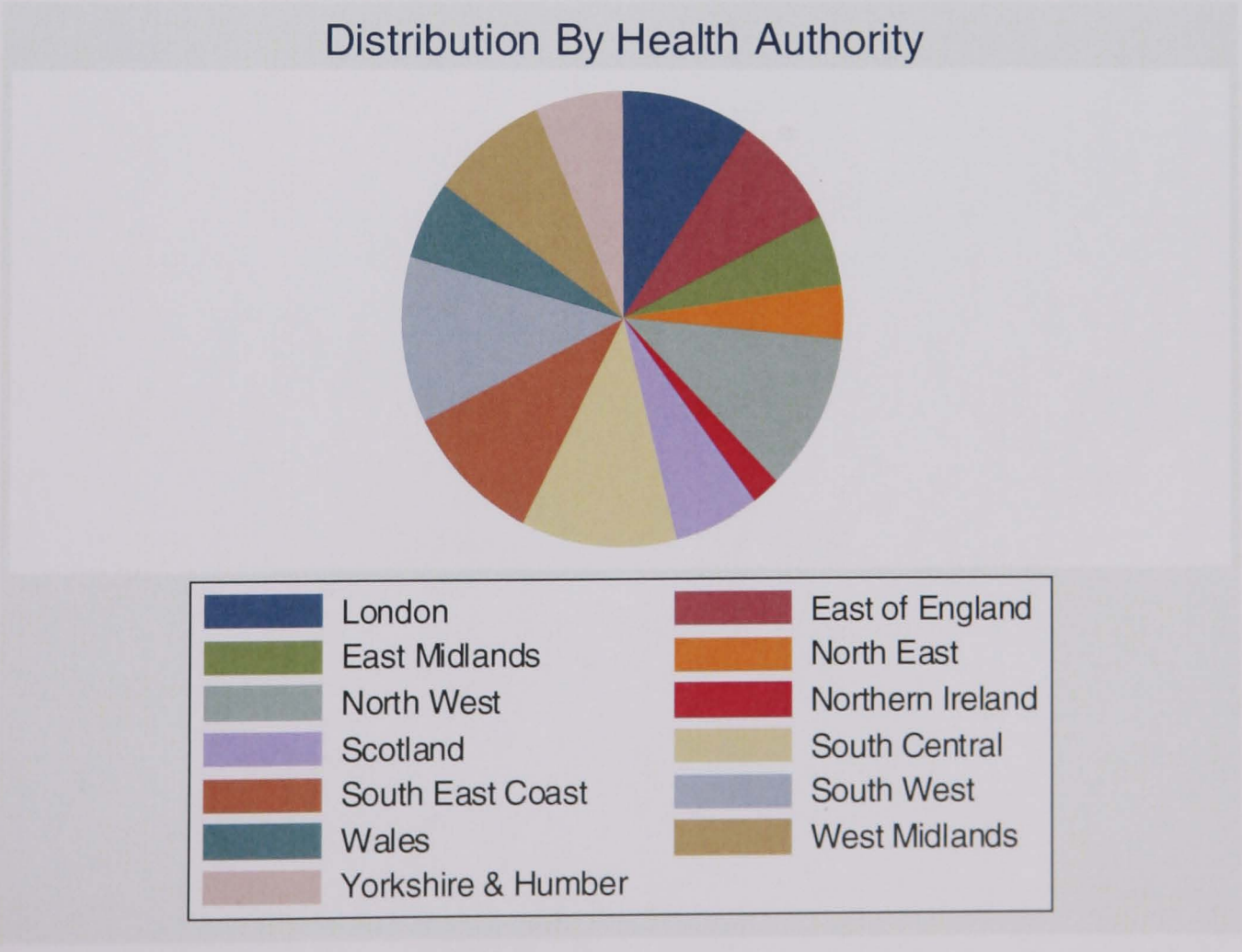


Figure 4.6: Health Authority Distribution of All Leukaemia Cases



4.5.1 Incidence

The crude incidence rates are shown in Table 4.2 and Table 4.3. CLL is the most common subtype and ALL the least common, with crude incidence rates of 4.20 and 0.49 per hundred thousand person years, respectively.

Table 4.2: Crude Incidence Rates Per 100 000 Person-years

	Person-years	ALL		CLL		Unsp. Lymph.		AML		CML		Unsp. Myel.	
		Cases	IR	Cases	IR	Cases	IR	Cases	IR	Cases	IR	Cases	IR
Overall:	36 993 338	182	0.49	1 554	4.20	356	0.96	602	1.63	265	0.72	257	0.70
Gender: Male	18 284 784	89	0.49	914	5.00	203	1.11	319	1.74	138	0.75	118	0.65
Female	18 708 554	93	0.50	640	3.42	153	0.82	283	1.51	137	0.73	139	0.74
Age at Diagnosis:													
<20	8 288 518	116	1.40	1	0.01	20	0.24	30	0.36	5	0.06	7	0.08
20-39	10 600 000	23	0.22	6	0.06	10	0.09	54	0.51	23	0.22	18	0.17
40-59	9 995 140	18	0.18	251	2.51	68	0.68	123	1.23	66	0.66	58	0.58
60-79	6 462 905	14	0.22	927	14.34	185	2.86	296	4.58	120	1.86	115	1.78
80+	1 668 108	11	0.66	369	22.12	73	4.38	99	5.93	61	3.66	59	3.54
Townsend Score:													
1	8 421 428	45	0.53	356	4.22	85	1.01	138	1.64	65	0.77	54	0.64
2	7 179 869	35	0.49	335	4.67	82	1.14	145	2.02	51	0.71	66	0.92
3	6 952 931	36	0.52	293	4.21	64	0.92	111	1.60	61	0.88	47	0.68
4	6 220 650	27	0.43	269	4.32	65	1.04	103	1.66	40	0.64	40	0.64
5	4 488 547	19	0.42	174	3.89	35	0.78	53	1.18	32	0.71	31	0.69
No Record	3 729 913	20	0.54	127	3.40	25	0.67	52	1.39	26	0.70	19	0.51

Table 4.3: Crude Incidence Rates Per 100 000 Person-years (Cont.)

Year of Diagnosis:												
1987-1991	3 323 812	10	0.30	62	1.87	29	0.87	21	0.63	16	0.48	13	0.39
1992-1996	8 945 190	31	0.35	230	2.57	102	1.14	82	0.92	66	0.74	49	0.55
1997-2001	11 800 000	54	0.46	449	3.80	104	0.88	164	1.39	75	0.63	88	0.74
2002-2006	12 900 000	81	0.63	733	5.68	113	0.88	295	2.29	108	0.84	103	0.80
Health Authority:													
London	4 417 180	25	0.57	162	3.67	31	0.70	63	1.43	28	0.63	15	0.34
East of England	3 227 730	14	0.43	120	3.72	31	0.96	53	1.64	26	0.81	17	0.53
East Midlands	1 798 239	6	0.33	84	4.67	16	0.89	25	1.39	15	0.83	17	0.95
North East	1 257 649	5	0.40	52	4.13	16	1.27	28	2.23	16	1.27	4	0.32
North West	4 001 708	24	0.60	171	4.27	34	0.85	67	1.67	29	0.72	26	0.65
Northern Ireland	1 156 921	6	0.52	31	2.68	11	0.95	11	0.95	4	0.35	8	0.69
Scotland	2 243 121	9	0.40	112	4.99	21	0.94	37	1.65	21	0.94	17	0.76
South Central	4 103 293	21	0.51	185	4.51	30	0.73	64	1.56	34	0.83	24	0.58
South East Coast	3 425 204	14	0.41	150	4.38	45	1.31	71	2.07	19	0.55	29	0.85
South West	3 733 229	20	0.54	191	5.12	25	0.67	78	2.09	37	0.99	36	0.96
Wales	1 909 249	12	0.63	78	4.09	30	1.57	25	1.31	12	0.63	16	0.84
West Midlands	3 518 793	20	0.57	117	3.32	34	0.97	47	1.34	19	0.54	28	0.80
Yorkshire & Humber	2 201 022	6	0.27	101	4.59	32	1.45	33	1.50	15	0.68	20	0.91

The age at diagnosis of leukaemia by subtype is shown in Figure 4.7. Most cases of ALL are diagnosed in early childhood, while most CLL is diagnosed in the elderly.

Figure 4.7: Age at Diagnosis

The gender differences in age at diagnosis for the 4 specific subtypes of interest are shown in Figure 4.8 to Figure 4.11. The more dramatic increase in incidence in men than women after the age of 40 (noted earlier for leukaemia overall (Figure 3.2)) is mostly accounted for by CLL, where this trend is most pronounced, and by AML.

Figure 4.8: ALL Incidence by Age and Gender

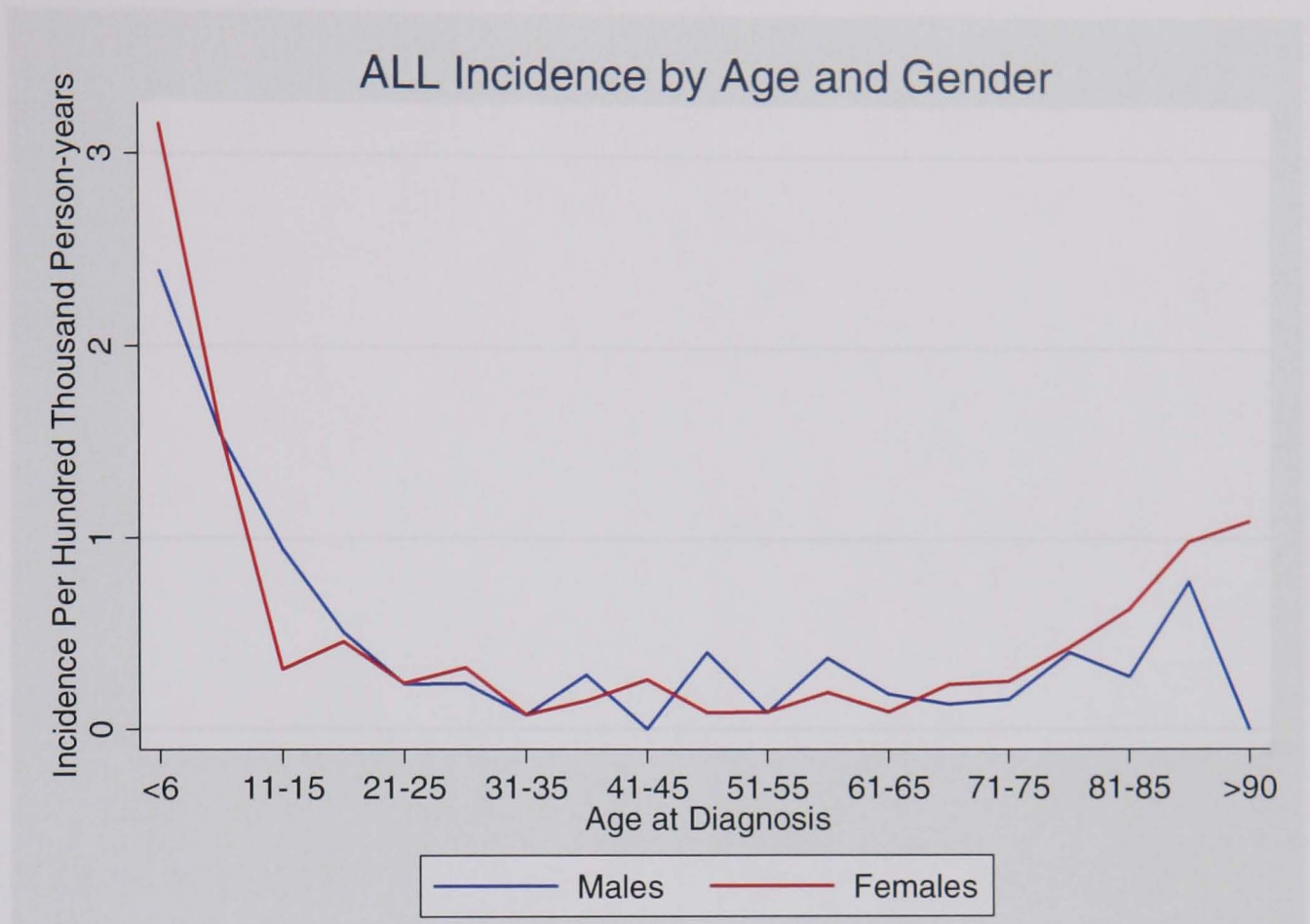


Figure 4.9: CLL Incidence by Age and Gender

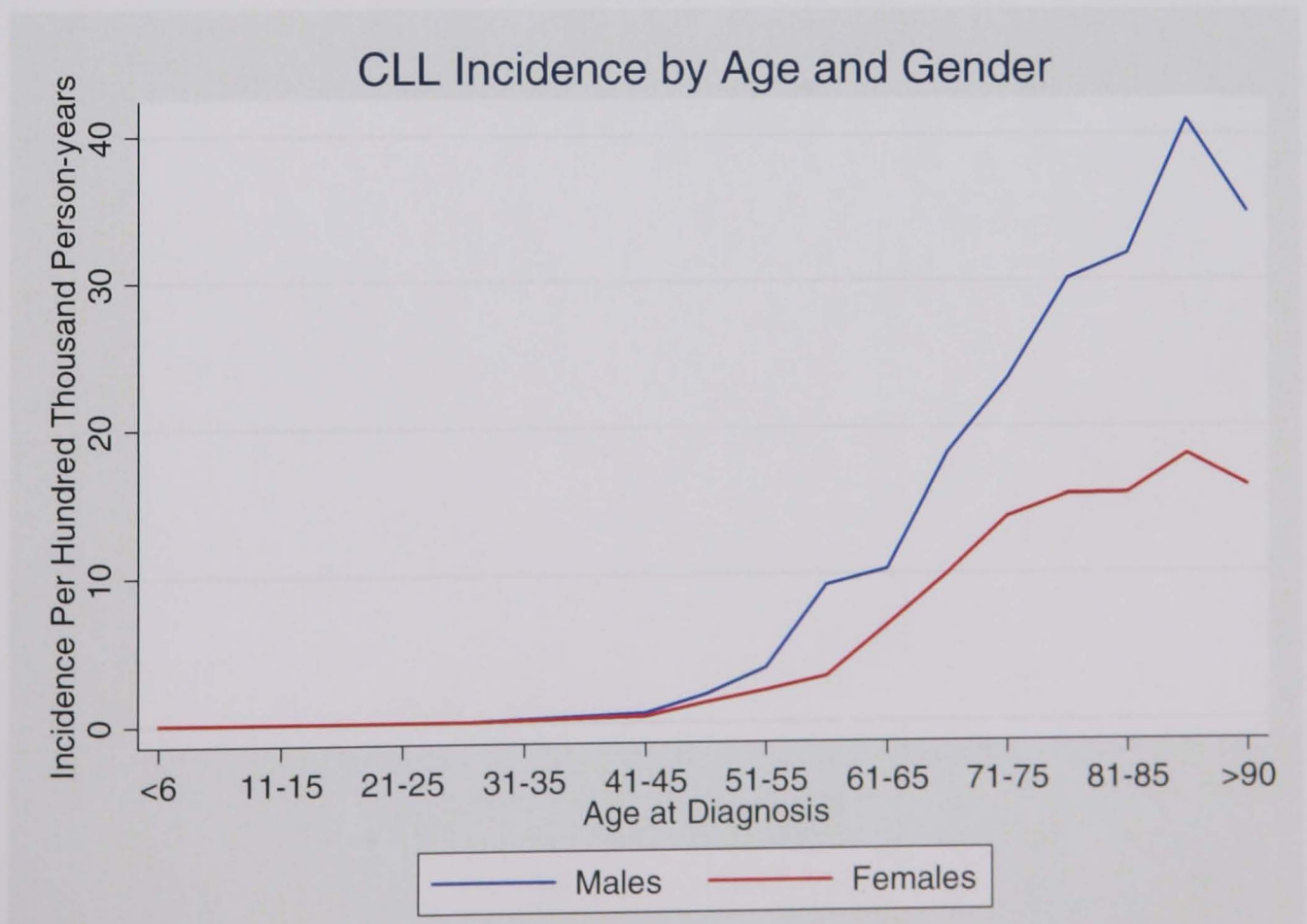


Figure 4.10: AML Incidence by Age and Gender

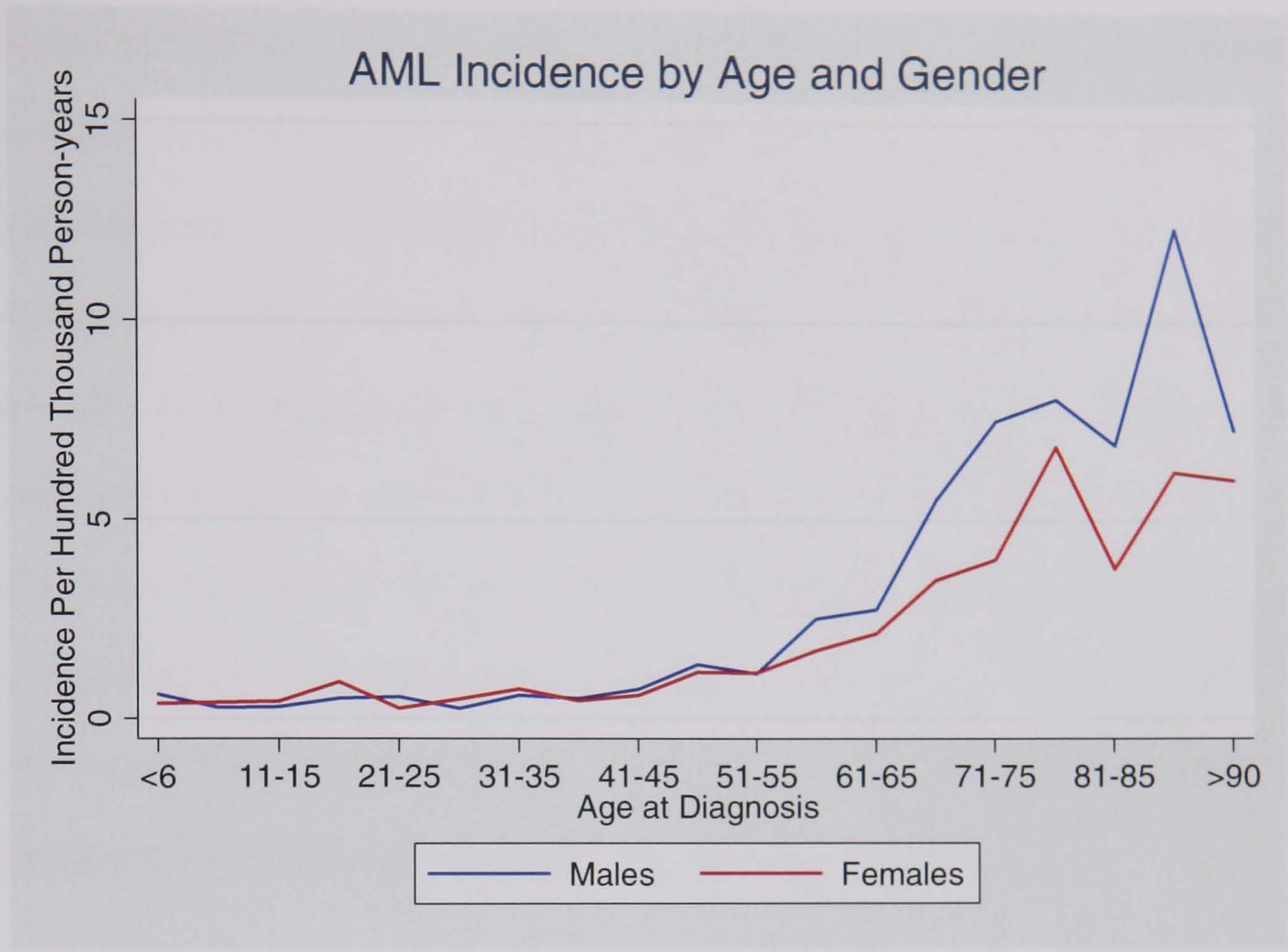
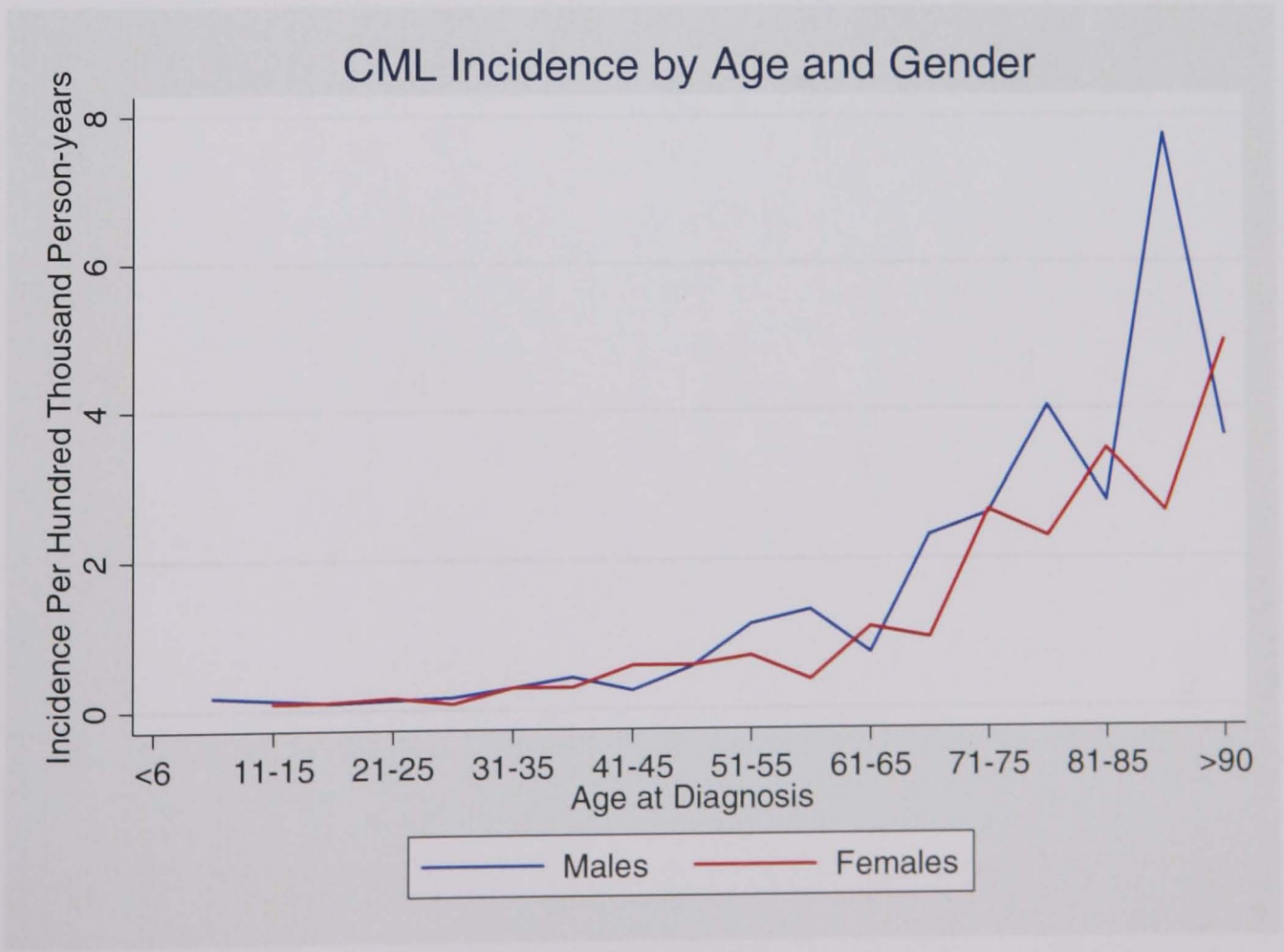


