THE EPIDEMIOLOGY OF HODGKIN'S DISEASE WITH SPECIAL REFERENCE TO EBV STATUS

DAVID JAMES LAWRENCE

Degree of Doctor of Philosophy

The University of Edinburgh

1999

Originality Statement

I declare that the contents of this thesis represent my own work; contributions from others have always been explicitly acknowledged.

Signature:

Date: 23/3/2000

CONTENTS

ACKNOWLEDGEMENT

I am very grateful to Dr. Freda Alexander, my principal supervisor, for her dedicated supervision, helpful suggestions and the encouragement she gave me, which were crucial to the completion of this thesis. I am also grateful to Dr. Harry Campbell, my second supervisor, for his help with style and thesis writing.

I would like to thank the team of research interviewers, under the supervision of Katharine Proudfoot, who collected the data for the SNEHD study. I would also like to thank Morag Leitch in the Department of Public Health Sciences for her help with SNEHD and more general administration and logistics.

Professor Ruth Jarrett and Dr Malcolm Taylor played central roles in the YHHCCS and SNEHD studies, being responsible for the histological review and EBV histopathology and the HLA typing respectively, for both studies.

I gratefully acknowledge the scholarship granted to me by the University of Edinburgh Medical School which enabled me to pursue this PhD.

LIST OF TABLES

LIST OF FIGURES

Glossary of Abbreviations and Definitions

95% CI	95% Confidence Interval
AIDS	Acquired Immuno-Deficiency Syndrome
Amp	Amplitude
AOS	All Other Subtypes
AT	Ataxia Telangiectasia
BMI	Body Mass Index
C/C study	Case/Control study
CHD	Coronary Heart Disease
CNS	Central Nervous System
Cos	Cosine
CSA	Common Services Agency
Dev	Deviance
DF	Degrees of Freedom
DNA	Deoxyribonucleic Acid
E	Expected value
EA	Early Antigen
EBER	Epstein-Barr virus encoded early RNA
EBNA	Epstein-Barr virus Nuclear Antigens
EBV	Epstein-Barr virus
EBV +ve	Evidence of EBV found in the HRS cells by EBER ISH or LMP-1 immunohistochemistry
EBV -ve	No evidence of EBV found in the HRS cells by EBER ISH or LMP-1 immunohistochemistry
EE	Energy Expenditure
FHS	Family History Score

FHSA	Family Health Services Authority
GNP	Gross National Product
GP	General Practitioner
HD	Hodgkin's Disease
HDL	High Density Lipoprotein
HIV	Human Immuno-deficiency Virus
HLA	Human Leukocyte Antigen
HR	Heart Rate
HRS	Hodgkin's Reed-Sternberg
HSV	Herpes Simplex virus
IDDM	Insulin Dependent Diabetes
IM	Infectious Mononucleosis
ISH	In situ hybridisation
Kcal	Kilocalories
LD	Lymphocyte Depleted
LMP	Latent Membrane Protein
LP	Lymphocyte Predominant
LRF	Leukaemia Research Fund
LRTI	Lower Respiratory Tract Infection
LTPA	Leisure-Time Physical Activity
MC	Mixed Cellularity
MET	Metabolic Equivalent
MRFIT	Multiple Risk Factor Intervention Trial
MS	Multiple Sclerosis
NA	Not Available
NCI	National Cancer Institute

NHANES	National Health And Nutrition Examination Survey
NHL	Non-Hodgkin's Lymphoma
NJ	New Jersey
NK	Natural Killer
NNA	Nearest Neighbour Analysis
NS	Nodular Sclerosing
NSW	New South Wales
NY	New York
NYHIP	New York Health Insurance Plan
O	Observed value
OPCS	Office Population, Censuses and Surveys
OR	Odds Ratio
PA	Pernicious Anaemia
PCR	Polymerase Chain Reaction
PMR	Proportional Mortality Ratio
PY	Person Years
RA	Rheumatoid Arthritis
RDD	Random Digit Dialling
REAL	Revised European American Lymphoma classification
RI	Relative Incidence
RNA	Ribonucleic Acid
RR	Relative Risk
SE	Standard Error
SEER	Surveillance, Epidemiology and End Results
SES	Socio-Economic Status
SF	San Francisco

Sin	Sine
SIR	Standardised Incidence Ratio
SMR	Standardised Mortality Ratio
SNEHD	Scotland and Newcastle Epidemiological study of Hodgkin's Disease
UK	United Kingdom
URTI	Upper Respiratory Tract Infection
USA	United States of America
VCA	Viral Capsid Antigen
$VO_{2\,max}$	Maximal Oxygen uptake
WAS	Wiskott-Aldrich Syndrome
WASR	World Age Standardised Rate
YHHCCS	Young adult Hodgkin's disease Case-Control Study

Abstract

Background:
Hodgkin's Disease (HD) is quite a rare cancer but is the most common malignancy in males age 16-34 years and in the top five for females of the same age in Scotland. Molecular evidence of the Epstein-Barr virus (EBV) has been found in 35% of HD cases (EBV +ve) in developed countries but the proportion is smallest for the 16-34 years age group. There is a variety of evidence to suggest that HD has an infectious aetiology, especially in the 16-34 years age group. As well as evidence suggesting an infectious aetiology other risk factors for HD have been reported including general and family health characteristics, immune and hormonal system variables.

Aim:
The aim of this study is to examine further the risk factors for HD both in total and in sub-groups by age and EBV status. The focus will be on whether the risk factors for EBV-status and/or age-at-diagnosis subgroups differ. Risk factors related to an infectious aetiology, family health and physical activity will be considered.

Methods:
1) Analysis of the seasonal presentation of HD using data from 2093 cases (age 0-34 years) from four Scandinavian cancer registries (Denmark, Finland, Norway, Sweden). Ecological analysis.
2) Analysis of infectious illness, general health, and family health risk factors for HD using data from a young adult (age 16-24 years) matched case-control study (YHHCCS) with EBV status known. Observational study.
3) Conduct of a large (493 cases, 512 controls) case-control study (SNEHD) with EBV status known. Role in writing, piloting, interviewing and analysing this study.
4) Analysis of infectious illness, general health, and family health risk factors for HD using data from a large (493 cases, 512 controls) case-control study (SNEHD) with EBV status known. Observational study.
5) Analysis of physical activity risk factors for HD using data from SNEHD. Observational study.

Results:
1) The results of the analysis of the Scandinavian cancer registry data show a statistically significant seasonal presentation of HD with a peak in January.
2) Greater numbers of childhood infectious illnesses were associated with a lower risk of HD in the total series and EBV +ve and EBV –ve HD in the YHHCCS study. Childhood infectious illnesses when aged ≥5 years were associated with an increased risk of EBV +ve HD and a decreased risk of EBV –ve HD. Infectious mononucleosis (IM) was associated with a significantly increased risk of HD in the total series and this effect was limited to EBV +ve HD. The difference between EBV subgroups was statistically significant for both of these variables.
3) SNEHD results show a statistically significant positive association of tonsillectomy with HD diagnosed age 16-34 years but not at older ages. HD (OR 5.17, 95% CI 1.12-23.79) or any haematological malignancy (OR 2.39, 95% CI 1.07-5.38) in a first-degree relative was associated with a significantly increased risk of HD. There were no significant differences between EBV-status subgroups for general or family health variables. Childhood infections aged ≥5 years were associated with an increased risk of EBV +ve HD and a decreased risk of EBV –ve HD. The difference between EBV-status subgroups was statistically significant (p=0.043). Infectious mononucleosis (IM) was associated

with a significantly increased risk of HD limited to the 16-34 years age group. The largest effect of IM was for EBV +ve HD presenting age 16-34 years.

4) The results of physical activity analysis show very little evidence of an effect of short-term activity (i.e. persisting for 3-10 years) on risk of HD. Consistent inactivity was associated with an increased risk of HD in the total series and all subgroups. Results did not differ significantly by EBV status.

Discussion:

The data presented here both confirm the literature and provide new results. Overall they support an infectious aetiology of HD, especially for the 16-34 years age group. However, there is no evidence of any specific infectious illness being involved in the aetiology of HD, apart from IM. The role of IM in risk of HD has been confirmed and , for the first time, evidence is provided that this relationship is strongest for EBV +ve HD in the 16-34 years age group. Support is given for tonsillectomy increasing risk of HD in the young adult (age 16-34 years) peak. The finding that risk of HD is increased with consistent inactivity is also novel and will require confirmation. In both the YHHCCS and SNEHD studies there was evidence of an association between the EBV status of the cases and childhood risk factors (childhood illness, IM). In future the epidemiological profile of HD by EBV status must be distinguished as a matter of course.

1. The epidemiology of Hodgkin's Disease.

1.1 Introduction:

Malignant lymphomas are divided into two groups: Hodgkin's Disease (HD) and Non-Hodgkin's lymphoma (NHL) (30% and 70% of lymphomas respectively). HD was first reported by Thomas Hodgkin in 1832 when he described a process originating in lymphatic tissues (Hodgkin, 1832). Hodgkin was the first person to contribute a constellation of pathologic and physiologic findings to a specific malignancy of the lymph nodes rather than an infectious disease or inflammation. Sternberg first described the cell that was thought to be pathognomic for the disease in 1898 and Reed, in 1902, further characterised this cell.

1.2 Definition and general background:

HD is a lymphoproliferative disorder characterised by the presence of large cells in the tumour tissue. These cells are called Hodgkin's cells if they are mono-nucleated and Reed-Sternberg cells in the case of multi-nuclearity (HRS cells). It is estimated that less than 1% of the total lymphoid cells in diseased tissues are HRS cells (Drexler, 1992). The lack of tumour cells means that the cellular origin of HRS cells remains controversial and has been allocated to almost every lympho-haemopoeitic cell type in the body (Gruss et al, 1997). HRS cells are not unique to HD and, therefore, to make the diagnosis HRS cells must be present in an appropriate cellular background (Jarrett, 1992).

The disease is characterised by progressive painless enlargement of the lymphoid tissues throughout the body. General features may include progressive weight loss, drenching night sweats and fever (Takvorian et al, 1993).

Traditionally HD has been subdivided by clinical staging (not discussed here, see Carbone et al, 1971) and histology.

Histological Classification:

The first histological classification to gain wide acceptance was that of Jackson & Parker (1944) which divided HD into three sub-groups: paragranuloma, granuloma, and sarcomatous disease. The categories of paragranuloma and sarcoma identified the most and least favourable prognostic groups, respectively, but 80% of cases remained in the indistinct middle category, granuloma. In the 1960s Lukes identified six subtypes of HD and viewed the various histologic sub-groups as indicative of differing host responses to a single disease process (Lukes & Butler, 1966; Lukes et al, 1966a). Almost immediately, this schema was simplified and a nodular sclerosis category added, becoming the Rye classification (Lukes et al, 1966b).

Based on the cellular composition and the histological appearance of the tumour the Rye classification groups HD into four types: nodular sclerosing (NS); mixed cellularity (MC); lymphocyte predominant (LP); and lymphocyte depleted (LD).

The vast majority of HD cases are NS and MC. NS is the most common subtype, accounting for 40-60% of cases and is becoming more common as the rate for this subtype increases (Medeiros & Greiner, 1995). NS occurs more frequently in females than the other subtypes and is rare over the age of 50. In general survival is good (84% at 5 years (Williams, 1990)). Approximately 30% of HD cases are MC. Patients generally present with advanced stage disease. This subtype can occur at any age but is proportionally more frequent in older people. 5 year survival is around 65% (Williams, 1990). The natural history and treatment of NS and MC are similar and for all intent and purposes are managed as the same disease.

Historically, LD accounted for 5-10% of cases but with current immuno-phenotyping this subtype is increasingly rare as NHLs, which may previously have been wrongly included as HD, are removed. LD is proportionally more frequent in older, male patients and is considered the most aggressive subtype. Patients generally present with constitutional

symptoms and widespread disseminated disease. 5 year survival is poor (32% survival at 5 years (Williams, 1990)). LP accounts for 5-10% of cases and is proportionally more common in males under 35. It generally presents with Stage I disease. There are two forms of LP: nodular and diffuse. Patients with LP rarely have systemic symptoms and the prognosis is usually favourable. (98% survival at 5 years (Williams, 1990)).

Problems with the classification of NHL led to the development of a new lymphoma classification system that removed the historical anachronism that separated HD from the other lymphomas. The Revised European American Lymphoma Classification (REAL) developed by the Lymphoma Group is a list of well-defined "real" lymphomas (Harris et al, 1994). Cases that do not fit into one of the defined categories are left unclassified, reflecting the fact that not everything is understood about them. The REAL classification takes account of evidence that suggests that LP differs morphologically, immunophenotypically, and clinically from classic HD (Table 1.1) and should, therefore, be treated as a separate entity. The REAL classification keeps the old categories of NS, MC, and LD (Classical HD). Nodular LP is no longer considered to be an HD subtype, while diffuse LP is classified within a new subtype (Lymphocyte rich Classical HD).

Other authors have suggested that HD could be classified into aetio-subgroups by age (MacMahon, 1966) or geography (Corea & O'Conor, 1971).

Classification of HD into aetio-subgroups:

Correa & O'Conor (1971) suggested that there are at least three different geographical epidemiological patterns of HD (see section, the geography of HD, p.10). MacMahon (1966) proposed three distinct forms of HD, with differing aetiologies, by age: a childhood form (0-14), a young adult form (15-34), and an older adult form (≥55 years).

Table 1.1: The REAL classification of HD.

	Classical HD (NS, MC & LD)	LP HD
Atypical cells	Diagnostic HRS cells, mono-nuclear or lacunar cells	LP variants ("L&H" (lymphocytic &/or histiocytic) or "popcorn" cells)
Diagnostic cells	always present	rare to absent
CD 15	usually +ve	-ve
CD 30	usually +ve	often +ve
CD 20	usually –ve	usually +ve
CD 45	usually –ve	+ve
EBV (in large cells)	often +ve (20-70%)	usually -ve

(Harris et al, 1994)

Search Methods:

MEDLINE and BIDS were searched back to 1990 to find English language studies that investigated the aetiology of HD using the following key words: Hodgkin's disease, aetiology, geography, socio-economic status (SES), birth order, sibship, education, infection, seasonal, EBV, LMP-1, EBNA, cluster, infectious mononucleosis, immune, immunodeficiency, HIV, AIDS, transplantation, heart, kidney, liver, tonsils, tonsillectomy, appendix, appendectomy, occupation, familial, parity. Additional references cited in primary material were also investigated, if pertinent, to get as complete a picture as possible of individual risk factors. Factors associated with risk of HD will be discussed following a section on the descriptive epidemiology of the disease.

1.3 Descriptive Epidemiology:

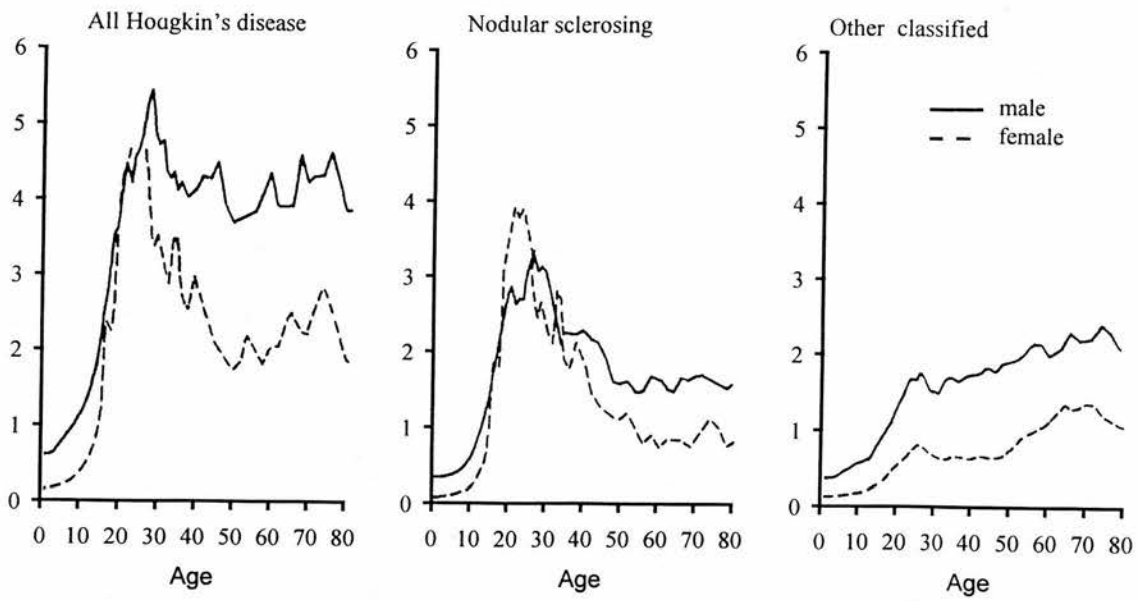
Incidence:

The annual age-adjusted (to the WASR) incidence rate of HD was 2.9 per 10^5 in Scottish males and 1.6 per 10^5 in Scottish females in 1996 (ISD, 1999).

Generally HD occurs in childhood and in older adults in developing countries but in developed countries there is a marked incidence peak in young adults (15-34 years) after which incidence remains fairly stable (Figure 1.1). This peak means that HD is one of the commonest malignancies in young adults, as data from Scotland make clear (Table 1.2). There is another peak in older adult years (≥50 years) in some regions and time periods. Incidence varies greatly with the highest rates in northern Italy and the USA and the lowest in Asia (Coleman et al, 1993) (Table 1.3).

In Europe HD incidence has been either stable or declining since 1970. Within this general decline, most countries had an increase in incidence in young adults. In each 5 year

Figure 1.1: Age-incidence curves of Hodgkin's disease by sex and subtype.



(reproduced with permission from Alexander et al, 1995a)

Table 1.2: Incidence and relative frequencies for the five most frequently diagnosed cancers in Scotland, by sex and age group, total 1986-95.

Males			**Females**		
Under 15 years	Incidence	Relative Frequency	**Under 15 years**	Incidence	Relative Freq.
Leukaemia	240	34.2%	Leukaemia	187	33.0%
Brain & CNS	144	20.5%	Brain & CNS	149	26.3%
Bone & Connective tissue	66	9.4%	Bone & Connective tissue	44	7.8%
NHL	66	9.4%	Kidney	32	5.7%
Kidney	42	6.0%	*Hodgkin's disease*	*27*	*4.8%*
Age 15-34 years	Incidence	Relative Frequency	**Age 15-34 years**	Incidence	Relative Freq.
Testis	852	32.1%	Cervix	701	20.8%
Hodgkin's disease	*301*	*11.3%*	Breast	661	19.6%
Malignant melanoma	256	9.6%	Malignant melanoma	520	15.5%
NHL	234	8.8%	*Hodgkin's disease*	*242*	*7.2%*
Brain & CNS	197	7.4%	Ovary	173	5.1%
Age 35-64 years	Incidence	Relative Frequency	**Age 35-64 years**	Incidence	Relative Freq.
Lung	9285	26.0%	Breast	15826	36.1%
Large Bowel	4507	12.6%	Lung	5104	11.7%
Bladder	2669	7.5%	Large Bowel	3902	8.9%
Stomach	1965	5.5%	Ovary	2481	5.7%
Prostate	1907	5.3%	Cervix	2309	5.3%

(Harris et al, 1998)

Table 1.3: Highest and lowest rates of HD in the world (per 10^5/yr) directly age standardised to the world standard population.

Sex	Highest rates (SE)		Lowest rates (SE)	
Male	Italy (Parma)	4.5 (0.7)	China (Qidong)	0.1 (0.1)
	USA (Connecticut) (white)	4.4 (0.2)	Japan (Yamagata)	0.1 (0.0)
Female	USA (Hawaii) (Chinese)	5.0 (2.6)	The Gambia	0.1 (0.1)
	Scotland (North)	4.4 (0.9)	Japan (Hiroshima)	0.1 (0.1)

(Crosignani et al, 1996)

period 1973-87 there was up to a 25% increase in male incidence at ages 15-44 in: Spain; France; UK; Scotland; Hungary; Yugoslavia; Israel; and Finland (Coleman et al, 1993).The same pattern is seen in female data from European registries e.g. France, Spain, Scotland but the increases were not as large.

Over three time periods (1973-77, 1978-82, 1983-87) the Surveillance, Epidemiology and End Results (SEER) programme of the National Cancer Institute (NCI) collected 9418 confirmed cases of HD in designated geographical regions considered representative of the USA as a whole. As in the European data the incidence of HD in the young adult group had increased since 1973 (Medeiros & Greiner, 1995). Chen et al (1997a) reported time trends of HD incidence 1935-92 in Connecticut alone and they also found that incidence of HD had increased dramatically in young adults (aged 20-44 years). The increase in young adults here and in the overall SEER data has been put down to a marked increase in the NS subtype which predominates in this age group (Figure 1.1).

1973-77 to 1983-87 the age-adjusted incidence rate (details of standard population used not given) of NS increased from 1.1 to 1.6 per 10^5. By 1983-87 NS accounted for 57.7% of all HD cases. The proportion of MC (23.4% of HD) and LP (6.0% of HD) remained stable, while LD (3.8% of HD) fell as more cases were classified as NHL (Medeiros & Greiner, 1995). Glaser & Swartz (1990) have observed this increase in the incidence of the NS subtype in other data from the USA 1969-1980 and Chen et al (1997a) in data from the Connecticut cancer registry 1935-92.

Mortality:

The annual age-adjusted (to the WASR) mortality rate of HD was 0.6 per 10^5 in Scottish males and 0.3 per 10^5 in Scottish females in 1996 (ISD, 1999).

La Vecchia et al (1991) found that from the 1960s onwards HD mortality has fallen in western European countries in all age groups, from 20-30% (Spain, Portugal, Greece) to

70% (Nordic countries). This decrease in mortality has been due to the success of therapy. The stable incidence rates and falling mortality have resulted in an increase in the incidence: mortality ratio e.g. for Denmark, Finland, Norway and Scotland combined the ratio was 1.35 in 1960 and 2.38 in 1980 (La Vecchia et al, 1991) (Figure 1.2).

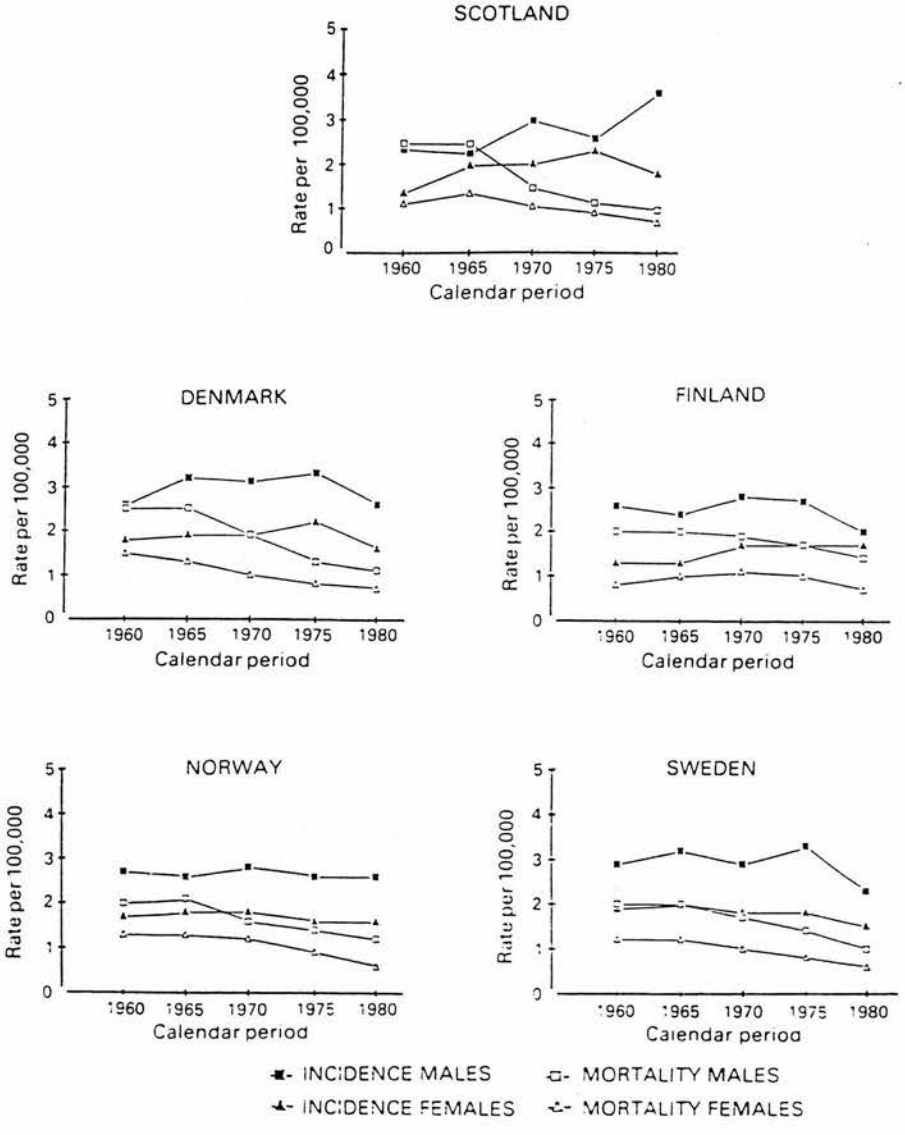
This general pattern of stable or falling incidence (except in young adults) and falling mortality is seen in most areas of the world including Asia, Oceania and North America.

The geography of HD:

There are geographical differences in incidence of HD and a variation in the distribution of histologic subtypes and age at presentation. Correa & O'Conor (1971) described three main types of male HD. Type I prevails in developing countries and is characterised by relatively high incidence and mortality rates in children, low incidence in the third decade of life, and another incidence peak in the older age groups. A great majority of cases are MC or LD. Type III prevails in developed countries and is characterised by very low rates in children and a pronounced initial peak in young adults. NS is the prevalent subtype. An intermediate pattern, Type II, is found in areas that lie between developed and developing e.g. central Europe. Additionally in some Asian countries e.g. Japan, Singapore, India there may be a Type IV pattern characterised by a relative paucity of cases in all age groups.

Correa & O'Conor (1971) observed a strong inverse relationship ($r = -0.9$) within individual selected countries between the incidence of HD in children (age 5-14 years) and the incidence of the disease in young adults (age 20-34 yrs). Where the childhood incidence of HD was high the incidence in young adults was reciprocally low (type III). However, MacFarlane et al (1995) using data from registries reporting in Parkin et al (1992) split into levels of economic development found little evidence of this inverse relationship in males

Figure 1.2: Comparison in trends in mortality and incidence rates in five European countries, circa 1960 to 1980.



(La Vecchia et al, 1991)

1983-87 (a similar finding to Cozen et al, 1992). HD rates are fairly low and it is possible that large variations in rates, especially among smaller registries may obscure this relationship. Therefore, MacFarlane et al (1995) restricted the analysis to the larger registries (those with at least $5x10^6$ person-years of observation) on a subset of the data for 1963-67. The original observation of a strong inverse relationship between the incidence of HD in children and the incidence of the disease in young adults in 1971 was confirmed, but in data from later periods this strong inverse relationship no longer existed (MacFarlane et al, 1995).

The bimodality of HD disease rates first described by MacMahon (1957) while still evident in the incidence rates from some registries is not always present. In the US the two incidence peaks have become less marked, 1973-80 to 1983-87 (Glaser, 1987). This is due primarily to increasing incidence in the age group 40-59 years and a fall in rates in older age groups. This flattening of the bimodal distribution has also been seen in the UK (McKinney et al, 1989).

Some Latin American countries have rates that are consistent with the type I pattern e.g. Brazil, Puerto Rico. However, others do not e.g. in Costa Rica rates increase consistently with age, while those in Cali, Columbia show little change with age. The absence in some instances of the defined type I pattern of rates associated with developing countries may be a result of continuing economic development with corresponding changes in the pattern of occurrence of disease to that associated with economically developed areas (type III). Correa & O'Conor (1971) predicted this.

1.4 Case Characteristics:

Age:

Attention was first called to the bimodality of HD incidence with peaks in young adults (15-34) and in the elderly by MacMahon (1957, 1966) and this has been confirmed by later work (Grufferman & Delzell, 1984) (Figure 1.1). Although these two peaks appear to be flattening there is little doubt that HD is still an unusual cancer in that the age pattern is not consistent with a simple increasing trend.

Sex:

A male excess of HD is apparent (Figure 1.1). Muir et al (1987) examined 77 HD incidence rates from 34 countries (excluding the USA). The mean male: female ratio was 1.7 (0.9-3.3). The age-specific curves for female HD for ALL countries resembles that of wealthy countries (Correa & O'Conor, 1971). The greatest male: female ratio occurs in children aged ≤10 years and in the later reproductive years (age 35-54 years).

Males account for the overwhelming majority of childhood cases (85%) (MacMahon, 1966). Glaser (1994) found the largest male excess of HD in early childhood, but by age 10-14 years the male and female rates converged and remained similar throughout adolescence. A possible explanation for this is that at least some forms of the disease have an infectious aetiology. Males are more susceptible than females to infectious diseases in childhood, especially in the first 5 years of life (Washburn et al, 1965). Resistance to infection may explain the lower risk of HD in female children. In the later reproductive years it may be that a protective effect of childbirth is responsible for lower rates in women.

Ernster et al (1979) observed that married white women aged 35-44 and 45-54 years had lower average annual incidence rates of HD than single women did in the USA ($2.2/10^5$ married vs. $3.7/10^5$ single aged 35-44 yrs; $2.5/10^5$ vs. $3.5/10^5$ respectively aged 45-54 yrs). Other lymphomas did not have this age-specific pattern. Hormonal factors have been a

neglected area of HD research but studies have recently been carried out looking at parity and HD risk (see section Hormonal factors and HD, p.50).

Race:

Whites in the USA have a higher incidence of HD than blacks and other ethnic sub-groups (Horn et al, 1996). This pattern could be due to the lower socio-economic status of non-whites in the USA. Orientals have the lowest incidence (Table 1.3). However, Mason & Fraumeni (1974) compared the HD mortality of Japanese in the USA and Japanese in Japan which is a low risk area (at this time mortality was an adequate proxy of incidence). They found that mortality was greater for Japanese Americans than Japanese in Japan. Table 1.3 also makes clear that Chinese women in Hawaii have the highest incidence of HD. Environmental and ethnic factors appear to interact.

1.5 Socio-Economic Status and related factors:

Socio-economic status (SES) is a measure of the social and economic standing of a person or a group of people. "The different patterns of the disease (HD) appear to be closely related to the economic stratification of the population at risk" (Correa & O'Conor, 1971).

In the poverty stricken communities of Columbia there is a high frequency of disease in male children. These cases usually belong to MC and LD subtypes of the disease with relatively poor prognosis. In contrast in wealthy communities the disease develops more frequently in young adults, at which time more favourable forms predominate, particularly NS (MacFarlane et al, 1995).

While there is a well-recognised international pattern of variation in the incidence of HD by SES the effect of SES within countries is less clear. Alexander et al (1991a) investigated 486 UK cases of HD in people aged <25 years diagnosed 1984-88. They found a RR of HD for high community SES vs. low of 1.21 (95% C.I. 1.01-1.46). Other authors have

found in developed countries that HD risk is higher in rural areas (Zambon et al, 1981; Hardell & Bengtsson, 1983) where SES may be expected to be lower depending on the country.

HD, especially in young adults, has been consistently associated with factors e.g. family size (small), number of playmates (few), high level of parental education, which suggest a high standard of living in childhood. High SES is important because it is a proxy for other factors e.g. late exposure to infection. Corea & O'Conor (1971) suggest that, in a given population, susceptibility to the agent or agents that cause HD is related to host response and immunity which, in turn, is dependent on environmental and socio-economic factors. In the more developed countries children are usually well nourished and increasingly protected from infectious disease. In these populations HD in children is uncommon but shows an initial peak in young adults. In developing countries and in communities with poverty and overcrowding nutrition is poor and susceptibility to all kinds of infection is high, especially in children. It is in such populations that HD in children is more common.

Childhood Social Class:

Studies investigating the risk of HD by childhood social class (as measured by various parental variables) are shown in Table 1.4. Few results are statistically significant but a higher level of parental education is related to an increased risk of HD in young adults in some studies. There is a different effect of social class on the risk of HD in children and young adults. Gutensohn & Shapiro (1982) found high parental social class to be protective for HD in children in the USA. In contrast, Zwitter et al (1996) investigated cases of HD aged 17-50 years in the USA and found higher parental class to be associated with a greater risk of HD.

Table 1.4: Risk of HD and parental social class.

Study	Numbers	Comparison	Risk Estimate	Significance
Gutensohn & Shapiro (1982) (USA)	C/C study 66 families with HD cases (age <15) & 186 control families	(i) family income 34-100 v. 0-33 (ii) family occ. class: high/ medium. v. low	(i) RR=0.28 (0.08-0.98) (ii) RR=0.76 (0.15-4.6)	(i) $p<0.05$ (ii) NS
Gutensohn & Cole (1981) (USA)	C/C study 225 cases (age 15-39) & 447 population controls	(i)high/middle s/c of parents v. low (ii)maternal edu. ≥12 yrs v. <12 yrs	(i) CRI=0.75 (ii) CRI=0.65	(i) NA (ii) NA
Zwitter et al (1996) (Slovenia)	Matched c/c study 192 female cases (aged 17-50) & 496 controls (age- & residence-matched)	(i) father high level edu v. elementary (ii) mother high level edu v. elementary (iii) white collar family v. industry	(i) OR=2.20 (1.10-4.40) (ii) OR=2.46 (0.89-6.79) (iii) OR=1.77 (1.10-2.67)	(i) p for trend 0.083 (ii) p for trend NA (iii) p for trend 0.053
Kirchoff et al (1980) (Brazil)	Matched c/c study 70 cases (age <79) & 70 controls (age- & sex-matched)	(i) mother's education. (ii) father's education.	(i) no sig. diff. (ii) no sig. diff.	(i) NS (ii) NS
Paffenbarger et al (1977) (USA)	Nested c/c study 45 male deaths (adult) & 180 surviving classmates.	(i) Family income (ii) Parental education. (iii) Parent occupation	(i) 'no effect' (ii) 'no effect' (iii) 'no effect'	(i) NS (ii) NS (iii) NS
Kravdal & Hansen (1993) (Norway)	441 female cases (age 15-56) diag. 1935-74 in a cohort 1.3m	relative incidence of HD in offspring by father's edu. level	(i) low 1.00 (ii) inter 0.93 (iii) high 1.03	(i) ref. (ii) NS (iii) NS

NS (not statistically significant) = $p>0.10$

Family size:

Studies investigating the risk of HD and sibship size can be seen in Table 1.5. All, apart from Zwitter et al (1996), show a decreasing risk of HD with increasing sibship size, although not all the results are statistically significant. Westergaard et al (1997) in a large record linkage study make it clear that large sibship size has opposite effects in children (age <15 years) and young adults (age >15 years). In Denmark large sibships gave rise to higher risks of HD in children but lower risks of HD in young adults. The trend among the young adults was statistically significantly different from that seen in children ($p<0.05$).

Birth Order:

Table 1.6 has the results of studies investigating risk of HD by birth order. Being later in the order is associated with a decreased risk of young adult HD. Again Westergaard et al (1997) found opposing trends in the young adults compared to the children. For HD diagnosed aged <15 years the adjusted RR of being born 4^{th} vs. 1^{st} was 1.50 (0.42-5.33), whereas for HD diagnosed aged >15 yrs the RR for the same category was 0.30 (0.1-0.97).

Own Education:

Studies investigating risk of HD by own educational level are in Table 1.7. All studies have investigated cases aged over 15 years and found an increased risk of HD with higher educational levels. However, another study that looked at young adult cases alone had statistically significant results. Most studies have looked at a wide range of ages. Mueller et al (1989) found that older cases had less education than controls, although this result was not statistically significant.

Socio-Economic Status and related factors: Conclusion:

There is a consistent body of evidence that HD is associated with factors in childhood that influence the age of infection with infectious agents and bacteria. The majority of work looking at childhood social environment has looked at all ages of HD

Table 1.5: Family size and risk of HD.

Study	Numbers	Comparison	Risk Estimate	Significance
Gutensohn et al (1975) (USA)	C/C study 136 cases (age 15-44) & 315 sibling & 78 spouse controls	(i) 1 sib (ii) 2 sibs (iii) 3 sibs (ref) (iv) 4 sibs (v) 5+ sibs	(i) RR=1.59 (ii) RR=1.07 (iii) RR=1.00 (iv) RR=0.90 (v) RR=0.81	Not given
Gutensohn & Cole (1981) (USA)	C/C study 225 cases (age 15-39) & 447 population controls	(i) 6+ sibs v. 1/2 (ii) 4/5 sibs v. 1/2 (iii) 3 sibs v. 1/2	(i) RR=0.56 (ii) RR=0.64 (iii) RR=0.64	p-trend 0.02
Zwitter et al (1996) (Slovenia)	Matched c/c study 192 female cases (aged 17-50) & 496 controls (age- & residence-matched)	(i) 2 sibs v. 1 (ii) 3 sibs v. 1 (iii) 4+ sibs v. 1	(i) OR=1.47 (0.77-2.80) (ii) OR=1.26 (0.64-2.49) (iii) OR=1.54 (0.81-2.92)	(i) NS (ii) NS (iii) NS
Kirchoff et al (1980) (Brazil)	Matched c/c study 70 cases (age <79) & 70 controls (age- & sex-matched)	mean sibship size cases v. controls	7.0 v. 6.6	NS
Bernard et al (1987) (UK)	Matched c/c study 248 cases (adult) & 489 hospital controls (age-, sex-, & health district-matched)	(i) 0/1 v. 5+ sibs (ii) 2 v. 5+ sibs (iii) 3 or 4 v. 5+ sibs	(i) RR=1.8 (1.1-2.7) (ii) RR=1.6 (0.9-2.6) (iii) RR=1.5 (0.9-2.4)	(i) p=0.02 (ii) p=0.06 (iii) NS
Bonelli et al (1990) (Italy)	C/C study 160 cases (age 15-78) & 185 hospital controls	sibship size (continuous variable)	RR=0.63 (0.46-0.86)	p for trend <0.01
Serraino et al (1991) (Italy)	C/C study 152 cases (age 15-77) & 613 hospital controls	(i) 4+ v. 1 sib (ii) as (i) but NS sub type only.	(i) RR=0.7 (0.3-1.4) (ii) RR=0.8 (0.3-2.3)	(i) p for trend NS (ii) p for trend NS
Paffenbarger et al (1977) (USA)	Nested c/c study 45 male deaths (adult) & 180 surviving classmates.	(i) 0 sibs vs. 1 sib (ii) 2+ sibs vs. 1 sib	(i) RR=1.4 (ii) RR=0.9	(i) NS (ii) NS
Westergaard et al (1997) (Denmark)	Cohort 2.1 million 1968-92 linked to Danish cancer registry => 378 HD (age 0-42)	(i) HD aged <15, 5+ sibs vs. 1 (ii) HD aged >15, 5+ sibs vs. 1	(i) RR=3.31 (1.36-8.02) (ii) RR=0.57 (0.30-1.08)	(i) p for trend 0.06 (ii) p for trend NS

NS (not statistically significant) = p>0.10

Table 1.6: Birth order and risk of HD.

Study	Numbers	Comparison	Risk Estimate	Significance
Gutensohn & Cole (1981) (USA)	C/C study 225 cases (age 15-39) & 447 population controls	4^{th}+ v. 1-3^{rd} (adjusted for sibship size)	RI = 0.66	p=0.07
Zwitter et al (1996) (Slovenia)	Matched c/c study 192 female cases (aged 17-50) & 496 controls (age- & residence-matched)	3^{rd} + v. 1^{st}	OR=0.86 (0.53-1.40)	NS
Kirchoff et al (1980) (Brazil)	Matched c/c study 70 cases (age <79) & 70 controls (age- & sex-matched)	Pairs of birth order	no sig. tendency for HD cases to be born early or late in birth sequence	NS
McKinney et al (1990) (UK)	Matched c/c study 248 cases (adult) & 489 hospital controls (age-, sex-, & health district-matched)	1^{st} v. later	OR=1.09 (0.76-1.56)	NS
Serraino et al (1991) (Italy)	C/C study 152 cases (age 15-77) & 613 hospital controls	(i) 3^{rd} + v. 1^{st} (all subtypes) (ii) 3^{rd} + v. 1^{st} (NS subtype only)	(i) RR=0.7 (0.4-1.2) (ii) RR=1.2 (0.5-2.9)	(i) p for trend NS (ii) p for trend NS
Westergaard et al (1997) (Denmark)	Cohort 2.1 million 1968-92 linked to Danish cancer registry => 378 HD (age 0-42)	(i) HD aged <15, 4^{th} + vs. 1^{st} (ii) HD aged >15, 4^{th} + vs. 1^{st}	(i) 1.50 (0.42-5.33) (ii) 0.30 (0.10-0.97)	(i) p for trend NS (ii) p for trend 0.07

NS (not statistically significant) = p>0.10

Table 1.7: HD risk and educational level.

Study	Numbers	Comparison	Risk Estimate	Significance
Gutensohn et al (1975) (USA)	C/C study 136 cases (age 15-44) & 315 sibling & 78 spouse controls	(i) no finish high school (ii) high school (iii) some college (ref) (iv) college grad. (v) grad. school	(i) RR=0.7 (ii) RR=1.1 (iii) RR=1.0 (iv) RR=1.6 (v) RR=1.8	Not given
Zwitter et al (1996) (Slovenia)	Matched c/c study 192 female cases (aged 17-50) & 496 controls (age- & residence-matched)	(i) middle school v. elementary (ii) high school v. elementary	(i) OR=1.13 (0.75-1.71) (ii) OR=1.15 (0.65-2.04)	(i) NS (ii) NS
Abramson et al (1978) (Israel)	Matched c/c study 343 cases (age ≥17) & population controls (sex-, age- & origin-matched)	high school graduate v. not	RR=1.7	p=0.007
Kirchoff et al (1980) (Brazil)	Matched c/c study 70 cases (age <79) & 70 controls (age- & sex-matched)	age left school	cases 9.2 v. controls 7.6	p=0.001
Bonelli et al (1990) (Italy)	C/C study 160 cases (age 15-78) & 185 hospital controls	(i) 14+ yrs edu. v. <5 yrs (ii) as (i) age 15-39 (iii) as (i) age 40+	(i) RR=6.68 (1.94-23.88) (ii) RR=44.48 (3.56-556.4) (iii) RR=10.61 (0.99-113.4)	(i) p for trend <0.001 (ii) p for trend <0.01 (iii) p for trend NS
Serraino et al (1991) (Italy)	C/C study 152 cases (age 15-77) & 613 hospital controls	(i) 14+ yrs edu. v. ≤5 yrs (adjusted for age & sex) (ii) as (i) NS subtype only	(i) RR=2.0 (1.0-3.9) (ii) RR=4.4 (1.8-11.0)	(i) p for trend NS (ii) p for trend 0.002
Mueller et al (1989) (USA, Norway)	Nested c/c study 43 cases (age 19-80) & 96 controls (sex-, race- & DOB-matched)	amount of education	(i) young adult HD more education than controls (ii) older HD less education than controls	(i) NS (ii) NS
Kravdal & Hansen (1993) (Norway)	441 female cases (age 15-56) diag. 1935-74 in a cohort 1.3m	relative incidence of HD by education level: (i) low; (ii) intermediate; (iii) high.	(i) RR=1.00 (ii) RR=1.12 (iii) RR=1.31	NS

NS (not statistically significant) = p>0.10

combined. When individual age groups are considered most work has been performed on young adults (age 15-34 years). There are limited data that suggest that children who develop HD are of lower SES (Gutensohn & Shapiro, 1982; Medeiros & Greiner, 1995; Westergaard et al, 1997). These children may be at increased risk of early infection. In contrast, among young adults the occurrence of HD is consistently associated with factors that suggest the delay or avoidance of childhood infection. For cases aged over 50 years there is no association of indicators of childhood social class with risk of HD.

1.6 HD and infection:

HD has a clinical course with many characteristics of an infection e.g. unexplained persistent fever, frequent sweating and weight loss. MacMahon (1966) proposed that an infectious agent caused HD in young adults (age 15-34 years) whereas at older ages (age ≥50 years) it was more of a traditional neoplasm.

Late-host-response model:

The evidence for an infectious aetiology of young adult HD is consistent with the late-host-response model under which exposure to an infectious agent (or group of agents) is relatively benign in early childhood but may lead to oncogenic change if delayed until school or adolescence (Gutensohn & Cole, 1977). After childhood, infection with several common viruses tends to be more clinically severe (Mueller et al, 1989). It has been proposed that if such infection results in unstable or incomplete immunity against a latent virus, persistent or intermittent viral activity may trigger a chain of molecular events that lead to the development of HD (Meuller et al, 1989). The infectious agent could be relatively common. If the agent was both ubiquitous and geographically uniform then this model would not require any case-to-case contact and would not result in case clustering.

The late-host response model is related to the paralytic polio model, under which, like paralytic polio, HD is the rare manifestation of a common infection with the probability of disease development increasing as the age at infection is delayed. Prevailing socio-economic conditions would determine the age of exposure to the causative agent and this determines the likelihood of those exposed developing the disease. The late host-response and the polio models agree for young adult HD but differ in their views of disease causation in children. In the paralytic polio model the same agent causes childhood and young adult HD.

If the late host-response model is correct factors which delay the age at which common infections are first encountered should be associated with increased risk of young adult HD e.g. few siblings, late birth order. These have been discussed above. In addition, a variety of evidence suggests HD, in particular young adult HD, could have an infectious aetiology: clustering investigations, the relationship of HD to various viruses, and its seasonal pattern of presentation.

Clustering:

There is a long history of research into disease clusters to find clues regarding causation but, there is no clearly accepted definition of a cluster. A cluster could be considered to be a focus of a higher than expected number of cases. The approach to clustering has undergone three eras (Boyle et al, 1996): the anecdotal report; the wide ranging search; and the statistical review. The statistical review can be divided up into space-time clustering and spatial clustering.

Space-time clustering: Vianna et al (1973) described 12 people with HD who were included among or could be "linked" by social contact to the members of one graduating high school class. This interlinked group was later extended to 31 cases. These authors stressed that social contacts were close and repeated and inferred that there was transmission of an aetiologic agent from case to case, either directly or through healthy contacts. It is difficult to assess the

significance of the findings in this report as the cases were identified retrospectively on the basis of shared exposures and thus, it was impossible to select a control population (Gutensohn & Cole, 1977). This deficiency was largely overcome by a later study in Nassau and Suffolk counties, New York.

Vianna & Polan's (1979) New York study was based on data from an area that had not previously been investigated for clustering. In 8 public high schools that had had a case of HD (student or teacher) 1960-64 there were 10 HD cases 1965-69. In the other 143 high schools of comparable size there were only 9 cases over the same time period. These authors also found that after the occurrence of a 'primary' case in 20 schools there were 21 cases in students and 7 in teachers in the same schools during the decade. The expected number of HD cases in students was 9.3 and in teachers it was 0.3. These results are of interest but have been criticised due to underascertainment of cases. The disease incidence that they reported was lower than the HD mortality for the same area 1960-69 (Grufferman et al, 1979).

Other attempts to detect space-time clustering have been made but two other studies have reported essentially negative results (Alderson & Nayak, 1971; Kryscio et al, 1969). However, if the period between contact and appearance of clinical symptoms is several years and variable and the population is mobile, space-time clustering may not be evident for clinical disease, even if its agent is transferred from person to person (Chen et al, 1984). In general studies of space-time interaction have given rise to inconclusive results (Kryscio et al, 1973; Greenberg et al, 1983). This is not surprising as these analyses have low statistical power for the putative aetiological models for HD (Chen et al, 1984).

Spatial clustering: More recent studies have concentrated on spatial clustering using the nearest neighbour test (NNA) and other methodologies. Alexander et al (1991b) used the NNA test and showed significant evidence of spatial clustering amongst younger cases (aged 0-34) but not in older cases (aged 50-79). The NNA test found 15% of cases aged 0-24 were

clustered ($p<0.01$) which lends support to young onset HD having an infectious aetiology. However, most evidence for this is in the 15-34 age group and this was not specifically addressed by the study. More recently Alexander et al (1995b) tested 494 young adult HD cases in the Yorkshire Health Region for spatial clustering. 18% of the cases in young people were classified as clustered ($p\leq0.05$) while no clustering was seen in the older cases.

There is significant evidence for weak clustering in HD cases. The pattern of clustering is consistent with an aetiology involving either delayed exposure to a common, but not ubiquitous, virus or infection with a virus with a long latent period (Jarrett, 1992).

Seasonal variation:

If an infectious agent is involved in the aetiology of HD a seasonal variation in presentation might be expected depending on the behaviour of the agent, the length and variability of the latent period and the method of spread. Thirteen studies have investigated this question (Cridland, 1961; Innes & Newell, 1961; Fraumeni & Li, 1962; Bjelke et al, 1969; Modan et al, 1969; Newell et al, 1972; Bogger-Goren et al, 1983; Alderson & Nayak, 1971; Newell et al, 1985; Nielly et al, 1995a & b; Douglas et al, 1996; Westerbeek et al, 1998; Douglas et al, 1998) (discussed in detail in Chapter 6). The first seven of these only observed the number of cases occurring each month with four studies concluding HD did have a seasonal presentation (Cridland, 1961; Innes & Newell, 1961; Fraumeni & Li, 1962; Bogger-Goren et al, 1983).

The latter six studies have used more formal statistical methodologies, either a test devised by Edwards (1961) or cosinor analysis. All six found a statistically significant seasonal presentation of HD. Interestingly, three of the studies divided cases by age and found the effect to be confined to the young adult age group (age 15-39 years Newell et al (1985); age <40 years Neilly et al (1995a & b) and Douglas et al (1996)). Douglas et al (1998) only found a significant seasonal presentation for NS subtype at all ages. However, the majority of

this effect was seen in the 0-44 years age group. The peak of presentation in all of these studies was at the beginning of the year (January-March).

The seasonal presentation of HD could reflect the seasonal pattern of a causal agent(s) or the pattern of a disclosing agent (one which promotes the diagnosis of an existing pre-clinical condition).

Infectious illness:

If the late-host-response model is correct then, assuming similar transmission routes to agents causing ordinary childhood infectious illness, young adult cases of HD should have fewer childhood infections. Infection may also occur at older ages. Studies of childhood infection and HD risk are presented in Table 1.8. Only one study, Paffenbarger et al (1977) found that men who reported common contagious disease in childhood were less likely to develop (fatal) HD. These results are particularly strong as it was a prospective study and, therefore, uninfluenced by recall bias.

It may be expected that cases would be infected older than controls, however, only one study (Evans & Gutensohn, 1984) found the age of onset of infection distinguished cases from controls, while three others (Paffenbarger et al, 1977; Kirschoff et al, 1980; Gutensohn & Cole, 1981) did not.

There is no consistent evidence linking any of the infections in Table 1.8 with HD. However, one virus has been consistently related to HD: Epstein-Barr virus (EBV).

Table 1.8: Risk of HD following childhood infection.

Study	Numbers	Comparison	Risk Estimate	Significance
Vianna et al (1971) (USA)	Matched c/c study 109 cases (age ≤40) & hospital. controls (age-, sex-, race-, residence-matched)	Past history of frequent colds, sore throats or systemic infection	No significant difference*	NS
Gutensohn & Cole (1981) (USA)	C/C study 225 cases (age 15-39) & 447 population controls	mean age of onset of measles or chicken pox	no significant difference*	NS
Newell et al (1973) (USA)	Matched c/c study 176 cases (aged ≥5) & controls (sex-, age-, social class-matched)	history of mumps history of chicken pox	RR=1.0 RR=1.1	NS NS
Kirchoff et al (1980) (Brazil)	Matched c/c study 70 cases (age <79) & 70 controls (age- & sex-matched)	(i) history of measles & chicken pox (ii) mean age of onset of measles & chicken pox	(i) no sig. difference* (ii) no sig. difference*	(i) NS (ii) NS
Evans & Gutensohn (1984) (USA)	C/C study 304 cases & 285 sibling controls	Antibody levels to varicella, measles, rubella	"quite similar" levels for cases & controls* mean measles age 6.5 vs. 5.7	NS p=0.01
Paffenbarger et al (1977) (USA)	Nested c/c study 45 male deaths (adult) & 180 surviving classmates.	pertussis scarlet fever measles varicella mumps influenza	RR=0.5 RR=0.4 RR=0.8 RR=0.6 RR=0.7 RR=0.6	p=0.04 p=0.07 NS NS NS NS
Andersson & Isager (1978) (Denmark)	Nested c/c study 65 cases (age 15-45) & 189 controls (age-, sex- & SES-matched)	measles	27 cases vs. 61 controls*	NS

NS (not statistically significant) = p>0.10

* No relative risks given in these papers.

Hodgkin's Disease and the Epstein-Barr virus:

EBV definition and background: EBV is a B-cell lymphotrophic virus with infectious and oncogenic properties (Hanto et al, 1981). Higginson et al (1992) estimated that about 95% of people have antibodies to this virus. In the great majority of EBV carriers infection is completely apathogenic, yet the virus is potentially oncogenic because a subset of its genes have the capacity to de-regulate cell growth (Rickinson, 1994). EBV is involved in the aetiology of African Burkitts lymphoma and nasopharyngeal carcinoma.

Primary infection generally occurs during the first few years of life, where it is usually asymptomatic. However, when infection is delayed until adolescence or later it often manifests clinically as infectious mononucleosis (IM). In IM the tonsils and cervical lymph nodes contain HRS cells showing strong latent membrane protein (LMP-1) expression (O'Grady et al, 1994). It is still not known why primary infection in infancy is almost always asymptomatic whereas later in life it can produce overt disease. It may relate to the size of the initial dose of the virus rather than to age per se. High doses of an orally transmitted virus are much more likely to be delivered by salivary exchange between consenting adults than from adult to child or child to child. Infection with a large virus dose makes it much more likely that virus driven expansion of the B-cell pool will exceed a critical threshold beyond which the atypical T-cell response is activated and disease ensues (Rickinson, 1994).

Infectious Mononucleosis and Hodgkin's Disease:

The first suggestion that EBV was associated with HD came from observations that the age incidence curves for IM and HD are very similar. The HRS cell, which is pathognomic for HD, has been observed in the lymph nodes of some IM patients. The results of cohort studies investigating the risk of HD following IM are shown in Table 1.9.

Table 1.9. Summary of results of prospective studies investigating IM & HD.

Study Name	Numbers	Diagnosis	Follow-up	Risk Estimate
Miller & Beebe (1973) (USA)	Cohort of 2437 WWII veterans with IM (2 HD)	Army diagnosis on discharge	1945-65	Total HD RR=2.0 (2/1*)
Connolly & Christie (1974) (USA)	Cohort of 4529 people (all ages) (5 HD)	Basis NK	1948-64	Male HD RR=2.0 (1/0.5*) Female HD RR=4.4 (4/0.9) Total HD RR=3.6 (5/1.4*)
Rosdahl et al (1974) (Denmark)	Cohort of 17073 (age <34) (17 HD)	+ve Paul-Bunnell test	test 1940-69, followed 1943-70	Male HD RR=4 (16/4.02) Female HD RR=0.50 (1/1.99) Total HD RR=2.8 (17/6.01)
Carter et al (1977) (USA)	Cohort 2282 former college students with IM (3 HD)	Unequivocal elevation of anti-body & other clinical illness	IM diagnosed 1949-69, followed 1949-onwards	Total HD RR=2.3 (3/1.3)
Munoz et al (1978) (Scotland, Sweden)	Cohort 9454 IM cases (7 HD)	Serologically confirmed	IM diagnosed 1950-73, followed to 1973	Male HD RR=1.9 (2/1.06) Female HD RR=7.0 (5/0.71) Total HD RR=4.0 (7/1.8)
Kvale et al (1979)	Cohort of 5840 IM (most age <34), (5 HD)	+ve Paul-Bunnell test	test 1961-72, followed 1961-75	Male HD RR=4.9 (4/0.81) Female HD RR=2.2 (1/0.46) Total HD RR=3.9 (5/1.27))
Lumio & Karjlainen (1993) (Finland)	Cohort 1234 severe IM 1954-83 (0 HD)	Confirmed from patient records	IM diagnosed 1954-83, followed to 1988	NA
Summary	41615			Male HD RR=3.60 (23/6.39) Female HD RR=2.70 (11/4.06) Total HD RR=3.05 (39/12.78)

* expected numbers estimated as half those for total lymphomas

Previous IM is associated with an increased risk of HD in all cohort studies (Miller & Beebe, 1973; Connolly & Christie, 1974; Rosdahl et al, 1974; Carter et al, 1977; Munoz et al, 1978; Kvale et al, 1979), with one exception (Kumio & Karjalainen, 1993). These cohort studies combined consist of almost 42000 people with, overall, an approximately 3-fold increase (39 HD observed/12.78 expected) in the risk of HD.

There are problems associated with studying the relationship between HD and IM. Risk of young adult HD is increased among those of higher SES, as is the incidence of IM. Thus, the findings of the cohort studies could simply reflect confounding by other risk factors related to susceptibility to late infection. Therefore, studies comparing HD incidence in IM and healthy cohorts should control for social class. In case-control studies which have controlled for social class the relationship between previous IM and HD remains.

Newell et al (1973) compared 176 cases and 176 matched population controls aged <44 years (matched for age, sex, race, and SES). Prior IM was associated with a non-significant increase in risk of HD (RR=1.4). Gutensohn & Cole (1981) also found more cases than controls aged 15-39 years reported IM one or more years prior to diagnosis with HD after controlling for birth order, family size and housing (relative incidence 1.8, 95% C.I lower boundary 1.1). Evans & Gutensohn (1984) found a RR of 1.5 (95% C.I lower boundary 0.96) for a previous history of IM in a study of 262 HD cases aged 15-54 years and 250 sibling controls.

Another problem which occurs when investigating the association of IM to HD is potential misclassification bias which arises in some people who, in the course of developing HD, are temporarily diagnosed with IM as the diseases have similar symptoms.

Combined the evidence from available studies suggests a role for previous IM in increasing HD risk and this in turn implicates EBV.

HD and EBV Serology:

Serological methods have been used to investigate the potential role of many viruses in HD (rubella, measles, parainfluenza, herpes simplex, varicella zoster, cytomegalovirus, Human Herpes Viruses 6 and 7) but the most consistent findings have come from EBV. Antibodies to viral capsid antigen (VCA) are found in all individuals who have been exposed to EBV and antibodies to early antigen (EA) are detected during primary infection and during viral reactivation.

Many case-control studies have been performed which show that patients with HD have elevated levels of antibodies to EBV (Henle & Henle, 1973; Hesse et al, 1977; Evans & Gutensohn, 1984). The majority of these have found increased antibody titres to both VCA and EA. However, treatment affects EBV antibody levels as does the presence of HD itself. Thus, the elevated antibody titres seen in HD could be due to viral reactivation secondary to therapy or onset of disease.

Mueller et al (1989) conducted a nested case-control study of 43 HD cases and 96 controls in a cohort with serum stored prior to diagnosis to test the association between elevated antibody titres to several EBV antigens. Relative risk of developing HD was significantly elevated in people with higher antibodies to the VCA and EA. These findings were confirmed by a second prospective study (Lehtinen et al, 1993). Elevated EBV antibody levels during a period before the onset of recognisable disease suggests EBV could have a role in the pathogenesis of HD. Alternatively, raised antibody levels could be a marker for infection by another virus (Jarrett, 1992).

HD and presence of EBV viral DNA:

The presence of the virus in current biological samples can be found by molecular methods including Southern blot analysis, Polymerase Chain Reaction (PCR), DNA and RNA in situ hybridisation (ISH) and immunohistochemistry (Armstrong at al, 1993). Southern

blotting has the advantage of determining the clonality of the infected tumour with respect to the viral terminal repeat sequences. ISH has the advantage of precise localisation of the virus in infected tumours (Gulley & Raab Traub, 1993). PCR is a very sensitive technique that tests for the presence of a virus in a biopsy sample, but not necessarily in the tumour cells. Given the ubiquitous nature of EBV, the PCR technique may amplify the EBV DNA from the reservoir of non-pathogenic cells. The results of PCR analysis have been criticised for over-estimating the presence of EBV in HD by a factor of two (Delsol et al, 1992).

The ideal assay for the analysis of HD samples should be reproducible, sensitive, permit cellular localisation and be applicable to the study of archive material (Armstrong et al, 1993). EBER ISH meets all these requirements and is the 'gold standard' of all molecular techniques for defining cases of EBV associated HD (Gulley & Roab Traub, 1993).

Table 1.10 summarises studies that have utilised the EBV EBER ISH assay or EBV LMP-1 immunohistochemistry in adult HD or HD at any age. These studies are emphasised as these techniques permit the localisation of EBV to the HRS cells. The proportion of EBV-associated HD ranges from 19% (Belkaid et al, 1995) to 94% (Chang et al, 1993) (EBV-associated HD means evidence of EBV has been found in the HRS cells). This variation may be because many of the series are small, of different ages, ethnicity and histology.

While many studies have used one or other of the techniques some have used both allowing a comparison of the results from each method to be made. Authors have found identical (Belkaid et al, 1995; Paulino et al, 1996) or very similar (Herbst et al, 1992; Chang et al, 1993; Zhou et al, 1993; Kanavaros et al, 1994; Park et al, 1994; Li et al, 1995; Kusada et al, 1998) proportions of EBV-associated cases using each method. However, Huh et al, (1996) found 60/87 (69%) of cases to be EBV-associated using EBV EBER ISH but only

Table 1.10. Summary of results of studies using EBER ISH and/or LMP-1 immunohistochemistry to investigate EBV-association rates in adult HD and HD at any age.

Study Name	Country	No. of cases (age in years)	Total EBV (%)	MC (%)	NS (%)
Brousset et al (1991)	France	54 (age 10-81)	16/54 (30)	12/26 (46)	4/22 (18)
Pallesen et al (1991)	Denmark	84 (no ages)	40/84 (48)	23/24 (96)	16/50 (32)
Weiss et al (1989)	USA	36 (no ages)	11/36 (31)	6/8 (75)	4/12 (33)
Delsol et al (1992)	France	107 (no ages)	36/107 (34)	32/55 (58)	4/40 (10)
Herbst et al (1992)	Germany	46 (no ages)	23/46 (50)	10/18 (56)	10/24 (42)
Khan et al (1993)	UK	77 (age 3-72)	25/77 (32)	15/22 (68)	9/38 (24)
Pinkus et al (1994)	USA	186	48/186 (26)	27/39 (69)	18/127 (14)
Poppema & Visser (1994)	Canada	72 (no ages)	19/72 (26)	12/14 (86)	7/53 (13)
Belkaid et al (1995)	France	21 (med. age 29)	4/21 (19)	2/NA*	2/NA*
Jarrett et al (1996)	UK	130 (all ages)	37/130 (28)	20/42 (48)	13/68 (19)
Kordeck et al (1996)	Poland	135 (all ages)	44/135 (33)	21/34 (62)	23/90 (26)
Leocini et al (1996)	Italy	65 (age 5-86)	31/65 (48)	12/13 (92)	17/37 (46)
Enblad et al (1997)	Sweden	107 (age 6-87)	27/107 (25)	8/13 (62)	18/81 (22)
Europe & USA		**1120**	**361/1120 (32)**	**198/308 (64)**	**143/642 (22)**
Zhou et al (1993)	China	28 (age 4-80)	17/28 (61)	10/11 (91)	6/14 (43)
Park et al (1994)	Korea	25 (no ages)	19/25 (76)	11/13 (85)	NA*
Chan et al (1995)	H. Kong	23 (age 5-82)	15/23 (65)	5/5 (100)	6/16 (56)
Li et al (1995)	Japan	40 (all ages)	17/40 (43)	NA*	NA*
Tomita et al (1996)	Japan	57 (age 5-80)	32/50 (64)	21/25 (84)	11/25 (44)*+
Huh et al (1996)	Korea	87 (mean age 38)	60/87 (69)	38/51 (75)	10/17 (59)
Paulino et al (1996)	Philipine	21 (age 11-64)	9/21 (43)	6/9 (67)	3/10 (30)
Benharroch et al (1997)	Israel	106 (all ages) 42 (age 15-34)	32/106 (30) 10/42 (23)*	15/33 (45)	14/64 (21)
Peh et al (1997)	Malaysia	55 (age ≥15)	26/51 (52)	19/22 (86)	6/27 (22)
Kusada et al (1998)	Japan	46 (age 6-82)	27/46 (49)	12/18 (67)	5/14 (36)
Liu et al (1998)	Taiwan	70 (age 7-75)	44/70 (63)	18/26 (69)	23/36 (64)
Mourad et al (1998)	S. Arabia	62 (age 3-72)	29/62 (47)	8/9 (89)	21/45 (47)
Asia/Middle East		**620**	**327/609 (54)**	**163/222 (73)**	**104/243 (43)**
Belkaid et al (1995)	Algeria	47 (med. age 26)	16/47 (44)	12/NA*	4/NA*
Weinrels et al (1996)	Kenya	48 aged ≥16	32/48 (67)	2/2 (100)	26/39 (67)
Leocini et al (1996)	Kenya	92 (age 2-76)	85/92 (92)	44/44 (100)	35/40 (88)
Kusada et al (1998)	Kenya	48 (age 4-65)	38/48 (79)	28/33 (85)	4/7 (57)
Africa		**235**	**171/235 (73)**	**74/79 (94)**	**65/86 (76)**
Chang et al (1993)	Peru	32 (age 2-75)	30/32 (94)	18/20 (90)	7/7 (100)
Quintanilla-Martinez et al (1995)	Mexico	39 (adults)	28/39 (72)	10/15 (67)	6/12 (50)
Zorate-Osorno et al (1995)	Mexico	27 (age 5-65)	18/27 (67)	7/7 (100)	6/13 (46)
Monterosso et al (1998)	Costa Rica	40 (age 6-84)	16/40 (40)	12/14 (86)	3/20 (15)
South America		**138**	**92/138 (67)**	**47/56 (84)**	**22/52 (42)**
TOTAL		**2113**	**951/2102 (45)**	**482/665 (72)**	**334/1023 (33)**

* not in totals. + AOS, all other subtypes.

39/87 (45%) to be associated using LMP-1 immunohistochemistry. The authors suggest that this difference could be result from technical variations in the formalin fixation they used. In the studies where there is a difference between the two methods ISH generally gives higher proportions of EBV-association, as it is a more sensitive method.

As well as the cellular localisation of the EBV genomes the clonality of the virus must be determined before a meaningful association between EBV and HD can be deduced. Monoclonality signifies that cells in the tissue specimen arose from one precursor cell and indicates a cancerous process whereby one cell divides continually without proper control. Polyclonality implies that the cells are distinct and arise secondary to inflammation indicating a benign reactive process. Several studies have assessed the clonality of the EBV genomes within HD tumours (Weiss et al, 1986, 1989; Anagnostopoulos et al, 1989; Boicchi et al, 1989; Staal et al, 1989; Gledhill et al, 1991; Jarrett et al, 1991). In the majority of cases the infected cells have been found to be clonal with respect to EBV.

The EBV genome has the capacity to encode for over 80 proteins and there is evidence that many of these are transcribed during viral replication (Baer et al, 1984). B cells which have been immortalised by EBV in vitro are known to express nine proteins: six nuclear proteins, the EBNAs; three membrane proteins, the latent membrane proteins (LMP-1, 2A, 2B). There are also two small RNAs, EBER-1 and 2, which may play a role in viral RNA processing. EBNA-1 is a DNA binding protein that enables the viral genome to exist in an episomal or plasmid state. EBNA-2 and LMP-1 proteins appear to play a critical role in cell immortalisation. Studies have demonstrated that EBV is transcriptionally active, as demonstrated by the presence of EBER RNA, EBNA-1, and LMP-1. To date the evidence for HD is that HRS cells show a latent infection with EBV comparable to that found in nasopharyngeal carcinoma (EBNA-1 +, LMP-1 +, EBER RNA +) (Brousset & Delsol, 1991).

The presence of monoclonal EBV genomes and localisation of the EBV viral DNA in the tumour cells supports the concept of an aetiological role of EBV in the pathogenesis of a significant proportion of HD cases (Herbst et al, 1992).

EBV association by histological subtype:

Table 1.10 also shows a comparison of the proportion of EBV-associated HD by MC and NS subtype. In all series, except Chang et al (1993), the MC subtype has a higher proportion of EBV-associated cases (overall 72% for MC and 33% for NS). The proportion of EBV-associated MC is statistically significantly higher than NS (Vestlev et al, 1992; Chang et al, 1993; Kordeck et al, 1996; Leocini et al, 1996; Benharroch et al, 1997; Peh et al, 1997; Enblad et al, 1997). Others have noted the smaller number of NS cases associated with EBV but statistical testing has not proved significant (Gledhill et al, 1991; Boiochhi et al, 1989).

Glaser et al (1997) produced a meta-analysis to discover the factors that predispose to EBV-associated HD summarising the findings of twelve research groups. They analysed 1546 cases of HD, including 368 unpublished cases. Of the 1546 cases, 618 (40%) were EBV-associated (compared with 45% overall in Table 1.10). The proportion of EBV-associated HD cases varied significantly across histologic subtype, age, sex, ethnicity, country of residence, and regional economic level (all $p \geq 0.001$) when entered separately as explanatory variables. These factors do not necessarily remain significant if entered into a model together (see p38). In the total series 70.4% of the MC subtype was EBV-associated compared with only 23.2% of NS. This difference across subtypes remained after stratification for age. In the summary of published studies in Table 1.10 72% of MC and 33% of NS cases were EBV-associated, a finding similar to Glaser et al (1997).

EBV-association by age:

The first authors to report the association of EBV status of HD by age were Jarrett et al (1991). These authors found 54% of HD cases aged less than 15 years and 71% of cases older than 50 years were EBV-associated. For cases aged 15-34 years the corresponding figure was less than 15%. Studies of EBV-association rates in paediatric HD can be seen in Table 1.11. Overall, 59% of published paediatric HD cases are associated with EBV (compared to 45% older cases (Table 1.10)). Within the 0-14 years age group several authors have found that the highest proportion of EBV +ve HD occurred at the youngest ages (0-4 years) and falls as age increases (Armstrong et al (1993); Precidio et al (1995); Jarrett et al (1996); and Andriko et al (1997). MC subtype again has the highest percentage of EBV-associated HD (83% MC vs. 53% NS). Where tests were performed the difference in EBV association between MC and NS was found to be statistically significant (Weinreb et al, 1992 & 1996).

Chang et al (1993) and Peh et al (1997) found the difference in EBV-association by age at diagnosis to be statistically significant. Li et al (1995), Chan et al (1995), Huh et al (1996), and Liu et al (1998) also observed higher EBV-association rates in children and the elderly although no statistical tests were performed. Gledhill et al (1991), Zhou et al (1993), and Tomita et al (1996) found an increasing trend for the proportion of EBV-associated HD cases increasing with age. In the meta-analysis by Glaser et al (1997) the proportion of HD cases associated with EBV differed by age, with the highest percentage in children aged ≤ 10 and adults aged ≥ 80 years and the lowest in young adults aged 15-29 years. Jarrett et al (1996) calculated that age at diagnosis was a statistically significant predictor of cases being associated with EBV, even after adjustment for histiologic subtype and sex.