Figure 4.11: CML Incidence by Age and Gender



Results of Poisson regression are shown in Table 4.4 and Table 4.5, in which incidence rate ratios are adjusted for all other variables in the table. Women have a lower incidence of CLL, unspecified Lymphocytic Leukaemia and AML than men. The incidence of all subtypes increases with increasing age at diagnosis (p for trend<0.01 for each subtype). A decreasing incidence of ALL with increasing deprivation is present, but this trend does not reach statistical significance (p for trend =0.13). There is no evidence of a socioeconomic class trend in incidence in any of the other subtypes. An increasing trend in incidence with calendar time is seen for ALL, CLL, AML and unspecified Myeloid Leukaemia. Statistically significant regional variation in incidence is only seen in unspecified Myeloid Leukaemia.

Table 4.4: Mutually Adjusted Incidence Rate Ratios (95% C.I.)

	ALL	CLL[1]	Unsp. Lymph.	AML	CML	Unsp. Myel.
Gender						
Males	1	1	1	1	1	1
Females	1.05(0.78-1.42) p=0.73	0.55(0.49-0.61) p<0.01	0.62(0.50-0.77) p<0.01	0.78(0.66-0.92) p<0.01	0.84(0.66-1.07) p=0.17	0.96(0.75-1.23) p=0.74
Diagnosis Age						
<20	1	}	1	1	1	1
20-39	0.14(0.09-0.23)	} 1	0.40 (0.19-0.85)	1.40(0.89-2.21)	3.32(1.25-8.81)	2.07(0.87-4.96)
40-59	0.11(0.06-0.18)	74.74 (33.24-168.02)	2.73(1.66-4.50)	3.28(2.18-4.93)	10.34(4.16-25.71)	6.86(3.13-15.02)
60-79	0.16(0.09-0.27)	433.54(194.22-967.72)	11.82(7.44-18.76)	12.34(8.41-18.09)	29.58(12.08-72.44)	20.28(9.44-43.54)
80+	0.47(0.25-0.87)*p<0.01	731.00(326.11-1638.61)*p<0.01	20.15(12.25-33.14)*p<0.01	16.23(10.67-24.67)*p<0.01	62.42(25.04-155.60)*p<0.01	41.68(19.00-91.43)*p<0.01
Townsend Score						
1	1	1	1	1	1	1
2	0.92(0.59-1.45)	1.01(0.86-1.17)	1.01(0.75-1.38)	1.15(0.90-1.47)	0.79(0.54-1.15)	1.32(0.92-1.89)
3	0.90(0.57-1.42)	0.95(0.81-1.12)	0.88(0.63-1.22)	0.92(0.71-1.20)	1.07(0.75-1.52)	1.02(0.69-1.52)
4	0.76(0.47-1.23)	1.01(0.86-1.19)	1.00(0.72-1.39)	1.01(0.77-1.32)	0.72(0.47-1.09)	1.01(0.66-1.53)
5	0.67(0.38-1.18)*p=0.13	0.93(0.77-1.13)*p=0.59	0.74(0.49-1.12)*p=0.25	0.73(0.52-1.03)*p=0.10	0.84(0.54-1.31)*p=0.32	1.19(0.75-1.87)*p=0.97
No Record	0.98(0.56-1.71)	0.88(0.72-1.09)	0.64(0.40-1.03)	0.86(0.61-1.21)	0.82(0.51-1.32)	0.86(0.50-1.49)

[1] Baseline age category for CLL is age<40. *p=test for trend across ordered categories.

Records with missing Townsend Scores were not included in trend analysis for Townsend Score.

Table 4.5: Mutually Adjusted Incidence Rate Ratios (95% C.I.) (Cont.)

Diagnosis Year						
1987-1991	1	1	1	1	1	1
1992-1996	1.15(0.56-2.34)	1.34(1.01-1.77)	1.25(0.83-1.89)	1.43(0.88-2.31)	1.49(0.86-2.58)	1.35(0.73-2.49)
1997-2001	1.49(0.76-2.92)	1.96(1.50-2.55)	0.95(0.63-1.44)	2.14(1.36-3.38)	1.27(0.74-2.18)	1.83(1.02-3.28)
2002-2006	2.02(1.04-3.90)*p<0.01	2.87(2.21-3.72)*p<0.01	0.92(0.61-1.39)*p=0.11	3.49(2.24-5.44)*p<0.01	1.64(0.97-2.78)*p=0.12	1.92(1.08-3.43) *p=0.01
Health Authority						
London	1	1	1	1	1	1
East of England	0.72(0.38-1.40)	0.84(0.66-1.07)	1.09(0.66-1.81)	0.98(0.67-1.43)	1.14(0.66-1.96)	1.37(0.68-2.75)
East Midlands	0.57(0.23-1.40)	1.06(0.80-1.39)	1.02(0.56-1.88)	0.79(0.48-1.29)	1.12(0.58-2.16)	2.48(1.23-5.00)
North East	0.69(0.26-1.82)	0.93(0.68-1.28)	1.56(0.85-2.85)	1.39(0.88-2.19)	1.75(0.93-3.30)	0.80(0.27-2.43)
North West	1.03(0.58-1.83)	0.94(0.75-1.72)	1.03(0.63-1.68)	0.97(0.68-1.40)	1.03(0.60-1.75)	1.69(0.89-3.21)
Northern Ireland	0.67(0.25-1.77)	0.72(0.49-1.06)	1.21(0.59-2.48)	0.66(0.35-1.26)	0.56(0.20-1.62)	1.94(0.82-4.60)
Scotland	0.69(0.32-1.49)	1.12(0.87-1.43)	1.28(0.74-2.23)	1.09(0.72-1.65)	1.46(0.83-2.59)	1.90(0.94-3.84)
South Central	0.81(0.44-1.50)	0.90(0.72-1.12)	0.77(0.46-1.29)	0.77(0.52-1.12)	1.09(0.65-1.84)	1.27(0.65-2.48)
South East Coast	0.61(0.31-1.21)	0.93(0.73-1.17)	1.40(0.88-2.24)	1.20(0.84-1.72)	0.63(0.33-1.18)	2.05(1.09-3.87)
South West	0.86(0.47-1.59)	0.97(0.78-1.21)	0.71(0.42-1.21)	1.07(0.75-1.52)	1.27(0.76-2.09)	2.14(1.16-3.95)
Wales	1.10(0.54-2.21)	0.87(0.66-1.15)	1.56(0.92-2.64)	0.72(0.44-1.71)	0.88(0.44-1.75)	2.07(1.02-4.21)
West Midlands	0.92(0.50-1.70)	0.79(0.62-1.01)	1.16(0.71-1.90)	0.83(0.56-1.23)	0.76(0.41-1.39)	2.11(1.12-3.99)
Yorks. & Humber	0.48(0.19-1.17)	0.97(0.75-1.26)	1.55(0.94-2.57)	0.89(0.57-1.37)	0.81(0.41-1.57)	2.21(1.12-4.33)

*p=test for trend across ordered categories.

4.5.2 Survival

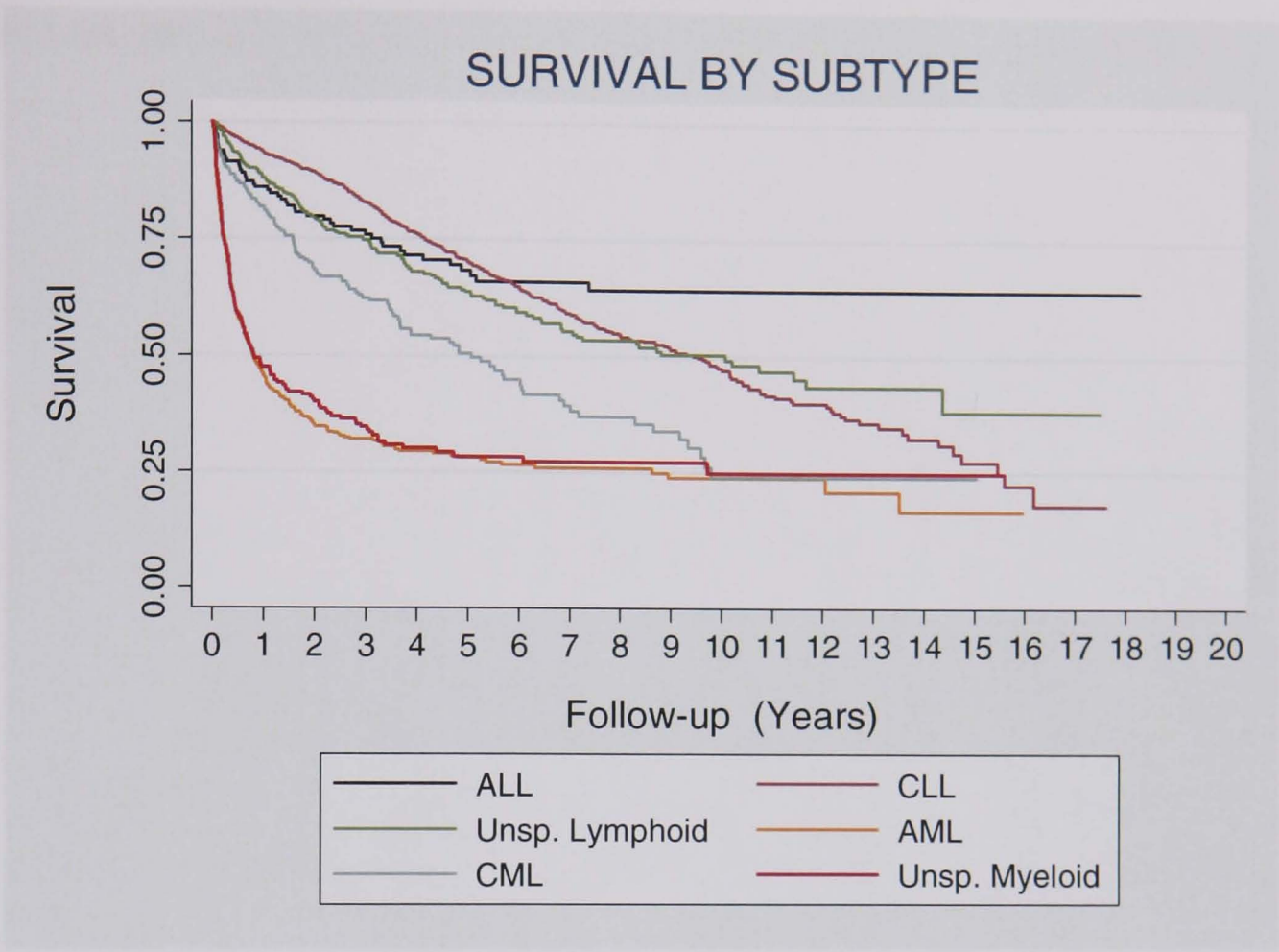
The median survival from each subtype is shown in Table 4.6 and Kaplan-Meier survival curves are plotted by subtype in Figure 4.12. ALL has the best prognosis, with more than 50% of cases surviving the end of the follow-up period. The 5-year survival for ALL was therefore calculated and is 69%. The poorest median survival, of only 9½ months, is seen in AML.

Table 4.6: Median Survival by Subtype

Subtype	Cases	Median Survival (Yrs)	Std. Error	95% C.I.
ALL*	180	-	-	-
CLL	1 549	9.53	0.30	8.20 - 10.18
Unspecified Lymphoid	352	10.05	0.54	6.83 - .
AML	593	0.79	0.03	0.64 - 1.00
CML	274	5.06	0.19	3.67 - 6.07
Unspecified Myeloid	250	0.81	0.04	0.59 - 1.39
OVERALL	3 198	6.82	0.37	6.16 - 7.46

* >50% of ALL cases survived follow-up period. (5-year mortality is 69%)

Figure 4.12: Kaplan-Meier Survival Plots by Subtype



Hazard ratios are presented in Table 4.7 and Table 4.8, and show a lower risk of death among women than men with CLL (HR 0.59, p<0.01, 95%, C.I. 0.49-0.72) and unspecified lymphoid leukaemia (HR 0.58, p=0.01, 95% C.I. 0.40-0.85), when adjusted for age-category, Townsend Score, year of diagnosis and health authority. The relative risk of death increases with increasing age at diagnosis in all subtypes (p for trend<0.01). In AML the risk of death increases with increasing deprivation (p for trend=0.01), but no such gradient is seen in any of the other subtypes. There has been no statistically significant trend in the risk of death over calendar time. The risk of death in AML in the North East was nearly twice that in London even when gender, age at diagnosis, Townsend Score and year of diagnosis was adjusted for,

and also significantly greater in the North West than in London (North East HR 1.94, 95% C.I. 1.10-3.42; North West HR 1.66, 95% C.I. 1.04-2.65).

Table 4.7: Mutually adjusted Hazard Ratios (95% C.I.)

	ALL	CLL[!]	Unsp.Lymph.	AML	CML	Unsp.Myel.
Gender						
Males	1	1	1	1	1	1
Females	1.34(0.71-2.56) p=0.37	0.59(0.49-0.72) p<0.01	0.58(0.40-0.85) p=0.01	0.89(0.72-1.10) p=0.29	0.89(0.61-1.30) p=0.56	1.02(0.72-1.43) p=0.94
Diagnosis Age						
<20	1	}	1	1	}	1
20-39	0.82(0.26-2.61)	} 1	4.01(0.35-46.57)	1.63(0.58-4.60)	} 1	0.88(0.17-4.54)
40-59	4.63(1.81-11.86)	0.36(0.08-1.57)	6.20(0.81-47.23)	4.64(1.79-11.16)	0.92(0.35-2.43)	1.38(0.31-6.10)
60-79	20.82(7.57-57.22)	1.08(0.26-4.48)	10.98(1.51-79.76)	10.02(4.09-24.55)	2.46(1.04-5.83)	3.88(0.91-16.53)
80+	4.96(1.43-17.20)*p<0.01	2.85(0.69-11.77)*p<0.01	29.51(4.00-217.82)*p<0.01	23.80(9.45-59.91)*p<0.01	6.86(2.74-17.22)*p<0.01	7.22(1.67-31.21)*p<0.01
Townsend Score						
1	1	1	1	1	1	1
2	1.05(0.40-2.80)	1.21(0.84-1.49)	0.75(0.44-1.29)	1.15(0.84-1.58)	0.60(0.31-1.15)	1.02(0.62-1.69)
3	1.21(0.45-3.28)	1.30(0.98-1.74)	1.94(1.16-3.24)	1.38(0.99-1.92)	0.86(0.49-1.50)	0.79(0.46-1.35)
4	0.89(0.32-2.5)	1.21(0.89-1.65)	0.67(0.37-1.23)	1.44(1.03-2.02)	0.91(0.48-1.73)	0.69(0.39-1.21)
5	0.25(0.05-1.17)*p=0.23	1.18(0.82-1.70)*p=0.31	0.76(0.39-1.50)*p=0.40	1.52(0.99-2.35)*p=0.01	0.73(0.36-1.51)*p=0.65	0.75(0.40-1.43)*p=0.13
No Record	0.79(0.23-2.63)	2.07(1.43-2.99)	1.29(0.59-2.83)	1.06(0.68-1.63)	1.34(0.65-2.75)	0.63(0.27-1.49)

[!] Baseline age category for CLL is age<40. *p=test for trend across ordered categories.

Records with missing Townsend Scores were not included in trend analysis for Townsend Score.

Table 4.8: Mutually adjusted Hazard Ratios (95% C.I.) (Cont.)

Diagnosis Year						
1987-1991	1	1	1	1	1	1
1992-1996	6.44(0.91-45.71)	0.88(0.60-1.27)	0.86(0.47-1.58)	0.60(0.34-1.08)	0.64(0.31-1.35)	1.31(0.52-3.31)
1997-2001	6.94(1.03-46.58)	1.02(0.71-1.48)	0.86(0.45-1.64)	0.52(0.30-0.90)	0.72(0.34-1.54)	1.13(0.46-2.78)
2002-2006	4.57(0.72-28.90)*p=0.83	1.09(0.74-1.61)*p=0.58	1.25(0.65-2.42)*p=0.77	0.54(0.32-0.93) *p=0.93	0.45(0.20-0.99)*p=0.07	1.10(0.45-2.69)*p=0.19
Health Authority						
London	1	1	1	1	1	1
East of England	1.00(0.23-4.38)	1.10(0.69-1.76)	1.14(0.43-3.06)	1.41(0.86-2.31)	0.85(0.37-1.97)	0.74(0.29-1.89)
East Midlands	0.37(0.04-3.83)	1.27(0.77-2.12)	1.89(0.69-5.20)	1.10(0.58-2.09)	0.89(0.33-2.41)	0.98(0.38-2.57)
North East	1.72(0.32-9.22)	1.32(0.74-2.36)	1.39(0.51-3.82)	1.94(1.10-3.42)	0.75(0.30-1.89)	0.95(0.26-3.46)
North West	0.86(0.27-2.76)	1.43(0.94-2.19)	1.69(0.74-3.85)	1.66(1.04-2.65)	0.98(0.43-2.19)	1.27(0.55-2.96)
Northern Ireland	6 cases, no deaths	1.12(0.49-2.54)	1.07(0.35-3.30)	1.12(0.43-2.92)	4 cases, no deaths	1.35(0.39-4.70)
Scotland	0.33(0.04-2.99)	0.84(0.50-1.42)	0.97(0.33-2.84)	1.21(0.68-2.14)	0.39(0.14-1.11)	0.96(0.36-2.59)
South Central	0.95(0.26-3.45)	1.21(0.79-1.86)	1.16(0.45-2.99)	1.14(0.70-1.85)	0.64(0.28-1.47)	0.90(0.36-2.24)
South East Coast	1.51(0.43-5.29)	1.38(0.88-2.15)	1.07(0.47-2.44)	1.43(0.89-2.29)	0.64(0.24-1.74)	0.80(0.31-2.07)
South West	1.05(0.33-3.35)	1.24(0.81-1.88)	0.87(0.33-2.31)	1.17(0.74-1.85)	0.57(0.25-1.32)	1.29(0.56-2.99)
Wales	1.80(0.41-7.92)	1.30(0.77-2.19)	1.34(0.53-3.41)	1.18(0.64-2.19)	0.85(0.28-2.55)	1.45(0.57-3.73)
West Midlands	2.13(0.65-6.94)	1.26(0.79-2.01)	1.87(0.81-4.29)	1.58(0.94-2.66)	0.41(0.13-1.33)	1.14(0.48-2.75)
Yorks. & Humber	6 cases, no deaths	1.56(0.97-2.53)	1.61(0.68-3.79)	1.05(0.59-1.87)	0.77(0.27-2.16)	0.88(0.35-2.21)

4.6 DISCUSSION

4.6.1 ALL

4.6.1.1 Incidence

No statistically significant difference in incidence between males and females was found in this study. Results of previously published studies are inconsistent in this regard. A study examining sex ratios in incident leukaemia found no gender difference in patients aged under 10 years, but a male predominance at older ages[7]. Many other studies have, however, found a male excess when only children were studied[5 8 35 64 67]. These differences may be explained by the different age ranges studied, different time periods covered by the respective studies, as well as the relative proportion of children in the respective study populations. Furthermore, studies have adjusted for different combinations of potential confounders from each other, and from this study. The age at diagnosis of ALL in this study is consistent with that published elsewhere, with a peak under the age of 5 years, followed by a relatively low incidence until after the 7th decade, when incidence increases again[35 64]. Several studies found no socioeconomic class trend in incidence, in keeping with findings here[9 10 35]. In terms of trends over time, results here are again consistent with those in the literature which show an increasing incidence of ALL starting in the 1950's through the 1970's[64],

followed by a plateau in incidence rates thereafter[67]. No significant regional variation was found, here in keeping with earlier UK studies[9].

4.6.1.2 Mortality

The survival patterns seen in ALL in these data are consistent with the findings of earlier published research. The 5-year survival has consistently been reported at around 70% for those diagnosed in the 1980's[13 14 16 68]. This research found no gender differences in mortality in ALL, while others have shown better survival in girls[68]. A review of UK children covering the period 1911-97 found greater mortality among boys than girls until the 1980s, after which mortality rates were similar, until mortality in girls was actually greater than in boys by 1997[11]. Survival was also found to be worse with increasing age at diagnosis[14 68]. Like this research, other studies have also consistently reported improvements in survival over the past 3 decades, both across Europe and within the UK[13 14 16 68]. No socioeconomic trend in mortality is seen in these data, nor are there significant regional differences in mortality. These findings are consistent with an earlier study of ALL in the UK [16].

4.6.2 CLL

4.6.2.1 Incidence

CLL is known to be the most commonly diagnosed leukaemia in adults in western countries[3 17], which concurs with the findings of this research. Again the gender and age distribution of incident CLL found here is consistent with previous studies, which show a male predominance[3 7 69], and greater incidence rates with increasing age[3 17 69]. An increasing incidence over time has also been noted before[11], although it is unclear whether this represents a true increase in disease incidence, or increased diagnosis as a result of routine blood testing in older patients, for example. This study found regional differences in CLL incidence, although no region had significantly different incidence rates from London. Again, this is consistent with the findings of others[17]. Socioeconomic class trends in CLL have not been published for the UK before.

4.6.2.2 Mortality

Survival in CLL ranges from months to decades, which is consistent with the fact that most patients will die from other, unrelated causes. The median survival in this study of 9½ years is consistent with the findings of others[3 17]. The finding that mortality increases with increasing age at diagnosis is also entirely expected, given that co-morbidity is likely to be greater with

increasing age and that patients die of co-morbid conditions more often than of CLL itself. The gender difference in mortality of about 2:1 (male: female) found here is consistent with previously published research[11 69 70]. Reasons for this phenomenon might include greater co-morbidity among men, better response to treatment in women, better tolerance to the adverse effects of treatment in women, and/or intrinsic gender differences in the disease at biological level. In addition to the combination of factors in play related to CLL, women also have a longer life expectancy than men in general. Mortality is reported to have remained fairly constant over the past 2 decades, again consistent with the results of this research[6 11]. No studies specifically examining the association of CLL mortality with socioeconomic class have been published. Given that over half of patients are reported to die from other causes[17], it seems that any associations with socioeconomic class will reflect the associations between the respective causes of death and socioeconomic class. The lack of association of socioeconomic class and mortality found in this research is thus not surprising. Regional mortality figures for CLL have not been published for UK patients, but this research showed no statistically significant differences in mortality in any region compared to London.

4.6.3 AML

4.6.3.1 Incidence

AML incidence is greater in men than women, and increases with increasing age at diagnosis in this study. Again this is entirely consistent with previous research findings[7 18 19 64 70]. These data show an increasing incidence over the period of this study. Other reports of time trends in AML incidence produce conflicting results, but are difficult to interpret since they cover different age ranges and different time periods[63]. There is no socioeconomic trend in incidence, nor any statistically significant regional difference in incidence in these data, which concurs with the findings of others[9 10].

4.6.3.2 Mortality

As is the case in CLL, mortality in AML is greater in men than in women, and increases with increasing age at diagnosis. Both these findings are in keeping with the published literature. Survival has previously been shown to decline with increasing age[2 18], and survival beyond 1 year is reported to be rare in those aged over 70[2]. In a study examining the effect of gender and age on survival in European children with AML, girls had better survival than boys overall, but there was no gender difference in survival under the age of 5 years[20]. The findings with regard to gender were similar across most of the European countries included in the study. A study of 507 adults with AML in the south-east of England, however, found no statistically significant

difference in 1- and 3-year survival between men and women[18]. While mortality in AML in the research presented here appears to have declined over the past 2 decades, the trend did not reach statistical significance. The MRC AML trial data show a consistent improvement in survival over time for younger, but not older, patients[71]. The fact that data in this study do not show a significant improvement in AML survival over time may reflect the overall age distribution of AML in this study, i.e. mostly older people, many of whom are not entered into clinical trials.

The statistically significant increase in mortality with increasing socioeconomic deprivation found here for AML, but not other subtypes, has not been shown before. Given that AML incidence is independent of socioeconomic class, the findings with regard to mortality suggest that diagnostic services are equally accessible across the spectrum of socioeconomic class, but that other factors influence the provision and/or uptake of therapeutic services. Furthermore, mortality in AML is worse in the North East and North West, even after adjusting for deprivation. Again, this is in the absence of regional variation in incidence. This phenomenon has also not been demonstrated before. Regional differences in mortality may similarly be explained by differences in treatment uptake, but may also reflect regional differences in the population's co-morbidity.

4.6.4 CML

4.6.4.1 Incidence

This research found a greater incidence of CML with increasing age, consistent with others' findings[2 4]. A male excess in incidence was found here, but the gender difference was not statistically significant. While the published literature mostly shows a male predominance[4 7 18], gender differences do not always reach statistical significance and the non-significant results found here may be due to the small number of CML cases in this study (265 cases). Although the incidence of CML in another study was found to increase with increasing socioeconomic deprivation among teenagers and young adults in England[9], this was not borne out in these data. The authors do not state the number of CML cases studied, however, so it is not clear whether that study is comparable to the one presented here in terms of size. The same study also found statistically significant differences in incidence across geographic regions, which is not evident in these data. Again, this difference is difficult to interpret due to the lack of information on the size of the study population. Differences in the age of the study populations and/or differences in study design may also account for the difference in results. No other reports were found in the literature describing the incidence of CML in the context of socioeconomic class or geographic region within the UK. No trend in the incidence of CML over time is seen in this study. A review reporting trends in incident CML among adults showed

an increase in men between 1970 and 1985, but this was followed by a decrease in incidence until the end of the study period in 1990, back to 1970 levels. In women the incidence rates were fairly similar in 1990 compared to what they were in 1970[11]. Reported CML trends need to be interpreted with caution since Chronic Myelomonocytic Leukaemia (CMML), previously included within CML, has been reclassified as a myelodysplastic syndrome. This reclassification occurred during the time period included in this study, and hence case ascertainment in this (and other studies incorporating this time period) will vary.

4.6.4.2 Mortality

Mortality in CML increases with age in these data, as is the case with the other leukaemia subtypes studied. This has been also been demonstrated before[18]. Mortality tends to be greater in men than women in this study, which is also consistent with other published work, which found better 1- and 5-year survival among women than men with both CML and CMML in a study of UK adults[18]. Gender differences did not reach statistical significance in the aforementioned study, but case numbers were small (180 CML cases, 99 CMML cases). The downward trend in mortality over the past 2 decades found here concurs with data published earlier, covering the period 1968 to 1997[11]. No association between mortality and socioeconomic class or geographic region is seen in this study. No national data in this regard has been published before for the UK.

In summary, the age and gender patterns of leukaemia incidence and mortality in the subtypes studied here are consistent with the published literature. Time trends in incidence and mortality must be interpreted with caution due to changes in case ascertainment and classification of subtypes over time. The increase in incidence of ALL, CLL and AML over the past 20 years has not been accompanied by improvements in mortality with time. Whether the increase in incidence over time is real or due to better recording, or both, remains unclear.

There is a paucity of published data examining socioeconomic class and regional variations in incidence and mortality for most leukaemia subtypes. Similar leukaemia incidence across social strata and geographic regions found in this research suggests that diagnostic services are equally accessible to all. The concomitant discrepancies in mortality, however, may suggest poorer uptake of or response to treatment among people with AML from lower social classes and those residing in the North East and the North West.

The socioeconomic class difference in survival may reflect a class bias in treatment offered to people with AML and/or greater co-morbidity in AML patients from lower socioeconomic classes. A class bias in access to treatments such as haemopoetic stem cell transplantation, for example, may however itself result from differences in co-morbidity, rather than a true class bias in treatment offered.

The following chapter examines social class differences in bone marrow transplantation and co-morbidity among patients with AML using hospital data.

MAIN FINDINGS

- **Diagnostic validity for leukaemia subtypes in THIN is good.**
- **Mortality in AML patients from lower socioeconomic classes is worse than among the better off, in the absence of a socioeconomic class trend in incidence. This is not seen for other leukaemia subtypes**.

QUESTIONS ARISING

- **Are patients from lower socioeconomic classes less likely to receive bone marrow transplantation than the better off?**
- **If so, is this accounted for by a socioeconomic class gradient in co-morbidity among patients with AML?**

5 Bone Marrow Transplantation and Socioeconomic Class in AML

5.1 ABSTRACT

Introduction: The previous study has shown that while the incidence of Acute Myeloid Leukaemia (AML) in the UK is similar across the spectrum of social class, mortality is nearly 50% greater among the most socio-economically deprived patients than among the most advantaged. **Aims and Objectives:** The aim of this study was to determine whether AML patients from lower socioeconomic classes are less likely to receive bone marrow transplantation (BMT) than those from higher socioeconomic classes, and whether any difference in access to BMT found is due to greater co-morbidity among more deprived patients than among the better off. **Methods:** Using Hospital Episode Statistics (HES) data, all incident cases of AML admitted to UK hospitals between 1998 and 2007 were identified, including those admitted as day-cases, along with all their co-existing medical diagnoses recorded in HES. All bone marrow transplants that these patients underwent during this period were also identified. The number of bone marrow transplantations undertaken in AML patients was calculated, and the results stratified by gender, age at diagnosis, year of diagnosis, degree of socioeconomic deprivation and co-morbidity. Logistic regression was used to

calculate odds ratios for bone marrow transplantation, adjusting for gender, age at diagnosis, year of diagnosis, degree of socioeconomic deprivation and co-morbidity score. **Results:** A total of 23 910 incident cases of AML were identified over this 10-year time period, of whom 1 140 (4.8%) underwent BMT. A smaller percentage of patients from lower socioeconomic classes had transplants than those from higher socioeconomic classes. Bone marrow transplantation declines with increasing socioeconomic deprivation (p for trend <0.01). Patients with AML in the most deprived socioeconomic quintile are 40% less likely to have a bone marrow transplantation than those in the most advantaged socioeconomic class (OR 0.60, p<0.01, 95% C.I. 0.49 - 0.73), even after adjusting for gender, age at diagnosis, year of bone marrow transplantation and co-morbidity. **Discussion:** This large cohort study demonstrates that AML patients from lower socioeconomic classes are less likely to undergo bone marrow transplantation than their better off counter-parts and that this phenomenon is independent of co-existing illness.

5.2 INTRODUCTION

The previous study showed that while the incidence of Acute Myeloid Leukaemia (AML) in the UK is similar across the spectrum of social class, mortality is nearly 50% greater among the most socio-economically deprived patients than among the most advantaged[72]. One possible reason for this difference in survival might be due to a class bias in treatment, such as bone marrow transplantation (BMT), and/or the result of greater co-morbidity among patients from lower socioeconomic classes. Trends in bone marrow transplantation across socioeconomic classes have not been investigated before.

Previous studies have, however, shown a social class bias in the use of chemotherapy and radiotherapy. A US study of breast cancer patients, for example, found that women with less than a high school education were 3 times more likely to receive smaller doses of adjunctive chemotherapy than those with a high school education, even after adjusting for co-morbidity, age, BMI, race and geographic region[73].

A similar phenomenon was found when lung cancer patients in Scotland were studied, with the poorest patients being 60% less likely to receive chemotherapy than their well-off counterparts[74]. Colorectal cancer patients in both the UK and the US are also up to 50% less likely to receive chemotherapy if they are from lower socioeconomic classes than if they are

better off[74 75]. Again, both of these studies were adjusted for a range of confounding factors, including co-morbidity.

5.3 AIMS AND OBJECTIVES

The aim of this study was to use Hospital Episode Statistics (HES) data to determine whether AML patients from lower socioeconomic classes are less likely to receive bone marrow transplantation than those from higher socioeconomic classes. A further aim was to determine whether any differences in BMT found was due to greater co-morbidity among more deprived patients than among the better off.

5.4 METHODS

Hospital episode statistics (HES) data were used in this study as bone marrow transplantation takes place in hospital and is likely to be more reliably recorded in HES than in THIN. As discussed earlier, these data include patient demographics such as date of birth, gender and region of residence, details of diagnoses and treatments received, as well as administrative details such as admission and discharge dates, along with the place patients were treated (NHS Trust or independent sector hospital, for example)[56]. Lower Super Output Area (LSOA), a measure of socio-economic deprivation of the patient is also included in HES data. I mapped the LSOA of patients to

Townsend Score for the purposes of this research to ensure consistency in the measure of socioeconomic class used.

Diagnosis dates are not held in HES; this was estimated for the purposes of this research. Patients were assumed to have been diagnosed with AML in the year they first had a code for AML in HES, provided they had at least one year of data in HES without any codes for AML. Patients who had a code for AML in HES from the first year they were represented in the data were censored. The time period of one year was chosen because the median survival of AML is relatively short (9½ months in this research [76]).

All incident cases of AML admitted to UK hospitals between 1998 and 2007, including those admitted as day-cases were identified. All co-existing medical diagnoses recorded in HES for these patients, along with all procedures, including bone marrow transplantation, which they underwent during all of their admissions over this period were also identified.

The number of bone marrow transplantations undertaken in AML patients was calculated, and the results stratified by gender, age at diagnosis, year of diagnosis, degree of socioeconomic deprivation and co-morbidity. Townsend Score was again used as the measure of socioeconomic deprivation. Recorded data on co-morbid illness were used to assign a co-morbidity score to each patient using the Charlson Co-morbidity Index described earlier. Logistic regression was used to calculate odds ratios for bone marrow transplantation, adjusting for gender, age at diagnosis, year of diagnosis,

degree of socioeconomic deprivation and co-morbidity score. All data management and statistical analyses were conducted using STATAv.10.0.

5.5 RESULTS

A total of 23 910 incident cases of AML were identified over this 10-year time period, of whom 1 140 (4.8%) underwent BMT. The numbers of bone marrow transplants performed across various strata in AML are shown in Table 5.1. A similar proportion of men and women, about 5%, underwent BMT. The frequency of bone marrow transplantation decreased with increasing age at diagnosis, with only 3 transplants recorded in those aged 71 or older. A smaller percentage of patients from lower socioeconomic classes had transplants than those from higher socioeconomic classes.

Table 5.1: Numbers of Bone Marrow Transplants

	BMT		TOTAL
	No	Yes (%)	
Overall	22 770	1 140 (4.8)	23 910
Gender			
Males	12 695	614 (4.6)	13 309
Females	10 008	525 (5.0)	10 533
Age at Diagnosis			
Up to 30	2 909	417 (13.0)	3 326
31 to 40	1 463	213 (13.0)	1 676
41 to 50	1 912	222 (10.0)	2 134
51 to 60	2 947	237 (7.0)	3 184
61 to 70	4 523	48 (1.0)	4 571
71 and older	9 013	3 (0.0)	9 016
Year of Diagnosis			
1997 to 1999	4 730	262 (5.2)	4 992
2000 to 2001	4 873	290 (5.6)	5 163
2002 to 2003	4 913	266 (5.1)	5 179
2004 to 2007	8 254	322 (3.8)	8 576
Townsend Score			
1	4 433	265 (5.6)	4 698
2	4 495	226 (5.6)	4 721
3	4 179	190 (4.3)	4 396
4	3 968	207 (4.9)	4 175
5	3 735	186 (4.7)	3 921
No Record	1 960	66 (3.3)	2 026
Charlson Co-morbidity Score			
0	16 174	855 (5.0)	17 029
1	3 238	108 (3.2)	3 346
2 or more	3 358	177 (5.0)	3 535

Table 5.2 shows the proportion of patients in each category of Charlson Score by socioeconomic class.

Table 5.2: Charlson Score of AML patients by Social Class

	Charlson Score					
	0		1		2 or more	
Townsend Score	%	95% C.I.	%	95% C.I.	%	95% C.I.
1	70	69 - 72	14	13 - 15	16	15 - 17
2	68	67 - 70	16	14 - 17	16	15 - 17
3	69	68 - 71	15	14 - 16	16	15 - 17
4	69	68 - 70	15	14 - 16	16	15 - 17
5	68	67 - 70	16	15 - 17	16	15 - 17
No Record	94	93 - 95	3	2 - 4	2	2 -3

Between 68 and 70% of AML patients have a Charlson Co-morbidity Score of 0, 14 -16% have a Score of 1, and 16% have a Score of 2 or more, with no significant differences across socioeconomic classes in any of the co-morbidity categories.

Odds Ratios for bone marrow transplantation in AML are shown in Table 5.3, where odds ratios are mutually adjusted for all other variables in the table (*p = test for trend across Townsend Scores 1 to 5). Bone marrow transplantation declines with increasing age at diagnosis after the age of 30 (p for trend <0.01), as well as with increasing socioeconomic deprivation (p for trend <0.01). Patients with AML in the most deprived socioeconomic quintile are 40% less likely to have a bone marrow transplantation than those in the most advantaged socioeconomic class (OR 0.60, p<0.01, 95% C.I. 0.49 - 0.73), even after adjusting for gender, age at diagnosis, year of bone marrow transplantation and co-morbidity.