However, Boiocchi et al (1989), Quintanilla-Martinez et al (1995), and Leocini et al (1996) failed to detect an association between EBV status and age. This failure may well

Table 1.11. Summary of results of studies using EBER ISH and/or EBV LMP-1 immunohistochemistry to investigate EBV-association rates in paediatric HD.

Study Name	Country	No. of cases (age in years)	Total EBV (%)	MC (%)	NS (%)
Jarrett et al (1991)	UK	24 (age <15)	13/24 (54)	NA*	NA*
Weinreb et al (1992)	UK	74 (age <15)	37/74 (50)	17/20 (85)	14/35 (39)
Ambinder et al (1993)	USA	25 (age <15)	9/25 (36)	6/7 (86)	2/15 (13)
Armstrong et al (1993)	UK	22 (age <15)	13/22 (59)	7/7 (100)	4/12 (33)
Khan et al (1993)	UK	24 (age <15)	6/24 (25)	2/4 (50)	4/19 (21)
Claviez et al (1994)	Germany	22 (age 4-17)	10/21 (48)	NA*	NA*
Kanavaros et al (1994)	Greece	22 (age 3-15)	12/22 (54)	8/11 (73)	4/10 (40)
Weinreb et al (1996)	UK	75 (age <16)	38/75 (51)	16/19 (84)	15/37 (41)
	Australia	16 (age <16)	11/16 (69)	3/3 (100)	8/11 (73)
	Greece	22 (age <16)	20/22 (91)	6/7 (86)	11/12 (92)
Razzouk et al (1997)	USA	26 (med. age 7.5)	15/26 (58)	NA*	NA*
Andriko et al (1997)	USA	44 (age <15)	17/44 (39)	3/4 (75)	5/13 (39)
Europe & USA		**396**	**167/395 (42)**	**68/82 (83)**	**67/164 (41)**
Armstrong et al (1993)	Saudi Arabia	8 (age <15)	7/8 (88)	5/5 (100)	2/2 (100)
Li et al (1994)	China	82 (age <15)	67/82 (82)	NA (91)*	NA*
Weinreb et al (1996)	Egypt	14 (age <16)	7/14 (50)	4/7 (57)	2/3 (67)
	Jordan	16 (age <16)	8/16 (50)	5/8 (63)	2/7 (29)
	UAE	10 (age <16)	6/10 (60)	3/4 (75)	NA*
	Iran	8 (age <16)	7/8 (88)	NA*	NA*
Benharroch et al (1997)	Israel	11 (age 0-14)	8/11 (73)	NA*	NA*
Peh et al (1997)	Malaysia	24 (age <15)	14/15 (93)	8/9 (89)	6/6 (100)
Kusada et al (1998)	Japan	4 (age <15)	4/4 (100)	3/3 (100)	1/1 (100)
Liu et al (1998)	Japan	6 (age <15)	4/6 (67)	NA*	NA*
Asia & Middle East		**183**	**132/174 (76)**	**28/36 (78)**	**13/19 (68)**
Weinreb et al (1996)	S. Africa	18 (age <16)	9/18 (50)	6/10 (60)	2/4 (50)
	Kenya	56 (age <16)	56/56 (100)	18/18 (100)	26/26 (100)
Kusada et al (1998)	Kenya	18 (age <15)	16/18 (89)	12/13 (92)	2/3 (67)
Africa		**92**	**81/92 (88)**	**36/41 (88)**	**30/33 (91)**
Ambinder et al (1993)	Honduras	11 (age <15)	11/11 (100)	6/6 (100)	3/3 (100)
Armstrong et al (1993)	Brazil	25 (age <15)	18/25 (72)	10/12 (83)	7/10 (70)
Precidio et al (1995)	Argentina	29 (age 3-15)	12/29 (41)	10/15 (67)	0/7 (0)
Weinreb et al (1996)	Costa Rica	42 (age <16)	34/42 (81)	10/10 (100)	19/24 (79)
Razzouk et al (1997)	Brazil	26 (med. age 9)	15/26 (58)	NA*	NA*
South America		**133**	**90/133 (68)**	**36/43 (84)**	**29/44 (66)**
TOTAL		**804**	**470/794 (59)**	**168/202 (83)**	**1396/260 (53)**

* not in total.

reflect the small number of paediatric and older cases likely to be present in non-selected series (Jarrett, 1992).

The study results shown in Table 1.11 and the results of Jarrett et al (1991), who found no association of EBV with sub-type after adjustment for age, support the hypothesis that HD in different age groups may have different aetiologies. EBV may have a pathogenic role in HD in children and older age groups but not in most young adults or HD not associated with EBV. If the virus were an innocuous passenger in HD it would be expected that the proportion of EBV-associated cases would be lower in childhood than adult disease. In fact this is not the case, even amongst tumours presenting in the first 10 years of life. This suggests that EBV plays an aetiological role in the development of childhood HD. De novo infection may be important in the paediatric group, a reactivation of latent infection, possibly as a result of decreasing T-cell immunity, more likely in older cases (Jarrett et al, 1991). Patients who are infected with HIV are also at an increased risk of developing EBV-associated HD and this lends some support to the suggestion that an impairment of T-cell function may precede EBV-associated HD (see HD and the immune system, p.40).

EBV-association and geography:

The proportion of EBV-associated HD can be related to geography. In the meta-analysis by Glaser et al (1997) country of residence was identified as a statistically significant predictor of the proportion of EBV-associated cases of HD. EBV has been identified in: 40-94% of cases from South America; 30-76% of cases from Asia and the Middle East; 34-92% from Africa. This compares to 19-50% from Europe and North America. Averaged across all studies in each region Europe and North America had the lowest proportion of EBV-associated HD (32%), followed by Asia and the Middle East (54%), South America (67%), and Africa (73%) (Table 1.10). The proportion of EBV-associated HD is high for all histologic subtypes in Africa and South America. In contrast the proportion of EBV-

associated HD is highest for the MC subtype in Europe/North America and Asia/Middle East. However, these geographical relationships could be due to factors such as the age and histological subtypes of the cases.

The larger proportion of EBV-associated HD from developing countries could be due to lower socio-economic status which may be associated with a chronic immunosuppressed state predisposing to viral infections (Chang et al, 1993). In Glaser et al's (1997) meta-analysis a variable for country SES based on world rank of GNP was used and found to significantly predict the proportion of EBV-associated HD ($p \leq 0.001$) but not after adjustment for age, subtype, sex, and ethnicity. Razzouk et al (1997) also found, after only adjusting for histologic subtype and age, the association between EBV and HD was independent of geographical location. Thus, all the potential variables explaining the proportion of EBV-associated HD should be entered into the predictive model together, especially age and subtype before a relationship is confirmed.

EBV Conclusion:

EBV has been implicated in the aetiology of HD ever since it was found that HD risk was increased following IM and that people with HD had raised anti-EBV titres compared with healthy controls. More direct evidence has come from numerous studies demonstrating the presence of EBV viral genomes in the HRS cells. However, there is a lack of correlation between raised antibody titres to EBV and EBV-associated HD (Brousset et al, 1991; Alexander et al, 1995b; Enblad et al, 1997).

The biological mechanism by which EBV could cause HD is not known. EBV may cause the activation of an unidentified oncogene; down regulate cellular genes which could be used to circumvent immunosurveillance; or induce growth advantages to neoplastic cells (Bai et al, 1994).

The majority of HD cases aged <15 and >50 years are EBV-associated, whereas the minority of cases aged 15-34 years are. The available data also suggest that MC cases are more likely to be EBV-associated than NS. There is now strong evidence to suggest that EBV is causally involved in a proportion of cases of HD (IARC, 1997). The variation of EBV-association by age and subtype is consistent with the multiple aetiology model first proposed by MacMahon (1966). Substantially increased risk of EBV-associated HD in children compared to young adults suggest that the timing of infection greatly affects the association of EBV with HD (Glaser et al, 1997). The variation in the magnitude of the age effect with regional economic level also points to the importance of socio-economic conditions in predicting association of EBV and HD. However, analyses that investigate this question should be performed with adjustment for the effects of age and histological subtype, as they are likely to be confounding variables. HD may be a rare response to underlying EBV infection with the nature of the response depending on other factors correlated with the degree of affluence of the population.

HD and Infection: Conclusion:

There is now a great deal of evidence supporting an infectious aetiology of HD. This evidence includes clustering and seasonal presentation of HD and applies particularly to the young adult age group and NS subtypes. However, EBV appears to play a role in children and older adults and the MC subtype. Other agents, possibly other viruses, may be involved in young adults (Jarrett, 1992). The late-host-response model may hold for the virus that is involved in the young adult form of HD. The identification of a specific infection as a cause of haematological malignancies in young adults could pave the way for effective preventative measures (Lehtinen & Lehtinen, 1998).

1.7 Hodgkin's Disease and the immune system:

HD patients show a severe impairment in their cellular immune responses (Slivnick et al, 1990). There is some evidence that immune deficiency is present prior to HD diagnosis (Staal et al, 1989; Bjorkholm et al, 1990) and, unlike NHL, immune deficiency persists even after cure (Staal et al, 1989). Altered cell-mediated immunity is characteristic of HD. It is possible that chronic antigenic stimulation together with defective regulation of the immune response may cause excessive lymphoid proliferation, which subsequently undergo genetic mutation, either occurring spontaneously or induced by oncogenic viruses or drugs and leading to the development of a lymphoma (Penn, 1981). If these assertions are true people who are immuno-suppressed would have an elevated risk of HD.

This section on HD and the immune system will be quite detailed as it forms the background to chapter 9 on physical activity and risk of HD.

HD in patients with HIV:

San Francisco County has a high prevalence of HIV infection. Medeiros & Greiner (1995) used SEER data to investigate the incidence of HD in this area. Age-adjusted incidence rates of HD in males increased from 5.0 per 10^5 1973-77 to 6.4 per 10^5 in 1983-87. Over the same time period incidence in females fell from 3.5 to 2.9. The percentage distribution of HD in females in San Francisco was similar to the distribution in all SEER regions. The distribution in men was skewed towards an increase in MC and miscellaneous HD while NS, LP, and LD remained stable. HD in HIV-infected people is usually MC or hard to classify and more common in males.

Several cohort studies have been performed which show the incidence of HD to be increased in the HIV-infected population, although not nearly as much as Kaposi's sarcoma or NHL (Table 1.12).

Table 1.12: Cohorts investigating the changing incidence of HD associated with HIV.

Study	Population	HD cases	SMR/SIR	Incidence rate
Hessol et al (1992)	6704 homosexual men in SF matched with N. California cancer registry data. Followed 1978-89	8	SMR=5 (2.0-10.3)	19.3/ 10^5 PY (SEER rates 5.3/10^5 1973, 3.99/10^5 1980)
Rabkin et al (1992)	1701 haemophiliacs (1065 HIV +ve) followed 12 yr.	2 in HIV +ve 1 in HIV –ve	SIR=6.6 (+ve) SIR=6.2 (–ve)	Not Given
Rabkin et al (1993)	Women aged 20-49 (at risk of HIV) linked to NY & NJ cancer and AIDS registries 1970s-88.	Not Given	Not Given	white-1976-78 4.2/10^5, 1987-88 4.1/10^5
Ragni et al (1993)	3041 haemophiliacs followed 1978-89	0	NA	NA
Reynolds et al (1993)	San Francisco popln based AIDS and cancer registries linked 1980-87.	16	SIR=8.8 (5-14.3) (SIR 1986-87 18.3)	Not Given
Rabkin & Yellin (1994)	Popln based cohort with high prevalence of HIV infection (83000 never married men aged SF 25-54). 1.39 million PY follow-up	100	SIR=2.0 (1.3-3.0) 1988-90	1973-79 cohort 3.9/10^5 vs. SEER 6/10^5. 1988-90 cohort 4.3/10^5 vs. SEER 8.8/10^5.
Lyter et al (1996)	5579 homosexual men in Pittsburgh multi centre AIDS cohort. Followed 1983 onwards	2	SIR HIV +ve=6.7	85/10^5 PY in HIV +ve (4.3/10^5 in male gen. popln)
Koblin et al (1996)	15565 homosexual men in NY and SF 1978-90	18	SIR=2.5 (1.5-3.9)	SF HIV +ve rate 23.4/10^5
Serraino et al (1997)	Italian HIV sero-conversion group. 1255 (906 m, 349 f) aged 20-49. 7075.4 PY follow-up	3	Total SIR=37.9 (7.8-110.6) Male SIR=50.9 (10.5-148.6) Female SIR=0.0	Not Given
Grulich et al (1997)	3616 AIDS cases up to Oct. 1995 matched with NSW Cancer registry.	10	RR HIV+ve=8.5 (4.1-16)	Not Given
Lacoste et al (1998)	French Aquitane hospital based cohort of 3897 adults. 13696 PY follow-up.	8	Total SIR=12.6 (5.4-24.8) Male SIR=17.4 (7.5-34.4)	58.5/10^5 PY (17.9-98.9)
Geodert et al (1998)	98336 AIDS cases age <70 in 9 regions of USA & Puerto Rico.	56 pre-AIDS 72 AIDS years 13 post-AIDS	RR=7.6 (4.1-13.1) after post-AIDS	Not Given
Franceshi et al (1998)	6067 AIDS cases linked to 13 Italian cancer registries until 1992	11	SIR=8.9 (4.4-16.0)	Not Given

Of thirteen cohorts studied only two (Rabkin et al, 1993; Ragni et al, 1993) have failed to show a statistically significant increase in HD in HIV-infected populations, and both of these were performed earlier in time and had shorter follow-up. Where calculated, incidence rates of HD in HIV-infected populations range from $6/10^5$ PY to $85/10^5$ PY. Goedert et al (1998) compared the risk of HD in the pre- and post-AIDS periods and found that risk had increased significantly (p-trend<0.0001). This assertion is supported by the work of other authors who have found the risk of HD to be increased in later time periods compared with earlier (Reynolds et al, 1993; Rabkin & Yellin, 1994; Koblin et al, 1996; Franceshi et al, 1998).

However, some of the HD cases in the HIV-infected population may be misdiagnosed B-cell lymphomas (Delabie & De Wolf-Peters, 1992). Thus, the apparent increase in HD may be due to pathological misclassification of NHLs, which have a greatly increased incidence in the HIV-infected population. False positive diagnosis of HD which is truly NHL does occur but this error is small in younger patients (aged <55 years) diagnosed in the 1980s (Glaser & Schwartz, 1990). Alternatively the increased incidence of HD in HIV infected cohorts could, at least partly, be independent of HIV i.e. HD occurs most frequently in young adults in the USA, the same age group for which HIV infection is most common. There is evidence that HD incidence is increasing in the young adult population but this is generally the NS subtype, not MC. Neither of these potential problems appear able to explain the increase in HD.

Evidence from cohort studies suggests that incidence of HD in the HIV-infected population is elevated. The increase is statistically significant, but the statistics are based on small numbers of cases and short time periods. The slight differences in the results of published studies may be due to differences in the populations studied and the risk groups analysed e.g. male vs. female, homosexual vs. haemophiliac (Diehl & Tesch, 1995).

Presentation of HD in patients with HIV:

There are many case reports of HD in HIV infected people. These range from single case reports at the beginning of the HIV epidemic (Robert & Schneidermann, 1984; Schieb & Seigel, 1985; Di Carlo et al, 1986) to large case series which have been reported more recently (Errante et al, 1994; Tirelli et al, 1995).

HD in HIV infected people has a particular presentation when compared with HD occurring in immuno-competent patients. There is an increased frequency of more advanced anatomical stages (Ann Arbor stages III and IV) (68% of 398 published cases with stage given). A predominance of MC subtype characterises HD in HIV patients (59% of 442 published cases with subtype given). HIV-infected patients have a more frequent presence of B symptoms and mediastinal involvement is unusual.

HD in HIV infected people is more frequently EBV-associated. 298 cases with details of EBV status have been published of which 243 (87%) are EBV-associated. Less than 45% of HD in the general population is EBV-associated (Table 1.10). Uccini et al (1990) found significantly more HD in HIV infected people to be related to positive EBV status than HD not associated with HIV ($p<0.01$).

HD in HIV infected people: Conclusion:

HD is still not recognised as an HIV-related malignancy (CDC 1992). However, the results of cohort studies suggest the incidence of HD is increased in HIV infected populations. Also, due to the profound immunologic disregulation that constitutes its background, HD in HIV infected people presents with morphologic, immunologic, pathogenetic, and clinical features that are characteristic and different from HD occurring in the general population (Ioachim et al, 1991). An association between HD and HIV seems to be well established, although with an SIR much lower than that for NHL (IARC, 1996).

HD following transplantation:

The existence of a relationship between immune deficiency and neoplasia had been recognised more than 10 years before the AIDS epidemic (Jaffe et al, 1983; Curran et al, 1984). The incidence of malignancies is markedly increased following transplantation; malignancies develop in 6% of renal transplant recipients, which is 100 times more than in the general population (Penn, 1978). Immunosuppression and cytotoxic therapy have been implicated as aetiologic agents. The majority of tumours developing in immune deficient individuals are lymphomas. NHLs in immune deficient patients differ from those in immune-competent patients in relation to their location, histology, natural history and response to treatment. This observation may also apply to HD. The development of HD has relatively rarely been reported in transplant recipients.

Several studies have analysed cohorts of patients who have undergone specific transplants e.g. heart (Weintraub & Warnke, 1982), bone marrow (Witherspoon et al, 1989), kidney (Kinlen et al, 1979), liver (Raymond et al, 1995; Hoover & Fraumeni, 1973). These studies combined consist of over 9500 transplant recipients in which only one case of HD has occurred. These results have led Kinlen (1996) to conclude that there is no increased risk of HD following transplant. B cell hyperplasia is not a feature of iatrogenic immunosuppression, as it is for HIV infection, which may explain why certain types of lymphoma e.g. HD and Burkitt's lymphoma, are relatively frequent in AIDS patients but not in transplant recipients (IARC, 1996).

However, HD does appear to present differently post-transplant. Case reports of thirty patients have been published in the literature (Doyle et al, 1983; Jarcard et al, 1994; Garnier et al, 1996; Cerilli et al, 1977; Sterling et al, 1974; Moreau et al, 1996; Goyal et al, 1996; Bedrossian et al, 1995; Bierman et al, 1996; Hood et al, 1996; Oldhafer et al, 1989). 57% were MC subtype, while NS was quite uncommon. Evidence for EBV involvement in

HD following transplantation has been found in 90% cases, a proportion similar to that in HIV-associated HD and much higher than in spontaneous HD.

Although HD is rare after transplantation, its association with EBV and the severity of the disease in these patients argues for a role of EBV infection and immunosuppression in the development of the disease (Garnier et al, 1996), although the relation to underlying immunosuppression may not be a strong one.

HD and Immunodeficiency diseases:

The risk of developing a malignant tumour is 4% in patients with genetic immune deficiencies, which represents an incidence 1000 times greater than that recorded in the age-matched general population (Curran et al, 1984; Fauci et al, 1984).

The majority of cancers developing in patients with congenital immunity are NHL. However, HD has an increased incidence among certain immunodeficiency states, particularly Ataxia Telangiectasia (AT), Wiskott-Aldrich Syndrome (WAS), and Bloom's syndrome (Mueller et al, 1995). In an international series of 12 cases from the Immunodeficiency Cancer Registry 9 (75%) were MC or LD subtype and only 1 (8%) NS (Robinson et al, 1987). Since each immune deficiency syndrome has its own unique genetic basis it seems plausible that the excess of HD is related to impaired immunity, with consequent enhanced susceptibility to EBV, rather than to the diverse range of underlying defects (Stiller, 1998).

Tonsillectomy:

The common site of early stage HD is in the region of the lymph nodes that drain the pharyngeal tonsil. It is well known that the tonsils can act as filter barriers to infective agents or can even facilitate the passage of some diseases (Vianna et al, 1971).

'Risk' of tonsillectomy is strongly related to childhood social class. Tonsillectomy is performed more frequently on children from more affluent homes (Wolman, 1956) and

medical practice. Using sibling controls or controlling for socio-economic factors e.g. place of residence in the analysis can negate the effect of socio-economic status.

Four published English language studies have not taken account of social class (Ruuskanen et al, 1971; Silingardi et al, 1982; Abramson et al, 1978; Bonelli et al, 1990). Other studies of HD following tonsillectomy have taken account of social class (Table 1.13). Of these, four found a statistically significant increase in risk of HD after tonsillectomy (Johnson & Johnson, 1972; Vianna et al, 1971 & 1974; Kirschoff et al, 1980). (The first three of these studies only analysed young adult cases.) Others found a non- significant increase in risk (Gutensohn et al, 1975; Newell et al, 1973; Hardell & Bengtsson, 1983; Anderson & Isager, 1978; Liaw et al, 1997). Mueller et al (1987), Serraino et al (1991), and Zwitter et al (1996) did not observe an increase at all.

Ablation of the tonsils may influence the development of HD in only a minority of those affected. Bock et al (1994) investigated selected cellular and humoral immune system parameters in 160 children who had undergone tonsillectomy and 302 age-matched controls and concluded that any changes were clinically insignificant. The results of the studies controlling for social class and the presence of a plausible biologic mechanism mean the role of tonsillectomy cannot be discounted, even though its association with HD is not strong or consistent.

Appendectomy:

The lymphoid tissue of the appendix, illium and colon may have a protective effect against viral antigens (Burnet, 1959). Removal of the appendix may remove a barrier to infection. Appendicitis may also be important because it could have an infectious aetiology (Andersson et al, 1995). If this is true appendicitis and, following this, appendectomy could be a proxy for late age at first infection, increasing risk of young adult HD. As with tonsillectomy rates of appendectomy can be related to social class.

Table 1.13. Summary of case-control studies investigating risk of HD following tonsillectomy, adjusted for SES.

Study	Numbers	Control Source	Comparison	Risk Estimate	Significance
Vianna et al (1971) (USA)	Matched c/c study 109 cases (age ≤40) & hospital. controls (age-, sex-, race-, residence-matched)	Hospital	Tonsillectomy yes v. no	RR = 2.9	$p < 0.01$
Johnson et al (1972) (USA)	Matched c/c study 174 cases (age 15-44) & controls	Siblings	Tonsillectomy yes v. no	RR = 2.0 (1.4-2.8)	$p<0.05$
Vianna et al (1974) (USA)	C/C study 95 cases (age <40) & controls	Siblings	Tonsillectomy yes v. no (all sibs) Tonsillectomy yes v. no (same sex sibs)	RR=2.0 (1.1-3.6) RR = 3.6 (1.3-9.7)	$p<0.05$ $p<0.02$
Gutensohn et al (1975) (USA)	C/C study 136 cases (age 15-44) & 315 sibling controls	Siblings	Tonsillectomy yes v. no	RR = 1.4 (0.8-2.6)	NS
Gledovic et al (1991) (Yugoslavia)	Matched c/c study, 113 cases (age 15-39) & 226 controls (age-, sex-, residence-matched	Neighbourhood/ Hospital	Tonsillectomy yes v. no	NA	NS
Zwitter et al (1996) (Slovenia)	Matched c/c study 192 female cases (aged 17-50) & 496 controls (age- & residence-matched)	General Popln	Tonsillectomy yes v. no	OR = 0.68 (0.4-1.01)	NS
Newell et al (1973) (USA)	Matched c/c study 176 cases (aged ≥5) & controls (sex-, age-, social class)	Hospital	Tonsillectomy yes v. no	RR = 1.2	NS
Kirchoff et al (1980) (Brazil)	Matched c/c study 70 cases (age <79) & 70 controls (age- & sex-matched)	Siblings	Tonsillectomy yes v. no (17% cases vs. 7% sibling controls)	RR=2.5 (1.0-6.0)	$p<0.05$
Hardell & Bengtsson (1983) (Sweden)	Matched c/c study 60 male cases (age 25-85) & 117 controls (age-, sex-, residence-matched)	General Popln	Tonsillectomy yes v. no	RR = 2.7 (0.6-11.6)	NS
Mueller et al (1987) (USA)	C/C study 563 cases (age ≥15) & 688 controls	Siblings	Tonsillectomy yes vs. no (age 15-39) Tonsillectomy yes vs. no (age 40-54)	RR=1.0 (0.72-1.4) RR=1.5 (0.07-3.3)	NS NS
Anderson & Isager (1978) (Denmark)	Nested c/c study 65 cases (age 15-45) & 189 controls (age-, sex- & ses-matched)	General Popln	Tonsillectomy yes v. no	RR=1.3 (0.5-3.3)	NS

NS (not statistically significant) = $p>0.10$

Many studies that have investigated the relationship between appendectomy and HD have not taken account of social class (Hyams & Wynder, 1968; Bierman, 1968; Ruuskanen et al, 1971; Abramson et al, 1978; Silingardi et al, 1982). The results of studies that did control for social class are in Table 1.14. None of these show a statistically significant increase in risk of HD following appendectomy. If appendectomy does relate to HD, it is not a major factor.

HD and the immune system: conclusion:

The results of analyses of HIV-infected cohorts suggest that people with this virus have a higher incidence of HD but there is little evidence of an increase associated with organ transplant or immuno-deficiency diseases. There is some evidence of an increased risk following removal of the tonsils but not the appendix. Further epidemiological studies will be needed to document a definite change in the incidence of HD among immunosuppressed patients (Knopf & Locker, 1995).

Table 1.14. Summary of studies investigating appendectomy & HD risk, adjusted for SES.

Study Name	Numbers	Comparison	Risk Estimate	Significance
Gledovic et al (1991) (Yugoslavia)	Matched c/c study, 113 cases (age 15-39) & 226 controls (age-, sex-, residence-matched	Appendectomy yes vs. no	NA	NS
Zwitter et al (1996) (Slovenia)	Matched c/c study 192 female cases (aged 17-50) & 496 controls (age- & residence-matched)	Appendectomy yes (10.1% cases, 11.6% controls) vs. no	OR = 0.78 (0.43-1.41)	NS
Newell et al (1973) (USA)	Matched c/c study 176 cases (aged ≥5) & controls (sex-, age-, social class-matched)	Appendectomy yes vs. no	RR = 1.2	NS
Hardell & Bengtsson (1983) (Sweden)	Matched c/c study 60 male cases (age 25-85) & 117 controls (age-, sex-, residence-matched)	Appendectomy yes vs. no	RR = 0.5 (0.2-1.3)	NS
Anderson & Isager (1978) (Denmark)	Nested c/c study 65 cases (age 15-45) & 189 controls (age-, sex- & ses-matched)	Appendectomy yes vs. no	0 cases with appendix removed	p=0.09

NS (not statistically significant) = $p>0.10$

1.8 Hormonal factors and HD:

La Vecchia et al (1991) observed a steady rise in the mortality rate for males aged 30-50 years compared with a flattening of the female mortality (and incidence) rates in the same age group. Incidence data have shown the peak male: female incidence ratio to be at the time when women have children (Muir et al, 1987). Alternatively the decline in female incidence that develops in young adulthood and becomes gradually more pronounced by middle age may be explained by sex differences in occupational exposures. However, for the few occupations associated with HD the attributable risk would be too low to explain the observed sex difference (Glaser, 1994).

As for HD and the immune system this section on hormonal factors and HD will be quite detailed as it also forms part of the background to chapter 9 on physical activity and risk of HD.

Parity

Although endocrinologic factors have been observed in several cancers of non-reproductive sites, they have not been considered for HD until recently (Glaser, 1994; Kravdal & Hansen, 1993). Chen et al (1997a) investigated the Connecticut Tumour Registry 1935-92 and found that between 1970 and 1992 the trend in HD incidence differed in males and females. In males HD incidence rose 2% every 5 years, whereas in females it rose at 11% and was expected to increase at this rate in the future. These authors obtained fertility data for Connecticut women aged 20-44 years for 1940-84. More women born after 1940 were having their first baby aged 30-39 years and fewer aged 20-24 years. Also the total number of births per mother had fallen since 1960. These data support the hypothesis of an increased risk of HD with lower parity and later age at first pregnancy. The results of studies looking at this topic are in Table 1.15.

Table 1.15. Summary of studies investigating parity and HD risk.

Study	Numbers	Comparison	Risk Estimate	Significance
Zwitter et al (1996) (Slovenia)	Matched c/c study 192 female cases (aged 17-50) & 496 controls (age- & residence-matched)	2 vs. 0 children 3 vs. 0 children	OR = 0.74 (0.4-1.3) OR = 0.83 (0.3-1.9)	NS NS
Abramson et al (1978) (Israel)	Matched c/c study 343 cases (age ≥17) & population controls (sex-, age- & origin-matched)	0 vs. 3 children 0 vs. 4 children 0 vs. 5 children	RR = 1.9 RR = 2.7 RR = 4.7	p = 0.022 p = 0.014 p = 0.001
Olsson et al (1990) (Sweden)	38 cases (age 17-85) & hospital controls	No. of cases with ≥ pregnancy vs. no. of controls	27% v. 23%	Not given
Franceschi et al (1991) (Italy)	C/C study 152 cases (age 15-77) & 613 hospital controls	1-2 children vs. 0 3+ children vs. 0	OR = 1.04 (0.5-2.3) OR = 0.77 (0.3-1.9)	NS NS
Zahm et al (1995) (USA)	C/C study 70 cases (35 m & 35 f) & 1432 popln control	Women with 3+ children vs. 0	OR = 0.4 (0.1-1.3)	NS
Tavani et al (1997) (Italy)	C/C study 68 cases (age 17-79) & 448 hospital controls	3+ pregnancies vs. 0 1+ abortion vs. 0 3 births vs. 0	OR=0.6 (0.3-1.4) OR=0.5 (0.2-1.1) OR=0.9 (0.4-2.3)	NS NS NS
Lambe et al (1998) (Sweden)	Record linkage study cancer & fertility registry. 917 HD born 1925-72 and 5 age-matched controls	ever parous vs. never	OR↓ (number of children in parous women unrelated to risk)	NS
Miller et al (1980) (Canada)	Cohort 11127 ever-married women	Low vs. high parity	RR = 1.32	p > 0.05
Kravdal & Hansen (1993) (Norway)	441 HD in cohort 1.3 million (f) aged 15-56	1 child vs. 0 2 children vs. 0 3+ children vs. 0	RI = 0.74 RI = 0.64 RI = 0.46	p < 0.05 p < 0.05 p < 0.05
Kravdal & Hansen (1996) (Norway)	Nested c/c study. 382 cases (age 15-56) & controls	1 child vs. 0 2 children vs. 0 3+ children vs. 0	OR=0.74 OR=0.64 OR=0.46	p<0.05 p<0.05 p<0.05
	Nested c/c study. 39 cases (age 40-56) & controls	1 child vs. 0 2 children vs. 0 3+ children vs. 0	OR = 0.86 OR = 0.59 OR = 0.53	NS NS NS

NS (not statistically significant) = p>0.10

Only two studies (Abramson et al, 1978; Kravdal & Hansen, 1993) found childbirth to be statistically significantly protective but all the other studies in Table 1.15 are suggestive of a protective effect. Abramson et al (1978) concluded that, "the male preponderance usually seen in middle and late adult life may be partly due to the protection afforded to women by their prior reproductive activities." This is supported by the fact that studies that have looked at risk of HD in men by parity have found no relationship e.g. Men: RR 1 child vs. 0 1.06, 2 vs. 0 1.04, ≥3 vs. 0 1.03; Women 0.74, 0.64, 0.46 respectively (Kravdal & Hansen, 1993).

A pregnancy is characterised by a surge of sex hormone secretion; high and low parity women have different histories of exposures to this hormonal environment. The mother's immune system is also altered during pregnancy with cell mediated immunity generally being depressed (Glaser, 1994). There is ample evidence that immune function is responsive to sex hormones.

Effects of social class:

Parity may be associated with HD risk because it is a correlate of social class, a demonstrated risk factor for HD. Fewer and later pregnancies would be predicted by the elevated adult social status of the women at higher risk of HD. An increase in HD risk associated with lower parity might reflect social class differences. Reported parity specific OR's may not be adjusted for identified social class confounders.

Zahm et al (1995) found the significant effect of childbirth was limited to women of higher socio-economic status (high SES OR ≥3 children vs. none = 0.1 (0.002-0.9)). However, Kravdal & Hansen (1996) adjusted for SES using either the case's or their father's educational level and these did not modify the findings of a statistically significant decreased risk of HD with increasing parity. Other studies (Franceshi et al, 1991; Zwitter et al, 1996; Tavani et al, 1997) have adjusted for SES variables and shown a non-significant protective effect for parity.

Age at first birth:

Observations of the data from the Connecticut cancer registry by Chen et al (1997a) suggest that women with a later age at first pregnancy have a greater risk of HD. This has been supported by the results of some case/control studies (Kravdal & Hansen, 1993. RR of first birth >25 vs. <21 yrs 1.22; Miller et al, 1980. RR of first birth >25 vs. <25 yrs 1.36 ($p<0.05$)). Lambe et al (1998) found some evidence of an increased risk of HD with late age at first birth in women aged less than 45 years at diagnosis. However, other studies have not detected this age effect (Franceschi et al, 1991; Zwitter et al, 1996).

Hormonal Factors and HD: Conclusion:

If parity were protective through a hormonal or immunologic process related to pregnancy, it would seem likely to affect women, regardless of social class but not men. If parity is protective for reasons of social class the effect should be seen in both genders. The evidence is for greater parity to decrease the risk of HD in women only, although not all results are statistically significant. Hormonal factors appear to be involved in the pathogenesis of HD, possibly operating through an effect on the immune system. The impact of the immune system has been supported by Tavani et al (1997). These authors suggest that since the pattern of risk of HD was similar for full-term pregnancies and for abortions that events early in pregnancy, including immuno-related events, rather than exposure to sex hormones (whose levels are higher at the end of pregnancy) are the likely mechanism.

1.9 Familial Hodgkin's Disease and inherited susceptibility:

There have been many reports of the multiple occurrence of HD within the same family. This familial aggregation may be as the result of sharing a common environment in childhood or of genetic factors (Higginson et al, 1992). Studies looking at familial aggregation of HD have found a range of risks. However, the best evidence for a genetic basis for HD comes from a study of twins by Mack et al (1995).

Mack et al (1995) investigated 432 sets of twins found through newspaper adverts. The number of cases of HD observed aged <50 years in the healthy monozygotic and dizygotic twins of patients with HD was compared with the number expected from national age-specific incidence rates. They observed that 0 of 187 pairs of dizygotic twins became concordant for HD, whereas 10 of 179 pairs of monozygotic twins did after 14 years follow up. This result compared with 0.1 (monozygotic) and 0.1 (dizygotic) expected cases in unaffected twins. Thus, monozygotic twins of patients with HD had a greatly increased risk (99-fold), whereas no increase in risk for dizygotic twins was seen. This difference in risk between monozygotic and dizygotic twins could be due to the effect of sharing the whole genome of the twin with HD rather than just half of it. These results do not appear to be an artefact of ascertainment or differential survival. The magnitude of the effect found in this study suggests a fundamental genetic influence on pathophysiology.

Ferraris et al (1997) reviewed twenty-eight articles published 1972-95 on familial HD. They also analysed 328 patients from 18 of the papers. The authors found a significant difference between sporadic and familial HD in the age at diagnosis. Familial HD only had evidence of one peak aged 15-34 years, whereas the sporadic HD had the characteristic bimodal appearance. This suggests familial HD could be linked with certain forms of HD. Mack et al (1995) observed that the majority of reported sibships with multiple cases of HD consist solely, or include multiple cases of, NS subtype.

To examine possible genetic markers of susceptibility to HD several authors have looked at human leukocyte antigens (HLA) in patients with the disease. HLA genes are of interest because resistance and susceptibility to infection are under their potential control. An association between HD and an HLA allele could be construed as providing support for an infectious cause of HD (Taylor et al, 1996). Chakravarti et al (1986) calculated that a recessive susceptibility gene linked to the HLA complex was responsible for 60% of HD in families. The remaining 40% could be due to other familial factors and/or environmental factors. There have been many investigations of association between specific HLA types and risk of HD e.g. Svejgaard et al (1975) found an increase in the frequency of HLA-A1, B5, B8, and B18 in an international series of 1500 HD cases. More recently authors have found a significant association between HD and HLA-DPB1*0301 (Oza et al, 1994; Taylor et al, 1996). The susceptibility might result from a particular DNA base sequence common to several alleles (Mack et al, 1995). There appears to be an association between HLA class II type and risk of HD but at present it is unlikely that any single allele or haplotype is responsible for susceptibility to HD.

The risk of HD developing in a first degree relative of a case of HD is increased but the precise risk is uncertain. Relative risks in a first degree relative range widely, from 2 (Abramson et al, 1977) to 17.1 (Grufferman et al, 1982; Bernard et al, 1987). Risk in identical twins appears to be increased 100-fold. However, familial HD is relatively rare, estimated at only 4.5% in published series (Ferraris et al, 1997). This may be due to the small family size associated with upper socio-economic groups at higher risk of developing HD, in part to the putative recessive pattern (Levine et al, 1995), and also to low penetrance. HLA association of HD and an increased risk in monozygotic twins constitute persuasive evidence of the role of genetic susceptibility for HD.

1.10 Occupation:

Most investigations on occupational risk of HD have been performed in three broad occupational categories: 1) employees in woodworking and wood related industries; 2) agricultural workers; 3) occupational groups at risk of developing HD as an infectious disease e.g. doctors, teachers. However, as occupational risk factors will not be investigated in later chapters only a brief discussion of occupational risks of HD for doctors and teachers will be given. A discussion of the findings for woodworking and agricultural occupation risks for HD can be found in Appendix A.

Doctors:

Vianna et al (1974) were the first to suggest that, if HD is infectious, physicians would be at an increased risk because they more frequently share a common environment with HD cases. They identified male physicians aged >25 years who died from HD in New York State 1960-72. The mortality rate in physicians was compared with that for Upstate New York and the four counties with the highest median income. There was a significant difference between the mortality rate in physicians and Upstate New York ($p<0.01$) and the four counties ($p<0.01$). Physicians had a RR of 1.8 of dying from HD.

Few other studies have investigated this relationship and none have reproduced these results (Smith et al, 1974; Matanoski et al, 1975; Grufferman et al, 1976).

Teachers:

Vianna & Polan (1973) found an increased risk of HD mortality for school teachers leading Milham (1974b) to look at the death records for all male residents of Washington State aged over 20 years dying 1950-71. The PMR of HD for schoolteachers of all ages was 246 (100 = all occupations) and for ages 30-49 years it was 315. The observed and expected number of deaths differed significantly (12 v. 4.9, $p<0.005$). However, schoolteachers may be at less risk of death from other causes e.g. murder, accidents, than other occupational groups

(Bahn, 1974). If the contribution of other causes of death to the all-cause death rate for teachers were small, the proportion of deaths due to HD would be relatively large.

In contrast to Milham (1974b) Hoover (1974) found a RR for teachers of 1.5, which is compatible with a social class effect. Also, no increased risk was observed in a 15 year incidence survey conducted by Grufferman et al (1976) (30 cases observed vs. 27 expected, RR=1.1, 0.7-1.6). The weight of evidence suggests that teachers as a group are unlikely to be at an increased risk of HD.

Occupational risks for HD: Conclusion:

Most investigations of HD occupational risk factors have been hindered by the rarity of the disease and by the lack of detailed information on work exposures (Grufferman & Delzell, 1984). For the most part, occupational exposure is classified by death certificate, medical records, or self-reported information obtained from questionnaires. There is little epidemiological evidence that adults with potential occupational exposure to HD patients have a significantly increased risk of the disease. If certain occupations are at a higher risk of HD the best candidates are woodworking and agriculture (Appendix A), although a consistently plausible agent has yet to be identified.

1.11 Epidemiology of Hodgkin's Disease: summary:

A number of risk factors have been investigated in relation to HD, the majority of these related to infection and immune function. However, few of these factors have a strong or consistent relationship with HD (Table 1.16). A great deal of work has been performed on the role of EBV in the aetiology of HD and it now appears that this virus plays a role in paediatric and older adult HD as well as the MC subtype. However, in the age group (15-34 years) and subtype (NS) with the most epidemiological evidence of an infectious origin EBV is least frequently seen. Also there has been little work performed on the relationship of individual risk factors to EBV status. The aetiology of HD remains elusive (Grufferman & Delzell, 1984; Glaser, 1994; Takvorian et al, 1993). HD is still the "enigmatic lymphoma" (Glaser et al, 1997).

Table 1.16: Summary of risk factors for HD.

Factor	Evidence	Consistency	Suggested subgroups affected	Biological mechanism
Childhood environment	strong	consistent	age 0-14: low SES ↑ risk age 15-34: high SES ↑ risk	Proxy for no. & timing of childhood infections
EBV in HRS cells	strong	consistent	age 0-14 & >50 MC subtype	Virus with capacity to de-regulate cell growth.
Family history	strong	consistent	age 15-34 NS subtype	Genetic susceptibility
Previous IM	moderate	consistent	age 15-34?	Proxy for delayed exposure to common viruses
HIV	moderate	consistent	MC & LD subtype	Immune suppression.
Parity (low)	moderate	consistent	None specified	Pregnancy alters female hormonal & immune systems
Tonsillectomy	moderate	inconsistent	None specified	Removal of barrier to infection
Occupation	moderate	consistent	None specified	Exposure to wood dust/chemicals. Exposure to pesticides/animal viruses
Transplantation	weak	consistent	None specified	Immune suppression
Appendicitis/ Appendectomy	weak	inconsistent	None specified	Removal of barrier to infection. Appendicitis caused by late exposure to infection?

2. Measurement of Physical Activity

2.1 Introduction:

Physical activity is defined as bodily movement due to skeletal muscle contractions that result in quantifiable energy expenditure (Thompson, 1994). There are three primary components of physical activity that can be varied and may have different effects on carcinogenesis: intensity (work-rate); duration (length per activity bout); frequency (times per week) of the activity (Thompson, 1994).

"The ability to relate physical activity to health depends on accurate, precise and reproducible measures" (Wilson et al, 1986; National Center for Health Statistics, 1989). However, no measurement methods are generally accepted so it has become the custom for investigators to write their own - often with scant regard to validity, and still less to comparability with other studies.

There is some evidence that the risk of HD is affected by physical activity (see chapter 3). It was decided to include a section on physical activity in the questionnaire for the Scotland and Newcastle study of Hodgkin's Disease (SNEHD). I wrote this section of the SNEHD questionnaire after performing a literature review of methods that have been used to measure physical activity. This chapter will summarise those methods found in the literature. Chapter 3 will summarise the potential cancer health benefits/consequences of physical activity.

2.2 Measurement of physical activity by direct monitoring:

Physical activity can be measured directly through behavioural observation, mechanical or electronic devices, or physiologic instruments. These approaches alleviate problems of poor memory and biased self-reporting but are limited by their high costs and

burden on participants. Thus, these methods have only been used in small studies and not for investigations into disease aetiology. However, they can potentially be used to validate self-reports of physical activity (p.70).

2.3 Measurement of physical activity by self-report:

The most commonly used and practical method of estimating the prevalence of physical activity in population studies has been by self-report (Washburn & Montoye, 1986). Respondents may be asked to recall leisure-time activities, occupational activities, or both over a particular time frame. Information obtained from self-report instruments is then converted into estimates of energy expenditure (kilocalories or kilojoules; metabolic equivalents (METs)) or some other summary measure that can be used to classify or rank a person by their activity level.

This chapter will focus on methods for assessing leisure time physical activity. This is because the majority of present-day jobs require little physical activity and it is questionable whether the kind of activity involved in one's occupation is comparable to that during aerobic exercise (Bartram & Wynder, 1989). There are also problems of misclassification inherent in using job titles for exposure assessment. For many women a major component of their activity is household activity rather than occupation. However, very few studies have attempted to include housework in their questionnaires (Freidenreich et al, 1998).

Methods for measuring leisure time physical activity can be divided into three broad groups: global surveys, recall surveys, and quantitative history surveys (Ainsworth et al, 1998). These three groupings are discussed below with examples that show there is great variety within each group.

1) Global Surveys:

Global surveys are the shortest type of physical activity questionnaire comprising of only one to four items.

The majority of global surveys were developed to investigate risk factors for CHD e.g. the New York Health Insurance Plan (NYHIP) questionnaire (Shapiro et al, 1965), the Harvard Alumni Study questionnaire (Paffenbarger et al, 1978), Lipid Research Clinics (Haskell et al, 1980), NHANES I questionnaire (Albanes et al, 1989). This kind of study asks very simple questions about activity: "Do you regularly engage in strenuous exercise or hard physical labour?" (Haskell et al, 1980).

Two of the above questionnaires have been used to investigate cancer risk as well as CHD risk (Paffenbarger et al, 1978; Albanes et al, 1989). Simple global surveys have also been developed specifically for cancer studies (Thune & Lund, 1996; Gerharrdson et al, 1988).

The simplest form of global survey refers simply to episodes of sweating. This method derives from the work of Professor R.A.Bruce (cited in Paffenbarger et al, 1993) who found that the best predictor of treadmill tested $VO_{2\ max}$ in normal subjects was the number of times they engaged in tasks long enough to work up a sweat. A sweating frequency question is easily administered and scored and has been used by La Porte et al (1985), Siconolfi et al (1985), Washburn et al (1990), Newcomb et al (1995), Booth et al (1996), and Min-Lee et al (1997). However, no single question instruments have been validated in a UK population.

The physical activity measures from global surveys provide a general impression of physical activity. Global surveys published in the literature provide good surrogate

measures of participation in vigorous physical activities (Kriska & Caspersen, 1997). Global surveys have been successfully validated by correlation with direct measures of physical fitness (NYHIP questionnaire) and physiological measures (Lipid Research Clinics questionnaire results have been related to higher levels of HDL cholesterol).

The majority of global surveys were developed for the investigation of CHD risk and not cancer. The sensitivity of global survey sweat questions are not sufficiently detailed to record changes in activity and the utility of self-reported sweating may be limited to distinguishing active from inactive subjects in epidemiological surveys (Washburn et al, 1990). Therefore, global surveys may not be detailed enough and of great practical use in cancer epidemiology.

2) Recall Surveys:

Recall surveys generally elicit more detailed information covering time periods of one day to one year.

The original Framingham Study (Kannel et al, 1979) produced a summary physical activity index based on the weighted sum of time the subject reported spending in five types of activity over 24 hours: basal e.g. sleeping; sedentary e.g. sitting or standing; slight e.g. walking; moderate e.g. gardening; heavy e.g. shovelling. Whittemore et al (1995) also used a 24 hour recall of a typical day but divided time spent into only four categories: sleeping, sitting, light or moderate activity, and vigorous activity.

The Tecumseh questionnaire was used to measure physical activity in the Tecumseh Community Health Study (Montoye, 1975). A personal interview lasting 1 to 1 1/2 hours was needed to complete the 36 occupation and 63 leisure time related questions on activity

over the previous year. Each activity was scored based on duration, frequency and intensity of activity. This instrument was later modified by Taylor et al (1978).

Baecke et al (1982) questioned subjects regarding work and leisure time activities in the previous year, subjects responded to a five point scale with descriptions ranging from never (point value 1) to always (point value 5). A sport activity index was then calculated based on intensity, time engaged, and the proportion of the year participating in each activity.

Sandler et al (1995) asked subjects to provide information on up to four sports that they engaged in during the year before diagnosis or a reference date. The intensity of specific activities was rated as low, medium, or high using codes based on a standard reference. For each sport the number of hours per week (time), the number of months of the year (proportion) they engaged in this activity were coded on a five-point scale.

The results of Baecke et al (1982), Montoye (1975), and Taylor et al (1978) questionnaires have all been significantly correlated with physiologic variables (lean body mass in males but not females; lower body fat and blood pressure; HDL cholesterol levels, BMI and heart rate, respectively). However, as with global surveys, recall surveys may not be detailed enough to be of great practical use in cancer epidemiology.

3) Quantitative History Surveys:

The studies mentioned in the previous section have looked at periods of time of a year or less and assumed the activity levels in this period are representative of lifetime activity. However, the assessment of physical activity at a single time to evaluate its association with cancer may be limited (Min-Lee et al, 1991). Such a measure may not adequately reflect activity over the longer term and may be inadequate when changes over

time are of primary interest (Schechtman et al, 1991). Studies are now attempting to produce life-time histories of activity.

Quantitative history surveys have many items that identify the intensity, duration, and frequency of activities performed over a lifetime or in specific periods of life where it is postulated that physical activity may have the most effect.

Bernstein et al (1994) assessed participation (at least 2 hours/week) in physical exercise of 545 female case/control pairs from menarche to breast cancer diagnosis date. These activities included participation as a member of a sports team, including practice; participation in individual sports; time spent at the gym and participation in exercise classes. Gerharrdsson et al (1990) asked separate questions about physical activity during working hours (day- time) and recreational hours (evenings and weekends) to gauge the level of activity in 1950, 1955, and every five years to 1985. Sturgeon et al (1993) elicited information on physical activity for each decade of life from the age of 20-29 to 70-79 yrs.

Alternatively questions pertaining to specific periods of life have been used, especially for female hormonal cancers. McTiernan et al (1996), Chen et al (1997b), and Gammon et al (1998) all emphasised activity around menarche in studies of breast cancer.

In studies of colon cancer lifetime questioning has been performed as it is unknown in which period physical activity may have the greatest effect. Slattery et al (1997) asked study participants to recall their activity pattern two years prior to diagnosis and ten and twenty years previously. White et al (1996) asked about physical activity in a ten year reference period that ended two years prior to diagnosis in a study of male and female colon cancer. Quantitative history surveys also allow the investigator to elicit changes in physical activity over time.

The studies described above probed for information about frequency and duration of activity episodes and some, (Slattery et al, 1997; Gammon et al, 1998), divided activities by intensity. Vigorous activities were "those that make you sweat or get out of breath", while moderate activities are "those activities which are done at a more moderate pace" (Slattery et al, 1997).

Although quantitative history surveys take longer to administer they can take account of seasonal and lifestyle changes in physical activity. However, when quantitative history surveys are validated the correlation between these surveys and direct measures of physical activity are modest (Kriska & Caspersen, 1997). This is not surprising as the activities could have been performed many years prior to direct measurements of physical activity or fitness.

2.4 Classification of Data:

An all-inclusive coding system is crucial to facilitate accurate interpretation of responses for analysis. The survey methods described previously use different methods of classifying data. Physical activity measures in global surveys are generally expressed as ordered category scores indicating higher or lower activity status e.g. sedentary, fairly active, and very active (Gerhardsson et al, 1990). In recall surveys and quantitative history surveys physical activity scores are based on ordinal scales, such as time spent exercising, kilocalories (kcal) expended per day or week in all activities or selected activities, or in other types of units that may reflect the time and intensity of activities (MET-hours).

The MET (metabolic equivalent) is usually the starting basis for calculating a subject's physical activity index. Resting metabolic rate is approximately 1 kcal/kg/hour and

is equivalent to 1 MET. The number of METs associated with a wide range of activities can be estimated from aggregated laboratory and field research. Until recently there was no 'gold standard' list. People made their own MET lists based on previous work with their own additions (Taylor et al, 1978; Wilson et al, 1986; Min-Lee et al, 1992).

Ainsworth et al (1993) derived the energy costs of specific activities from the best available published and unpublished data to form a comprehensive list that most authors use today. However, this classification is primarily based on previously published data and as such may not reflect the exact average energy cost of all physical activities. The MET values of some activities were not derived from actual measurements of oxygen consumption; instead they were estimated from the energy cost of activities having similar movement patterns

The process of ageing illustrates a more general problem with MET scores. As people age their maximal oxygen uptake (a measure of cardiorespiratory fitness) decreases. Activity of a given MET (absolute intensity) therefore requires a greater percentage of their maximal oxygen uptake (relative intensity). A walk of 4 mph may be light exercise for a 20 year old, moderate for a 60 year old, and vigorous for an 80 year old although all would be classed as 4 METs (US Dept Health, 1996, p32).

The use of METs allows different intensities of physical activity to be investigated (bearing the above caveat in mind). For example, McTiernan et al (1996) looked specifically at high intensity exercise (≥7.0 METs) while White et al (1996) divided activities into low, moderate, and high intensity (<4.5, 4.5-5.5, and ≥6.0 METs respectively). However, there are problems of consistency as one author considered high intensity activity to be ≥7.0 METs and one ≥6.0 METs. These studies, therefore, would not be directly comparable. There is no accepted MET definition of each activity level.

Kilocalories of energy expended each week in a particular activity are calculated by multiplying the MET score of that activity by the weight in kilograms of the person by the number of hours the activity is performed. An index of total energy expended per week can be calculated by adding the kilocalorie scores for all activities together. Subjects can then be categorised into levels of kilocalorie expenditure. However, the investigator may not know a subject's body weight, especially if the activity took place in the distant past. Therefore, Min-Lee et al (1992) estimated energy expenditure without the weight parameter. Sports of intensity of <4.5 METs were rated at 5 kcals per minute whilst those of ≥4.5 METs 10 kcals per minute. The energy expenditure thus calculated was very highly correlated with energy expenditure calculated using weight and MET scores (Spearman rank correlation co-efficient 0.99). Errors can arise with methods that take account of body weight because these estimates of energy expenditure may more closely reflect body weight rather than energy expenditure (Ainsworth et al, 1993).

At present the most frequent method of constructing summary scores is to use MET-hours per day or week. These are calculated by multiplying the reported time spent on each activity by its MET score. This method has the advantage of excluding body weight and, thus, avoids confounding energy expenditure by body weight (Giovannucci et al, 1995).

The coding system developed can also allow analysis of changes in activity level over time (if the questionnaire has been written to attempt this or has been administered more than once). Min-Lee et al (1997) administered the same questionnaire 3 years apart and created 4 categories of exercise defined as: 1) inactive at study entry and after 3 years; 2) inactive at entry and then active; 3) active at entry and then inactive; 4) active at both measurements. Besides allowing for the evaluation of physical activity patterns over time a joint assessment can also increase the precision of activity measurement i.e. people who

provide more than one report of being active (or inactive) are likely to be those who are truly active (or inactive). It is also possible that obtaining long-term activity patterns enable the detection of the relevant time period when physical activity may be related to the development of cancer (Ainsworth, 1998).

2.5 Problems with the recall of physical activity:

In the most general sense physical activity is defined by level of caloric expenditure. While this has been useful in heart disease studies it is somewhat like defining diet simply by calorie intake. Evidence is emerging that specific types and patterns of physical activity, rather than absolute levels, may have differential associations for human health (La Porte et al, 1985). Recall of these characteristics is, therefore, important.

Cognitive survey research has shown that individuals report usual patterns of events more accurately than events that occur unusually since these patterns can be more readily recalled from generic memory (Jobe et al, 1993). Thus, occupational activities are constant over a longer period of time and are, therefore, easier to remember. Exercise or sporting activities, which are more subject to variation, are more difficult to recall.

Jacobs et al (1993) evaluated the reliability and validity of 10 commonly used physical activity questionnaires in 78 men and women aged 20-59 yrs. Most questionnaires, even the very simple ones, were related to the performance of heavy intensity physical activity. Few of them were related to light or moderate activities. Freidenreich et al (1998) found heavy activity was more easily remembered and reported. Light activities appear to be the most difficult to recall. Unfortunately they are also the most common. Most existing simple questionnaires probe predominantly for heavy intensity activities (Jacobs et al,

1993). This makes linguistic sense since most people asked about performance of regular exercise seem to think about vigorous or organised activities, not routine activities.

The stability and timeframe of behaviour involved link in with the intensity of the activity. For highly stable activities like sleep, heavy intensity leisure activities and household chores the timeframe of the question and the details of the scope of the activity seem to be of little importance. For highly variable activities, like light or moderate, the issue of recent vs. habitual looms large (Jacobs et al, 1993).

2.6 Validation of Measurement:

To be valid an instrument must measure what it is intended to measure (LaPorte et al, 1985). Physical activity is very difficult to measure. Therefore, it is of great concern how well self-reported physical activity accurately represents a person's habitual activity status. The principle difficulty in establishing the validity of a physical activity measure is the lack of an accepted 'gold standard' criterion measure for comparison. Investigators have, therefore, resorted to indirect validation approaches. Most commonly, questionnaires have been compared with direct monitoring of physical activity e.g. cardiorespiratory fitness.

Given the uncertainty over methods of validation field tests of repeatability at least are highly desirable and essential good practice.

2.7 Reliability of Measurement:

Testing the reliability of activity questionnaires is problematic. It is difficult with repeated measures of physical activity to determine whether you are measuring the

unreliability of the measuring instrument or true variation in the physical activity of an individual. Information regarding the stability of physical activity patterns in the population is scarce. Booth et al (1996) assessed repeatability on 115 participants recall of activity over 2 different time periods and a repeat measure of the same time period. The results of this study suggest that variations in repeatability coefficients between recall of the same 2 week period and activity recalled over different 2 week periods was due to actual variation in physical activity participation over different time periods, not poor recall or poor measurement characteristics.

The main problem with using repeatability as a criterion for the quality of a physical activity assessment is that a large proportion of the study population might report no participation at both test and retest measurements (Booth et al, 1996). These respondents would have identical zero values for both measurements, potentially inflating the measure of reliability. An assessment of the reliability of any measure, therefore, should consider the statistics derived from both the whole study sample and from the subset of respondents who reported participating in the activity of interest on at least one occasion.

2.8 Measurement of physical activity: Conclusion:

As with other complex behaviours, an accurate assessment of lifetime patterns of physical activity is difficult to achieve. Few people have stable activity patterns throughout their lives. Instead most people have activity patterns that vary daily, seasonally, or during different periods of their life. Also it is difficult to assess activity during the aetiological window of time when the activity is most important because the time period of exposure that influences cancer risk is not known. The optimal approach for cancer epidemiology is to use multiple assessments over many years to help combat this problem and increase the

precision of the measurement (Oliveria & Christos, 1997). Alternatively a quantitative history survey covering lifetime physical activity could be used. Case-control studies more completely assess lifetime activity levels, or levels at different periods of life; cohort studies are not usually appropriate as they generally rely on activity levels recorded at baseline.

The methods of measuring and classifying physical activity in epidemiological studies are not standardised. The methods can be divided into broad categories but there is still great variety within each category. Currently there is no accepted 'gold standard'. No one method of physical activity measurement is valid, reliable, practical, and non-interfering (LaPorte et al, 1985). Techniques that are the most valid and reliable (direct monitoring) are not practical or can actually alter the behaviour of the subject and influence the activity being measured. Conversely, those techniques that are practical and/or do not interfere with normal activity e.g. global, recall, or quantitative history surveys, are likely to be less valid and reliable. These surveys rely heavily on recall processes to obtain information about physical activity at some point in the past. A weakness of all surveys is recall bias (Ainsworth, 1998).

A physical activity questionnaire should attempt to cover lifetime activity of all types (occupational, household, and recreational) and focus on the intensity, duration, and frequency of each activity. The section of the SNEHD questionnaire that records physical activity was written to take account of the above points. It covers a long period of time and records the intensity, duration and frequency of each activity. Analysis of data from these questions is presented in chapter 9.

It is only with the development of quantitative history survey methods that detailed records of physical activity are being collected. These methods should be used for cancer epidemiology in the future.

3. Health Consequences of Physical Activity

3.1 Introduction:

"Physical activity reduces the risk of premature mortality in general, and of CHD, hypertension and diabetes mellitus in particular. Physical activity also improves mental health and is important for the health of muscles, bone, and joints" (US Dept Health, 1996, p4). The influence of physical activity on cancer risk is less clear. This may be because cancer is "a matrix of diseases differing in aetiology, site, timing, symptoms and course" (Paffenbarger et al, 1987).

Cherry (1922) was first to propose that increased physical activity protected against the development of cancer. He observed that primitive societies had lower cancer rates than more civilised cultures and that the amount of occupational activity was inversely associated with cancer mortality.

The effect of physical activity on cancer risk could act through an immune or a hormonal pathway or both. This chapter will discuss the impact of physical activity on the immune and hormonal systems and the results of studies investigating cancer risk.

3.2 Physical activity and cancer:

Study search methods:

The next section will focus on studies that have investigated the effect of physical activity on cancer risk. MEDLINE and BIDS were searched back to 1990 to find English language studies with the following key words: exercise, physical activity, cancer, aetiology, colon, breast, prostate, endometrium, ovarian, testicular, lung, Hodgkin's disease. Once these were found citations of these were followed up, as were the references in each study to get as full a picture as possible of the effect of physical activity on cancer risk. Although an attempt was made to find all studies those investigating leisure time activity or total activity will be

emphasised. Those relating to occupational activity will not be discussed in detail for reasons given in Chapter 2.

Physical activity and all-site cancer:

Eight cohort studies have been published in the literature. Of these one was based on occupational physical activity alone and this found no significant effect of activity levels (Paffenbarger et al, 1978). Three of the remaining seven have looked at fitness (Blair et al, 1989; Wannamethee et al, 1993; Steenland et al, 1995) while the remaining four have focused on recreational or total activity. Details of these studies are shown in Table 3.1.

Three of four studies looking at activity levels found a statistically significantly lower risk of cancer in those undertaking more activity (Paffenbarger et al, 1986; Albanes et al, 1989; Wannamethee et al, 1993). These results were significant for males only. Three others showed a non-significant decreased risk of cancer (Garfinkel & Stellman, 1988 (males only); Blair et al, 1989 (males only); Steenland et al, 1995 (both sexes combined)). Garfinkel & Stellman (1988) found an increased risk of cancer in women with increasing activity level. Leon et al (1991) found a non-significant increase in risk of cancer in males but suggest that the predominance of cigarette smokers in the MRFIT cohort may have negated any possible protective effect of leisure time physical activity against cancer. Linsted et al (1991) and Arraiz et al (1992) found no relationship between physical activity and risk of cancer.

Three studies investigated risk of all site cancer with fitness level. These all showed a decreased risk of cancer with higher level of fitness, but this effect was only statistically significant in one (Steenland et al, 1995).

All of these studies are prospective cohorts and are in the form of global surveys (see Chapter 2), based on a single simple activity question at baseline e.g. "How much exercise do you get (at work or play)?" (Garfinkel & Stellman, 1988). These studies take no

Table 3.1: Physical activity and all-site cancer.

Study	Study sample	Comparison	Risk Estimate	Significance
Paffenbarger et al (1986) (USA)	Cohort 16936 male Harvard alumni aged 35-74 followed 1962/66 to 1978 (446 cancer deaths)	post college activity. trend of ↓ risk with ↑ activity	NA	p=0.007
Garfinkel & Stellman (1988) (USA)	Cohort 868620 men & women in the American Cancer Society CPS II (1355 male & 993 female cancer death)	composite of work & recreational activity in 1982. heavy exercisers vs. moderate exercisers	SMR Male=99 SMR Fem=120	NS p>0.05
Albanes et al (1989) (USA)	Cohort NHANES I, 5138 men & 7407 female aged 25-74 (460 male & 399 female cancer case)	recreational activity at baseline. quite inactive vs. very active	RR Male=1.8 (1.4-2.4) RR Female=1.3 (1.0-1.8)	p<0.05 p=0.05
Blair et al (1989) (USA)	Cohort Cooper Clinic study of 10224 (m) & 3120 (w) questioned 1970-81, followed 8 yrs (64 male & 18 female cancer death)	physical fitness score. trend of ↓ death rate in higher quintiles of fitness	M slope -3.5 F slope -7.5	NS NS
Leon et al (1991) (USA)	Cohort of 12138 male middle-aged men at risk of CHD in MRFIT trial. 10.5 yr follow-up (265 cancer deaths)	level of habitual ltpa. i. tertile 2 activity vs. tertile 1 ii. tertile 3 activity vs. tertile 1	i. RR=1.22 (0.91-1.63) ii. RR=1.06 (0.78-1.44)	i. NS ii. NS
Wannamethee et al (1993) (UK)	Cohort 7735 males aged 40-59 examined 1978-80, followed to 1989 (225 cancer deaths)	heart rate & physical activity. i. light/mod activity v. inactive ii. vigorous/mod vig vs. inactive	i. RR=0.84 (0.63-1.15) ii. RR=0.59 (0.38-0.92)	i. NS ii. p<0.05
Steenland et al (1995) (USA)	Cohort 14407 men & women in National Health & Nutrition Survey 1970s-1987 (657 male & 593 female cancer death)	pulse rate at baseline. highest pulse quartile vs. lowest	RR=1.27 (1.04-1.57)	p<0.05

ltpa = leisure time physical activity
NS (not statistically significant) = p>0.10

account of changes in activity levels or ask for detailed activity information. Studies of all-site cancer also combine cancers that may be influenced by physical activity with those that are not. These problems may account for essentially null results.

Physical Activity and Colon Cancer:

Colon cancer is the malignancy for which most data is available relating to physical activity. Animal studies have shown that chemically induced colon cancer is reduced by physical activity, including both forced strenuous activity which led to weight loss (Klurfield et al, 1987), and moderate voluntary activity which did not result in weight loss (Andrianpoulos et al, 1987; Reddy et al, 1988).

A literature search produced 35 studies that investigated the risk of colon cancer by activity level. Ten of these studies used only measures of occupational activity (Milham et, 1983; Vena et al, 1985; Fredriksson et al, 1989; Brownson et al, 1989; Peters et al, 1989; Brownson et al, 1991; Dosemeci et al, 1992; Arbman et al, 1993; Vineis et al, 1993; Chow et al, 1994). All of these showed an increased risk of colon cancer with lower activity levels and five of the results were statistically significant (Milham et al, 1983; Vena et al, 1985; Fredriksson et al, 1989; Brownson et al, 1991; Chow et al, 1994). The effect was generally stronger in males if the sexes were analysed separately. The results of occupation based studies are remarkably consistent, especially when few have evaluated confounders.

The remaining studies in the literature investigated recreational or total activity. Of these thirteen attempted to adjust for at least one of the crucial confounding effects of diet or body size. The results of these studies are shown in Table 3.2.

Again the results are consistent. Ten of the studies had a statistically significant reduction in risk of colon cancer with increasing activity levels. The only exception to this reduction was Marcus et als' (1994) finding that risk of colon cancer was increased in women

Table 3.2: Physical activity and Colon cancer

Study	Numbers	Comparison	Risk Estimate	Significance
Severson et al (1989) (USA)	Cohort of 8006 Japanese men (age 46-65) living in Hawaii in 1965, followed for 20 years (191 cases)	total physical activity, HR at baseline. i. highest vs. lowest tertile activity ii. high HR vs. low	i. RR=0.71 (0.51-0.99) ii. RR=1.37 (0.97-1.93)	i. p for trend NS ii. p for trend <0.05
Ballard-Barbash et al (1990) (USA)	Cohort Framingham heart study, 1906 male & 2308 female aged 30-62, followed 28 yrs (73 male & 79 female cases)	physical activity composite. lowest vs. highest tertile activity	RR Male=1.8 (1.0-3.2) RR Female=1.1 (0.6-1.8)	p=0.05 NS
Slattery et al (1988) (USA)	Nested C/C study 229 cases (age 40-79) & 384 popln controls	physical activity 2 yr. prior to interview. highest v. lowest quartile of EE	OR male=0.70 (0.38-1.29)* OR female=0.48 (0.27-0.87)*	NA NA
Giovannucci et al (1995) (USA)	Cohort 47723 male health professionals (age 40-75) questioned 1986 followed 1992 (201 cases)	recreational physical activity. highest quintile EE/day vs. lowest	RR=0.53 (0.32-0.88)	p for trend 0.03
Min-Lee et al (1997) (USA)	Cohort 21807 Physician's Health study (age 40-84) enrolled 1982 followed 1994 (217 cases)	physical activity measured at baseline & 3 years later. inactive at both assessments vs. active at both	OR=1.3 (0.9-2.0)	NA
Marcus et al (1994) (USA)	C/C study 536 female cases (age <74) & 2315 popln controls	total strenuous activity age 14-22 any strenuous activity vs. none	OR=1.2 (0.82-1.27)	p for trend NS
White et al (1996) (USA)	C/C study 444 male & female cases (age 30-62) & 427 popln controls	total activity in 10 yrs < diagnosis. moderate/high intensity activity (≥4.5 METs) 5 hrs/wk vs. 0	RR=0.78 (0.55-1.10)	p for trend NS

Table 3.2: Physical activity and Colon cancer (continued).

Slattery et al (1997) (USA)	Matched C/C study 2073 cases (age 30-79) & 2466 controls (age- & sex-matched)	ltpa lowest amount of activity vs. highest	OR male=1.63 (1.26-2.12) OR female=1.59 (1.21-2.10)	p>0.05 p<0.05
Longnecker et al (1995) (USA)	C/C study of 163 cases (age ≥32) & 703 controls	vigorous ltpa ≥2 hrs/wk vs. 0	OR=0.57 (0.33-0.97)	p for trend 0.06
Thune & Lund (1996) (Norway)	Cohort 53242 males & 28274 females (age 20-69) screened 1972-78, av. follow-up 15 yrs (236 male & 99 female cases)	total activity active vs. sedentary	RR male= RR female=0.63 (0.39-1.04)	p for trend 0.04 p for trend 0.04
Thun et al (1992) (USA)	Nested C/C study of 1150 cases (av. age 57) & 5746 controls (age-, sex-, race-matched)	ltpa heavy exercise vs. none	RR male=0.60 (0.28-1.27) RR female=0.90 (0.41-1.96)	NS NS
Gerharrdsson et al (1990) (Sweden)	C/C study of 452 cases (age 40-80) & 624 popln controls	total & ltpa activity sedentary total activity vs. v.active	OR=1.8 (1.0-3.4)	p=0.05

NS (not statistically significant) = p>0.10
HR = heart rate. EE = energy expenditure. ltpa = leisure time physical activity.
* 90% confidence intervals.

with increasing activity level. In the studies that looked at the sexes separately the protective effect was generally stronger in males than females. This could be due to the narrower range of activity exhibited by women. Of the studies that tested for a dose-response, six found a statistically significant decrease in colon cancer risk as activity levels increased (Slattery et al, 1988; Severson et al, 1989; Gerhardsson et al, 1990; Giovannucci et al, 1995; Longnecker et al, 1995; Thune & Lund, 1996).

There are several mechanisms by which physical activity could reduce the risk of colon cancer. The most widely held hypothesis is that physical activity stimulates colon peristalsis thereby decreasing the time that dietary factors, endogenous secretions such as bile acids, and carcinogens reside in the bowel (Burkitt, 1971; Burkitt et al, 1972). It is a well-known clinical observation that physical inactivity, particularly bed rest, is associated with constipation. At the other extreme long distance runners have been reported to develop diarrhoea and even faecal incontinence (Sullivan & Wong, 1992). Several studies have shown that exercise does decrease transit time (Holdstock et al, 1970; Cordain et al, 1986; Oettle et al, 1991; Koffler et al, 1992). However, others have not (Bingham & Cummings, 1989; Lampe et al, 1991. Coenen et al, 1992). Also transit time is not a well-established risk factor for colon neoplasia.