Table 5.3: Odds Ratios for Bone Marrow Transplantation

	O. R.*	Std. Err.	P>z	95% C.I.
Gender				
Males	1	-	-	-
Females	0.91	0.04	0.03	0.83 - 1.99
Age at Diagnosis				
Up to 30	1	-	-	-
31 to 40	0.86	0.08	0.10	0.72 - 1.03
41 to 50	0.65	0.06	<0.01	0.54 - 0.77
51 to 60	0.40	0.04	<0.01	0.34 - 0.48
61 to 70	0.05	0.01	<0.01	0.04 - 0.07
71 and older	0.00	0.00	<0.01	0.00 - 0.00
Year of Diagnosis				
1997 to 1999	1	-	-	-
2000 to 2001	1.12	0.10	0.22	0.93 - 1.34
2002 to 2003	1.11	0.10	0.28	0.92 - 1.33
2004 to 2007	0.84	0.07	0.05	0.70 - 1.00
Townsend Score				
1	1	-	-	-
2	0.85	0.08	0.10	0.70 - 1.03
3	0.76	0.08	0.01	0.63 - 0.93
4	0.78	0.08	0.01	0.64 - 0.95
5	0.60	0.06	<0.01 *p<0.01	0.49 - 0.73
No Record	0.18	0.03	<0.01	0.14 - 0.24
Charlson Score				
0	1	-	-	-
1	1.03	0.11	0.80	0.83 - 1.27
2 or more	1.58	0.14	<0.01	1.33 - 1.89

*All odds ratios are adjusted for all other variables in the table.

No statistically significant interaction was found between Townsend score and age, or year of diagnosis. There was, however, statistically significant interaction (p=0.01) between Townsend score and gender, with greater differences in odds ratios across the socioeconomic class gradient in women than in men. Results of logistic regression (Odds Ratios) stratified by gender are shown in Table 5.4.

Table 5.4: Odds Ratios for Bone Marrow Transplantation by Gender

	O. R.**	Std. Err.	P>z	95% C.I.
MALES				
Townsend Score				
1	1	-	-	-
2	0.83	0.11	0.16	0.63 - 1.08
3	0.77	0.11	0.07	0.58 - 1.02
4	0.83	0.11	0.17	0.63 - 1.08
5	0.68	0.09	<0.01 *p=0.01	0.52 - 0.88
No Record	0.18	0.03	<0.01	0.12 - 0.26
FEMALES				
Townsend Score				
1	1	-	-	-
2	0.86	0.12	0.29	0.65 - 1.14
3	0.74	0.11	0.04	0.56 - 0.98
4	0.73	0.11	0.03	0.54 - 0.97
5	0.51	0.08	<0.01 *p<0.01	0.38 - 0.70
No Record	0.21	0.05	<0.01	0.14 - 0.34

**Adjusted for Age at diagnosis, year of diagnosis and Charlson Score

*p = test for trend across Townsend Scores 1 to 5.

5.6 DISCUSSION

This study has shown that AML patients from more deprived socioeconomic classes are less likely to undergo bone marrow transplantation than their counter-parts from more advantaged social classes, even after adjusting for the presence of recorded co-existing disease.

The main strength of this study is the large size of the study population. It has been possible to study over 23 000 incident cases of AML in the UK using data derived from hospital records. Data also included the co-existing medical conditions of AML patients, which made adjusting for co-morbidity possible. By using available data rather than questionnaires or interviews, any bias in the reporting of socioeconomic class and any social class bias in participation in the study has been eliminated. Bone marrow transplantation recording is likely to be accurate in hospital records given the highly specialised nature of the procedure.

One potential weakness of this study is that it was not possible to adjust for the cytogenetic risk group of AML patients. Based on cytogenetics at presentation, AML patients are classified into good, intermediate, and adverse risk groups, each with very different long term outcomes[77]. Good risk patients in their first remission are not transplanted, whereas adverse risk patients are almost always transplanted (subject to fitness and donor availability). Any bias introduced by this into these results will, however,

have applied across all social classes, unless patients from lower social classes are more likely to be in the good risk group than those from higher socioeconomic classes, for which there is no evidence. Although there is evidence that patients from lower social classes present later with disease symptoms in general, it is unlikely that late presentation is an important factor in AML survival given the acute presentation of the disease and its relatively poor prognosis.

The accuracy of social class classification is imperfect given that Townsend Score is not an individual measure of deprivation. This will have introduced a non-differential bias into these results, if any, i.e. both patients who had had a bone marrow transplant and those who had not will have been similarly affected. Such a bias will have moved odds ratios closer to '1'. It seems then that if it had been had been possible to perfectly adjust for socioeconomic deprivation, our results will have shown an even greater class bias.

The validity of co-morbidity recorded in HES data may also be imperfect. Any inaccuracies would, however, apply equally across all social class strata and so is unlikely to have introduced bias into these results. Furthermore, these results showed no difference in recorded co-morbidity across the social classes. Residual confounding cannot be ruled out completely, however, since only co-morbidity recorded in the hospital episode data has been taken into account. Other co-morbidities not related to hospital admission or not recorded during admission may have existed which would have resulted in incomplete adjustment for co-morbidity.

To my knowledge no studies examining the association between bone marrow transplantation and socioeconomic class have previously been published. Studies have, however, examined the associations between social class and chemotherapy in a number of cancers. Several studies found that lower socioeconomic class predicted under-use of chemotherapy in colorectal cancer (CRC) [74 75], breast cancer[73] and lung cancer[74]. Two North American studies found that low socioeconomic status was associated with under-use of adjuvant chemotherapy in both breast[73] and colorectal cancers[75], and postulated that this was in part due to a combination of poor access to care, financial barriers and physicians' assumptions and biases regarding patients from lower socioeconomic classes. Examples of the latter include assumptions that adequate social and monetary support may not be available, that patients from lower socioeconomic classes have lower expectations of treatment, and that they are less likely to comply with treatment. A further study concluded that lower incomes, absent or limited insurance cover and poorer education reduced access to high-quality adjuvant chemotherapy, which in turn reduced survival in breast cancer[78].

In the UK, a Scottish study showed that patients from the poorest deprivation quintile were less likely to receive chemotherapy for lung cancer and colorectal cancer than the most advantaged patients even after adjusting for age, tumour stage at diagnosis, health authority and distance from oncology centre[74]. Delay between referral and treatment was similar across all social classes and so did not explain the findings. Although this study did not adjust for co-morbidity, another Scottish study which had done so also found poorer

survival in colorectal cancer patients from the most deprived socioeconomic quintiles, in a study population which showed no correlation between socioeconomic deprivation and co-morbidity[79]. The findings of these studies, in the UK healthcare setting where access to treatment is equal and free, suggest that decision-making (by both physicians and patients) regarding chemotherapy may be influenced by non-clinical factors.

This large cohort study demonstrates that AML patients from lower socioeconomic classes are less likely to undergo bone marrow transplantation than their better off counter-parts and that this phenomenon is independent of co-existing illness. It is very likely that a similar gradient for BMT exists in other leukaemia subtypes, which have not been explored here. While this apparent gradient may be the result of incomplete adjustment for co-morbidity, non-clinical factors such as those discussed above offer possible alternative explanations for this phenomenon. Bone marrow transplantation is widely available across the UK, at transplant centres in both district general hospitals and tertiary care hospitals[80], and so regional variation in availability seems an unlikely explanation for these findings. (See Appendix C)

Furthermore, a socio-economic class gradient in chemotherapy uptake in leukaemia is also likely. While such a gradient has been found for other cancers as discussed above, it was not possible to investigate this as part of this research, given that details of in-patient prescriptions (such as those for chemotherapy) are not held in THIN or HES data.

MAIN FINDING

- **AML patients from lower socioeconomic classes are less likely to undergo bone marrow transplantation than their better off counter-parts, even when co-morbidity is accounted for.**

6 Non-steroidal Anti-inflammatory Drugs and Leukaemia Risk

6.1 ABSTRACT

Introduction: The use of non-steroidal anti-inflammatory drugs (NSAIDs) is associated with a reduced risk of developing colorectal and lung cancer. While there is some evidence that similar associations exist in leukaemia, studies have thus far been small and underpowered. **Aims and Objectives:** The aim of this large population-based case-control study, using general practice data, was to determine whether the use of NSAIDs is associated with a reduced risk of acute and chronic leukaemias, and whether their use has any impact on survival in these patients. **Methods:** The incident cases of leukaemia identified in the previous studies were used as the cases in this study. In addition at least 4 controls per case, matched by age, gender and general practice, were randomly selected from the same dataset, at the same time. Conditional logistic regression was used to determine odds ratios for NSAID prescription rates and the risk of developing several leukaemia subtypes. Cox regression was then used to determine the association between NSAID prescription rate and risk of death in leukaemia. Hazard ratios were adjusted for gender, age at diagnosis, smoking status and Townsend Score. **Results:** The risk of leukaemia overall appears to increase with increased use of NSAIDs prior to diagnosis. This is not seen when

individual leukaemia subtypes are examined, however, except perhaps in CLL where patients who had received 2-5 prescriptions/year are 29% more likely to be diagnosed with CLL than those who had not had any NSAID prescriptions (O.R. 1.29, p=0.05, 95% C.I. 1.00-1.67). There was no statistically significant association between exposure to NSAIDs prior to leukaemia diagnosis and survival. **Discussion:** This study provides strong evidence that the use of NSAIDs does not reduce the risk of developing leukaemia, nor do they improve survival.

6.2 INTRODUCTION

Observational studies have suggested that the use of non-steroidal anti-inflammatory drugs (NSAIDs) is associated with a reduced risk of developing and/or dying from a number cancers including gastrointestinal and lung cancers[81-86].

A UK population-based case-control study published in 2000, for example, found a significant inverse relationship between the incidence of oesophageal, gastric and colon cancer, and the number NSAID prescriptions patients received in the 2 to 3 years prior to their cancer diagnosis[81]. The size of the protective effects in that study ranged from a 26% reduction in the risk of colon cancer to a 50% reduction in risk of gastric cancer when patients received at least 7 NSAID prescriptions per year. US studies had earlier reported a 22% reduction in the risk of colon cancer when aspirin was used at least twice per week in a cohort of men[82], and a reduction in the risk of fatal gastric and colon cancers of over 40% in patients who had aspirin more than 16 times per month compared to those who took no aspirin[83]. Subsequent randomised control trials have confirmed that NSAIDs have a protective effect in the risk of developing colorectal cancer, with quoted risk reductions of between 26 and 42%[87 88].

In terms of lung cancer, a Danish study found a 50% reduction in risk in people who had at least 4 NSAID prescriptions/year in the 3 years prior to

their diagnosis[84]. A meta-analysis found risk reductions in lung cancer of between 21 and 32% in people who used NSAIDs compared to those who did not, with greater effects seen with increased duration of use[85].

While there is some provisional evidence that the use of NSAIDs may also reduce the risk of developing certain leukaemias[37 38], these studies have been hampered by small numbers of cases. A case-control study of 412 cases of AML appeared to show a risk reduction of 50% in those who used NSAIDs for at least 4 weeks in the 8 to 10 years prior to diagnosis, but results did not reach statistical significance[37]. A cohort study of 81 leukaemia cases found a 55% reduction in incidence when aspirin was taken at least twice/week[38]. The reported risk reduction was 70% in AML, and 62% in CLL, but there were only 5 AML cases and 8 CLL cases in that study. Others reported finding no association between NSAID use and risk of leukaemias[83], or a protective effect with aspirin, but not non-aspirin NSAIDs[38].

NSAIDs have the potential to protect against malignancy in a number of possible ways. The mechanism of action in colorectal cancer is believed to be through cyclo-oxygenase (COX) inhibition, which in turn inhibits prostaglandin synthesis and prostaglandin-induced cellular immunity[38 88]. Inhibition of angiogenesis, induction of apoptosis, disruption of signal-transduction pathways, and inhibition of oxidative DNA damage have also been suggested as mechanisms whereby NSAIDs may play a chemo-protective role any malignancy[38].

6.3 AIMS AND OBJECTIVES

This large population-based case-control study utilised general practice data to determine whether the use of NSAIDs is associated with a reduced risk of acute and chronic leukaemias, and whether their use has any impact on survival in these patients.

6.4 METHODS

'The Health Improvement Network' (THIN) dataset was used as before. Data held include patient demographic data, Townsend score of socioeconomic deprivation, as well as patients' medical and prescribed-drug histories[43].

The incident cases of leukaemia identified in the dataset for the previous studies were used as the cases in this study. At least 4 controls for each case were randomly selected from the same dataset at the same time, and were matched to controls by age (to within 1 year), gender and general practice. Controls had to be alive on the date the index case was diagnosed with leukaemia, and had to have been contributing data for at least 12 months prior to that date.

The NSAID prescription rate in the period prior to diagnosis (or index date for controls) was calculated as the number of NSAID prescriptions per year of follow-up. Prescriptions within the 12 months prior to diagnosis (or index date

for controls) were excluded from the analysis to minimize any potential bias introduced by those NSAIDs prescribed for symptoms of leukaemia, such as bone pain, for example. Exposure to NSAIDs was then categorised into 4 groups for analyses: No NSAID prescriptions, less than 2 prescriptions per year, 2 to 5 prescriptions per year, and more than 5 prescriptions per year.

Smoking status was included as a potential confounder given that the literature on its association with leukaemia remains inconclusive. Smoking status was categorised into 2 groups: those who had never smoked, and those who were current or ex-smokers. The reason for this categorisation relates to the recording of smoking status in THIN, which is discussed in greater detail in the next chapter.

Townsend Score was again used as the measure of socioeconomic deprivation.

Co-morbidity was also adjusted for, as the use of NSAIDs may be associated with co-morbid conditions. As THIN data include the medical histories of patients, it was possible to identify co-morbidity in all cases and controls. I used the Charlson Co-morbidity Index to assign a co-morbidity score to each case and control. Co-morbidity was then categorized into 4 ordered groups and included in the analysis as a categorical variable.

Conditional logistic regression was used to determine odds ratios for NSAID prescription rates and the risk of leukaemia overall, as well as for 4 subtypes

of interest. Odds ratios were adjusted for Townsend Score, smoking status and Charlson Co-morbidity Index.

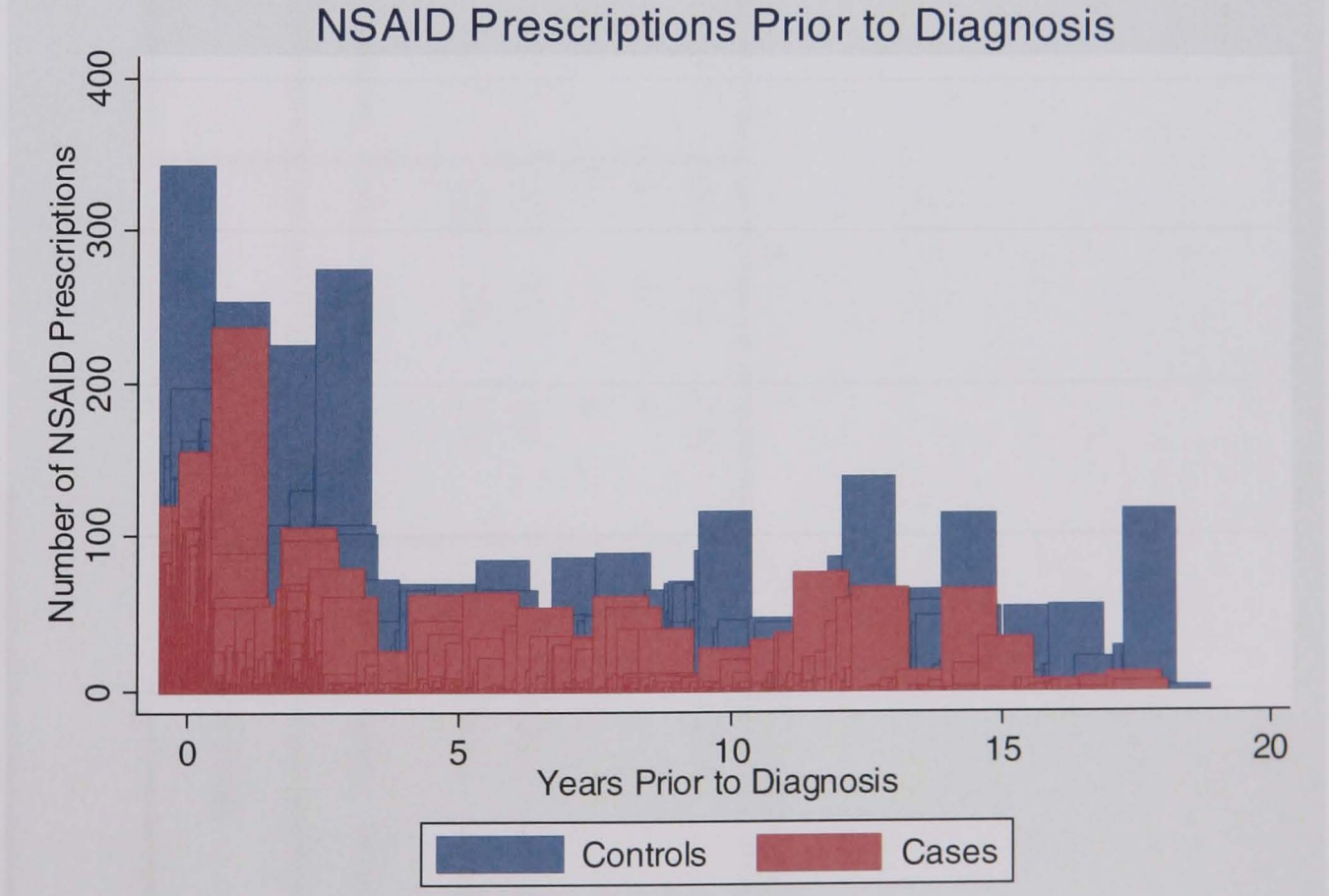
Cox regression was then used to determine the association between NSAID prescription rate and risk of death in people with leukaemia overall, and in those diagnosed with one of the leukaemia subtypes of interest. Hazard ratios were adjusted for gender, age at diagnosis, smoking status and Townsend Score.

Ethical approval was obtained from the Nottingham Research Ethics Committee.

6.5 RESULTS

3 226 patients had been identified as having one of the following diagnoses: Acute Lymphocytic Leukaemia (ALL); Chronic Lymphocytic Leukaemia; Unspecified Lymphocytic Leukaemia; Acute Myeloid Leukaemia; Chronic Myeloid Leukaemia; and Unspecified Myeloid Leukaemia, which accounted for 78% of all leukaemias identified in the dataset. The number of NSAID prescriptions prior to leukaemia diagnosis (or index date for controls) is shown in Figure 6.1 for these 3 226 patients.

Figure 6.1: NSAID Prescriptions Prior to Leukaemia Diagnosis



NSAID prescription rates prior to leukaemia diagnoses are shown in Table 6.1 for cases and controls.

Table 6.1: NSAID Prescription Rates in Leukaemia

Prescription Rate*	Overall			ALL			CLL			AML			CML		
	Controls	Cases	Total	Controls	Cases	Total	Controls	Cases	Total	Controls	Cases	Total	Controls	Cases	Total
No NSAIDs	5 915	1 436	7 351	506	127	633	2 713	630	3 343	1 058	288	1 346	510	129	639
< 2 /year	4 228	1 153	5 381	105	34	139	2 201	587	2 788	846	225	1 070	338	90	428
2-5 /year	577	175	752	8	1	9	310	95	405	107	36	143	43	11	54
> 5/year	462	117	579	6	3	9	227	55	282	88	20	108	47	11	58
Total	11 182	2 881	14 063	625	165	790	5 451	1 367	6 818	2 099	568	2 667	938	241	1 179

*Number of prescriptions per year.

Odds ratios for leukaemia adjusted for Townsend Score, smoking status and Charlson Score are shown in Table 6.2 for leukaemias overall, Table 6.3 for ALL, Table 6.4 for CLL, Table 6.5 for AML and Table 6.6 for CML. The risk of leukaemia overall appears to increase with increased use of NSAIDs, up to 5 prescriptions per year prior to diagnosis. This is not seen at higher prescription rates, however. In bi-variate analyses, Townsend Score and smoking do not alter odds ratios, but the inclusion of co-morbidity tends to decrease ORs toward 1. Multi-variate analysis shows that the risk of leukaemia overall appears to increase with increased use of NSAIDs, up to 5 prescriptions per year prior to diagnosis. Townsend Score, smoking and Charlson Score were all included in the multi-variate analysis in order to be consistent with subtype analyses.

The use of NSAIDs is not associated with incident leukaemia in any of the subtypes studied here, except perhaps in CLL where patients who had received 2-5 prescriptions/year are 29% more likely to be diagnosed with CLL than those who had not had any NSAID prescriptions (O.R. 1.29, p=0.05, 95% C.I. 1.00-1.67). No association is seen at higher NSAID prescription rates though. As is seen in these tables, the association of incident leukaemia with Townsend Score, smoking and Charlson Score varies between level of NSAID exposure, and by leukaemia subtype. For this reason all 3 these variables were in included in all the multivariate analyses to ensure consistency.

Table 6.2: Odds Ratios for Leukaemias Overall

	Odds Ratio	P>z	95 % C.I.
Uni-variate Analyses			
Number of Prescriptions			
No NSAIDS	1	-	-
<2/year	1.15	0.00	1.05-1.27
2-5/year	1.28	0.01	1.06-1.55
>5/year	1.06	0.59	0.85-1.33
Townsend Score			
1	1	-	-
2	1.10	0.14	0.97-1.25
3	1.01	0.88	0.88-1.16
4	0.99	0.92	0.86-1.15
5	0.90	0.25	0.76-1.07
No Record	1.06	0.67	0.80-1.42
Smoking			
Never Smoked	1	-	-
Current/Ex	0.95	0.25	0.86-1.04
No Record	0.93	0.36	0.79-1.09
Charlson Score			
0	1	-	-
1	1.17	0.01	1.04-1.31
2-5	1.51	<0.001	1.36-1.69
6+	1.49	0.00	1.17-1.92
Bi-variate Analyses			
Adjusted for Townsend Score:			
No NSAIDS	1	-	-
<2/year	1.15	0.00	1.05-1.27
2-5/year	1.28	0.01	1.06-1.55
>5/year	1.06	0.59	0.85-1.33
Adjusted for smoking:			
No NSAIDS	1	-	-
<2/year	1.15	0.00	1.05-1.27
2-5/year	1.28	0.01	1.06-1.55
>5/year	1.06	0.58	0.85-1.33
Adjusted for Charlson Score:			
No NSAIDS	1	-	-
<2/year	1.12	0.03	1.01-1.23
2-5/year	1.23	0.03	1.02-1.48
>5/year	1.00	0.99	0.80-1.25
Multi-variate Analyses*			
No NSAIDS	1	-	-
<2/year	1.13	0.02	1.02-1.24
2-5/year	1.23	0.03	1.02-1.49
>5/year	1.00	0.97	0.80-1.25

*Adjusted for Townsend Score, smoking and Charlson Score.

Table 6.3: Odds Ratios for ALL

	Odds Ratio	P>z	95 % C.I.
Uni-variate Analyses			
Number of Prescriptions			
No NSAIDS	1	-	-
<2/year	1.36	0.26	0.80-2.31
2-5/year	0.44	0.46	0.05-3.82
>5/year	2.55	0.22	0.57-11.35
Townsend Score			
1	1	-	-
2	0.81	0.44	0.47-1.39
3	0.98	0.93	0.54-1.75
4	0.58	0.08	0.31-1.07
5	0.82	0.59	0.40-1.69
No Record	1.13	0.83	0.37-3.46
Smoking			
Never Smoked	1	-	-
Current/Ex	0.68	0.18	0.38-1.20
No Record	1.35	0.29	0.78-2.28
Charlson Score			
0	1	-	-
1	0.86	0.59	0.51-1.47
2-5	3.84	<0.001	2.00-7.38
6+	3.80	0.16	0.58-24.70
Bi-variate Analyses			
Adjusted for Townsend Score:			
No NSAIDS	1	-	-
<2/year	1.33	0.30	0.78-2.26
2-5/year	0.45	0.47	0.05-3.92
>5/year	2.30	0.28	0.51-10.39
Adjusted for smoking:			
No NSAIDS	1	-	-
<2/year	1.51	0.13	0.88-2.60
2-5/year	0.45	0.48	0.05-4.00
>5/year	2.65	0.20	0.59-11.84
Adjusted for Charlson Score:			
No NSAIDS	1	-	-
<2/year	1.26	0.43	0.72-2.20
2-5/year	0.73	0.77	0.08-6.24
>5/year	2.54	0.26	0.50-12.87
Multi-variate Analyses*			
No NSAIDS	1	-	-
<2/year	1.39	0.25	0.79-2.45
2-5/year	0.72	0.77	0.08-6.42
>5/year	2.18	0.36	0.42-11.28

*Adjusted for Townsend Score, smoking and Charlson Score.

Table 6.4: Odds Ratios for CLL

	Odds Ratio	P>z	95 % C.I.
Uni-variate Analyses			
Number of Prescriptions			
No NSAIDS	1	-	-
<2/year	1.18	0.02	1.02-1.35
2-5/year	1.40	0.01	1.09-1.82
>5/year	1.10	0.57	0.80-1.51
Townsend Score			
1	1	-	-
2	1.08	0.45	0.89-1.30
3	1.04	0.73	0.85-1.26
4	1.05	0.66	0.85-1.30
5	0.99	0.95	0.78-1.27
No Record	0.87	0.52	0.57-1.33
Smoking			
Never Smoked	1	-	-
Current/Ex	0.99	0.91	0.87-1.13
No Record	0.64	0.00	0.49-0.84
Charlson Score			
0	1	-	-
1	1.26	0.01	1.06-1.49
2-5	1.74	<0.001	1.49-2.02
6+	2.13	<0.001	1.54-2.94
Bi-variate Analyses			
Adjusted for Townsend Score:			
No NSAIDS	1	-	-
<2/year	1.18	0.02	1.02-1.35
2-5/year	1.40	0.01	1.08-1.81
>5/year	1.10	0.57	0.80-1.51
Adjusted for smoking:			
No NSAIDS	1	-	-
<2/year	1.15	0.05	1.00-1.32
2-5/year	1.38	0.02	1.06-1.78
>5/year	1.08	0.65	0.79-1.49
Adjusted for Charlson Score:			
No NSAIDS	1	-	-
<2/year	1.13	0.09	0.98-1.30
2-5/year	1.31	0.04	1.01-1.70
>5/year	1.02	0.96	0.74-1.41
Multi-variate Analyses*			
No NSAIDS	1	-	-
<2/year	1.12	0.13	0.97-1.28
2-5/year	1.29	0.05	1.00-1.67
>5/year	1.01	0.96	0.73-1.39

*Adjusted for Townsend Score, smoking and Charlson Score.

Table 6.5: Odds Ratios for AML

	Odds Ratio	P>z	95 % C.I.
Uni-variate Analyses			
Number of Prescriptions			
No NSAIDS	1	-	-
<2/year	0.99	0.89	0.80-1.22
2-5/year	1.22	0.36	0.80-1.84
>5/year	0.84	0.52	0.50-1.43
Townsend Score			
1	1	-	-
2	1.21	0.19	0.91-1.60
3	0.93	0.63	0.69-1.26
4	0.99	0.96	0.71-1.38
5	0.70	0.08	0.47-1.04
No Record	1.58	0.15	0.84-2.97
Smoking			
Never Smoked	1	-	-
Current/Ex	0.90	0.31	0.73-1.11
No Record	1.30	0.13	0.92-1.84
Charlson Score			
0	1	-	-
1	0.96	0.75	0.74-1.24
2-5	0.91	0.43	0.71-1.16
6+	0.40	0.03	0.18-0.90
Bi-variate Analyses			
Adjusted for Townsend Score:			
No NSAIDS	1	-	-
<2/year	1.00	0.96	0.80-1.23
2-5/year	1.25	0.30	0.82-1.90
>5/year	0.84	0.51	0.49-1.42
Adjusted for smoking:			
No NSAIDS	1	-	-
<2/year	1.02	0.87	0.82-1.26
2-5/year	1.27	0.27	0.83-1.93
>5/year	0.88	0.64	0.53-1.50
Adjusted for Charlson Score:			
No NSAIDS	1	-	-
<2/year	1.00	0.99	0.81-1.24
2-5/year	1.24	0.31	0.82-1.89
>5/year	0.86	0.58	0.50-1.46
Multi-variate Analyses*			
No NSAIDS	1	-	-
<2/year	1.04	0.73	0.84-1.29
2-5/year	1.32	0.20	0.86-2.01
>5/year	0.88	0.64	0.51-1.51

*Adjusted for Townsend Score, smoking and Charlson Score.

Table 6.6: Odds Ratios fro CML

	Odds Ratio	P>z	95 % C.I.
Uni-variate Analyses			
Number of Prescriptions			
No NSAIDS	1	-	-
<2/year	1.05	0.76	0.76-1.46
2-5/year	1.03	0.94	0.51-2.07
>5/year	0.83	0.61	0.41-1.68
Townsend Score			
1	1	-	-
2	0.71	0.16	0.44-1.15
3	1.22	0.40	0.77-1.96
4	0.83	0.47	0.50-1.38
5	0.81	0.51	0.44-1.51
No Record	1.96	0.18	0.74-5.23
Smoking			
Never Smoked	1	-	-
Current/Ex	0.69	0.02	0.50-0.95
No Record	0.64	0.10	0.37-1.09
Charlson Score			
0	1	-	-
1	1.41	0.09	0.95-2.10
2-5	1.86	0.00	1.27-2.72
6+	1.30	0.59	0.50-3.42
Bi-variate Analyses			
Adjusted for Townsend Score:			
No NSAIDS	1	-	-
<2/year	1.05	0.79	0.75-1.46
2-5/year	1.01	0.99	0.50-2.04
>5/year	0.90	0.77	0.44-1.83
Adjusted for smoking:			
No NSAIDS	1	-	-
<2/year	1.04	0.82	0.74-1.46
2-5/year	0.96	0.92	0.48-1.96
>5/year	0.84	0.63	0.41-1.71
Adjusted for Charlson Score:			
No NSAIDS	1	-	-
<2/year	1.01	0.96	0.72-1.40
2-5/year	1.00	1.00	0.49-2.03
>5/year	0.73	0.40	0.36-1.50
Multi-variate Analyses*			
No NSAIDS	1	-	-
<2/year	0.99	0.96	0.70-1.40
2-5/year	0.91	0.81	0.45-1.88
>5/year	0.82	0.59	0.39-1.69

*Adjusted for Townsend Score, smoking and Charlson Score.

The number of deaths in the study population is shown by leukaemia subtype in Table 6.7. These figures exclude those for whom smoking status was not recorded, as these patients were excluded in the subsequent calculations of hazard ratios (which are adjusted for smoking status).

Table 6.7: Number of Deaths by Leukaemia Sub-type

	Number of Cases*	Number of Deaths
ALL	78	29
CLL	1277	349
Unspecified Lymphoid Leukaemia	249	93
AML	493	301
CML	216	99
Unspecified Myeloid Leukaemia	201	128

* Excludes people for whom smoking status is not known.

Hazard ratios for death in leukaemia, adjusted for gender, age at diagnosis, Townsend Score and smoking status, are shown in Table 6.8. There is no statistically significant association between exposure to NSAIDs prior to leukaemia diagnosis and survival, except in CLL where people who had up to 2 prescriptions per year had a 25% lower mortality than those who were not prescribed any NSAIDs (H.R. 0.75, p=0.01, 95% C.I. 0.60-0.94).

Table 6.8: Adjusted Hazard Ratios for Risk of Death in Leukaemia

Prescription Rate***	Overall			ALL			CLL			AML			CML		
	H.R.*	P>z	95% C. I.	H.R.*	P>z	95% C. I.	H.R.*	P>z	95% C. I.	H.R.*	P>z	95% C. I.	H.R.*	P>z	95% C. I.
No NSAIDs	1	-	-	1	-	-	1	-	-	1	-	-	1	-	-
< 2 /year	0.96	0.58	0.84-1.10	2.63	0.07	0.92-7.54	0.75	0.01	0.60-0.94	1.17	0.21	0.91-1.50	1.00	0.99	0.64-1.56
2-5 /year	0.95	0.73	0.73-1.25	n/a**	-	-	0.74	0.21	0.46-1.18	1.54	0.07	0.96-2.47	1.68	0.21	0.75-3.78
> 5 /year	1.02	0.88	0.77-1.35	2.83	0.23	0.52-15.49	0.97	0.89	0.63-1.50	1.37	0.26	0.79-2.38	0.81	0.73	0.25-2.61

*Hazard Ratios are adjusted for gender, age at diagnosis, Townsend Score and smoking status.

**There were no deaths in this group.

***Number of prescriptions per year.

6.6 DISCUSSION

This study demonstrates that the use of NSAIDs is not associated with a reduced risk of developing leukaemia and is, if anything, associated with a small increased risk of leukaemia overall and CLL specifically. Furthermore, NSAIDs have no survival benefit in the acute and chronic leukaemias studied, except in CLL, where low prescription rates appear to be beneficial.

The main strength of this study is that the large number of incident leukaemia cases for whom there are data has facilitated a statistically powerful study. The preceding studies suggest that the leukaemia diagnoses in THIN are valid in terms of age, incidence and survival. THIN data are also a good source of population controls from which matched controls for this study could be selected. Obtaining prescription information directly from THIN, rather than relying on self-reporting by study subjects, has avoided recall bias with respect to NSAID use.

One potential weakness of this study may be inaccuracies in prescription information held within THIN. This has not had an important impact on these results given that the validity of general practice data, including that of prescription data, has been demonstrated elsewhere[44 51]. A further potential weakness of this study is that the use of over-the-counter NSAIDs has not been taken into account. This is unlikely to have biased these results since cases and controls are equally likely to have used over-the-counter NSAIDs

prior to diagnosis (or index date for controls). Furthermore, since people over the age of 60 are entitled to free prescriptions, their general practice prescription records are likely to accurately reflect their NSAID use.

Few studies have explored the association of NSAID use and risk of leukaemia. A case-control study of 412 AML cases found NSAIDs had a protective effect on the risk of one AML subtype (M2), but results did not reach statistical significance[37]. Another study of 81 post-menopausal women with leukaemia (including 35 cases of CLL, 28 of AML and 5 of CML) found a statistically significant protective effect with NSAID use in leukaemia overall, as well as with aspirin use in leukaemia overall, AML and CLL[38]. However, both these studies relied on self-reported drug histories (in the former subjects were required to recall their drug intake up to 10 years prior to their diagnosis) subjecting results to recall bias.

The small positive association of NSAID prescriptions and incident CLL found here is likely to be due to reverse causation, i.e. where NSAIDs have been prescribed for symptoms of leukaemia. The reason this is apparent in CLL, but not other subtypes studied, may be because CLL progresses more slowly than the other subtypes, and so excluding prescriptions during the 12 months prior to diagnosis (or index date for controls) may not have sufficiently eliminated the problem of reverse causation in the CLL cases.

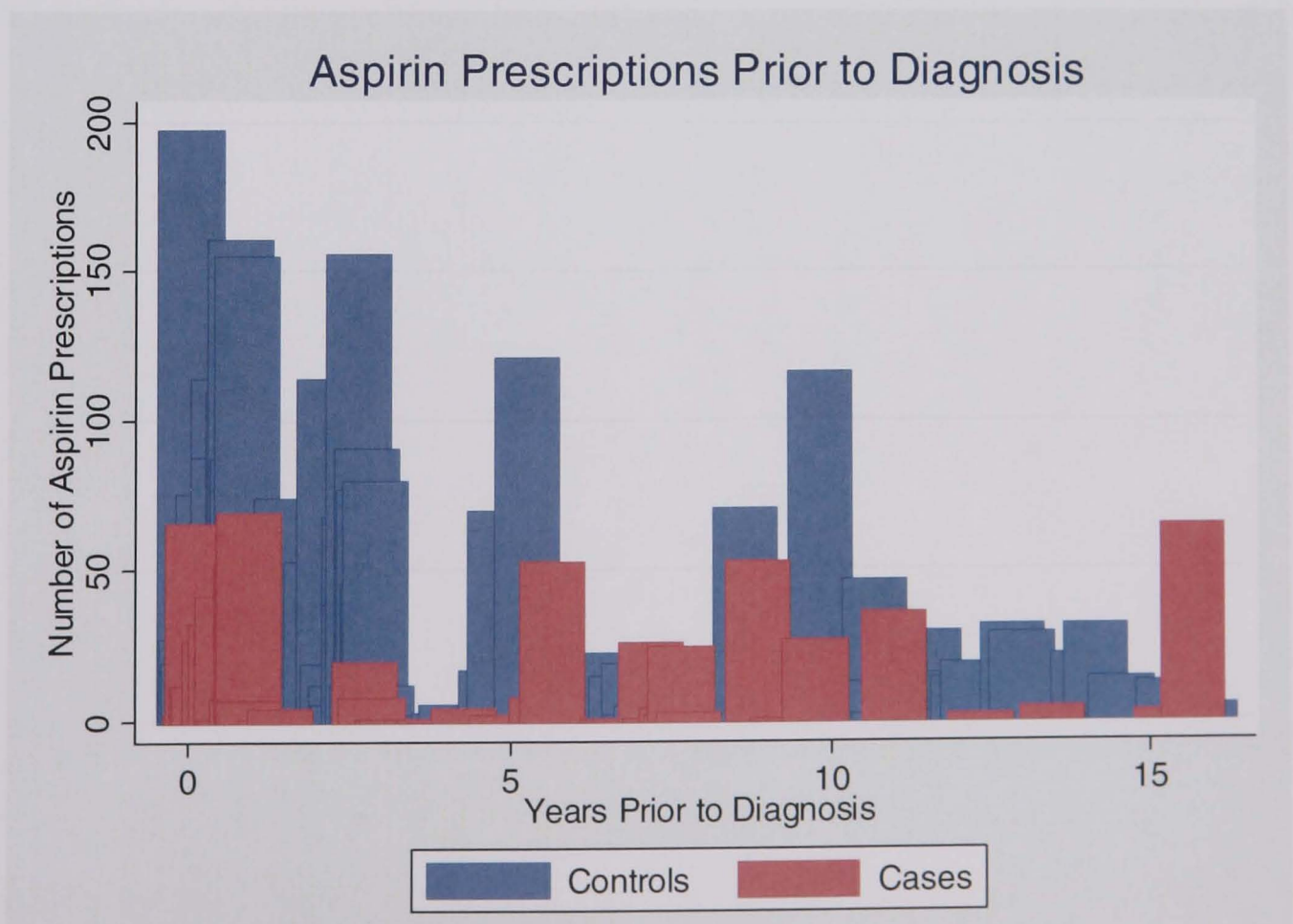
In a secondary analysis, examining only aspirin prescription rates, no association with the incidence of and mortality from leukaemia overall was

found (results not shown), which contrasts with the findings of others[38]. Aspirin prescription rates among the cases in this study are too low to allow similar analyses by leukaemia subtype (see Table 6.9 and Figure 6.2).

Table 6.9: Aspirin Prescription Rates in Leukaemias Overall

Prescription Rate	Controls	Cases	Total
No Aspirin	12 614	3 164	15 778
< 2 prescriptions/year	90	24	114
2-5 prescriptions/year	27	10	37
> 5 prescriptions/year	39	4	43
Total	12 770	3 202	15 972

Figure 6.2: Aspirin Prescriptions Prior to Leukaemia Diagnosis



This study provides strong evidence that NSAIDs do not reduce the risk of leukaemia, and in all probability do not offer any survival benefit in these cancers. In addition, no statistically significant association with aspirin use and the incidence or mortality from leukaemia overall was found in this study.

MAIN FINDING

- **NSAID use is not associated with a reduced risk of Leukaemia, nor do they offer any survival benefit.**

7 Alcohol Consumption and Risk of Leukaemia

7.1 ABSTRACT

Introduction: Alcohol consumption may produce chemo-protective effects in leukaemia via several mechanisms. The relationship between alcohol consumption and the risk of leukaemia has been explored in a number of studies, but results thus far have been inconsistent. **Aims and Objectives:** This population-based case-control study investigates the association between alcohol consumption and the incidence of and mortality from the acute and chronic leukaemias. **Methods:** The incident cases of leukaemia and the matched controls identified in THIN in the previous studies were used for this study. Conditional logistic regression was used to determine odds ratios for alcohol consumption and risk of developing leukaemia and several leukaemia subtypes. Cox regression was then used to determine the association between alcohol consumption and the risk of death in leukaemia. Hazard ratios were adjusted for gender, age at diagnosis, smoking status, Townsend Score and co-morbidity. **Results:** There is no statistically significant association between alcohol consumption and risk of developing leukaemia overall, nor with any of the leukaemia subtypes studied here. Alcohol consumption is associated with a lower risk of death in leukaemia overall (HR 0.83, p=0.04, 95% C.I. 0.69 - 0.99), as well as in ALL (HR 0.14, p<0.01, 95% C.I. 0.04 - 0.44) and CLL (HR 0.71, p=0.02, 95% C.I. 0.53 -

0.96), when compared to those who had not consumed any alcohol.