Exercise may affect prostaglandin synthesis, specifically a type of prostaglandin called F_2 alpha. Animal studies have demonstrated that F_2 alpha inhibits tumour growth in the colon, and strenuous physical activity appears to increase the levels of prostaglandin F_2 alpha (Bennett & Del Tacca, 1975; Tutton & Barkla, 1980). Prostaglandin E_2 is synthesised in greater quantities in human cancer cells (Bennett & Del Tacca, 1975). Inhibitors of prostaglandin synthesis are known to suppress colon cancer development and observations have revealed that exercise may reduce prostaglandin E_2 levels.

Alternatively, the decrease in risk could be due to associated beneficial lifestyle changes, such as a low fat diet (Sandler et al, 1995).

Together the research on physical activity and colon cancer strongly suggests that activity has a protective effect against the risk of developing this cancer.

Physical activity and cancers associated with the hormonal system:

Prostate Cancer One of the main actiologic hypotheses for prostate cancer is prolonged androgenic stimulation (Le Marchand et al, 1991). Since physical training lowers testosterone levels (see p.94) physical activity may protect against the development of prostatic disease through a favourable effect on hormone profiles (Hartman et al, 1998). However, in most studies possible confounding by diet, body size, race or socio-economic status have not been considered.

Seven published studies have investigated only occupational activity. Three studies described a statistically significantly increased risk of prostate cancer with lower activity levels (Milham et al, 1983; Vena et al, 1987; Brownson et al, 1991). Three more described an increase in risk with low levels of occupational activity, but the results were not significant (Paffenbarger et al, 1987; Dosemeci et al, 1993; Hsing et al, 1994).

The eighteen studies of recreational or total physical activity, or cardiorespiratory fitness, have also produced inconsistent results. Only eight of these studies made any attempt to control for confounding by diet or body mass. These can be considered the best studies and their results are presented in Table 3.3 (although diet has only been adjusted for in one study: Ilic et al, 1996). One study (Albanes et al, 1989) found leisure time physical activity to be statistically significantly protective for prostate cancer. Two studies found an inverse relationship (Oliveria & Blair, 1993; Thune & Lund, 1994) but the results were not significant. In the study by Thune & Lund (1994) the relationship was limited to those men aged over 60 years. Two studies found no effect of recreational activity or recreational and

Table 3.3: Physical activity and prostate cancer

Study	Numbers	Comparison	Risk Estimate	Significance
Albanes et al (1989) (USA)	Cohort NHANES I 5138 (m) aged 25-74 (95 cases)	ltpa. little/no exercise v. much	RR=1.80 (1.0-3.3)	p for trend <0.05
Ilic et al (1996) (Serbia)	Matched C/C study 101 cases & 202 controls (age-, residence-, hospital entry date-matched)	lifetime ltpa.	Not given	NS
Thune & Lund (1994) (Norway)	Cohort 53242 (age 19-50 at entry), av. follow-up 15 yrs (220 cases)	total activity at baseline. walking at work & moderate ltpa vs. sedentary walking at work & regular rec training vs. sedentary	RR=0.61 (0.36-1.01) RR=0.45 (0.20-1.01)	p for trend 0.03
Oliveria et al (1996) (USA)	Cohort 12975 men (age 20-80) examined 1970-89 (94 cases)	fitness quartile 1982. i. 2 vs. 1 ii. 3 vs. 1 iii. 4 vs. 1	i. RR=1.1 (0.63-1.77) ii. RR=0.73 (0.41-1.29) iii. RR=0.26 (0.10-0.63)	p for trend <0.004
Gann et al (1995) (USA)	Cohort 22380 men enrolled 1967-73, av. follow-up 19.2 yrs	heart rate at baseline. quintiles of heart rate (lowest to highest)	(i) RR=1.00 (ref) (ii) RR=1.55 (iii) RR=1.85 (iv) RR=2.18 (v) RR=2.69	p for trend 0.006
Cerhan et al (1997) (USA)	Cohort 1050 men (age 65-101) followed 1982-93 (71 cases)	physical activity level high activity level vs. inactive	RR=1.9	p for trend 0.01
Severson et al (1989) (USA)	Cohort 8006 Japanese men living in Hawaii 1965, followed 20yrs (206 cases)	total & recreational activity. i. highest total tertile vs. lowest ii. as above ltpa	i. RR=1.05 (0.75-1.47) ii. RR=1.05 (0.73-1.51)	i. NS ii. NS

NS (not statistically significant) = p>0.10
ltpa = leisure time physical activity.

occupational activity combined (Severson et al, 1989; Ilic et al, 1996) and one found that greater activity levels to be associated with an increased risk (Cerhan et al, 1997).

Studies of the association with cardiorespiratory fitness are more consistent. Both Gann et al (1995) and Oliveria et al (1996) observed a statistically significantly protective effect for higher fitness level. Both of these studies had a statistically significant dose-response. However, in contrast to the results of Thune & Lund (1994), the effect of fitness in the Oliveria et al (1996) was limited to men aged <60 years.

The prostate is primarily an androgen dependent gland controlled essentially by levels of plasma testosterone, of which 90-95% of the body's production is synthesised in the testes. Physical activity may protect against the development of prostate cancer through reductions in the resting levels of endogenous sex hormones, particularly testosterone (Ross et al, 1988; Gittes, 1991). The prostate gland is dependent on these hormones for growth and development and is affected by hormonal stimulation. A recent case-control study conducted within the Physician's Health Study (Gann et al, 1996) found that high levels of circulating testosterone was associated with an increased risk of prostate cancer. Men with prostate cancer have higher levels of endogenous testosterone than non-cancerous men, and cancerous tissues have been reported to have higher levels of testosterone than healthy tissue (Ahluwahia et al, 1981). If exercise does decrease the risk of prostate cancer, a chronic decrease in testosterone levels would be important in influencing risk as compared with transient change, suggesting that long-term exercise may be necessary for decreased prostate cancer risk.

The body of evidence to date shows no consistent relationship between prostate cancer and physical activity. Further studies are needed to investigate the frequency, intensity, and duration of activity, as well as the type and period of life during which activity might be beneficial. However, it is reasonable that exercise may be a potential factor that can be modified to protect from prostate cancer.

Breast Cancer The intensity and duration of exercise has been found to affect the development of experimentally induced breast cancer in rats (Thompson, 1994). Several mechanisms for this effect have been hypothesised by Kramer & Wells (1996). The relationship could be due to: 1) the maintenance of low body fat and modulation of extra-glandular oestrogen; 2) a reduction in the number of ovulatory cycles and subsequent diminuition of lifetime exposure to endogenous oestrogen; 3) enhancement of natural immune function; 4) association with other healthy lifestyle habits.

Most aetiologic theories have focused on oestrogen because of the numerous menstrual and reproductive risk factors for breast cancer e.g. women who have an early menarche, late menopause, and low parity are at increased risk. Physical training has been shown to decrease oestrogen levels (see p.91). Many breast cancer risk factors that are believed to operate through an oestrogen pathway are culturally or personally determined by a woman's lifestyle. In an effort to reduce risk, however, few of these factors are easily modifiable, with the exception of physical activity (Gammon et al, 1996).

27 studies on the relationship of physical activity to breast cancer risk have been published. Seven of these have focused on only occupational activity (Milham et al, 1983; Vihko et al, 1992; Dosemeci et al, 1993; Pukkala et al, 1993; Zheng et al, 1993; Coogan et al, 1997; Mezzetti et al, 1998). Of these seven studies four (Milham et al, 1983; Pukkala et al, 1993; Zheng et al, 1993; Mezzetti et al, 1998) found a statistically significant increase in risk of breast cancer with low levels of occupational activity. The others also found the same pattern of risk but the results were not significant.

The remaining 21 studies have investigated total or recreational activity. The majority of these have made an effort to control for confounding by SES, reproductive history, body mass, and diet. Eighteen controlled for female reproductive history, 16 for body mass, and 14

adjusting for dietary factors (of these 14 5 only considered alcohol intake). The results from these studies are shown in Table 3.4.

Six studies analysed all women together. Two studies found a statistically significantly increased risk of breast cancer with lower levels of activity (Frisch et al, 1985; Fraser et al, 1997), two found a non-significant protective effect (Dorgan et al, 1994; Freidenreich et al, 1995; D'Avonzo et al, 1996), one found no effect (Albanes et al, 1989), and one found activity to increase risk (Dorgan et al, 1994). Nine studies presented results by menopausal status. Nine studies had results for pre-menopausal women with three having a significant inverse relationship (Bernstein et al, 1994; Hirose et al, 1995; Thune et al, 1997). In two studies there was no effect (Chen et al, 1997b; Gammon et al, 1998) while two studies revealed that a lower level of activity was associated with a lower risk of breast cancer in pre-menopausal women (Albanes et al, 1989; Hu et al, 1997). The pattern of risk for post-menopausal women is rather more consistent with seven of the nine studies having an inverse risk. Of these studies four were statistically significant (Hirose et al, 1995; McTiernan et al, 1996; Cerhan et al, 1996; Carpenter et al, 1996) and three were not (Albanes et al, 1989; Freidenreich et al, 1995; Hu et al, 1997). In the remaining two studies activity level had no effect on breast cancer risk (Taoli et al, 1995; Thune et al, 1997).

It is possible that physical activity during adolescence and young adulthood may protect against the later development of breast cancer. Eight of the studies in Table 3.4 examined this possibility. Among the eight studies, two found a strong and statistically significant reduction in risk (Frisch et al, 1985; Mittendorf et al, 1995), one found a non-significant reduction in risk (D'Avonzo et al, 1996), and four no association (McTiernan et al, 1996; Chen et al, 1997b; Rockhill et al, 1998; Gammon et al, 1998). However, it is not

Table 3.4: Physical activity and Breast cancer.

Study	Numbers	Comparison	Risk Estimate	Significance
Frisch et al (1985) (USA)	Cohort 5398 former college students (age 21-80), 2622 athletes & 2776 non-athletes (69 cases)	college sport. non-athletes vs. athletes	RR=1.86 (1.00-3.47)	p=0.05
Rockhill et al (1998) (USA)	Cohort Nurses Health Study II (age 25-42) enrolled 1989, 6 years follow-up (372 cases)	strenuous ltpa age 18-22 & no. hours ltpa at baseline. i. age 18-22 2/wk 10-12 months p.a. vs. none ii. 1989 ltpa ≥7 hrs/wk vs. <1 hr	i. RR=1.1 (0.8-1.6) ii. RR=1.1 (0.8-1.5)	i. NS ii. NS
McTiernan et al (1996) (USA)	C/C study 537 cases (age 50-64) & 492 controls	ltpa 2yrs <interview. high intensity exercise 3hrs/wk vs. 0	i. OR=0.6 (0.4-1.0)	i. p for trend 0.07
Albanes et al (1989) (USA)	Cohort NHANES I 7407 women (age 25-74 at enrolment) followed 1971-75 to 1982-84 (122 cases)	ltpa. all women sedentary vs. most active	RR=1.0 (0.6-1.6)	NS
Bernstein et al (1994) (USA)	C/C study 545 cases (age <40) & 545 controls	lifetime ltpa. ≥3.8 hr/wk exercise vs. 0	OR=0.42 (0.27-0.64)	p for trend <0.05
Friedenreich et al (1995) (Australia)	C/C study 444 cases (age 20-74) & 444 matched ward controls	ltpa. all women 4000 kcal/wk vs. 0	OR=0.73 (0.51-1.05)	NS
Cerhan et al (1996) (USA)	Cohort 1825 (age 65-102 at entry) followed 1982-93 (46 cases)	activity at baseline. i. moderate activity vs. none ii. high activity vs. none	i. RR=0.5 (0.2-1.0) ii. RR=0.2 (0.1-0.9)	p for trend 0.007
Thune et al (1997) (Norway)	Cohort 25614 (age 20-49 at enrolment) screened 1974-83, av. follow-up 13.7yrs (351 cases)	ltpa. all women regular exercise vs. sedentary	RR=0.63 (0.42-0.95)	p for trend 0.04

Table 3.4: Physical activity and Breast cancer (continued).

D'Avonzo et al (1996) (Italy)	C/C study 2569 cases (age 23-74) & 2588 hospital controls	ltpa. i. lowest v. highest quintile 15-19 ii. lowest v. highest quintile 30-39 iii. lowest v. high quintile 50-59	i. RR=0.95 (0.77-1.18) ii. RR=0.76 (0.55-1.05) iii. RR=0.66 (0.30-1.25)	i. NS ii. NS iii. NS
Chen et al (1997b) (USA)	C/C study 747 cases (age 21-45) & 961 controls	ltpa 2yr < ref. date. 4+ episodes/wk vs 0	OR=0.93 (0.71-1.22)	p for trend NS
Gammon et al (1998) (USA)	C/C study 1668 cases (age <45) & 1505 controls	total activity age 12-13, 20 & year <interview. i. lowest v. highest quartile 12-13 ii. lowest v. highest quartile 20 iii. lowest v. high quartile last year	i. OR=0.98 (0.79-1.21) ii. OR=1.03 (0.83-1.27) iii. OR=1.15 (0.93-1.43)	ii. p for trend NS iii. p for trend 0.04
Hirose et al (1995) (Japan)	C/C study 1186 cases (age 20-80) & 23163 hospital controls	current ltpa. i. pre-menopausal 2 vs. 0 episodes/wk ii. post-menopausal as above	i. RR=0.64 (0.48-0.84) ii. RR=0.71 (0.53-0.96)	i. p<0.05 ii. p<0.05
Mittendorf et al (1995) (USA)	C/C study 6888 cases (age 17-74) & 9539 controls	strenuous ltpa age 14-22. daily vs. none	RR=0.5 (0.4-0.7)	p for trend <0.05
Fraser et al (1997) (USA)	Cohort 20341 Adventist Health Study (age 24-90 at enrolment) (218 cases)	total activity. low v. high level total activity	RR=1.46 (1.11-1.92)	p<0.05
Dorgan et al (1994) (USA)	Cohort 2298 women (age 35-68) in Framingham heart study (117 cases)	ltpa. highest vs. lowest quartile	RR=1.6 (0.9-2.9)	p for trend <0.05 (opposite direction)
Carpenter et al (1996) (USA)	Matched C/C study 1362 cases (age>55) & controls (age-, race-, neighbour-matched)	lifetime strenuous ltpa. 3.7hrs/wk vs. 0	OR=0.52 (0.32-0.95)	p<0.05
Taioli et al (1995) (USA)	C/C study 617 hospital cases (all ages) & 531 hospital controls	ltpa. pre-menopausal ≥3hrs/wk vs. none	OR=0.7 (0.4-1.4)	NS

NS (not statistically significant) = p>0.10
ltpa = leisure time physical activity.

easy to compare the results of these studies as they all looked at slightly different age groups, from 12-13 years (Gammon et al, 1998) to college athletes (Frisch et al, 1985).

Breast cancer has a poorly understood aetiology. Because the effect of physical activity on breast cancer is likely to be modest, or vary throughout lifetime, measurement of physical activity needs to be very accurate to minimise the possibility that an effect is missed because of measurement error (Freidenreich et al, 1998). Non-significant results may be due to the low levels of physical activity reported by women who did exercise and the low prevalence of reported physical activity. Future studies should include the entire period from childhood to diagnosis and include measurement of recreational, household and occupational activity. Although the evidence is not conclusive that physical activity has a protective effect on breast cancer there is sufficient evidence to warrant more research on this topic.

Ovarian cancer Three studies have investigated an effect of activity on ovarian cancer risk (Table 3.5). Of these two looked only at occupational activity with both finding a slightly elevated risk of ovarian cancer with inactive occupations (Pukkala et al, 1993; Zheng et al, 1993). Mink et al (1995) found a positive association of ovarian cancer with high levels of physical activity. This is inconsistent with existing theories of ovarian cancer pathogenesis.

Endometrial Cancer There are five studies in the literature investigating endometrial cancer risk. One of these was based on occupational activity only (Kalandidi et al, 1996). This study found a decreased risk of endometrial cancer in jobs requiring manual labour ($p=0.03$). The remaining four investigated both leisure time activity and occupational activity and adjusted for the major potential identified confounding factors. The results of these studies are given in Table 3.5. None of the results for leisure time activity were statistically significant except for a moderate amount of vigorous exercise in young adulthood and 20 years before interview in the study by Olson et al (1997). However, the

Table 3.5: Physical activity and other female hormonal cancers.

Study	Numbers	Comparison	Risk Estimate	Significance
Endometrial cancer				
Sturgeon et al (1993) (USA)	C/C study 405 cases (age 20-74) & 297 controls	lifetime ltpa. always inactive vs. always active	RR=1.5 (0.7-3.2)	NS
Olson et al (1997) (USA)	C/C study 145 cases (ages not given) & 298 popln controls	lifetime total activity. top level activity 2yr <interview	OR=0.67 (0.42-1.09)	NS
Levi et al (1993) (Switzerland, Italy)	C/C study 274 cases (age 31-74) & 572 hospital controls	total activity. lowest level vs. highest	RR=2.4 (1.0-5.8)	p for trend <0.05
Shu et al (1993) (China)	C/C study 268 cases (age 18-74) & 268 controls	ltpa lowest quartile vs. highest	RR=0.8 (0.8-1.3)	NS
Ovarian cancer				
Mink et al (1996) (USA)	Cohort 31396 post-menopausal women (97 cases)	total activity most active vs. least	RR=1.97 (1.22-3.19)	p for trend 0.006 (opposite direction)

NS (not statistically significant) = $p>0.10$
ltpa = leisure time physical activity.

results all point in the same direction, a protective effect of recreational activity on endometrial cancer risk. Of the two studies that looked for a dose-response (Levi et al, 1993; Olson et al, 1997) only one found a statistically significant relationship (Levi et al, 1993).

Many observations implicate oestrogen as having a role in cancers of the reproductive tissues: the breast; endometrium; cervix; and ovary. This evidence includes observations that the cancer risks are modified in association with age at menarche, number of ovulatory cycles, and age at menopause. There is some evidence that suggests that physical activity can influence at least some of these aspects of menstrual history.

Strenuous exercise has been found to delay menarche in many cases (Warren, 1980; Frisch et al, 1981) and cause oligomenorrhea and secondary amenorrhea (Dale et al, 1979; Frisch et al, 1981). These disturbances to the menstrual cycle may result in a reduced lifetime exposure to ovarian hormones. However, even participation in moderate levels of activity can have discernible effects on the frequency of ovulatory cycles. Bernstein et al (1987) found moderate physical activity (as defined by >600 kcal energy expenditure per week) was associated with a three-fold increase in risk of anovulatory cycles.

More active women have lower levels of oestrogen (see p.91) or increased metabolism of oestrogen to less potent forms.

The life period during which activity could effect breast cancer risk is also not known. Thus, while exercise does have an effect on female sex steroid hormones it is unclear whether this is responsible for the lower risk of female reproductive cancers seen in some studies.

The body of work on ovarian and endometrial cancer does not show a consistent statistically significant relationship between other hormone dependent cancers in women and physical activity.

Physical activity and Other Cancers:

Testicular cancer Forman et al (1994) and the United Kingdom Testicular Cancer Study Group (UKTCSG) (1994) found exercise to be moderately protective for testicular cancer. Thun & Lund (1994) found no such effect. Coldman et al (1982) and Haughey et al (1989) have found increased risk of testicular cancer associated with cycling and horse riding, but (UKTCSG) found no evidence of such an effect.

Lung Cancer Exercise has been found to protect against lung cancer in 6 studies (Garfinkel & Stellman, 1988; Albanes et al, 1989; Paffenbarger et al, 1987; Paffenbarger et al, 1978; Paffenbarger et al, 1986; Severson et al, 1989). However, only one of these (Albanes et al, 1989) was statistically significant and adjusted for smoking status and cigarette pack year history.

3.3 Physical activity and the endocrine system:

The effects of physical activity on the endocrine system can be divided by sex and also the effects produced by a single episode of activity versus the effect of more long-term training.

Female: Single episode:

Most interest has been on the effect of exercise on the levels of progesterone and oestrogen but most studies have looked at the chronic (p.91), rather than the acute, effects. Jurkowski (1978), Bonen et al (1979), and Fahey et al (1997) have found significantly elevated levels of oestradiol and progesterone after a single activity episode. Jurkowski (1978) investigated three activity levels and found the increase in progesterone was related to exercise intensity as did Fahey et al (1997) with the significant increase in progesterone only seen after high intensity exercise. Jurkowski (1978) also found that oestradiol levels were only

statistically significantly increased after heavy or exhaustive exercise and not light activity. In contrast, Fahey et al (1997) found oestradiol to be significantly increased after 1 hour of low and 1 hour of high intensity activity. Bonen et al (1979) found heavy exercise in untrained subjects produced significant increments in ovarian hormones; no such increments were observed in trained subjects at the same absolute exercise workload.

Several studies have shown a statistically significant increase in the level of female testosterone (Sutton et al, 1973; Shangold et al, 1981; Obminski et al, 1997). However, the level of testosterone remained well within the normal range for women. Sutton et al (1973) found that only maximal exercise had a statistically significant effect on testosterone levels and not sub-maximal effort. Kraemer et al (1993) found no relationship between a single episode of activity and testosterone.

Studies investigating the pituitary hormones lutenizing hormone and follicle stimulating hormone have found essentially no change (Sutton et al, 1973; Jurkowski et al, 1978; Bonen et al, 1979; Bonen et al, 1981).

Cortisol (Obminski et al, 1997; Kraemer et al, 1998a), prolactin (Shangold et al, 1981; Bullen et al, 1982), and growth hormone (Hakkinen & Pakarinen, 1995; Kraemer et al, 1998a) levels are increased after acute exercise.

Female: Training:

Most investigations again focused on the long term exercise effects on the levels of the ovarian hormones oestrogen and progesterone. The majority of studies find a lower level of oestradial, oestrogen, or oestrone in regular exercisers (Dale et al, 1979; Boyden et al, 1983; Bullen et al, 1985; Nelson et al, 1988; Cauley et al, 1989; Newcomb et al, 1995; Nagata et al, 1997; De Souza et al, 1998). This effect is seen in both pre- and post-menopausal women

(McTiernan et al, 1996). However, not all studies have found a decreased level of oestrogen in trained women (Bonen et al, 1979; Schwartz et al, 1981; Bullen et al, 1982). Bullen and colleagues (1982) looked only at moderate activity and found no effect whereas De Souza et al (1998) found moderate physical activity statistically significantly decreased oestrone levels. Boyden et al (1983) suggested that it is particularly endurance running that lowers the circulating concentration of oestrogen.

The majority of evidence points to regular exercise lowering the concentration of progesterone with most (Dale et al, 1979; Bonen et al, 1981; Ellison & Lager, 1986), but not all (Bonen et al, 1979) reductions being statistically significant.

The majority of studies have found no effect of training on testosterone levels in women (Baker et al, 1981; Schwartz et al, 1981; Boyden et al, 1983; Weiss et al, 1983; Hakkinen et al, 1990; Kraemer et al, 1991). Only one study suggested that testosterone levels are statistically significantly elevated in established female runners (Dale et al, 1979) but the testosterone levels were still well within normal levels for women.

As with single episodes of exercise little work has been performed on cortisol and prolactin levels. Bullen et al (1982) suggest that moderate activity may increase the levels of circulating cortisol but have no effect on prolactin. More studies are needed to investigate these relationships.

Male: Single episode:

The focus of an exercise effect on the endocrine system in men has been testosterone. The majority of studies have found that acute exercise increases testosterone levels in men (Sutton et al, 1973; Hakkinen & Pakarinen, 1995; Volek et al, 1997; Obminski et al, 1997; Fahrner & Hackney, 1998; Kraemer et al, 1998a). This relationship has been found in old

men as well as young (Kraemer et al, 1998b; Hakkinen & Pakarinen, 1995) and for more moderate (Fahrner & Hackney, 1998) and high intensity activity (Sutton et al, 1973; Obminski et al, 1997; Volek et al, 1997). In contrast, Fournier et al (1997) found testosterone levels to significantly fall over the course of a 110km ultra-marathon, possibly due to the very long duration of intense exercise. Kuoppasalmi et al (1980) found maximal short term running to have little effect on testosterone levels but a moderate intensity run of 90 minutes led to a decreased level. Short-term intense exercise appears to increase testosterone levels but prolonged exercise of at least 90 minutes typically results in a reduced level, regardless of intensity. These levels tend to return to normal within 24 hours. More investigation is needed into the effect of sub-maximal exercise on endogenous testosterone levels as such a relationship is still to be fully described.

The mode of exercise can also affect the exercise-induced change in testosterone. In contrast to the literature on running (Kuoppasalmi et al, 1976), cycling (Cumming et al, 1986), and resistance exercise (Kraemer, 1988) Cumming et al (1987) observed a decrease in testosterone levels following intense swimming exercises. Jensen et al (1991) observed no difference in testosterone levels following intense endurance exercise compared to strength exercise.

Growth hormone rises during short-term exercise after an initial delay of about 10 minutes but plateaus after 30 minutes (Sutton et al, 1973; Felsing et al, 1992; Richter et al, 1994; Hakkinen & Pakarinen, 1995). This change in hormone level is not statistically significant if physical activity is performed at a reduced level. However, Fukatsu et al (1996) did not observe an increase in GH, even after prolonged intense exercise.

Cortisol is consistently elevated following acute exercise (Hartley et al, 1972; Sutton et al, 1973; Hakkinen & Pakarinen, 1995; Fukatsu et al, 1996; Farmer et al, 1997; Obminski

et al, 1997; Kraemer et al, 1998a; Kraemer et al, 1998; Fahrner & Hackney, 1998). In studies that compared maximal and sub-maximal activity the effect was limited to the former (Sutton et al, 1973; Hakkinen & Pakarinen, 1995).

Male: Training:

The focus of studies of the effect of long-term exercise on resting hormone levels has again been testosterone. In general, if sedentary people undergo a training programme or athletes increase their training level, the resting blood level of testosterone falls (Webb et al, 1984; Wheeler et al, 1984; Urhausen et al, 1987; Flynn et al, 1994; Duclos et al, 1996; Gulledge & Hackney, 1996; Flynn et al, 1997; Hackney et al, 1998). Other authors have not found a difference in resting testosterone levels between trained athletes and sedentary controls (Lucia et al, 1996) or no change in testosterone levels over the course of 8-12 weeks of training (Lucia et al, 1996). Sutton et al (1973) and Kraemer et al (1998a) found an increased level of testosterone.

Flynn et al (1997) has found a greater decrease in levels of free testosterone than total testosterone after a doubling of mileage in male distance runners. A finding confirmed by Wheeler et al (1984), Flynn et al (1994), and Hackney et al (1998). The greater effect on circulating levels of free testosterone is important. Free testosterone more accurately reflects the biologic activity of circulating testosterone and suggests a more notable decrease in androgen bioavailability than would be apparent from the measurement of total testosterone alone.

There is no agreement if these changes in testosterone mean that those who train regularly have testosterone levels outside of the physiological range. Although the mean level of total and non-specifically bound testosterone was statistically significantly lower in runners

it was still within the physiologic range in studies by Wheeler et al (1984) and Hackney et al (1998).

Although trained athletes appear to have lower serum levels of testosterone at rest, a relationship between physical activity and testosterone level has not been confirmed in older men and the effect of moderate and light activity on hormonal status has not been evaluated.

Very little work has been performed on the long-term effect of physical activity on other hormones in men. Little effect of activity has been seen on resting levels of cortisol (Wheeler et al, 1984; Gulladge & Hackney, 1996; Mujika et al, 1996; Flynn et al, 1997), or prolactin (Gulladge & Hackney, 1996; Mujika et al, 1996; Hackney et al, 1998).

Physical activity and the endocrine system: Conclusion:

Improvements in laboratory techniques have allowed research related to exercise and endocrinology to flourish. The emerging literature, however, is often inconsistent or contradictory. Exercise intensity, duration, mode, frequency, and volume may each have specific effects on the endocrine changes seen with exercise and training. Furthermore, hormonal responses to exercise are dependent on initial training status and fitness level. Careful considerations should be given to the biological relevance of any statistically significant result.

Physical activity has not been strongly related to hormone concentrations. The only consistent relationship of training is with decreases in resting levels of oestrogen and progesterone in women and testosterone in men. It is not clear if these decreases are clinically important as the levels reported are usually within the physiologic range. It is also not clear whether all absolute levels of exercise have the same impact. However, the sympathetic and medullary responses to exercise are directly proportional to the relative intensity of the

exercise (Deuster et al, 1989). In general if individuals perform exercise bouts at an absolute intensity untrained individuals will elicit a more profound endocrine change than their trained counterparts (Luger et al, 1987). A blunted hormone response is particularly evident in endurance trained athletes (Vasankari et al ,1993).

3.4 Physical activity and the immune system:

It has long been known that the immune system plays an important role in recognising and killing tumour cells (Roitt et al, 1991). Any change in immune function could affect the body's ability to defend itself against the initiation and promotion of a tumour. During the last 95 years 629 papers (60% in the 1990s) dealing specifically with exercise and immunology have been published (Nieman, 1997).

Published papers can be divided into those that have looked at the effect of a single bout of exercise; those that have compared trained people with sedentary controls; and changing immune parameters over the course of a training programme. For the following section moderate amounts of physical activity are defined as 3-5 times per week, 15-60 minutes per session at 40-60% VO_2 max. High intensity activity is performed above 60% VO_2 max and may take place more frequently for longer periods.

Single Episode:

Acute exercise is the most frequently studied area of exercise immunology and has revealed that a rapid interchange of immune cells between peripheral lymphoid cells and the lymphoid organs occurs. The response depends on many factors including the intensity, duration and mode of exercise. Of cells involved in immune response Natural Killer (NK)

cells, neutrophils, and macrophages (of the innate immune system) appear to be the most responsive to the effects of acute exercise, both in terms of numbers and function.

NK cells NK cells possess the ability to initiate spontaneous cytolytic activity against virally infected cells and malignant cells (O'Shea & Ortaldo, 1992; Welsh & Vargas-Cortes, 1993; Strasner et al, 1997). Animal studies have shown that natural cytotoxicity is important in the immune surveillance against cancer (Garelick et al, 1982).

Despite convincing evidence from animal studies, the role of NK cells in immunologic surveillance against cancer in humans is poorly defined. Circumstantial evidence for the importance of natural cytotoxicity in humans is provided by studying NK deficiency states such as Chediak-Higashi syndrome (Roder et al, 1980), X-linked proliferative syndrome (Sullivan et al, 1980), or people who have had a renal transplant (Gui et al, 1983) and are pharmacologically immunosuppressed. These individuals have reduced NK cell activity and an increased risk of cancer.

A consistent biphasic NK cell response to acute exercise has emerged from the literature. There is a transient increase in the absolute and relative function of blood mononuclear cells expressing characteristic NK cell markers immediately after exercise, immediately followed by a drop below baseline levels within 2 hours of recovery (Brahmi et al, 1985; Pedersen et al, 1988; Nieman et al, 1991; Nieman et al, 1992; Shinkai et al, 1992). The magnitude of both phases of response seems to correlate with exercise intensity (Strasner et al, 1997). The temporary enhancement of NK activity occurs whether the exerciser is young or old, male or female, trained or untrained, fit or unfit.

There is debate as to whether the delayed decrease in total NK cell cytotoxic activity after prolonged intense exercise represents a true suppression of killing activity or is due to the redistribution of circulating lymphocyte subsets during exercise (Mackinnon et al, 1988; Berk

et al, 1990; Mackinnon, 1992; Nieman et al, 1993a). Again this effect is related to the relative exercise intensity.

The changing number of NK cells and their changing activity level could be brought about through a hormonal mechanism. NK cell activity is negatively correlated with serum cortisol levels (Nieman & Nehlson-Cannarella, 1991) so that the exercise-induced surge in cortisol secretion could contribute to the late suppression of NK cell activity. The precise physiologic role of steroid hormones on NK activity requires further clarification. Pross & Bainen (1982), Thyss et al (1984), and Holbrook et al (1983) showed that NK cell activity did not vary within the menstrual cycle and was not affected by the menstrual status of women implying no effect of female steroid hormones (Holbrook et al, 1983). In contrast Sulke et al (1985) measured suspected NK cell activity in 18 healthy women and showed a significant fall in activity in the peri-ovulatory period. However, NK cell activity was not related to oestradiol concentrations in the individual women, and oestradiol at physiologic concentrations may not directly affect the function of NK cells (Sulke et al, 1985).

Two main, and clinically relevant, questions remain to be answered: 1) Do the small exercise-induced modulations of immune cell function reach clinical relevance; 2) what amount of exercise can improve immunity without impairing it?

Neutrophils In general high intensity exercise suppresses most neutrophil functions, both acutely and chronically, while the results of moderate exercise have been conflicting. Despite the limitations of studies conducted in this field so far, there is no doubt that exercise triggers functional responses in neutrophils. The effect of duration or type of activity at a fixed intensity have received little attention. The physiological implications of exercise-induced changes in neutrophil function remains unclear but the overall trend fits in with the

epidemiological evidence that intense exercise increases vulnerability to infection while moderate exercise may be immuno-potentiating (Smith, 1997).

Conclusion In general acute exercise bouts of moderate duration (<60 minutes) and intensity (<60% $_{VO2\ max}$) are associated with fewer perturbations and less stress to the immune system and appear to transiently improve immunity. Prolonged high intensity exercise appears to put general stress on the immune system and depress immunity.

Training:

Studies of training effects on the immune system can be divided into two kinds: studies which focus on the different immune parameters in athletes and non-athletes at one point in time (cross-sectional studies); and those which follow a group of people over the course of a training programme (longitudinal studies).

Natural killer cells: The majority of cross-sectional studies support enhanced resting NK cell activity in athletes compared to non-athletes, in both younger and older age groups (Pedersen et al, 1989; Tvede et al, 1991; Nieman et al, 1993b; Nieman et al, 1995). Nieman et al (1995) found NK cell activity was 57% higher in experienced marathon runners than sedentary controls. Tvede et al (1991) found higher levels of NK cell activity in elite cyclists during the summer months (intensive training period) when compared to winter (low training).

Longitudinal studies have given inconsistent results. The paucity of longitudinal data likely reflects the confounding of training and fitness with acute exercise effects and various compliance issues inherent in longitudinal training studies. Several prospective studies utilising moderate endurance training regimes of 8-15 weeks have reported enhanced NK cell activity relative to sedentary controls (Fiatrone et al, 1989; Crist et al, 1989; Hoffman-Goetz et al, 1990; Nieman et al, 1990a; Pedersen et al, 1990). Other authors have not reported this

enhanced NK cell activity with training (Baslund et al, 1993; Nieman et al, 1993b). Indeed, Gleeson et al (1995) reported NK cell numbers to be decreased over an 8-month training season in elite swimmers. It may be that exercise has to be performed for a prolonged period of time (i.e. years) before NK cell activity is affected and this may explain the lack of an effect in studies with short training periods.

Neutrophils During periods of high intensity training neutrophil function is suppressed in athletes (Green et al, 1981; Lewicki et al, 1987; Hack et al, 1992; Hack et al, 1994; Baj et al, 1994; Pyne et al, 1995). Hack et al (1994) and Baj et al (1994) observed neutrophil function in athletes was similar to controls during periods of low training workloads, but significantly suppressed during the summer months of intensive training. As neutrophils have phagocytic activity, suppression of neutrophil function during periods of heavy training is probably a significant factor explaining increased upper respiratory tract infection (URTI) risk among athletes (Nieman, 1997).

Exercise, immune system, and URTI Two models have been put forward to explain the relation ship between infection and exercise. They are not necessarily mutually exclusive. Nieman & Nehlsen-Cannarella (1992) proposed a j-shaped curve to describe the relationship between amount of exercise and URTI. This model suggests that individual's who exercise moderately exhibit a lower incidence of URTI compared to the sedentary population; in contrast, athletes undertaking strenuous training would exhibit an increased risk. However, the precise shape and activity levels involved in this model are not known.

Pedersen & Ullum (1994) suggested an 'open window' hypothesis to describe the time period after intensive training during which the athlete had an increased risk of infection. Moderate exercise stimulates immune function during and for a short time following exercise. In contrast intensive exercise causes an initial stimulation followed by a longer lasting (i.e.

hours) suppression. It is during this open window that the athlete is at an increased risk of contracting an infection. Some elite athletes train intensively at least once a day and, thus, may spend a significant amount of time within this 'open window'.

Data from epidemiological studies on long-distance runners are consistent with a relationship between excessive training and susceptibility to URTI. Compared with non-athletes the risk of URTI is higher in athletes training for competition and competing in long distance events such as marathons or ultra-marathons (Peters & Bateman, 1983; Nieman et al, 1990a; Peters et al, 1993). The risk of URTI appears to be related to training volumes and possibly intensity.

Conclusion In response to long-term exercise training, the only finding to date reported with consistency between investigators is a significant elevation of NK cell activity. There is some evidence that neutrophil function is suppressed during periods of heavy training. Limited data suggest that unusually heavy chronic exercise may increase risk of URTI. Again, this is relatively heavy exercise i.e. an unfit person undertaking heavy exercise will be more affected than a fit person exercising at the same absolute level. However, a threshold of activity seems to be needed to cause a breakdown in immune function. The optimal level of exercise that enhances but does not suppress immune function is not yet known.

Physical activity and the immune system: Conclusion:

The immune response to exercise varies with exercise mode, intensity, and duration. Moderate amounts of exercise training has little, if any, chronic effect on immune function at rest (Nieman et al, 1990b; Nehlsen-Cannarella et al, 1991a). Thus, any protective effect of immuno-surveillance that comes from moderate exercise training is probably the result of changes during each exercise episode (Nieman et al, 1991; Nehlsen-Cannarella et al, 1991b). Moderate training reduces the amount of effect of a given bout of exercise at a relative

intensity and this decreases the likelihood of an adverse immune response (Shepherd & Shek, 1994). Athletes reveal positive chronic effects of exercise training with an enhanced NK cell activity. However, they also have a suppression of neutrophil function. Problems occur when training becomes over-training. This causes adverse immune responses and may be responsible for an increased risk of URTI.

Ideally to test the effect of regular physical activity on immune function a large group of individuals, randomised to sedentary control groups and exercise groups, should be followed for at least one year with multiple immune measures being taken before, during and after the study (Nieman, 1997). This study has not yet been performed and would require a great deal of money. At present only a few small longitudinal and cross-sectional studies are available.

On the basis of these Shepherd & Shek (1995) concluded that exercise can induce changes in the activity of the cells of the immune system. This could be relevant to HD and other cancers of the immune system.

3.5 Physical activity and Hodgkin's disease:

There has been very little work performed on the impact of physical activity on HD. Two studies have been found in the literature (Table 3.6) that suggest there may be some inverse relationship. However, these are both cohort studies with very few cases of HD. This relationship will be investigated in chapter 9 using data from the Scotland and Newcastle study of Hodgkins Disease (SNEHD). There is a possibility that moderate

Table 3.6: Hodgkin's Disease

Study	Numbers	Comparison	Risk Estimate	Significance
Whittemore et al (1985) (USA)	Cohort 51977 male college alumni & 4706 female alumni (age 35-70), 1.8m PY follow-up (52 cases)	college sport ≥5 hrs/wk vs. <5 hr/wk	RR=0.73	NS
Paffenbarger et al (1987) (USA)	Cohort 16936 college alumni (age 35-74), followed 1962/66-1978 (10 HD deaths)	post-college ltpa <500 kcal/wk 500-1999 kcal/wk ≥2000 kcal/wk	RR=0.9 RR=0.4 RR=0.3	p for trend 0.116

NS (not statistically significant) = $p>0.10$
ltpa = leisure time physical activity.

levels of physical activity may be associated with a lower risk of HD through its impact on the endocrine and immune systems while strenuous activity could increase risk.

3.6 Health Consequences of Physical Activity: Conclusion:

Physical activity has been found to confer beneficial effects on the immune system as long as activity is not performed at too high an intensity for too long. Physical activity also has an impact on the endocrine system changing the levels of various hormones, particularly lowering the resting levels of oestrogen and progesterone in women and testosterone in men. In contrast to the effect on the immune system physical activity appears to need to be quite intense for these hormonal effects to be seen. In theory these effects could have an impact on cancer risk.

Calabrase (1990) concluded that data supporting the hypothesis that exercise protects against cancer through an immune pathway was lacking, and that several lines of reasoning argue against such a relationship. Chief among these is the observation that the cancers found in immuno-suppressed individuals such as AIDS patients (lymphoreticular origin) are not the same as those affected by physical activity (colon, female reproductive, prostate). However, Hoffman-Goetz & Husted (1995) have proposed that although various exercise induced mechanical and hormonal changes best explain the change in risk of colon, breast, and prostate cancer in the physically active, several potential immunological effects may be contributing factors.

At present physical activity appears to be protective for colon cancer, although there is no agreement on how this protective effect is conferred. Data regarding a relationship between physical activity and breast, endometrial, ovarian, prostate, testicular cancers, and Hodgkin's disease are too limited and inconsistent to support any firm conclusions.

4. Data for Analysis

The results presented in the following five chapters are based on data from three separate sources.

4.1 Scandinavian Cancer Registry:

The analysis of the seasonal presentation of HD in Chapter 6 is based on data from four Scandinavian cancer registries (Denmark, Finland, Norway, and Sweden). The data consisted of over 2000 cases of HD aged 0-34 years with details of age, sex, and month and year of HD registration provided. I had no role in the collection of these data.

4.2 Young Adult Hodgkins Disease Case-Control Study (YHHCCS):

YHHCCS is a matched case-control study of HD in young adults residing in the counties of West Yorkshire, Avon, Cornwall, Cumbria, Devon, Dorset, Gloucestershire, Humber, Lancashire, North Yorkshire, and Somerset. Eligible cases were those aged 16-24 years, diagnosed 1/10/1991 to 31/5/1995, living in the parts of the counties listed above included in the Leukaemia Research Fund (LRF) Data Collection Study (DCS) area (Cartwright et al, 1990). Controls were matched on age, sex, and administrative area (FHSA) of residence with the aim of having two controls per case. The matching criteria was fulfilled in most cases, the only exceptions being one set with only one control and two sets with three controls. Controls that did not give consent or who could not be traced were replaced with further random selections based on the same matching criteria.

The fieldwork for this study was the responsibility of the LRF centre and I had no role in the data collection. Reports of my data analysis are presented in Chapter 7.

4.3 Scotland and Newcastle Study of Hodgkins Disease (SNEHD):

SNEHD is a study of HD in adults residing in the Northern health region of England and all of Scotland with the exception of the Western Isles and Dumfries and Galloway. Eligible cases were those aged 16-74 years on 1/1/1993 with HD diagnosed 1/1/1993 to 31/3/1997 living in Newcastle/North Tyneside, Gateshead/South Tyneside, Cumbria, Northumberland, Durham, Sunderland, and Cleveland FHSAs and Lothian, Greater Glasgow, Argyll & Clyde, Borders, Fife, Forth Valley, Grampian, Highland, Tayside, Lanarkshire, and Ayrshire & Arran health boards. Controls in 10 year age groups were selected randomly from FHSA and health board lists (for more details of control selection see Chapter 5). Controls that did not give consent or who could not be traced were replaced with further random selections from the same GP practice based on the same sex and age-group criteria.

The aim of the SNEHD study was to identify risk factor profiles for HD in total and also for HD subgroups classified by age group at diagnosis and EBV in the tumour cells. The fieldwork for this study was the responsibility of Dr. Freda Alexander in Edinburgh and a team of six interviewers (of which I was one) and support staff. I was involved in the writing of the sport and athletics section of the questionnaire (based on the research presented in Chapter 2) and the piloting of the total questionnaire. I was responsible for the calculation of control numbers and the administrative work involved in control selection and contact. I analysed the results of general and family health sections with advice from Dr Alexander (Chapter 8) and also the section on sport and athletics without advice from Dr Alexander (Chapter 9).

5 SNEHD Control Selection

5.1 SNEHD Control Number Calculation:

The age distribution of HD cases is bimodal (see Chapter 1) whereas the age distribution of the general population is closer to a normal curve. Therefore, in order to take account of this difference the number of controls required for SNEHD was to be frequency matched with the number of cases by area (FHSA)/Health board), age group, and sex. To estimate expected numbers I combined the data for HD registration for the last year data were available by age and sex (1989 for England & Wales, 1990 for Scotland) with the populations for the same years. This allowed the calculation of an approximate HD rate per 10^5 for the UK by age and sex (Table 5.1).

These rates were then applied to the most up-to-date population estimates available for the SNEHD region (estimates for England provided by the Office Population Censuses and Surveys, for Scotland by the General Registry Office for Scotland). Approximate expected number of cases arising in the SNEHD region each year were calculated (Table 5.2). Multiplying this annual number of cases by four (the planned number of years of the study) gives a figure of around 630 expected HD cases. Thus, 630 controls would be required for SNEHD.

The expected age distribution of the cases is shown in Table 5.3 and reveals evidence of the bimodality of the distribution of HD and the greater frequency of males over female cases. The expected age distribution of controls based on a standard calculation (630/N where N is the number of people in each age/sex/area group) are also in Table 5.3. There are too many female controls in all age groups over 24 years and not enough male controls to allow a fine level of analysis. Thus, the aim of

Table 5.1. Approximate rates of Hodgkin's Disease in England & Wales and Scotland by 10-year age group.

	Age Group						
Sex	15-24	25-34	35-44	45-54	55-64	65-74	TOTAL
Male	3.83	3.83	3.58	2.98	2.79	4.07	2.88
Female	3.53	2.29	1.66	1.12	1.69	2.35	1.88

Table 5.2. Approximate expected number of cases per year in the SNEHD study region by 10-year age group.

	Age Group						
Sex	16-24	25-34	35-44	45-54	55-64	65-74	TOTAL
Male	18.1	23.5	18.8	13.9	10.7	12.5	97.5
Female	15.9	13.9	8.7	5.3	7.0	9.1	59.9

Table 5.3. Expected numbers of cases in the SNEHD region and the number of controls using the original formula by 10-year age group.

	Age Group					
	16-24		25-34		35-44	
Sex	Cases	Controls	Cases	Controls	Cases	Controls
Male	72.4	58.4	94	76.5	75.2	64.9
Female	63.6	55.9	55.6	75.1	34.4	65.5
	Age Group					
	45-54		55-64		65-74	
Sex	Cases	Controls	Cases	Controls	Cases	Controls
Male	55.6	58	42.8	47.8	50	38.3
Female	21.2	58.7	28	51.6	36.4	48.4

frequency matching the cases to the controls is not met by this basic calculation. The distribution of controls, therefore, was adjusted to make sure that the frequency of controls in each age/sex group was as close as possible. Once this adjustment was made the number of controls required was very close to the expected number of cases.

However, a few further adjustments to the control numbers were made. Only 12 controls for Orkney and Shetland combined were required and these were randomly allocated to Orkney or Shetland and to age/sex groups. It was decided that the number of selected controls was too similar to the number of expected cases and to avoid the risk of having too few controls recruited in any age/sex group due to refusals the number of controls in Scotland was inflated by 15%. One year into the study Ayrshire and Arran was included. Therefore, control numbers for Ayrshire and Arran were calculated by the same method as before. The final number of controls required in each age/sex group after all the adjustments are shown in Tables 5.4 (Scotland) and 5.5 (England).

Overall, the total number of controls needed for SNEHD was calculated to be 750.

5.2 Patient details:

Control information for England and Scotland came in different forms.

England:

Each Family Health Service Authority (FHSA) covered by the study (Cleveland, Cumbria, Newcastle/North Tyneside, Gateshead/South Tyneside, Sunderland, Northumberland and Durham) sent a tape with a 1 in a 100 random sample of patients from their lists. Random numbers were then generated and controls were picked from the sample until the required numbers in each age/ sex group had been filled. Random numbers also used to generate a year of pseudo-diagnosis (1993-1996) for each control.

Table 5.4. Total control numbers required for Scotland including 15% increase and Ayrshire & Arran.

	Age Group						
Sex	16-24	25-34	35-44	45-54	55-64	65-74	TOTAL
Male	57	75	60	44	32	37	305
Female	53	42	28	20	24	25	192

Table 5.5. Total control numbers required for the North of England.

	Age Group						
Sex	16-24	25-34	35-44	45-54	55-64	65-74	TOTAL
Male	26	38	31	22	18	20	155
Female	24	20	14	12	12	16	98

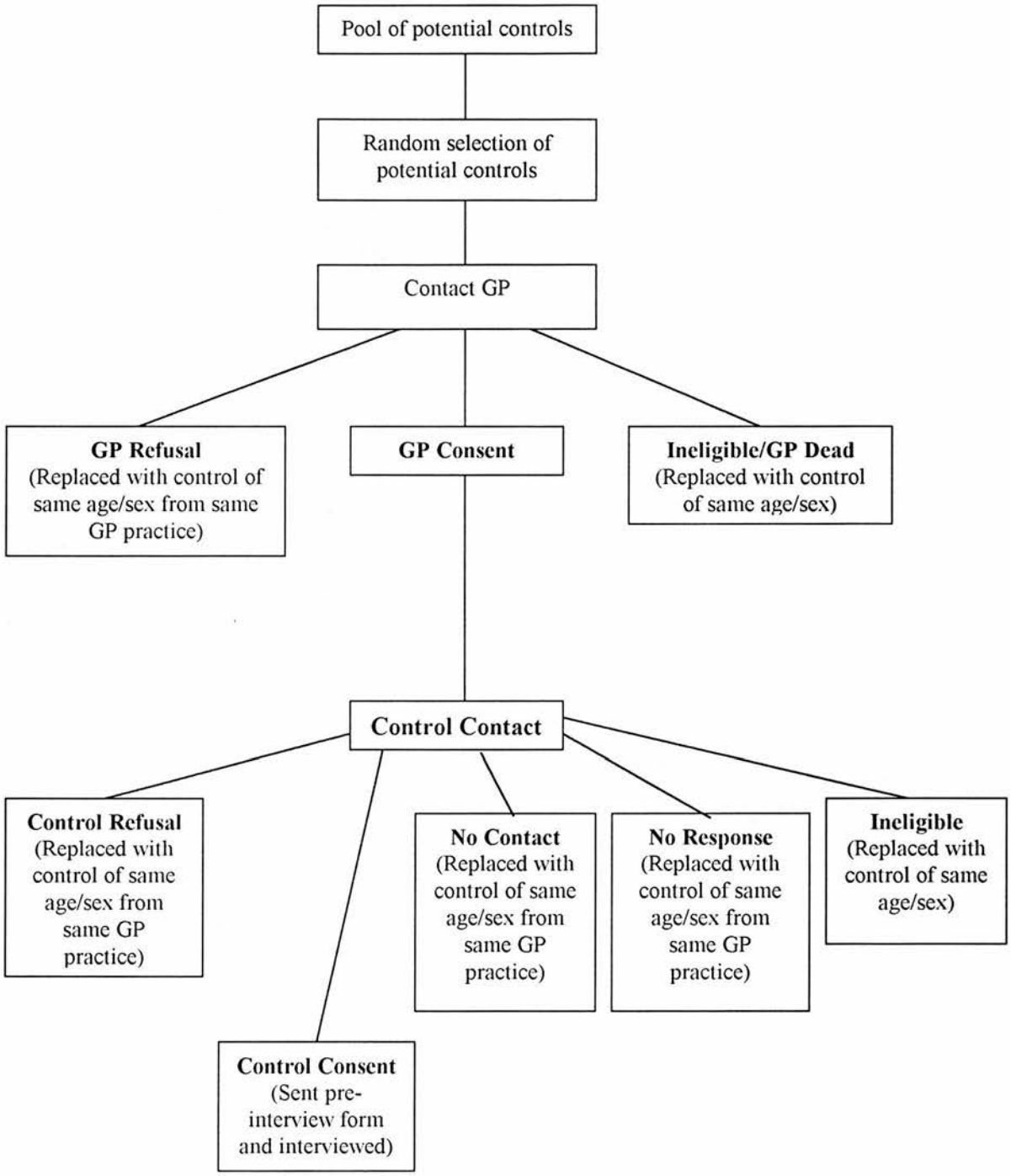
Scotland:

The method of extracting patient details in Scotland was different to that for England. Random numbers were then used to calculate a date of pseudo-diagnosis (1993-1996). These pseudo-diagnosis dates were applied to the overall control numbers giving the required controls for each year of pseudo-diagnosis in each age/sex group in each health board. A computer programme written by the Common Services Agency of the NHS in Dundee selected patients at random from each health board list with specific details supplied from the study centre (Edinburgh). Requests for controls were sent to each individual health board in Scotland (Borders, Lothian, Lanarkshire, Argyll & Clyde, Greater Glasgow Health Board, Fife, Forth Valley, Tayside, Grampian, Highland, Orkney, Shetland and later Ayrshire and Arran) approximately every 3 months from March 1996 until the complement of controls was filled. Control details were sent to Edinburgh and then mounted onto Excel.

Control Approaches:

The stages of contacting a potential control are shown in Figure 5.1. No control was contacted without the permission of their GP. If the GP refused permission to contact the control was replaced with a control of the same age/sex group from the same GP practice to avoid geographical SES bias. (It was only possible to replace control refusals from the same GP practice in Scotland as the format of the English data prevented this.) If the GP provided information that the potential control was ineligible (dead, not aged 16-74 years, not born in the UK, not mentally competent to perform the interview) the control was replaced with another from the same age/sex group. If GPs did not reply in 2-3 weeks a reminder letter was sent with a further letter 2-3 weeks after that if a reply had still not been received. GPs were also contacted by telephone.

Figure 5.1. Flow diagram of contact route for potential controls.



For definitions of bold text see Table 5.6

Once the GP had consented one of two types of letter were sent out to the potential control. The SNEHD study compared two methods of approaching controls. One method was more formal and was signed by Dr. Alexander (office contact), and the other was slightly more friendly and informal and signed by the interviewer for that control's region (interviewer contact). The two methods were assigned randomly. *The design and conduct of this randomised study was my responsibility.* Simple randomisation could have been used, but by chance this may lead to long runs of the same contact type. To avoid this and to ensure that the number of controls in each contact group was fairly similar at all times block randomisation was used with a block size of four. This ensures that of every four controls, two would be contacted using the office letter and two with the interviewer letter. With two possible letters (I interviewer contact; O office contact) and a block size of four there are six possible combinations: OOII, OIOI, OIIO, IIOO, IOIO, and IOOI. The order of the block was assigned using random numbers.

A potential control was given 2-3 weeks to reply to our letter, after which their address was checked by telephoning or faxing the health board/FHSA in which they were registered. If a telephone number was available an attempt was made by a research interviewer to contact them by telephone. This was a quick method for getting a response (positive or negative) from the potential control.

If the address was correct and telephone details were not available, a reminder letter was sent (again from the interviewer or the office). If the potential control again did not reply after a further 2-3 weeks a second reminder letter was sent out. After three letters in total most people had replied. For those people who had not replied the interviewer for their area, where feasible, made a visit in person (cold calling). If cold calling was unsuccessful the control was coded 'No Response' on our database and treated as a refusal for replacement purposes.

Every effort was made to contact the potential control but controls could only be traced if they were registered with a GP. If the patient had moved and not re-registered there was no other way of contacting that person and they were entered on the database as 'No Contact'.

5.3 Control Response Rate Methods:

Controls were contacted using the methods described above and assigned a code depending on their response (Table 5.6). At the end of the study period characteristics of the controls who agreed to take part (DONE) were compared with those who did not (all other codes except INELIGIBLE). The first step was to analyse the proportion of positive responses for first choice controls only and the total series (including replacements). As mentioned above two methods of control contact were used to see if either had a better effect on control response. Control response is also likely to affected by demographic variables (age, sex, Carstairs index of address at contact, and country). The proportion of positive responses were calculated with and without the inclusion of controls considered to be 'No Contact', as it could not be confirmed that these controls were eligible. The proportions of positive case responses are also given including and excluding 'No Contact'.

Once the proportion of positive responses had been calculated in subgroups based on demographic variables and contact type more formal statistical methods were used to compare those who agreed to take part with those who did not. Logistic regression analyses were performed on the dependent variable YES/NO (i.e. control agreed to take part or not) using the following explanatory variables: sex, area (England or Scotland), contact (Interviewer or Office) (all dichotomous variables), age group (both as a trend and as 6 10 year categories), Carstairs index (both as a trend and 7 categories). Univariate logistic

Table 5.6. Control status codes used for SNEHD.

Code	Definition
GP CONTACT	One or more letters sent to GP
GP CONSENT	GP consents
GP DEAD	GP reports control dead (INELIGIBLE)
GP REFUSAL	GP refuses permission to contact
CTRL CONTACT	Control approached
CTRL CONSENT	Control consents, interview pending
CTRL DEAD	Control found to be dead (INELIGIBLE)
CTRL REFUSAL	Control refuses
DONE	Interview done
INELIGIBLE	Control dead, wrong age, not competent
NO CONTACT	Not possible to find the potential control
NO RESPONSE	Control address correct, no reply

regression analyses were performed first. Those variables that were statistically significant were then put in multivariate models.

5.4 Results:

The overall proportion of positive responses for cases and controls are shown in Table 5.7. It can be seen that they differ, with approximately 75% of eligible cases taking part compared with less than 50% of first choice controls. Table 5.7 also makes it clear that the number of cases and controls considered to be 'NO CONTACT' was different. Only 8 cases (1.2%) were so categorised compared with 57 controls (7.7%). The proportion of male and female cases that could not be contacted was approximately the same (7.7% and 7.4% respectively). However, the proportion did vary by the age of the control. The majority of controls who could not be contacted were in the younger age groups (age 16-24 and 25-34). By age group the 'NO CONTACT' rates were: age 16-24 11.1%, age 25-34 10.1%, age 35-44 6.8%, age 45-54 5.3%, age 55-64 3.1%, and age 65-74 yrs 4.0%. The youngest age groups are more mobile and may be less likely to re-register with GPs.

The control positive response varied by sex, with women more likely to agree than men (55.6% vs. 45.9%). This was not the case for cases with approximately equal proportions of males and females agreeing to participate. When looking at positive responses by age the youngest (age 16-24 years) and oldest (age 65-74 years) age groups were less likely to agree to take part with the other age groups having similar rates. The cases had a smooth falling off of proportion of positive responses with age. In the controls an interesting pattern is revealed. The youngest males and the oldest females are the least likely to agree to take part (27.8% and 37.9% respectively). Indeed, the oldest age group is the only one in which the proportion of positive female responses is lower then the male. This pattern was not seen for the cases.

Table 5.7. Response rates for cases and first choice controls by demographic variables and contact type.

Comparison		Case (inc. 'No Contact)	Case (not inc. 'No Contact')	Control (inc. 'No Contact')	Control (no inc 'No Contact')
Total		493/666 (**74.0**)	493/658 (**74.9**)	340/739 (**46.0**)	340/682 (**49.9**)
Female		208/277 (**75.1**)	208/274 (**75.9**)	153/297 (**51.5**)	153/275 (**55.6**)
Male		285/389 (**73.3**)	285/384 (**74.4**)	187/442 (**42.3**)	187/407 (**45.9**)
Age	16-24	115/145 (**79.3**)	115/143 (**80.4**)	79/208 (**38.0**)	79/185 (**42.7**)
	25-34	123/158 (**77.8**)	123/156 (**78.8**)	76/149 (**51.1**)	76/134 (**56.7**)
	35-44	71/87 (**81.6**)	71/85 (**83.5**)	49/118 (**41.5**)	49/110 (**44.5**)
	45-54	64/80 (**80.0**)	64/79 (**81.0**)	53/93 (**57.0**)	53/88 (**60.2**)
	55-64	59/87 (**67.8**)	59/87 (**67.8**)	53/96 (**55.2**)	53/93 (**57.0**)
	65-74	54/90 (**60.0**)	54/90 (**60.0**)	30/75 (**40.0**)	30/72 (**41.7**)
Male	16-24	64/81 (**79.0**)	64/79 (**81.0**)	27/107 (**20.6**)	27/97 (**27.8**)
	25-34	69/90 (**76.7**)	69/90 (**76.7**)	48/95 (**50.5**)	48/83 (**57.8**)
	35-44	45/57 (**78.9**)	45/55 (**81.8**)	31/81 (**38.3**)	31/74 (**41.9**)
	45-54	41/53 (**77.4**)	41/52 (**78.8**)	33/61 (**54.1**)	33/56 (**58.9**)
	55-64	42/55 (**76.4**)	42/55 (**76.4**)	29/54 (**53.7**)	29/54 (**53.7**)
	65-74	21/40 (**52.5**)	21/40 (**52.5**)	19/44 (**43.2**)	19/43 (**44.2**)
Female	16-24	51/64 (**79.7**)	51/64 (**79.7**)	52/101 (**51.5**)	52/88 (**59.1**)
	25-34	54/68 (**79.4**)	54/66 (**81.8**)	28/54 (**51.9**)	28/51 (**54.9**)
	35-44	26/30 (**86.7**)	26/30 (**86.7**)	18/37 (**48.6**)	18/36 (**50.0**)
	45-54	23/27 (**85.2**)	23/27 (**85.2**)	20/32 (**62.5**)	20/32 (**62.5**)
	55-64	17/32 (**53.1**)	17/32 (**53.1**)	24/42 (**57.1**)	24/39 (**61.5**)
	65-74	33/50 (**66.0**)	33/50 (**66.0**)	11/31 (**35.5**)	11/29 (**37.9**)
England		170/219 (**77.6**)	170/217 (**78.3**)	113/241 (**46.9**)	113/229 (**49.3**)
Scotland		323/447 (**72.3**)	323/441 (**73.2**)	227/498 (**45.6**)	227/453 (**50.1**)
Office letter		NA	NA	172/369 (**46.6**)	172/342 (**50.3**)
Intrviewr letter		NA	NA	168/370 (**45.4**)	168/340 (**49.4**)
Carstair's	1	NA	NA	23/38 (**60.5**)	23/35 (**65.7**)
	2	NA	NA	34/68 (**50.0**)	34/58 (**58.6**)
	3	NA	NA	70/140 (**50.0**)	70/130 (**53.8**)
	4	NA	NA	68/133 (**51.1**)	68/124 (**54.8**)
	5	NA	NA	67/146 (**45.9**)	67/138 (**48.6**)
	6	NA	NA	42/106 (**39.6**)	42/96 (**43.8**)
	7	NA	NA	22/61 (**36.1**)	22/55 (**40.0**)

There was little difference in the proportion of positive responses for cases or controls by country (78.3% versus 73.2% and 49.3% versus 50.1% in England and Scotland respectively). There was also little difference in the response generated by the two approach letters. 50.3% of first choice controls who received the formal letter from Edinburgh agreed to take part compared with 49.4% of controls who received the letter signed by an interviewer.

There was a decreasing likelihood of potential controls agreeing to take part with lower socio-economic status (as assessed by Carstair's index of address from health board). (Response by Carstair's index was not calculated for cases as postcode was missing for 379 (56.9%) of them.) The most affluent (Carstair's index 1) were the most likely to agree to take part (65.7%) and the least affluent (Carstair's index 7) were the least likely (40.0%). This socio-economic pattern was consistent across sexes and also for the type of contact letter received. It was not consistent across countries. In Scotland the most and least affluent groups were the most likely to agree to take part with the lowest response rates in the middle groups. In England there was no consistent socio-economic pattern but the numbers in the denominators were very small. It was not possible to say much about the effect of socio-economic status by age as, again, the denominator numbers were very small.

The results of the logistic regression analyses to find variables that predict a control's willingness to take part are shown in Tables 5.8 (first choice controls) and 5.9 (total controls). These results tend to support the observations of the response rates. When first choice controls only are examined the significant predictors in univariate analysis were: Carstair's index (as a trend, p=0.004), age group (as a categorical variable, p=0.005), and sex (p=0.014). The same variables are also significant predictors for the total control series. However, in this analysis the area variable was also significant. This is not surprising

Table 5.8. Results of regression models for variables affecting positive response (1st choice only).

Model	-2logl	diff dev	df	p
Univariate models				
constant	1019.76	---	---	---
constant + SEX	1013.70	6.06	1	0.01
constant + CONTACT	1019.65	0.11	1	0.74
constant + AREA	1019.65	0.11	1	0.74
constant + AGRGROUP (trend)	1016.68	3.08	1	0.08
constant + AGRGROUP (categorical)	1002.98	16.78	5	0.005
constant + CARSTAIRS (trend)[a]	948.82	8.18	1	0.004
constant + CARSTAIRS (categorical)[a]	947.163	9.839	6	0.13
Multivariate models				
constant + AGE GROUP (cat) + CARSTAIRS (trend)[b]	934.56	14.26	5	0.01
constant + AGE GROUP (cat) + CARSTAIRS (trend) + AGE GRP*CARSTAIRS interaction[c]	928.65	5.91	5	0.32
constant + AGE GROUP (cat) + CARSTAIRS (trend) + SEX[c]	924.39	10.17	1	0.001
constant + AGE GROUP (cat) + CARSTAIRS (trend) + SEX + AGE GRP*SEX interaction[d]	914.08	10.31	5	0.07
constant + AGE GROUP (cat) + CARSTAIRS (trend) + SEX + CARSTAIRS*SEX interaction[d]	923.55	0.84	1	0.33

a. Compared with model of constant only for those not missing Carstair's index base equals 957.00. Differences are from this base rather than 1019.76.
b. Compared with model of constant + CARSTAIRS (trend) base of 948.82.
c. Compared with model of constant + AGE GRP (cat) + CARSTAIRS (trend).
d. Compared with model of constant + AGE GRP (cat) + CARSTAIRS (trend) + SEX.

Table 5.9. Results of regression models for variables affecting positive response (Total series)

Model	-2logl	diff	df	p
Univariate models				
constant	1713.673	---	---	---
constant + SEX	1702.442	11.231	1	0.0008
constant + CONTACT	1713.668	0.005	1	0.94
constant + AREA[b]	1708.913	4.759	1	0.03
constant + AGRGRP (trend)	1712.390	1.282	1	0.26
constant + AGRGRP (categorical)	1700.150	13.522	5	0.02
constant + CARSTAIRS (trend)[a]	1590.512	7.733	1	0.005
constant + CARSTAIRS (categorical)[a]	1589.165	9.079	6	0.17
Multivariate models				
constant + AREA + AGE GRP (cat) + CARSTAIRS (trend)[b]	1569.54	9.11	5	0.11
constant + AREA + AGE GRP (cat) + CARSTAIRS (trend) + AGE GRP*CARSTAIRS interaction[c]	1563.96	5.59	5	0.35
constant + AREA + AGE GRP (cat) + CARSTAIRS (trend) + SEX[c]	1551.94	17.61	1	<0.0001
constant + AREA + AGE GRP (cat) + CARSTAIRS (trend) + SEX + AGE GRP*SEX interaction[d]	1535.34	16.60	5	0.005
constant + AREA + AGE GRP (cat) + CARSTAIRS (trend) + SEX + CARSTAIRS*SEX interaction[d]	1551.60	0.33	1	0.56

a. Compared with model of constant only for those not missing Carstair's index base equals 1598.24. Differences are from this base rather than 1713.67.
b. Compared with model of constant + CARSTAIRS (trend) base of 1590.51.
c. Compared with model of constant + AGE GRP (cat) + CARSTAIRS (trend).
d. Compared with model of constant + AGE GRP (cat) + CARSTAIRS (trend) + SEX.

as there were many more controls selected for Scotland. (Area was kept in the multivariate models as it had been statistically significant, although meaningless in terms of prediction). In the total series age group does not result in a statistically significantly better fitting model when added to a model including Carstair's index and area. However, sex does have a statistically significant result. Of the statistically significant individual variables the only statistically significant interactions were for age group and sex in the total series ($p=0.0053$). The interaction of age group and sex was almost statistically significant when looking at the first choice controls alone ($p=0.07$).

5.5 Discussion:

The term response rate is generally used to show the proportion of people agreeing to take part in a study, although this is not strictly speaking a rate. However, the proportion of positive responses will be referred to as a response rate for this discussion. The calculation of response rates is important as they provide an estimate of the proportion of the targeted population who participate (Groves & Lyberg, 1988). The success of any survey depends on achieving a high response rate. Thus, the biggest threat to the validity of a case-control study is non-participation of subjects and the resulting potential for selection bias (Austin et al, 1994). Although the cases have the condition of interest, it is the controls that constitute the key methodological part of the case-control method (Janerech et al, 1979).

Cases and controls should be "representative of the same base experience" (Meittenen, 1985). The base is the set of people in which diseased subjects become cases. Typically in chronic disease epidemiology membership of the base is dynamic, in the sense that the subject may be in the base at certain times and out of it at other times (Wacholder et al, 1992). The simplest way to satisfy this principle is to choose a random sample of individuals from the same source as the cases. In simple random sampling controls are

selected randomly from the base. Therefore, each eligible member has the same probability of selection as a control, and the sampling is independent i.e. the presence of a specific subject in the sample does not make the presence of any other more or less likely. This was the original method used for SNEHD but it did not result in a frequency-matched distribution of cases and controls. SNEHD control numbers were calculated using stratified sampling and frequency matching, the base being sub-divided up into strata determined by age, sex, and area with the sampling fraction allowed to vary across strata.

Participation rates in case-control studies are frequently low with the different results for cases and controls, as is the case for SNEHD. Low participation rates will create selection bias if participants and non-participants have a different exposure distribution and if participation rates differ for cases and controls (Austin et al, 1994). The potential for such bias increases in relation to the proportion of non-participants. This can be a great problem when, as in SNEHD, population controls are used. However, bias can occur even if the case and control rates are similar, since the reasons that cases and controls agree to take part, or refuse to, may differ and may relate to the exposure of interest (Austin et al, 1994).

Hartge et al (1984) reported the control response rate for four case-control studies in the United States that used random digit dialling (RDD). The household screening response rates were 80%, 86%, 88%, and 89%. The corresponding subject interview response rates were 78%, 72%, 84%, and 78%. The overall response rate is the product of these two values and ranged from 62-74%. These overall response rates are quite low. Stech (1981) has documented a considerable increase in refusal rates 1952-79 in survey research. Several studies conducted in the mid-1980s have reported even lower response rates. The Male Breast Cancer Study (Thomas et al, 1992) reported a 65% response rate for RDD screening and a 69% response rate for the field phase, giving an overall response rate of 45%. In the Oral and Pharyngeal Study (Blot et al, 1988) the response rate for the screening phase of RDD was

79% and the response rate for the field phase was 73%, giving an overall response rate of 57%.

Slattery et al (1995) surveyed 102 epidemiologists and biostatisticians involved in cancer epidemiology research. 61% replied to queries about response rates. Despite the method of response rate calculation, use of incentives, and the length of the interview, over 60% of the researchers who replied believed that it was more difficult to obtain high response rates in 1993 compared to 5 years previously.

Low control participation rates are a particularly serious threat to validity. The reasons for which controls agree to participate are unclear, but may differ to those of cases. Some may be motivated by an increased sense of altruism. Some studies have shown that non-respondents are less healthy than respondents are (Doll & Hill, 1964; Criqui et al, 1978), and smoke more (Criqui et al, 1978). Another study has found that respondents were more likely to have a history of screening procedures (Olsen et al, 1992). Non-respondents in telephone surveys are frequently less educated, lower in social class and older than respondents (O'Neil, 1979). Austin et al (1981), as part of a heart disease risk factor program, collected information from subjects who refused to participate as well as those who did. They demonstrated that the most common response pattern was that subjects with risk factors tended to participate more often than those without risk factors, and subjects with disease tended to participate less often than those without.