Discussion: Alcohol consumption is positively associated with survival from leukaemia even after adjusting for co-morbid illness. Better recording of alcohol consumption in THIN will enhance its value in investigating associations of alcohol consumption and disease, or other outcomes.

7.2 INTRODUCTION

A number of studies have examined the relationship between alcohol consumption and the risk of leukaemia, but results have thus far been inconsistent[24-27]. Several of these studies have found inverse associations with low or moderate alcohol intake and positive associations with high alcohol intake and leukaemia incidence, but results were often not statistically significant[24 26 27]. When individual leukaemia subtypes were investigated, studies tended to be underpowered, and again showed insignificant results[24 25 27]. To my knowledge, the impact of alcohol consumption on mortality in leukaemia has not been investigated.

Alcohol has the potential to exert its effects, both protective and damaging, via several mechanisms. There is evidence, for example, of its immunomodulatory effects whereby moderate consumption elicits an improved cellular and humoral response while heavy drinking impairs immune function[89]. Anti-oxidants contained in wine and beer, as well as the improvements in insulin sensitivity ascribed to moderate alcohol consumption may be further mechanisms through which alcohol potentially exerts anti-carcinogenic effects.

7.3 AIMS AND OBJECTIVES

This population-based case-control study investigates the association between alcohol consumption and the incidence of and mortality from the acute and chronic leukaemias.

7.4 METHODS

The incident cases of leukaemia and the controls identified in the THIN dataset for the previous study were used for this study.

Alcohol consumption was categorised into 2 groups: those who never consumed alcohol and those who were recorded as being current or ex-drinkers prior to their diagnosis of leukaemia (or index date for controls). The reason for this relatively simplistic classification is due to the incomplete recording of alcohol in THIN, which will be discussed later.

As in the previous study, smoking status was categorised into those who never smoked and those who were current or ex-smokers, again because of incomplete recording. Cases and controls were assigned the most recent smoking and alcohol status in their records prior to their diagnosis (or index date for controls). Townsend Score was also adjusted for as before.

Conditional logistic regression was used to determine odds ratios for alcohol consumption and the risk of leukaemia overall, as well as for 4 subtypes of interest. Odds ratios were adjusted for smoking status, Townsend Score as well as co-morbidity.

Cox regression was then used to determine the association between alcohol consumption and risk of death in people with leukaemia overall and in those diagnosed with one of the leukaemia subtypes of interest. Hazard ratios were adjusted for gender, age at diagnosis, smoking status, Townsend Score smoking and co-morbidity.

A secondary analysis was then conducted to quantify any bias introduced into the initial results by co-existing illness among study participants. As THIN data include the medical histories of patients, it was possible to identify co-morbidity in all cases and controls. The Charlson Co-morbidity Index[90] was used to assign a co-morbidity score to each case and control. Co-morbidity was then categorized into 4 ordered groups and included in the analysis as a categorical variable. Conditional logistic regression and Cox regression was then performed as before, but additionally adjusting for co-morbidity.

Ethical approval was obtained from the Nottingham Research Ethics Committee.

7.5 RESULTS

7.5.1 Primary Analysis

Alcohol consumption was recorded in THIN for 8 735 out of 11 182 controls, (78%), and for 2 227 out of 2 881 cases (77%). The alcohol intake of cases and controls are shown in Table 7.1.

Table 7.1: Alcohol Records

	Controls	Cases	Total
Alcohol			
Never	1 295	346	1 641
Ex	95	24	119
Current	7 345	1 857	9 202
No Data	2 447	654	3 101
Total	11 182	2 881	14 063

Alcohol consumption is not associated with an increased risk of leukaemia overall, nor with any of the subtypes studied here, when adjusted for smoking status and Townsend Score (Table 7.2).

Table 7.2: Odds Ratios for Leukaemia

	OVERALL			ALL			CLL			AML			CML		
	O.R.	P	95% C.I.	O.R.	P	95% C.I.	O.R.	P	95% C.I	O.R.	P	95% C.I	O.R.	P	95% C.I
Alcohol															
Never	1	-	-	1	-	-	1	-	-	1	-	-	1	-	-
Ever	0.95	0.49	0.83 - 1.09	1.46	0.38	0.62 - 3.44	1.05	0.65	0.86 - 1.27	0.84	0.28	0.62 - 1.15	0.79	0.33	0.49 - 1.27
No Record	1.05	0.61	0.87 - 1.26	1.16	0.75	0.45 - 3.01	1.22	0.15	0.93 - 1.59	1.13	0.53	0.76 - 1.69	0.77	0.44	0.40 - 1.49
Smoking Status															
Never	1	-	-	1	-	-	1	-	-	1	-	-	1	-	-
Ever	0.96	0.35	0.87 - 1.05	0.68	0.20	0.38 - 1.22	0.99	0.88	0.87 - 1.13	0.94	0.58	0.76 - 1.17	0.70	0.03	0.51 - 0.97
No Record	0.87	0.17	0.72 - 1.06	1.40	0.28	0.76 - 2.57	0.57	0.00	0.41 - 0.78	1.07	0.75	0.71 - 1.62	0.65	0.21	0.33 - 1.28
Townsend Score															
1	1	-	-	1	-	-	1	-	-	1	-	-	1	-	-
2	1.10	0.14	0.97 - 1.26	0.82	0.49	0.48 - 1.43	1.08	0.43	0.89 - 1.30	1.22	0.17	0.92 - 1.62	0.70	0.16	0.43 - 1.14
3	1.01	0.86	0.88 - 1.16	1.00	0.99	0.56 - 1.79	1.05	0.66	0.86 - 1.28	0.94	0.68	0.69 - 1.27	1.23	0.40	0.76 - 1.97
4	0.99	0.94	0.86 - 1.15	0.60	0.12	0.32 - 1.13	1.05	0.63	0.85 - 1.30	1.00	0.99	0.72 - 1.39	0.85	0.52	0.51 - 1.41
5	0.91	0.27	0.76 - 1.08	0.86	0.69	0.42 - 1.78	1.00	0.97	0.78 - 1.29	0.71	0.09	0.47 - 1.05	0.80	0.49	0.43 - 1.50
No Record	1.07	0.66	0.80 - 1.42	1.07	0.91	0.34 - 3.32	0.90	0.64	0.59 - 1.39	1.51	0.20	0.80 - 2.84	2.13	0.14	0.78 - 5.76

Hazard Ratios for death in leukaemia adjusted for gender, age at diagnosis, Townsend Score and smoking status are shown in Table 7.3. These data show that alcohol consumption is associated with a lower risk of death in leukaemia overall (HR 0.83, p=0.04, 95% C.I. 0.69 - 0.99), as well as in ALL (HR 0.14, p<0.01, 95% C.I. 0.04 - 0.44) and CLL (HR 0.71, p=0.02, 95% C.I. 0.53 - 0.96), when compared to those who had not consumed any alcohol.

Table 7.3: Hazard Ratios for death in Leukaemia

	OVERALL			ALL			CLL			AML			CML		
Alcohol	**H.R.***	**p**	**95% C.I.**	**H.R.***	**p**	**95% C.I.**	**H.R.***	**p**	**95% C.I.**	**H.R.***	**p**	**95% C.I.**	**H.R.***	**p**	**95% C.I.**
Never	1	-	-	1	-	-	1	-	-	1	-	-	1	-	-
Ever	0.83	0.04	0.69 - 0.99	0.14	<0.01	0.04 - 0.44	0.71	0.02	0.53 - 0.96	1.17	0.38	0.82 - 1.67	0.63	0.09	0.37 - 1.07
No Record	1.11	0.39	0.88 - 1.39	0.12	<0.01	0.03 - 0.50	0.98	0.93	0.67 - 1.43	1.30	0.24	0.84 - 2.01	1.39	0.37	0.68 - 2.85

*Hazard ratios are adjusted for gender, age at diagnosis, Townsend Score, smoking status and co-morbidity.

7.5.2 Secondary Analysis

The relationships between Charlson Co-morbidity Index and alcohol consumption patterns among cases and controls are shown in Table 7.4 and Table 7.5, respectively, and graphically in Figure 7.1.

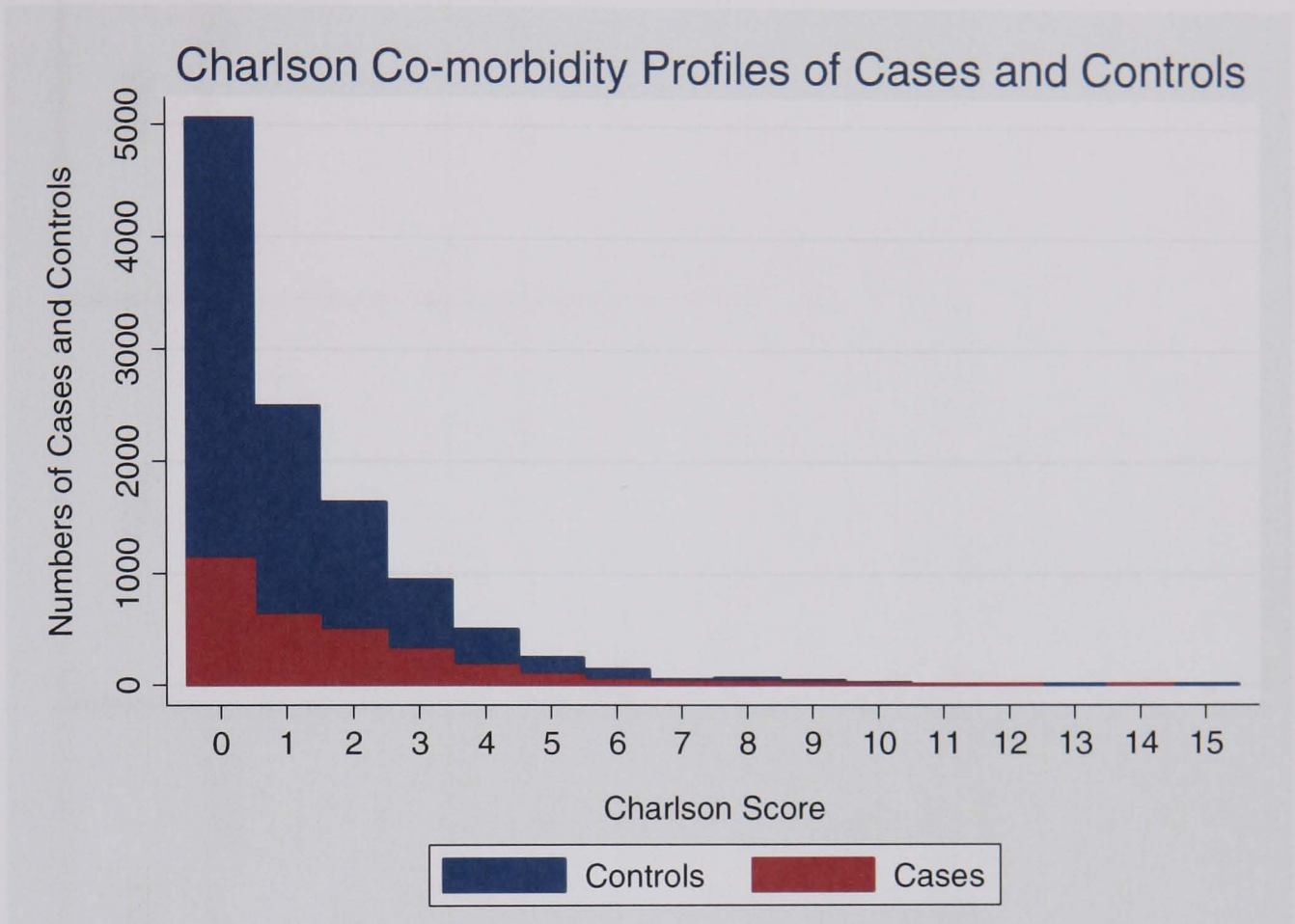
Table 7.4: Alcohol Consumption and Charlson Scores for Cases

Charlson Score	Cases Alcohol Consumption				
	No Data	Never	Ex	Current	Total
0	337	110	4	675	1 126
1	128	83	9	394	614
2 - 5	176	141	11	717	1 045
6+	13	12	0	71	96
Total	654	346	24	1 857	2 881

Table 7.5: Alcohol consumption and Charlson Score for Controls

Charlson Score	Control Alcohol Consumption				
	No Data	Never	Ex	Current	Total
0	1 463	493	22	3 075	5 053
1	483	311	16	1 684	2 494
2 - 5	459	439	47	2 372	3 317
6+	47	52	10	214	318
Total	2 447	1 295	95	7 345	11 182

Figure 7.1: Charlson Co-morbidity Profiles of Cases and Controls



The hazard ratios defining the associations between alcohol consumption and survival by leukaemia subtype are shown in Table 7.6. This shows the positive association between alcohol consumption and survival persisted in leukaemia overall, ALL and CLL, when co-morbidity is adjusted for.

Table 7.6: Hazard Ratios for death in Leukaemia (adjusted for CCI)

	OVERALL			ALL			CLL			AML			CML		
	H.R.*	P	95% C.I.	H.R.*	P	95% C.I.	H.R.*	P	95% C.I.	H.R.*	P	95% C.I.	H.R.*	P	95% C.I.
Alcohol															
Never	1	-	-	1	-	-	1	-	-	1	-	-	1	-	-
Ever	0.83	0.04	0.69 - 0.99	0.16	<0.01	0.05 - 0.53	0.72	0.03	0.54 - 0.97	1.20	0.33	0.84 - 1.71	0.65	0.12	0.37 - 1.12
No Record	1.12	0.32	0.89 - 1.41	0.12	<0.01	0.03 - 0.51	1.03	0.87	0.71 - 1.51	1.37	0.16	0.88 - 2.13	1.46	0.33	0.69 - 3.09
Charlson Score															
0	1	-	-	1	-	-	1	-	-	1	-	-	1	-	-
1	1.07	0.44	0.91 - 1.26	0.68	0.45	0.25 - 1.83	1.91	<0.01	1.37 - 2.66	1.01	0.97	0.76 - 1.34	1.24	0.45	0.71 - 2.19
2 - 5	1.06	0.44	0.92 - 1.22	0.45	0.10	0.17 - 1.18	1.81	<0.01	1.35 - 2.42	1.31	0.03	1.02 - 1.69	1.21	0.47	0.72 - 2.05
6+	1.45	0.01	1.11 - 1.90	1.69	0.64	0.19 - 15.13	3.09	<0.01	2.04 - 4.66	1.90	0.11	0.86 - 4.20	1.51	0.45	0.52 - 4.38

*Hazard ratios are also adjusted for gender, age at diagnosis, Townsend Score and smoking status.

7.6 DISCUSSION

This study found no statistically significant association between alcohol consumption and the risk of developing leukaemia. Interestingly though, alcohol consumption was found to be positively associated with a reduced risk of death from leukaemia overall, ALL and CLL.

The main strength of this study lies in the large size of the dataset from which cases were derived. Greater numbers of cases allowed leukaemia subtypes to be examined with greater statistical power than in previous studies. THIN data are also a source of general population controls, which makes large case-control studies easier and cheaper to conduct than studies that require interviews and/or questionnaires. The scope of data held within THIN, for example, age, gender and general practice, allows matching of controls to cases thus reducing potential bias. Furthermore, the availability of information on socioeconomic class, medical and smoking histories makes adjusting for these potential confounders possible. As these data are extracted directly from THIN, rather than via interviews and questionnaires, recall bias in terms of exposures is also avoided.

The greatest weakness of this study is the number of patients for whom alcohol consumption is incompletely and/or unreliably recorded. Alcohol is recorded in 2 ways in THIN: The status of patients' alcohol consumption is recorded as current, ex, or never, at various points in time within the dataset.

In addition, the number of units of alcohol consumed per week is also recorded. These both represented sources of weakness in this study.

Firstly, inconsistencies were noted in the classification of patients' alcohol consumption status, whereby people who were recorded as being drinkers or ex-drinkers at one point in time were, for example, recorded as 'never drinkers' at a later date. There were also people who were recorded as being current drinkers, when they had earlier been recorded as being ex-drinkers, with no way of knowing whether these records were erroneous or whether patients had given up alcohol and then resumed its consumption. It is for these reasons that current drinkers and ex-drinkers were grouped together into a single category for the purposes of this study. Smoking records in THIN have similar inaccuracies and hence a similar, relatively simplistic categorisation was adopted for these studies. Only people who were consistently recorded as being 'never drinkers' were classified as such. While this approach improved the reliability of the classification of patients' alcohol consumption status, it meant that 'current ' exposure to alcohol could not be distinguished from ‘previous’ exposure to alcohol in the analyses. This is an important consideration as any influence alcohol may have aetiologically or with respect to survival, may differ between current and ex-drinkers.

Incomplete recording of the number of units of alcohol consumed per week further weakened this study. While 2 443 out of 11 182 controls, i.e. 22% of controls have no record of alcohol consumption status and no record of the number of units of alcohol consumed, an additional 2 673 of the 7 349

controls i.e. 36% of controls recorded as being current drinkers, also have no record of the number of units consumed per week. Records are similarly incomplete for cases with 654 out of 2 881 cases, i.e. 23% having no record of alcohol consumption status and no record of the number of units of alcohol consumed, and a further 672 out of 1 857 cases i.e. 36% of cases recorded as being current drinkers, also having no record of the number of units consumed per week. One can conclude that since records are equally incomplete for cases and controls, bias in terms of whether or not the amount of alcohol consumed is recorded for cases versus controls is not likely to be an important issue. The paucity of these data does, however, mean that investigating the existence of a dose-response relationship between alcohol consumption and leukaemia incidence and/or survival reliably is difficult.

The relationship between alcohol consumption and leukaemia has been examined previously, but results have been inconsistent. In a case-control study of 578 white men with leukaemia, alcohol consumption was positively associated with ALL, but not leukaemia overall or CML[24]. Odds ratios did not reach statistical significance, however, nor did investigators find a dose response gradient with the quantity of alcohol consumed, a pattern which may have suggested causality, had it been found.

In another case-control study of 164 case-control pairs, Pogoda et al. found that alcohol consumption was associated with a decreased risk of AML in adults, but results did not reach statistical significance[25]. The authors acknowledge that alcohol consumption in their study population was related

to higher socioeconomic class, and that their controls were also of higher socioeconomic class than their cases. As only education was controlled for in the analysis, this combination of factors may have resulted in residual confounding with respect to socioeconomic class and consequently produced spuriously low odds ratios.

In another case-control study of 765 incident cases of acute de novo leukaemia in adults and 618 controls, regular drinkers had a reduced relative risk of leukaemia compared to non-drinkers[26]. When consumption patterns and different types of alcohol were examined, light and moderate beer intake was associated with a reduced risk of leukaemia that was statistically significant, while moderate or heavy wine intake was associated with an increased relative risk of leukaemia, although results for the latter did not reach statistical significance. As in the study by Pagoda et al.[25], alcohol consumption was more prevalent among controls than cases and education was the only marker of socioeconomic class adjusted for. Again this may have resulted in downwardly biased odds ratios if the controls were of higher socioeconomic class than the cases in this study. Unfortunately, the similarity or otherwise of cases and controls with respect to socioeconomic class is not reported in the paper. Selection bias related to social class may also explain the inverse association found with beer if, for example, beer intake is inversely associated with socioeconomic class and controls were of lower socioeconomic class than cases.

A more recent multi-centre case-control study exploring the associations of alcohol intake with a number of leukaemia subtypes was also inconclusive[27]. In this study of 649 cases and 1 771 controls, any alcohol intake was found to be associated with a reduced risk of leukaemia overall, ALL and CLL, compared to those who never drank alcohol. The contrary was true for AML and CML. None of these odds ratios reached statistical significance, however. This study also examined the type and quantity of alcohol consumed and found an inverse association with leukaemia overall and moderate intake of all alcohol, wine, beer and spirits. Similar inverse associations were seen in ALL, CLL and AML with all alcohol. A positive association with leukaemia overall was found for high-level consumption of all alcohol, wine and beer. Again, results did not reach statistical significance. CML was positively associated with alcohol consumption at all levels, but odds ratios again did not reach statistical significance.

While the finding that survival from leukaemia overall is better in those who currently consume or have previously consumed alcohol, than in those who never did may be a true phenomenon, it may also be due to bias or confounding. Recall bias has been avoided in this study by utilising alcohol consumption records held in the dataset, rather than conducting interviews or questionnaires retrospectively, as was the case in many of the studies reported above. The study design has minimized reporting bias with respect to alcohol consumption by using alcohol consumption records that pre-date patients' leukaemia diagnosis. Potential confounding by gender, age at diagnosis, Townsend Score and smoking status have been adjusted for in

the analyses. Better survival among current and ex-drinkers may be explained by the 'healthy drinker' effect, whereby patients with less debilitating or less symptomatic co-morbid illness may be more likely to continue to drink alcohol.

A secondary analysis was therefore conducted adjusting for confounding by co-morbidity (using the Charlson Co-morbidity Score as before), which demonstrated that the positive association between alcohol consumption and survival persisted in leukaemia overall, ALL and CLL.

This adjustment for confounding by co-morbid disease does not, of course, take into account the stage of leukaemia or the severity of leukaemia symptoms. The negative association of alcohol with mortality may still, therefore, be explained by the 'healthy drinker' effect.

The alternative explanation, that alcohol really is protective, has some biological plausibility in that alcohol is known to be toxic to bone marrow and white blood cells[91 92], with the latter study showing that although alcohol was toxic to both leukaemic and non-leukaemic lymphocytes, its toxic effect was greater on leukaemic lymphocytes.

In summary, this study has shown that incident leukaemia is independent of alcohol consumption, but that the latter is associated with a lower risk of death in leukaemia overall, ALL and CLL. There is under-recording of the amount of alcohol patients consume in this dataset, in which over a third of

current drinkers have no record of the number of units of alcohol they consume. This is an important consideration as any influence alcohol may have aetiologically or with respect to survival, may differ between current and ex-drinkers. Improving this will facilitate the investigation of any dose-response relationship with incidence and mortality in leukaemia, as well as other diseases, at population level.

MAIN FINDINGS

- **Alcohol consumption is not associated with incident leukaemia, but it is associated with reduced mortality in leukaemia overall, ALL and CLL.**
- **The amount of alcohol patients consume is under-recorded in THIN, and improvements in this will further enhance THIN as resource for epidemiology research.**

8 CONCLUSIONS

8.1 SUMMARY OF FINDINGS

• THIN data is a valuable resource for leukaemia research with diagnostic validity comparable to cancer registries.

• Mortality in leukaemia patients overall from lower socioeconomic classes is worse than among the better off, despite no class gradient in incidence.

• Mortality in AML patients from lower socioeconomic classes is worse than among the better off, in the absence of a socioeconomic class trend in incidence. This is not seen for other leukaemia subtypes.

• AML patients from lower socioeconomic classes are less likely to undergo bone marrow transplantation than their better off counter-parts, even after adjusting for co-morbidity.

• NSAID use is not associated with a reduced risk of leukaemia, nor do they offer any survival benefit.

• Alcohol consumption is not associated with incident leukaemia, but it is associated with reduced mortality in leukaemia overall, ALL and CLL.

• The amount of alcohol patients consume is under-recorded in THIN, and improvements in this will further enhance THIN as resource for epidemiology research.

8.2 CLINICAL IMPLICATIONS AND FUTURE RESEARCH

This research firstly provides reassurance that diagnostic services for leukaemia in the UK are accessible to all patients regardless of socioeconomic class and geographic region. Mortality, however, is greater among socio-economically disadvantaged leukaemia patients, specifically those with AML.

While greater mortality among patients from lower socioeconomic classes is by no means unique to leukaemia, the reasons for lower bone marrow transplantation rates among poorer patients with AML does warrant further investigation. Lower bone marrow transplantation rates in this group do not necessarily explain the greater mortality seen, since mortality is influenced by multiple factors, including cytogenetic profile as discussed earlier, and the use of and response to chemotherapy, as well as other risk factors that were able to be investigated here. Cytogenic information is not held in THIN, nor are details of in-patient prescriptions, such as chemotherapy, hence neither of these factors could be studied in this project.

The fact that lower uptake of bone marrow transplantation appears to be related to non-clinical factors rather than reflecting greater co-morbidity in patients from lower socioeconomic classes is of concern. While a social class gradient in uptake of bone marrow transplantation has not been shown before, others have found such gradients in the uptake of chemotherapy, as

discussed in Chapter 5. Possible reasons suggested by these investigators include: lower incomes, absent or limited insurance cover and poorer education result in reduced access to these therapies; and that health professionals' assumptions that adequate social and monetary support may not be available; that patients from lower socioeconomic classes have lower expectations of treatment; and that they are less likely to comply with treatment.

While factors such as health insurance do not apply in the UK setting, the other factors that may influence uptake of bone marrow transplants (and chemotherapy) mentioned above are relevant in the UK and can and should be investigated. One way of doing this would be for transplant centres to conduct audits to assess local patterns of bone marrow transplantation. Centres should establish whether these social class trends in bone marrow transplantation are apparent locally. If so, an examination into their own practice would be justified to establish the reasons for any such trend. Local audits could also examine whether there are social class trends in chemotherapy uptake, as they will have access to in-patient prescription data. Only once in-patient prescription data are linked to other databases will it be possible to examine this issue efficiently at a national level.

This work has also demonstrated that THIN data provide a valuable resource for conducting epidemiological research. While the diagnostic validity in THIN has been demonstrated for many acute and chronic diseases, the additional health-related data, such as co-morbidity, prescription data, and information

on smoking and alcohol consumption, held in the dataset also allow a range of variables to be included in study designs and analyses. This represents a distinct advantage over registry data as a resource for epidemiological research, at least until such time as linkage of these data sources has been completed.

The impact of several lifestyle factors, such as alcohol consumption, smoking and obesity is easier, quicker and cheaper to investigate in these data without the need for questionnaires and interviews. Using routinely collected data rather than the latter methods reduces reporting bias and excludes recall bias, both of which improve the quality of epidemiological studies. This research has shown that any dose-response effects of alcohol and/or smoking consumption will be more accurately ascertained if these data were more completely recorded, and represents an area of current weakness in general practice datasets, although with improvements in recording of these additional health data, the range and quality of studies possible using these data will be enhanced.

THIN data are also a valuable resource for studies of pharmacoepidemiology given the robustness of prescription data held. This is an area where this dataset is under-utilized despite its huge potential. In terms of patient safety, for example, these data represent 'real-world' use of medications in an unselected population compared to clinical trial data. These data are not only useful for calculating background incidence rates of disease or symptoms in

the general population, but is also invaluable in determining these rates in sub-populations of patients taking specific treatments, for example.

While this research only presents one such example, the potential protective effects of a range of drugs can be investigated with relative ease using these data. Similarly, where there is speculation surrounding the safety of a particular drug or drug class, this can be addressed by pharmacoepidemiology studies demonstrating its safety retrospectively in large numbers of exposed patients.

Another potential area of pharmacoepidemiology research that is of public health importance is in demonstrating the safety of medicines retrospectively. Women of childbearing potential and children, for example, are excluded from most clinical trials of new medicines. This results in a lack of safety data for these populations and they are consequently not prescribed potentially beneficial medication. Where such medication is prescribed 'off licence', this exposure will be captured in these data and safety in these populations can then be demonstrated retrospectively. Examples include the use of steroids in pregnancy for exacerbations of asthma and inflammatory bowel disease, amongst others.

A range of fairly complex research questions can therefore be answered by using general practice data such as THIN. Furthermore, these studies can be conducted relatively quickly and cheaply using these data, and more resources should be made available to facilitate this.

REFERENCES

1. Cancer Research UK. UK Leukaemia Statistics. *CancerStats*. London: Cancer Research UK, 2007.
2. Hughes-Jones NC, Wickramasinghe SN, Hatton C. *Haematology*. 7th ed. Oxford: Blackwell Publishing, 2004.
3. Ghia P, Ferreri AM, Galigaris-Cappio F. Chronic lymphocytic leukemia. *Critical Reviews in Oncology-Hematology* 2007;64(3):234-46.
4. Goldman JM. Chronic Myeloid Leukaemia. In: Whittaker JA, Holmes JA, editors. *Leukaemia and Related Disorders*. Oxford: Blackwell Science Ltd, 1998:252 - 269.
5. Donadieu J, Auclerc MF, Baruchel A, Leblanc T, Landman-Parker J, Perel Y, et al. Critical study of prognostic factors in childhood acute lymphoblastic leukaemia: differences in outcome are poorly explained by the most significant prognostic variables. Fralle group. French Acute Lymphoblastic Leukaemia study group. *British Journal of Haematology* 1998;102(3):729-39.
6. Cartwright RA. Epidemiology. In: Whittaker JA, Holmes JA, editors. *Leukaemia and Related Disorders*. Oxford: Blackwell Science Ltd, 1998:1- 30.
7. Cartwright RA, Gurney KA, Moorman AV. Sex ratios and the risks of haematological malignancies. *British Journal of Haematology* 2002;118(4):1071-7.
8. Stiller CA, Kroll ME, Boyle PJ, Feng Z. Population mixing, socioeconomic status and incidence of childhood acute lymphoblastic leukaemia in England and Wales: analysis by census ward. *British Journal of Cancer* 2008;98(5):1006-11.
9. Alston R, Rowan S, Eden T, Moran A, Birch J. Cancer Incidence Patterns by Region and Socioeconomic deprivation in Teenagers and Young Adults in England. *British Journal of Cancer* 2007;96(11):1760-1766.
10. Dockerty JD, Draper G, Vincent T, Rowen SD, Bunch KJ. Case-control study of parental age, parity and socio-economic level in relation to childhood cancers. *International Journal of Epidemiology* 2001;30(6):1428-1437.
11. Swerdlow A, Silva I, Doll R. *Cancer Incidence and Mortality in England and Wales: trends and risk factors.* Oxford: Oxford University Press, 2001.
12. Bailey CC, Kinsey SE. Acute Lymphoblastic Leukaemias. In: Whittaker JA, Holmes JA, editors. *Leukaemia and Related Disorders*. Oxford: Blackwell Science Ltd, 1998:87 - 104.
13. Chessells JM, Richards SM, Bailey CC, Lilleyman JS, Eden OB. Gender and treatment outcome in childhood lymphoblastic leukaemia: report from the MRC UKALL trials. *British Journal of Haematology* 1995;89(2):364-72.

14. Coebergh JW, Pastore G, Gatta G, Corazziari I, Kamps W, Group EW. Variation in survival of European children with acute lymphoblastic leukaemia, diagnosed in 1978--1992: the EUROCARE study. *European Journal of Cancer* 2001;37(6):687-94.
15. Gatta G, Corazziari I, Magnani C, Peris-Bonet R, Roazzi P, Stiller C, et al. Childhood cancer survival in Europe. *Annals of Oncology* 2003;14 Suppl 5:v119-27.
16. Schillinger JA, Grosclaude PC, Honjo S, Quinn MJ, Sloggett A, Coleman MP. Survival after acute lymphocytic leukaemia: effects of socioeconomic status and geographic region. *Arch Dis Child* 1999;80(4):311-317.
17. Hamblin TJ, Oscier DG. Chronic Lymphocytic Leukaemias. In: Whittaker JA, Holmes JA, editors. *Leukaemia and Related Disorders*. Oxford: Blackwell Science Ltd, 1998:105 - 135.
18. Phekoo KJ, Richards MA, Moller H, Schey SA, South Thames Haematology Specialist C. The incidence and outcome of myeloid malignancies in 2,112 adult patients in southeast England. *Haematologica* 2006;91(10):1400-4.
19. Jarvisalo J, Tola S, Korkala ML, Jarvinen E. A cancer register-based case study of occupations of patients with acute myeloid leukemia. *Cancer* 1984;54(4):785-90.
20. Gatta G, Luksch R, Coleman MP, Corazziari I, Group EW. Survival from acute non-lymphocytic leukaemia (ANLL) and chronic myeloid leukaemia (CML) in European children since 1978: a population-based study. *European Journal of Cancer* 2001;37(6):695-702.
21. Brandt L. Environmental factors and leukemia. *Medical Oncology and Tumour Pharmacotherapy* 1985;2(1):7-10.
22. Descatha A, Jenabian A, Conso F, Ameille J. Occupational exposures and haematological malignancies: Overview on human recent data. *Cancer Causes and Control* 2005;16(8):939-953.
23. Brandt L. Exposure to organic solvents and risk of haematological malignancies. *Leukaemia Research* 1992;16(1):67-70.
24. Brown LM, Gibson R, Burmeister LF, Schuman LM, Everett GD, Blair A. Alcohol consumption and risk of leukemia, non-Hodgkin's lymphoma, and multiple myeloma. *Leukemia Research* 1992;16(10):979-984.
25. Pogoda JM, Nichols PW, Preston-Martin S. Alcohol consumption and risk of adult-onset acute myeloid leukemia: results from a Los Angeles County case-control study. *Leukemia Research* 2004;28(9):927-931.
26. Rauscher GH, Shore D, Sandler DP. Alcohol intake and incidence of de novo adult acute leukemia. *Leukemia Research* 2004;28(12):1263-1265.
27. Gorini G, Stagnaro E, Fontana V, Miligi L, Ramazzotti V, Nanni O, et al. Alcohol consumption and risk of leukemia: A multicenter case-control study. *Leukemia Research* 2007;31(3):379-386.
28. Ringborg U. Alcohol and risk of cancer. *Alcoholism: Clinical & Experimental Research* 1998;22(7 Suppl):323S-328S.
29. Williams RR, Horm JW. Association of cancer sites with tobacco and alcohol consumption and socioeconomic status of patients: interview

study from the Third National Cancer Survey. *Journal of the National Cancer Institute* 1977;58(3):525-47.
30. Severson RK, Davis S, Heuser L, Daling JR, Thomas DB. Cigarette Smoking and Acute Nonlymphocytic Leukaemia. *Am. J. Epidemiol.* 1990;132(3):418-422.
31. Brownson RC, Novotny TE, Perry MC. Cigarette Smoking and Adult Leukemia: A Meta-analysis. *Arch Intern Med* 1993;153(4):469-475.
32. Wald N. Smoking and leukaemia. *BMJ* 1988;297(6649):638-.
33. Pogoda JM, Preston-Martin S, Nichols PW, Ross RK. Smoking and Risk of Acute Myeloid Leukemia: Results from a Los Angeles County Case-Control Study. *Am. J. Epidemiol.* 2002;155(6):546-553.
34. Hursting SD, Margolin BH, Switzer BR. Diet and human leukemia: an analysis of international data. *Preventive Medicine* 1993;22(3):409-22.
35. Roman E, Simpson J, Ansell P, Lightfoot T, Mitchell C, Eden TOB. Perinatal and reproductive factors: a report on haematological malignancies from the UKCCS. *European Journal of Cancer* 2005;41(5):749-59.
36. Ross JA, Xie Y, Davies SM, Shu X-o, Pendergrass TW, Robison LL. Prescription medication use during pregnancy and risk of infant leukemia (United States). *Cancer Causes & Control* 2003;14(5):447-51.
37. Pogoda JM, Katz J, McKean-Cowdin R, Nichols PW, Ross RK, Preston-Martin S. Prescription drug use and risk of acute myeloid leukemia by French-American-British subtype: results from a Los Angeles County case-control study. *International Journal of Cancer* 2005;114(4):634-8.
38. Kasum CM, Blair CK, Folsom AR, Ross JA. Non-steroidal anti-inflammatory drug use and risk of adult leukemia. *Cancer Epidemiology, Biomarkers & Prevention* 2003;12(6):534-7.
39. Cerhan J, Vachon C, Habermann T, Ansell S, Witzig T, Kurtin P, et al. Hormone Replacement Therapy and Risk of Non-Hodgkin Lymphoma and Chronic Lymphocytic Leukemia. *Cancer Epidemiol Biomarkers Prev* 2002;11(11):1466-1471.
40. Ross J, Sinner P, Blair C, Cerhan J, Folsom A. Hormone Replacement Therapy is not associated with an increased risk of Leukaemia (United States) *Cancer Causes and Control* 2005;16:483-488.
41. Bourke A, Dattani H, Robinson M. Feasibility study and methodology to create a quality-evaluated database of primary care data. *Informatics in Primary Care* 2004;12(3):171-7.
42. Netting an important database *The Lancet* 2001;357(9257):649.
43. Epic. THIN Data from EPIC: A Guide for Researchers. London: EPIC, 2007.
44. Jick H, Jick SS, Derby LE. Validation of information recorded on general practitioner based computerised data resource in the United Kingdom. *BMJ* 1991;302(6779):766-8.
45. Kaye JA, Derby LE, del Mar Melero-Montes M, Quinn M, Jick H. The incidence of breast cancer in the General Practice Research Database compared with national cancer registration data. *British Journal of Cancer* 2000;83(11):1556-8.

46. Meal A, Leonardi-Bee J, Smith C, Hubbard R, Bath-Hextall F. Validation of THIN data for non-melanoma skin cancer. *Quality in Primary Care* 2008;16(1):49-52.
47. Fombonne E, Heavey L, Smeeth L, Rodrigues LC, Cook C, Smith PG, et al. Validation of the diagnosis of autism in general practitioner records. *BMC Public Health* 2004;4:5.
48. Lawrenson R, Todd JC, Leydon GM, Williams TJ, Farmer RD. Validation of the diagnosis of venous thromboembolism in general practice database studies. *British Journal of Clinical Pharmacology* 2000;49(6):591-6.
49. Lewis JD, Brensinger C, Bilker WB, Strom BL. Validity and completeness of the General Practice Research Database for studies of inflammatory bowel disease. *Pharmacoepidemiology & Drug Safety* 2002;11(3):211-8.
50. Soriano JB, Maier WC, Visick G, Pride NB. Validation of general practitioner-diagnosed COPD in the UK General Practice Research Database. *European Journal of Epidemiology* 2001;17(12):1075-80.
51. Jick SS, Kaye JA, Vasilakis-Scaramozza C, Garcia Rodriguez LA, Ruigomez A, Meier CR, et al. Validity of the general practice research database. *Pharmacotherapy* 2003;23(5):686-9.
52. Lewis JD, Schinnar R, Bilker WB, Wang X, Strom BL. Validation studies of the health improvement network (THIN) database for pharmacoepidemiology research. *Pharmacoepidemiology & Drug Safety* 2007;16(4):393-401.
53. Gribbin J, Hubbard RB, Le Jeune I, Smith CJP, West J, Tata LJ. Incidence and mortality of idiopathic pulmonary fibrosis and sarcoidosis in the UK. *Thorax* 2006;61(11):980-5.
54. MHRA. The Medicines and Healthcare products Regulatory Agency London: Crown, 2010.
55. QResearch. QResearch. Nottingham: QRESEARCH, 2010.
56. NHS Information Centre. HESonline: Health and Social Care Information Centre, 2007.
57. Dixon J, Sanderson C, Elliott P, Walls P, Jones J, Petticrew M. Assessment of the reproducibility of clinical coding in routinely collected hospital activitydata: a study in two hospitals. *J Public Health* 1998;20(1):63-69.
58. Aylin P, Bottle A, Majeed A. Use of administrative data or clinical databases as predictors of risk of death in hospital: comparison of models. *BMJ* 2007;334(7602):1044-.
59. The National Archives. The National Digital Archive of Datasets. London: The National Archives, 1999.
60. Hall W, Ramachandran R, Narayan S, Jani A, Vijayakumar S. An electronic application for rapidly calculating Charlson comorbidity score. *BMC Cancer* 2004;4(1):94.
61. Hewitt P. Health Secretary announces new architecture of the local NHS. London: Department of Health, 2006:Press release.
62. Adamson P. Leukaemia. In: Quinn M, Wood H, Cooper N, Rowan R, editors. *Cancer Atlas of the United Kingdom and Ireland 1991-2000.* Basingstoke: Palgrave Macmillan, 2005.