There are not only problems associated with people who refuse to participate but also with those who cannot be contacted for any reason (labelled as 'NO CONTACT' in SNEHD). Thus, some of the non-response may be due to letters that never reach study subjects rather than non-participation. To maximise response rates it is important to try as hard as possible to find subjects. In general, there is a clear danger of bias resulting from less successfully tracing migrant subjects, because migration and health are related. In British mortality data it has

been shown that those who move longer distances have lower mortality, suggesting a 'healthy migrant' effect not unlike the 'healthy worker' effect (Fox & Goldblatt, 1982). When accurate addresses are available for each study subject investigators can be reasonably confidant that letters that are not returned are due to non-response on the part of study subjects (Sandler & Holland, 1990). However, if an address list is no longer current the investigator has no assurance that the letter has reached the intended subject, unless it is returned. High rates of mobility, divorce, unemployment, and inaccurate data sources can combine to render follow-up costly and uncertain (Alderman, 1996).The people who are lost may or may not be eligible to take part in the study. In SNEHD everything possible was done to trace non-responders including telephoning, if possible, checking addresses with health boards and checking health board registrations. However, if a potential control had moved and failed to re-register with a new GP there was no further way of contacting them.

To have a response rate that reflects a representative sample of the population it is necessary to include individuals who require more effort to locate as well as those who are easily located. However, in general those with unknown eligibility are often excluded from the denominator of any calculations, leaving what Slattery et al (1995) have described as a 'co-operation rate' rather than a response rate. This was the main reason for calculating the SNEHD response rates twice, once including 'NO CONTACT' and once without.

Attention has focused on methods for encouraging potential controls to agree to take part. A potential mechanism to improve study response is the use of incentives. Holt et al (1991) found that 134 of 177 women who had originally refused to take part in a study agreed to do so after re-contact and the offer of $10 for their time. However, the use of incentives would hugely increase the cost of studies and is not frequently used in the UK and has possible ethical problems.

The conduct of population-based case-control studies requires special care to make sure that response rates are high and comparable for cases and controls (West et al, 1984). The need to assess comparability and to minimise misclassification, non-response and observation bias is clear, but the techniques for doing so are not. The problem is unresolvable short of obtaining information on the exposure distribution of the non-participants, an impossible task in most situations. Only a few variables were available for the SNEHD controls to compare those that did agree to take part with those that did not. Response rate differed by age, sex, and socio-economic status (based on the Carstair's index). These factors can be adjusted for in the analysis, but these are only basic things. In the case of SNEHD some ethical committees requested the removal from control literature of anything which might 'put pressure' on the potential. The study did as much as possible to get people to take part and trace those who had moved but even so had quite poor control response rates.

6 Seasonality of Hodgkin's Disease.

6.1 Introduction:

Speculation about the possible infectious nature of Hodgkin's Disease (HD) began soon after its original description by Thomas Hodgkin in 1832. There is a variety of evidence that HD could have an infectious origin including clustering investigations, HD's relationship with EBV, and the seasonal presentation of HD. The majority of these were discussed in detail in chapter 1.

If the aetiology of HD involves recent exposure to an infectious agent or allergens one might expect a seasonal variation in onset, depending on the length and variability of the latent period and the method of transmission. However, little attention has been paid to the observation first made by Cridland in 1961 that there is seasonal variation in the incidence of HD. Although, some investigators have reported fluctuations in onset by month, the peak being in the northern hemisphere late winter/ early spring (Cridland, 1961; Innes & Newell, 1961; Fraumeni & Li, 1962; Newell et al, 1985; Neilly et al, 1995a & b; Douglas et al, 1996; Westerbeek et al, 1998; Douglas et al, 1998), others have not confirmed this (Bjelke, 1969; Modan et al, 1969; Newell, 1972).

This study uses data from the Danish, Finnish, Norwegian and Swedish cancer registries in order to search for seasonal patterns in the occurrence of HD over 1, 2, 3, and 4 years. This is the first study to look at seasonal effects over a period greater than a year. This may be important because several viruses have epidemic patterns repeating over periods of several years.

6.2 Methods:

Source of cases Incident HD case data for Denmark, Finland, Norway, and Sweden were supplied by the national cancer registries. For each case of HD the month and year of birth and the month and year of diagnosis was provided; from these the age at diagnosis could be calculated. The dates of diagnosis varied by country: Denmark provided high quality data for 1978-92; Finland 1953-93; Norway 1980-92; and Sweden 1980-91.

For each country cases aged 0-34 years were used as a single series and in later analyses divided into two (0-14, 15-34 years) or three (0-9, 10-19, 20-34 years) age groups. This second method was used because some authors have found that the difference between paediatric and young adult HD in terms of EBV-association focuses on children under age 10 (Razzouk et al, 1997).

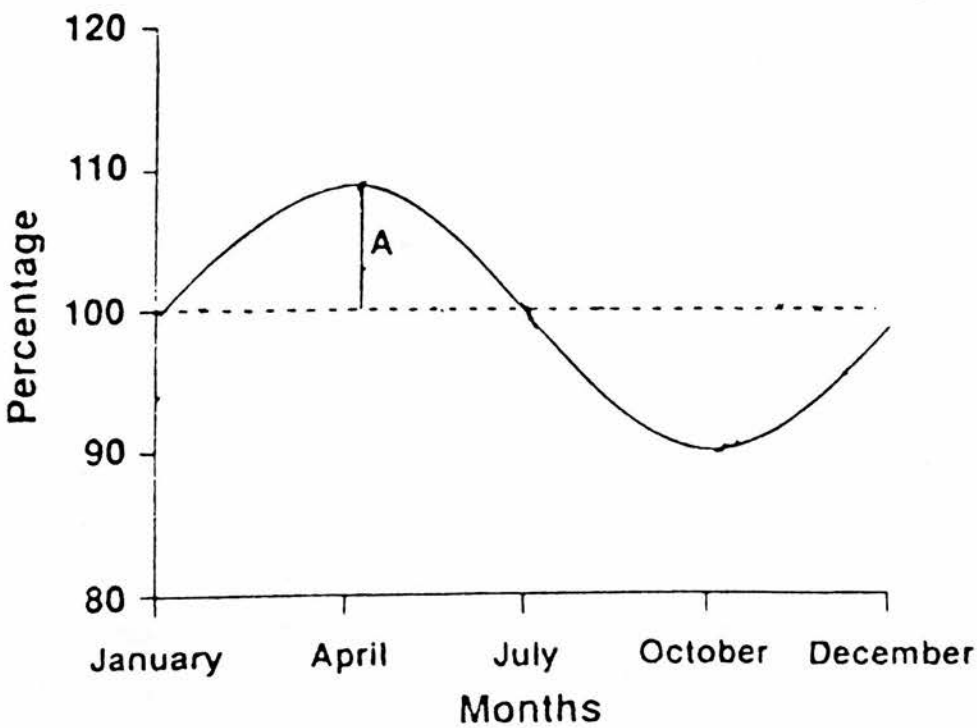
Statistical and Computational Analysis A separate Excel spreadsheet was constructed for each country to record the distribution of observed HD cases by month of diagnosis. Standard age-specific reference rates were computed using rates taken from Parkin et al (1992); these were the average of the rates for each of the four countries in each of seven age groups: 0-4, 5-9, 10-14, 15-19, 20-24, 25-29, 30-34 years. Expected numbers of cases for each month in each country were calculated by applying the reference rates to person-years at risk derived from census data (Denmark, Norway, Sweden) or annual population estimates for the whole period (Finland). The spreadsheet was completed with values for the calendar year and sine and cosine values corresponding to 1, 2, 3, and 4 year cycles i.e. one year = $360^{o}/n$ (n=1, 2, 3, 4). Each month is then assigned an angular value e.g. for a one year cycle January=0^{o}, February=30^{o} ... December=330^{o}. Three separate spreadsheets were constructed: one with the total observed and expected cases of HD; one with the cases divided into two age groups (0-14, 15-34 years); and one with the cases divided into three age groups (0-9, 10-19, 20-34 years). A total series was also constructed

with the records from Denmark, Finland, Norway, and Sweden combined and aligned across calendar year.

The statistical model for the cyclic pattern over n years was log (O/E) = $(\sin + \cos)_n$ + ε where ε is the random error. Seasonal analysis of monthly frequencies of disease diagnosis was performed using Poisson regression with the link function = log. The dependent variable was the observed number of cases of HD, the explanatory variables were annual trend, sex, age group and the sine and cosine values for 1, 2, 3, and 4 year cycles $(\sin + \cos)_n$ $(n=1,2,3,4)$. (For example, $(\sin + \cos)_1$ is shorthand for entering sine and cosine values for a one year cycle into the regression model.) The offset was the expected number of cases. This was implemented using SAS 6.11 PROC GENMOD. Univariate analyses were used to find the model that best explained the distribution of observed cases using analysis of deviance with asymptotic chi-square distribution applied to the difference in deviances in nested models. The comparison involved the addition of $(\sin + \cos)_n$ to the null model of trend+sex. The differences in deviance were always tested on 2 degrees of freedom. $(\text{Sin} + \cos)_n$ were always included in the models as a pair rather than individually since no prior hypothesis specified the zero of the cycle. Once the best model was found the sine and cosine coefficients could be used to calculate where the series peak lay (the acrophase) and also the magnitude of the seasonal fluctuation, the amplitude (Figure 6.1).

The effect age group was then investigated as well as the interaction of sex and age group with $(\sin + \cos)_n$ for all four countries individually and the total series.

Figure 6.1: Example seasonal curve showing the amplitude A and the position of the peak of the curve. On the vertical axis the monthly data are given as a percentage above and below the mean.



6.3 Results:

Between 1978-92 there were 802 registrations of HD in Denmark, 773 registrations 1980-91 in Sweden, 1733 registrations 1953-93 in Finland, and 487 1980-92 in Norway. These data can be seen in Tables 6.1 to 6.4. Data from each country has the young adult peak seen usually in HD and the rates are similar to those found in Parkin et al (1992).

The results of the step up models for the four countries and the total series can be seen in Tables 6.5 to 6.9. For Denmark the best model was that with $(\sin + \cos)_1$ included; indeed this was the only seasonal effect. In Finland, $(\sin + \cos)_1$ was again statistically significant, as were $(\sin + \cos)_2$. $(\text{Sin} + \cos)_4$ was statistically significant in Norway while Sweden was the only country not to show any evidence of seasonality. When the four countries were analysed together there was statistically significant evidence of a one-year cycle but not a two, three or four year effect.

More detailed results of the Poisson regressions for the one year cycle are in Table 6.10. In both Denmark and Finland $(\sin + \cos)_1$ variables made a statistically significant improvement to the model. For Denmark the peak incidence was $\theta = 27.9^{o}$ and for Finland at $\theta = 25.7^{o}$ i.e. in late January for both. Neither Norway nor Sweden showed a statistically significant result for $(\sin + \cos)_1$. However, the total series did, the likelihood ratio test being highly statistically significant ($p<0.001$) with peak towards the end of January ($\theta = 29.0^{o}$). The country-seasonality figure shows the result of adding country*$(\sin + \cos)_1$ interaction terms to the total model. There was no statistically significant interaction, therefore the data are consistent with patterns in all the countries being the same. (No interaction test was performed for $(\sin + \cos)_{2,3,4}$ as these main effects were not statistically significant in the total series). The amplitude (extent of seasonal

Table 6.1. Hodgkin's Disease - Denmark 1978-92

	Nos.				Rates per 10^5 population			
Age	Male	Female	Total	% total	Male	Female	Total	M/F
0-4	4	0	4	0.50	0.16	0.00	0.08	∞
5-9	8	4	12	1.50	0.31	0.16	0.24	1.94
10-14	28	20	48	6.00	1.01	0.75	0.88	1.35
15-19	81	44	125	15.6	2.80	1.60	2.20	1.75
20-24	118	93	211	26.3	3.86	3.21	3.54	1.20
25-29	117	97	214	26.7	4.00	3.49	3.75	1.15
30-34	123	65	188	23.4	4.41	2.44	3.43	1.81
Total	479	323	802	100.0	2.36	1.66	2.02	1.42

Table 6.2. Hodgkin's Disease - Sweden 1980-91

	Nos.				Rates per 10^5 population			
Age	Male	Female	Total	% total	Male	Female	Total	M/F
0-4	4	0	4	0.50	0.13	0.00	0.08	∞
5-9	16	8	24	3.10	0.51	0.27	0.39	1.89
10-14	33	27	60	7.80	1.00	0.85	0.93	1.18
15-19	87	69	156	20.2	2.47	2.05	2.26	1.20
20-24	83	101	184	23.8	2.33	2.97	2.65	0.78
25-29	106	74	180	23.3	2.95	2.16	2.56	1.36
30-34	99	66	165	21.3	2.66	1.86	2.26	1.44
Total	428	345	773	100.0	1.72	1.45	1.59	1.21

Table 6.3. Hodgkin's Disease - Finland 1953-93

	Nos.				Rates per 10^5 population			
Age	Male	Female	Total	% total	Male	Female	Total	M/F
0-4	11	3	14	0.81	0.15	0.04	0.10	3.75
5-9	44	19	63	3.64	0.58	0.26	0.42	2.23
10-14	62	50	112	6.46	0.78	0.65	0.72	1.20
15-19	147	141	288	16.6	1.85	1.84	1.85	1.01
20-24	243	163	406	23.4	3.11	2.18	2.65	1.43
25-29	266	164	430	24.8	3.58	2.29	2.94	1.56
30-34	248	172	420	24.2	3.46	2.48	2.97	1.40
Total	1021	712	1733	100.0	1.93	1.39	1.66	1.39

Table 6.4. Hodgkin's Disease - Norway 1980-92

	Nos.				Rates per 10^5 population			
Age	Male	Female	Total	% total	Male	Female	Total	M/F
0-4	1	1	2	0.40	0.06	0.06	0.06	1.00
5-9	5	3	8	1.60	0.28	0.17	0.23	1.65
10-14	17	16	33	6.80	0.86	0.84	0.85	1.02
15-19	41	39	80	16.4	1.91	1.91	1.91	1.00
20-24	81	57	138	28.3	3.76	2.78	3.27	1.35
25-29	71	49	120	24.6	3.37	2.45	2.91	1.38
30-34	69	37	106	21.8	3.30	1.87	2.59	1.76
Total	285	202	487	100.0	1.93	1.44	1.69	1.34

Table 6.5. Univariate analysis of cyclic pattern: Denmark.

Model no.	Factors add.	Compared	Diff in dev.	Diff in df	p
1	**$(\sin + \cos)_1$**	**trend, sex**	**7.27**	**2**	**0.026**
2	$(\sin + \cos)_2$	trend, sex	4.03	2	0.133
3	$(\sin + \cos)_3$	trend, sex	1.26	2	0.533
4	$(\sin + \cos)_4$	trend, sex	4.02	2	0.134

Table 6.6. Univariate analysis of cyclic pattern: Finland.

Model no.	Factors add.	Compared	Diff in dev.	Diff in df	p
1	**$(\sin + \cos)_1$**	**trend, sex**	**27.01**	**2**	**p<0.001**
2	**$(\sin + \cos)_2$**	**trend, sex**	**9.12**	**2**	**0.011**
3	$(\sin + \cos)_3$	trend, sex	2.04	2	0.361
4	$(\sin + \cos)_4$	trend, sex	1.57	2	0.456

Table 6.7. Univariate analysis of cyclic pattern: Norway.

Model no.	Factors add.	Compared	Diff in dev.	Diff in df	p
1	$(\sin + \cos)_1$	trend, sex	3.34	2	0.188
2	$(\sin + \cos)_2$	trend, sex	3.92	2	0.141
3	$(\sin + \cos)_3$	trend, sex	3.11	2	0.211
4	**$(\sin + \cos)_4$**	**trend, sex**	**14.81**	**2**	**p<0.001**

Table 6.8. Univariate analysis of cyclic pattern: Sweden.

Model no.	Factors add.	Compared	Diff in dev.	Diff in df	p
1	$(\sin + \cos)_1$	trend, sex	1.32	2	0.517
2	$(\sin + \cos)_2$	trend, sex	1.10	2	0.577
3	$(\sin + \cos)_3$	trend, sex	2.03	2	0.362
4	$(\sin + \cos)_4$	trend, sex	0.28	2	0.869

Table 6.9. Univariate analysis of cyclic pattern: Total series.

Model no.	Factors add.	Compared	Diff in dev.	Diff in df	p^+
1	**$(\sin + \cos)_1$**	**country, trend, sex**	**27.47**	**2**	**p<0.001**
2	$(\sin + \cos)_2$	country, trend, sex	3.51	2	0.173
3	$(\sin + \cos)_3$	country, trend, sex	4.35	2	0.114
4	$(\sin + \cos)_4$	country, trend, sex	0.60	2	0.741

Table 6.10. Results of one-year seasonality analysis.

	Regression coefficient (95% CI)		Statistical testing		Interpretation	
count	sin	cos	dev[a]	p[b]	peak	amp
Den	**0.063 (-0.04-0.16)**	**0.12 (0.02-0.22)**	**7.27**	**0.026**	**27.9°**	**13.5%**
Fin	**0.077 (0.01-0.14)**	**0.16 (0.09-0.23)**	**27.01**	**p<0.001**	**25.7°**	**17.7%**
Nor	0.008 (-0.12-0.13)	0.12 (-0.009-0.24)	3.34	0.188	4.1°	11.7%
Swe	0.045 (-0.05-0.14)	-0.037 (-0.14-0.06)	1.32	0.517	129.4°	5.8%
Tot	**0.059 (0.01-0.10)**	**0.11 (0.06-0.15)**	**27.47**	**p<0.001**	**29°**	**12.1%**

Country*seasonality interaction 0.10>p>0.05

a. dev = difference in deviance v. trend+sex model, except for total = v. country+trend+sex.
b. on 2 degrees of freedom (= difference in degrees of freedom v. trend+sex model, except for total = v. country+trend+sex.)

variation) for Denmark, Finland, and the total series were 13.5%, 17.7% and 12.1% respectively (Figures 6.2-6.5). Thus, while there is statistically significant evidence of seasonal presentation of HD, the curve is quite shallow.

Due to the statistical significance of the results of the one year cycle for presentation of HD in the total dataset, without statistically significant evidence of between-country heterogeneity, patterns of 2, 3, and 4 year cycles were calculated again adjusted for one year (Tables 6.11-6.13). Finland was the only country to show statistically significant evidence of a two-year cycle. $(\text{Sin} + \cos)_3$ was not statistically significant in any of the countries studied or the total series. Norway showed statistically significant evidence of a four-year cycle of HD presentation persisting after adjustment.

A step-down model process was applied to make more parsimonious models containing $(\sin + \cos)_n$ adjusted for age, trend and sex with results in Table 6.14. The process of finding the best model for Finland was the most complicated, as the data showed independent evidence of both one and two year seasonality. The best model consisted of both trend and sex, $(\sin + \cos)_{1,2}$ and an interaction term of trend and sex. Once the seasonal pattern of HD presentation had been analysed an attempt was made to find out if the seasonal/cyclical pattern was different by sex or age group. Results are in Tables 6.15-6.20. Analyses were performed for Sweden because of the significant seasonality results observed for the other countries. The Swedish data showed the only significant interaction with age group and the seasonal effect was limited to those cases aged 10-19 years. This significant result may only be occurring by chance as there was no statistically significant evidence of seasonality in the original Swedish data.

While the seasonal effect did not interact significantly with age group, the seasonal effects could more easily be ascribed to separate age groups. In Denmark,

Figure 6.2: Hodgkin's Disease in Denmark 1978-92. On the vertical axis the monthly data are given as a percentage above and below the mean. The horizontal axis is in months.

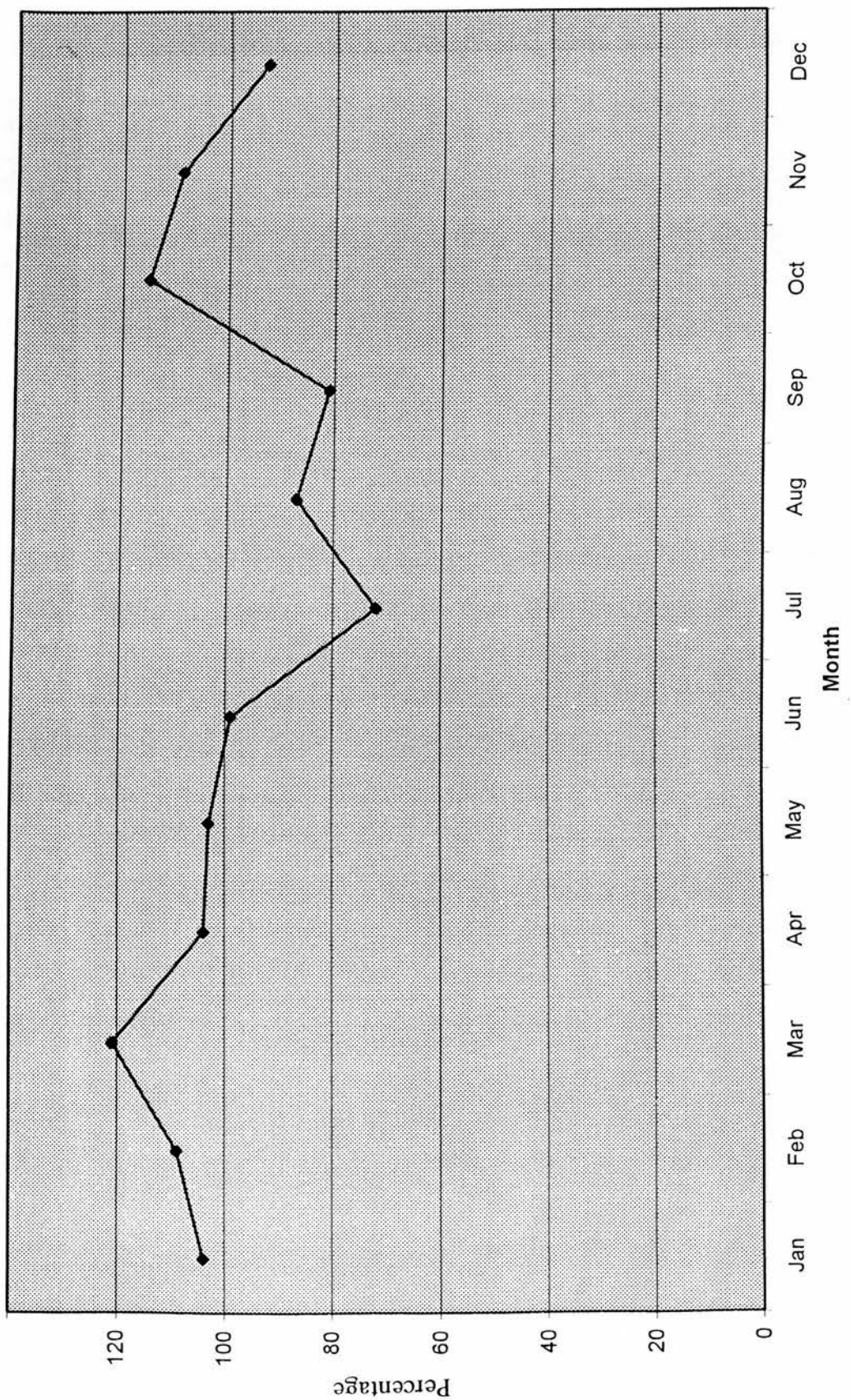


Figure 6.3: Hodgkin's Disease in Finland 1953-93. On the vertical axis the monthly data are given as a percentage above and below the mean. The horizontal axis is in months.

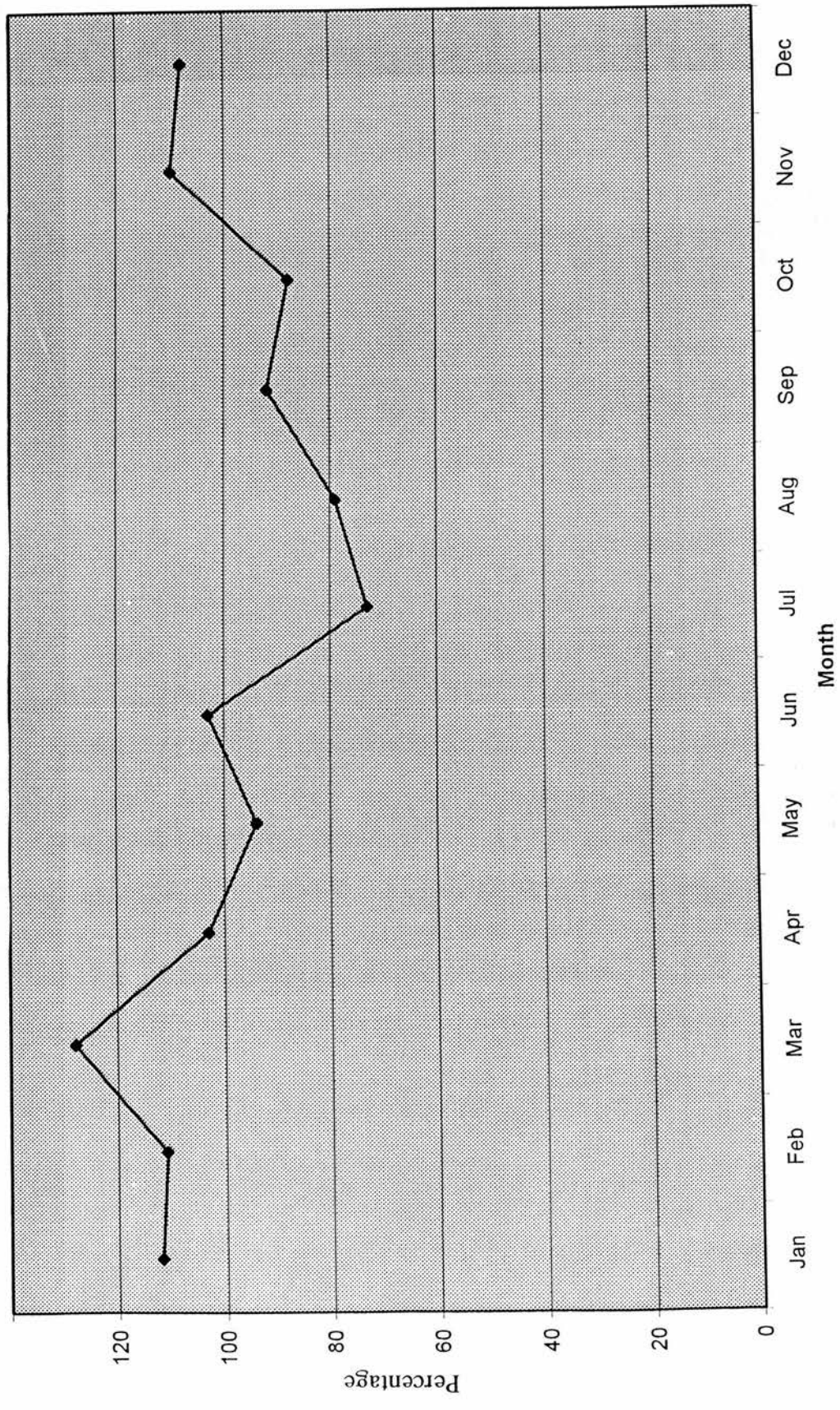


Figure 6.4: Hodgkin's Disease in Norway 1980-92. On the vertical axis the monthly data are given as a percentage above and below the mean. The horizontal axis is in months.

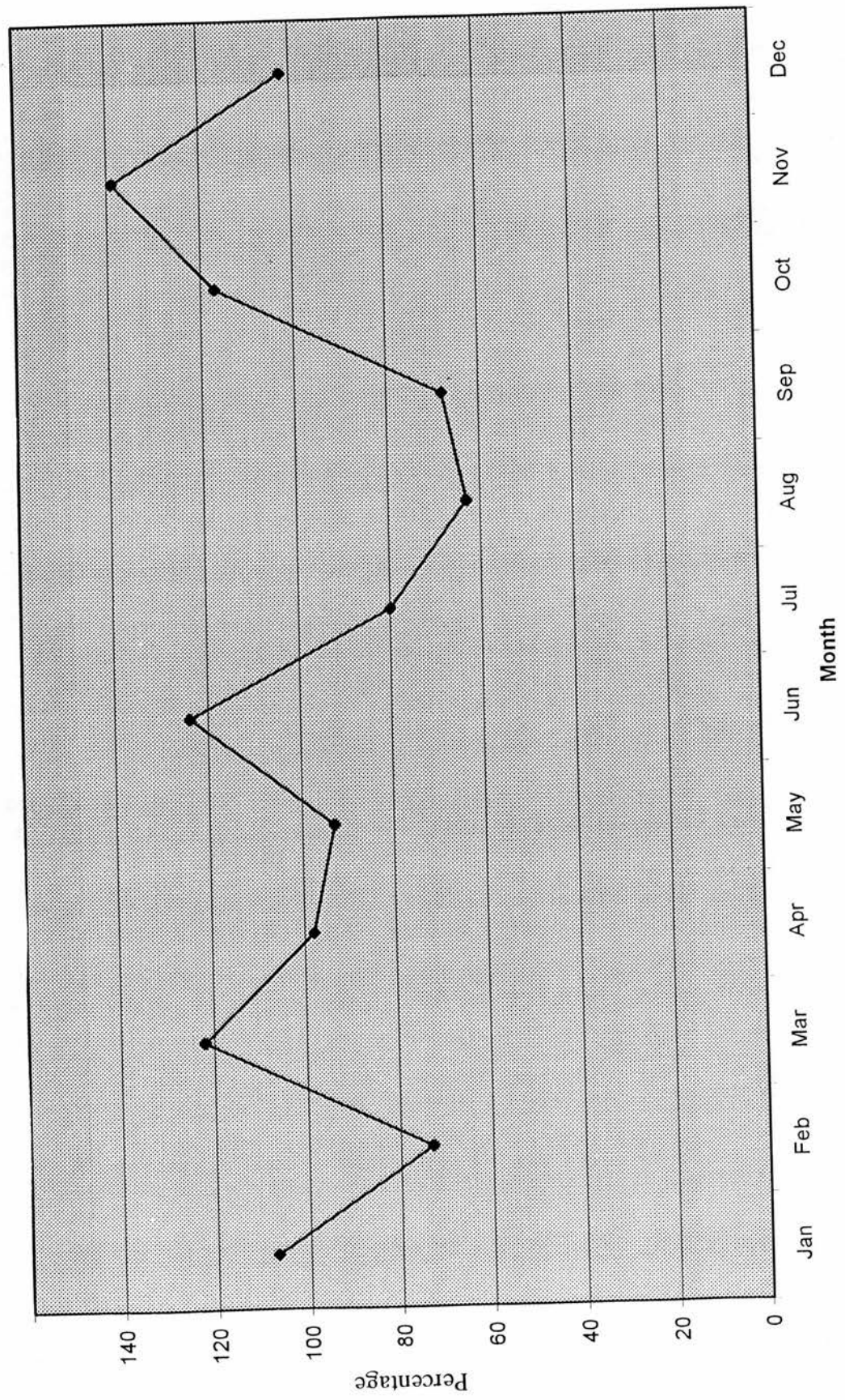


Figure 6.5: Hodgkin's Disease in Sweden 1980-91. On the vertical axis the monthly data are given as a percentage above and below the mean. The horizontal axis is in months.

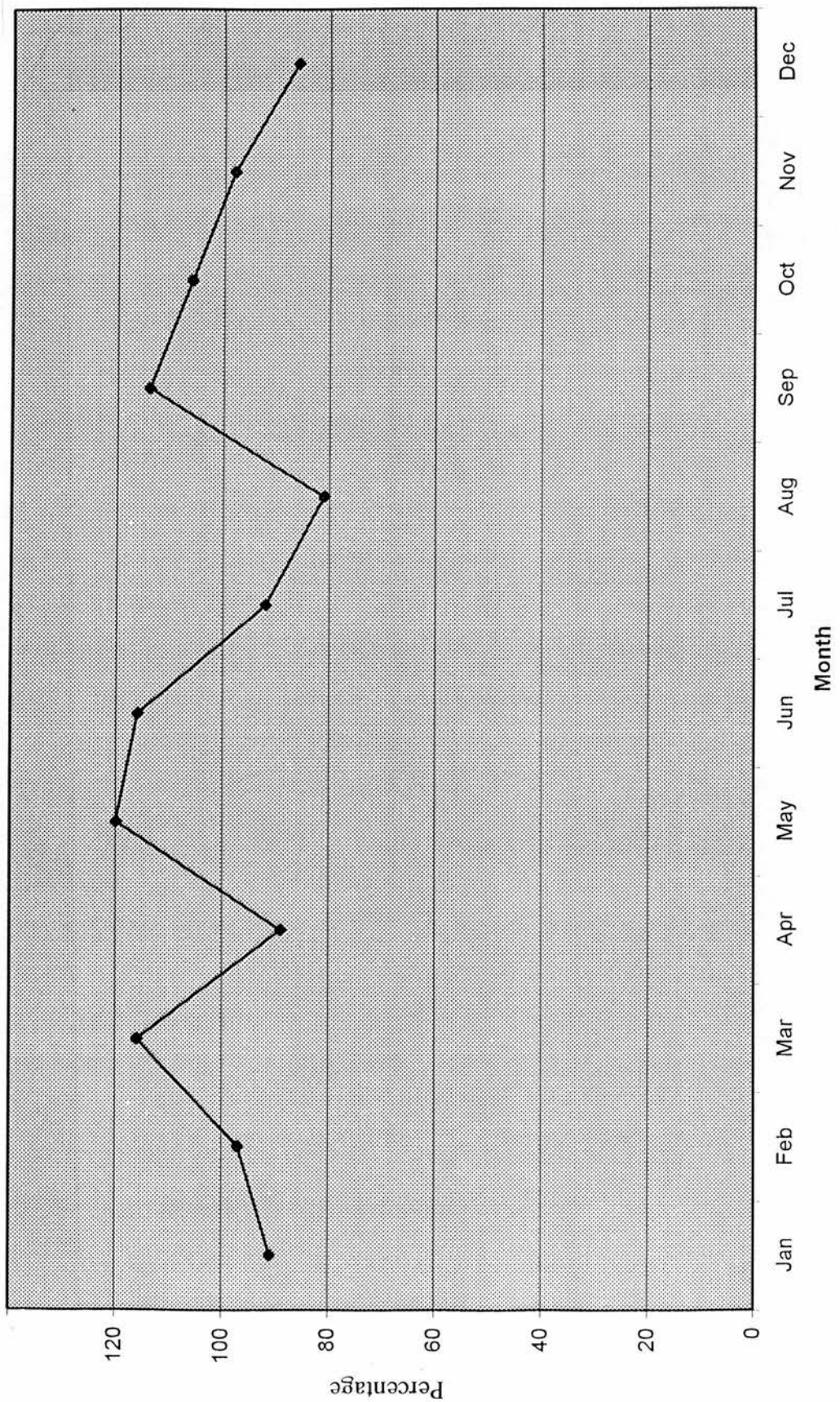


Table 6.11. Results of two-year seasonality analysis (after adjustment for one year).

	Regression coefficient (95% CI)		Statistical testing	
count	sin	cos	dev[a]	p[b]
Den	0.034 (-0.07-0.14)	0.085 (-0.01-0.18)	3.54	0.170
Fin	**0.076 (0.01-0.15)**	**-0.078 (-0.14- -0.01)**	**9.90**	**0.007**
Nor	0.077 (-0.05-0.21)	-0.105 (-0.23-0.02)	4.13	0.127
Swe	0.053 (-0.05-0.15)	0.001 (-0.10-0.10)	1.09	0.580
Tot	0.008 (–0.04-0.05)	-0.043 (-0.08-0.004)	3.23	0.199

a. dev = difference in deviance trend+sex +sin1+cos1+sin2+cos2 v. trend+sex+sin1+cos1 model, except for total = v. country+trend+sex+sin1+cos1.
b. on 2 degrees of freedom (= difference in degrees of freedom trend+sex +sin1+cos1+sin2+cos2 v. trend+sex+sin1+cos1 model, except for total = v. country+trend+sex+sin1+cos1.)

Table 6.12. Results of three-year seasonality analysis (after adjustment for one year).

	Regression coefficient (95% CI)		Statistical testing	
count	sin	cos	dev[a]	p[b]
Den	-0.055 (-0.15-0.04)	-0.015 (-0.08-0.08)	1.26	0.533
Fin	-0.031 (-0.10-0.04)	0.036 (-0.03-0.10)	1.94	0.379
Nor	0.109 (-0.02-0.24)	0.039 (-0.09-0.17)	3.25	0.197
Swe	0.016 (-0.09-0.12)	0.072 (-0.03-0.17)	2.06	0.357
Tot	0.002 (-0.04-0.05)	0.047 (0.001-0.09)	4.19	0.123

a. dev = difference in deviance trend+sex +sin1+cos1+sin3+cos3 v. trend+sex+sin1+cos1 model, except for total = v. country+trend+sex+sin1+cos1.
b. on 2 degrees of freedom (= difference in degrees of freedom trend+sex +sin1+cos1+sin3+cos3 v. trend+sex+sin1+cos1 model, except for total = v. country+trend+sex+sin1+cos1.)

Table 6.13. Results of four-year seasonality analysis (after adjustment for one year).

	Regression coefficient (95% CI)		Statistical testing	
count	sin	cos	dev[a]	p[b]
Den	0.023 (-0.08-0.12)	-0.102 (-0.20- -0.002)	4.23	0.121
Fin	0.038 (-0.03-0.1)	-0.024 (-0.09-0.04)	1.68	0.432
Nor	**0.234 (0.10-0.36)**	**-0.111 (-0.24--.02)**	**15.04**	**p<0.001**
Swe	0.014 (-0.09-0.12)	-0.023 (-0.12-0.08)	0.27	0.874
Tot	0.015 (-0.03-0.06)	0.011 (-0.03-0.56)	0.62	0.733

a. dev = difference in deviance trend+sex +sin1+cos1+sin4+cos4 v. trend+sex+sin1+cos1 model, except for total = v. country+trend+sex+sin1+cos1.
b. on 2 degrees of freedom (= difference in degrees of freedom trend+sex +sin1+cos1+sin4+cos4 v. trend+sex+sin1+cos1 model, except for total = v. country+trend+sex+sin1+cos1.)

Table 6.14. Univariate analysis to test for the significance of sex and trend.

Country	Model no.	Factors add.	Compared	Diff in dev	Diff in df	p
Denmark	1	trend,	$(\sin + \cos)_1$	0.27	1	NS
	2	sex,	$(\sin + \cos)_1$	0.02	2	NS
	3	trend, sex	$(\sin + \cos)_1$	1.46	3	NS
Finland	1	trend	$(\sin + \cos)_1$	3.51	1	0.061
	2	sex	$(\sin + \cos)_1$	0.11	2	NS
	3	**trend, sex**	**$(\sin + \cos)_1$**	**19.64**	**3**	**p<0.001**
	4	**trend**	**sex, $(\sin + \cos)_1$**	**19.52**	**1**	**p<0.001**
	5	**sex**	**trend, $(\sin + \cos)_1$**	**16.12**	**2**	**p<0.001**
	6	**trend*sex**	**trend, sex, $(\sin + \cos)_1$**	**9.61**	**1**	**0.002**
Norway	1	trend,	$(\sin + \cos)_4$	0.004	1	NS
	2	sex,	$(\sin + \cos)_4$	0.18	2	NS
	3	trend, sex	$(\sin + \cos)_4$	0.50	3	NS
Sweden[a]	1	sex	trend	6.84	1	0.009
	2	trend	sex	9.47	2	0.009
	3	$(\sin + \cos)_1$	trend, sex	1.32	2	NS
Total	1	sex	country	0.00	1	NS
	2	**trend**	**country**	**25.17**	**1**	**p<0.001**
	3	**trend**	**country, sex**	**25.17**	**1**	**p<0.001**
	4	sex	country, trend	0.0007	1	NS
	5	sex	country, trend, $(\sin + \cos)_1$	0.0007	1	NS

NS (not statistically significant) = p>0.10

a. No effect of $(\sin + \cos)_n$ in Sweden.

Table 6.15. Multivariate analysis: main effect of sex and interaction of sex with cyclic patterns.

Country	Factors add.	Compared	Diff in dev	Diff in df	p
Denmark	sex	$(\sin + \cos)_1$	0.02	2	NS
	*sex	$(\sin + \cos)_1$, sex	0.12	2	NS
Finland	*sex	trend, sex, trend*sex, $(\sin + \cos)_1$, $(\sin + \cos)_2$	3.72	4	NS
Norway	sex	$(\sin + \cos)_4$	0.18	2	NS
	*sex	$(\sin + \cos)_4$, sex	0.14	2	NS
Sweden	1) *sex	1) trend,sex, $(\sin + \cos)_1$	0.5	2	NS
	2) *sex	2) trend,sex, $(\sin + \cos)_2$	0.66	2	NS
	3) *sex	3) trend,sex, $(\sin + \cos)_4$	0.97	2	NS
Total	sex	country, trend, $(\sin + \cos)_1$	0.0007	1	NS
	*sex	country, sex trend, $(\sin + \cos)_1$	0.26	2	NS

NS (not statistically significant) = $p>0.10$

*sex = interaction of cyclic pattern with sex.

Table 6.16. One-year seasonality by sex.

		Regression coefficient (95% CI)		Interpretation	
Country	sex	sin	cos	peak	amp%
Denmark	Male	0.052 (-0.08-0.18)	0.111 (-0.02-0.24)	25.7°	12.3
	Female	0.081 (-0.07-0.24)	0.132 (-0.02-0.29)	31.5°	15.5
Finland+	Male	0.117 (0.03-0.20)	0.149 (0.06-0.24)	38.1°	18.9
	Female	0.029 (-0.08-0.13)	0.180 (0.08-0.29)	8.8°	18.2
Norway	Male	-0.012 (-0.18-0.15)	0.126 (-0.04-0.29)	275.4°	12.7
	Female	0.037 (-0.16-0.23)	0.104 (-0.09-0.30)	19.3°	11.0
Sweden	Male	0.031 (-0.10-0.17)	-0.008 (-0.14-0.13)	104.4°	3.2
	Female	0.062 (-0.09-0.21)	-0.074 (-0.22-0.08)	140.0°	9.7
Total	Male	0.068 (0.01-0.13)	0.607 (0.55-0.67)	6.3°	61.1
	Female	0.045 (-0.02-0.11)	0.103 (0.03-0.17)	23.5°	11.2

(none of the interactions of seasonality with sex are significant on 2 df, see Table 15)
+ trend included in model

Table 6.17. Multivariate analysis: age as main effect and interaction of age in two groups (0-14, 15-34 years) with cyclic pattern (*age).

Country	Factors add.	Compared	Diff in dev	Diff in df	p
Denmark	age	$(\sin + \cos)_1$	0.02	1	NS
	*age	$(\sin + \cos)_1$, age	0.68	2	NS
Finland	age	trend, sex, trend*sex, $(\sin + \cos)_1$, $(\sin + \cos)_2$	0.05	1	NS
	*age	trend, sex, trend*sex, $(\sin + \cos)_1$, $(\sin + \cos)_2$, age	8.71	4	NS
Norway	age	$(\sin + \cos)_4$	0.19	1	NS
	*age	$(\sin + \cos)_4$, age	0.65	2	NS
Sweden	1) age	1) trend,sex, sin1, cos1	0.002	1	NS
	1) *age	1) trend,sex, $(\sin + \cos)_1$, age	0.9	2	NS
	2) age	2) trend,sex, $(\sin + \cos)_2$	0.002	1	NS
	2) *age	2) trend,sex, $(\sin + \cos)_2$, age	0.9	2	NS
	3) age	3) trend,sex, $(\sin + \cos)_4$	0.002	1	NS
	3) *age	3) trend,sex, $(\sin + \cos)_4$, age	3.09	2	NS
Total	age	country, trend, $(\sin + \cos)_1$	0.04	1	NS
	*age	country, trend, $(\sin + \cos)_1$, age	3.03	2	NS

NS (not statistically significant) = $p>0.10$

Table 6.18. Multivariate analysis: age as main effect and interaction of age in three groups (0-9, 10-19, 20-34 years) with cyclic pattern (*age).

Country	Factors add.	Compared	Diff in dev	Diff in df	p
Denmark	age	$(\sin + \cos)_1$	0	2	NS
	*age	$(\sin + \cos)_1$, age	3.24	4	NS
Finland	age	trend, sex, trend*sex, $(\sin + \cos)_1$, $(\sin + \cos)_2$	0.11	2	NS
	*age	trend, sex, trend*sex, $(\sin + \cos)_1$, $(\sin + \cos)_2$, age	10.88	8	NS
Norway	age	$(\sin + \cos)_4$	0.18	2	NS
	*age	$(\sin + \cos)_4$, age	2.91	4	NS
Sweden	1) age	1) trend,sex, $(\sin + \cos)_1$	0.002	2	NS
	1) *age	1) trend,sex, $(\sin + \cos)_1$, age	7.13	4	NS
	2) age	2) trend,sex, $(\sin + \cos)_2$	0.22	2	NS
	2) *age	2) trend,sex, $(\sin + \cos)_2$, age	3.70	4	NS
	3) age	3) trend,sex, $(\sin + \cos)_4$	0.002	2	NS
	3) *age	3) trend,sex, $(\sin + \cos)_4$, age	11.94	4	0.025>p >0.01
Total	age	country, trend, $(\sin + \cos)_1$	0.07	2	NS
	*age	country, trend, $(\sin + \cos)_1$, age	6.30	4	NS

NS (not statistically significant) = p>0.10

Table 6.19. One-year seasonality by age group at two levels.

		Regression coefficient (95% CI)		Interpretation	
Country	Age group	sin	cos	peak	amp%
Denmark	0-14	0.072 (-0.28-0.42)	-0.02 (-0.37-0.23)	105.5°	7.5
	15-34	0.063 (-0.04-0.17)	0.132 (0.03-0.23)	25.5°	14.6
Finland	0-14	0.129 (-0.07-0.33)	-0.109 (-0.31-0.09)	130.2°	16.9
	15-34	0.075 (0.004-0.15)	0.194 (0.12-0.27)	21.2°	20.8
Norway	0-14	-0.153 (-0.58-0.28)	0.396 (-0.04-0.83)	291.1°	42.5
	15-34	0.024 (-0.11-0.16)	0.091 (-0.04-0.22)	14.6°	9.4
Sweden	0-14	-0.085 (-0.38-0.21)	-0.004 (-0.30-0.29)	182.7°	8.5
	15-34	0.062 (-0.04-0.17)	-0.041 (-0.15-0.06)	123.5°	7.4
Total	0-14	0.036 (-0.11-0.18)	-0.011 (-0.26-0.03)	107.0°	3.8
	15-34	0.061 (0.01-0.11)	0.119 (0.07-0.17)	27.0°	13.3

(none of the interactions of seasonality with age group are significant on 2 df, see Table 6.17)

Table 6.20. Main effect of age in three groups and interaction of age group with one-year seasonality.

		Regression coefficient (95% CI)		Interpretation	
Country	Age group	sin	cos	peak	amp%
Denmark	0-9	0.630 (-0.12-1.37)	-0.183 (-0.90-0.53)	106.2°	65.5
	10-19	0.072 (-0.14-0.28)	0.155 (-0.06-0.37)	24.9°	17.1
	20-34	0.047 (-0.07-0.16)	0.118 (0.005-0.23)	21.7°	12.7
Finland	0-9	0.237 (-0.09-0.56)	-0.058 (-0.38-0.26)	103.8°	24.4
	10-19	0.169 (0.03-0.31)	0.112 (-0.03-0.25)	56.5°	20.3
	20-34	0.043 (-0.04-0.12)	0.190 (0.11-0.27)	12.9°	19.5
Norway	0-9	-0.414 (-1.32-0.49)	0.204 (-0.69-1.1)	333.8°	46.2
	10-19	0.120 (-0.14-0.38)	0.336 (0.07-0.60)	19.7°	35.7
	20-34	-0.015 (-0.14-0.38)	0.05 (-0.10-0.19)	286.7°	5.2
Sweden	0-9	-0.526 (-0.108-0.03)	-0.122 (-0.66-0.41)	193.1°	54.0
	10-19	0.191 (0.001-0.38)	-0.0006 (-0.19-0.19)	90.0°	19.1
	20-34	0.015 (-0.11-0.14)	-0.048 (-0.17-0.07)	162.6°	5.0
Total	0-9	0.077 (-0.17-0.32)	-0.055 (-0.3-0.19)	125.6°	9.4
	10-19	0.145 (0.05-0.24)	0.118 (0.03-0.21)	50.8°	18.7
	20-34	0.030 (-0.02-0.08)	0.109 (0.06-0.16)	15.4°	11.3

(none of the interactions of seasonality with age group are significant on 2 df, see Table 6.18)

Finland, and the total series the effects of $(\sin + \cos)_1$ are limited to the older age group (age 15-34 years). In these countries, as well as Norway, the peak of the curves in the older age group are all within January ($0\text{-}30^0$) while peaks for the 0-14 years age group are at totally different times of the year (April for Denmark and the total series, May for Finland). Norway and Sweden show no evidence for any seasonal effect being limited to a particular age group. When the data is divided into three age groups there is again some evidence that the seasonal effect is limited to the young adults in Denmark, Finland, and the total series. In these countries the peak in the 20-34 years age group is again January.

6.4 Discussion:

As with many other epidemiological data at a national level covering many years, involving many reporting centres there may be weaknesses in the quality of the data used in the analyses presented here. The figures on HD for the four countries were collected over different time periods and reporting and diagnostic criteria may have changed during this time, especially in Finland as the data from this country covers a longer time period (40 years) than the other countries (10 years). Alexander et al (1989) reported the results of cross-checking 1986 notifications of haematopoetic cancers from Yorkshire, Trent and South West regional cancer registries received by the Leukaemia Research Fund (LRF) in October 1987 against LRF registrations for 1986. The cancer registry only contained 56.8% of the total HD registrations within 9 months. Since an important source of data for the registries is death certificates it is likely that their ascertainment will be less good for cancers with low mortality rates e.g. HD. However, in the context of cancer registration Alexander et als' time period of 9 months is a short one. The Office of Population, Censuses and Surveys (OPCS) did not publish data on 1971 registrations until 1978; more recently

this time gap has been reduced to 3 years. The data for this study goes back to 1993 and, therefore, should be reasonably complete.

The data used here are for children and young adults (age 0-34 years) and not the older cases (age 50+ years). Therefore, changes in diagnosis over time which mean that cases that may previously been classified as HD may now be classified as NHL in the 50+ age group could not affect these data.

Month of histologic confirmation should be considered the most objective indicator of disease onset and this is the data used in the analysis presented here and the majority of studies. Onset of symptoms can vary according to the individual symptoms and, therefore, have a large subjective component involving recall and other factors. The date of first symptom is much weaker (Douglas et al, 1998). In the analysis in this chapter the date used was that of HD registration which would normally approximate closely to histologic confirmation.

The seasonal presentation of HD cases from the Scandinavian cancer registries has been analysed using formal statistical methodologies. However, this is not always the case for studies of this question. Seven of thirteen studies published (Table 6.21) have only described the pattern of presentation of HD. More recent studies have made use of two methods of analysis: Edward's (1961) and cosinor analysis.

Edward's (1961) method used a square root transformation of the monthly frequencies of disease incidence to quantify the degree of cyclic variation. The degree of cyclic trend is based on the ratio of the peak annual frequency to the lowest annual frequency, estimated from a fitted sine curve. The time of year that the peak occurs can be estimated as well.

Table 6.21. Studies investigating the seasonal presentation of Hodgkin's Disease.

Author	Cases	Test	Age	Peak	p
Cridland (1961) (England)	106	None	NA	Dec	NA
Innes & Newell (1961) (Scotland)	104	None	NA	Jan	NA
Fraumeni & Li (1962) (USA)	314	None	<15	Nov/Dec Dec/Jan	NA NA
Bjelke (1969) (Norway)	820	None	All	NA	NA
Modan et al (1969) (Israel)	204	No details	All	Jun	<0.01
Newell (1972) (USA)	282	None	<40	Feb/Apr/Oct	NA
	276		>40	Sep	NA
Bogger-Goren et al (1983) (Israel)	21	None	2.5-14.5	NA	NA
Alderson & Nayak (1971) (England)	737	Edwards 1961	All	May/Jun	0.005>p> 0.001
Newell et al (1985) (USA)	117	Edwards 1961	0-14m	Mar	NS
	83		0-14f	Oct	NS
	981		15-39m	Feb	0.002
	702		15-39f	Feb	NS
	1057		>40m	Mar	NS
	767		>40f	Oct	NS
Westerbeek et al (1998) (England)	166	Edwards 1961	0-14 symptoms	Dec	0.002
			0-14 diagnosis	Feb	NS

Table 6.21 (cont.). Studies investigating the seasonal presentation of Hodgkin's Disease.

Author	Cases	Test	Age	Peak	p
Neilly et al (1995a & b) (UK)	446	Cosinor analysis	<40m	Feb	<0.05
	350		<40f	Mar	<0.05
	327		>40m	Feb	NS
	216		>40f	Apr	NS
Douglas et al (1996) (Scotland)	1580	Cosinor analysis	<40tot	Feb/Mar	<0.05
	882		<40m	Jan/Feb	<0.05
	697		<40f	Mar/Apr	<0.05
	1864		>40tot	May	NS
	983		>40m	Apr	NS
	881		>40f	Jun	NS
Douglas et al (1998) (England)	2959	Cosinor analysis	0-79	Apr	0.063
	1260		0-79 NS	Mar	0.009
	1894		0-44tot	Mar	NS
	1098		0-44m	Feb	NS
	796		0-44f	Mar	NS
	1066		45-79tot	May	NS
	654		45-79m	Mar	NS
	410		45-79f	Apr	NS

NS (not statistically significant) = p>0.10

For cosinor analysis the year length is considered to be 360° and the mid-point of each month of the year is assigned an angular value t, for January (15°) through to December (345°). In order to establish the presence of a seasonal presentation in the data a cosinor model of the form $Y = \beta_0 + \beta_1 \sin(t) + \beta_2 \cos(t) + \varepsilon$ is fitted, where ε is random error. Cosinor analysis also gives the angular position in the year (converted to the nearest month) where the fitted regression line has its highest value. The methods used in the analysis presented here is very similar to cosinor analysis in that both fit a function for (sin + cos) in the regression model.

The earliest reports of a seasonal presentation of HD go back to 1961 and generally just described the seasonal pattern of presentation of HD. In 1961 Cridland looked for a seasonal pattern in the clinical onset of HD. In 106 out of 269 case histories seen at the Royal Marsden Hospital 1945-69 there was reason to suppose that the clinical onset was closely related to the time of apparent onset (precise criteria given). Of the 106 patients Cridland studied 23 (21.7%) had a clinical onset in December. Innes & Newell (1961) used the same method to look at a series of patients in Edinburgh. Using this methodology meant that only 104 out of a possible 440 patients were used in the analysis. Of these 104, 43 (41.3%) had a clinical onset in January-March (18% in January). The methodology in both of these studies can be criticised.

Cridland (1961) and Innes & Newell (1961) found, for a highly selected group of HD cases, constituting only 40% and 24% of their entire series respectively, that a 'significant' number had clinical onset in the winter. However, these results are entirely dependent on the ability of patients to recall the month in which they first noticed large lymph nodes (taken to be month of onset). As mentioned previously specification of the month of onset is difficult and involves many uncertainties. The month of diagnosis on the other hand, can usually be fixed objectively. If there were marked seasonal variation in the

clinical onset of HD, one might expect these to be reflected in the distribution of cases by month of diagnosis (Bjelke, 1969).

Fraumeni et al (1962) observed diagnosis to be more common in December and January (64/285 = 22.5%) in a study of *children aged under 15 years* in the USA. This study also suggested that the clinical onset of HD peaked in December and January. Bogger-Goren et al (1983), again looking at children, found 13/21 (61.9%) cases of HD experienced onset of symptoms in the cold season (October-March). These cases diagnosed in the cold season had more acute symptoms and were generally stage III or IV.

In contrast to studies which have found a seasonal presentation of HD Modan et al (1969), Bjelke et al (1969), and Newell (1972) found no evidence of a seasonal presentation. "As a whole, variation in the distribution of cases by month of diagnosis could easily be ascribed to chance" (Bjelke, 1969). Bjelke (1969) also made special mention of the fact that, in particular, no seasonal pattern was observed for the cases aged 15-34 years.

Alderson & Nayak (1971) used Edwards' method and showed the peak number of diagnoses of HD in England took place in May/June. This 'spring peak', as they described it, was seen especially in males and, more surprisingly, in those cases aged over 50 years. Newell et al (1985) using the same method investigated the occurrence of HD by month of diagnosis and sex for 1969-71 and 1973-75 in seven regions of the United States and Puerto Rico. These authors found that seasonal variation was not statistically significant for boys or girls aged under 14 for either time period. There was also little monthly fluctuation in adults over 40. However, in males aged 15-39 years diagnosis peaked in February (spring) in four of the six years and in an adjacent month in the other two. For females the diagnosis peaked in January or February for three of the six years. The results for all six years combined can be seen in Table 6.20. More recently Westerbeek et al (1998) investigated the seasonal presentation of various haematological malignancies, including HD, in *children aged 0-14*

years in England. These authors investigated the data for evidence of seasonality in the date of first symptoms and date of diagnosis. Westerbeek et al (1998) found a striking seasonal variation in the date of first symptom for HD, with a highly statistically significant peak in December, with an amplitude of 41.0%. There was no significant seasonal variation in the date of diagnosis of HD. Westerbeek et al (1998) suggested that the date of first symptoms more closely reflects the event that precipitates the clinical onset of disease than date of diagnosis. However, the majority of studies have used date of diagnosis as the variable of interest and many have found statistically significant seasonal diagnosis of HD.

Edwards' method was used to check for seasonal presentation of HD in the total combined series for Scandinavia presented in this chapter. The results of the Edwards' method and the Poisson regression method used here produce almost identical results (both giving a statistically significant one-year seasonality with peaks of 28.6^{o} and 29.0^{o} respectively).

Neilly et al (1995a & b) analysed 1359 cases of HD from the Scotland and Newcastle Lymphoma Group (SNLG) registry 1979-92 by cosinor analysis. A March peak was evident when all cases were analysed ($p<0.01$). When analysed separately NS and MC HD showed the March peak ($p<0.05$) while LP and LD HD had no demonstrable seasonal variation. However, there were far fewer LP and LD cases (13 and 6.7% of the total series respectively) and, therefore, the lack of observed seasonality for LP and LD could be due to low power. In a breakdown by age and sex there was evidence of seasonality in both sexes but only in those under 40 years of age. High grade NHL analysed in the same way did not demonstrate seasonality of presentation.

Douglas et (1996) also used cosinor analysis to investigate the seasonality of 3444 cases of HD (including the cases used by Neilly et al, 1995 a & b) taken from the cancer registry for Scotland. They found significant seasonality in the total data and in both sexes

and the sexes combined and in cases diagnosed under 40 years. The peak in males was about a month earlier than in females (Jan/Feb v. Mar/Apr) (both $p<0.05$). Above the age of 40 years the seasonality in each sex was no longer significant.

Douglas et al (1998) analysed population-based data from parts of England and Wales over 10 years. These authors examined 2959 cases by age (0-44 and 45-79 years) and sex groups and also by Rye type. Patients with NS HD had a highly statistically significant seasonal presentation of HD with a low amplitude and a peak in March. A significant seasonal peak was also seen for LP HD except this had a high amplitude and a peak in August. There were no statistically significant results for any individual age and sex group analyses.

The findings of the analysis presented here agree with those of previous work in that diagnosis appears to peak in spring in the Northern hemisphere. Seasonality is a feature of infectious illness but at present there is not any conclusive evidence that HD is caused by an infection, although there is a great deal of circumstantial evidence. It is possible that tumours caused by an infection could have a relatively constant incubation period giving a seasonal rhythm for tumour presentation.

However, there are alternative explanations for a seasonal presentation of HD; it need not be related to infection. The human immune response also has seasonal features (Abo et al, 1978; Canon et al, 1986; Levi et al, 1988). Presentation of HD appears to coincide with an increased period of T cell activity in normal adults which may be related to viral exposure (Neilly et al, 1995b). Seasonal variation in HD could be due to the attention of patients being drawn to previously unrecognised enlarged nodes by a coincidental infection of the upper respiratory tract. However, Cridland (1961) found only 10.4% of HD patients presented with coincidental or recent upper respiratory tract infection and that these were not most prevalent in the month of highest incidence of HD.

If an infectious agent is involved in the aetiology of HD a seasonal variation might be expected depending on the behaviour of the agent, the length and variability of the latent period and method of spread. A lack of seasonal onset or diagnosis does not rule out the possibility of an infectious origin, as low incidence in infected people combined with a very long and/or variable latent period could blur its episodic appearance. Alternatively a causative agent might not vary seasonally. Conversely, a significant seasonal presentation is not conclusive proof of an infectious origin.

In agreement with Newell et al (1985), Neilly et al (1995a & b), and Douglas et al (1996) this analysis found statistically significant evidence of seasonality in the presentation of HD. The effect of age group and the interaction of age group with seasonality did not reach statistical significance but the effect was most evident in the young adults. No significant interaction with sex was found. The peak month in this study was slightly earlier in the individual countries and the total series (January) than in other reports (February/March) (Newell et al, 1985; Neilly et al, 1995a & b; Douglas et al, 1996). The findings are consistent with the hypothesis that clinical HD is the rare sequellae of exposure to a very prevalent but seasonal environmental agent e.g. a virus of low pathogenecity to which infection is most frequent at the beginning of the year.

7 Results of YHHCCS General & Family Health Analyses

7.1 Introduction:

The aetiology of HD remains unclear. There is a body of evidence relating to past own medical history and its effect on risk of HD. The effect of infections are especially of interest because evidence has amassed for the 'late-host-response' model of HD in young adults, under which the disease is a (rare) sequella to late first infection by one or more unknown infectious agents. Evidence includes studies showing a positive association between HD in young adults and proxies for risk for late first exposure to infection and the relationship of IM to HD (see Chapter 1). General support for an infectious aetiology of HD is provided by reports linking specific HLA class II alleles (Bodmer et al, 1989; Klitz et al, 1994). The best documented association is a positive one with HLA DPB1*0301 phenotype (Taylor et al, 1996). It would be expected that people who had more infections would have a lower risk of young adult HD, especially if these occurred early in life. If infection was delayed risk of young adult HD may be increased. Although there have been studies looking at specific infections (Vianna et al, 1971; Newell et al, 1973; Paffenbarger et al, 1977; Andersson & Isager, 1978; Kirschoff et al, 1980; Gutensohn & Cole, 1981; Evans & Gutensohn, 1984) none have considered the total number of childhood infectious illnesses.

Own medical history factors including childhood infectious illnesses (chicken pox, German measles, measles, mumps, and whooping cough) individually and in total at specific ages are tested here on data from the YHHCCS study for their effect on HD risk in young adults. These infectious illnesses are being interpreted as proxies for increased probability of exposure to a wide spectrum of agents with and without clinical symptoms.

EBV has now been found in around 35% of HD cases in the developed world (see Chapter 1) and its role is generally agreed to be causal. Sub-classification of cases by the

presence of EBV (EBV +ve) or absence (EBV –ve) of EBV in the HRS cells provides a classification which may identify aetiological subgroups. (EBV +ve is analogous to EBV-associated, used in Chapter 1.) Only one case series has attempted to compare epidemiological risk factors by EBV status (Sleckman et al, 1998). Risk factors for HD are considered separately for subgroups based on EBV status, histological status, and HLA DPB1 type.

There is also evidence that certain aspects of family history are risk factors for HD. These variables will be tested on data from the YHHCCS study. In addition an attempt was made to develop a basic scoring system for cancer risk and infectious mononucleosis (IM) risk in family members. Many studies of family history define exposure by the number of affected relatives without taking account of the age of family members.

7.2 Materials and Methods:

The YHHCCS study is described in Chapter 4.

Interview Data:

Trained interviewers conducted face-to-face interviews. The period of interest was from birth up to a reference date which was the date of diagnosis of the case in each matched set. Information was requested on proxies for exposure to infection, past history of infectious illness, past medical history of the index and their family, and a limited history of infectious illness in friends and family.

Categories of interest:

Index General Health analyses: The effect of appendicitis at any age, at age younger than the mean of controls (aged <12 yrs), and older than the mean age of controls (≥12 yrs) was investigated. Removal of Tonsils/adenoids was analysed in the same way (mean age of controls 5.5 yrs). Thyroid disease and previous transplant were also recorded. The questionnaire included

questions on previous malignancies and auto-immune diseases, but these could not be analysed as there were no positive responses.

Infectious illness analyses: The presence of specific childhood infectious illnesses (chicken pox, German measles, measles, mumps, whooping cough) were recorded on the questionnaire. These infectious illnesses were analysed as a simple yes/no variable at any age and as yes/no in age groups: 0-4, 5-9, ≥10 years. These data were then combined to give a total number of childhood infections at all ages and at ages 0-4, 5-9 and ≥10 years. A blind assessment (to case/control status) of the distribution of frequencies of childhood infections in the total series led to sensible strata for the number of episodes for analysis (Table 7.1). The total number of infections at all ages, and aged 0-4 and 5-9 years were also analysed as dichotomies in which the lowest level (generally none) is taken as a reference group and the rest are combined. If no age was given the data was considered to be missing in the analysis by age of individual infections. For total infectious illness, missing values were treated as zeros in the sums, unless the data were missing for all individual infections, then the total was set to missing. To check if this assumption was correct the analyses were repeated using data from people who answered yes/no for 5 infections, for 4, for 3, for 2, and for 1 only. The results for all of these were compared. There was very little difference in the magnitude of the odds ratios produced and the directions were constant. Therefore, only the results treating missing values as zeros will be presented.

The presence of IM in the index was elicited from the questionnaire, however, laboratory confirmation was not obtained. The question on IM had three possible answers: no, suspected, and yes. IM was analysed in four ways. Separate analyses were performed treating suspected IM as yes and suspected IM as no. To remove the possible confusion arising from IM in the year prior to diagnosis (symptoms of IM can arise as a preliminary indication of HD) IM

Table 7.1: Categories for analysis of number of all childhood infections combined.

Age group	Number of levels	Definition
All ages	4	1 (0 or 1), 2 (2), 3 (3), 4 (≥4)
Aged <5 years	3	0 (0), 1 (1), 2 (≥2)
Aged 5-9 years	4	0 (0), 1 (1), 2 (2), 3 (≥3)
Aged ≥10 years	2	0 (0), 1 (≥1)

up to one year prior to diagnosis was analysed separately (also twice once treating suspected IM as yes and one as no).

Family health analyses: Presence of malignancy in a first-degree relative was recorded on the questionnaire. For analysis malignancies were divided into three groups: haematological malignancy (ICD 9 200-204); any haematological or young onset (aged <50 years) malignancy; any malignancy. The number of relatives with the condition of interest was analysed as either two levels (≥1v.0) or three levels (0, 1, ≥2). In the literature, most analyses of cancer in a first-degree relative adjust for the number of relatives. However, this may not be enough due to the differing risk of cancer by age and sex. Therefore, in the analysis presented here the risk of cancer in a family member was weighted by age. Incidence rates of all the cancers affecting YHHCCS first degree relatives were taken from Cancer Incidence in 5 Continents (Parkin et al, 1992) allowing the construction of an approximate ratio of cancer cases aged 0-24: 25-49: ≥50 years for the UK. As age increases there are more cases of cancer in an approximate ratio of 1:8:47. Weighted sums of relatives at risk equal Σw_I across relatives where $w_I = 1$ for relatives age 0-24, 8 for relatives age 25-49, and 47 for relatives age ≥50 years. The weighted sum of relatives at risk was then divided into quartiles. Those with larger, but especially older families had a greater risk of cancer and, therefore, larger scores.

The presence in the family of multiple sclerosis (MS), pernicious anaemia (PA), rheumatoid arthritis (RA), and thyroid disease were investigated. The answers to these questions were then combined to give a total auto-immune disease variable. Again, if possible the number of relatives with the particular disease were analysed with two (≥1, 0) or three levels (≥2, 1, 0).

Although there was probably a lack of ascertainment of IM in first degree relatives IM in these relatives was analysed twice as for index IM, once considering suspected as yes and once as no. As for cancer the number of relatives were weighted by risk of IM by age. Douglas

et al (1996) investigated reports of IM and found it to be diagnosed most frequently aged 15-24 years (GP reports 48% of IM diagnosed aged 15-24 years, laboratory data 51.6%). Only 2.3% IM was diagnosed in 0-4 year olds by GP reports, 3.4% from laboratory data; 20.5% and 14.0% respectively in 5-14 year olds. Therefore, children aged 0-4 years are 2.3/48 or 3.4/51.6 (mean=0.06) times as likely to get IM as those aged 15-24 years, children aged 5-14 years are 20/48 or 14/51.6 (mean 0.35) times as likely. The weighted sum of relatives at risk equalled Σw_I across relatives where $w_I = 0.06$ for children age 0-4 years, 0.35 for children age 5-14 years and 1 for all other relatives. IM was considered as a continuous variable: number of episodes of IM in first degree relative / weighted sum of first degree relatives at risk. Thus IM in a family member was always adjusted for the number of 'at risk' relatives.

IM in someone who shared a house and IM in a close friend were also considered as exposure categories. Finally, a category of EBV illness/contact which was positive for subjects who had had previous IM and/or had a close friend with IM. Again IM was considered in two ways, one taking suspected as yes and one as no.

Factors controlled for in the analysis:

Carstairs index is a measure of socio-economic status (SES) based on residence at diagnosis (cases) or pseudo-diagnosis (controls) and is used in the analysis of YHHCCS because participation between cases and controls was biased re current SES. The address has been used to give the Carstairs index (Jarman et al, 1991) as an indicator at 5 levels of SES using data derived from the 1991 census. However, Carstair's index was unavailable for a number of cases and controls and, therefore, all analyses that adjusted for this variable involved the exclusion of a small amount of data.

Additionally risk of HD is associated with SES, young adults of higher SES have a higher risk of HD possibly related to delayed exposure to infection (see Chapter 1). Many of the risk factors analysed e.g. tonsillectomy, appendectomy, are also associated with SES.

Analyses were also performed adjusted for potential confounding factors that reflect childhood social class: parental education, sibship size, and birth order (Table 7.2).

Family health analyses were additionally adjusted for the number of first degree relatives or the weighted sum of relatives at risk.

EBV Status and Histopathology:

Paraffin embedded biopsy material was retrieved from cases and histopathological review was performed by the staff at the LRF Virus centre in Glasgow under the leadership of Professor Ruth Jarrett. EBV status was available for 103 cases and histological review for 105 cases, including all but one of those with EBV status known. Cases are described as EBV +ve if the HRS cells scored positive for either EBER ISH or LMP-1 immunohistochemistry.

HLA DPB1 typing:

Dr Malcolm Taylor in Manchester carried out typing of DPB1 alleles. Type is known for 81 cases and 73 controls but for the case and at least one matched control in only 46 matched sets. In the analysis reported here the subjects have been classified as to whether HLA DPB1*0301 was present or absent (0301yes or 0301no respectively).

Statistical methods:

Conditional logistic regression was used for all analyses on SAS 6.11 using PROC PHREG using the specification TIES=DISCRETE for N:M matched data. The results are reported as Odds Ratios (OR) with 95% Confidence Intervals (95% CI). Analyses were applied to the total series of all subjects, the subset for which EBV status was known (EBV +ve vs. EBV –ve), the subset with histological status known (NS vs. AOS (all other subtypes)), and the subset with HLA DPB1 type known (0301yes vs. 0301no). Formal tests of interaction were performed to see if any of the sub-groups were statistically significantly different from one another i.e. if EBV +ve and –ve groups had different results.

Table 7.2: Potential social class confounders related to socio-economic status.

1) Birth order: 1st, 2nd, 3rd or later
2) Number of full- and half-siblings: 0, 1, 2 or more (higher SES => smaller families).
3) Siborder: combination of number of siblings and birth order, 0 (1st born with no siblings), 1 (1st born with siblings), 2 (later born with or without siblings)
4) Housing density: number of people living in each home before aged 16 years/ number of rooms in each home averaged for each home lived in under age 16 years and divided into quartiles 1 (lowest number of people per room and, therefore, highest SES)...4 (highest number of people per room and, therefore, lowest SES)
5) Age father and mother left school: ≤15, 16, ≥17 years for biological parents only, higher age left school, higher SES in childhood
6) Educational level reached by mother and father: 1 ('O' level or lower), 2 (A level or equivalent – university diploma/certificate, HNC/HND, City and Guilds full technology certificate, BEC/TEC/BTEC/SCOTBEC/SCOTTECH higher, nursing qualification (not degree), non-degree qualification, qualification from college of technology or professional institution, BTEC foundation), 3 (degree). Higher level, higher SES in childhood.

For the total series all analyses were performed unadjusted, adjusted for Carstairs index, and then with adjustment for Carstairs index and each of the eight other potential social class confounders. To reduce the total number of tests performed only those childhood social class variables that resulted in a statistically significantly better fitting model and changed the odds ratio (OR) by more than 10% were included in the subgroup analyses. For ease of presentation results will only be given for the unadjusted and adjusted for Carstairs index analyses.

The number of positive responses for some of the variables of interest were small, with some variables having no cases or controls exposed. In these cases SAS 6.11 returns ORs of infinity and could not calculate confidence intervals. Exact methods are needed to calculate these confidence intervals.

All testing of statistical significance for conditional logistic regression modelling has examined the deviance difference against its asymptomatic chi-square distribution under the null hypothesis (Clayton & Hills, 1993). When a main effect was examined the chi-square distribution had one degree of freedom and the chi-square for the interaction (e.g. with EBV status) has a further 1 degree of freedom. This procedure is appropriate if the hierarchy of hypotheses to be considered is: no association, association common to both subgroups, association different by case subgroups. The data were also examined for the statistical significance of association of each risk factor with the cases in at least one subgroup (e.g. the deviance difference when terms for the (possibly different) effects in EBV +ve and EBV –ve HD were added to the model were tested against the chi-square distribution with 2 degrees of freedom). This is appropriate if the prior hypothesis was that the association with the exposure would be specific to EBV +ve or EBV –ve HD.

7.3 Results:

The main characteristics of cases and controls are shown in Table 7.3. Cases are predominantly of NS subtype and EBV –ve. There are slightly more male cases than female.

The results of the general health variables for the total series are in Table 7.4 and by subgroups in Tables 7.5-7.7. None of the variables analysed give rise to any significant results. The EBV +ve, AOS, and 0301yes subgroups are small and, therefore, the ORs have very wide confidence intervals. The case that had received a transplant was EBV +ve and of AOS subtype.

The results for infectious illness analyses for the total series are in Table 7.8. Combined infections (at any age) were associated with a statistically significant reduction in risk of HD when considered as a dichotomy. However, the trends across levels were not statistically significantly different from the null for combined infectious illness at any age (results not shown). Of the individual infections only measles had any statistically significant results. The results for the other individual infections were very unimpressive and are not presented. Measles was associated with a statistically significant reduction in risk of HD if occurring at any age and age 5-9 years .

The results for childhood infection within subgroups are in Tables 7.9-7.11. The total number of infections at any age were associated with a statistically significant lower risk of EBV +ve and EBV –ve HD. None of the trends across levels were statistically significant. When split by age group more infections aged ≥5 were associated with an increased risk of EBV +ve HD but in the opposite direction for EBV –ve HD. Statistical testing of the interaction by EBV status confirm that the association between EBV +ve and EBV –ve cases differ significantly. When divided by histological subtype the risk of NS HD was statistically significantly lower with more childhood infections at any age. None of the effects remained significant when split by age. The interaction by histological status was statistically significant for childhood infection at any age,

Table 7.3: YHHCCS cases and controls for analysis with selected characteristics.

Category		Cases	Controls
Total		118	235
Age (mean)		16-25 (20.8)	16-25 (20.8)
Sex	Male	62 (52.5)	125 (52.7)
	Female	56 (47.5)	112 (47.3)
Histology	LP	14 (11.9)	NA
	MC	14 (11.9)	NA
	LD	1 (1.7)	NA
	NS	81 (68.6)	NA
	NOS*	4 (3.4)	NA
EBV status	+ve	19 (16.1)	NA
	-ve	84 (71.2)	NA
	not done/missing	15 (12.7)	NA
0301 Allele	yes	21 (17.8)	NA
	no	60 (50.8)	NA
	missing	37 (31.4)	NA

* NOS (not otherwise specified)

Table 7.4: Total series results for general health

	Case		Control		Unadjusted	Adj. for Carstairs Index
Risk factor	No	No (%) yes	No	No (%) yes	OR (95%CI)	OR (95%CI)
Appendix removed at any age	118	7 (5.9)	236	17 (7.2)	0.78 (0.31-1.96)	0.97 (0.32-2.93)
Appendix removed before age 12	118	2 (1.7)	236	8 (3.4)	0.50 (0.11-2.36)	1.06 (0.19-5.87)
Appendix removed after age 12	118	5 (4.2)	236	9 (3.8)	1.06 (0.34-3.31)	0.91 (0.23-3.58)
Tonsils removed at any age	118	20 (16.9)	236	33 (14.0)	1.27 (0.68-2.37)	1.16 (0.59-2.29)
Tonsils removed before age 5.5	118	10 (8.5)	236	16 (6.8)	1.29 (0.55-3.00)	1.00 (0.39-2.59)
Tonsils removed after age 5.5	118	9 (7.6)	236	16 (6.8)	1.13 (0.49-2.60)	1.19 (0.50-2.86)
Thyroid disease	118	1 (0.8)	236	1 (0.4)	2.00 (0.13-31.98)	0.00 (0.00-∞)
Transplant[1]	118	1 (0.8)	236	0 (0.0)	∞ (0.00-∞)	∞ (0.00-∞)

numbers and percentages for unadjusted analysis only

1. Exact methods required to compute confidence intervals.

Table 7.5: General Health factors and EBV status

	EBV +ve				EBV -ve			
	Case	Control	Unadjusted	Adj Carstairs	Case	Control	Unadjusted	Adj Carstairs
Risk factor	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)
appendix removed any age	2 (10.5)	2 (5.4)	2.00 (0.28-14.20)	4.18 (0.37-47.07)	4 (4.8)	14 (8.3)	0.51 (0.16-1.64)	0.73 (0.18-2.92)
appendix removed age <12	0 (0.0)	0 (0.0)	NA	NA	2 (2.4)	7 (4.1)	0.57 (0.12-2.75)	1.11 (0.20-6.91)
appendix removed age ≥12	2 (10.5)	1 (2.7)	4.00 (0.36-44.11)	4.18 (0.37-47.07)	2 (2.4)	7 (4.1)	0.49 (0.09-2.53)	0.45 (0.05-3.97)
tonsils removed any age	4 (21.0)	7 (18.9)	1.18 (0.30-4.63)	0.77 (0.16-3.62)	13 (15.5)	21 (12.4)	1.30 (0.60-2.81)	1.18 (0.51-2.73)
tonsils removed age <5.5	2 (10.5)	3 (8.1)	1.33 (0.22-7.98)	0.67 (0.07-6.49)	8 (9.5)	11 (6.5)	1.57 (0.57-4.33)	1.36 (0.45-4.18)
tonsils removed age ≥5.5	2 (10.5)	4 (10.8)	1.00 (0.18-5.46)	0.92 (0.16-5.38)	5 (6.0)	9 (5.3)	1.12 (0.36-3.49)	1.12 (0.34-3.63)
thyroid disease[1]	0 (0.0)	0 (0.0)	NA	NA	1 (1.2)	1 (0.6)	2.00 (0.13-31.98)	0.00 (0.00-∞)
transplant[1]	1 (5.3)	0 (0.0)	∞ (0.00-∞)	∞ (0.00-∞)	0 (0.0)	0 (0.0)	NA	NA

numbers and percentages for unadjusted analysis only

1. Exact methods required to compute confidence intervals.

Table 7.6: General Health factors and histological status

	NS				AOS			
	Case	Control	Unadjusted	Adj Carstairs	Case	Control	Unadjusted	Adj Carstairs
Risk factor	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)
appendix removed any age[1]	4 (4.9)	13 (8.0)	0.57 (0.18-1.80)	0.74 (0.37-47.07)	2 (6.9)	3 (5.2)	1.44 (0.19-11.12)	∞ (0.00-∞)
appendix removed age<12[1]	2 (2.5)	5 (3.1)	0.80 (0.16-4.12)	1.52 (0.25-9.42)	0 (0.0)	3 (5.2)	0.00 (0.00-∞)	0.00 (0.00-∞)
appendix removed age≥12[1]	2 (2.5)	8 (4.9)	0.43 (0.09-2.15)	0.36 (0.04-3.09)	2 (6.9)	0 (0.0)	∞ (0.00-∞)	∞ (0.00-∞)
tonsils removed any age	12 (14.8)	22 (13.6)	1.12 (0.51-2.47)	1.04 (0.44-2.45)	6 (20.7)	9 (15.5)	1.40 (0.46-4.30)	1.20 (0.35-4.15)
tonsils removed age <5.5	7 (8.6)	12 (7.4)	1.20 (0.43-3.35)	1.01 (0.32-3.20)	3 (10.3)	3 (5.2)	2.00 (0.40-9.91)	1.31 (0.22-7.90)
tonsils removed age ≥5.5	5 (6.2)	9 (5.6)	1.12 (0.36-3.49)	1.19 (0.36-3.87)	3 (10.3)	6 (10.3)	1.00 (0.25-4.00)	1.08 (0.25-4.58)
thyroid disease[1]	1 (1.2)	1 (0.6)	2.00 (0.13-31.98)	0.00 (0.00-∞)	0 (0.0)	0 (0.0)	NA	NA
transplant[1]	0 (0.0)	0 (0.0)	NA	NA	1 (3.4)	0 (0.0)	∞ (0.00-∞)	∞ (0.00-∞)

numbers and percentages for unadjusted analysis only

1. Exact methods required to compute confidence intervals.

Table 7.7: General Health factors and DPB1*0301 allele

	0301yes				0301no			
	Case	Control	Unadjusted	Adj Carstairs	Case	Control	Unadjusted	Adj Carstairs
Risk factor	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)
appendix removed any age[1]	0 (0.0)	2 (4.8)	0.00 (0.00-∞)	0.00 (0.00-∞)	3 (5.0)	12 (9.9)	0.43 (0.11-1.63)	0.73 (0.13-4.23)
appendix removed age<12[1]	0 (0.0)	0 (0.0)	NA	NA	1 (1.7)	7 (5.8)	0.29 (0.04-2.32)	0.75 (0.08-7.37)
appendix removed age≥12[1]	0 (0.0)	2 (4.8)	0.00 (0.00-∞)	0.00 (0.00-∞)	2 (3.3)	5 (4.1)	0.68 (0.12-3.96)	0.79 (0.08-7.86)
tonsils removed any age	3 (14.3)	6 (14.3)	1.00 (0.23-4.35)	1.30 (0.26-6.39)	9 (15.0)	17 (14.0)	1.07 (0.44-2.65)	0.78 (0.29-2.12)
tonsils removed age <5.5	1 (4.8)	3 (7.1)	0.59 (0.05-7.43)	0.53 (0.04-7.70)	6 (10.0)	7 (5.8)	1.80 (0.57-5.69)	1.16 (0.32-4.27)
tonsils removed age ≥5.5	2 (9.5)	3 (7.1)	1.33 (0.22-7.98)	2.32 (0.30-17.85)	3 (5.0)	9 (7.4)	0.64 (0.17-2.50)	0.63 (0.16-2.52)
thyroid disease[1]	0 (0.0)	0 (0.0)	NA	NA	1 (1.7)	0 (0.0)	∞ (0.00-∞)	NA
transplant[1]	0 (0.0)	0 (0.0)	NA	NA	1 (1.7)	0 (0.0)	∞ (0.00-∞)	∞ (0.00-∞)

numbers and percentages for unadjusted analysis only

1. Exact methods required to compute confidence intervals.

Table 7.8: Total series results for infectious illness

	Case		Control		Unadjusted	Adj. for Carstairs Index
Risk factor	No	No (%) yes	No	No (%) yes	OR (95%CI)	OR (95%CI)
Total infections 2 v.1	118	30 (25.4)	236	85 (36.0)	0.36 (0.19-0.70)	0.33 (0.26-1.00)
3 v.1	118	31 (26.3)	236	70 (29.7)	0.49 (0.26-0.90)	0.51 (0.25-1.16)
4 v.1	118	20 (16.9)	236	38 (16.1)	0.56 (0.28-1.15)	0.54 (0.81-1.22)
2, 3, or 4 v.1	118	81 (68.6)	236	193 (81.8)	0.46 (0.27-0.78)	0.45 (0.25-0.83)
Total infections aged <5 1 v.0	118	29 (24.6)	236	64 (27.1)	0.86 (0.51-1.45)	0.91 (0.51-1.62)
2 v.0	118	15 (12.7)	236	32 (13.6)	0.89 (0.46-1.75)	0.87 (0.42-1.79)
1 or 2 v.0	118	44 (37.3)	236	96 (40.7)	0.87 (0.55-1.38)	0.89 (0.54-1.49)
Total infections aged 5-9 1 v.0	118	42 (35.6)	236	91 (38.6)	0.75 (0.43-1.31)	0.87 (0.48-1.59)
2 v.0	118	25 (21.2)	236	57 (24.2)	0.69 (0.35-1.36)	0.65 (0.30-1.40)
3 v.0	118	19 (16.1)	236	37 (15.7)	0.81 (0.39-1.67)	0.89 (0.41-1.95)
1, 2, or 3 v.0	118	86 (72.9)	236	185 (78.5)	0.75 (0.45-1.24)	0.82 (0.47-1.42)
Total infections aged ≥10 1 v.0	118	20 (16.9)	236	48 (20.3)	0.80 (0.44-1.44)	0.85 (0.44-1.63)
Measles at any age	111	42 (37.8)	225	120 (53.3)	0.52 (0.32-0.83)	0.53 (0.32-0.90)
Measles aged <5	106	18 (17.0)	209	42 (20.1)	0.72 (0.39-1.33)	0.86 (0.44-1.67)
Measles aged 5-9	106	18 (17.0)	209	60 (28.7)	0.48 (0.25-0.94)	0.49 (0.24-0.99)
Measles aged ≥10	106	1 (0.9)	209	2 (1.0)	1.00 (.09-11.03)	0.00 (0.00-∞)

numbers and percentages for unadjusted analysis only

Total infections = sum of chicken pox, German measles, measles, mumps and whooping cough episodes. For level explanation see Table 3.

Table 7.9: Infectious illness by EBV status

	EBV +ve				EBV -ve			
	Case	Control	Unadjusted	Adj Carstairs	Case	Control	Unadjusted	Adj Carstairs
Risk factor	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)
Total infect at any age 2 v.1	3 (15.8)	12 (32.4)	0.13 (0.02-0.91)	0.11 (0.01-0.92)	22 (26.2)	59 (34.9)	0.37 (0.17-0.79)	0.32 (0.13-0.80)
3 v.1	4 (21.1)	14 (37.8)	0.15 (0.02-0.99)	0.10 (0.01-0.90)	21 (25.0)	51 (30.2)	0.44 (0.21-0.91)	0.49 (0.22-1.08)
4 v.1	5 (26.3)	6 (16.2)	0.46 (0.06-3.44)	0.61 (0.07-5.32)	13 (15.5)	28 (16.6)	0.48 (0.21-1.11)	0.42 (0.17-1.03)
2,3, or 4 v.1	12 (63.2)	32 (86.5)	0.19 (0.04-0.95)	0.18 (0.03-0.95)	56 (66.7)	138 (81.7)	0.42 (0.23-0.80)	0.43 (0.21-0.86)
Total infections aged <5 1 v.0	4 (21.1)	13 (35.1)	0.41 (0.10-1.74)	0.28 (0.05-1.54)	22 (26.2)	44 (26.0)	1.04 (0.57-1.91)	1.06 (0.54-2.11)
2 v.0	1 (5.3)	3 (8.1)	0.52 (0.05-5.40)	0.45 (0.04-5.21)	12 (14.3)	23 (13.6)	1.08 (0.49-2.37)	0.92 (0.39-2.18)
1 or 2 v.0	5 (26.3)	16 (43.2)	0.43 (0.11-1.63)	0.32 (0.07-1.49)	34 (40.5)	67 (39.6)	1.05 (0.61-1.82)	1.01 (0.55-1.88)
Total infections aged 5-9 1 v.0	8 (42.1)	9 (24.3)	5.72 (0.62-52.52)	6.65 (0.69-63.64)	27 (31.1)	69 (40.8)	0.48 (0.25-0.92)	0.52 (0.26-1.07)
2 v.0	3 (15.8)	12 (32.4)	1.64 (0.15-18.15)	1.15 (0.09-15.25)	18 (21.4)	37 (21.9)	0.55 (0.25-1.23)	0.53 (0.21-1.32)
3 v.0	5 (26.3)	5 (13.5)	11.90 (0.82-173.75)	13.35 (0.86-206.54)	11 (13.1)	31 (18.3)	0.40 (0.17-0.96)	0.45 (0.18-1.15)
1,2, or 3 v.0	16 (84.2)	26 (70.3)	4.54 (0.53-38.66)	4.70 (0.55-40.22)	56 (66.7)	137 (81.1)	0.48 (0.27-0.86)	0.51 (0.27-0.98)
Total infections aged ≥10 1 v.0	6 (31.6)	8 (21.6)	1.65 (0.46-5.84)	2.31 (0.57-9.37)	13 (15.5)	37 (21.9)	0.65 (0.32-1.33)	0.62 (0.27-1.39)
Measles	9 (47.4)	21 (58.3)	0.48 (0.13-1.78)	0.48 (0.13-1.81)	27 (34.6)	84 (51.9)	0.53 (0.31-0.91)	0.41 (0.27-0.90)
Measles <5	3 (17.6)	7 (21.9)	0.30 (0.04-2.48)	0.31 (0.04-2.59)	13 (17.3)	28 (18.5)	0.88 (0.43-1.79)	0.92 (0.42-1.99)
Measles 5-9	4 (23.5)	9 (28.1)	0.77 (0.13-4.48)	0.93 (0.14-6.03)	10 (13.3)	44 (29.1)	0.40 (0.18-0.88)	0.40 (0.17-0.92)
Measles ≥10	0 (0.0)	1 (3.1)	0.00 (0.00-∞)	0.00 (0.00-∞)	1 (1.3)	1 (0.7)	2.00 (0.13-31.98)	0.00 (0.00-∞)

Numbers and percentages for unadjusted analysis only.

Table 7.10: Infectious illness by histological status

	NS				AOS			
	Case	Control	Unadjusted	Adj Carstairs	Case	Control	Unadjusted	Adj Carstairs
Risk factor	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)
Total infections at any age 2 v.1	19 (23.5)	54 (33.3)	0.28 (0.12-0.65)	0.24 (0.09-0.65)	8 (27.6)	23 (39.7)	0.45 (0.14-1.50)	0.40 (0.11-1.51)
3 v.1	21 (25.9)	51 (31.5)	0.35 (0.16-0.76)	0.42 (0.18-0.98)	6 (20.7)	18 (31.0)	0.47 (0.14-1.56)	0.34 (0.09-1.37)
4 v.1	12 (14.8)	27 (16.7)	0.40 (0.16-0.97)	0.40 (0.16-1.04)	7 (24.1)	8 (13.8)	1.07 (0.28-4.16)	0.97 (0.22-4.19)
2,3, or 4 v.1	52 (64.2)	132 (81.5)	0.34 (0.17-0.67)	0.37 (0.17-0.77)	21 (72.4)	49 (84.5)	0.54 (0.20-1.45)	0.46 (0.16-1.35)
Total infections aged <5 1 v.0	19 (23.4)	47 (29.0)	0.74 (0.39-1.39)	0.77 (0.37-1.58)	10 (34.5)	15 (25.9)	1.45 (0.53-3.99)	1.51 (0.50-4.53)
2 v.0	9 (11.1)	19 (11.7)	0.85 (0.36-2.04)	0.87 (0.33-2.29)	4 (13.8)	10 (17.2)	0.94 (0.27-3.27)	0.72 (0.18-2.82)
1 or 2 v.0	28 (34.6)	66 (40.7)	0.77 (0.43-1.36)	0.79 (0.41-1.53)	14 (48.3)	25 (43.1)	1.24 (0.50-3.07)	1.13 (0.44-2.90)
Total infections aged 5-9 1 v.0	28 (34.6)	63 (39.9)	0.58 (0.28-1.18)	0.71 (0.33-1.56)	12 (41.4)	19 (32.8)	1.36 (0.50-3.73)	1.24 (0.43-3.57)
2 v.0	18 (22.2)	40 (24.7)	0.56 (0.24-1.28)	0.53 (0.20-1.39)	5 (17.2)	14 (24.1)	0.72 (0.18-2.89)	0.77 (0.17-3.48)
3 v.0	11 (13.6)	27 (16.7)	0.50 (0.20-1.26)	0.70 (0.26-1.88)	5 (17.2)	9 (15.5)	1.18 (0.30-4.70)	0.90 (0.21-3.85)
1,2, or 3 v.0	57 (70.4)	130 (80.2)	0.56 (0.29-1.07)	0.66 (0.32-1.38)	22 (75.9)	42 (72.4)	1.16 (0.46-2.93)	1.05 (0.40-2.75)
Total infections aged ≥10 1 v.0	15 (18.5)	33 (20.4)	0.89 (0.45-1.78)	0.94 (0.43-2.06)	4 (13.8)	12 (20.7)	0.55 (0.14-2.19)	0.58 (0.13-2.51)
Measles	26 (34.2)	81 (51.9)	0.48 (0.27-0.84)	0.54 (0.29-1.00)	11 (40.7)	31 (57.4)	0.54 (0.21-1.39)	0.41 (0.14-1.23)
Measles <5	10 (14.1)	28 (19.4)	0.60 (0.27-1.34)	0.74 (0.31-1.76)	6 (22.2)	12 (23.5)	0.77 (0.25-2.37)	0.82 (0.23-2.92)
Measles 5-9	10 (14.1)	40 (27.8)	0.41 (0.18-0.96)	0.50 (0.20-1.21)	5 (18.5)	15 (29.4)	0.60 (0.18-1.98)	0.43 (0.11-1.64)
Measles ≥10	1 (1.4)	1 (0.7)	2.00 (0.13-31.98)	0.00 (0.00-∞)	0 (0.0)	1 (2.0)	0.00 (0.00-∞)	0.00 (0.00-∞)

Numbers and percentages for unadjusted analysis only.

Table 7.11: Infectious illness by DPB1*0301 status

	0301 yes				0301 no			
	Case	Control	Unadjusted	Adj Carstairs	Case	Control	Unadjusted	Adj Carstairs
Risk factor	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)
Total infections at any age 2 v.1	5 (23.8)	14 (33.3)	0.25 (0.04-1.47)	0.17 (0.02-1.28)	14 (23.3)	43 (35.5)	0.51 (0.21-1.24)	0.50 (0.18-1.39)
3 v.1	5 (23.8)	14 (33.3)	0.32 (0.07-1.45)	0.25 (0.05-1.30)	19 (31.7)	31 (25.6)	1.00 (0.44-2.29)	1.03 (0.40-2.68)
4 v.1	4 (19.0)	7 (16.7)	0.59 (0.12-3.06)	0.51 (0.09-2.80)	10 (16.7)	19 (15.7)	0.84 (0.31-2.25)	0.75 (0.25-2.24)
2,3, or 4 v.1	14 (66.7)	35 (83.3)	0.37 (0.10-1.34)	0.30 (0.08-1.19)	43 (71.7)	93 (76.9)	0.76 (0.37-1.54)	0.74 (0.33-1.67)
Total infections aged <5 1 v.0	4 (19.0)	16 (38.1)	0.36 (0.08-1.54)	0.38 (0.08-1.96)	18 (30.0)	24 (19.8)	1.73 (0.83-3.61)	1.68 (0.72-3.94)
2 v.0	2 (9.5)	2 (4.8)	1.09 (0.12-9.54)	2.93 (0.23-38.02)	7 (11.7)	18 (14.9)	0.91 (0.34-2.41)	0.76 (0.27-2.15)
1 or 2 v.0	6 (28.6)	18 (42.9)	0.43 (0.11-1.70)	0.57 (0.14-2.30)	25 (41.7)	42 (34.7)	1.39 (0.73-2.66)	1.23 (0.60-2.53)
Total infections aged 5-9 1 v.0	9 (42.3)	15 (35.7)	0.37 (0.07-2.10)	0.25 (0.03-2.41)	20 (33.3)	43 (35.5)	0.94 (0.43-2.07)	1.09 (0.46-2.56)
2 v.0	3 (14.3)	10 (23.8)	0.16 (0.02-1.28)	0.06 (0.004-0.89)	15 (25.0)	34 (28.1)	0.89 (0.35-2.25)	0.90 (0.30-2.65)
3 v.0	3 (14.3)	11 (26.2)	0.13 (0.01-2.41)	0.06 (0.004-0.93)	11 (18.3)	15 (12.4)	1.46 (0.53-4.04)	1.55 (0.51-4.74)
1,2, or 3 v.0	15 (71.4)	36 (85.7)	0.27 (0.05-1.47)	0.14 (0.02-1.28)	46 (76.7)	92 (76.0)	1.01 (0.50-2.07)	1.12 (0.51-2.45)
Total infections aged ≥10 1 v.0	4 (19.0)	8 (19.0)	1.00 (0.29-3.50)	0.86 (0.23-3.23)	8 (13.3)	20 (16.5)	0.77 (0.32-1.82)	0.61 (0.21-1.77)

Numbers and percentages for unadjusted analysis only.

even though the ORs for NS and AOS HD were both in the same direction. Measles was again associated with a lower risk of HD in both subgroups but the OR was only statistically significant for NS HD. Due to the small number of cases with HLA DPB1 type known none of the results were statistically significant.

The impact of IM on risk of HD in the total series is shown in Table 7.12. IM in the index was associated with an increased risk of HD whether suspected IM was treated as yes or no or if IM in the year prior to diagnosis was included or not. However, the only result that was statistically significant was that for IM at any time with suspected IM treated as no. IM in a family member, in a member of the household, or in a close friend was also associated with an elevated risk of HD but none of the results were statistically significant.

The results for IM by subgroup are in Tables 7.13-7.15. IM is associated with a statistically significantly elevated risk of EBV +ve HD but has little association with EBV –ve HD. Statistical testing of the interaction by EBV status confirms that the associations with EBV +ve and EBV –ve cases differ significantly. The effects of IM in a family member were in the opposite direction with a decreased risk of EBV +ve HD and an increased risk of EBV –ve but the interaction by EBV status was not significant. Risk of both EBV +ve and EBV –ve HD was increased following IM in a household member or a close friend. There was a statistically significant increase in risk of EBV +ve HD following EBV illness/contact (OR 5.80, 95% CI 1.20-28.17) but not for EBV –ve HD. None of the results for IM by histological subtype were statistically significant. IM in the index, in a family member, or the household were all associated with an increased risk of 0301yes HD. However, the results were only statistically significant when treating suspected IM in the index as yes. EBV exposure was statistically significantly associated with an increased risk of 0301yes HD. There was little effect of any of the IM variables on risk of HD when *0301 was not present. The interaction for the EBV contact with allele subgroups was statistically significant.

Table 7.12: Total series results for infectious mononucleosis

	Case		Control		Unadjusted	Adj. for Carstairs Index
Risk factor	No	No (%) yes	No	No (%) yes	OR (95%CI)	OR (95%CI)
IM[1]	118	25 (21.2)	233	32 (13.7)	1.69 (0.95-3.00)	1.87 (0.95-3.66)
IM[2]	118	19 (16.1)	233	21 (9.0)	2.03 (1.01-4.08)	2.43 (1.10-5.33)
IM (not including year prior to diagnosis)[1]	118	21 (17.8)	233	30 (12.9)	1.46 (0.80-2.66)	1.46 (0.72-2.94)
IM (not including year prior to diagnosis)[2]	118	16 (13.6)	233	19 (8.2)	1.80 (0.88-3.71)	1.93 (0.87-4.28)
IM in family[1, 3]	118	22 (18.6)	237	35 (14.8)	2.07 (0.59-7.31)	2.21 (0.56-8.78)
IM in family[2, 3]	118	24 (20.3)	237	30 (12.7)	2.82 (0.77-10.31)	3.22 (0.77-13.54)
IM in household[1]	115	25 (21.7)	230	43 (18.7)	1.21 (0.70-2.10)	1.45 (0.78-2.69)
IM in household[2]	115	24 (20.9)	230	38 (16.1)	1.46 (0.84-2.34)	1.66 (0.87-3.15)
IM in close friend	118	41 (34.7)	236	75 (31.8)	1.15 (0.72-1.85)	
EBV illness/contact[1]	118	43 (36.4)	236	68 (28.8)	1.43 (0.88-2.30)	1.61 (0.94-2.77)
EBV illness/contact[2]	118	36 (30.5)	236	54 (22.9)	1.46 (0.84-2.40)	1.77 (1.02-3.08)

numbers and percentages for unadjusted analysis only

1. suspected IM = yes.
2. suspected IM = no.
3. Analysis adjusted for numbers of relatives at risk (see methods).

Table 7.13: Infectious mononucleosis by EBV status

	EBV +ve				EBV -ve			
	Case	Control	Unadjusted	Adj Carstairs	Case	Control	Unadjusted	Adj Carstairs
Risk factor	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)
IM[1,3]	6 (31.6)	3 (8.1)	9.16 (1.07-78.31)	∞ (0.00-∞)	16 (19.0)	22 (13.3)	1.51 (0.76-3.02)	1.50 (0.68-3.33)
IM (not previous year)[1,3]	6 (31.6)	2 (5.4)	∞ (0.00-∞)	∞ (0.00-∞)	12 (14.3)	21 (12.7)	1.14 (0.55-2.38)	1.00 (0.42-2.41)
IM[2,3]	6 (31.6)	3 (8.1)	9.16 (1.07-78.31)	∞ (0.00-∞)	11 (13.1)	15 (9.0)	1.53 (0.66-3.54)	1.60 (0.63-4.07)
IM (not previous year)[2,3]	6 (31.6)	2 (5.4)	∞ (0.00-∞)	∞ (0.00-∞)	8 (9.5)	14 (8.4)	1.13 (0.46-2.76)	1.09 (0.40-2.97)
IM in family[1, 4]	4 (21.1)	4 (10.8)	0.55 (0.03-11.79)	0.87 (0.04-21.89)	15 (17.9)	28 (16.5)	2.15 (0.42-10.92)	2.59 (0.46-14.57)
IM in family[2, 4]	4 (21.1)	4 (10.8)	0.55 (0.03-11.79)	0.87 (0.04-21.89)	15 (17.9)	23 (13.5)	3.67 (0.67-20.27)	4.87 (0.76-31.38)
IM in household[1]	5 (26.3)	6 (16.2)	1.76 (0.45-6.86)	2.23 (0.49-10.09)	17 (21.0)	33 (20.1)	1.09 (0.57-2.08)	1.23 (0.63-2.41)
IM in household[2]	5 (26.3)	6 (16.2)	1.76 (0.45-6.86)	2.23 (0.49-10.09)	16 (19.8)	28 (17.1)	1.41 (0.69-2.87)	1.67 (0.79-3.55)
IM in close friend	5 (26.3)	9 (24.3)	1.15 (0.33-3.97)	1.30 (0.35-4.83)	30 (35.9)	57 (33.7)	1.09 (0.63-1.91)	0.87 (0.46-1.65)
EBV illness /contact[1]	9 (47.4)	7 (18.9)	3.96 (1.03-15.21)	5.80 (1.20-28.17)	30 (35.7)	51 (30.2)	1.29 (0.74-2.26)	1.45 (0.76-2.74)
EBV illness /contact[2]	9 (47.4)	7 (18.9)	3.96 (1.03-15.21)	5.80 (1.20-28.17)	24 (28.6)	41 (24.3)	1.23 (0.70-2.19)	1.44 (0.79-2.82)

numbers and percentages for unadjusted analysis only

1. suspected IM = yes.

2. suspected IM = no.
3. Exact methods required to compute confidence intervals.
4. Analyses adjusted for numbers of relatives at risk (see methods)

Table 7.14: Infectious mononucleosis by histologic status

	NS				AOS			
	Case	Control	Unadjusted	Adj Carstairs	Case	Control	Unadjusted	Adj Carstairs
Risk factor	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)
IM[1]	13 (16.0)	19 (11.9)	1.44 (0.68-3.03)	1.66 (0.72-3.83)	9 (31.0)	8 (14.0)	2.47 (0.86-7.08)	2.24 (0.60-8.30)
IM (not previous year)[1]	11 (13.6)	19 (11.9)	1.19 (0.55-2.57)	1.31 (0.55-3.10)	8 (27.6)	6 (10.5)	3.16 (0.93-10.78)	2.28 (0.52-10.05)
IM[2]	10 (12.3)	11 (6.9)	2.10 (0.81-5.45)	2.23 (0.81-6.13)	6 (20.7)	7 (12.3)	1.70 (0.54-5.39)	2.10 (0.54-8.27)
IM (not previous year)[2]	8 (9.9)	11 (6.9)	1.56 (0.59-4.15)	1.61 (0.57-4.50)	6 (20.7)	5 (8.8)	2.51 (0.69-9.09)	2.70 (0.62-11.69)
IM in family[1, 3]	14 (17.3)	29 (17.8)	0.53 (0.08-3.50)	0.55 (0.07-4.28)	5 (17.2)	5 (8.6)	6.55 (0.57-75.68)	23.83 (0.79-719.77)
IM in family[2, 3]	14 (17.3)	24 (14.7)	0.89 (0.14-5.90)	1.02 (0.13-7.84)	5 (17.2)	5 (8.6)	6.55 (0.57-75.68)	23.83 (0.79-719.77)
IM in household[1]	17 (21.8)	32 (20.1)	1.11 (0.58-2.14)	1.35 (0.66-2.76)	5 (17.2)	9 (16.1)	1.13 (0.35-3.70)	2.01 (0.52-7.84)
IM in household[2]	16 (20.5)	27 (17.0)	1.26 (0.64-2.49)	1.61 (0.76-3.41)	5 (17.2)	9 (16.1)	1.13 (0.35-3.70)	2.01 (0.53-7.84)
EBV illness /contact[1]	27 (33.3)	49 (30.2)	1.15 (0.65-2.03)	1.39 (0.73-2.64)	12 (41.4)	12 (20.7)	3.00 (1.01-8.89)	3.02 (0.89-10.23)
EBV illness /contact[2]	23 (28.4)	38 (23.5)	1.26 (0.70-2.26)	1.54 (0.81-2.93)	9 (31.0)	11 (19.0)	2.00 (0.68-5.88)	2.89 (0.83-10.05)

numbers and percentages for unadjusted analysis only

1. suspected IM = yes.

2. suspected IM = no.

3. Analyses adjusted for numbers of relatives at risk (see methods)

Table 7.15: Infectious mononucleosis by DPB1*0301 status

	0301 yes				0301 no			
	Case	Control	Unadjusted	Adj Carstairs	Case	Control	Unadjusted	Adj Carstairs
Risk factor	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)
IM[1]	8 (38.1)	4 (9.5)	6.61 (1.37-31.76)	5.65 (1.15-27.92)	12 (20.0)	18 (15.1)	1.40 (0.62-3.16)	1.42 (0.52-3.83)
IM (not previous year)[1]	7 (33.3)	4 (9.5)	5.62 (1.14-27.79)	4.68 (0.91-23.94)	9 (15.0)	17 (14.3)	1.05 (0.44-2.49)	0.82 (0.28-2.45)
IM[2]	6 (31.6)	4 (9.5)	4.65 (0.91-23.82)	4.54 (0.89-23.23)	9 (15.0)	10 (8.4)	2.02 (0.73-5.60)	2.46 (0.73-5.60)
IM (not previous year)[2]	5 (23.8)	4 (9.5)	3.68 (0.68-19.86)	3.59 (0.67-19.27)	7 (11.7)	9 (7.6)	1.57 (0.56-4.42)	1.36 (0.37-4.98)
IM in family[1, 3]	5 (23.8)	4 (9.5)	10.47 (0.17-656.10)	86.19 (0.53-14153)	9 (15.0)	21 (17.2)	0.95 (0.18-5.11)	0.74 (0.11-5.05)
IM in family[2, 3]	5 (23.8)	3 (7.1)	18.61 (0.24-1453.7)	225.43 (0.86-58867)	9 (15.0)	18 (17.2)	1.27 (0.24-6.76)	1.07 (0.16-7.08)
IM in household[1]	5 (23.8)	6 (14.3)	1.67 (0.51-5.46)	2.58 (0.66-10.08)	12 (20.7)	24 (20.5)	1.07 (0.48-2.41)	1.26 (0.49-3.21)
IM in household[2]	5 (23.8)	5 (11.9)	2.00 (0.58-6.91)	3.37 (0.78-14.66)	11 (19.0)	21 (17.9)	1.14 (0.50-2.60)	1.38 (0.52-3.64)
EBV illness /contact[1]	12 (57.1)	8 (19.0)	4.65 (1.46-14.76)	5.69 (1.62-19.99)	20 (33.3)	39 (32.0)	1.08 (0.53-2.17)	1.06 (0.46-2.43)
EBV illness /contact[2]	10 (47.6)	7 (16.7)	3.48 (1.17-10.34)	5.10 (1.48-17.54)	16 (26.7)	29 (23.8)	1.19 (0.57-2.47)	1.28 (0.55-3.02)

numbers and percentages for unadjusted analysis only

1. suspected IM = yes.

2. suspected IM = no.

3. Analyses adjusted for numbers of relatives at risk (see methods)

Family health results for the total series are in Table 7.16. None of the results are statistically significant. Results of the family health variables by subgroup are in Tables 7.17-7.19. Due to the small number of EBV +ve cases the results for EBV –ve HD are very similar to those for the total series.

Table 7.16: Total series results for family health

	Case		Control		Unadjusted	Adj. for Carstairs Index
Risk factor	No	No (%) yes	No	No (%) yes	OR (95%CI)	OR (95%CI)
Haematological malignancy in a 1st degree relative[1]	118	2 (1.7)	237	0 (0.0)	∞ (0.00-∞)	∞ (0.00-∞)
Young onset or haematological malignancy in a 1st degree relative	118	14 (11.9)	237	23 (9.7)	1.30 (0.63-2.69)	1.23 (0.53-2.87)
Any malignancy in a 1st degree relative	118	14 (11.9)	237	24 (10.1)	1.23 (0.60-2.53)	1.15 (0.50-2.66)
Multiple sclerosis in a 1st degree relative[1]	117	2 (1.7)	236	0 (0.0)	∞ (0.00-∞)	∞ (0.00-∞)
Pernicious anaemia in a 1st degree relative	116	1 (0.9)	231	1 (0.4)	2.00 (0.13-31.98)	1.50 (0.09-25.65)
Rheumatoid arthritis in a 1st degree relative	116	7 (6.0)	231	13 (5.6)	1.11 (0.42-2.93)	1.62 (0.53-4.93)
Thyroid disease in a 1st degree relative	116	4 (3.4)	231	8 (3.5)	1.00 (0.29-3.50)	0.68 (0.17-2.70)
Combined auto-immune disease in a 1st degree relative	118	13 (11.0)	236	21 (8.9)	1.31 (0.61-2.78)	1.24 (0.53-2.91)

numbers and percentages for unadjusted analysis only

1. Exact methods required to compute confidence intervals.

Table 7.17: Family Health factors by EBV status

	EBV +ve				EBV –ve			
	Case	Control	Unadjusted	Adj Carstairs	Case	Control	Unadjusted	Adj Carstairs
Risk factor	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)
haematological malignancy[1]	0 (0.0)	0 (0.0)	NA	NA	2 (2.4)	0 (0.0)	∞ (0.00-∞)	∞ (0.00-∞)
haemato/young onset malignancy	2 (10.5)	4 (10.8)	1.00 (0.18-5.46)	0.96 (0.17-5.35)	10 (11.9)	15 (8.8)	1.49 (0.60-3.70)	1.90 (0.61-5.03)
any malignancy	2 (10.5)	4 (10.8)	1.00 (0.18-5.46)	0.96 (0.17-5.35)	10 (11.9)	16 (9.4)	1.37 (0.56-3.32)	1.68 (0.56-5.03)
multiple sclerosis[1]	0 (0.0)	0 (0.0)	NA	NA	2 (2.4)	0 (0.0)	∞ (0.00-∞)	∞ (0.00-∞)
pernicious anaemia	0 (0.0)	0 (0.0)	NA	NA	1 (1.2)	1 (0.6)	2.00 (0.13-31.98)	1.59 (0.09-27.78)
rheumatoid arthritis[1]	1 (5.9)	1 (2.9)	2.00 (0.13-31.98)	∞ (0.00-∞)	5 (6.0)	11 (6.6)	0.92 (0.30-2.85)	1.59 (0.47-5.40)
thyroid disease[1]	0 (0.0)	1 (2.9)	0.00 (0.00-∞)	0.00 (0.00-∞)	4 (4.8)	5 (3.0)	1.71 (0.42-7.05)	1.22 (0.25-5.84)
auto-immune dis. combined	1 (5.3)	2 (5.4)	1.00 (0.09-11.03)	1.81 (0.08-40.22)	11 (13.1)	16 (9.5)	1.52 (0.64-3.62)	1.63 (0.62-4.31)

numbers and percentages for unadjusted analysis only

1. Exact methods required to compute confidence intervals.

Table 7.18: Family Health factors by histological status

	NS				AOS			
	Case	Control	Unadjusted	Adj Carstairs	Case	Control	Unadjusted	Adj Carstairs
Risk factor	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)
haematological malignancy[1]	1 (1.2)	0 (0.0)	∞ (0.00-∞)	∞ (0.00-∞)	0 (0.0)	0 (0.0)	NA	NA
haemato/young onset malignancy	10 (12.3)	12 (7.4)	1.97 (0.76-5.12)	1.90 (0.56-6.41)	3 (10.3)	6 (10.3)	1.00 (0.25-4.00)	1.27 (0.29-5.56)
any malignancy	10 (12.3)	13 (8.0)	1.76 (0.70-4.45)	1.63 (0.51-5.20)	3 (10.3)	6 (10.3)	1.00 (0.25-4.00)	1.27 (0.29-5.56)
multiple sclerosis[1]	2 (2.5)	0 (0.0)	∞ (0.00-∞)	∞ (0.00-∞)	0 (0.0)	0 (0.0)	NA	NA
pernicious anaemia	1 (1.2)	1 (0.6)	2.00 (0.13-31.98)	1.55 (0.09-26.87)	0 (0.0)	0 (0.0)	NA	NA
rheumatoid arthritis	4 (4.9)	8 (5.0)	1.04 (0.30-3.64)	1.56 (0.40-6.10)	2 (7.4)	4 (7.1)	1.00 (0.16-6.42)	2.08 (0.12-36.89)
thyroid disease	2 (2.5)	5 (3.1)	0.77 (0.13-4.48)	0.36 (0.04-3.54)	2 (7.4)	2 (3.6)	2.00 (0.28-14.20)	1.69 (0.23-12.50)
auto-immune dis. combined	8 (9.9)	13 (8.0)	1.33 (0.46-3.59)	1.16 (0.38-3.56)	4 (13.8)	6 (10.3)	1.38 (0.36-5.29)	1.76 (0.33-9.33)

numbers and percentages for unadjusted analysis only

1. Exact methods required to compute confidence intervals.

Table 7.19: Family Health factors by DPB1*0301 allele status

	0301yes				0301no			
	Case	Control	Unadjusted	Adj Carstairs	Case	Control	Unadjusted	Adj Carstairs
Risk factor	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)
haematological malignancy[1]	1 (4.8)	0 (0.0)	∞ (0.00-∞)	∞ (0.00-∞)	1 (1.7)	0 (0.0)	∞ (0.00-∞)	∞ (0.00-∞)
haemato/young onset malignancy	2 (9.5)	5 (11.9)	0.77 (0.13-4.48)	0.71 (0.12-4.20)	8 (13.3)	12 (9.8)	1.50 (0.55-4.14)	1.24 (0.36-4.29)
any malignancy	2 (9.5)	5 (11.9)	0.77 (0.13-4.48)	0.71 (0.12-4.20)	8 (13.3)	13 (10.7)	1.35 (0.50-3.61)	1.11 (0.33-3.70)
multiple sclerosis	0 (0.0)	0 (0.0)	NA	NA	0 (0.0)	0 (0.0)	NA	NA
pernicious anaemia[1]	0 (0.0)	0 (0.0)	NA	NA	0 (0.0)	1 (0.8)	0.00 (0.00-∞)	0.00 (0.00-∞)
rheumatoid arthritis[1]	1 (5.0)	1 (2.4)	2.00 (0.13-31.98)	0.00 (0.00-∞)	3 (5.1)	8 (6.8)	0.74 (0.17-3.14)	1.75 (0.33-9.27)
thyroid disease[1]	0 (0.0)	1 (2.4)	0.00 (0.00-∞)	0.00 (0.00-∞)	4 (6.8)	3 (2.5)	3.35 (0.59-18.88)	2.13 (0.33-13.65)
auto-immune dis. combined[1]	1 (4.8)	2 (4.8)	1.00 (0.09-11.03)	0.00 (0.00-∞)	7 (11.7)	11 (9.1)	1.38 (0.48-3.95)	1.91 (0.56-6.59)

numbers and percentages for unadjusted analysis only

1. Exact methods required to compute confidence intervals.

7.4 Discussion:

Many different a priori hypotheses related to past medical history have been investigated by YHHCCS. Very few of these have given results that were statistically significant. In part this is due to the age of the cases and controls, there has not been enough time to develop the condition of interest e.g. previous cancer. Also, the small numbers in some of the calculations, especially within sub-groups, mean that many of the comparisons lack statistical power. In these instances the absence of statistically significant effects does not necessarily reflect the absence of a causal relationship. Certain aspects of the results should be interpreted with care, especially those which are non-significant and have very large or small point estimates for the odds ratio. While bearing these caveats in mind the results from YHHCCS can be compared with published work. Of the variables studied four are most often reported in the literature: IM, appendectomy, removal of tonsils/adenoids, and family history of cancer.

Previous IM has been implicated as a risk factor for HD for a variety of reasons (see Chapter 1). YHHCCS data confirmed the statistically significant increased risk of HD following IM in the total series, with and without adjustment for Carstairs index. If IM in the year prior to diagnosis was excluded the results for the total series were no longer statistically significant. When analysed as separate HD subtypes the effect of prior IM was seen for all groups after adjustment, but the results for EBV –ve HD were weakest. When IM in the year prior to diagnosis was removed associations were only seen for EBV +ve HD and AOS HD. Household exposure to IM significantly elevated the risk of HD in the total series and greatly so in the EBV +ve subgroup, but risk of EBV –ve HD was not significantly elevated. The association of IM was not completely restricted to EBV +ve HD, but the results are strongest in this subgroup. This is in contrast to the results of a case series comparison by Sleckman et al (1998) which found no association between prior IM and EBV status. This study analysed similar numbers to YHHCCS but with a much broader age range at diagnosis (16-55 years). The results from

YHHCCS suggest a specific causal association of recent EBV exposure with EBV +ve HD. The data are also consistent with a weak positive association of EBV –ve HD with previous IM; this could be interpreted in terms of lifestyles and environments predisposing to late exposure to EBV and agents with similar transmission routes.

The possible effect of appendectomy on risk of HD is discussed in Chapter 1. However, no studies have found statistically significant results. YHHCCS analyses are based on reports of appendicitis and also show no evidence of an effect of appendicitis on HD risk in the total series.

Six published studies have investigated risk of HD following tonsillectomy in young adults after adjusting for SES. Of these four have found a statistically significant increase in risk of HD (see Chapter 1). The tonsils can act as filter barriers to infective agents (Vianna et al, 1971). Data from YHHCCS showed a very slight non-significant increased risk of HD following tonsillectomy at any age after controlling for SES. Some authors have suggested that the role of tonsilar tissue in immune function is most important in childhood and decreases with age (Brandzaeg, 1987; Andersson et al, 1994; Liaw et al, 1997) but the YHHCCS total series produced the opposite pattern, with a non-statistically significant increase in risk following late removal. YHHCCS data does contribute further information on HD risk following tonsillectomy. The results of studies controlling for social class and the presence of a plausible biologic mechanism mean the risk of tonsillectomy can still not be discounted, even thought the association is not strong.

It is interesting to note that the only person to receive a transplant was a case that was EBV +ve and of AOS subtype. This concurs with the case reports of HD following transplantation (see Chapter 1).

Very few studies have attempted to assess risk of HD by number of childhood infections and none have attempted to form composites of infections in total or at specific age groups. Paffenbarger et al (1977) found risk of HD to be decreased following several childhood

infections but the only statistically significant result was for whooping cough. The combined variable for childhood infections at any age and in the 3 age groups of interest (0-4, 5-9, 10+ years) is protective in the total YHHCCS series. The same pattern is seen for measles alone and it is possible that measles is driving the result for the combined infections. There are not many differences in the effect of childhood infections by subgroup apart from the finding that EBV +ve HD is associated with more infections aged ≥5 years and EBV –ve HD is associated with fewer infections in the same age group. This finding could, however, be due to chance as a large number of statistical tests were performed.

The presence of HD in a first degree relative is relatively rare, estimated at only 4.5% in published series (Ferraris et al, 1997) but it is usually found to be a statistically significant risk factor in epidemiological studies. The presence of a haematological malignancy in a first degree relative conferred increased risk of HD in the total YHHCCS series. It is interesting to note that the subtype most associated with an increased risk following a haematological or young onset malignancy or any malignancy at all is NS HD. Thus, YHHCCS results agree with observations by Ferraris et al (1997) and Mack et al (1995) who thought familial HD could be linked to certain kinds of HD, especially NS HD.

Most studies of family history of cancer have used the number of affected relatives in the family to calculate relative risk but they have not considered the heterogeneity of the familial risk of cancer in a systematic way. Yang et al (1998) compared a simple classification of family history of breast cancer (yes/no) to the method of using a quantitative family history score (FHS) based on a comparison of the observed number of cases in a family compared with the expected number during the observation period, taking into account some of the co-variates of the family, in this cases age, sex, race, birth cohort, as predictors of breast cancer mortality. With the use of the FHS, about one third of the women with a positive family history of breast cancer were at no higher risk for breast cancer mortality than those without a family history of the disease. As a

quantitative relative risk for each family history, FHS gave a better fit to the data, and it provided an incremental improvement in the predictive accuracy of developing fatal breast cancer. In YHHCCS a FHS was constructed based only on the age of family members. This score allowed the production of better fitting models compared with adjusting only for the total number of close relatives. In the total series adjusting for the total number of close relatives did not significantly improve the fit of any of the models of family history of cancer after adjustment for Carstairs index. However, adjusting for the risk of cancer in the family, weighted by age only, statistically significantly improved the model's fit in every case ($0.01<p<0.001$). Thus, even a simple attempt at creating a FHS can be effective.

Case-control studies can both test and generate hypotheses. YHHCCS data support a role for IM in increasing risk of HD. YHHCCS results suggest a specific causal association of recent EBV exposure (IM) with EBV +ve HD. Another new finding is the protective effect of childhood infectious illness. YHHCCS data to a lesser extent agrees with previous work on family history of cancer as a risk factor for HD. The role of appendectomy, tonsillectomy and auto-immune diseases could not be adequately analysed due to small numbers of positive responses. Larger studies of a wider range of ages are needed to make more conclusive statements about these variables.

8 Results of SNEHD General & Family Health Analyses

8.1 Introduction:

There is a body of evidence relating to past own and family medical history and their effect on HD risk. Own and family medical history factors are tested here on data from the Scotland and Newcastle study of Hodgkin's Disease (SNEHD). The analysis of SNEHD data has the ability to test hypotheses generated by YHHCCS. Investigations were performed for the same childhood infectious illnesses (chicken pox, German measles, measles, mumps, and whooping cough) individually and in total at specific ages of occurrence. SNEHD data also included a much larger number of infections which allowed a more comprehensive analysis of the impact of infectious illness on HD risk. The importance of infectious illness on risk of HD is discussed in Chapter 1. Several studies have looked at specific infections (Vianna et al, 1971; Newell et al, 1973; Paffenbarger et al, 1977; Andersson & Isager, 1978; Kirchoff et al, 1980; Gutensohn & Cole, 1981; Evans & Gutensohn, 1984). However, prior to the YHHCCS analysis none assessed a combined number of infectious illnesses.

SNEHD covers a wide range of ages (16-74 years) compared with YHHCCS (16-24 years) (see chapter 7) and allows the investigation of risk factors in subgroups based on age at diagnosis. A causal role for EBV has now been found in around 35% of HD cases in the developed world (see Chapter 1). As for YHHCCS, SNEHD analysis will compare the epidemiological risk factors by EBV status. Risk factors for HD will be analysed in the total series and sub-groups based on age group (16-34 and 50+ years being the main groups of interest) and EBV status. Age group 16-24 years was also analysed as a direct comparison for YHHCCS but the results will not be presented in detail. A basic family history score (FHS) for cancer risk in family members based on age was calculated for the YHHCCS analysis. For SNEHD a more complicated FHS was developed based on age and sex to see if this resulted in better fitting regression models.

8.2 Materials and Methods:

The SNEHD study is described in Chapter 4.

Interview Data:

Trained interviewers conducted face-to-face interviews. The period of interest was from birth up to the date of diagnosis of cases and a date of pseudo-diagnosis for controls. Information was requested on proxies for exposure to infection, past history of infectious illness, past medical history of the index and family, a limited history of infectious illness in friends and family, and sporting activities.

Categories of interest:

The sections of the questionnaire used to record details of general health, family health, and infectious illnesses are in Appendix B. All analyses are based on reports of illness at interview, without confirmation from medical records.

Index General Health analyses: Information on previous malignancy in the index was recorded at interview. These reports were divided into any previous malignancy and previous haematological malignancy and analysed as dichotomous variables (yes/no). The questionnaire allowed the reporting of several auto-immune diseases using closed questions on insulin dependent diabetes, thyroid disease, rheumatoid arthritis, pernicious anaemia, multiple sclerosis, systemic lupus and Sjogrens syndrome. These variables were analysed as dichotomous responses (yes/no) but due to the small number of positive responses the results are not presented. A combined auto-immune disease variable was constructed from a combination of the answers to all the above closed questions and an open question on 'immune disorders'. The answers to closed questions on eczema and asthma were analysed individually and also combined with other allergies (taken from the answers to open questions on 'serious or unusual illness') to form a variable called atopic.

The effect of appendicitis and removal of the appendix were analysed separately at any age, at age younger than the mean of controls (aged <16 yrs), and older than the mean age of controls (≥16 yrs). Removal of Tonsils/adenoids was analysed at any age, age <5.5,

age 5.5-10, and ≥11 years. Three categories were used because of difficulties in splitting cases and controls by age as some people reported a school stage e.g. when at primary school, rather than as a year or an age. Receiving a blood transfusion was analysed as a dichotomy (yes/no). An open question on chronic infections was asked with responses analysed as dichotomies (yes/no) at any age, age 0-4, 5-10, ≥11 or ≥5 years. The numbers of specific chronic infections were small, however, answers were divided into general groups for analysis: lower respiratory tract infection (lrti), upper respiratory tract infection (urti), chronic tonsillitis/throat infection, chronic tonsillitis/throat infection/lymph glands, urti plus these, or 'other' chronic infections.

Infectious illness analyses: Childhood infectious illnesses (chicken pox, German measles, measles, mumps and whooping cough) were recorded as a simple yes/no variable in the questionnaire. Data were then analysed as dichotomies at any age and in age groups: 0-4, 5-10, ≥11, and ≥5 years. The answers to questions on these infectious illnesses were combined to give a total number of childhood infections at all ages and at ages 0-4, 5-10, ≥11, and ≥5 years. The numbers of childhood infections were split into levels the same as those for YHHCCS. The total number of infections at all ages, and aged 0-4, 5-10, ≥11, and ≥5 years were also analysed as dichotomies in which the lowest level (generally none) is taken as a reference group with the rest combined. In addition reports of other individual infectious illness were recorded in closed questions on the questionnaire: HSV-1, HSV-2, shingles, and pneumonia. These were analysed in the same way as the childhood infections. There was also an open question on 'other serious infectious disease', however only scarlet fever was recorded with any frequency. For the herpes viruses (including chicken pox) the effect of very late exposure (in this case meaning first exposure after leaving school) was also investigated. The herpes viruses were combined to give a total number of herpes virus infections at any age and in the age groups described previously. A variable of 'additional' infections was constructed as a combination of several infections reported more frequently in

response to the open question on 'Any other serious infectious disease'. 'Additional' infections included: tuberculosis, diphtheria, meningitis, malaria, scarlet fever, rheumatic fever and the closed question on pneumonia. 'Additional' and childhood infections were also combined for analysis. Finally a variable combining childhood infections, herpes viruses, 'additional' infections, and all chronic infections was constructed (based on a summation of the answers for individual infectious illnesses to prevent double counting e.g. for chicken pox). Blind (to the case/control status) assessment of the distribution of frequencies of infectious illness led to sensible strata for the number of episodes for analysis.

To allow for any confusion in reporting close to diagnosis or pseudo-diagnosis all infectious illnesses, separately or in combination, were analysed including or excluding the year prior to the date of interest. However, the results for both analyses were very similar and therefore the results including the last year are the only ones presented. If no age was given the data were considered to be missing when analysed by age of illness. For total infectious illnesses missing values were treated as zeros in the sums, unless the data were missing for all individual infections, when the total was set to missing.

Interviewees could respond to the question on IM in three ways: yes, no, not sure. If the interviewee answered positively a further question was asked to elicit if the IM had been confirmed by a blood test. IM was analysed in 3 separate ways. Definite IM was IM confirmed by a blood test i.e. the interviewee had to respond yes to having IM and yes to it being confirmed by a blood test, all other responses were considered as no. Suspected IM allowed IM that had not been confirmed by a blood test or if the index was not sure to be yes with only interviewees responding specifically no being considered as no. IM in levels was also analysed with 1 equal to definitely no, 2 IM not sure or not confirmed by a blood test, 3 equal to IM confirmed by a blood test.

Family health analyses: the presence of malignancy in first degree relatives was elicited from the questionnaire. Responses to this question were divided into HD, haematological malignancy, any young onset cancer (aged <50 yrs), any haematological or young onset

malignancy, breast, uterine, ovarian or any female cancers combined, or any malignancy for analysis. If numbers permitted the number of relatives with the condition of interest was tested as either two (≥1v.0) or three levels (0, 1, ≥2). However, the number of people with ≥2 relatives was very small and, therefore, only the results for the dichotomy are presented. In the literature most analyses of cancer in a first degree relative adjust for the number of relatives. However, as described in the YHHCCS chapter, this may not be enough due to the differing risk of cancer by age and sex. The cancer variables were adjusted for number of first degree relatives and for risk of cancer in a first degree relative adjusted for age in 3 age groups (0-24, 25-49, ≥50 yrs) (as for YHHCCS). In addition a finer weighting was constructed using 10 year age groups and sex by the method outlined in Chapter 7.

The questionnaire asked for information on presence in the family of insulin dependent diabetes (IDDM), multiple sclerosis (MS), pernicious anaemia (PA), rheumatoid arthritis (RA) (at any age and age <50 yrs), and thyroid disease. The answers to these closed questions on auto-immune diseases were combined with the response to the open questions on 'immune disorders', 'inherited illness', 'unusual illness', and 'anything else which runs in the family' to form a composite auto-immune disease variable. Again, if possible the number of relatives with the particular disease were analysed as two (≥1, 0) or three levels (≥2, 1, 0).

IM in a first degree relative was analysed. IM was analysed as a continuous variable: number of episodes of IM in family member / number of close blood relatives (weighted by age for risk of IM) as for YHHCCS. Also in order to make the analysis results more easily understood, a separate variable was constructed as a dichotomy (no relatives with IM vs. one or more relatives.

Factors controlled for in the analysis:

Risk of HD is associated with socio-economic status (SES), age and sex. HD has a bimodal pattern of presentation and males tend to have a higher risk than females (see Chapter 1). Young adults of higher SES have a higher risk of HD possibly related to delayed

exposure to infection (see Chapter 1). Many of the risk factors analysed e.g. tonsillectomy, appendectomy, IM, are also associated with SES.

All analyses were performed adjusted for age (in 10 year groups) and sex. They were additionally adjusted for Carstair's index. Carstair's index is a measure of SES based on residence at diagnosis (cases) or pseudo-diagnosis (controls). Carstair's index is constructed using the index postcode to give a census area (ward) code of the address and from this the Carstair's index can be calculated based on four census variables: unemployment, overcrowding, non-car ownership, low social class (Jarman et al, 1991). The index has 7 levels ranging from 1 (least deprived) to 7 (most deprived). As for YHHCCS Carstair's index was not available for a number of cases and controls and all analyses that are adjusted for this variable involve the loss of a small amount of data.

Analyses were also performed adjusted for potential confounding factors that reflect childhood social class: parental education, sibship size, and birth order (as for YHHCCS).

All family health analyses were performed adjusted for age and sex, adjusted for age, sex, and Carstair's index, and for the number of first degree relatives. Cancer questions only were adjusted for the weighted sum of first degree relatives at risk based on age or age and sex to find if this resulted in a better fit of the model than adjusting for number of relatives alone.

EBV Status and Histopathology:

Paraffin embedded biopsy material was retrieved from cases and histopathological review is being performed by the staff of the LRF virus centre in Glasgow under the leadership of Professor Ruth Jarrett. However, the review had not been finished when the analyses presented here were performed (data available for 173 (35.0%) of cases). Therefore, none of the analyses have been performed on histological subgroups.

Cases are described as EBV +ve if the HRS cells scored positive for EBER ISH or LMP-1 immunohistochemistry. EBV status was available for 323 (65.5%) cases. EBV

testing had not been finished at the time of this analysis. It is anticipated that EBV status will become available for almost all cases.

HLA DPB1 typing:

Typing of the HLA DPB1 alleles by Dr Malcolm Taylor was not complete when these analyses were performed and therefore no comparison with YHHCCS results is possible.

Statistical methods:

Logistic regression was used for all analyses on SAS 6.11 using PROC LOGISTIC. The results are reported as Odds Ratios (OR) with 95% Confidence Intervals (95% CI). Analyses were applied to the total series of all subjects, the series split by age at presentation in groups (16-34 and 50+ yrs and also age 16-24 yrs for comparison with YHHCCS results), and the subset for which EBV status was known, and with EBV status known by age subgroup (16-34 and 50+ yrs). The age groups 16-34 and 50+ years were first suggested to be separate aetio-subgroups by MacMahon (1966) (see Chapter 1).

To reduce the total number of tests performed only confounder variables that resulted in a statistically significantly better fitting model were considered in the sub-group analyses.

To investigate whether associations with case status differed by age-at-diagnosis group (16-34 vs. 50+) models were produced comparing the variable of interest if associated with case status age 16-34 or age 50+ yrs in the series of the two age groups combined with a model including the variable if it occurred age 16-34 & age 50+ yrs. The deviance difference between this and the model with the single variable of interest (common effect across age groups) was then compared with the chi-square distribution on 1 degree of freedom.

Similarly to find if risk factors differed by EBV status a series consisting of all cases with available EBV status was constructed and the variable of interest was analysed with EBV status as the dependent variable. If the OR was positive it meant the risk of EBV +ve

HD was higher and vice versa (if the OR was statistically significant this signified a significant difference in effect between the EBV subgroups).

8.3 Results:

The main characteristics of the cases and controls are given in Table 8.1. The majority of cases were male and in the young adult peak (age 16-34 yrs). The table also shows how successful the methods used in calculating the number of controls (Chapter 5) were in producing a similar distribution of controls to the expected number of cases. This observation applies as well to the age groups when observed separately. The only slight difference between the cases and controls is in the Carstair's index where more cases than controls are in the 'most deprived' category and more controls than cases in the 'least deprived' category. EBV data were only available for 323 of the cases and one third of these were EBV +ve. In the age groups the proportion of EBV +ve cases was smaller (26.2%) in the 16-34 years age group than in the 50+ years age group (45.2%).

General Health:

The results for index general health in the total series are in Table 8.2. There were very few statistically significant results. Eczema is associated with a statistically significant increase in risk of HD. There is little evidence of an effect on risk of removal of the tonsils or appendix, or auto-immune diseases in the total series. Chronic infections (age ≥11 and ≥5 yrs) are associated with a statistically significant increased risk of HD.

The results for index general health by age-at-diagnosis subgroups are in Table 8.3. Again there are very few statistically significant results, especially for HD presenting after age 50 years. The results for age 50+ years are very similar to those for the total series. In the 16-34 years age group there are few statistically significant results due to the small number of positive responses. However, receiving a blood transfusion and removal of the tonsils are both associated with a statistically significant increase in risk of HD. The significant effect of

Table 8.1. Characteristics of the SNEHD cases and controls

	Cases	Controls
Characteristic	Number (%)	Number (%)
Sex: Male	281 (57.0)	290 (56.6)
Female	212 (43.0)	222 (43.3)
Age: 16-24 years	114 (23.1)	99 (19.3)
25-34 years	127 (25.8)	121 (24.5)
35-44 years	74 (15.0)	79 (15.4)
45-54 years	63 (12.8)	72 (14.1)
55-64 years	62 (12.6)	73 (14.3)
65-74 years	53 (10.8)	68 (13.3)
Carstair's index: 1 (least deprived)	20 (4.1)	33 (6.4)
2	47 (9.5)	46 (9.0)
3	80 (16.2)	100 (19.5)
4	90 (18.3)	100 (19.5)
5	101 (20.5)	96 (18.8)
6	74 (15.0)	74 (14.5)
7 (most deprived)	39 (7.9)	31 (6.1)
EBV status: +ve	101 (31.3)	NA
-ve	222 (59.7)	NA
Histopathological Review completed	173	NA

Table 8.2: Total series results for general health

	Case		Control		Adj.age & sex	Adj. age, sex, Carstairs
Risk factor	No	No (%) yes	No	No (%) yes	OR (95%CI)	OR (95%CI)
Any previous cancer	493	11 (2.2)	512	10 (2.0)	1.38 (0.57-3.35)	1.64 (0.66-4.10)
Any previous haematological malignancy	493	3 (0.6)	512	2 (0.4)	1.85 (0.31-11.21)	1.91 (0.31-11.66)
Eczema	492	71 (14.4)	511	50 (9.8)	1.49 (1.01-2.20)	1.51 (1.00-2.27)
Received blood transfusion	492	38 (7.7)	508	48 (9.4)	0.88 (0.56-1.39)	0.91 (0.56-1.45)
Appendix removed at any age	493	51 (10.3)	512	57 (11.1)	0.99 (0.65-1.48)	1.02 (0.66-1.57)
Appendix removed < age 16	493	24 (4.9)	512	28 (5.5)	0.84 (0.47-1.49)	0.84 (0.46-1.53)
Appendix removed ≥ age 16	493	26 (5.3)	512	29 (5.7)	1.09 (0.63-1.90)	1.16 (0.65-2.09)
Tonsils removed at any age	493	123 (24.9)	512	123 (24.0)	1.11 (0.83-1.49)	1.15 (0.85-1.56)
Tonsils removed before age 5.5	493	43 (8.7)	510	42 (8.2)	1.09 (0.70-1.70)	1.18 (0.74-1.89)
Tonsils removed age 5.5-10	493	51 (10.3)	510	47 (9.2)	1.18 (0.78-1.80)	1.25 (0.81-1.94)
Tonsils removed ≥ age 11	493	30 (6.1)	510	36 (7.1)	0.91 (0.55-1.51)	0.82 (0.48-1.40)
Any Auto-immune disease	493	31 (6.2)	510	34 (6.6)	0.96 (0.70-1.32)	0.97 (0.70-1.34)
Atopic illness	493	120 (24.4)	512	106 (20.7)	1.14 (0.89-1.47)	1.19 (0.92-1.55)
Lower respiratory tract infection	493	26 (5.3)	512	13 (2.5)	2.24 (1.13-4.42)	2.15 (1.08-4.27)

Table 8.2: Total series results for general health (cont)

	Case		Control		Adj.age & sex	Adj. age, sex, Carstairs
Risk factor	No	No (%) yes	No	No (%) yes	OR (95%CI)	OR (95%CI)
Chronic infection at any age	493	66 (13.4)	512	47 (9.2)	1.52 (1.02-2.27)	1.44 (0.96-2.17)
Chronic infection age 0-4	492	10 (2.0)	512	7 (1.4)	1.43 (0.54-3.78)	1.26 (0.46-3.43)
Chronic infection age 5-10	492	13 (2.6)	512	13 (2.5)	1.01 (0.46-2.21)	1.00 (0.46-2.20)
Chronic infection age ≥11	492	44 (8.9)	512	27 (5.3)	1.80 (1.09-2.96)	1.72 (1.03-2.89)
Chronic infection age ≥5	492	56 (11.4)	512	39 (7.6)	1.57 (1.02-2.41)	1.50 (0.96-2.34)

numbers and percentages for unadjusted analysis only

Table 8.3: General Health factors and age groups (16-34 & 50+)

	16-34				50+			
	Case	Control	Adj. age, sex	Adj.age, sex, Carstairs	Case	Control	Adj.age & sex	Adj. age, sex, Carstairs
Risk factor	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)
Previous cancer	0 (0.0)	1 (0.5)	0.00 (0.00-∞)	0.00 (0.00-∞)	10 (6.8)	8 (4.8)	1.57 (0.59-4.16)	1.61 (0.60-4.29)
Previous haem. malig	0 (0.0)	0 (0.0)	NA	NA	3 (2.0)	2 (1.2)	1.70 (0.28-10.42)	1.73 (0.28-10.73)
Eczema	42 (17.4)	33 (15.0)	1.19 (0.72-1.96)	1.16 (0.69-1.95)	15 (10.2)	7 (4.3)	2.58 (1.02-6.54)	2.31 (0.89-6.00)
Received blood transfusion	14 (5.8)	5 (2.3)	2.70 (0.96-7.65)	3.30 (1.05-10.36)	20 (13.5)	30 (18.2)	0.74 (0.40-1.38)	0.68 (0.36-1.30)
appendix removed any age	15 (6.2)	14 (6.4)	1.07 (0.48-2.37)	1.32 (0.55-3.16)	27 (18.2)	25 (15.2)	1.27 (0.70-2.32)	1.27 (0.68-2.37)
appendix removed age <16	8 (3.3)	8 (3.6)	0.68 (0.23-1.98)	0.66 (0.21-2.13)	11 (7.4)	8 (4.8)	1.57 (0.61-4.04)	1.65 (0.64-4.25)
appendix removed age ≥16	7 (2.9)	6 (2.7)	1.86 (0.55-6.28)	3.50 (0.73-16.76)	15 (10.1)	17 (10.3)	1.00 (0.48-2.11)	0.95 (0.43-2.08)
tonsils removed any age	47 (19.5)	25 (11.4)	1.92 (1.13-3.25)	1.88 (1.09-3.23)	46 (31.1)	53 (32.1)	0.94 (0.58-1.52)	1.05 (0.63-1.73)
tonsils removed age <5.5	17 (7.1)	10 (4.5)	1.62 (0.72-3.62)	1.50 (0.66-3.41)	12 (8.1)	13 (8.0)	0.98 (0.43-2.24)	1.17 (0.50-2.72)
tonsils removed age 5.5-10	20 (8.3)	8 (3.6)	2.50 (1.07-5.84)	2.98 (1.22-7.28)	17 (11.5)	21 (12.9)	0.82 (0.41-1.64)	0.93 (0.45-1.91)
tonsils removed age ≥11	10 (4.1)	7 (3.2)	1.29 (0.48-3.48)	1.01 (0.36-2.87)	17 (11.5)	19 (11.7)	1.04 (0.51-2.10)	1.01 (0.49-2.09)
Any Auto-immune disease	7 (2.9)	6 (2.8)	0.96 (0.51-1.80)	0.96 (0.51-1.80)	16 (10.8)	20 (12.2)	0.94 (0.60-1.45)	0.92 (0.59-1.46)
Atopic illness	75 (31.1)	65 (29.5)	1.08 (0.79-1.48)	1.10 (0.79-1.52)	25 (16.9)	21 (12.7)	1.29 (0.75-2.19)	1.31 (0.75-2.27)

Table 8.3: General Health factors and age groups (16-34 & 50+) (cont)

	16-34				50+			
	Case	Control	Adj.age & sex	Adj. age, sex, Carstairs	Case	Control	Adj.age & sex	Adj. age, sex, Carstairs
Risk factor	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)
Lower respiratory tract infection	9 (3.7)	2 (0.9)	4.20 (0.90-19.66)	3.94 (0.84-18.59)	11 (7.4)	7 (4.2)	1.75 (0.66-4.66)	1.59 (0.58-4.35)
Chronic infection at any age	30 (12.4)	20 (9.1)	1.44 (0.79-2.63)	1.32 (0.71-2.44)	17 (11.5)	15 (9.1)	1.24 (0.59-2.59)	1.08 (0.50-2.31)
Chronic infection age 0-4	6 (2.5)	4 (1.8)	1.37 (0.38-4.91)	1.12 (0.29-4.25)	2 (1.4)	1 (0.6)	2.16 (0.19-24.41)	2.20 (0.19-25.21)
Chronic infection age 5-10	7 (2.9)	7 (3.2)	0.91 (0.31-2.66)	0.87 (0.30-2.56)	4 (2.7)	2 (1.2)	2.26 (0.40-12.63)	2.30 (0.41-12.90)
Chronic infection age ≥11	16 (6.7)	10 (4.5)	1.54 (0.68-3.47)	1.43 (0.62-3.29)	12 (8.1)	11 (6.7)	1.18 (0.50-2.78)	0.98 (0.40-2.41)
Chronic infection age ≥5	23 (9.6)	16 (7.3)	1.38 (0.70-2.69)	1.29 (0.65-2.56)	15 (10.1)	13 (7.9)	1.26 (0.57-2.76)	1.08 (0.48-2.44)

numbers and percentages for unadjusted analysis only

tonsil removal remained after adjustment for total number of childhood infections (age 0-4 and ≥5 yrs) and IM.