63. Little J. *Epidemiology of Childhood Cancer*. Lyon: International Agency for Research on Cancer, 1999.
64. Greenberg RS, Shuster JLJ. Epidemiology of Cancer in Children. *Epidemiol Rev* 1985;7(1):22-48.
65. Poole C, Greenland S, Luetters C, Kelsey JL, Mezei G. Socioeconomic status and childhood leukaemia: a review. *Int. J. Epidemiol.* 2006;35(2):370-384.
66. Coleman MP, Rachet B, Woods LM, Mitry E, Riga M, Cooper N, et al. Trends and socioeconomic inequalities in cancer survival in England and Wales up to 2001. *British Journal of Cancer* 2004;90(7):1367-73.
67. Westergaard T, Andersen PK, Pedersen JB, Olsen JH, Frisch M, Sorensen HT, et al. Birth characteristics, sibling patterns, and acute leukemia risk in childhood: A population-based cohort study. *Journal of the National Cancer Institute* 1997;89(13).
68. Stiller CA, Eatock EM. Patterns of care and survival for children with acute lymphoblastic leukaemia diagnosed between 1980 and 1994. *Archives of Disease in Childhood* 1999;81(3):202-8.
69. Redaelli A, Laskin BL, Stephens JM, Botteman MF, Pashos CL. The clinical and epidemiological burden of chronic lymphocytic leukaemia. *European Journal of Cancer Care* 2004;13(3):279-87.
70. Molica S. Sex differences in incidence and outcome of chronic lymphocytic leukemia patients. *Leukemia & Lymphoma* 2006;47(8):1477-80.
71. Newland A. Progress in the Treatment of Acute Myeloid Leukaemia in Adults. *Int J Hematol.* 2002;76(Suppl 1):253-8.
72. Bhayat F, Das-Gupta E, Smith C, McKeever T, Hubbard R. The incidence of and mortality from leukaemias in the UK: a general population-based study. *BMC Cancer* 2009;9(1):252.
73. Griggs JJ, Culakova E, Sorbero MES, van Ryn M, Poniewierski MS, Wolff DA, et al. Effect of patient socioeconomic status and body mass index on the quality of breast cancer adjuvant chemotherapy.[see comment]. *Journal of Clinical Oncology* 2007;25(3):277-84.
74. Campbell NC, Elliott AM, Sharp L, Ritchie LD, Cassidy J, Little J. Impact of deprivation and rural residence on treatment of colorectal and lung cancer. *British Journal of Cancer* 2002;87(6):585-90.
75. McGory ML, Zingmond DS, Sekeris E, Bastani R, Ko CY. A patient's race/ethnicity does not explain the underuse of appropriate adjuvant therapy in colorectal cancer. *Diseases of the Colon & Rectum* 2006;49(3):319-29.
76. Bhayat F, Das-Gupta E, Smith C, McKeever TM, Hubbard R. The incidence of and mortality from leukaemias in the UK: a general population-based study. *BMC Cancer* 2009;9(1):252.
77. Grimwade D, Walker H, Oliver F, Wheatley K, Harrison C, Harrison G, et al. The Importance of Diagnostic Cytogenetics on Outcome in AML: Analysis of 1,612 Patients Entered Into the MRC AML 10 Trial. *Blood* 1998;92(7):2322-2333.
78. Muss HB. Factors used to select adjuvant therapy of breast cancer in the United States: an overview of age, race, and socioeconomic status. *Journal of the National Cancer Institute* 2001;Monographs.(30):52-5.

79. Munro AJ, Bentley AHM. Deprivation, comorbidity and survival in a cohort of patients with colorectal cancer. *European Journal of Cancer Care* 2004;13(3):254-62.
80. British Society of Blood and Marrow Transplantation. British Society of Blood and Marrow Transplantation. London: BSBMT, 2010.
81. Langman MJ, Cheng KK, Gilman EA, Lancashire RJ. Effect of anti-inflammatory drugs on overall risk of common cancer: case-control study in general practice research database. *BMJ* 2000;320(7250):1642-6.
82. Giovannucci E, Rimm EB, Stampfer MJ, Colditz GA, Ascherio A, Willett WC. Aspirin use and the risk for colorectal cancer and adenoma in male health professionals. *Annals of Internal Medicine* 1994;121(4):241-6.
83. Thun MJ, Namboodiri MM, Calle EE, Flanders WD, Heath CW, Jr. Aspirin use and risk of fatal cancer. *Cancer Research* 1993;53(6):1322-7.
84. Olsen JH, Friis S, Poulsen AH, Fryzek J, Harving H, Tjonneland A, et al. Use of NSAIDs, smoking and lung cancer risk. *British Journal of Cancer* 2008;98(1):232-7.
85. Khuder SA, Herial NA, Mutgi AB, Federman DJ. Nonsteroidal antiinflammatory drug use and lung cancer: a metaanalysis. *Chest* 2005;127(3):748-54.
86. Moysich KB, Menezes RJ, Ronsani A, Swede H, Reid ME, Cummings KM, et al. Regular aspirin use and lung cancer risk. *BMC Cancer* 2002;2:31.
87. Flossmann E, Rothwell PM. Effect of aspirin on long-term risk of colorectal cancer: consistent evidence from randomised and observational studies. *Lancet* 2007;369(9573):1603-13.
88. Baron JA, Cole BF, Sandler RS, Haile RW, Ahnen D, Bresalier R, et al. A randomized trial of aspirin to prevent colorectal adenomas. *New England Journal of Medicine* 2003;348(10):891-9.
89. Diaz LE, Montero A, Gonzalez-Gross M, Vallejo AI, Romeo J, Marcos A. Influence of alcohol consumption on immunological status: a review. *European Journal of Clinical Nutrition* 2002;56 Suppl 3:S50-3.
90. de Groot V, Beckerman H, Lankhorst GJ, Bouter LM. How to measure comorbidity: a critical review of available methods. *Journal of Clinical Epidemiology* 2003;56(3):221-229.
91. Chanarin I. Haemopoesis and Alcohol. *Br Med Bull* 1982;38(1):81-86.
92. Schrek R, Stefani SS. Effects of alcohol and hyperthermia on normal and leukemic lymphocytes. *Oncology* 1981;38(2):69-71.

APPENDICES

Appendix A. Publications Based On This Research

1. F. Bhayat, E. Das-Gupta, C. Smith, T McKeever, R. Hubbard. The incidence of and mortality from leukaemias in the UK: a general population-based study. ***BMC Cancer 2009, 9:252 DOI:10.1186/1471-2407-9-252***

2. F. Bhayat, E. Das-Gupta, C. Smith, R. Hubbard. NSAID use and risk of leukaemia: a population-based case-control study. ***Pharmacoepidemiology and drug safety* 2009; 18: 833–836 *DOI:10.1002/pds.1789***

3. F. Bhayat, E. Das-Gupta, R. Hubbard. Bone marrow transplantation in AML, and socio-economic class. ***Submitted, under review.***

Appendix B. Leukaemia READ Codes

READ CODE	DESCRIPTION
B64..00	Lymphoid leukaemia
B64..11	Lymphatic leukaemia
B640.00	Acute lymphoid leukaemia
B641.00	Chronic lymphoid leukaemia
B641.11	Chronic lymphatic leukaemia
B642.00	Subacute lymphoid leukaemia
B64y.00	Other lymphoid leukaemia
B64y000	Aleukaemic lymphoid leukaemia
B64y100	Prolymphocytic leukaemia
B64y200	Adult T-cell leukaemia
B64yz00	Other lymphoid leukaemia NOS
B64z.00	Lymphoid leukaemia NOS
B65..00	Myeloid leukaemia
B650.00	Acute myeloid leukaemia
B651.00	Chronic myeloid leukaemia
B651.11	Chronic granulocytic leukaemia
B651000	Chronic eosinophilic leukaemia
B651200	Chronic neutrophilic leukaemia
B651z00	Chronic myeloid leukaemia NOS
B652.00	Subacute myeloid leukaemia

B65y.00	Other myeloid leukaemia
B65y000	Aleukaemic myeloid leukaemia
B65y100	Acute promyelocytic leukaemia
B65yz00	Other myeloid leukaemia NOS
B65z.00	Myeloid leukaemia NOS
B66..00	Monocytic leukaemia
B66..11	Histiocytic leukaemia
B66..12	Monoblastic leukaemia
B660.00	Acute monocytic leukaemia
B661.00	Chronic monocytic leukaemia
B662.00	Subacute monocytic leukaemia
B66y.00	Other monocytic leukaemia
B66y000	Aleukaemic monocytic leukaemia
B66yz00	Other monocytic leukaemia NOS
B66z.00	Monocytic leukaemia NOS
B67..00	Other specified leukaemia
B670.00	Acute erythraemia and erythroleukaemia
B671.00	Chronic erythraemia
B672.00	Megakaryocytic leukaemia
B672.11	Thrombocytic leukaemia
B673.00	Mast cell leukaemia
B67y.00	Other and unspecified leukaemia
B67y000	Lymphosarcoma cell leukaemia

B67yz00	Other and unspecified leukaemia NOS
B67z.00	Other specified leukaemia NOS
B68..00	Leukaemia of unspecified cell type
B680.00	Acute leukaemia NOS
B681.00	Chronic leukaemia NOS
B682.00	Subacute leukaemia NOS
B68y.00	Other leukaemia of unspecified cell type
B68z.00	Leukaemia NOS
B69..00	Myelomonocytic leukaemia
B690.00	Acute myelomonocytic leukaemia
B691.00	Chronic myelomonocytic leukaemia
B692.00	Subacute myelomonocytic leukaemia
BBr..00	[M]Leukaemias
BBr0.00	[M]Leukaemias unspecified
BBr0000	[M]Leukaemia NOS
BBr0100	[M]Acute leukaemia NOS
BBr0111	[M]Blast cell leukaemia
BBr0112	[M]Blastic leukaemia
BBr0113	[M]Stem cell leukaemia
BBr0200	[M]Subacute leukaemia NOS
BBr0300	[M]Chronic leukaemia NOS
BBr0400	[M]Aleukaemic leukaemia NOS
BBr0z00	[M]Leukaemia unspecified, NOS

BBr1.00	[M]Compound leukaemias
BBr1000	[M]Compound leukaemia
BBr1011	[M]Mixed leukaemia
BBr1z00	[M]Compound leukaemia NOS
BBr2.00	[M]Lymphoid leukaemias
BBr2000	[M]Lymphoid leukaemia NOS
BBr2011	[M]Lymphatic leukaemia
BBr2100	[M]Acute lymphoid leukaemia
BBr2200	[M]Subacute lymphoid leukaemia
BBr2300	[M]Chronic lymphoid leukaemia
BBr2400	[M]Aleukaemic lymphoid leukaemia
BBr2500	[M]Prolymphocytic leukaemia
BBr2600	[M]Burkitt's cell leukaemia
BBr2700	[M]Adult T-cell leukaemia/lymphoma
BBr2z00	[M]Other lymphoid leukaemia NOS
BBr3.00	[M]Plasma cell leukaemias
BBr3000	[M]Plasma cell leukaemia
BBr3z00	[M]Plasma cell leukaemia NOS
BBr4.00	[M]Erythroleukaemias
BBr4000	[M]Erythroleukaemia
BBr4200	[M]Chronic erythraemia
BBr4z00	[M]Erythroleukaemia NOS
BBr5.00	[M]Lymphosarcoma cell leukaemias

BBr5000	[M]Lymphosarcoma cell leukaemia
BBr5z00	[M]Lymphosarcoma cell leukaemia NOS
BBr6.00	[M]Myeloid leukaemias
BBr6000	[M]Myeloid leukaemia NOS
BBr6011	[M]Granulocytic leukaemia NOS
BBr6100	[M]Acute myeloid leukaemia
BBr6200	[M]Subacute myeloid leukaemia
BBr6300	[M]Chronic myeloid leukaemia
BBr6311	[M]Naegeli-type monocytic leukaemia
BBr6400	[M]Aleukaemic myeloid leukaemia
BBr6500	[M]Neutrophilic leukaemia
BBr6600	[M]Acute promyelocytic leukaemia
BBr6700	[M]Acute myelomonocytic leukaemia
BBr6800	[M]Chronic myelomonocytic leukaemia
BBr6z00	[M]Other myeloid leukaemia NOS
BBr7.00	[M]Basophilic leukaemias
BBr7000	[M]Basophilic leukaemia
BBr7z00	[M]Basophilic leukaemia NOS
BBr8.00	[M]Eosinophilic leukaemias
BBr8000	[M]Eosinophilic leukaemia
BBr8z00	[M]Eosinophilic leukaemia NOS
BBr9.00	[M]Monocytic leukaemias
BBr9000	[M]Monocytic leukaemia NOS

BBr9011	[M]Histiocytic leukaemia
BBr9012	[M]Schilling-type monocytic leukaemia
BBr9100	[M]Acute monocytic leukaemia
BBr9200	[M]Subacute monocytic leukaemia
BBr9300	[M]Chronic monocytic leukaemia
BBr9400	[M]Aleukaemic monocytic leukaemia
BBr9z00	[M]Other monocytic leukaemia NOS
BBrA.00	[M]Miscellaneous leukaemias
BBrA000	[M]Mast cell leukaemia
BBrA100	[M]Megakaryocytic leukaemia
BBrA111	[M]Thrombocytic leukaemia
BBrA400	[M]Hairy cell leukaemia
BBrA500	[M]Acute megakaryoblastic leukaemia
BBrAz00	[M]Miscellaneous leukaemia NOS
BBrz.00	[M]Leukaemia NOS
ByuD500	[X]Other lymphoid leukaemia
ByuD600	[X]Other myeloid leukaemia
ByuD700	[X]Other monocytic leukaemia
ByuD800	[X]Other specified leukaemias
ByuD900	[X]Other leukaemia of unspecified cell type
ZV10600	[V]Personal history of leukaemia
ZV10611	[V]Personal history of lymphoid leukaemia

Appendix C: UK Bone Marrow Transplant Centres

344	ABERDEEN	Royal Infirmary, Aberdeen	Dr Dominic Culligan	Auto and Allograft	Adult
736	BANGOR	Ysbyty Gwynned Hospital, Bangor	Dr David Edwards	Auto only	Adult
619	BATH	Royal United Hospital, Bath	Dr Christopher Knechtli	Auto only	Adult
268	BELFAST	Belfast City Hospital, Belfast	Dr Mary Frances McMullin	Auto and Allograft	Adult
781	BIRMINGHAM	Children's Hospital, Birmingham	Dr Sarah Lawson	Auto and Allograft	Paediatric
405	BIRMINGHAM	Dudley Hospital, Birmingham	Dr Savio.Fernandes	Auto only	Adult
284	BIRMINGHAM	Heartlands Hospital, Birmingham	Dr Don Milligan	Auto and Allograft	Adult
387	BIRMINGHAM	Queen Elizabeth Hospital, Birmingham	Prof. Charles Craddock	Auto and Allograft	Adult
832	BLACKPOOL	Blackpool Victoria Hospital, Blackpool	Dr M Macheta	Auto only	Adult
765	BOURNEMOUTH	Royal Bournemouth Hospital, Bournemouth	Dr Sally Killick	Auto only	Adult
386	BRISTOL	Avon Haematology Unit, Bristol	Dr Jenny Bird	Auto and Allograft	Adult
386	BRISTOL	Bristol Children's Hospital, Bristol	Prof. David Marks	Auto and Allograft	Adult
386	BRISTOL	Bristol	Dr Jacqueline	Auto	Paediatric

		Children's Hospital, Bristol	Cornish OBE	and Allograft	
566	CAMBRIDGE	Addenbrookes Hospital, Cambridge	Dr Charles Crawley	Auto and Allograft	Both
303	CARDIFF	University Hospital of Wales, Cardiff	Dr Keith Wilson	Auto and Allograft	Adult
398	CHELTENHAM	Cheltenham General Hospital, Cheltenham	Dr Eve Blundell	Auto only	Adult
322	COVENTRY	University Hospital, Coventry	Dr Waqas Bokhari	Auto only	Adult
774	DUBLIN	Our Lady's Hospital, Dublin. Rep. Ireland	Dr Anne O'Meara	Auto and Allograft	Paediatric
719	DUNDEE	Ninewells Hospital, Dundee	Dr David Meiklejohn	Auto only	Adult
228	EDINBURGH	Western General Hospital, Edinburgh	Dr Peter Johnson	Auto and Allograft	Adult
571	EXETER	Royal Devon and Exeter Hospital, Exeter	Dr Claudius Rudin	Auto only	Adult
408	GALWAY	University College Hospital Galway, Galway, Rep Ireland	Dr Patrick Hayden	Auto only	Adult
244	GLASGOW	Royal Infirmary, Glasgow	Dr Grant McQuaker	Auto and Allograft	Adult
707	GLASGOW	Yorkhill Children's Hospital, Glasgow	Dr Brenda Gibson	Auto and Allograft	Paediatric

128	IPSWICH	Ipswich Hospital, Ipswich	Dr Nick Dodds	Auto only	Adult
254	LEEDS	St James' University Hospital, Leeds	Dr Maria Gilleece	Auto and Allograft	Both
713	LEICESTER	Leicester Royal Infirmary, Leicester	Dr Ann Hunter	Auto and Allograft	Adult
773	LIVERPOOL	Alder Hey Children's, Liverpool	Dr Mark Caswell	Auto and Allograft	Paediatric
501	LIVERPOOL	Royal Hospital, Liverpool	Dr Richard Clark	Auto and Allograft	Adult
243	LONDON	Great Ormond Street Children's Hospital, London	Dr Paul Veys	Auto and Allograft	Paediatric
721	LONDON	Guy's Hospital, London	Dr Majid Kamzi	Auto and Allograft	Adult
205	LONDON	Hammersmith Hospital, London	Prof. Jane Apperley	Auto and Allograft	Adult
763	LONDON	King's College Hospital, London	Dr Tony Pagliuca	Auto and Allograft	Adult
216	LONDON	Royal Free Hospital, London	Prof. Stephen Mackinnon	Auto and Allograft	Both
218	LONDON	Royal Marsden, London	Dr Mike Potter	Auto and Allograft	Both
768	LONDON	St Bartholomew's Hospital London	Prof. John Gribben	Auto and Allograft	Adult
539	LONDON	St Georges Hospital, London	Dr Mickey Koh	Auto and Allograft	Both
866	LONDON	St Mary's Hospital.	Dr Josu de la Fuente	Auto and	Paediatric

		London		Allograft	
263	LONDON	The London Clinic, Harley Street, London	Dr Mike Potter	Auto only	Adult
450	LONDON	Parkside Hospital, London	Prof Ray Powels	Auto only	Adult
224	LONDON	University College Hospital, London	Dr Kirsty Thomson	Auto and Allograft	Both
780	MANCHESTER	Christie Hospital, Manchester	Dr Adrian Bloor	Auto and Allograft	Adult
521	MANCHESTER	Manchester Children's Hosp, Manchester	Dr Robert Wynn	Auto and Allograft	Paediatric
601	MANCHESTER	Manchester Royal Infirmary, Manchester	Prof. John Yin	Auto and Allograft	Adult
276	NEWCASTLE	Royal Victoria Infirmary, Newcastle	Dr Graham Jackson	Auto and Allograft	Adult
276	NEWCASTLE	Royal Victoria Infirmary, Newcastle	Dr Roderick Skinner	Auto and Allograft	Paediatric
391	NORWICH	Norfolk and Norwich Hospital, Norfolk	Dr Matthew Lawes	Auto only	Adult
717	NOTTINGHAM	Nottingham City Hospital, Nottingham	Prof Nigel Russell	Auto and Allograft	Adult
255	OXFORD	John Radcliffe Hospital, Oxford	Dr Tim Littlewood	Auto and Allograft	Adult
823	PLYMOUTH	Derriford Hosp, Plymouth	Dr Simon Rule	Auto and Allograft	Adult
765	POOLE	Dorset Cancer Centre, Poole	Dr Andrew Bell	Auto only	Adult
757	SALISBURY	District General	Dr Jonathan	Auto	Adult

		Hospital, Salisbury	Cullis	only	
778	SHEFFIELD	Children's Hospital, Sheffield	Dr Ajay Vora	Auto and Allograft	Paediatric
778	SHEFFIELD	Sheffield Teaching Hospitals NHS Foundation Trust	Dr John Snowden	Auto and Allograft	Adult
704	SOUTHAMPTON	Southampton General Hospital, Southampton	Dr Kim Orchard	Auto and Allograft	Both
394	STOKE	North Staffordshire Hospital, Staffs	Dr Richard Chasty	Auto only	Adult
554	SWANSEA	Singleton Hospital, Swansea	Dr Saad Al-Ismail	Auto only	Adult
608	SWINDON	Great Western Hospital, Swindon	Dr Norbert Blesing	Auto only	Adult
708	TAUNTON	Taunton and Somerset Hospital, Taunton	Dr Simon Bolam	Auto only	Adult

http://www.bsbmt.org/pages/42-Transplant_Centre_List