The results for the total series by EBV status are in Table 8.4. Again very few of the results are statistically significant. The association with eczema is limited to EBV +ve HD. The only other statistically significant result is for atopic illness in EBV +ve HD. This latter result may be due to the fact that the atopic variable includes eczema. There are no statistically significant differences between EBV +ve and EBV –ve HD apart from those for eczema (OR EBV +ve vs. –ve cases 2.23, 1.11-4.49) and atopic illness (OR EBV +ve vs. –ve cases 2.06, 1.28-3.30). The results for general health variables by EBV status and age at diagnosis are in Tables 8.5 and 8.6. In the 16-34 age group EBV +ve HD was statistically significantly associated with eczema but none of the other results were significant. EBV –ve HD in this age group was associated statistically significantly with removal of the tonsils and lrti. The only statistically significant difference between EBV +ve and EBV –ve HD in the 16-34 age group was for eczema with EBV +ve HD 3 times more likely to be associated with this illness (OR 2.97, 1.23-7.13). In the 50+ age group none of the results in either EBV subgroup were statistically significant.

Infectious Illness:

The results for combined infectious illness in the total series are in Table 8.7. All the age groups of infection for childhood infectious illness have been given to allow a comparison with YHHCCS. In the total series more childhood infections at any age is associated with a lower risk of HD; if infection occurs early risk is lower but if it occurs later risk is elevated. However, none of the results are statistically significant for the dichotomous variables. Only the results for infectious illness at any age are presented for the other groups of infections. Greater frequency of clinical illness due to herpes viruses is statistically significantly protective. When split by age group this protective effect is seen for each age group but especially so if the analysis is limited to first episode of herpes viruses after

Table 8.4: General Health factors by EBV status

	EBV +ve				EBV –ve			
	Case	Control	Adj.age & sex	Adj. age, sex, Carstairs	Case	Control	Adj.age & sex	Adj. age, sex, Carstairs
Risk factor	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)
Previous cancer	3 (3.0)	10 (2.0)	1.69 (0.44-6.48)	1.66 (0.42-6.61)	3 (1.4)	10 (2.0)	0.97 (0.26-3.66)	1.12 (0.29-4.33)
Previous haem. malig	0 (0.0)	2 (0.4)	0.00 (0.00-∞)	0.00 (0.00-∞)	1 (0.5)	2 (0.4)	1.82 (0.16-20.57)	1.84 (0.16-20.95)
Eczema	19 (18.8)	50 (9.8)	2.23 (1.23-4.04)	2.37 (1.27-4.41)	27 (12.2)	50 (9.8)	1.11 (0.67-1.84)	1.10 (0.64-1.87)
Received blood transfusion	13 (12.9)	48 (9.4)	1.53 (0.77-3.04)	1.40 (0.68-2.90)	14 (6.3)	48 (9.4)	0.80 (0.42-1.51)	0.83 (0.43-1.63)
appendix removed any age	11 (10.9)	57 (11.1)	0.91 (0.44-1.86)	0.83 (0.38-1.86)	16 (7.2)	57 (11.1)	0.68 (0.37-1.25)	0.72 (0.38-1.35)
appendix removed age <16	7 (6.9)	28 (5.5)	1.07 (0.43-2.67)	1.00 (0.37-2.69)	6 (2.7)	28 (5.5)	0.43 (0.16-1.13)	0.46 (0.17-1.23)
appendix removed age ≥16	3 (3.0)	29 (5.7)	0.55 (0.16-1.86)	0.43 (0.10-1.89)	10 (4.5)	29 (5.7)	1.01 (0.47-2.15)	1.08 (0.48-2.42)
tonsils removed any age	27 (26.7)	123 (24.0)	1.18 (0.72-1.93)	1.22 (0.73-2.05)	59 (26.6)	123 (24.0)	1.31 (0.90-1.90)	1.40 (0.95-2.05)
tonsils removed age <5.5	10 (9.9)	42 (8.2)	1.19 (0.58-2.47)	1.38 (0.65-2.90)	22 (9.9)	42 (8.2)	1.35 (0.78-2.34)	1.33 (0.74-2.39)
tonsils removed age 5.5-10	11 (10.9)	47 (9.2)	1.24 (0.61-2.51)	1.32 (0.63-2.79)	26 (11.7)	47 (9.2)	1.44 (0.86-2.42)	1.66 (0.98-2.81)
tonsils removed age ≥11	7 (6.9)	36 (7.1)	1.03 (0.44-2.40)	0.75 (0.28-1.99)	11 (5.0)	36 (7.1)	0.77 (0.38-1.56)	0.79 (0.39-1.61)
Any Auto-immune disease	1 (1.0)	14 (2.7)	1.03 (0.62-1.72)	0.93 (0.53-1.91)	7 (3.2)	20 (3.9)	1.07 (0.72-1.59)	1.09 (0.73-1.63)
Atopic illness	26 (25.7)	91 (17.8)	1.62 (1.11-2.38)	1.70 (1.14-2.53)	46 (20.8)	106 (20.7)	0.89 (0.64-1.24)	0.92 (0.65-1.29)

Table 8.4: General Health factors by EBV status (cont)

	EBV +ve				EBV –ve			
	Case	Control	Adj.age & sex	Adj. age, sex, Carstairs	Case	Control	Adj.age & sex	Adj. age, sex, Carstairs
Risk factor	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)
Lower respiratory tract infection	4 (4.0)	13 (2.5)	1.58 (0.50-4.97)	1.31 (0.41-4.17)	11 (5.0)	13 (2.5)	2.28 (0.99-5.23)	2.30 (0.99-5.34)
Chronic infection at any age	13 (12.9)	47 (9.2)	1.58 (0.81-3.07)	1.24 (0.61-2.52)	29 (13.1)	47 (9.2)	1.40 (0.85-2.31)	1.34 (0.80-2.25)
Chronic infection age 0-4	1 (1.0)	7 (1.4)	0.71 (0.09-5.91)	0.68 (0.08-5.63)	6 (2.7)	7 (1.4)	1.71 (0.56-5.21)	1.45 (0.45-4.680
Chronic infection age 5-10	5 (5.0)	13 (2.5)	2.06 (0.71-5.96)	1.97 (0.68-5.76)	5 (2.3)	13 (2.5)	0.77 (0.27-2.21)	0.77 (0.27-2.22)
Chronic infection age ≥11	7 (6.9)	27 (5.3)	1.53 (0.64-3.68)	1.04 (0.38-2.82)	19 (8.6)	27 (5.3)	1.70 (0.91-3.16)	1.70 (0.90-3.21)
Chronic infection age ≥5	12 (11.9)	39 (7.6)	1.81 (0.90-3.62)	1.42 (0.67-3.01)	23 (10.4)	39 (7.6)	1.35 (0.78-2.34)	1.34 (0.76-2.35)

numbers and percentages for unadjusted analysis only

Table 8.5: General Health factors age 16-34 by EBV status

	EBV +ve				EBV –ve			
	Case	Control	Adj.age & sex	Adj. age, sex, Carstairs	Case	Control	Adj.age & sex	Adj. age, sex, Carstairs
Risk factor	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)
Previous cancer	0 (0.0)	1 (0.5)	0.00 (0.00-∞)	0.00 (0.00-∞)	0 (0.0)	1 (0.5)	0.00 (0.00-∞)	0.00 (0.00-∞)
Previous haem. malig	0 (0.0)	0 (0.0)	NA	NA	0 (0.0)	0 (0.0)	NA	NA
Eczema	13 (29.5)	33 (15.0)	2.26 (1.05-4.87)	2.53 (1.15-5.58)	18 (14.5)	33 (15.0)	0.91 (0.49-1.71)	0.89 (0.46-1.71)
Received blood transfusion	2 (4.5)	5 (2.3)	2.57 (0.45-14.67)	4.69 (0.75-29.45)	7 (5.6)	5 (2.3)	2.39 (0.73-7.75)	2.71 (0.74-10.01)
appendix removed any age	3 (6.8)	14 (6.4)	0.90 (0.19-4.25)	0.56 (0.07-4.67)	6 (4.8)	14 (6.4)	0.71 (0.24-2.06)	0.94 (0.31-2.91)
appendix removed age <16	3 (6.8)	8 (3.6)	1.36 (0.27-6.77)	0.80 (0.09-6.85)	3 (2.4)	8 (3.6)	0.43 (0.09-2.08)	0.49 (0.10-2.43)
appendix removed age ≥16	0 (0.0)	6 (2.7)	0.00 (0.00-∞)	0.00 (0.00-∞)	3 (2.4)	6 (2.7)	1.25 (0.27-5.74)	2.47 (0.40-15.14)
tonsils removed any age	8 (18.2)	25 (11.4)	1.66 (0.69-4.03)	1.48 (0.58-3.79)	28 (22.6)	25 (11.4)	2.30 (1.27-4.18)	2.35 (1.28-4.33)
tonsils removed age <5.5	3 (6.8)	10 (4.5)	1.50 (0.39-5.81)	1.41 (0.36-5.60)	9 (7.3)	10 (4.5)	1.71 (0.67-4.37)	1.52 (0.58-4.01)
tonsils removed age 5.5-10	2 (4.5)	8 (3.6)	1.41 (0.28-7.08)	1.67 (0.32-8.62)	14 (11.3)	8 (3.6)	3.46 (1.40-8.55)	4.10 (1.59-10.57)
tonsils removed age ≥11	3 (6.8)	7 (3.2)	1.89 (0.45-7.98)	1.24 (0.23-6.66)	5 (4.0)	7 (3.2)	1.21 (0.37-3.93)	1.21 (0.37-3.93)
Any Auto-immune disease	0 (0.0)	6 (2.8)	0.00 (0.00-∞)	0.00 (0.00-∞)	4 (3.2)	6 (2.8)	1.06 (0.52-2.15)	1.06 (0.52-2.15)
Atopic illness	19 (43.2)	65 (29.5)	1.42 (0.84-2.39)	1.47 (0.85-2.53)	33 (26.6)	65 (29.5)	0.92 (0.62-1.36)	0.93 (0.62-1.40)

Table 8.5: General Health factors age 16-34 by EBV status (cont)

	EBV +ve				EBV –ve			
	Case	Control	Adj.age & sex	Adj. age, sex, Carstairs	Case	Control	Adj.age & sex	Adj. age, sex, Carstairs
Risk factor	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)
Lower respiratory tract infection	2 (4.5)	2 (0.9)	3.47 (0.45-26.88)	2.64 (0.33-21.06)	6 (4.8)	2 (0.9)	5.61 (1.11-28.38)	5.44 (1.06-27.81)
Chronic infection at any age	7 (15.9)	20 (9.1)	2.20 (0.83-5.82)	1.72 (0.61-4.84)	19 (15.3)	20 (9.1)	1.75 (0.89-3.43)	1.64 (0.82-3.28)
Chronic infection age 0-4	1 (2.3)	4 (1.8)	1.09 (0.11-10.56)	1.15 (0.12-11.40)	5 (4.1)	4 (1.8)	2.26 (0.59-8.61)	1.80 (0.44-7.38)
Chronic infection age 5-10	2 (4.5)	7 (3.2)	1.50 (0.29-7.85)	1.30 (0.24-7.00)	4 (3.3)	7 (3.2)	0.94 (0.27-3.31)	0.91 (0.26-3.24)
Chronic infection age ≥11	4 (9.1)	10 (4.5)	2.93 (0.82-10.48)	2.06 (0.49-8.58)	9 (7.3)	10 (4.5)	1.60 (0.63-4.06)	1.58 (0.62-4.06)
Chronic infection age ≥5	6 (13.6)	16 (7.3)	2.53 (0.88-7.25)	1.87 (0.60-5.85)	13 (10.6)	16 (7.3)	1.43 (0.66-3.10)	1.41 (0.65-3.09)

numbers and percentages for unadjusted analysis only

Table 8.6: General Health factors age 50+ by EBV status

	EBV +ve				EBV –ve			
	Case	Control	Adj.age & sex	Adj. age, sex, Carstairs	Case	Control	Adj.age & sex	Adj. age, sex, Carstairs
Risk factor	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)
Previous cancer	3 (7.1)	8 (4.8)	1.69 (0.42-6.81)	1.36 (0.32-5.69)	2 (3.9)	8 (4.8)	0.93 (0.19-4.64)	0.95 (0.19-4.87)
Previous haem. malig	0 (0.0)	2 (1.2)	0.00 (0.00-∞)	0.00 (0.00-∞)	1 (2.0)	2 (1.2)	1.74 (0.15-20.88)	1.79 (0.15-21.93)
Eczema	5 (11.9)	7 (4.3)	3.00 (0.90-10.00)	2.23 (0.61-8.17)	3 (5.9)	7 (4.3)	1.50 (0.37-6.14)	1.54 (0.38-6.34)
Received blood transfusion	9 (21.4)	30 (18.2)	1.33 (0.56-3.12)	0.96 (0.38-2.41)	6 (11.8)	30 (18.2)	0.69 (0.27-1.79)	0.70 (0.26-1.85)
appendix removed any age	7 (16.7)	25 (15.2)	1.18 (0.46-2.99)	1.00 (0.37-2.74)	8 (15.7)	25 (15.2)	1.12 (0.46-2.71)	1.21 (0.49-2.98)
appendix removed age <16	3 (7.1)	8 (4.8)	1.56 (0.39-6.17)	1.52 (0.38-6.16)	2 (3.9)	8 (4.8)	0.79 (0.16-3.93)	0.80 (0.16-4.03)
appendix removed age ≥16	3 (7.1)	17 (10.3)	0.69 (0.19-2.52)	0.45 (0.10-2.13)	6 (11.8)	17 (10.3)	1.29 (0.47-3.57)	1.44 (0.51-4.07)
tonsils removed any age	13 (31.0)	53 (32.1)	0.95 (0.45-1.99)	1.15 (0.53-2.49)	15 (29.4)	53 (32.1)	0.88 (0.44-1.76)	0.98 (0.48-2.03)
tonsils removed age <5.5	4 (9.5)	13 (8.0)	1.18 (0.36-3.83)	1.40 (0.41-4.73)	4 (7.8)	13 (8.0)	0.95 (0.29-3.09)	1.08 (0.33-3.58)
tonsils removed age 5.5-10	6 (14.3)	21 (12.9)	1.11 (0.42-2.96)	1.43 (0.52-3.96)	7 (13.7)	21 (12.9)	0.93 (0.37-2.38)	1.06 (0.40-2.81)
tonsils removed age ≥11	3 (7.1)	19 (11.7)	0.61 (0.17-2.18)	0.62 (0.17-2.29)	4 (7.8)	19 (11.7)	0.76 (0.24-2.43)	0.78 (0.24-2.50)
Any Auto-immune disease	5 (11.9)	20 (12.2)	1.12 (0.62-2.05)	1.00 (0.51-1.95)	6 (11.8)	20 (12.2)	1.01 (0.55-1.86)	1.05 (0.57-1.93)
Atopic illness	10 (23.8)	21 (12.7)	1.72 (0.84-3.49)	1.59 (0.76-3.33)	5 (9.8)	21 (12.7)	0.74 (0.29-1.85)	0.79 (0.31-1.99)

Table 8.6: General Health factors age 50+ by EBV status (cont)

	EBV +ve				EBV –ve			
	Case	Control	Adj.age & sex	Adj. age, sex, Carstairs	Case	Control	Adj.age & sex	Adj. age, sex, Carstairs
Risk factor	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)
Lower respiratory tract infection	2 (4.8)	7 (4.2)	1.08 (0.21-5.47)	0.78 (0.15-4.08)	4 (7.8)	7 (4.2)	1.95 (0.54-7.03)	2.00 (0.55-7.32)
Chronic infection at any age	4 (9.5)	15 (9.1)	1.00 (0.31-3.24)	0.60 (0.16-2.26)	4 (7.8)	15 (9.1)	0.81 (0.25-2.59)	0.81 (0.25-2.62)
Chronic infection age 0-4	0 (0.0)	1 (0.6)	0.00 (0.00-∞)	0.00 (0.00-∞)	0 (0.0)	1 (0.6)	0.00 (0.00-∞)	0.00 (0.00-∞)
Chronic infection age 5-10	2 (4.8)	2 (1.2)	3.73 (0.50-27.79)	3.41 (0.45-25.53)	1 (2.0)	2 (1.2)	1.74 (0.15-20.88)	1.81 (0.15-22.00)
Chronic infection age ≥11	2 (4.8)	11 (6.7)	0.67 (0.14-3.20)	0.27 (0.03-2.21)	4 (7.8)	11 (6.7)	1.13 (0.34-3.78)	1.13 (0.33-3.81)
Chronic infection age ≥5	4 (9.5)	13 (7.9)	1.17 (0.35-3.87)	0.73 (0.19-2.77)	4 (7.8)	13 (7.9)	0.92 (0.28-3.02)	0.93 (0.28-3.05)

numbers and percentages for unadjusted analysis only

Table 8.7: Total series index total infections

	Case		Control		Adj.age & sex	Adj. age, sex, Carstairs
Risk Factor	No	No (%) yes	No	No (%) yes	OR (95%CI)	OR (95%CI)
Total childhood infections at any age 2,3,4 vs. 1	482	358 (74.3)	509	403 (79.2)	0.76 (0.56-1.02)	0.81 (0.59-1.11)
Total childhood infections at any age 2 vs. 1	482	135 (28.0)	509	144 (28.3)	0.75 (0.53-1.06)	0.79 (0.55-1.13)
3 vs. 1		125 (25.9)		127 (25.0)	0.76 (0.54-1.09)	0.85 (0.59-1.24)
4 vs. 1		97 (20.1)		131 (25.7)	0.62 (0.43-0.89)	0.66 (0.45-0.96)
Total childhood infections age 0-4 1,2 vs. 0	482	150 (31.1)	509	171 (33.6)	0.89 (0.68-1.17)	0.90 (0.68-1.19)
Total childhood infections age 0-4 1 vs. 0	482	93 (19.3)	509	98 (19.3)	0.97 (0.70-1.34)	0.93 (0.67-1.31)
2 vs. 0		57 (11.8)		73 (14.3)	0.79 (0.54-1.15)	0.85 (0.57-1.26)
Total childhood infections age 5-10 1,2,3 vs. 0	482	366 (75.9)	509	388 (76.2)	0.97 (0.72-1.30)	1.06 (0.78-1.45)
Total childhood infections age 5-10 1 vs. 0	482	142 (29.5)	509	125 (24.6)	1.16 (0.81-1.64)	1.33 (092-1.93)
2 vs. 0		100 (20.7)		126 (24.8)	0.80 (0.55-1.16)	0.89 (0.61-1.31)
3 vs. 0		124 (25.7)		137 (26.9)	0.95 (0.66-1.35)	0.99 (0.68-1.43)
Total childhood infections age ≥11 1,2 vs. 0	482	120 (24.9)	509	115 (22.6)	1.12 (0.84-1.51)	1.16 (0.85-1.58)
Total childhood infections age ≥11 1 vs. 0	482	110 (22.8)	509	85 (16.7)	1.35 (0.98-1.35)	1.35 (0.97-1.88)
2 vs. 0		10 (2.1)		30 (5.9)	0.37 (0.18-0.76)	0.45 (0.21-0.94)

Table 8.7: Total series index total infections (cont)

	Case		Control		Adj.age & sex	Adj. age, sex, Carstairs
Risk Factor	No	No (%) yes	No	No (%) yes	OR (95%CI)	OR (95%CI)
Total childhood infections age ≥5 2,3,4 vs 1	482	264 (54.8)	509	309 (60.7)	0.78 (0.61-1.01)	0.82 (0.63-1.07)
Total childhood infections age ≥5 2 vs. 1	482	113 (23.4)	509	129 (25.3)	0.76 (0.55-1.04)	0.82 (0.59-1.14)
3 vs. 1		95 (19.7)		100 (19.6)	0.82 (0.59-1.16)	0.86 (0.60-1.23)
4 vs. 1		56 (11.6)		80 (15.7)	0.65 (0.44-0.96)	0.67 (0.44-1.00)
Total herpes virus infections at any age 1,2,3 vs. 0	490	408 (83.3)	512	453 (88.5)	0.59 (0.41-0.86)	0.63 (0.43-0.93)
Total herpes virus infections at any age 1 vs. 0	490	262 (53.5)	512	277 (54.1)	0.59 (0.40-0.87)	0.64 (0.43-0.96)
2 vs. 0		123 (25.1)		151 (29.5)	0.53 (0.35-0.80)	0.58 (0.38-0.90)
3 vs. 0		23 (4.7)		25 (4.9)	0.64 (0.33-1.24)	0.70 (0.36-1.38)
Total additional infections at any age 1,2 vs. 0	491	86 (17.5)	512	73 (14.3)	1.46 (1.02-2.08)	1.50 (1.04-2.18)
Total additional infections at any age 1 vs. 0	491	76 (15.5)	512	64 (12.5)	1.44 (0.99-2.08)	1.54 (1.05-2.27)
2 vs. 0		10 (2.0)		9 (1.8)	1.48 (0.59-2.75)	1.23 (0.46-3.29)

Table 8.7: Total series index total infections (cont)

	Case		Control		Adj.age & sex	Adj. age, sex, Carstairs
Risk Factor	No	No (%) yes	No	No (%) yes	OR (95%CI)	OR (95%CI)
Total childhood + additional infections at any age 2,3,4 vs. 1	491	370 (75.4)	512	409 (79.9)	0.77 (0.57-1.04)	0.82 (0.60-1.13)
Total childhood + additional infections at any age 2 vs. 1	491	123 (25.1)	512	123 (24.0)	0.82 (0.57-1.18)	0.86 (0.59-1.26)
3 vs. 1		130 (26.5)		145 (28.3)	0.75 (0.52-1.16)	0.86 (0.59-1.24)
4 vs. 1		117 (23.8)		141 (27.5)	0.72 (0.50-1.03)	0.75 (0.52-1.10)
Total all combined infections at any age 2,3,4,5 vs. 1	493	410 (83.2)	512	445 (86.9)	0.75 (0.53-1.06)	0.86 (0.60-1.25)
Total all combined infections at any age 2 vs. 1	493	82 (16.6)	512	92 (18.0)	0.70 (0.45-1.09)	0.79 (0.50-1.24)
3 vs. 1		134 (27.2)		121 (23.6)	0.88 (0.59-1.33)	1.07 (0.70-1.64)
4 vs. 1		97 (19.7)		104 (20.3)	0.75 (0.49-1.15)	0.83 (0.53-1.29)
5 vs. 1		97 (19.7)		128 (25.0)	0.64 (0.42-0.98)	0.76 (0.49-1.18)

Childhood infections: chicken pox, German measles, measles, mumps, whooping cough
Additional infections: tuberculosis, diphtheria, meningitis, malaria, scarlet fever, rheumatic fever, pneumonia
Herpes virus infections: chicken pox, HSV-1, HSV-2, shingles
Total combined infections: Sum of all of the above

leaving school. 'Additional' infections are associated with a statistically significant increased risk of HD. When childhood and 'additional' infections are combined the effects of each are diminished. Analyses of combined infections by age of infection were uninformative.

The results of total infections by age at diagnosis group are in Table 8.8. In the 16-34 years age group there are no statistically significant results, except for a greater number of infectious illnesses aged ≥5 years being associated with a lower risk of HD. Herpes viruses combined are associated with a lower risk of HD, when the levels are examined those cases with the most herpes infections have a statistically significantly reduced risk of HD. When split by age group of infection a higher risk of HD is associated with more late herpes infections (2vs. 0 age ≥11 yrs OR 3.01, 1.29-7.01; 3 vs. 0 age ≥5 yrs OR 2.30, 0.42-12.58; 2 vs. 0 first episode after leaving school OR 6.89, 0.83-57.09).

In the 50+ years age at diagnosis group the total number of childhood infectious illnesses at any age is associated with a statistically significant lower risk of HD. Total herpes virus infections at all ages is associated with a statistically significant lower risk of HD. More 'additional' infectious illness were associated with a statistically significant increase in risk of HD if they occurred age ≥5 (OR 1.85. 1.06-3.25).

The results for the total series by EBV status are in Table 8.9. None of the total infectious illness variables are statistically significantly associated with risk of EBV +ve or EBV –ve HD. The EBV subgroups were not significantly different from each other.

The results for the EBV subgroups for age at diagnosis groups 16-34 yrs are in Table 8.10. Again there are no statistically significant results for total childhood infectious illness. However, when the number of childhood infections aged ≥5 are examined the difference between the EBV +ve and –ve subgroups is statistically significant with a larger number of infections in this age group associated with twice the risk of EBV +ve HD (p=0.034). When childhood and 'additional' infectious illnesses are combined the effect is again very similar in both EBV subgroups for all ages of infection. However, again there is a statistically

Table 8.8: Index total infections by age group (16-34 & 50+)

	16-34				50+			
	Case	Control	Adj.age & sex	Adj. age, sex, Carstairs	Case	Control	Adj.age & sex	Adj. age, sex, Carstairs
Risk factor	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)
Tot child infect any age 2,3,4 v. 1	180 (75.3)	168 (76.4)	0.96 (0.62-1.47)	1.07 (0.69-1.68)	94 (67.1)	130 (79.8)	0.52 (0.31-0.88)	0.52 (0.30-0.90)
Tot child infect any age 2 vs. 1	70 (29.3)	70 (31.8)	0.86 (0.52-1.42)	1.00 (0.60-1.69)	36 (25.7)	38 (23.3)	0.61 (0.32-1.15)	0.61 (0.32-1.18)
3 vs. 1	71 (29.7)	60 (27.3)	1.02 (0.62-1.69)	1.19 (0.70-2.03)	27 (19.3)	38 (23.3)	0.44 (0.23-0.84)	0.46 (0.23-0.93)
4 vs. 1	39 (16.3)	38 (17.3)	0.88 (0.49-1.58)	0.93 (0.51-1.70)	31 (22.1)	54 (33.1)	0.36 (0.19-0.67)	0.36 (0.19-0.68)
Tot child infect age 0-4 1,2 vs. 0	87 (36.4)	69 (31.4)	1.26 (0.85-1.86)	1.25 (0.84-1.88)	34 (24.3)	61 (37.4)	0.55 (0.33-0.91)	0.60 (0.36-1.00)
Tot child infect age 0-4 1 vs. 0	51 (21.3)	40 (18.2)	1.27 (0.79-2.04)	1.27 (0.78-2.08)	21 (15.0)	38 (23.3)	0.55 (0.30-1.00)	0.60 (0.32-1.11)
2 vs. 0	36 (15.1)	29 (13.2)	1.24 (0.72-2.12)	1.23 (0.70-2.15)	13 (9.3)	23 (14.1)	0.56 (0.27-1.16)	0.60 (0.29-1.27)
Tot child infect ag 5-10 1,2,3 v. 0	182 (76.2)	170 (77.3)	0.93 (0.60-1.44)	1.08 (0.68-1.69)	102 (72.9)	121 (74.2)	0.88 (0.53-1.48)	0.91 (0.53-1.57)
Tot child infect age 5-10 1 vs. 0	80 (33.5)	52 (23.6)	1.34 (0.80-2.25)	1.47 (0.86-2.53)	41 (29.3)	38 (23.3)	1.12 (0.60-2.10)	1.29 (0.67-2.51)
2 vs. 0	50 (20.9)	68 (30.9)	0.64 (0.38-1.09)	0.80 (0.46-1.38)	28 (20.0)	31 (19.0)	0.94 (0.47-1.85)	0.96 (0.48-1.94)
3 vs. 0	52 (21.8)	50 (22.7)	0.91 (0.53-1.57)	1.02 (0.57-1.80)	33 (23.6)	52 (31.9)	0.67 (0.36-1.26)	0.64 (0.33-1.23)
Tot child infect age ≥11 1,2 vs. 0	69 (28.9)	51 (23.2)	1.35 (0.88-2.05)	1.25 (0.81-1.93)	34 (24.3)	30 (18.4)	1.49 (0.85-2.62)	1.53 (0.85-2.77)
Tot child infect age ≥11 1 vs. 0	67 (28.0)	40 (18.2)	1.35 (0.88-2.05)	1.25 (0.81-1.93)	29 (20.7)	18 (11.0)	2.01 (1.05-3.87)	2.02 (1.02-3.99)
2 vs. 0	2 (0.8)	11 (5.0)	NA	NA	5 (3.6)	12 (7.4)	0.48 (0.16-1.41)	0.57 (0.19-1.71)

Table 8.8: Index total infections by age group (16-34 & 50+) (cont)

	16-34				50+			
	Case	Control	Adj.age & sex	Adj. age, sex, Carstairs	Case	Control	Adj.age & sex	Adj. age, sex, Carstairs
Risk factor	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)
Tot child infect age ≥5 2,3,4 vs 1	126 (52.7)	136 (61.8)	0.69 (0.48-1.00)	0.77 (0.52-1.14)	72 (51.4)	96 (58.9)	0.73 (0.46-1.16)	0.68 (0.42-1.09)
Tot child infect age ≥5 2 vs. 1	60 (25.1)	70 (31.8)	0.63 (0.40-0.98)	0.74 (0.47-1.18)	30 (21.4)	30 (18.4)	0.89 (0.49-1.64)	0.83 (0.44-1.56)
3 vs. 1	55 (23.0)	46 (20.9)	0.88 (0.54-1.43)	0.94 (0.56-1.55)	16 (11.4)	30 (18.4)	0.47 (0.23-0.94)	0.43 (0.21-0.89)
4 vs. 1	11 (4.6)	20 (9.1)	0.40 (0.18-0.89)	0.45 (0.20-1.00)	26 (18.6)	36 (22.1)	0.63 (0.34-1.16)	0.59 (0.32-1.11)
Tot herp infect any age 1,2,3 v. 0	217 (90.4)	201 (91.4)	0.90 (0.48-1.71)	1.07 (0.54-2.10)	101 (69.2)	138 (83.6)	0.40 (0.23-0.70)	0.39 (0.21-0.70)
Tot herp infect at any age 1 vs. 0	147 (61.3)	135 (61.4)	0.67 (0.43-1.05)	0.80 (0.51-1.27)	58 (39.7)	79 (47.9)	0.38 (0.21-0.70)	0.37 (0.20-0.69)
2 vs. 0	59 (24.6)	63 (28.6)	0.89 (0.55-1.44)	0.94 (0.57-1.56)	33 (22.6)	41 (24.8)	0.42 (0.21-0.84)	0.43 (0.21-0.88)
3 vs. 0	11 (4.6)	3 (1.4)	0.41 (0.19-0.99)	0.46 (0.21-1.03)	10 (6.8)	18 (10.9)	0.28 (0.11-0.71)	0.28 (0.11-0.72)
Tot add infect at any age 1,2 vs. 0	20 (8.3)	12 (5.5)	1.56 (0.74-3.28)	1.48 (0.70-3.14)	46 (31.3)	43 (26.1)	1.49 (0.89-2.49)	1.58 (0.92-2.69)
Tot add infect at any age 1 vs. 0	19 (7.9)	11 (5.0)	1.62 (0.75-3.48)	1.54 (0.71-3.35)	38 (25.9)	35 (21.2)	1.48 (0.86-2.57)	1.66 (0.94-2.95)
2 vs. 0	1 (0.4)	1 (0.5)	0.90 (0.06-14.63)	0.84 (0.05-13.65)	8 (5.4)	8 (4.8)	1.36 (0.49-3.81)	1.24 (0.43-3.60)

Table 8.8: Index total infections by age group (16-34 & 50+) (cont)

	16-34				50+			
	Case	Control	Adj.age & sex	Adj. age, sex, Carstairs	Case	Control	Adj.age & sex	Adj. age, sex, Carstairs
Risk factor	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)
Tchild+add infect any age 2-5 vs. 1	181 (75.4)	169 (76.8)	0.93 (0.61-1.44)	1.04 (0.67-1.63)	103 (70.1)	133 (80.6)	0.56 (0.33-0.94)	0.56 (0.32-0.97)
Tchild+add infect any age 2 vs. 1	66 (27.5)	66 (30.0)	0.86 (0.52-1.42)	1.00 (0.59-1.69)	34 (23.1)	25 (15.2)	0.95 (0.48-1.89)	0.93 (0.45-1.92)
3 vs. 1	70 (29.2)	63 (28.6)	0.95 (0.57-1.58)	1.13 (0.66-1.92)	31 (21.1)	46 (27.9)	0.47 (0.25-0.89)	0.54 (0.28-1.06)
4 vs. 1	45 (18.8)	40 (18.2)	0.96 (0.55-1.70)	0.99 (0.55-1.79)	38 (25.9)	62 (37.6)	0.43 (0.23-0.79)	0.42 (0.22-0.79)
Tot comb infect any age 2-5 vs. 1	197 (81.7)	187 (85.0)	0.80 (0.49-1.31)	0.93 (0.55-1.56)	120 (81.1)	140 (84.8)	0.76 (0.42-1.48)	0.89 (0.47-1.67)
Tot comb infect any age 2 vs. 1	46 (19.1)	53 (24.1)	0.65 (0.36-1.19)	0.74 (0.40-1.38)	19 (12.8)	16 (9.7)	1.06 (0.45-2.49)	1.27 (0.52-3.12)
3 vs. 1	73 (30.3)	61 (27.7)	0.91 (0.52-1.60)	1.17 (0.64-2.12)	38 (25.7)	27 (16.4)	1.21 (0.59-2.53)	1.50 (0.69-3.27)
4 vs. 1	45 (18.7)	47 (21.4)	0.73 (0.40-1.35)	0.81 (0.43-1.53)	22 (14.9)	29 (17.6)	0.66 (0.30-1.45)	0.77 (0.34-1.73)
5 vs. 1	33 (13.7)	26 (11.8)	0.96 (0.48-1.91)	1.06 (0.52-2.16)	41 (27.7)	68 (41.2)	0.53 (0.27-1.04)	0.60 (0.29-1.21)

Childhood infections: chicken pox, German measles, measles, mumps, whooping cough
Additional infections: tuberculosis, diphtheria, meningitis, malaria, scarlet fever, rheumatic fever, pneumonia
Herpes virus infections: chicken pox, HSV-1, HSV-2, shingles
Total combined infections: Sum of all of the above

Table 8.9: Index total infections by EBV status

	EBV +ve				EBV –ve			
	Case	Control	Adj.age & sex	Adj. age, sex, Carstairs	Case	Control	Adj.age & sex	Adj. age, sex, Carstairs
Risk factor	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)
Tot child infect any age 2,3,4 v. 1	70 (71.4)	403 (79.2)	0.68 (0.42-1.12)	0.77 (0.46-1.29)	170 (77.6)	403 (79.2)	0.91 (0.62-1.34)	0.93 (0.62-1.38)
Tot child infect any age 2 vs. 1	24 (24.5)	144 (28.3)	0.62 (0.34-1.11)	0.63 (0.34-1.17)	64 (29.2)	144 (28.3)	0.88 (0.56-1.39)	0.94 (0.59-1.50)
3 vs. 1	22 (22.4)	129 (25.1)	0.62 (0.34-1.13)	0.75 (0.39-1.41)	63 (28.8)	129 (25.1)	0.98 (0.63-1.55)	1.07 (0.67-1.73)
4 vs. 1	24 (24.5)	131 (25.7)	0.68 (0.38-1.24)	0.76 (0.41-1.42)	43 (19.6)	131 (25.7)	0.72 (0.45-1.17)	0.75 (0.45-1.24)
Tot child infect age 0-4 1,2 vs. 0	31 (31.6)	171 (33.6)	0.95 (0.60-1.52)	1.04 (0.64-1.68)	74 (33.8)	171 (33.6)	0.99 (0.71-1.39)	0.99 (0.69-1.41)
Tot child infect age 0-4 1 vs. 0	24 (24.5)	98 (19.3)	1.28 (0.76-2.14)	1.34 (0.78-2.30)	41 (18.7)	98 (19.3)	0.97 (0.64-1.48)	0.93 (0.60-1.44)
2 vs. 0	7 (7.1)	73 (14.3)	0.51 (0.22-1.16)	0.59 (0.26-1.36)	33 (15.1)	73 (14.3)	1.02 (0.64-1.61)	1.08 (0.66-1.72)
Tot child infect ag 5-10 1,2,3 v. 0	72 (73.5)	388 (76.2)	0.87 (0.53-1.43)	1.12 (0.66-1.91)	166 (75.8)	388 (76.2)	0.95 (0.65-1.38)	0.99 (0.67-1.46)
Tot child infect age 5-10 1 vs. 0	21 (21.4)	125 (24.6)	0.79 (0.42-1.48)	1.05 (0.54-2.04)	67 (30.6)	125 (24.6)	1.17 (0.75-1.82)	1.30 (0.82-2.07)
2 vs. 0	22 (22.4)	126 (24.8)	0.82 (0.43-1.53)	1.09 (0.56-2.11)	41 (18.7)	126 (24.8)	0.69 (0.43-1.12)	0.74 (0.45-1.22)
3 vs. 0	29 (29.6)	137 (26.9)	0.99 (0.55-1.79)	1.20 (0.64-2.24)	58 (26.5)	137 (26.9)	0.99 (0.63-1.55)	0.95 (0.59-1.52)
Tot child infect age ≥11 1,2 vs. 0	26 (26.5)	115 (22.6)	1.24 (0.75-2.04)	1.16 (0.68-1.95)	54 (24.7)	115 (22.6)	1.10 (0.76-1.60)	1.13 (0.77-1.67)
Tot child infect age ≥11 1 vs. 0	23 (23.5)	85 (16.7)	1.45 (0.85-2.45)	1.26 (0.72-2.21)	48 (21.9)	85 (16.7)	1.27 (0.85-1.90)	1.29 (0.85-1.95)
2 vs. 0	3 (3.1)	30 (5.9)	0.53 (0.16-1.79)	0.67 (0.20-2.29)	6 (2.7)	30 (5.9)	0.51 (0.21-1.26)	0.62 (0.25-1.54)

Table 8.9: Index total infections by EBV status (cont)

	EBV +ve				EBV –ve			
	Case	Control	Adj.age & sex	Adj. age, sex, Carstairs	Case	Control	Adj.age & sex	Adj. age, sex, Carstairs
Risk factor	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)
Tot child infect age ≥5 2,3,4 vs 1	55 (56.1)	309 (60.7)	0.82 (0.53-1.28)	0.94 (0.59-1.49)	121 (55.3)	309 (60.7)	0.80 (0.58-1.11)	0.81 (0.58-1.14)
Tot child infect age ≥5 2 vs. 1	17 (17.3)	129 (25.3)	0.58 (0.32-1.07)	0.69 (0.37-1.29)	51 (23.3)	129 (25.3)	0.75 (0.50-1.13)	0.82 (0.54-1.22)
3 vs. 1	24 (24.5)	100 (19.6)	1.04 (0.60-1.81)	1.21 (0.68-2.16)	44 (20.1)	100 (19.6)	0.87 (0.57-1.34)	0.85 (0.54-1.35)
4 vs. 1	14 (14.3)	80 (15.7)	0.79 (0.41-1.52)	0.81 (0.41-1.61)	26 (11.9)	80 (15.7)	0.73 (0.44-1.22)	0.74 (0.43-1.26)
Tot herp infect any age 1,2,3 v. 0	81 (80.2)	453 (88.5)	0.53 (0.30-0.94)	0.62 (0.34-1.14)	187 (85.4)	453 (88.5)	0.65 (0.40-1.04)	0.67 (0.41-1.10)
Tot herp infect at any age 1 vs. 0	50 (49.5)	277 (54.1)	0.53 (0.29-0.97)	0.62 (0.33-1.18)	123 (56.2)	277 (54.1)	0.62 (0.38-1.01)	0.67 (0.40-1.12)
2 vs. 0	27 (26.7)	151 (29.5)	0.54 (0.28-1.04)	0.61 (0.30-1.23)	56 (25.6)	151 (29.5)	0.54 (0.32-0.92)	0.60 (0.34-1.05)
3 vs. 0	4 (4.0)	25 (4.9)	0.52 (0.16-1.68)	0.59 (0.18-1.96)	8 (3.7)	25 (4.9)	0.56 (0.23-1.41)	0.62 (0.24-1.56)
Tot add infect at any age 1,2 vs. 0	20 (19.8)	73 (14.3)	1.54 (0.87-2.75)	1.54 (0.85-2.81)	38 (17.3)	73 (14.3)	1.65 (1.05-2.60)	1.64 (1.02-2.63)
Tot add infect at any age 1 vs. 0	16 (15.8)	4 (4.0)	1.41 (0.76-2.63)	1.41 (0.74-2.69)	34 (15.5)	4 (4.0)	1.61 (1.00-2.59)	1.63 (1.00-2.67)
2 vs. 0	64 (12.5)	9 (1.8)	2.61 (0.75-9.05)	2.55 (0.72-8.99)	4 (1.8)	9 (1.8)	1.67 (0.49-5.68)	1.71 (0.50-5.85)

Table 8.9: Index total infections by EBV status (cont)

	EBV +ve				EBV –ve			
	Case	Control	Adj.age & sex	Adj. age, sex, Carstairs	Case	Control	Adj.age & sex	Adj. age, sex, Carstairs
Risk factor	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)
Tchild+add infect any age 2-5 vs. 1	73 (72.3)	409 (79.9)	0.68 (0.42-1.11)	0.75 (0.45-1.25)	177 (80.5)	409 (79.9)	1.05 (0.70-1.57)	1.08 (0.71-1.64)
Tchild+add infect any age 2 vs. 1	23 (22.8)	123 (24.0)	0.71 (0.39-1.32)	0.76 (0.37-1.32)	63 (28.6)	123 (24.0)	1.12 (0.70-1.79)	1.20 (0.73-1.95)
3 vs. 1	21 (20.8)	145 (28.3)	0.55 (0.29-1.02)	0.65 (0.34-1.25)	62 (28.2)	145 (28.3)	0.96 (0.60-1.53)	1.10 (0.67-1.75)
4 vs. 1	29 (28.7)	141 (27.5)	0.80 (0.45-1.43)	0.90 (0.49-1.65)	52 (23.6)	141 (27.5)	0.92 (0.57-1.49)	0.94 (0.56-1.56)
Tot comb infect any age 2-5 vs. 1	84 (83.2)	445 (86.9)	0.77 (0.43-1.37)	0.88 (0.48-1.64)	188 (84.7)	445 (86.9)	0.84 (0.53-1.32)	0.92 (0.57-1.47)
Tot comb infect any age 2 vs. 1	18 (17.8)	92 (18.0)	0.78 (0.37-1.63)	0.78 (0.35-1.71)	36 (16.2)	92 (18.0)	0.73 (0.41-1.28)	0.74 (0.41-1.36)
3 vs. 1	25 (24.8)	121 (23.6)	0.84 (0.42-1.67)	1.08 (0.52-2.21)	65 (29.3)	121 (23.6)	1.00 (0.60-1.68)	1.18 (0.67-2.03)
4 vs. 1	18 (17.8)	104 (20.3)	0.70 (0.34-1.46)	0.77 (0.36-1.67)	44 (19.8)	104 (20.3)	0.85 (0.49-1.47)	0.86 (0.48-1.53)
5 vs. 1	23 (22.8)	128 (25.0)	0.74 (0.37-1.50)	0.89 (0.42-1.86)	43 (19.4)	128 (25.0)	0.75 (0.43-1.31)	0.84 (0.47-1.49)

Childhood infections: chicken pox, German measles, measles, mumps, whooping cough
Additional infections: tuberculosis, diphtheria, meningitis, malaria, scarlet fever, rheumatic fever, pneumonia
Herpes virus infections: chicken pox, HSV-1, HSV-2, shingles
Total combined infections: Sum of all of the above

Table 8.10: Index total infections age 16-34 by EBV status

	EBV +ve				EBV –ve			
	Case	Control	Adj.age & sex	Adj. age, sex, Carstairs	Case	Control	Adj.age & sex	Adj. age, sex, Carstairs
Risk factor	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)
Tot child infect any age 2,3,4 v. 1	35 (79.5)	168 (76.4)	1.20 (0.54-2.70)	1.46 (0.62-3.44)	98 (79.7)	168 (76.4)	1.19 (0.69-2.05)	1.32 (0.75-2.34)
Tot child infect any age 2 vs. 1	11 (25.0)	70 (31.8)	0.80 (0.34-2.35)	1.01 (0.36-2.82)	40 (32.5)	70 (31.8)	1.11 (0.60-2.06)	1.34 (0.70-2.55)
3 vs. 1	15 (34.1)	60 (27.3)	1.45 (0.58-3.64)	1.92 (0.72-5.13)	39 (31.7)	60 (27.3)	1.27 (0.68-2.38)	1.50 (0.77-2.90)
4 vs. 1	9 (20.5)	38 (17.3)	1.39 (0.50-3.90)	1.67 (0.57-4.90)	19 (15.4)	38 (17.3)	0.99 (0.47-2.05)	1.03 (0.48-2.23)
Tot child infect age 0-4 1,2 vs. 0	15 (34.1)	69 (31.4)	1.17 (0.58-2.35)	1.17 (0.56-2.44)	51 (41.5)	69 (31.4)	1.31 (0.98-1.75)	1.30 (0.96-1.76)
Tot child infect age 0-4 1 vs. 0	12 (27.3)	40 (18.2)	1.61 (0.74-3.51)	1.58 (0.70-3.59)	27 (22.0)	40 (18.2)	1.38 (0.78-2.43)	1.35 (0.75-2.43)
2 vs. 0	3 (6.8)	29 (13.2)	0.55 (0.16-1.96)	0.60 (0.16-2.16)	24 (19.5)	29 (13.2)	1.69 (0.91-3.09)	1.65 (0.87-3.13)
Tot child infect ag 5-10 1,2,3 v. 0	38 (86.4)	170 (77.3)	1.78 (0.70-4.51)	3.05 (1.01-9.22)	90 (73.2)	170 (77.3)	0.78 (0.47-1.30)	0.86 (0.50-1.47)
Tot child infect age 5-10 1 vs. 0	11 (25.0)	52 (23.6)	1.62 (0.55-4.80)	2.62 (0.76-9.07)	42 (34.1)	52 (23.6)	1.20 (0.65-2.19)	1.28 (0.68-2.40)
2 vs. 0	12 (27.3)	68 (30.9)	1.40 (0.48-4.03)	2.54 (0.74-8.71)	21 (17.1)	68 (30.9)	0.45 (0.23-0.87)	0.54 (0.27-1.08)
3 vs. 0	15 (34.1)	50 (22.7)	2.47 (0.87-6.97)	4.11 (1.23-13.76)	27 (22.0)	50 (22.7)	0.79 (0.42-1.51)	0.84 (0.42-1.64)
Tot child infect age ≥11 1,2 vs. 0	14 (31.8)	51 (23.2)	1.59 (0.77-3.28)	1.28 (0.59-2.76)	33 (26.8)	51 (23.2)	1.27 (0.76-2.12)	1.18 (0.70-2.00)
Tot child infect age ≥11 1 vs. 0	13 (29.5)	40 (18.2)	1.79 (0.84-3.79)	1.38 (0.62-3.09)	33 (26.8)	40 (18.2)	1.61 (0.94-2.74)	1.52 (0.88-2.63)
2 vs. 0	1 (2.3)	11 (5.0)	0.66 (0.08-5.49)	0.72 (0.09-6.03)	0 (0.0)	11 (5.0)	0.00 (0.00-∞)	0.00 (0.00-∞)

Table 8.10: Index total infections age 16-34 by EBV status (cont)

	EBV +ve				EBV –ve			
	Case	Control	Adj.age & sex	Adj. age, sex, Carstairs	Case	Control	Adj.age & sex	Adj. age, sex, Carstairs
Risk factor	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)
Tot child infect age ≥5 2,3,4 vs 1	30 (68.2)	136 (61.8)	1.28 (0.64-2.59)	1.66 (0.78-3.53)	61 (49.6)	136 (61.8)	0.69 (0.39-0.95)	0.66 (0.42-1.05)
Tot child infect age ≥5 2 vs. 1	10 (22.7)	70 (31.8)	0.77 (0.32-1.88)	1.03 (0.40-2.63)	30 (24.4)	70 (31.8)	0.57 (0.34-0.99)	0.68 (0.39-1.19)
3 vs. 1	19 (43.2)	46 (20.9)	2.61 (1.17-5.82)	3.41 (1.43-8.12)	24 (19.5)	46 (20.9)	0.68 (0.37-1.24)	0.69 (0.37-1.30)
4 vs. 1	1 (2.3)	20 (9.1)	0.31 (0.04-2.57)	0.37 (0.04-3.130	7 (5.7)	20 (9.1)	0.48 (0.19-1.21)	0.53 (0.21-1.36)
Tot herp infect any age 1,2,3 v. 0	38 (86.4)	201 (91.4)	0.55 (0.20-1.50)	0.73 (0.25-2.19)	113 (91.9)	201 (91.4)	1.07 (0.48-2.40)	1.18 (0.50-2.75)
Tot herp infect at any age 1 vs. 0	25 (56.8)	135 (61.4)	0.51 (0.18-1.45)	0.71 (0.23-2.21)	79 (64.2)	135 (61.4)	1.00 (0.45-2.22)	1.22 (0.52-2.89)
2 vs. 0	11 (25.0)	63 (28.6)	0.52 (0.16-1.64)	0.62 (0.18-2.15)	30 (24.4)	63 (28.6)	0.81 (0.34-1.94)	0.99 (0.39-2.52)
3 vs. 0	2 (4.5)	3 (1.4)	2.46 (0.32-18.94)	2.83 (0.36-22.42)	4 (3.3)	3 (1.4)	1.95 (0.36-10.60)	2.40 (0.43-13.40)
Tot add infect at any age 1,2 vs. 0	4 (9.1)	12 (5.5)	1.64 (0.48-5.56)	1.55 (0.45-5.38)	11 (8.9)	12 (5.5)	1.70 (0.72-3.99)	1.54 (0.64-3.70)
Tot add infect at any age 1 vs. 0	4 (9.1)	11 (5.0)	1.83 (0.55-6.48)	1.82 (0.52-6.42)	10 (8.1)	11 (5.0)	1.64 (0.67-4.00)	1.50 (0.60-3.74)
2 vs. 0	0 (0.0)	1 (0.5)	0.00 (0.00-∞)	0.00 (0.00-∞)	1 (0.8)	1 (0.5)	2.19 (0.13-35.74)	2.04 (0.13-33.45)

Table 8.10: Index total infections age 16-34 by EBV status (cont)

	EBV +ve				EBV -ve			
	Case	Control	Adj.age & sex	Adj. age, sex, Carstairs	Case	Control	Adj.age & sex	Adj. age, sex, Carstairs
Risk factor	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)
Tchild+add infect any age 2-5 vs. 1	35 (79.5)	169 (76.8)	1.15 (0.51-2.58)	1.37 (0.58-3.22)	99 (80.5)	169 (76.8)	1.23 (0.71-2.13)	1.36 (0.77-2.42)
Tchild+add infect any age 2 vs. 1	10 (22.7)	66 (30.0)	0.81 (0.30-2.18)	0.89 (0.31-2.52)	38 (30.9)	66 (30.0)	1.16 (0.62-2.16)	1.40 (0.72-2.69)
3 vs. 1	13 (29.5)	63 (28.6)	1.18 90.46-3.02)	1.52 (0.56-4.13)	40 (32.5)	63 (28.6)	1.26 (0.67-2.36)	1.53 (0.79-2.95)
4 vs. 1	12 (27.3)	40 (18.2)	1.69 (0.64-4.48)	2.00 (0.72-5.55)	21 (17.1)	40 (18.2)	1.06 (0.5202.18)	1.07 (0.50-2.28)
Tot comb infect any age 2-5 vs. 1	38 (86.4)	187 (85.0)	1.07 (0.41-2.77)	1.39 (0.49-3.90)	102 (82.3)	187 (85.0)	0.80 (0.44-1.45)	0.89 (0.48-1.65)
Tot comb infect any age 2 vs. 1	10 (22.7)	53 (24.1)	0.93 (0.30-2.85)	1.10 (0.33-3.68)	21 (16.9)	53 (24.1)	0.59 (0.28-1.23)	0.63 (0.29-1.36)
3 vs. 1	8 (18.2)	61 (27.7)	0.69 (0.22-2.18)	1.07 (0.31-3.66)	45 (36.3)	61 (27.7)	1.08 (0.56-2.11)	1.31 (0.65-2.62)
4 vs. 1	13 (29.5)	47 (21.4)	1.58 (0.53-4.67)	1.93 (0.60-6.22)	17 (13.7)	47 (21.4)	0.50 (0.23-1.10)	0.54 (0.24-1.22)
5 vs. 1	7 (15.9)	26 (11.8)	1.37 (0.40-4.66)	1.71 (0.47-6.27)	19 (15.3)	26 (11.8)	1.09 (0.49-2.46)	1.12 (0.48-2.60)

Childhood infections: chicken pox, German measles, measles, mumps, whooping cough
Additional infections: tuberculosis, diphtheria, meningitis, malaria, scarlet fever, rheumatic fever, pneumonia
Herpes virus infections: chicken pox, HSV-1, HSV-2, shingles
Total combined infections: Sum of all of the above

significant difference between the EBV subgroups by number of infections aged ≥5 yrs. These are associated with a hugely increased risk of EBV +ve HD and essentially a null finding for EBV –ve HD. The difference between the EBV subgroups is statistically significant (p=0.043). When all infectious illnesses are combined the results are very similar for both EBV subgroups, but again there is a different effect of total infections age ≥5 yrs. More total infections at this age are associated with a lower risk of EBV –ve HD (OR 2-5 vs. 1 0.59, 0.36-0.96) and an increased risk of EBV +ve HD (OR 1.97, 0.81-4.83). The differences between EBV +ve and –ve HD are statistically significant with the greater number of infections age ≥5 yrs associated with 3 times the risk of EBV +ve HD (p=0.012).

The results for EBV status by age at diagnosis 50+ years are in Table 8.11. More total childhood infections at any age are associated with a statistically significant decreased risk of both EBV +ve and EBV –ve HD. However, none of the other results were statistically significant and none of the comparisons between EBV +ve and EBV –ve HD are statistically significant.

The results for the analysis of individual infections are shown in Table 8.12 for the total series and in Table 8.13 for the age at diagnosis groups. Only results for individual infections at any age have been presented as there were very few significant results. Full results for measles are given to compare with YHHCCS. In the total series scarlet fever statistically significantly increased risk of HD. The rest of the ORs were near to the null.

When the age groups (16-34 and 50+ yrs) were analysed separately there were again very few significant results (Table 8.13). Chicken pox was associated with a statistically significant increased risk of HD age 16-34 yrs, especially if chicken pox occurred age ≥11 (OR 1.65, 1.01-2.68). In this age group measles age ≥5 was associated with a statistically significantly decreased risk of HD. Most of the other ORs are near to the null. In the 50+ age at diagnosis group chicken pox is statistically significantly protective. In the total series and the 16-34 and 50+ age at diagnosis groups when all childhood infectious illnesses (age 0-4,

Table 8.11: Index total infections age 50+ by EBV status

	EBV +ve				EBV –ve			
	Case	Control	Adj.age & sex	Adj. age, sex, Carstairs	Case	Control	Adj.age & sex	Adj. age, sex, Carstairs
Risk factor	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)
Tot child infect any age 2,3,4 v. 1	24 (61.5)	130 (79.8)	0.38 (0.18-0.83)	0.43 (0.19-0.96)	31(63.3)	130 (79.8)	0.45 (0.22-0.93)	0.41 (0.20-0.85)
Tot child infect any age 2 vs. 1	9 (23.1)	38 (23.3)	0.44 (0.17-1.14)	0.46 (0.17-1.21)	12 (24.5)	38 (23.3)	0.59 (0.25-1.40)	0.57 (0.23-1.40)
3 vs. 1	4 (10.3)	38 (23.3)	0.18 (0.06-0.60)	0.22 (0.07-0.76)	11 (22.4)	38 (23.3)	0.47 (0.19-1.13)	0.44 (0.17-1.12)
4 vs. 1	11 (28.2)	54 (33.1)	0.37 (0.15-0.90)	0.39 (0.15-0.98)	8 (16.3)	54 (33.1)	0.25 (0.10-0.65)	0.26 (0.10-0.67)
Tot child infect age 0-4 1,2 vs. 0	13 (33.0)	61 (37.4)	0.86 (0.41-1.81)	1.00 (0.46-2.14)	9 (18.4)	61 (37.4)	0.39 (0.18-0.87)	0.42 (0.19-0.95)
Tot child infect age 0-4 1 vs. 0	10 (25.6)	38 (23.3)	1.05 (0.47-2.41)	1.24 (0.53-2.89)	5 (10.2)	38 (23.3)	0.35 (0.13-0.96)	0.39 (0.14-1.08)
2 vs. 0	3 (7.7)	23 (14.1)	0.53 (0.15-1.91)	0.60 (0.16-2.21)	4 (8.2)	23 (14.1)	0.47 (0.14-1.42)	0.48 (0.15-1.50)
Tot child infect ag 5-10 1,2,3 v. 0	25 (64.1)	121 (74.2)	0.58 (0.27-1.24)	0.67 (0.30-1.48)	36 (73.5)	121 (74.2)	0.93 (0.44-1.94)	0.89 (0.42-1.89)
Tot child infect age 5-10 1 vs. 0	9 (23.1)	38 (23.3)	0.67 (0.26-1.74)	0.84 (0.31-2.27)	16 (32.7)	38 (23.3)	1.30 (0.55-3.10)	1.42 (0.58-3.47)
2 vs. 0	7 (17.9)	31 (19.0)	0.63 (0.22-1.77)	0.68 (0.24-1.96)	9 (18.4)	31 (19.0)	0.88 (0.33-2.34)	0.79 (0.28-2.18)
3 vs. 0	9 (23.1)	52 (31.9)	0.49 (0.19-1.27)	0.54 (0.19-1.47)	11 (22.4)	52 (31.9)	0.68 (0.27-1.69)	0.60 (0.23-1.54)
Tot child infect age ≥11 1,2 vs. 0	10 (25.6)	30 (18.4)	1.54 (0.67-3.52)	1.54 (0.64-3.97)	11 (22.4)	30 (18.4)	1.32 (0.59-2.92)	1.44 (0.64-3.23)
Tot child infect age ≥11 1 vs. 0	8 (20.5)	18 (11.0)	1.92 (0.76-4.90)	1.71 (0.63-4.64)	8 (16.3)	18 (11.0)	1.58 (0.63-3.99)	1.76 (0.69-4.51)
2 vs. 0	2 (5.1)	12 (7.4)	0.69 (0.15-3.24)	0.88 (0.18-4.26)	3 (6.1)	12 (7.4)	0.78 (0.21-2.95)	0.89 (0.23-3.47)

Table 8.11: Index total infections age 50+ by EBV status (cont)

	EBV +ve				EBV –ve			
	Case	Control	Adj.age & sex	Adj. age, sex, Carstairs	Case	Control	Adj.age & sex	Adj. age, sex, Carstairs
Risk factor	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)
Tot child infect age ≥5 2,3,4 vs 1	17 (43.6)	96 (58.9)	0.52 (0.26-1.06)	0.53 (0.25-1.11)	26 (53.1)	96 (58.9)	0.80 (0.42-1.54)	0.71 (0.36-1.39)
Tot child infect age ≥5 2 vs. 1	4 (10.3)	30 (18.4)	0.36 (0.11-1.12)	0.34 (0.11-1.10)	11 (22.4)	30 (18.4)	1.05 (0.45-2.44)	0.98 (0.41-2.35)
3 vs. 1	4 (10.3)	30 (18.4)	0.35 (0.11-1.10)	0.40 (0.12-1.28)	7 (14.3)	30 (18.4)	0.61 (0.24-1.59)	0.51 (0.18-1.43)
4 vs. 1	9 (23.1)	36 (22.1)	0.67 (0.28-1.61)	0.65 (0.26-1.66)	8 (16.3)	36 (22.1)	0.60 (0.24-1.48)	0.60 (0.24-1.50)
Tot herp infect any age 1,2,3 v. 0	29 (69.0)	138 (83.6)	0.41 (0.19-0.92)	0.39 (0.17-0.91)	34 (69.4)	138 (83.6)	0.40 (0.19-0.86)	0.38 (0.17-0.83)
Tot herp infect at any age 1 vs. 0	17 (40.5)	79 (47.9)	0.42 (0.18-0.99)	0.40 (0.16-1.00)	21 (42.9)	79 (47.9)	0.36 (0.16-1.80)	0.35 (0.15-0.82)
2 vs. 0	10 (23.8)	41 (24.8)	0.48 (0.18-1.30)	0.46 (0.16-1.30)	10 (20.4)	41 (24.8)	0.36 (0.13-0.93)	0.38 (0.14-1.03)
3 vs. 0	2 (4.8)	18 (10.9)	0.21 (0.04-1.09)	0.20 (0.03-1.07)	3 (6.1)	18 (10.9)	0.25 (0.06-0.99)	0.25 (0.06-1.01)
Tot add infect at any age 1,2 vs. 0	15 (35.7)	43 (26.1)	1.76 (0.83-3.72)	1.74 (0.79-3.84)	17 (34.0)	43 (26.1)	2.03 (0.97-4.25)	1.99 (0.94-4.23)
Tot add infect at any age 1 vs. 0	11 (26.2)	35 (21.2)	1.57 (0.69-3.57)	1.55 (0.65-3.71)	14 (28.0)	35 (21.2)	1.93 (0.89-4.19)	2.01 (0.90-4.49)
2 vs. 0	4 (9.5)	8 (4.8)	2.62 (0.71-9.60)	2.55 (0.67-9.66)	3 (6.0)	8 (4.8)	1.82 (0.44-7.58)	1.92 (0.46-8.09)

Table 8.11: Index total infections age 50+ by EBV status (cont)

	EBV +ve				EBV –ve			
	Case	Control	Adj.age & sex	Adj. age, sex, Carstairs	Case	Control	Adj.age & sex	Adj. age, sex, Carstairs
Risk factor	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)
Tchild+add infect any age 2-5 vs. 1	27 (64.3)	133 (80.6)	0.42 (0.20-0.89)	0.44 (0.20-0.97)	36 (72.0)	133 (80.6)	0.63 (0.30-1.32)	0.58 (0.27-1.22)
Tchild+add infect any age 2 vs. 1	10 (23.8)	25 (15.2)	0.83 (0.32-2.18)	0.78 (0.29-2.12)	14 (28.0)	25 (15.2)	1.22 (0.49-3.03)	1.23 (0.47-3.18)
3 vs. 1	4 (9.5)	46 (27.9)	0.17 (0.05-0.58)	0.21 (0.06-0.71)	11 (22.0)	46 (27.9)	0.48 (0.19-1.21)	0.51 (0.20-1.31)
4 vs. 1	13 (31.0)	62 (37.6)	0.43 (0.18-1.04)	0.45 (0.18-1.11)	11 (22.0)	62 (37.6)	0.38 (0.16-0.95)	0.36 (0.14-0.93)
Tot comb infect any age 2-5 vs. 1	33 (78.6)	140 (84.8)	0.64 (0.27-1.52)	0.72 (0.29-1.80)	41 (80.4)	140 (84.8)	0.75 (0.33-1.71)	0.82 (0.34-1.93)
Tot comb infect any age 2 vs. 1	6 (14.3)	16 (9.7)	1.02 (0.30-3.45)	1.17 (0.33-4.13)	7 (13.7)	16 (9.7)	1.10 (0.34-3.55)	1.11 (0.32-3.82)
3 vs. 1	11 (26.2)	27 (16.4)	1.08 (0.38-3.07)	1.23 (0.41-3.68)	12 (23.5)	27 (16.4)	1.07 (0.39-2.96)	1.30 (0.46-3.73)
4 vs. 1	4 (9.5)	29 (17.6)	0.37 (0.10-1.35)	0.43 (0.11-1.66)	10 (19.6)	29 (17.6)	0.87 (0.31-2.47)	0.86 (0.29-2.58)
5 vs. 1	12 (28.6)	68 (41.2)	0.47 (0.17-1.28)	0.51 (0.18-1.47)	12 (23.5)	68 (41.2)	0.47 (0.17-1.21)	0.51 (0.19-1.41)

Childhood infections: chicken pox, German measles, measles, mumps, whooping cough
Additional infections: tuberculosis, diphtheria, meningitis, malaria, scarlet fever, rheumatic fever, pneumonia
Herpes virus infections: chicken pox, HSV-1, HSV-2, shingles
Total combined infections: Sum of all of the above

Table 8.12: Total series index individual infections

	Case		Control		Adj.age & sex	Adj. age, sex, Carstairs
Risk Factor	No	No (%) yes	No	No (%) yes	OR (95%CI)	OR (95%CI)
Chicken Pox at any age	449	346 (77.1)	486	375 (77.2)	0.91 (0.67-1.25)	1.03 (0.74-1.44)
German Measles at any age	438	170 (38.8)	463	189 (40.8)	0.90 (0.69-1.18)	0.87 (0.66-1.16)
Measles at any age	449	314 (69.9)	484	362 (74.8)	0.85 (0.63-1.14)	0.89 (0.65-1.21)
Measles age 0-4	439	93 (21.2)	478	99 (20.7)	1.05 (0.76-1.45)	1.05 (0.75-1.47)
Measles age 5-10	439	202 (46.0)	478	242 (50.6)	0.88 (0.67-1.15)	0.89 (0.67-1.17)
Measles age ≥11	439	10 (2.3)	478	16 (3.3)	0.67 (0.30-1.51)	0.88 (0.38-2.04)
Measles age ≥5	439	211 (48.1)	478	257 (53.8)	0.84 (0.65-1.10)	0.87 (0.66-1.15)
Mumps at any age	460	233 (50.7)	488	264 (54.1)	0.88 (0.68-1.14)	0.95 (0.73-1.24)
Whooping Cough at any age	457	80 (17.5)	494	106 (21.5)	0.85 (0.61-1.19)	0.87 (0.61-1.23)
Herpes Simplex 1 at any age	488	170 (34.8)	511	214 (41.9)	0.76 (0.58-0.98)	0.75 (0.57-0.97)
Herpes Simplex 2 at any age	489	3 (0.6)	512	2 (0.4)	1.58 (0.26-9.56)	∞ (0.00-∞)
Shingles at any age	486	59 (12.1)	510	63 (12.4)	1.08 (0.73-1.59)	1.13 (0.76-1.68)
Pneumonia at any age	487	37 (7.6)	511	42 (8.2)	1.01 (0.63-1.61)	1.02 (0.62-1.67)
Scarlet Fever at any age	491	39 (7.9)	512	24 (4.7)	1.95 (1.14-3.33)	1.99 (1.15-3.46)

Table 8.13: Index individual infections by age group (16-34 & 50+)

	16-34				50+			
	Case	Control	Adj.age & sex	Adj. age, sex, Carstairs	Case	Control	Adj.age & sex	Adj. age, sex, Carstairs
Risk factor	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)
Chicken Pox at any age	203 (86.4)	175 (81.8)	1.42 (0.85-2.27)	1.59 (0.93-2.72)	69 (58.5)	108 (70.6)	0.56 (0.34-0.94)	0.61 (0.36-1.05)
German Measles at any age	96 (42.5)	82 (39.8)	1.13 (0.77-1.66)	1.60 (0.73-1.63)	38 (31.7)	52 (35.6)	0.79 (0.46-1.34)	0.72 (0.42-1.26)
Measles at any age	134 (58.8)	129 (60.3)	0.96 (0.65-1.41)	1.03 (0.69-1.53)	104 (81.3)	135 (88.2)	0.54 (0.28-1.07)	0.56 (0.27-1.14)
Measles age 0-4	52 (23.0)	33 (15.4)	1.69 (1.04-2.75)	1.65 (0.99-2.73)	25 (20.2)	36 (24.0)	0.81 (0.45-1.44)	0.87 (0.48-1.59)
Measles age 5-10	74 (32.7)	91 (42.5)	0.66 (0.45-0.98)	0.72 (0.48-1.08)	74 (59.7)	92 (61.3)	0.88 (0.53-1.44)	0.83 (0.49-1.39)
Measles age ≥11	6 (2.7)	6 (2.8)	0.95 (0.30-3.01)	0.97 (0.30-3.12)	1 (0.8)	4 (2.7)	0.34 (0.04-3.07)	0.43 (0.05-3.96)
Measles age ≥5	80 (35.4)	96 (44.9)	0.67 (0.46-0.99)	0.74 (0.49-1.10)	75 (60.5)	96 (64.0)	0.81 (0.49-1.34)	0.78 (0.46-1.32)
Mumps at any age	106 (45.7)	107 (49.5)	0.86 (0.59-1.24)	0.93 (0.63-1.67)	69 (52.3)	89 (58.2)	0.78 (0.49-1.25)	0.78 (0.48-1.27)
Whooping Cough at any age	22 (9.5)	20 (9.2)	1.02 (0.54-1.93)	0.92 (0.48-1.78)	39 (31.0)	58 (37.9)	0.75 (0.45-1.24)	0.76 (0.45-1.28)
Herpes Simplex 1 at any age	74 (30.8)	84 (38.2)	0.73 (0.49-1.08)	0.74 (0.49-1.10)	55 (37.9)	71 (43.3)	0.77 (0.48-1.23)	0.70 (0.42-1.15)
Herpes Simplex 2 at any age	2 (0.8)	0 (0.0)	∞ (0.00-∞)	∞ (0.00-∞)	1 (0.7)	0 (0.0)	∞ (0.00-∞)	∞ (0.00-∞)
Shingles at any age	19 (8.0)	11 (5.0)	1.72 (0.80-3.73)	1.95 (0.88-4.35)	30 (20.8)	36 (22.1)	0.97 (0.56-1.70)	1.02 (0.58-1.79)
Pneumonia at any age	6 (2.5)	7 (3.2)	0.80 (0.26-2.42)	0.82 (0.27-2.49)	18 (12.5)	25 (15.2)	0.84 (0.44-1.64)	0.93 (0.47-1.86)
Scarlet Fever at any age	12 (5.0)	4 (1.8)	2.79 (0.88-8.80)	2.44 (0.76-7.84)	20 (13.6)	13 (7.9)	2.15 (1.01-4.60)	2.06 (0.94-4.49)

≥5 yrs) are put in a multivariate model none of the infections have any statistically significant results. The statistically significant effect of scarlet fever seen in the total series is confined to the 50+ years age at diagnosis group.

The results for the total series by EBV status for individual infections are in Table 8.14. There are very few statistically significant results for the individual infections and the ORs are very similar in both EBV subgroups. Formal comparison of the EBV subgroups reveals none of the differences were statistically significant.

The results for EBV subgroups for age 16-34 years age at diagnosis are in Table 8.15. Again very few results were statistically significant. Measles is associated with an increased risk of EBV +ve HD and a null finding for EBV –ve. Measles aged ≥5 years is associated with an increased risk of EBV +ve HD and a statistically significantly decreased risk of EBV –ve HD. Measles was the only infection that differed statistically significantly between the EBV subgroups with late measles (age ≥5 yrs) significantly more likely to be associated with EBV +ve HD (OR 3.26, 1.49-7.11, p=0.003).

The results by EBV status by 50+ age at diagnosis group are in Table 8.16. The risk of HD associated with measles was similar by EBV subgroup and close to the null. Whooping cough is strongly associated with both EBV +ve and EBV –ve HD (OR 7.21, 1.09-47.58 EBV +ve; OR 5.87, 0.93-37.13 EBV –ve). There were no statistically significant differences between the EBV subgroups.

The main individual infection of interest is IM, both in the index and first degree relatives. The results for the total series are in Table 8.17 and by age at diagnosis group in Table 8.18. From Table 8.17 it can be seen that IM in the index, however it is defined, is associated with a statistically significant increased risk of HD. IM in a first degree relative is also associated with an increased risk but this is not statistically significant. When HD is split by age at diagnosis IM is associated with a statistically significantly increased risk of

Table 8.14: Index individual infections by EBV status

	EBV +ve				EBV -ve			
	Case	Control	Adj.age & sex	Adj. age, sex, Carstairs	Case	Control	Adj.age & sex	Adj. age, sex, Carstairs
Risk factor	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)
Chicken Pox at any age	69 (75.8)	375 (77.2)	0.90 (0.53-1.54)	1.10 (0.63-1.95)	167 (81.1)	375 (77.2)	1.11 (0.73-1.68)	1.13 (0.73-1.74)
German Measles at any age	33 (37.9)	189 (40.8)	0.93 (0.58-1.50)	0.93 (0.57-1.53)	87 (42.2)	189 (40.8)	1.02 (0.73-1.43)	0.91 (0.64-1.30)
Measles at any age	70 (76.9)	362 (74.8)	1.16 (0.67-2.00)	1.26 (0.71-2.24)	138 (67.3)	362 (74.8)	0.82 (0.57-1.19)	0.86 (0.58-1.27)
Measles age 0-4	17 (19.1)	99 (20.7)	0.92 (0.52-1.64)	0.97 (0.54-1.76)	45 (22.3)	99 (20.7)	1.12 (0.75-1.67)	1.17 (0.77-1.77)
Measles age 5-10	51 (57.3)	242 (50.6)	1.32 (0.83-2.09)	1.32 (0.82-2.14)	85 (42.1)	242 (50.6)	0.80 (0.57-1.13)	0.78 (0.55-1.12)
Measles age ≥11	0 (0.0)	16 (3.3)	0.00 (0.00-∞)	0.00 (0.00-∞)	6 (3.0)	16 (3.3)	0.94 (0.36-2.45)	1.19 (0.44-3.21)
Measles age ≥5	51 (57.3)	257 (53.8)	1.16 (0.73-1.84)	1.20 (0.74-1.93)	90 (44.6)	257 (53.8)	0.78 (0.56-1.10)	0.79 (0.55-1.13)
Mumps at any age	48 (50.5)	264 (54.1)	0.85 (0.55-1.32)	0.93 (0.59-1.48)	105 (49.8)	264 (54.1)	0.88 (0.63-1.22)	0.93 (0.66-1.31)
Whooping Cough at any age	17 (17.7)	106 (21.5)	0.83 (0.46-1.51)	0.83 (0.44-1.57)	34 (16.5)	106 (21.5)	0.88 (0.56-1.37)	0.93 (0.59-1.47)
Herpes Simplex 1 at any age	37 (36.6)	214 (41.9)	0.81 (0.52-1.26)	0.76 (0.48-1.20)	68 (31.2)	214 (41.9)	0.66 (0.47-0.92)	0.69 (0.49-0.98)
Herpes Simplex 2 at any age	0 (0.0)	2 (0.4)	0.00 (0.00-∞)	NA	0 (0.0)	2 (0.4)	0.00 (0.00-∞)	NA
Shingles at any age	10 (10.1)	63 (12.4)	0.83 (0.40-1.71)	0.83 (0.40-1.73)	24 (11.0)	63 (12.4)	1.05 (0.63-1.77)	1.04 (0.61-1.77)
Pneumonia at any age	10 (10.1)	42 (8.2)	1.25 (0.59-2.64)	1.36 (0.63-2.92)	15 (6.9)	42 (8.2)	1.06 (0.56-1.99)	1.00 (0.51-1.95)
Scarlet Fever at any age	9 (8.9)	24 (4.7)	2.12 (0.94-4.79)	1.96 (0.82-4.68)	20 (9.1)	24 (4.7)	2.48 (1.31-4.70)	2.56 (1.33-4.93)

Table 8.15: Index individual infections age 16-34 by EBV status

	EBV +ve				EBV -ve			
	Case	Control	Adj.age & sex	Adj. age, sex, Carstairs	Case	Control	Adj.age & sex	Adj. age, sex, Carstairs
Risk factor	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)
Chicken Pox at any age	38 (86.4)	175 (81.8)	1.34 (0.52-3.44)	1.73 (0.62-4.83)	106 (88.3)	175 (81.8)	1.66 (0.86-3.21)	1.78 (0.90-3.54)
German Measles at any age	17 (41.5)	82 (39.8)	1.12 (0.56-2.23)	1.11 (0.54-2.29)	53 (44.9)	82 (39.8)	1.23 (0.78-1.95)	1.13 (0.70-1.84)
Measles at any age	29 (69.0)	129 (60.3)	1.53 (0.74-3.15)	1.85 (0.86-3.99)	69 (59.0)	129 (60.3)	0.96 (0.60-1.54)	0.98 (0.61-1.60)
Measles age 0-4	6 (14.3)	33 (15.4)	1.01 (0.39-2.62)	1.03 (0.39-2.74)	34 (29.3)	33 (15.4)	2.28 (1.31-3.95)	2.21 (1.25-3.92)
Measles age 5-10	23 (54.8)	91 (42.5)	1.64 (0.83-3.24)	1.96 (0.96-4.00)	31 (26.7)	91 (42.5)	0.50 (0.30-0.82)	0.52 (0.31-0.87)
Measles age ≥11	0 (0.0)	6 (2.8)	0.00 (0.00-∞)	0.00 (0.00-∞)	3 (2.6)	6 (2.8)	1.03 (0.25-4.24)	1.01 (0.24-4.22)
Measles age ≥5	23 (54.8)	96 (44.9)	1.48 (0.75-2.91)	1.73 (0.85-3.54)	34 (29.3)	96 (44.9)	0.52 (0.32-0.84)	0.54 (0.33-0.89)
Mumps at any age	24 (55.8)	107 (49.5)	1.21 (0.62-2.36)	1.38 (0.68-2.77)	56 (46.3)	107 (49.5)	0.88 (0.56-1.37)	0.96 (0.60-1.52)
Whooping Cough at any age	4 (9.3)	20 (9.2)	1.04 (0.32-3.31)	0.85 (0.23-3.14)	10 (8.5)	20 (9.2)	0.90 (0.40-2.00)	0.91 (0.41-2.02)
Herpes Simplex 1 at any age	13 (29.5)	84 (38.2)	0.70 (0.34-1.44)	0.62 (0.29-1.32)	34 (27.6)	84 (38.2)	0.62 (0.38-1.01)	0.61 (0.37-1.01)
Herpes Simplex 2 at any age	0 (0.0)	0 (0.0)	NA	NA	0 (0.0)	0 (0.0)	NA	NA
Shingles at any age	2 (4.7)	11 (5.0)	0.98 (0.21-4.68)	1.12 (0.23-5.51)	11 (8.9)	11 (5.0)	1.78 (0.74-4.29)	2.03 (0.82-5.02)
Pneumonia at any age	1 (2.3)	7 (3.2)	0.78 (0.09-6.75)	0.80 (0.09-7.06)	3 (2.5)	7 (3.2)	0.74 (0.19-2.93)	0.75 (0.19-2.98)
Scarlet Fever at any age	3 (6.8)	4 (1.8)	3.55 (0.73-17.37)	3.19 (0.63-16.19)	7 (5.7)	4 (1.8)	3.36 (0.95-11.82)	2.84 (0.78-10.40)

Table 8.16: Index individual infections age 50+ by EBV status

	EBV +ve				EBV -ve			
	Case	Control	Adj.age & sex	Adj. age, sex, Carstairs	Case	Control	Adj.age & sex	Adj. age, sex, Carstairs
Risk factor	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)
Chicken Pox at any age	20 (62.5)	108 (70.6)	0.65 (0.29-1.47)	0.74 (0.32-1.73)	23 (54.8)	108 (70.6)	0.48 (0.24-0.98)	0.47 (0.22-0.99)
German Measles at any age	11 (34.4)	52 (35.6)	0.88 (0.38-2.06)	0.83 (0.34-2.02)	14 (31.8)	52 (35.6)	0.82 (0.38-1.75)	0.72 (0.33-1.57)
Measles at any age	31 (88.6)	135 (88.2)	0.95 (0.29-3.11)	0.85 (0.25-2.86)	33 (73.3)	135 (88.2)	0.33 (0.14-0.77)	0.33 (0.13-0.81)
Measles age 0-4	9 (25.7)	36 (24.0)	1.12 (0.48-2.62)	1.17 (0.49-2.81)	6 (13.6)	36 (24.0)	0.47 (0.18-1.27)	0.52 (0.20-1.36)
Measles age 5-10	22 (62.9)	92 (61.3)	1.00 (0.46-2.18)	0.89 (0.40-1.98)	25 (56.8)	92 (61.3)	0.79 (0.39-1.60)	0.76 (0.37-1.56)
Measles age ≥11	0 (0.0)	4 (2.7)	0.00 (0.00-∞)	0.00 (0.00-∞)	1 (2.3)	4 (2.7)	1.09 (0.12-10.30)	1.19 (0.12-11.48)
Measles age ≥5	22 (62.9)	96 (64.0)	0.89 (0.41-1.95)	0.82 (0.37-1.83)	26 (59.1)	96 (64.0)	0.80 (0.39-1.62)	0.76 (0.37-1.59)
Mumps at any age	14 (37.8)	89 (58.2)	0.43 (0.21-0.90)	0.47 (0.22-1.02)	23 (48.9)	89 (58.2)	0.65 (0.34-1.28)	0.61 (0.31-1.21)
Whooping Cough at any age	13 (34.2)	58 (37.9)	0.87 (0.41-1.85)	0.95 (0.43-2.12)	11 (26.2)	58 (37.9)	0.62 (0.29-1.33)	0.67 (0.30-1.46)
Herpes Simplex 1 at any age	17 (40.5)	71 (43.3)	0.91 (0.44-1.86)	0.74 (0.34-1.60)	17 (34.7)	71 (43.3)	0.70 (0.35-1.38)	0.72 (0.35-1.46)
Herpes Simplex 2 at any age	0 (0.0)	0 (0.0)	NA	NA	0 (0.0)	0 (0.0)	NA	NA
Shingles at any age	6 (14.6)	36 (22.1)	0.63 (0.24-1.62)	0.61 (0.23-1.60)	10 (20.4)	36 (22.1)	1.03 (0.46-2.30)	1.01 (0.45-2.29)
Pneumonia at any age	8 (20.0)	25 (15.2)	1.48 (0.60-3.63)	1.68 (0.66-4.26)	7 (14.3)	25 (15.2)	1.14 (0.45-2.92)	1.06 (0.39-2.86)
Scarlet Fever at any age	6 (14.3)	13 (7.9)	2.16 (0.75-6.21)	1.88 (0.59-6.00)	8 (16.0)	13 (7.9)	3.01 (1.11-8.18)	3.06 (1.11-8.40)

Table 8.17: Total series IM in index and family members

	Case		Control		Adj.age & sex	Adj. age, sex, Carstairs
Risk Factor	No	No (%) yes	No	No (%) yes	OR (95%CI)	OR (95%CI)
Index Definite IM	493	30 (6.1)	512	13 (2.5)	2.34 (1.20-4.56)	2.42 (1.21-4.88)
Index Suspected IM	493	49 (9.9)	512	25 (4.9)	2.07 (1.25-3.41)	2.17 (1.30-3.64)
Index Level IM	493	49 (9.9)	512	25 (4.9)	1.57 (1.15-2.14)	1.62 (1.18-2.24)
Family IM (continuous)	493	52 (10.8)	512	47 (9.3)	5.21 (0.98-27.77)	7.35 (1.29-42.06)
Family IM (dichotomy)	493	52 (10.8)	512	47 (9.3)	1.16 (0.76-1.76)	1.28 (0.83-1.98)

Table 8.18: IM in index and family members by age group

	16-34				50+			
	Case	Control	Adj.age & sex	Adj. age, sex, Carstairs	Case	Control	Adj.age & sex	Adj. age, sex, Carstairs
Risk factor	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)
Index Definite IM	21 (8.7)	7 (3.2)	2.91 (1.21-6.99)	2.90 (1.21-7.00)	1 (0.7)	1 (0.6)	0.95 (0.06-15.56)	1.07 (0.07-17.79)
Index Suspected IM	30 (12.4)	13 (5.9)	2.27 (1.15-4.48)	2.31 (1.16-4.57)	7 (4.7)	6 (3.6)	1.16 (0.37-3.59)	1.25 (0.40-3.89)
Index Level IM	30 (12.4)	13 (5.9)	1.68 (1.12-2.53)	1.69 (1.12-2.55)	7 (4.7)	6 (3.6)	1.10 (0.43-2.80)	1.17 (0.46-3.00)
Family IM (continuous)	31 (13.1)	13 (6.0)	41.34 (3.25-526.16)	39.40 (3.02-514.64)	11 (7.6)	15 (9.3)	0.34 (0.00-43.68)	1.50 (0.01-242.87)
Family IM (dichotomy)	31 (13.1)	13 (6.0)	2.37 (1.20-4.66)	2.35 (1.19-4.66)	11 (7.6)	15 (9.3)	0.76 (0.33-1.71)	0.96 (0.41-2.25)

HD aged 16-34 but has not aged 50+ years (Table 8.18). This contrast is also seen for IM in a first degree relative.

Results for IM by EBV status in the total series are in Table 8.19 and for separate age at diagnosis groups in Tables 8.20 (16-34 yrs) and 8.21 (50+ yrs). The effect of IM in the index was not limited to the EBV +ve cases in the total series or in either age at diagnosis group. However, in the 16-34 years age at diagnosis group definite IM is associated with a greater risk of EBV +ve HD than EBV –ve, but the difference between the subgroups is not statistically significant (case series comparison EBV +ve vs, EBV –ve OR-1.82, 0.54-6.11). There is little effect of IM in the 50+ years age group (EBV +ve or EBV –ve). IM in a family member is associated with a greater risk of EBV +ve HD in the total series and the 16-34 age group but not the 50+ years. It is only in the 16-34 years age at diagnosis group that the effect is statistically significantly different between EBV +ve and –ve HD with the chances of EBV +ve HD following IM in a first degree relative being 3 times that of EBV –ve.

The effect of IM was considered after adjustment for the presence of the five childhood infections (plus scarlet fever in the 50+ yrs age at diagnosis group) age 0-4 and ≥5 years to see if IM was merely a marker for late age at first infection. While none of the childhood infections were individually statistically significant when adjusted for on another, IM in the total series in all its forms was associated with a statistically significantly increased risk of HD. This relationship was seen in the 16-34 but not the 50+ years age group. The effect of IM in the EBV subgroups was also analysed after adjustment for the childhood infectious illnesses and removal of the tonsils. The results for IM in both EBV subgroups at all ages, age 16-34, and age 50+ years was unaffected by these adjustments.

Family Health:

Results for family health variables for the total series are in Table 8.22. There are very few statistically significant results. There is evidence that malignancy in a family member is associated with a statistically significant increased risk of index HD. This is

Table 8.19: IM in index and family members by EBV status

	EBV +ve				EBV –ve			
	Case	Control	Adj.age & sex	Adj. age, sex, Carstairs	Case	Control	Adj.age & sex	Adj. age, sex, Carstairs
Risk factor	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)
Index Definite IM	5 (5.0)	13 (2.5)	1.99 (0.68-5.82)	2.12 (0.71-6.33)	13 (5.9)	13 (2.5)	2.11 (0.96-4.66)	2.37 (1.05-5.35)
Index Suspected IM	11 (10.9)	25 (4.9)	2.30 (1.08-4.88)	2.65 (1.22-5.75)	19 (8.6)	25 (4.9)	1.69 (0.91-3.16)	1.84 (0.97-3.47)
Index Level IM	11 (10.9)	25 (4.9)	1.60 (0.99-2.57)	1.71 (1.05-2.78)	19 (8.6)	25 (4.9)	1.42 (0.98-2.06)	1.50 (1.03-2.20)
Family IM (continuous)	14 (14.0)	47 (9.3)	27.69 (2.32-330.54)	40.86 (2.88-579.81)	24 (11.1)	47 (9.3)	5.56 (0.61-50.53)	7.94 (0.83-76.16)
Family IM (dichotomy)	14 (14.0)	47 (9.3)	1.60 (0.84-3.04)	1.78 (0.90-3.51)	24 (11.1)	47 (9.3)	1.25 (0.74-2.11)	1.35 (0.78-2.33)

Table 8.20: IM in index and family members by EBV status age 16-34

	EBV +ve				EBV –ve			
	Case	Control	Adj.age & sex	Adj. age, sex, Carstairs	Case	Control	Adj.age & sex	Adj. age, sex, Carstairs
Risk factor	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)
Index Definite IM	5 (11.4)	7 (3.2)	3.32 (0.97-11.35)	3.09 (0.90-10.60)	9 (7.3)	7 (3.2)	2.40 (0.87-6.65)	2.46 (0.89-6.82)
Index Suspected IM	7 (15.9)	13 (5.9)	2.60 (0.95-7.15)	2.65 (0.95-7.37)	13 (10.5)	13 (5.9)	1.91 (0.85-4.27)	1.97 (0.87-4.43)
Index Level IM	7 (15.9)	13 (5.9)	1.82 (1.01-3.28)	1.79 (0.99-3.24)	13 (10.5)	13 (5.9)	1.52 (0.94-2.46)	1.55 (0.96-2.50)
Family IM (continuous)	12 (27.3)	13 (6.0)	783.04 (25.75-∞)	793.08 (22.27-∞)	12 (10.0)	13 (6.0)	12.41 (0.50-310.01)	14.76 (0.58-378.00)
Family IM (dichotomy)	12 (27.3)	13 (6.0)	5.35 (2.21-12.96)	5.33 (2.11-13.46)	12 (10.0)	13 (6.0)	1.74 (0.76-3.98)	1.83 (0.92-3.23)

Table 8.21: IM in index and family members by EBV status age 50+

	EBV +ve				EBV –ve			
	Case	Control	Adj.age & sex	Adj. age, sex, Carstairs	Case	Control	Adj.age & sex	Adj. age, sex, Carstairs
Risk factor	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)
Index Definite IM	0 (0.0)	1 (0.6)	0.00 (0.00-∞)	0.00 (0.00-∞)	0 (0.0)	1 (0.6)	0.00 (0.00-∞)	0.00 (0.00-∞)
Index Suspected IM	3 (7.1)	6 (3.6)	1.80 (0.41-7.90)	2.22 (0.48-10.29)	2 (3.9)	6 (3.6)	0.80 (0.15-4.27)	0.81 (0.15-4.33)
Index Level IM	3 (7.1)	6 (3.6)	1.39 (0.38-5.11)	1.61 (0.42-6.12)	2 (3.9)	6 (3.6)	0.71 (0.16-3.11)	0.72 (0.16-3.13)
Family IM (continuous)	2 (4.8)	15 (9.3)	0.00 (0.00-169.98)	0.00 (0.00-∞)	5 (10.2)	15 (9.3)	4.64 (0.01-∞)	11.97 (0.03-∞)
Family IM (dichotomy)	2 (4.8)	15 (9.3)	0.48 (0.10-2.18)	0.64 (0.13-3.07)	5 (10.2)	15 (9.3)	1.05 (0.35-3.08)	1.21 (0.40-3.71)

Table 8.22: Total series presence of illness in first degree relative

	Case		Control		Adj.age & sex	Adj. age, sex, Carstairs
Risk Factor	No	No (%) yes	No	No (%) yes	OR (95%CI)	OR (95%CI)
HD	490	10 (2.0)	508	2 (0.4)	5.17 (1.12-23.79)	4.77 (1.02-22.36)
Haematological malignancy	490	19 (3.9)	508	9 (1.8)	2.39 (1.07-5.36)	2.58 (0.1.10-6.05)
Young onset cancer	490	55 (11.2)	508	34 (6.7)	1.82 (1.16-2.85)	2.15 (1.32-3.49)
Haematological/young onset cancer	490	60 (12.2)	508	39 (7.7)	1.75 (1.14-2.68)	2.02 (1.28-3.19)
Any cancer	490	139 (28.4)	508	134 (26.4)	1.24 (0.92-1.66)	1.40 (1.02-1.90)
Any Auto-immune disease	493	93 (18.9)	512	97 (18.9)	1.10 (0.84-1.22)	0.99 (0.82-1.20)

numbers and percentages for unadjusted analysis only
Results are presented unadjusted for number of first degree relatives as this adjustment did not result in statistically significantly better fitting models

particularly so for HD in a first degree relative (OR 5.17, 1.12-23.79) or haematological malignancy combined (OR 2.39, 1.07-5.36). The patterns of risk for family health variables are the same for the age at diagnosis subgroups (Table 8.23).

Results for family health variables by EBV status are in Table 8.24. The results for cancer in a first degree relative are similar in the EBV subgroups. However, when all malignancies in first degree relatives are considered there is a statistically significant increase in risk of EBV +ve HD but little effect on risk of EBV –ve HD. When the two are formally compared the excess risk of EBV +ve HD, compared with EBV –ve HD, is almost statistically significant (p=0.064). Results for the 16-34 and 50+ years age at diagnosis groups by EBV status are in Tables 8.25 and 8.26 but none are statistically significant.

Table 8.23: Presence of illness in first degree relatives by age group (16-34 & 50+)

	16-34				50+			
	Case	Control	Adj.age & sex	Adj. age, sex, Carstairs	Case	Control	Adj.age & sex	Adj. age, sex, Carstairs
Risk factor	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)
HD	6 (2.5)	1 (0.5)	5.46 (0.65-45.86)	4.35 (0.50-37.83)	3 (2.1)	0 (0.0)	∞ (0.00-∞)	∞ (0.00-∞)
Haematological malignancy	7 (2.9)	1 (0.5)	6.38 (0.78-52.57)	5.06 (0.60-42.75)	8 (5.5)	6 (3.7)	1.63 (0.55-4.82)	1.99 (0.63-6.32)
Young onset cancer	23 (9.5)	11 (5.0)	2.00 (0.95-4.21)	2.71 (1.18-6.25)	17 (11.7)	15 (9.2)	1.33 (0.63-2.77)	1.29 (0.60-2.81)
Haem/young onset cancer	23 (9.5)	12 (5.5)	1.82 (0.88-3.76)	2.39 (1.07-5.34)	21 (14.5)	17 (10.4)	1.48 (0.74-2.94)	1.45 (0.71-2.97)
Any cancer	35 (14.5)	22 (10.0)	1.56 (0.88-2.76)	1.75 (0.95-3.21)	6 (4.1)	4 (2.5)	1.19 (0.76-1.87)	1.24 (0.78-1.99)
Any Auto-immune disease	35 (14.5)	32 (14.5)	0.97 (0.72-1.30)	0.97 (0.72-1.31)	30 (20.3)	39 (23.7)	0.96 (0.70-1.30)	0.96 (0.69-1.32)

numbers and percentages for unadjusted analysis only

Results are presented unadjusted for number of first degree relatives as this adjustment did not result in statistically significantly better fitting models

Table 8.24: Presence of illness in first degree relatives by EBV status

	EBV +ve				EBV -ve			
	Case	Control	Adj.age & sex	Adj. age, sex, Carstairs	Case	Control	Adj.age & sex	Adj. age, sex, Carstairs
Risk factor	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)
HD	3 (3.0)	2 (0.4)	6.61 (1.08-40.35)	6.62 (1.05-41.68)	5 (2.3)	2 (0.4)	5.38 (1.02-28.42)	4.33 (0.77-24.36)
Haematological malignancy	4 (4.0)	9 (1.8)	2.18 (0.65-7.30)	2.14 (0.61-7.47)	6 (2.7)	9 (1.8)	1.81 (0.63-5.26)	1.70 (0.54-5.36)
Young onset cancer	11 (10.9)	34 (6.7)	1.71 (0.83-3.52)	2.26 (1.06-4.81)	22 (10.0)	34 (6.7)	1.65 (0.94-2.92)	1.81 (0.97-3.36)
Haem/young onset cancer	13 (12.9)	39 (7.7)	1.79 (0.91-3.50)	2.26 (1.12-4.56)	23 (10.5)	39 (7.7)	1.52 (0.87-2.62)	1.62 (0.90-2.93)
Any cancer	34 (33.7)	134 (26.7)	1.46 (0.89-2.37)	1.83 (1.10-3.06)	49 (22.3)	134 (26.4)	1.00 (0.67-1.48)	1.07 (0.70-1.62)
Any Auto-immune disease	22 (21.8)	97 (18.7)	1.10 (0.81-1.51)	1.09 (0.79-1.50)	42 (19.0)	97 (18.9)	1.00 (0.79-1.27)	0.98 90.76-1.26)

numbers and percentages for unadjusted analysis only

Results are presented unadjusted for number of first degree relatives as this adjustment did not result in statistically significantly better fitting models

Table 8.25: Presence of illness in first degree relatives age 16-34 by EBV status

	EBV +ve				EBV -ve			
	Case	Control	Adj.age & sex	Adj. age, sex, Carstairs	Case	Control	Adj.age & sex	Adj. age, sex, Carstairs
Risk factor	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)
HD	1 (2.3)	1 (0.5)	3.69 (0.22-66.70)	4.11 (0.25-68.43)	4 (3.2)	1 (0.5)	9.21 (1.00-85.14)	6.43 (0.64-64.69)
Haematological malignancy	1 (2.3)	1 (0.5)	3.69 (0.22-66.70)	4.11 (0.25-68.43)	4 (3.2)	1 (0.5)	9.21 (1.00-85.14)	6.43 (0.64-64.69)
Young onset cancer	3 (6.8)	11 (5.0)	1.29 (0.33-4.99)	2.08 (0.51-8.49)	10 (8.1)	11 (5.0)	1.75 (0.72-4.27)	2.11 (0.78-5.66)
Haem/young onset cancer	3 (6.8)	12 (5.5)	1.11 (0.29-4.26)	1.76 (0.44-7.05)	10 (8.1)	12 (5.5)	1.60 (0.67-3.86)	1.88 (0.72-4.92)
Any cancer	7 (15.9)	22 (10.0)	2.01 (0.77-5.24)	2.69 (0.99-7.31)	14 (11.3)	22 (10.0)	1.14 (0.56-2.33)	1.12 (0.52-2.42)
Any Auto-immune disease	5 (13.3)	32 (14.5)	0.82 (0.44-1.54)	0.87 (0.47-1.66)	21 (16.9)	32 (14.5)	1.07 (0.77-1.50)	1.06 (0.75-1.50)

numbers and percentages for unadjusted analysis only

Results are presented unadjusted for number of first degree relatives as this adjustment did not result in statistically significantly better fitting models

Table 8.26: Presence of illness in first degree relatives age 50+by EBV status

	EBV +ve				EBV -ve			
	Case	Control	Adj.age & sex	Adj. age, sex, Carstairs	Case	Control	Adj.age & sex	Adj. age, sex, Carstairs
Risk factor	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)	No (%) yes	No (%) yes	OR (95%CI)	OR (95%CI)
HD	2 (4.8)	0 (0.0)	∞ (0.00-∞)	∞ (0.00-∞)	0 (0.0)	0 (0.0)	NA	NA
Haematological malignancy	3 (7.1)	6 (3.7)	2.02 (0.48-8.44)	1.69 (0.36-7.87)	1 (2.0)	6 (3.7)	0.58 (0.07-4.96)	0.69 (0.08-6.29)
Young onset cancer	7 (16.7)	15 (9.2)	1.92 (0.73-5.08)	2.19 (0.80-6.05)	4 (8.2)	15 (9.2)	0.92 (0.29-2.94)	0.74 (0.20-2.73)
Haem/young onset cancer	9 (21.4)	17 (10.4)	2.30 (0.94-5.64)	2.41 (0.95-6.13)	5 (10.2)	17 (10.4)	1.04 (0.36-3.00)	0.87 (0.27-2.79)
Any cancer	23 (54.8)	69 (42.3)	1.63 (0.82-3.23)	2.02 (0.98-4.17)	20 (40.8)	69 (42.3)	0.87 (0.45-1.69)	0.87 (0.44-1.72)
Any Auto-immune disease	12 (28.5)	39 (23.7)	1.21 (0.79-1.86)	1.16 (0.74-1.81)	6 (11.7)	39 (23.7)	0.60 (0.33-1.10)	0.67 (0.39-1.15)

numbers and percentages for unadjusted analysis only

Results are presented unadjusted for number of first degree relatives as this adjustment did not result in statistically significantly better fitting models

8.4 Discussion:

The effects of many health variables on the risk of HD have been analysed on data from the SNEHD case-control study. In contrast to YHHCCS the effect of each variable has been analysed in the total series (age 16-74 years) as well as different age groups (16-34 and 50+ yrs). These two age groups were described by MacMahon (1966) (see Chapter 1). However, as with YHHCCS, few of the analyses have given rise to statistically significant results. In the total series and the analysis by age at diagnosis groups the results are unlikely to be due to a lack of positive responses. However, when looking at EBV status the number of cases with these data available were limited as when analysis began EBV status was not confirmed for all cases. EBV status is only available at present for 323 cases and therefore, the statistical power of EBV-status subgroup analyses is very limited. This applies particularly to EBV-status subgroups in young adults and older people respectively.

Infectious Mononucleosis:

IM has frequently been reported to be associated with an increased risk of HD (see Chapter 1). The SNEHD results support this relationship with a statistically significant increased risk of HD following IM in the total series whether IM was confirmed or suspected. This statistically significant effect appears to be limited to the young adult peak (age 16-34 yrs). This supports the results of YHHCCS which found an increased risk of HD in a subset of the young adult peak (age 16-24 yrs). The total number of HD cases with a history of IM and EBV status known in this study is very small. The increased risk of HD associated with IM was found in both EBV subgroups in the total series but this effect was driven by the results in the young adult peak. The odds ratios associated with EBV +ve HD were higher, especially in the young adult age group. However, none of the ORs were statistically significantly different from one another. When IM in a family member was analysed the differences between the age at diagnosis groups and the EBV status groups are even clearer. IM in a family member as a continuous variable is associated with a substantially increased risk of HD in the young adult peak, but not in the 50+ years age at

diagnosis group. The difference in the effect family IM is limited to EBV +ve HD, especially in the 16-34 years age at diagnosis group. In this age group the OR of EBV +ve vs EBV –ve HD is almost 4. This finding for family IM in the young adult peak is the opposite to that found for YHHCCS (where risk of EBV -ve HD was increased and EBV +ve HD was decreased) and much more plausible.

As for YHHCCS the role of IM is not completely limited to EBV +ve HD but the suggestion of a relationship is strongest in this group. Also the age at diagnosis subgroup analyses and comparisons lend support to hypothesis that IM is only associated with HD in the 16-34 years age group. The SNEHD results contrast with the results of a case series by Sleckman et al (1998). These authors found no association between IM and EBV status of HD. The broader age range (16-55 yrs) that Sleckman et al analysed may have masked any differences related to the effect of IM being limited to a specific EBV subgroup AND age group. The results of SNEHD and YHHCCS suggest a specific causal association of recent EBV exposure with EBV +ve HD and a weak positive association of EBV –ve HD with prior IM.

Infectious illness:

Analysis of the SNEHD study allowed the testing of hypotheses generated by YHHCCS e.g. more childhood infections are associated with a lower risk of HD in the young adult peak and the same finding for measles. The wide range of ages in the SNEHD study allowed the effect of these and other infectious illnesses to be analysed in the total series and in specific age groups (16-34 and 50+ yrs) and EBV subgroups. The statistically significant effect of childhood infectious illnesses observed in the YHHCCS analysis was not seen in the young adult peak (age 16-34 yrs) in SNEHD, nor in the 16-24 years age group (results not presented). There was some suggestion that more childhood infectious illnesses aged ≥5 years was associated with a decreased risk of HD age 16-34 years. Under the late-host-response model one would expect more late infections in this age group to increase risk.

However, there is no definition of what would constitute early or late infection. In the 50+ years age at diagnosis group more childhood infectious illnesses were associated with a decreased risk of HD, especially if they took place age 0-4. More infections (age ≥11 yrs) were associated with an increased risk. The effect of the combined herpes virus on increasing risk is limited to the 16-34 years age group. The effect of 'additional' infections led to a universal increase in risk of HD in the total series and in each age group. The problem of combining several infections, some of which have opposite effects, is seen when looking at the total infection variable in SNEHD. Most of these results are very close to the null. This could suggest that it is not the overall total number of infections that are associated with an increased risk of HD but rather it is the result of specific infections.

Sleckman et al (1998) took as a prior hypothesis that EBV +ve cases would show evidence of childhood experience conducive to late exposure to infectious agents but failed to find any supportive evidence. Interestingly the finding from YHHCCS that EBV +ve HD is associated with more childhood infectious illness aged ≥5 years and EBV –ve HD is associated with fewer infections at the same age is confirmed in the 16-34 years age group in SNEHD. The difference between the subgroups in this case was statistically significant. In SNEHD this effect is also seen when the childhood and 'additional' infections are combined and the total number of infections are analysed. Thus, the results from SNEHD and YHHCCS support the hypothesis that EBV +ve HD in the young adult peak should have evidence of late (interpreted as age ≥5 yrs) exposure to infectious agents. These are the only statistically significant differences between the EBV subgroups in SNEHD for infectious illness.

The analysis of individual infections in YHHCCS revealed very little about the effect of childhood infectious illness on the risk of HD apart from a protective role for measles. Measles' protective effect in the young adult peak was again seen in the SNEHD data but only if measles occurred age ≥5 years. However, when all the childhood infectious

illnesses were entered together into a multivariate regression model none of the infections had a statistically significant effect on risk of HD. In contrast to what was said above for total infections this suggests that no one infection has an individual effect on risk of HD but there may be a combined effect.

A great number of infectious illnesses have been analysed in SNEHD and with the age at diagnosis and EBV status subgroups this has resulted in a large number of regression models. Thus, it could be that the results that are statistically significant have occurred by chance and are false positives. However, a number of the hypotheses tested here were stated a priori e.g. the effect of childhood infectious illness, the effect of IM by age and EBV status. These are the results that should be considered most sound. Many of the results should be regarded as hypothesis generation for future research e.g. total infections, individual infections.

It seems likely that it is the combined number of infections that is more important than any specific infection in the impact of risk of HD. This is particularly suggested for EBV +ve HD where it is the late infections in total that increase risk while the single infection results are quite unimpressive. The only exception to this is IM whose significant effect remained (in the 16-34 years age group) after adjustment for the presence of all the childhood infections.

Removal of the Tonsils and the Appendix:

Removal of the tonsils and the appendix are two of the most analysed general health factors for their effect on the risk of HD (see Chapter 1). As with all published studies of HD that adjust for SES there is no effect of appendectomy on risk of HD in SNEHD. The analysis of SNEHD allowed a comparison of the effect of appendicitis and appendectomy. The ORs associated with each were identical and very close to the null. In agreement with published studies the results of SNEHD confirm a role for tonsillectomy in increasing risk of HD in the young adult peak. Four of six published studies have found the risk of HD in young adults to be increased following tonsillectomy after adjustment for SES. Of these,

three had statistically significant results (Vianna et al, 1971; Johnson et al, 1972; Vianna et al, 1974). The effect of tonsillectomy remained after adjustment for the five childhood infectious illnesses and IM. However, in agreement with YHHCCS when the analysis of tonsillectomy are limited to the 16-24 years age group in SNEHD there was no effect. The tonsils can act as filter barriers to infection (Vianna et al, 1971) and it appears that the removal of this barrier has the most effect in the young adult peak where the evidence of an infectious aetiology for HD is greatest.

Family Health Variables:

There is little evidence of an effect of family health factors on the risk of HD apart from those relating to IM and cancer in first degree relatives. Familial HD is usually found in the literature as a risk factor for HD and SNEHD concurs with this finding. All familial cancers and haematological malignancies alone had the same pattern of risk in the total series and both age at diagnosis groups. However, only the ORs for the total series were statistically significant. There was no statistically significant difference between the EBV subgroups when analysed for family cancer. Unfortunately due to the lack of completion of the histological review it was not possible to check the assertion by Ferraris et al (1997) and Mack et al (1995) that familial HD is more closely related to the NS subtype. In the analysis of YHHCCS the construction of a family history score (FHS) based on the age of the relative resulted in a significantly better fitting model than using the number of relatives alone. SNEHD analyses of cancer in first degree relatives were adjusted in the same way and also using a more comprehensive FHS. However, neither FHS resulted in better fitting models than those using the number of first degree relatives. In fact none of the adjustments for number of first degree relatives or either FHS resulted in statistically significantly better fitting models.

The analysis of the SNEHD case-control study has allowed the investigation of many risk factors for HD including some new hypotheses generated by the YHHCCS analysis. Some of the results seen in YHHCCS have been reproduced but others have not,

even when the subset of 16-24 years age at diagnosis cases and controls has been compared. The large size and age range of the SNEHD study has allowed the investigation of many variables in age at diagnosis groups and EBV subgroups. It appears that different risk factors act differently between the age at diagnosis groups of HD e.g. IM, infectious illness, tonsillectomy, and cancer in first degree relatives. These risk factors appear to have effects more concentrated in the young adult peak. There appear to be fewer differences associated with EBV status in the total age at diagnosis groups. This maybe because de novo EBV infection could be important in the paediatric group, a reactivation of latent infection, possibly as a result of decreasing T-cell immunity more likely in older cases (Jarrett et al, 1991). Young adult EBV +ve HD may be close in time to EBV exposure. The main differences in SNEHD associated with IM and total number of infections age ≥5 years in the 16-34 years age group. Other studies are needed to test these findings in other settings. The results of SNEHD and YHHCCS emphasise the need to include age group and EBV status in any study investigating risk factors for the development of HD.

9 Risk of HD associated with physical activity

9.1 Introduction:

There is a variety of evidence that suggests that HD risk is affected by variables associated with the hormonal and immune systems. The evidence of an effect on the hormonal system includes the differing incidence between men and women and the lower risk of women who have had children (see chapter 1). An impact on the immune system is inferred from the fact that people with HIV infection have a much greater risk of HD and, in the young adult peak, removal of the tonsils is also associated with increased risk (see chapter 1). Physical activity can have an effect on both the hormonal and immune systems. It appears that moderate levels of activity boost the immune system but prolonged activity at high intensities suppresses immune function (discussed in detail chapter 3). Physical activity changes the concentration of various hormones including oestrogen and testosterone (discussed in detail chapter 3). It seemed appropriate to include in SNEHD an investigation of the hypothesis that physical activity would impact on the risk of HD; if so the probable biological pathway would be via the hormonal and immune systems. It was hypothesised that the effect would be that those people who exercised moderately would have a lower risk of HD while those who were very active would be at increased risk. However, very little work has been published on the risk of HD associated with physical activity. Only two studies have appeared in the literature (Whittemore et al, 1985; Paffenbarger et al, 1987). Both of these studies found a non-significantly decreased risk of HD associated with greater physical activity. This chapter will investigate the hypothesis outlined above on data from the SNEHD study in total and also in subgroups based on age, EBV status and sex.

9.2 Methods:

A full description of the SNEHD study is given in Chapter 4.

Interview data:

Trained interviewers conducted face-to-face interviews. The period of interest for the total interview was from birth up to the date of diagnosis for each case and a randomly allocated date of pseudo-diagnosis for controls. Information was requested on many health variables but this chapter will concentrate on physical activity.

Questionnaires:

The actual pages used for eliciting information on physical activity are in Appendix C. It can be seen that the period of interest was from age 7 to 39 years. It was decided not to ask about lifetime activity as many of the SNEHD participants were over 50 years of age and the level of detail for each activity was quite great resulting in recall problems. Splitting the questions into three sections allowed easier directions to be given to the respondents. For physical activity at school ages normal timetabled physical education was not included as it was assumed that all children would have approximately the same levels of this activity. Within each section participants could report up to 5 activities. For each activity the participant had to report the number of hours per week they took part in the activity in that age group. Respondents were guided when calculating these hours of participation. The maximum number of hours averaged over a three-month period in the age range was recorded. Hours were recorded to the nearest half-hour (rounding down). Also the proportion of the year the activity took place was elicited: habitually (>8 months per year), seasonally (4-8 months per year), or occasionally (<4 months per year). If the respondent had not reached any of the age groups on a page then these ages were considered to be 'not applicable'. However, if the index was one or more years into that age range the activity was recorded and used for analysis.

In addition questions were asked about more recent physical activity, in the last 5 years before diagnosis/ pseudo-diagnosis. These were much more simple questions and

asked people to rate their own activity and also a sweating question. Using these different methods had the added benefit of allowing a comparison of physical activity scores based on in-depth questions about activities and these more simple questions based on self-assessment and sweat episodes in cases and controls aged less than 40 years.

Categories of interest:

As mentioned in Chapter 2 there are a variety of methods of measuring physical activity in questionnaires. There are also a great number of ways of analysing the data once it has been collected. It was decided to compare four methods that have been found from the literature on the SNEHD data to see if the different methods give rise to different results. Three of the methods involve the use of MET (metabolic equivalents) scores. The MET scores used in this analysis come from the work of Ainsworth et al (1993) and this compendium is frequently cited by other authors in this field. The use of MET scores has the advantage of allowing activity intensity to be investigated. In this study MET scores of 4.5-5.5 were considered to be moderate activity and those of over 6 to be vigorous (White et al, 1996). The other physical activity score was based on time.

In general energy expended per week is calculated by multiplying the MET scores of an activity by the weight in kilograms of the person by the number of hours the activity is performed and summing across activities. However, in SNEHD the subject's body weight was not known. This is not a problem, however, as methods that take account of body weight can produce errors because estimates of energy expenditure may more closely reflect body weight rather than activity levels (Ainsworth et al, 1993).

A common calculation for all the methods was to make some allowance for the amount of time during the year that the activity took place e.g. habitual, seasonal, occasional. The number of hours spent in habitual activities (>8 months of the year) was multiplied by 10/12, hours spent in seasonal activities (4-8 months of the year) was multiplied by 6/12, and the hours spent in occasional activities (1-4 months of the year) by 2/12.

Method 1: following White et al (1996). Activities were divided into three groups: low, medium, and high intensity based on their MET score (<4.5, 4.5-5.5, ≥6 respectively). These three groups were then allocated kilocalorie scores of 4.5, 5, and 6 kilocalories per minute. These values were than multiplied by the number of minutes spent on each activity in each age group. All activities were than summed within each age group. The continuous number of kilocalories was then split into those people with no activity in an age group and quartiles of activity based only on those subjects who had performed any activity. This method of division was used for each age group.

Method 2: following Min-Lee et al (1992). Sports of <4.5 METS were rated at 5 kcal/min while those of ≥4.5 METS were rated at 10 kcal/min. The number of hours spent in physical activity were summed for each age group and split as for method 1.

Method 3: The most frequently used method used to score physical activities at present is to use MET hours/ week. These are calculated by multiplying the time (in hours) spent on each activity by its MET score. The number of MET hours spent on each activity are summarised within each age group and summed for each age group and split as for method 1.

Method 4: It was also decided to use a method of summation that took no account of the intensity of the activity. The number of hours spent in physical activity were summed for each age group and split as for method 1.

The activities from the three pages of the questionnaire were combined into the six age groups. This involved the addition of the school activities age 7-11, 12-14, 15-leaving from page 1 to the 'other' activities in the same age groups from page 3. The competitive activities from page 2 age leave-21, 22-29, and 30-39 years were added to the 'other' activities from page 3 in the same age groups. For the total series and the 16-34 years age at diagnosis group details for the school period of activity (ages 7-11, 12-14, 15-leave school) only were used to prevent the dropping of data in those subjects who had not reached the older ages (leave school-21, 22-29, 30-39 yrs). In contrast all the activity ages were used for

the 50+ years age at diagnosis group. The odds ratios presented for each period of activity are adjusted for the other activity periods used for that age group i.e. in the total series activity age 7-11 years is adjusted for activity at ages 12-14 and 15-leave school.

The effect of consistent activity (top quartile in each age group (7-11, 12-14, 15-leave school, leave school-21, 22-29, and 30-39 yrs)) or inactivity (no activity in each age group) was also investigated. As for the activity levels only school activity was used in the total series and the 16-34 yrs age group and all 6 age groups for the 50+ yrs age group. However, no one remained in the top quartile for the six activity periods of interest and therefore, it was only possible to investigate the effects of long-term inactivity in the 50+ yrs age group.

The answers to the questions from page 4 were analysed separately with a three level variable of activity levels in the five years prior to interview: little or none, moderate, a lot. A four level sweat episode variable was created from the answers to question 2: no sweating, less often, weekly, two or more times per week.

Potential Confounders:

A large number of confounders were considered in the analysis of SNEHD health variables in Chapter 8. However, only age, sex, and Carstair's index were used in the analysis of physical activity. This was because the other potential confounders were found to have very little effect in the previous chapter and there is no evidence in the literature that they are associated with level of physical activity.

EBV Status and Histopathology:

See chapter 8.

Statistical Methods:

Logistic regression was used for all analyses on SAS 6.11 using PROC LOGISTIC. The results are reported as Odds Ratios (OR) with 95% confidence intervals (95% CI). Analyses were applied to the total series of subjects, the series split by age at presentation (16-34 and 50+ years), EBV status, and sex.

9.3 Results:

The questionnaire was written to elicit as much detail about physical activities as possible. However, only around half of all those in each age group performed any activity at all (Table 9.1). The greatest proportion of active people was seen in the 12-14 and leaving school-21 years age groups. The observations for total activity apply as well to vigorous activity but even fewer people performed any moderate activity (20% or less, results not shown). The comparison of activity levels by case/control status are in Figure 9.1. The activity levels by status are very similar in all age groups but there is slight evidence of more controls performing no activity and more cases in the top quartile of activity. As with the results for any activity vs. none the results for vigorous activity are very similar to those for total activity. The numbers of people in the moderate activity are much smaller and very variable (results not shown).

There are large differences in the activity levels of male and female subjects. There are a higher number of females with no activity in all of the age groups of interest and a higher number of males in the top quartile of activity.

It has been suggested that there is a greater stability in the reporting of extremes of activity (see Chapter 2). This observation applies to SNEHD data as can be seen from Table 9.2. In every age group the highest proportion of activity reported is in the no activity and the top quartile of activity groups. There is also evidence that the activity reported in one age group is highly correlated with that in all other age groups (Table 9.3).

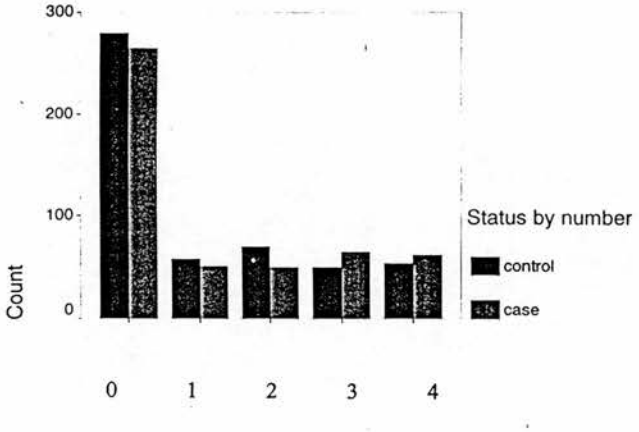
The two single questions on page 4 of the sports questionnaire were correlated together to see if the answers to each were comparable. The answers to these questions were highly correlated ($p<0.01$). Thus, two questions that appear to measure different things are providing very similar information.

The results of regression models for activity in the last 5 years are in Table 9.4. There are no significant trends in the level of activity in the total series or the 16-34 and 50+ yrs age

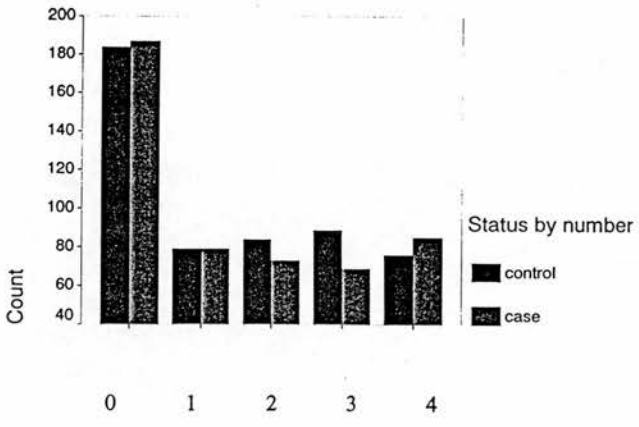
Table 9.1: Any vs. no activity at each age

Age Group	Any Activity (%)	No Activity (%)	Total
Age 7-11	545 (54.2)	460 (45.8)	1005
Age 12-14	371 (36.9)	634 (63.1)	1005
Age 15-leave school	536 (53.3)	469 (46.7)	1005
Age leave school-21	467	499	966
Age 22-29	448	461	909
Age 30-39	385	298	683

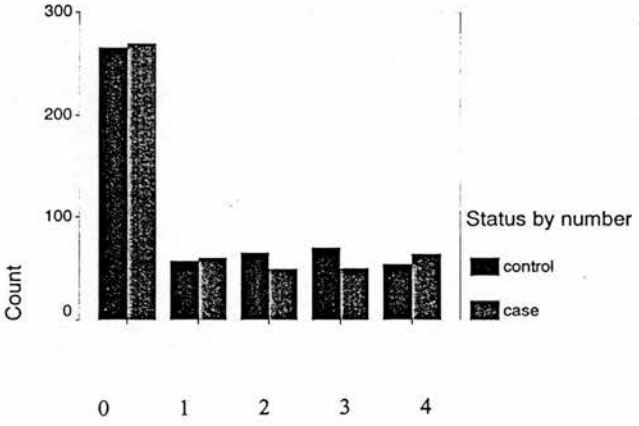
Figure 9.1. Bar chart comparison of case/control activity levels.



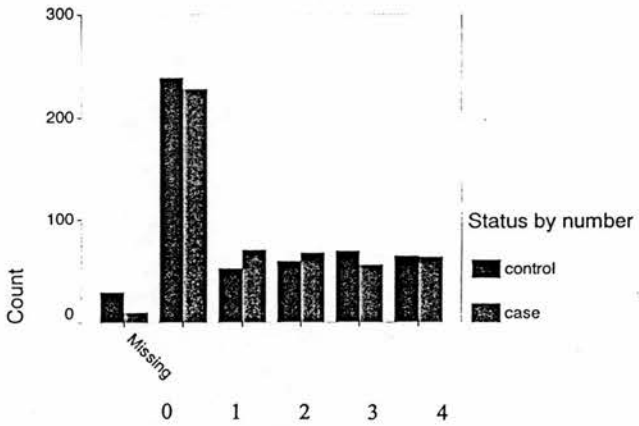
MET hours per week in levels age 7-11



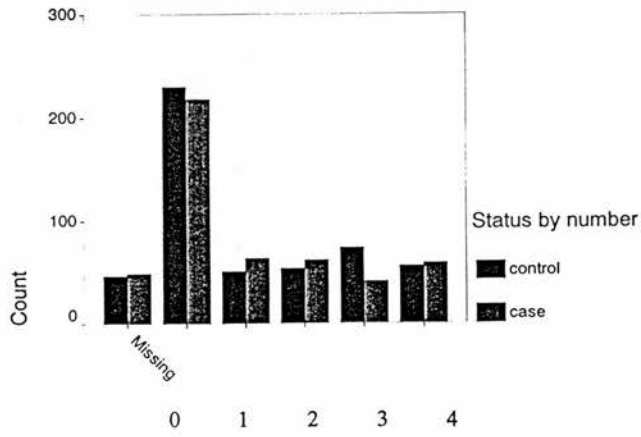
MET hours per week in levels age 12-14



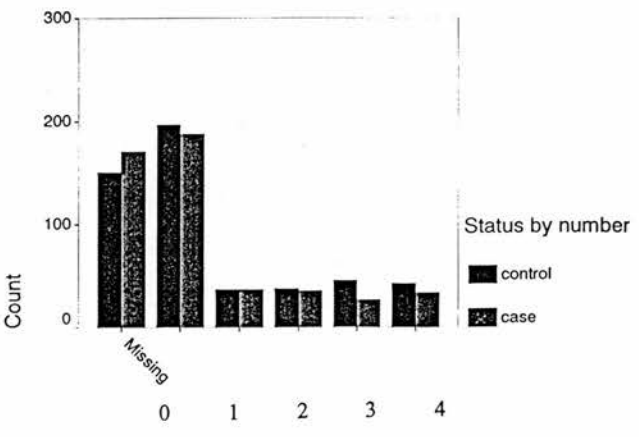
MET hours per week in levels age 15-leave school



MET hours per week in levels age leave school-21



MET hours per week in levels age 22-29



MET hours per week in levels age 30-39

Table 9.2: Between individual consistency of activity recording across age groups

		MET hours per week in levels age 12-14					
MET hours per week in levels age 7-11		0	1	2	3	4	Total
	0	326 (59.8)					545
	1		35 (32.1)				109
	2			51 (42.5)			120
	3				53 (46.1)		115
	4					97 (83.6)	116
		MET hours per week in levels age 15-leave school					
MET hours per week in levels age 12-141		0	1	2	3	4	Total
	0	347 (93.5)					371
	1		56 (35.4)				158
	2			59 (37.6)			157
	3				66 (41.8)		158
	4					92 (57.1)	161
		MET hours per week in levels age leave school-21					
MET hours per week in levels age 15-leave school		0	1	2	3	4	Total
	0	357 (68.3)					523
	1		37 (32.0)				112
	2			27 (25.5)			106
	3				40 (34.8)		115
	4					55 (50.0)	110
		MET hours per week in levels age 22-29					
MET hours per week in levels age leave school-21		0	1	2	3	4	Total
	0	349 (77.9)					448
	1		39 (35.5)				110
	2			42 (36.5)			115
	3				54 (44.6)		121
	4					70 (60.9)	115
		MET hours per week in levels age 30-39					
MET hours per week in levels age 22-29		0	1	2	3	4	Total
	0	318 (87.4)					364
	1		30 (35.3)				85
	2			30 (37.5)			80
	3				31 (40.8)		76
	4					49 (62.8)	78

Table 9.3: Correlation of Activity in each age group with all other age groups

		MET hours per week age 7-11	MET hours per week age 12-14	MET hours per week age 15-leave	MET hours per week age leave-21	MET hours per week age 22-29	MET hours per week age 30-39
MET hours per week age 7-11	Correlation Sig (2 tailed) N	1.00 NA 1005	0.68 p<0.0001 1005	0.49 p<0.0001 1005	0.34 p<0.0001 966	0.27 p<0.0001 909	0.16 p<0.0001 683
MET hours per week age 12-14	Correlation Sig (2 tailed) N	0.68 p<0.0001 1005	1.00 NA 1005	0.71 p<0.0001 1005	0.45 p<0.0001 966	0.35 p<0.0001 909	0.26 p<0.0001 683
MET hours per week age 15-leave school	Correlation Sig (2 tailed) N	0.49 p<0.0001 1005	0.71 p<0.0001 1005	1.00 NA 1005	0.49 p<0.0001 966	0.39 p<0.0001 909	0.29 p<0.0001 683
MET hours per week leave school-21	Correlation Sig (2 tailed) N	0.34 p<0.0001 966	0.45 p<0.0001 966	0.49 p<0.0001 966	1.00 NA 966	0.62 p<0.0001 909	0.46 p<0.0001 683
MET hours per week age 22-29	Correlation Sig (2 tailed) N	0.27 p<0.0001 909	0.35 p<0.0001 909	0.39 p<0.0001 909	0.62 p<0.0001 909	1.00 NA 909	0.73 p<0.0001 683
MET hours per week age 30-39	Correlation Sig (2 tailed) N	0.16 p<0.0001 683	0.26 p<0.0001 683	0.29 p<0.0001 683	0.46 p<0.0001 683	0.73 p<0.0001 683	1.00 NA 683

Table 9.4: Results for activity in the 5 years to diagnosis/ pseudo-diagnosis

	Case		Control			
Risk Factor	No	No (%) yes	No	No (%) yes	OR (95%CI)[1]	OR (95%CI)[2]
Total Series						
Trend	493	297 (60.2)	512	323 (63.1)	1.00 (0.84-1.19)	1.03 (0.86-1.24)
Levels moderate amount vs. little or none	493	206 (41.8)	512	247 (48.2)	0.79 (0.60-1.04)	0.84 (0.63-1.12)
a lot vs. little or none		91 (18.5)		76 (14.8)	1.12 (0.77-1.61)	1.18 (0.80-1.73)
Age 16-34						
Trend	241	160 (66.4)	220	151 (68.6)	1.08 (0.84-1.40)	1.10 (0.84-1.44)
Levels moderate amount vs. little or none	241	100 (41.5)	220	112 (50.9)	0.75 (0.49-1.15)	0.81 (0.52-1.25)
a lot vs. little or none		60 (24.9)		39 (17.7)	1.28 (0.76-2.17)	1.31 (0.76-2.26)
Age 50+						
Trend	148	81 (54.7)	165	100 (60.6)	0.90 (0.65-1.23)	0.94 (0.67-1.31)
Levels moderate amount vs. little or none	148	59 (39.9)	165	72 (43.6)	0.86 (0.52-1.41)	0.87 (0.52-1.48)
a lot vs. little or none		22 14.9)		28 (17.0)	0.82 (0.52-1.59)	0.92 (0.46-1.83)

1. Adjusted for age and sex
2. Adjusted for age, sex, and the Carstairs index

groups. When divided into levels subjects with a moderate level of activity had a slightly lower risk of HD and those in the highest activity category had a slightly increased risk. This pattern is seen in the total series and the 16-34 years age group but none of the results are statistically significant and therefore are not shown. The results for sweat episodes per week in the 5 years to diagnosis/ pseudo-diagnosis are in Table 9.5. Again there is no statistically significant trend of risk in the total series or the age groups and all the odds ratios are very close to the null. The results by level make it clear that those who work up a sweat less than weekly have a lower risk of HD than those who do so more frequently (significantly so in the total series and the 50+ years age group). The pattern of risk across levels for this variable is j-shaped with a lower risk of HD for those who performed some physical activity vs. none.

The results for MET hours in the total series and the 16-34 and 50+ years age groups are in Table 9.6. The trends of activity in the total series and the age groups are all very close to the null. When the activity descriptions are reduced to dichotomies the results are again very close to the null with the exception of a statistically significant increased risk of HD age 50+ yrs for any activity age 7-11 (Table 9.7).

The results of activity presented separately for males and females are in Tables 9.8-9.11 for the total series and the 16-34 and 50+ years age groups respectively. Very few results are statistically significant and the pattern of risk is similar for males and females.

The effect of physical activity was compared in EBV +ve and EBV –ve HD in the total series and by age group. There were no differences between the effect of activity in the EBV subgroups (Tables 9.12-9.15). Very few of the results were statistically significant and all the confidence intervals overlap.

The numbers of subjects who were consistently active (or inactive) for the total series and by age group were quite limited. The results of the constant activity regression models are in Table 9.16 for the total series. Inactivity all through school is significantly more associated with HD, while there is little effect of being in the top quartile of activity over the same time

Table 9.5: Results for sweat episodes in the 5 years to diagnosis/ pseudo-diagnosis

	Case		Control			
Risk Factor	No	No (%) yes	No	No (%) yes	OR (95%CI)[1]	OR (95%CI)[2]
Total Series						
Trend	493	260 (52.7)	512	288 (56.3)	0.95 (0.86-1.04)	0.97 (0.88-1.07)
Levels less than weekly vs. none	493	11 (2.2)	512	27 (5.3)	0.36 (0.17-0.74)	0.36 (0.17-0.76)
weekly vs. none		53 (10.8)		55 (10.7)	0.83 (0.54-1.28)	0.95 (0.60-1.50)
at least twice a week vs. none		196 (39.8)		206 (40.2)	0.83 (0.63-1.10)	0.87 (0.65-1.17)
Age 16-34						
Trend	241	161 (66.8)	220	147 (66.8)	1.01 (0.88-1.16)	1.03 (0.89-1.19)
Levels less than weekly vs. none	241	7 (2.9)	220	10 (4.5)	0.62 (0.22-1.73)	0.66 (0.23-1.84)
weekly vs. none		32 (13.3)		28 (12.7)	1.02 (0.56-1.87)	1.29 (0.68-2.47)
at least twice a week vs. none		122 (50.6)		109 (49.5)	0.99 (0.65-1.52)	1.05 (0.67-1.62)
Age 50+						
Trend	148	51 (34.5)	165	160 (97.0)	0.88 (0.75-1.04)	0.88 (0.74-1.05)
Levels less than weekly vs. none	148	2 (1.4)	165	10 (6.1)	0.32 (0.06-1.73)	0.32 (0.06-1.48)
weekly vs. none		7 (4.7)		55 (33.3)	0.58 (0.21-1.63)	0.55 (0.18-1.62)
at least twice a week vs. none		42 (28.4)		95 (57.6)	0.70 (0.42-1.15)	0.70 (0.41-1.18)

1. Adjusted for age and sex
2. Adjusted for age, sex, and the Carstairs index

Table 9.6: Results for trend of total MET hours of activity in the age groups under consideration

	Case		Control			
Risk Factor	No	No (%) yes	No	No (%) yes	OR (95%CI)[1]	OR (95%CI)[2]
Total Series						
Trend MET hours age 7-11	493	228 (46.2)	512	232 (45.3)	1.10 (0.98-1.24)	1.13 (1.00-1.28)
Trend MET hours age 12-14	493	306 (62.1)	512	328 (64.1)	0.94 (0.82-1.09)	0.90 (0.78-1.05)
Trend MET hours age 15-leave school	493	223 (45.2)	512	246 (48.0)	0.95 (0.84-1.08)	1.01 (0.89-1.15)
Age 16-34						
Trend MET hours age 7-11	241	127 (52.5)	220	132 (60.0)	0.97 (0.82-1.14)	1.00 (0.84-1.18)
Trend MET hours age 12-14	241	157 (65.1)	220	153 (69.5)	0.97 (0.80-1.18)	0.92 (0.74-1.13)
Trend MET hours age 15-leave school	241	130 (53.9)	220	126 (57.3)	0.98 (0.82-1.18)	1.06 (0.87-1.29)
Age 50+						
Trend MET hours age 7-11	148	59 (39.9)	165	46 (27.9)	1.52 (1.17-1.96)	1.57 (1.19-2.05)
Trend MET hours age 12-14	148	82 (55.4)	165	89 (53.9)	0.76 (0.56-1.02)	0.73 (0.54-0.99)
Trend MET hours age 15-leave school	148	46 (31.1)	165	48 (29.1)	1.09 (0.86-1.40)	1.15 (0.88-1.49)
Trend MET hours age leave school-21	148	66 (44.6)	165	72 (43.6)	0.97 (0.77-1.21)	0.94 (0.75-1.19)
Trend MET hours age 22-29	148	65 (43.9)	165	70 (42.4)	1.02 (0.76-1.37)	1.00 (0.73-1.37)
Trend MET hours age 30-39	148	54 (36.5)	165	70 (42.4)	0.90 (0.71-1.15)	0.93 (0.72-1.21)

1. Adjusted for age and sex
2. Adjusted for age, sex, and the Carstairs index

Table 9.7: Results for any activity dichotomy in the age groups under consideration

	Case		Control			
Risk Factor	No	No (%) yes	No	No (%) yes	OR (95%CI)[1]	OR (95%CI)[2]
Total Series						
Any activity yes vs. no age 7-11	493	228 (46.2)	512	232 (45.3)	1.07 (0.79-1.44)	1.20 (0.88-1.65)
Any activity yes vs. no age 12-14	493	306 (62.1)	512	328 (64.1)	0.96 (0.68-1.37)	0.83 (0.57-1.20)
Any activity yes vs. no age 15-leave school	493	223 (45.2)	512	246 (48.0)	0.82 (0.59-1.14)	0.93 (0.66-1.31)
Age 16-34						
Any activity yes vs. no age 7-11	241	127 (52.5)	220	132 (60.0)	0.73 (0.46-1.13)	0.79 (0.50-1.26)
Any activity yes vs. no age 12-14	241	157 (65.1)	220	153 (69.5)	0.95 (0.54-1.64)	0.85 (0.47-1.51)
Any activity yes vs. no age 15-leave school	241	130 (53.9)	220	126 (57.3)	1.02 (0.61-1.71)	1.09 (0.64-1.87)
Age 50+						
Any activity yes vs. no age 7-11	148	59 (39.9)	165	46 (27.9)	2.33 (1.27-4.26)	2.67 (1.47-5.03)
Any activity yes vs. no age 12-14	148	82 (55.4)	165	89 (53.9)	0.69 (0.37-1.29)	0.59 (0.31-1.14)
Any activity yes vs. no age 15-leave school	148	46 (31.1)	165	48 (29.1)	1.06 (0.56-1.99)	1.20 (0.61-2.36)
Any activity yes vs. no age leave school-21	148	66 (44.6)	165	72 (43.6)	0.79 (0.40-1.56)	0.81 (0.40-1.64)
Any activity yes vs. no age 22-29	148	65 (43.9)	165	70 (42.4)	1.63 (0.78-3.42)	1.42 (0.65-3.11)
Any activity yes vs. no age 30-39	148	54 (36.5)	165	70 (42.4)	0.53 (0.28-1.00)	0.60 (0.30-1.18)

1. Adjusted for age and sex
2. Adjusted for age, sex, and the Carstairs index

Table 9.8: Results for activity in the 5 years to diagnosis/ pseudo-diagnosis by sex

	Male				Female			
	Case	Control			Case	Control		
Risk factor	No (%) yes	No (%) yes	OR (95%CI)[1]	OR (95%CI)[2]	No (%) yes	No (%) yes	OR (95%CI)[1]	OR (95%CI)[2]
Total Series								
Trend	179 (63.7)	190 (65.5)	1.09 (0.87-1.37)	1.13 (0.88-1.43)	118 (55.7)	133 (59.9)	0.87 (0.65-1.15)	0.88 (0.66-1.18)
Moderate amount vs. little or none	112 (39.9)	144 (49.7)	0.74 (0.51-1.08)	0.83 (0.56-1.24)	94 (44.3)	103 (46.4)	0.86 (0.58-1.29)	0.87 (0.57-1.31)
A lot vs. little or none	67 (23.8)	46 (15.9)	1.34 (0.84-2.15)	1.39 (0.85-2.28)	24 (11.3)	30 (13.5)	0.75 (0.41-1.39)	0.79 (0.42-1.50)
Age 16-34								
Trend	97 (74.0)	85 (73.3)	1.26 (0.89-1.78)	1.24 (0.86-1.78)	63 (57.3)	66 (63.5)	0.92 (0.62-1.37)	0.95 (0.63-1.43)
Moderate amount vs. little or none	53 (40.5)	61 (52.6)	0.77 (0.42-1.43)	0.87 (0.46-1.67)	47 (42.7)	51 (49.0)	0.77 (0.43-1.39)	0.81 (0.44-1.48)
A lot vs. little or none	44 (33.6)	24 (20.7)	1.59 (0.79-3.21)	1.55 (0.74-3.22)	16 (14.5)	15 (14.4)	0.95 (0.41-2.19)	1.01 (0.43-2.38)
Age 50+								
Trend	50 (58.1)	57 (60.0)	1.20 (0.79-1.83)	1.29 (0.82-2.01)	31 (50.0)	43 (61.4)	0.61 (0.36-1.03)	0.65 (0.38-1.13)
Moderate amount vs. little or none	33 (38.4)	41 (43.2)	1.18 (0.58-2.37)	1.28 (0.61-2.71)	26 (41.9)	31 (44.3)	0.70 (0.34-1.47)	0.69 (0.32-1.51)
A lot vs. little or none	17 (19.8)	16 (16.8)	1.45 (0.61-3.42)	1.66 (0.69-4.11)	5 (8.1)	12 (17.1)	0.33 (0.10-1.07)	0.40 (0.12-1.37)

1. Adjusted for age and sex
2. Adjusted for age, sex, and the Carstairs index

Table 9.9: Results for sweat episodes in the 5 years to diagnosis/ pseudo-diagnosis by sex

	Male				Female			
	Case	Control			Case	Control		
Risk factor	No (%) yes	No (%) yes	OR (95%CI)[1]	OR (95%CI)[2]	No (%) yes	No (%) yes	OR (95%CI)[1]	OR (95%CI)[2]
Total Series								
Trend	167 (59.4)	182 (62.8)	0.94 (0.83-1.07)	0.97 (0.86-1.11)	93 (43.9)	106 (47.7)	0.95 (0.83-1.10)	0.96 (0.83-1.11)
Less than weekly vs. none	7 (2.5)	17 (5.9)	0.34 (0.13-0.97)	0.36 (0.13-0.97)	4 (1.9)	10 (4.5)	0.38 (0.11-1.24)	0.40 (0.12-1.31)
Weekly vs. none	28 (10.0)	31 (10.7)	0.73 (0.41-1.32)	1.03 (0.54-1.95)	25 (11.8)	24 (10.8)	0.98 (0.52-1.83)	0.92 (0.48-1.78)
At least twice a week vs. none	132 (47.0)	134 (46.2)	0.80 (0.55-1.16)	0.87 (0.59-1.29)	64 (30.2)	72 (32.4)	0.84 (0.55-1.30)	0.87 (0.55-1.36)
Age 16-34								
Trend	101 (77.1)	90 (77.6)	1.03 (0.84-1.26)	1.05 (0.85-1.30)	60 (54.5)	57 (54.8)	0.99 (0.82-1.21)	1.01 (0.83-1.24)
Less than weekly vs. none	4 (3.1)	7 (6.0)	0.49 (0.13-1.89)	0.58 (0.14-2.29)	3 (2.7)	3 (2.9)	1.02 (0.19-5.36)	1.11 (0.21-5.90)
Weekly vs. none	15 (11.5)	15 (12.9)	0.87 (0.36-2.12)	1.38 (0.49-3.85)	17 (15.5)	13 (12.5)	1.29 (0.56-2.95)	1.26 (0.54-2.93)
At least twice a week vs. none	82 (62.6)	68 (58.6)	1.00 (0.54-1.87)	1.10 (0.58-2.12)	40 (36.4)	41 (39.4)	0.95 (0.52-1.72)	1.01 (0.55-1.85)
Age 50+								
Trend	36 (41.9)	47 (49.5)	0.89 (0.72-1.10)	0.91 (0.73-1.13)	15 (24.2)	23 (32.9)	0.89 (0.67-1.17)	0.85 (0.64-1.15)
Less than weekly vs. none	1 (1.2)	4 (4.2)	0.17 (0.02-1.66)	0.17 (0.02-1.70)	1 (1.6)	1 (1.4)	1.19 (0.07-20.50)	1.20 (0.07-20.73)
Weekly vs. none	5 (5.8)	5 (5.3)	0.80 (0.21-3.09)	1.03 (0.25-4.28)	2 (3.2)	5 (7.1)	0.46 (0.08-2.57)	0.22 (0.02-2.00)
At least twice a week vs. none	30 (34.9)	38 (40.0)	0.69 (0.36-1.31)	0.72 (0.37-1.40)	12 (19.4)	17 (24.3)	0.74 (0.31-1.73)	0.69 (0.29-1.69)

1. Adjusted for age and sex
2. Adjusted for age, sex, and the Carstairs index

Table 9.10: Results for trend of total MET hours of activity in the age groups under consideration by sex

	Male				Female			
	Case	Control			Case	Control		
Risk factor	No (%) yes	No (%) yes	OR (95%CI)[1]	OR (95%CI)[2]	No (%) yes	No (%) yes	OR (95%CI)[1]	OR (95%CI)[2]
Total Series								
Trend MET hours age 7-11	154 (54.8)	151 (52.1)	1.15 90.99-1.34)	1.16 (0.99-1.37)	74 (34.9)	81 (36.5)	1.00 (0.82-1.22)	1.05 (0.86-1.28)
Trend MET hours age 12-14	188 (66.9)	199 (68.6)	0.95 (0.79-1.14)	0.92 (0.75-1.11)	118 (55.7)	129 (58.1)	0.93 (0.74-1.16)	0.89 (0.71-1.14)
Trend MET hours age 15-lve school	144 (51.2)	162 (55.9)	0.90 (0.78-1.05)	0.97 (0.83-1.14)	79 (37.3)	84 (37.8)	1.06 (0.85-1.33)	1.08 (0.85-1.38)
Age 16-34								
Trend MET hours age 7-11	80 (61.1)	79 (68.1)	0.97 (0.78-1.20)	1.01 (0.80-1.28)	47 (42.7)	53 (51.0)	0.95 (0.74-1.22)	0.96 (0.75-1.24)
Trend MET hours age 12-14	92 (70.2)	86 (74.1)	1.02 (0.78-1.35)	0.92 (0.68-1.25)	65 (59.1)	67 (64.4)	0.90 (0.67-1.22)	0.90 (0.66-1.22)
Trend MET hours age 15-lve school	84 (64.1)	76 (65.5)	0.97 (0.77-1.23)	1.09 (0.85-1.41)	46 (41.8)	50 (48.1)	0.99 (0.73-1.35)	1.00 (0.73-1.39)
Age 50+								
Trend MET hours age 7-11	41 (47.7)	33 (34.7)	1.51 (1.10-2.07)	1.45 (1.05-2.00)	18 (29.0)	13 (18.6)	1.63 (0.97-2.75)	1.95 (1.09-3.48)
Trend MET hours age 12-14	53 (61.6)	57 (60.0)	0.76 (0.52-1.12)	0.79 (0.53-1.18)	29 (46.8)	32 (45.7)	0.74 (0.44-1.24)	0.64 (0.36-1.14)
Trend MET hours age 15-lve school	31 (36.0)	35 (36.8)	1.02 (0.77-1.36)	1.04 (0.77-1.40)	15 (24.2)	13 (18.6)	1.27 (0.75-2.14)	1.41 (0.78-2.52)
Trend MET hours age lve school-21	47 (54.7)	55 (57.9)	0.86 (0.66-1.13)	0.86 (0.65-1.13)	19 (30.6)	17 (24.3)	1.30 (0.82-2.08)	1.23 (0.73-2.03)
Trend MET hours age 22-29	46 (53.5)	49 (51.6)	1.16 (0.81-1.64)	1.14 (0.79-1.65)	19 (30.6)	21 (30.0)	0.79 (0.42-1.50)	0.72 (0.36-1.46)
Trend MET hours age 30-39	36 (41.9)	46 (48.4)	0.93 (0.70-1.24)	0.95 (0.70-1.29)	18 (29.0)	24 (34.3)	0.82 (0.48-1.40)	0.93 (0.53-1.61)

1. Adjusted for age and sex
2. Adjusted for age, sex, and the Carstairs index

Table 9.11: Results for any activity dichotomy in the age groups under consideration by sex

	Male				Female			
	Case	Control			Case	Control		
Risk factor	No (%) yes	No (%) yes	OR (95%CI)[1]	OR (95%CI)[2]	No (%) yes	No (%) yes	OR (95%CI)[1]	OR (95%CI)[2]
Total Series								
Any activity yes vs. no age 7-11	154 (54.8)	151 (52.1)	0.66 (0.35-1.24)	0.75 (0.38-1.48)	74 (34.9)	81 (36.5)	0.81 (0.42-1.54)	0.81 (0.42-1.57)
Any activity yes vs. no age 12-14	188 (66.9)	199 (68.6)	0.96 (0.40-2.27)	0.73 (0.28-1.87)	118 (55.7)	129 (58.1)	0.90 (0.44-1.87)	0.90 (0.43-1.90)
Any activity y vs. n age 15-l school	144 (51.2)	162 (55.9)	1.14 (0.54-2.44)	1.38 (0.61-3.14)	79 (37.3)	84 (37.8)	0.95 (0.46-1.97)	0.96 (0.46-2.04)
Age 16-34								
Any activity yes vs. no age 7-11	80 (61.1)	79 (68.1)	1.14 (0.76-1.71)	1.22 (0.79-1.89)	47 (42.7)	53 (51.0)	0.94 (0.60-1.49)	1.08 (0.67-1.74)
Any activity yes vs. no age 12-14	92 (72.2)	86 (74.1)	1.04 (0.63-1.74)	0.89 (0.52-1.54)	65 (59.1)	67 (64.4)	0.89 (0.54-1.46)	0.81 (0.48-1.36)
Any activity y vs. n age 15-l school	84 (64.1)	76 (65.5)	0.68 (0.44-1.05)	0.82 (0.52-1.31)	46 (41.8)	50 (48.1)	1.05 (0.64-1.74)	1.08 (0.64-1.83)
Age 50+								
Any activity yes vs. no age 7-11	41 (47.7)	33 (34.7)	2.38 (1.07-5.29)	2.28 (1.00-5.20)	18 (29.0)	13 (18.6)	2.00 (0.73-5.51)	3.12 (1.04-9.39)
Any activity yes vs. no age 12-14	53 (61.6)	57 (60.0)	0.69 (0.27-1.77)	0.67 (0.25-1.77)	29 (46.8)	32 (45.7)	0.69 (0.29-1.68)	0.56 (0.22-1.42)
Any activity y vs. n age 15-l school	31 (36.0)	35 (36.8)	0.84 (0.37-1.89)	0.92 (0.39-2.16)	15 (24.2)	13 (18.6)	1.75 (0.59-5.16)	1.94 (0.59-6.41)
Any activity y vs. n age l school-21	47 (54.7)	55 (57.9)	0.62 (0.26-1.48)	0.70 (0.29-1.72)	19 (30.6)	17 (24.3)	1.30 (0.36-4.35)	1.06 (0.27-4.21)
Any activity yes vs. no age 22-29	46 (53.5)	49 (51.6)	1.92 (0.77-4.78)	1.66 (0.63-4.32)	19 (30.6)	21 (30.0)	1.08 (0.27-4.38)	0.95 (0.20-4.40)
Any activity yes vs. no age 30-39	36 (41.9)	46 (48.4)	0.59 (0.26-1.34)	0.66 (0.27-1.58)	18 (29.0)	24 (34.3)	0.57 (0.19-4.89)	0.77 (0.24-2.47)

1. Adjusted for age and sex
2. Adjusted for age, sex, and the Carstairs index

Table 9.12: Results for activity in the 5 years to diagnosis/ pseudo-diagnosis by EBV status

	EBV +ve				EBV -ve			
	Case	Control			Case	Control		
Risk factor	No (%) yes	No (%) yes	OR (95%CI)[1]	OR (95%CI)[2]	No (%) yes	No (%) yes	OR (95%CI)[1]	OR (95%CI)[2]
Total Series								
Trend	57 (56.4)	323 (63.1)	0.86 (0.62-1.17)	0.89 (0.64-1.23)	137 (61.7)	323 (63.1)	1.04 (0.83-1.31)	1.08 (0.85-1.37)
Moderate amount vs. little or none	42 (41.6)	247 (48.2)	0.73 (0.46-1.16)	0.81 (0.50-1.31)	94 (42.3)	247 (48.2)	0.82 (0.57-1.16)	0.87 (0.60-1.26)
A lot vs. little or none	15 (14.9)	76 (14.8)	0.85 (0.45-1.63)	0.90 (0.46-1.76)	43 (19.4)	76 (14.8)	1.17 (0.74-1.86)	1.26 (0.78-2.03)
Age 16-34								
Trend	31 (70.5)	151 (68.6)	1.11 (0.69-1.79)	1.09 (0.88-1.65)	83 (66.9)	151 (68.6)	1.21 (0.88-1.65)	1.24 (0.90-1.72)
Moderate amount vs. little or none	19 (43.2)	112 (50.9)	0.84 (0.38-1.82)	0.95 (0.42-2.12)	50 (40.3)	112 (50.9)	0.75 (0.45-1.25)	0.82 (0.48-1.39)
A lot vs. little or none	12 (27.3)	39 (17.7)	1.47 (0.60-3.58)	1.48 (0.58-3.76)	33 (26.6)	39 (17.7)	1.42 (0.77-2.61)	1.54 (0.82-2.90)
Age 50+								
Trend	21 (50.0)	100 (60.6)	0.62 (0.36-1.06)	0.70 (0.40-1.23)	30 (58.8)	100 (60.6)	1.11 (0.71-1.75)	1.16 (0.72-1.86)
Moderate amount vs. little or none	19 (45.2)	72 (43.6)	0.86 (0.41-1.79)	1.02 (0.47-2.21)	21 (41.2)	72 (43.6)	1.10 (0.54-2.27)	1.08 (0.51-2.29)
A lot vs. little or none	2 (4.8)	28 (17.0)	0.23 (0.05-1.07)	0.30 (0.06-1.41)	9 (17.6)	28 (17.0)	1.24 (0.49-3.13)	1.37 (0.53-3.57)

1. Adjusted for age and sex
2. Adjusted for age, sex, and the Carstairs index

Table 9.13: Results for sweat episodes in the 5 years to diagnosis/ pseudo-diagnosis by EBV status

	EBV +ve				EBV -ve			
	Case	Control			Case	Control		
Risk factor	No (%) yes	No (%) yes	OR (95%CI)[1]	OR (95%CI)[2]	No (%) yes	No (%) yes	OR (95%CI)[1]	OR (95%CI)[2]
Total Series								
Trend	45 (44.6)	288 (56.3)	0.85 (0.72-1.00)	0.85 (0.72-1.00)	124 (55.9)	288 (56.3)	0.98 (0.87-1.10)	1.00 (0.85-1.13)
Less than weekly vs. none	1 (0.9)	27 (5.3)	0.15 (0.02-1.10)	0.15 (0.02-1.15)	8 (3.6)	27 (5.3)	0.57 (0.25-1.31)	0.54 (0.22-1.31)
Weekly vs. none	8 (7.9)	55 (10.7)	0.57 (0.25-1.28)	0.70 (0.31-1.61)	22 (9.9)	55 (10.7)	0.77 (0.44-1.35)	0.89 (0.49-1.62)
At least twice a week vs. none	36 (35.6)	206 (40.2)	0.68 (0.43-1.10)	0.69 (0.42-1.12)	94 (42.3)	206 (40.2)	0.89 (0.42-1.27)	0.95 (0.66-1.37)
Age 16-34								
Trend	29 (65.9)	147 (66.8)	0.93 (0.72-1.20)	0.95 (0.73-1.24)	83 (66.9)	147 (66.8)	1.04 (0.88-1.24)	1.08 (0.90-1.29)
Less than weekly vs. none	0 (0.0)	10 (4.5)	0.00 (0.00-∞)	0.00 (0.00-∞)	7 (5.6)	10 (4.5)	1.24 (0.44-3.51)	1.40 (0.49-3.99)
Weekly vs. none	6 (13.6)	28 (12.7)	0.97 (0.34-2.77)	1.27 (0.43-3.75)	12 (9.7)	28 (12.7)	0.76 (0.35-1.66)	1.01 (0.44-2.31)
At least twice a week vs. none	23 (52.3)	109 (49.5)	0.97 (0.47-1.99)	1.03 (0.48-2.18)	64 (51.6)	109 (49.5)	1.04 (0.64-1.71)	1.18 (0.71-1.98)
Age 50+								
Trend	11 (26.2)	70 (42.4)	0.79 (0.60-1.03)	0.76 (0.57-1.01)	20 (39.2)	70 (42.4)	0.96 (0.76-1.21)	0.94 (0.74-1.19)
Less than weekly vs. none	0 (0.0)	5 (3.0)	0.00 (0.00-∞)	0.00 (0.00-∞)	0 (0.0)	5 (3.0)	0.00 (0.00-∞)	0.00 (0.00-∞)
Weekly vs. none	1 (2.4)	10 (6.1)	0.27 (0.03-2.27)	0.38 (0.04-3.27)	3 (5.9)	10 (6.1)	0.79 (0.20-3.14)	0.79 (0.19-3.23)
At least twice a week vs. none	10 (23.8)	55 (33.3)	0.54 (0.25-1.20)	0.48 (0.21-1.11)	17 (33.3)	55 (33.3)	0.88 (0.44-1.75)	0.82 (0.40-1.68)

1. Adjusted for age and sex
2. Adjusted for age, sex, and the Carstairs index

Table 9.14: Results for trend of total MET hours of activity in the age groups under consideration by EBV status

	EBV +ve				EBV –ve			
	Case	Control			Case	Control		
Risk factor	No (%) yes	No (%) yes	OR (95%CI)[1]	OR (95%CI)[2]	No (%) yes	No (%) yes	OR (95%CI)[1]	OR (95%CI)[2]
Total Series								
Trend MET hours age 7-11	47 (46.5)	232 (45.3)	1.25 (1.01-1.54)	1.26 (1.01-1.57)	106 (47.7)	232 (45.3)	1.00 (0.86-1.16)	1.03 (0.88-1.20)
Trend MET hours age 12-14	58 (57.4)	328 (64.1)	0.95 (0.74-1.21)	0.94 (0.73-1.21)	146 (65.8)	328 (64.1)	0.94 (0.79-1.13)	0.91 (0.75-1.10)
Trend MET hours age 15-lve school	35 (34.7)	246 (48.0)	0.76 (0.61-0.94)	0.81 (0.65-1.01)	113 (50.9)	246 (48.0)	1.07 (0.90-1.26)	1.12 (0.94-1.34)
Age 16-34								
Trend MET hours age 7-11	26 (59.1)	132 (60.0)	1.10 (0.83-1.47)	1.09 (0.81-1.48)	67 (54.0)	132 (60.0)	0.96 (0.79-1.17)	0.99 (0.81-1.21)
Trend MET hours age 12-14	32 (72.7)	153 (69.5)	1.00 (0.70-1.43)	0.92 (0.62-1.37)	83 (37.4)	153 (69.5)	0.87 (0.68-1.12)	0.83 (0.64-1.09)
Trend MET hours age 15-lve school	25 (56.8)	126 (57.3)	0.82 (0.59-1.14)	0.93 (0.65-1.34)	70 (56.5)	126 (57.3)	1.16 (0.91-1.47)	1.24 (0.96-1.59)
Age 50+								
Trend MET hours age 7-11	20 (47.6)	46 (27.9)	1.71 (1.11-2.64)	1.78 (1.13-2.80)	16 (31.4)	46 (27.9)	1.17 (0.81-1.69)	1.21 (0.83-1.78)
Trend MET hours age 12-14	21 (50.0)	89 (53.9)	0.89 (0.55-1.43)	0.90 (0.54-1.48)	27 (52.9)	89 (53.9)	0.80 (0.52-1.21)	0.77 (0.50-1.19)
Trend MET hours age 15-lve school	7 (16.7)	48 (29.1)	0.71 (0.47-1.08)	0.75 (0.49-1.14)	17 (33.3)	48 (29.1)	1.13 (0.78-1.65)	1.18 (0.80-1.76)
Trend MET hours age lve school-21	18 (42.9)	72 (43.6)	0.99 (0.71-1.38)	0.96 (0.68-1.35)	23 (45.1)	72 (43.6)	0.96 (0.69-1.33)	0.95 (0.67-1.33)
Trend MET hours age 22-29	18 (42.9)	70 (42.4)	1.28 (0.79-2.07)	1.28 (0.77-2.13)	22 (43.1)	70 (42.4)	0.97 (0.64-1.46)	0.90 (0.59-1.39)
Trend MET hours age 30-39	12 (28.6)	60 (36.4)	0.59 (0.38-0.92)	0.60 (0.37-0.96)	20 (39.2)	60 (36.4)	1.01 (0.72-1.42)	1.07 (0.75-1.52)

1. Adjusted for age and sex
2. Adjusted for age, sex, and the Carstairs index

Table 9.15: Results for any activity dichotomy in the age groups under consideration by EBV status

	EBV +ve				EBV -ve			
	Case	Control			Case	Control		
Risk factor	No (%) yes	No (%) yes	OR (95%CI)[1]	OR (95%CI)[2]	No (%) yes	No (%) yes	OR (95%CI)[1]	OR (95%CI)[2]
Total Series								
Any activity yes vs. no age 7-11	47 (46.5)	232 (45.3)	0.72 (0.33-1.53)	0.70 (0.32-1.55)	106 (47.7)	232 (45.3)	0.79 (0.47-1.33)	0.80 (0.51-1.52)
Any activity yes vs. no age 12-14	58 (57.4)	328 (64.1)	1.61 (0.61-4.22)	1.45 (0.52-4.00)	146 (65.8)	328 (64.1)	0.89 (0.46-1.72)	0.78 (0.46-1.72)
Any activity y vs. n age 15-l school	35 (34.7)	246 (48.0)	0.65 (0.27-1.56)	0.76 (0.30-1.90)	113 (50.9)	246 (48.0)	1.23 (0.66-2.32)	1.23 (0.66-2.32)
Age 16-34								
Any activity yes vs. no age 7-11	26 (59.1)	132 (60.0)	1.26 (0.74-2.11)	1.34 (0.78-2.30)	67 (54.0)	132 (60.0)	0.96 (0.66-1.40)	1.06 (0.71-1.57)
Any activity yes vs. no age 12-14	32 (72.7)	153 (69.5)	0.96 (0.54-1.71)	0.93 (0.51-1.71)	83 (37.4)	153 (69.5)	1.04 (0.66-2.03)	0.88 (0.54-1.42)
Any activity y vs. n age 15-l school	25 (56.8)	126 (57.3)	0.48 (0.27-0.83)	0.55 (0.30-0.98)	70 (56.5)	126 (57.3)	0.99 (0.65-1.52)	1.10 (0.70-1.73)
Age 50+								
Any activity yes vs. no age 7-11	18 (42.9)	46 (27.9)	3.61 (1.39-9.41)	4.06 (1.51-10.91)	16 (31.4)	46 (27.9)	1.05 (0.71-1.54)	1.07 (0.72-1.59)
Any activity yes vs. no age 12-14	21 (50.0)	89 (53.9)	0.63 (0.24-1.65)	0.62 (0.23-1.68)	27 (52.9)	89 (53.9)	1.06 (0.70-1.60)	1.01 (0.66-1.56)
Any activity y vs. n age 15-l school	7 (16.7)	48 (29.1)	0.41 (0.14-1.18)	0.50 (0.17-1.52)	17 (33.3)	48 (29.1)	0.85 (0.57-1.28)	0.88 (0.57-1.35)
Any activity y vs. n age l school-21	18 (42.9)	72 (43.6)	0.86 (0.31-2.43)	0.89 (0.30-2.68)	23 (45.1)	72 (43.6)	0.85 (0.57-1.26)	0.83 (0.55-1.25)
Any activity yes vs. no age 22-29	18 (42.9)	70 (42.4)	2.43 (0.79-7.50)	2.40 (0.74-7.78)	22 (43.1)	70 (42.4)	1.18 (0.77-1.80)	1.13 (0.73-1.74)
Any activity yes vs. no age 30-39	12 (28.6)	70 (42.4)	0.32 (0.12-0.88)	0.32 (0.11-0.91)	20 (39.2)	70 (42.4)	0.95 (0.67-1.35)	0.94 (0.64-1.36)

1. Adjusted for age and sex
2. Adjusted for age, sex, and the Carstairs index

Table 9.16: Results for consistent activity in the age groups under consideration

	Case		Control			
Risk Factor	No	No (%) yes	No	No (%) yes	OR (95%CI)[1]	OR (95%CI)[2]
Total Series						
Top quartile vs. all others age 7-11, 12-14, & 15-leave school	493	27 (5.5)	512	28 (5.5)	0.95 (0.55-1.65)	0.99 (0.56-1.76)
No activity vs. all others age 7-11, 12-14, 15-leave school	493	164 (33.3)	512	144 (28.1)	1.37 (1.04-1.81)	1.32 (0.99-1.76)
All activity combinations vs. no activity age 7-11, 12-14, 15-leave school	493	302 (61.3)	512	340 (66.4)	0.73 (0.51-0.96)	0.75 (0.56-1.01)
Top quartile vs. no activity age 7-11, 12-14, 15-leave school		27 (5.5)		28 (5.5)	0.75 (0.42-1.36)	0.80 (0.43-1.48)
Age 16-34						
Top quartile vs. all others age 7-11, 12-14, & 15-leave school	241	17 (7.1)	220	13 (5.9)	1.18 (0.55-2.50)	1.40 (0.63-3.10)
No activity vs. all others age 7-11, 12-14, 15-leave school	241	71 (29.5)	220	42 (19.1)	1.82 (1.17-2.82)	1.81 (1.15-2.86)
All activity combinations vs. no activity age 7-11, 12-14, 15-leave school	241	153 (63.5)	220	165 (75.0)	0.54 (0.35-0.84)	0.53 (0.34-0.84)
Top quartile vs. no activity age 7-11, 12-14, 15-leave school		17 (7.1)		13 (5.9)	0.73 (0.32-1.62)	0.87 (0.36-2.07)
Age 50+						
No activity vs. all others age 7-11, 12-14, 15-leave school	148	43 (29.1)	165	43 (26.1)	1.28 (0.76-2.14)	1.31 (0.76-2.25)

1. Adjusted for age and sex
2. Adjusted for age, sex, and the Carstairs index

period. When these variables are analysed as levels those people who had performed some activity at some time whilst at school had a statistically significantly lower risk of HD than those who had been inactive. This pattern is repeated in the 16-34 years age group with more extreme odds ratios, but not the 50+ years age group.

Consistent activity was analysed by sex in the total series and the 16-34 and 50+ years age groups (Table 9.17). There are some differences between the sexes with males showing the same pattern of risk as that in the total series. However, in females those subjects that remained in the top quartile of activity had a decreased risk of HD (not statistically significant) compared with an increased risk for males. Females are the same as males in that those who did not perform any activity had a higher risk of HD. However, in females the results were statistically significant and the odds ratios larger. In the 50+ years age group the effect of no activity was no longer present in females but it was in males.

The results for consistent activity in the EBV subgroups are in Table 9.18. The pattern of risk is similar to that in the total series is repeated with the most extreme odds ratios seen in the 16-34 years age group. Those who are in the top quartile of activity have a doubling in risk of EBV +ve HD age 16-34 years (after adjustment for age, sex, and Carstair's index). Also the people who have done some activity had a lower risk than those who had been in the top quartile at all ages had an increased risk (although the results were not statistically significant). There were only very small differences in the odds ratios for EBV +ve and EBV –ve HD.

Table 9.17: Results for consistent activity in the age groups under consideration by sex

	Male				Female			
	Case	Control			Case	Control		
Risk factor	No (%) yes	No (%) yes	OR (95%CI)[1]	OR (95%CI)[2]	No (%) yes	No (%) yes	OR (95%CI)[1]	OR (95%CI)[2]
Total Series								
Top quartile vs. all others	24 (8.5)	22 (7.6)	1.08 (0.59-1.98)	1.09 (0.57-2.06)	3 (1.4)	6 (2.7)	0.50 (0.12-2.03)	0.61 (0.14-2.62)
No activity vs. all others	78 (27.8)	69 (23.8)	1.37 (0.93-2.01)	1.29 (0.86-1.93)	86 (40.6)	75 (33.8)	1.37 (0.92-2.05)	1.36 (0.90-2.05)
All activity comb vs. no activity	179 (63.7)	199 (68.6)	0.72 (0.49-1.07)	0.76 (0.51-1.15)	123 (58.0)	141 (63.5)	0.74 (0.50-1.10)	0.75 (0.49-1.13)
Top quartile vs. no activity	24 (8.5)	22 (7.6)	0.84 (0.43-1.66)	0.89 (0.44-1.81)	3 (1.4)	6 (2.7)	0.40 (0.10-1.69)	0.50 (0.11-2.20)
Age 16-34								
Top quartile vs. all others	14 (10.7)	8 (6.9)	1.48 (0.59-3.70)	1.64 (0.62-4.31)	3 (2.7)	5 (4.8)	0.58 (0.13-2.50)	0.70 (0.15-3.29)
No activity vs. all others	30 (22.9)	18 (15.5)	1.68 (0.88-3.24)	1.77 (0.89-3.54)	41 (37.3)	24 (23.1)	1.96 (1.07-3.57)	1.94 (1.05-3.56)
All activity comb vs. no activity	87 (66.4)	90 (77.6)	0.56 (0.29-1.09)	0.53 (0.26-1.06)	66 (60.0)	75 (72.1)	0.52 (0.28-0.95)	0.52 (0.28-0.96)
Top quartile vs. no activity	14 (10.7)	8 (6.9)	0.94 (0.32-2.71)	0.99 (0.32-3.03)	3 (2.7)	5 (4.8)	0.37 (0.08-1.69)	0.45 (0.09-2.24)
Age 50+								
No activity vs. all others	19 (22.1)	18 (18.9)	1.46 (0.69-3.09)	1.42 (0.64-3.14)	24 (38.7)	25 (35.7)	1.06 (0.51-2.19)	1.08 (0.51-2.31)

1. Adjusted for age and sex
2. Adjusted for age, sex, and the Carstairs index

Table 9.18: Results for consistent activity in the age groups under consideration by EBV status

	EBV +ve				EBV -ve			
	Case	Control			Case	Control		
Risk factor	No (%) yes	No (%) yes	OR (95%CI)[1]	OR (95%CI)[2]	No (%) yes	No (%) yes	OR (95%CI)[1]	OR (95%CI)[2]
Total Series								
Top quartile vs. all others	7 (6.9)	28 (5.5)	1.17 (0.49-4.78)	1.31 (0.54-3.18)	10 (4.5)	28 (5.5)	0.76 (0.36-1.61)	0.87 (0.40-1.87)
No activity vs. all others	39 (38.6)	144 (28.1)	1.76 (1.10-2.80)	1.52 (0.93-2.47)	67 (30.2)	144 (28.1)	1.26 (0.88-1.81)	1.26 (0.87-1.82)
All activity comb vs. no activity	55 (54.5)	340 (66.4)	0.55 (0.34-0.89)	0.64 (0.39-1.05)	145 (65.3)	340 (66.4)	0.80 (0.56-1.15)	0.80 (0.55-1.16)
Top quartile vs. no activity	7 (6.9)	28 (5.5)	0.76 (0.30-1.93)	0.95 (0.37-2.45)	10 (4.5)	28 (5.5)	0.64 (0.29-1.43)	0.74 (0.33-1.66)
Age 16-34								
Top quartile vs. all others	5 (11.4)	13 (5.9)	1.66 (0.55-5.05)	2.10 (0.65-6.75)	7 (5.6)	13 (5.9)	1.00 (0.38-2.62)	1.24 (0.46-3.36)
No activity vs. all others	11 (25.0)	42 (19.1)	1.85 (0.83-4.12)	1.79 (0.77-4.13)	35 (28.2)	42 (19.1)	1.63 (0.97-2.74)	1.62 (0.95-2.77)
All activity comb vs. no activity	28 (63.6)	165 (75.0)	0.51 (0.23-1.14)	0.52 (0.22-1.22)	82 (66.1)	165 (75.0)	0.61 (0.36-1.08)	0.60 (0.35-1.04)
Top quartile vs. no activity	5 (11.4)	13 (5.9)	0.96 (0.27-3.49)	1.23 (0.32-4.77)	7 (5.6)	13 (5.9)	0.68 (0.24-1.94)	0.85 (0.29-2.48)
Age 50+								
No activity vs. all others	13 (31.0)	43 (26.1)	1.34 (0.63-2.85)	1.21 (0.32-4.77)	16 (31.4)	43 (26.1)	1.53 (0.74-3.14)	1.58 (0.75-3.33)

1. Adjusted for age and sex
2. Adjusted for age, sex, and the Carstairs index

9.4 Discussion:

The SNEHD study comprised 493 cases and 512 controls. This was a large study of HD, however, almost half of the subjects did not report any physical activity at all. This lack of numbers of people reporting any activity combines with the problem of ascertaining an individual's activity level. In SNEHD the problem of recall of physical activity looms large, as we were most interested in activity at school and afterwards up to age 39. Many of the cases and controls are older than 50 years of age and thus may have difficulty recalling activity levels many years previously, especially as part of a longer general health questionnaire. This is a problem as exercise or sporting activities are subject to variation and, therefore, are more difficult to recall (Jacobs et al, 1993). For more stable activity e.g. heavy intensity activity, recall is not affected by the timeframe of the questionnaire, but for light or moderate activities it is more difficult to recall the timing of the activity (Jacobs et al, 1993). Respondents were directed using consistently applied interviewer rules at face-to-face interview. The amount of information that could be recorded was the most it was felt reasonable to collect in one section of a questionnaire of over 50 pages lasting more than an hour. Respondents were asked specific questions about each activity performed to elicit detail about the duration, intensity, and frequency (consistency) of each activity. These are the three primary components of physical activity that can be varied and may have different effects on carcinogenesis (Thompson, 1994).

The analysis presented here only takes account of recreational activity i.e. that performed outside working hours. The results, therefore, do not reflect total activity and could be criticised for this. A full employment history was recorded for each SNEHD subject and this, in theory, could be used to calculate energy expenditure at work. However, due to time constraints it was not possible to perform this analysis. In general occupations are more sedentary than in the past and it is now considered that most of an individual's activity takes place outside of the workplace (see Chapter 2) but SNEHD subjects cover a wide range of

ages and this occupational activity may have been important for many of the older subjects. An analysis of total activity in the future could be interesting.

When collecting the SNEHD data it was considered to be better to under-estimate activity rather than over-report. This is why averages for the time spent in an activity were the average of a three-month period to avoid a short burst of activity being misinterpreted as continuous activity across a 10-year age group. Also, use of the seasonal variable allowed the scaling down of activity levels to take account the amount of the year that the activity was performed. Finally all the MET scores used were the average or general scores available in the Ainsworth et al (1993) compendium. However, it could be that the methods were too conservative and have resulted in the under-estimation of activity levels. This could especially apply to individual's that performed more than 5 activities but time constraints during the interview prevented the recording of any more information.

The analysis of SNEHD data attempted to compare four different methods of collating physical activity data: two based on the calculation of weekly expenditure of kilocalories , one based on MET hours and one based on time only. The use of MET scores allowed the impact of activity intensity to be analysed. However, there were very few reports of moderate activities and the odds ratios were unstable with wide confidence intervals. As mentioned in Chapter 2 there is also no agreed definition of which MET score applies to which activity level. The results for each of the four methods were very similar suggesting that risk of HD may be only related to the amount of activity rather than the intensity of activity

The results of the SNEHD analysis for physical activity are mostly negative, indeed the ORs for activity in the five years prior to diagnosis and the investigation of activity levels up to age 39 years are all very close to the null. This observation applies to the total series and the series split by age, EBV status, and sex. It is only when consistent activity is considered that the results become more compelling. There was an increase in the risk of HD associated with consistent inactivity in the total series and all the subgroups. The effect of

being in the top quartile of activity was muted but based on small numbers. Those people in the top level of activity at all ages had a higher risk of HD than those in all the other combinations of activity combined (although the difference was not statistically significant). Surprisingly this pattern is not seen in females. Females analysed alone had the same increased risk of HD associated with consistent inactivity but those females who were consistently in the top quartile of activity had the lowest risk of HD (odds ratios not statistically significant). This pattern was seen in the total series and the 16-34 years age group but could not be analysed in the 50+ years age group due to lack of numbers in the top category. This result could suggest that physical activity is protective for females through a hormonal mechanism analogous to pregnancy. Those in the top category may have fewer menses due to high levels of activity. Alternatively activity may lower oestrogen levels without affecting the menstrual cycle resulting in a lower risk of HD. Exercise training has been found to lower resting levels of oestrogen in most studies (Dale et al, 1979; Boyden et al, 1983; Bullen et al, 1985; Nelson et al, 1988; Cauley et al, 1989; Newcombe et al, 1995; Nagata et al, 1997; De Souza et al, 1998). In theory the effect of inactivity will not be seen in the 50+ years age group for females but this could not be analysed due to lack of numbers.

When looking at EBV subgroups there was the doubling of risk of EBV +ve HD associated with being consistently in the top quartile of activity. In the young adult peak (age 16-34 years) the risk of EBV +ve HD is associated with consistently high activity levels is highest. It may be that immunosuppression caused by consistently high levels of activity leads to inadequate immune response to new or latent EBV infection. Thus, the effect of physical activity on EBV could be similar to the j-shaped curve proposed by Nieman & Nehlson-Cannarella (1992) to describe the relationship between physical activity and URTI. This model suggests that individuals who exercise moderately have a lower risk of URTI compared to the sedentary population; in contrast athletes undergoing strenuous training would exhibit an increased risk of URTI. However, there is no evidence that model is correct for EBV and HD in the presence of physical activity.

There has been very little work performed on the impact of physical activity on the risk of HD. There are only two studies in the literature. Both of these studies have some evidence of an inverse relationship of physical activity and HD risk. Whittemore et al (1985) found a non-significant decreased risk of HD associated with over 5 hours of college sports (RR=0.73). Paffenbarger et al (1987) found a non-significant trend in risk of HD over three activity levels in a cohort study of college alumni. However, neither found evidence of high activity levels increasing risk.

The results of these two studies and the results from SNEHD give some credence to the suggestion that physical activity has an impact on the risk of HD. The results from SNEHD are the first to be statistically significant. Further support for a role of physical activity comes from the presence of a plausible biologic mechanism. What appears to be required is a study of people who perform more consistent levels of physical activity e.g. athletes, dancers. Alternatively studies should take account of total activity, both occupational and recreational. If this latter option is chosen some attempt should be made to take account of home activities e.g. cleaning, washing, which are often ignored in studies of physical activity. It would be necessary to use a specific activity questionnaire and concentrate only on this topic to get as much detail as possible about all activities. A more detailed study could also allow a more thorough investigation of the proposed j-shaped curve of risk.

In conclusion consistent inactivity is associated with an statistically significant increased risk of HD while consistently high activity levels may or may not increase risk. Those who undertake moderate levels of activity have a reduced risk of HD, although this relationship was not statistically significant. The relationship of physical activity and risk of HD requires further investigation.

10. Conclusions.

Analyses presented in this thesis examined risk factors for HD both in total and in sub-groups by EBV status, age and sex. Risk factors related to an infectious aetiology, family health and physical activity have been considered.

In agreement with Newell et al (1985), Neilly et al (1995a & b), and Douglas et al (1996) the analysis presented in chapter 6 found statistically significant evidence of seasonality in the presentation of HD with a peak in January. This finding supports a possible infectious aetiology of HD. The effect of age group and the interaction of age group with seasonality did not reach statistical significance but seasonal presentation was most evident in young adults (age 15-34 years), the age group with most evidence of an infectious aetiology.

IM has frequently been reported to be associated with an increased risk of HD (see Chapter 1). The SNEHD (chapter 8) and YHHCCS (chapter 7) results support this relationship with a statistically significant increased risk of HD following IM. Age at diagnosis subgroup analyses and comparisons in SNEHD suggest that the statistically significant effect of IM is limited to the young adult peak (16-34 years). The increased risk of HD associated with IM was found in both EBV subgroups in the young adult peak. However, the odds ratios associated with EBV +ve HD were higher, especially in the young adult age group in both studies but the difference (EBV +ve vs. EBV –ve) was only statistically significant in YHHCCS. In both SNEHD and YHHCCS the role of IM is not completely limited to EBV +ve HD but the suggestion of a relationship is strongest in this group. The results of the SNEHD and YHHCCS analyses suggest a specific causal association of recent EBV exposure with EBV +ve HD and a weak positive association of EBV –ve HD with prior IM in the young adult peak. Thus, for the first time evidence is presented that the effect of IM is limited to a specific EBV subgroup AND age group.

Results from SNEHD and YHHCCS suggest it is more likely that it is the combined number of infections that is more important than any specific infection in the impact of risk of HD. This is because, apart from IM, no other individual infections, with the possible exception of measles, had an effect on risk of HD. Very few studies have attempted to assess the effect of childhood infectious illnesses on the risk of HD and none have attempted to form composites of infections in total or at specific age groups. The statistically significant protective effect of combined childhood infectious illnesses observed in the YHHCCS (age 16-24 years) analysis was not seen in the young adult peak in SNEHD (age 16-34 years). There was some suggestion that more childhood infectious illnesses aged ≥5 years was associated with a decreased risk of HD age 16-34 years.

Sleckman et al (1998) took as a prior hypothesis that EBV +ve cases would show evidence of childhood experience conducive to late exposure to infectious agents but failed to find any supportive evidence. EBV +ve HD was associated with more childhood infectious illnesses aged ≥5 years and EBV –ve HD was associated with fewer infections at the same age in both SNEHD and YHHCCS. In both studies the difference between the EBV subgroups was statistically significant. In SNEHD this effect is also seen when the childhood and 'additional' infections are combined and the total number of infections are analysed. Thus, for the first time the results from SNEHD and YHHCCS support the hypothesis that EBV +ve HD in the young adult peak has evidence of late (interpreted as age ≥5 years) exposure to infectious agents.

The results for tonsillectomy and physical activity give some support to the idea that the immune system has an impact on risk of HD. A statistically significant increased risk of HD following tonsillectomy in the young adult peak was found in SNEHD remained after adjustment for the five childhood infectious illnesses and IM. Tonsilar tissue is part of the immune system and the tonsils can act as filter barriers to infection (Vianna et al, 1971). It

appears that the removal of this barrier has the most effect in the young adult peak where the evidence of an infectious aetiology for HD is greatest.

An increased risk of HD is associated with consistent inactivity in the total series and all subgroups (chapter 9). Those people in the top level of activity at all ages had a higher risk of HD than those in all the other activity levels combined (although the difference was not statistically significant). This pattern was not seen in females which could also suggest that physical activity is protective for females through a hormonal mechanism analogous to pregnancy.

When looking at EBV subgroups there was the doubling of risk of EBV +ve HD associated with being consistently in the top quartile of activity. In the young adult peak (age 16-34 years) the risk of EBV +ve HD is associated with consistently high activity levels is highest. It may be that immunosuppression caused by consistently high levels of activity leads to inadequate immune response to new or latent EBV infection. Thus, the effect of physical activity on EBV could be similar to the j-shaped curve proposed by Nieman & Nehlson-Cannarella (1992) to describe the relationship between physical activity and URTI. This is the first statistically significant evidence of an effect of physical activity on risk of HD and the first results that suggest, albeit tentatively, that the effect may be limited to specific EBV status, age and sex subgroups.

Different risk factors act differently between the age at diagnosis groups of HD. The effects of IM, infectious illness, tonsillectomy, and possibly physical activity are concentrated in the young adult peak. There are fewer differences associated with EBV status. The main differences in SNEHD and YHHCCS by EBV subgroup are IM and total number of infections age ≥5 years in the 16-34 years age group. Other studies are needed to test these findings in other settings. The results of the analyses presented emphasise the need to include age group and EBV status in any study investigating risk factors for the development of HD.

Bibliography

Abramson JH, Pridon H, Sacks MJ et al (1978) "A case-control study of Hodgkin's Disease in Israel." *JNCI* 61: 307-314.

Acheson ED, Pippard EC, Winter PD (1984) "Mortality of English furniture workers." *Scand J Work Environ Health* 10: 211-217.

Ahluwalia B, Jackson MA, Jones GW et al (1981) "Blood hormone profiles in prostate cancer patients in high-risk and low-risk populations." *Cancer* 48: 2267-2273.

Ainsworth BE, Haskell WL, Leon AS et al (1993) "Compendium of physical activities: classification of energy costs of human activities." *Med Sci Sports Exerc* 25: 71-80.

Ainsworth BE, Sternfield B, Slattery ML et al (1998) "Physical activity and breast cancer: evaluation of physical activity assessment methods." *Cancer* 83: 611-620.

Albanes D, Blair A, TaylorPR (1989) "Physical activity and risk of cancer in the NHANES I population." *Am J Public Health* 79: 744-750.

Alderman BW (1996) "Intensive follow-up of control subjects: is it necessary? (letter)." Am J Public Health 86: 744-745.

Alderson MR, Nayak R (1971) "Study of space-time clustering in Manchester." *Br J Prec Soc Med* 77: 168-173.

Alexander FE, Ricketts TJ, McKinney PA et al (1991a) "Community lifestyle characteristics and incidence of Hodgkin's Disease in young people." *Int J Cancer* 48: 10-14

Alexander FE, McKinney PA, Williams J et al (1991b) "Epidemiological evidence for the "two disease hypothesis" in Hodgkin's Disease." *Int J Epidemiol* 20: 354-361.

Alexander FE (1995a) "Epidemiology: overview and spatial clustering." In: Jarrett RF (ed) "*The Etiology of Hodgkin's Disease.*" New York: Plenum Press pp1-13.

Alexander FE, Daniel CP, Armstrong AA et al (1995b) "Case clustering, Epstein-Barr virus Reed-Sternberg cell status and Herpes virus serology Hodgkin's disease – Results of a case control study." *Eur J Cancer* 31A: 1479-1484.

Ambinder RF, Browning RJ, Lorenzano I et al (1993) "Epstein-Barr virus and childhood Hodgkin's Disease in Honduras and the United States." *Blood* 81: 462-467.

Anagnostopoulos I, Herbst H, Neidobitek G et al (1989) "Demonstration of monoclonal EBV genomes in Hodgkin's disease and KI-1-positive anaplastic large cell lymphoma by combined Southern blot and in situ hybridisation." *Blood* 74: 810-816.

Andersson E, Isager H (1978) "Pre-morbid factors in Hodgkin's disease. II BCG-vaccination status, tuberculosis, infectious diseases, tonsillectomy, and appendectomy." *Scand J Haematol* 21: 273-277.

Andersson R, Hugander A, Thulin A et al (1995) "Clusters of acute appendicitis: further evidence for an infectious aetiology." *Int J Epidemiol* 24: 829-833.

Andrianpopulos G, Nelson RL, Bombeck CT et al (1987) "The influence of physical activity in 1,2 dimethylhydrazine induced colon carcinogenesis in rats." *Anticancer Res* 7: 849-52.

Andriko JAW, Aguiera NS, Nandedkar MA et al (1997) "Childhood Hodgkin's disease in the United States: an analysis of histologic subtypes and association with Epstein-Barr virus." *Modern Pathology* 10: 366-371.

Arbman G, Axelson O, Fredriksson M et al (1993) "Do occupational factors influence risk of colon and rectal cancer in different ways?" *Cancer* 72: 2543-2549.

Armstrong AA, Alexander FE, Paes RP et al (1993) "Association of Epstein-Barr virus with paediatric Hodgkin's Disease." *Am J Pathol* 142: 1683-1688.

Arraiz GA, Wigle DT, Yang Mao (1992) "Risk assessment of physical activity and physical fitness in the Canada Health Survey mortality follow-up study." *J Clin Epidemiol* 45: 419-428.

Austin MA, Criqui MH, Barrett-Connor E et al (1981) "The effect of response bias on the odds ratio." Am J Epidemiol 114: 137-143.

Austin H, Hill HA, Flanders WD et al (1994) "Limitations in the application of case-control methodology." Epidemiologic Reviews 16: 65-76.

Baecke JAH, Burema J, Fritters JER (1982) "A short questionnaire for the measurement of habitual physical activity in epidemiological studies." *Am J Clin Nutr* 36: 932-942.

Baer R, Bankier AT, Biggin AT, et al (1984) "DNA sequence and expression of the B95-8 Epstein-Barr virus genome." *Nature* 310: 207-211.

Bahn AK (1974) "Hodgkin's Disease in schoolteachers (letter)." *N Engl J Med* 291: 207.

Bai MC, Jiwa NM, Horstman A et al "Decreased expression of cellular markers in Epstein-Barr virus positive Hodgkin's Disease"

Baj Z, Kantorski J, Majewskia E et al (1994) "Immunological status of competitive cyclists before and after a training season." *Int J Sports Med* 15: 319-324.

Baker ER, Mather RS, Kirk RF et al (1981) "Female nurses and secondary amenorrhea: correlation with age, parity, mileage, and plasma sex binding globulin concentration." *Fertil Steril* 36: 183.

Ballard-Barbash R, Schatzkin A, Albanes D et al (1990) "Physical activity and risk of large bowel cancer in the Framingham Study." *Cancer Res* 56: 3610-3613.

Bartram HP, Wynder EL (1989) "Physical activity and colon cancer risk? Physiological considerations" *Am J Gastroenterology* 84: 109-112.

Baslund B, Lyngberg K, Andersen V et al (1993) "Effect of 8 weeks of bicycle training on the immune system of patients with rheumatoid arthritis." *J Appl Physiol* 75: 1691-1695.

Bedrossian J, Metivier F, Jaccard A et al (1995) "Hodgkin's disease and cadaveric renal transplantation." *Transplantation Proc* 27: 1783-1784.

Belkaid MI, Zerhouni F, Beljord K et al (1995) "Epstein-Barr virus markers and Hodgkin's disease: comparison of results between Algerian and French patients." *Bulletin du Cancer* 82: 357-363.

Benharroch D, Brousset P, Goldstein J et al (1997) "Association of the Epstein-Barr virus with Hodgkin's disease in southern Israel." *Int J Cancer* 71: 138-141.

Bennet A, DelTacca M (1975) "Prostaglandins in human colonic carcinoma." *Gut* 16: 409.

Berk LS, Nieman DC, Youngberg WS et al (1990) "The effect of long endurance running on natural killer cells in marathoners." *Med Sci Sports Exerc* 22: 207-212.

Berk LS, Tan SA, Nieman DC et al (1985) "The effect of stress from acute exhaustive exercise on lymphocyte-T helper-suppresser cell ratio in athletes and non-athletes." *Med Sci Sports Exerc* 17: 492.

Bernard SM, Cartwright RA, Darwin CM et al (1987) "Hodgkin's Disease: case-control epidemiological study in Yorkshire." *Br J Cancer* 55: 85-90.

Bernstein L, Henderson BE, Hanisch R et al (1994) "Physical exercise and reduced risk of breast cancer in young women." *JNCI* 86: 1403-1408.

Bernstein L, Ross RK, Lobo RA et al (1987) "The effects of moderate physical activity on menstrual cycle patterns in adolescence: implications for breast cancer." *Br J Cancer* 55: 681-5.

Bierman HR (1968) "Human appendix and neoplasia." *Cancer* 21: 109-118.

Bierman PJ, Vose JM, Langnas AN et al (1996) "Hodgkin's disease following solid organ transplantation." *Annals Oncology* 7: 265-270.

Bingham SA, Cummings JH (1989) "Effects of exercise and physical fitness on large intestinal function." *Gastroenterology* 97: 1389-1399.

Bjelke E (1969) "Hodgkin's Disease in Norway." *Acta Med Scand* 185: 73-81.

Bjorkholm M, Holm G, Mellstedt H (1990) "Immunocompetance on patients with Hodgkin's disease." In: Lacher MJ, Redman JR (eds) "*Hodgkin's disease: consequence of survival.*" Philadelphia: Lea & Febiger pp112-150.

Blair A, Hayes HM (1982) "Mortality patterns among US veterinarians, 1947-77: an expanded study." *Int J Epidemiol* 11: 391-397.

Blair SN, Kohl HW, Paffenbarger RS et al (1989) "Physical fitness and all cause mortality: a prospective study of healthy men and women." *JAMA*. 262: 2395-2401.

Blot WJ, McLaughlin JK, Winn DM et al (1988) "Smoking and drinking in relation to oral and pharyngeal cancer." Cancer Res 48: 3282-3287.

Bock A, Popp W, Herkner RR (1994) "Tonsillectomy and the immune system – a long-term follow-up comparison between tonsilectomised and non-tonsilectomised children." *Eur Archives of Oto-Rhino-Laryngology* 251: 423-427.

Bogger-Goren S, Zaizov R, Vogel R et al (1983) "Clinical and virological observations in childhood Hodgkin's disease in Israel." *Israel J Medical Science* 19: 989.

Boiocchi M, Carbone A, De Re V et al (1989) "Is the Epstein-Barr virus involved in Hodgkin's Disease?" *Tumori* 75: 345-350.

Bonelli L, Vitale V, Bistolfi F et al (1990) "Hodgkin's disease in adults: association with social factors and age at tonsillectomy." *Int J Cancer* 45: 423-427.

Bonen A, Long WY, MacIntyre KP (1979) "Effects of exercise on the serum concentration of FSH, LH, progesterone and oestradial." *Eur J Appl Physiol* 42: 15-23.

Bonen A, Belcastro AN, Lingl NY et al (1981) "Profiles of selected hormones during menstrual cycle of teenage athletes." *J Appl Physiol* 50: 545-551.

Booth ML, Owen N, Baumas AE et al (1996) "Retest reliability of recall measures of leisure-time physical activity in Australian adults." *Int J Epidemiol* 25: 153-159.

Boyden TW, Pamenter RW, Stanforth P et al (1981) "Sex steroids and endurance running in women." *Fertility Sterility* 39: 629-632.

Boyle P, Walker AM, Alexander FE (1996) "Historical aspects of leukaemia clusters." In: Alexander FE, Boyle P (eds) "*Methods for investigating localised clustering of disease.*" Lyon, France: IARC Scientific Publication No 135 pp xi.

Brahmi Z, Thomas JE, Park M et al (1985) "The effect of acute exercise on natural killer cell activity of trained subjects and sedentary human beings." *J Clin Immunol* 5: 321-328.

Brandtzaeg P (1987) "Immune function and immunopathology of palatine and nasopharyngeal tonsils." In: Bernstein J, Ogra P (eds) *"Immunology of the ear."* Raven Press, New York.

Brousset P, Chittal S, Schlaifer D et al (1991) "Detection of Epstein-Barr virus messenger RNA in Reed-Sternberg cells of Hodgkin's disease by in situ hybridisation with biotinylated probes on specially modified Acetone Methyl Benzoate Xylene (ModAMex section)." *Blood* 77: 1781-1786.

Brownson RC, Reif JS (1988) "A cancer registry based study of occupational risk for lymphoma, multiple myeloma and leukaemia." *Int J Epidemiol* 17: 27-32.

Brownson RC, Zahm SH, Chang JC et al (1989) "Occupational risk of colon cancer: an analysis by anatomic sub-site." *Am J Epidemiol* 130: 675-687.

Brownson RC, Chang JC, Davis JR et al (1991) "Physical activity on the job and cancer in Missouri." *Am J Public Health* 81: 639-642.

Bullen BA, Skrinar GS, McArthur W et al (1982) "Exercise effect upon plasma melatonin levels in women: possible physiological significance." *Can J Appl Sports Sci* 7: 90-97.

Bullen BA, Skrinar GS, Beitnis IZ et al (1985) "Introduction of menstrual disorders by strenuous exercise in untrained women." *N Eng J Med* 315: 1349-53.

Burkitt DP (1971) "Epidemiology of cancer of the colon and rectum." *Cancer* 28: 3-13.

Burkitt DP, Walker ARP, Painter NS (1972) "Effects of dietary fibre on stools and transit-times and its role in the causation of disease." *Lancet* 2: 1408-1411.

Calabrase LH (1990) "Exercise, immunity, cancer and infection." In: Boucchard C, Shepherd RJ, Stephens T et al (eds) "*Exercise, fitness, and health: a consensus of current knowledge.*" Champagne, Illinois: Human Kinetics Publishing pp 567-579.

Carbone PP, Kaplan HS, Musshoff K et al (1971) "Report of the committee on Hodgkin's Disease staging." *Cancer Res* 31: 1860-1861.

Carter CD, Brown TM, Herbst JT et al (1977) "Cancer incidence following infectious mononucleosis." *Am J Epidemiol* 105: 30-36.

Cartwright RA, Alexander FE, McKinney PA et al (1990) *"Leukaemia and Lymphoma: An atlas of the distribution of cases within areas of England and Wales, 1984-88."* Leukaemia Research Fund, London.

Cauley JA, Gutai JP, Kuller LH et al (1989) "The epidemiology of serum sex hormones in post-menopausal women." *Am J Epidemiol* 129: 1120-1131.

Centres for Disease Control (1992) "Revised classification system for HIV infection and expanded case definition for AIDS among adolescents and adults." *MMWR* 41: 1.

Cerhan J, Chiu B, Wallace R et al (1996) "Systolic blood pressure, physical activity and risk of breast cancer in elderly women." *Am J Epidemiol* 143: SS Abstract 284.

Cerhan et al (1997) "Association of smoking, body mass, and physical activity with risk of prostate cancer in the Iowa 65+ rural health study (US)." *Cancer Causes Control* 8: .

Cerilli J, Rynasiewicz JJ, Lemus LB et al (1977) "Hodgkin's disease in renal transplantation." *Am J Surgery* 133: 182-184.

Chakravarti A, Hallaran SL, Bale SJ et al (1986) "Etiologic heterogeneity in HD: HLA linked and unlinked determinants of susceptibility independent of histological considerations." *Genet Epidemiol* 3: 407-415.

Chan JK, Yip TT, Tsang WY, et al (1995) "Detection of Epstein-Barr virus in Hodgkin's Disease occurring in an Oriental population." *Human Pathology* 26;314-318.

Chang KL, Albujar PF, Chen Y-Y et al (1993) "High prevalence of Epstein-Barr virus in the Reed-Sternberg cells of Hodgkin's Disease occuring in Peru." *Blood* 81: 496-501.

Chen R, Mantel N, Klingberg MA (1984) "A study of three techniques for time-space clustering in Hodgkin's disease." *Statistics in Medicine* 3: 173-184.

Chen YT, Zheng T, Chou MC et al (1997a) "The increase of Hodgkin's disease incidence among young adults. Experience of Connecticut 1935-92." *Cancer* 79: 2209-2218.

Chen C-L, White E, Malone KE et al (1997b) "Leisure-time physical activity in relation to breast cancer among young women (Washington, United States)." *Cancer Causes Control* 8: 77-84.

Cherry (1922) "A theory of cancer." *Med J Aust* 1: 425-438.

Chow WH, Malker HSR, Hsing HW et al (1994) "Occupational risks for colon cancer in Sweden." *J Occup Med* 36: 647-651.

Claviez A, Tiemann M, Peters J et al (1994) "The impact of EBV, proliferation rate, and BCL-2 expression in Hodgkin's disease in childhood." *Annals Hematology* 68: 61-66.

Coenen C, Wegener M, Wedmann B et al (1992) "Does physical activity influence bowel transit time in healthy young men?" *Am J Gastroenterol* 87: 292-295.

Coldman AJ, Elwood JM, Gallagher RP (1982) "Sports activities and risk of testicular cancer." *Br J Cancer* 46: 749-56.

Coleman MP, Esteve J, Damiecki P et al (1993) "*Trends in cancer incidence and mortality.*" Lyon: IARC Scientific Publications No 121.

Connelly RR, Christine BW (1974) "A cohort study of cancer following infectious mononucleosis." *Cancer Res* 34: 1172-1178.

Coogan PF, Newcomb PA, Clapp RW et al (1997) "Physical activity in usual occupation and risk of breast cancer (United States)." *Cancer Causes Control* 8: 626-631.

Cordain L, Latin RW, Behnke JJ (1986) "The effect of an aerobic running program on bowel transit time." *J Sports Med* 26: 101-104.

Correa P, O'Conor GT (1971) "Epidemiologic patterns of Hodgkin's Disease." *Int J Cancer* 8: 192-201.

Cozen W, Katz J, Mack T (1992) "Risk patterns of Hodgkin's disease in Los Angeles vary by cell type." *Cancer Epidemiol Biomarkers Prev* 1: 261-268.

Cridland MD (1961) "Seasonal incidence of clinical onset of Hodgkin's Disease." *BMJ* 2: 621-623.

Criqui MH, Barrett-Connor E, Austin M (1978) "Differences between respondents and non-respondents in a population-based cardiovascular disease study." Am J Epidemiol 108: 367-372.

Crist DM, Mackinnon LT, Thompson RF et al (1989) "Physical exercise increases natural killer mediated tumour cytotoxicity in elderly women." *Gerontol* 35: 66-71.

Crosagnini P, Demicheli V, Fontana A et al (1996) "Incidence and time trends for lymphomas, leukaemias and myelomas: hypothesis generation." *Leuk Res* 20: 285-290.

Cumming DC, Brunsting III LA, Strich G et al (1986) "Reproductive hormones increase in response to acute exercise in men." *Med Sci Sports Exerc* 18: 369-373.

Cumming DC, Wall SR, Quinney SA et al (1987) "Decrease in serum testosterone levels with maximal intensity swimming exercise in trained male and female swimmers." *Endocrin Res* 13: 31-41.

Curran JW, Lawrence DN, Jaffe HW et al (1984) "Aquired immunodeficiency syndrome (AIDS) associated with transfusions." *N Engl J Med* 310: 69-75.

D'Avonzo B, Nanni O, La Vecchia C et al (1996) "Physical activity and breast cancer risk." *Cancer Epidemiology Biomarkers Prev* 5: 155-160.

Dale E, Gerlach DH, Wilhite AL (1979) "Menstrual disfunction among distance runners." *Obstetrics and Gynecology* 54: 47-55.

De Souza MJ, Miller BE, Loucks AB et al (1998) "High frequency of luteal phase deficiency and anovulation in recreational women runners: blunted elevation in follicle-stimulating hormone observed during luteal-follicular transition." *J Clin Endocrinol Metab* 83: 4220-4232.

Delabie J, De Wolf-Peeters C (1992) "Human immunodeficiency virus associated Hodgkin's disease (letter)." 69: 601.

Delsol G, Brousset P, Chittal S et al (1992) "Correlation of the expression of Epstein-Barr virus latent membrane protein and in situ hybridisation with biotinylated Bam H1-W probes in Hodgkin's Disease." *Am J Pathol* 140: 247-253.

Deuster PA, Chouos GP, Luger A et al (1989) "Hormonal and metabolic responses of untrained, moderately trained, and highly trained men to three exercise intensities." *Metabolism* 38: 141-148.

Di Carlo EF, Amberson JB, Metroka CE et al (1986) "Malignant lymphomas and the Acquired Immunodeficiency Syndrome: evaluation of 30 cases using a working formulation." *Arch Pathol Lab Med* 110: 1012-1016.

Diehl V, Tesch H (1995) "Hodgkin's disease and HIV-infection (letter)." *N Engl J Med* 333: 65-66.

Doll R, Hill AB (1964) "Mortality in relation to smoking: 10 years' observation of British doctors." *Br Med J* 1: 1399-1410.

Dorgan JF, Brown C, Barrett M et al (1994) "Physical activity and risk of breast cancer in the Framingham Heart Study." *Am J Epidemiol* 139: 662-9.

Dosemeci M, Hayen RB, Vetter R et al (1993) "Occupational physical activity, socioeconomic status and risk of 15 cancer sites in Turkey." *Cancer Causes Control* 4: 313-321.

Douglas AS, Brown T, Reid D (1996) "Infectious mononucleosis and Hodgkin's disease: a similar seasonality." *Leuk Lymph* 23: 323-331.

Douglas S, Cortina-Barja M, Cartwright R (1998) "Seasonal variation in the incidence of Hodgkin's disease." *Br J Haematol* 103: 653-662.

Doyle TJ, Kumarapuram KV, Koichi M et al (1983) "Hodgkin's Disease in renal transplant recipients." *Cancer* 51: 245-247.

Drexler HG (1992) "Recent results on the biology of Hodgkin's and Reed-Sternberg cells. I Biopsy material." *Leuk Lymph* 8: 283-313.

Duclos M, Corcuff JB, Rashedi M et al (1996) "Does functional alteration of the gonadotropic axis occur in endurance trained athletes during and after exercise/ a preliminary study." *Eur J Appl Physiol Occ Physiol* 73: 427-433.

Edwards JH (1961) "The recognition and estimation of cyclic trends." *Ann Human Genet* 25; 83-87.

Ellison PT, Lager C (1986) "Moderate recreational running is associated with low salivary progesterone profiles in women." *Am J Obstet Gynecol* 143: 1000-1003.

Enblad G, Sandvej K, Lennette E et al (1997) "Lack of correlation between EBV serology and presence of EBV in the Hodgkin's and Reed-Sternberg cells of patients with Hodgkin's disease." *Int J Cancer* 72: 394-397.

Ernster VL, Sacks ST, Selvin S et al (1979) "Cancer incidence by marital status: United States third National Cancer Survey." *JNCI* 63: 567-585.

Errante D, Zagonel V, Vaccher E et al (1994) "Hodgkin's disease in patients with HIV-infection and in the general population – comparison of clinicopathological features and survival." *Annals of Oncology* 5: S37-S40.

Evans AS, Gutensohn N (1984) "A population-based case-control study of EBV and other viral antibodies among persons with Hodgkin's disease and their siblings." *Int J Cancer* 34: 149-157.

Fahey TD, Gates SL, Colvin W et al (1997) "The effects of prolonged intense exercise on oestrdiol, progesterone, LH, and FSH concentration during the mid-menstrual cycle." *Biology of Sport* 14: 175-183.

Fahroner CL, Hackney AC (1998) "Effects of endurance exercise on free testosterone concentration and binding effect of sex-hormone binding globulin (SHBG)." *Int J Sports Med* 19: 12-15.

Fauci AS, Macher AB, Longo DL et al (1984) "Aquired immunodeficiency syndrome: eoidemiologic, immunologic and therapeutic considerations." *Ann Intern Med* 100: 92-106.

Felsing NE, Brasel JA, Cooper DM (1992) "Effect of low and high intensity exercise on circulating growth hormone in men." *J Clin Endocrinol Metab* 75: 157-167.

Ferraris AM, Racchi O, Rapezzi D et al (1997) "Familial HD: a disease of young adulthood?" *Ann Hemat* 74: 131-134.

Fiatrone MA, Marley JE, Bloom ET et al (1989) "The effect of exercise on natural killer cell activity in young and old subjects." *J Gerontol* 44: M37-M45.

Flynn MG, Pizza FX, Boone JB Jr et al (1994) "Incidence of training stress during competitive running and swimming seasons" *Int J Sports Med* 18: 21-26.

Flynn MG, Pizza FX, Brolinson PG (1997) "Hormonal responses to excessive training: influence of cross-training." *Int J Sports Med* 18: 191-196.

Fonte R, Grigis L, Grigis P et al (1982) "Chemicals and Hodgkin's disease." *Lancet* 2: 50.

Forman D, Pike MC, Davey G et al. (1994) "Aetiology of testicular cancer with congenital abnormalities, age at puberty, infertility and exercise." *BMJ* 308: 1393-1399.

Fournier P-E, Stalder J, Bermillod B et al (1997) "Effects of 110 kilometre ultra-marathon race on plasma hormone levels." *Int J Sports Med* 18: 252-256.

Fox AJ, Goldblatt PO (1982) "Longitudinal study 1975: sociodemographic mortality differences." Series LS, no. 1 London: HMSO.

Franceschi S, Serraino D, La Vecchia C et al (1991) "Occupation and risk of Hodgkin's Disease in north-east Italy." *Int J Cancer* 48: 831-835.

Franceschi S, Dal Maso L, Arniani S et al (1998) "Risk of cancer other than Kaposi's sarcoma and non-Hodgkin's lymphoma in persons with AIDS in Italy." *Br J Cancer* 78: 966-970.

Fraser GE, Shavlick D (1997) "Risk factors, lifetime risk and age at onset of breast cancer." *Ann Epidemiol* 7: 375-382.

Fraumeni JF, Li FP (1962) "Hodgkin's Disease in childhood: an epidemiologic study." *JNCI* 42; 681-691.

Fredriksson M, Bengtsson N-O, Hardell L et al (1989) "Colon cancer, physical activity and occupational exposures. A case control study." *Cancer* 63: 1838-1842.

Friedenreich CM, Rohan TE (1995) "A review of physical activity and breast cancer." *Epidemiology* 6: 311-317.

Friedenreich CM, Courneya KS, Bryant HE (1998) "The lifetime physical activity questionnaire: development and reliability." *Med Sci Sports Exerc* 30: 266-274.

Frisch RE, Gotx-Welbergen N, McArthur JW al (1981) "Delayed menarche and amenorrhea among college athletes in relation to age of onset of training." *JAMA* 246: 1559-1563.

Frisch RE, Wynshak G, Albright NL et al (1985) "Lower prevalence of breast cancer and cancers of the reproductive system among former college athletes compared to non-athletes." *Br J Cancer* 52: 885-891.

Fukatsu A, Sato N, Shimizu H (1996) "50 mile walking race suppresses neutrophil bactericidal function by inducing increases in cortisol and ketone bodies." *Life Sciences* 58: 2337-2343.

Gammon MD, Britton JA, Teitelbaum SL (1996) "Does physical activity reduce the risk of breast cancer? – a review of the epidemiological literature." *Menopause* 3: 172-180.

Gammon MD, Schoenberg JB, Britton JA et al (1998) "Recreational physical activity and breast cancer risk among women under 45 years." *Am J Epidemiol* 147: 273-280.

Gann PH, Daviglus ML, Dyer AR (1995) "Heart rate and prostate cancer mortality – results of a prospective analysis." *Cancer Epidemiology Biomarkers Prev* 4: 611-616.

Gann PH, Hennekins CH, Ma J et al (1996) "Prospective study of sex hormone levels and risk of prostate cancer." *JNCI* 88: 1118-1126.

Garelick E, Wittrout RH, Okmura K et al (1982) "Role of Natural Killer cells in the control of metastatic spread and growth of tumour cells in mice." *Int J Cancer* 30: 107-112.

Garfinkel, Stellman (1988) "Mortality by relative weight and exercise." *Cancer* Supplement: 1844-50.

Garnier J-L, Lebranchu Y, Dantal J et al (1996) "Hodgkin's Disease after transplantation." *Transplantation* 61: 71-76.

Geodert JJ, Cote TR, Virgo P et al (1998) "Spectrum of AIDS-associated malignant disorders." *Lancet* 351: 1833-1839.

Gerharrdsson M, Norell ST, Kiviranta H et al (1986) "Sedentary jobs and colon cancer." *Am J Epidemiol* 123: 775-780.

Gerharrdsson M, Steinbeck G, HagmanU et al (1990) "Physical activity and colon cancer: a case-referent study in Stockholm." *Int J Cancer* 46: 985-989.

Giovannucci E, Ascherio A, Rimm EB et al (1995) "Physical activity, obesity, and risk of colon cancer and adenoma in men." *Annals Intern Med* 122: 327-334.

Gittes RF (1991) "Carcinoma of the prostate." *N Engl J Med* 324: 236-245.

Glaser SL (1987) "Regional variation in Hodgkin's Disease incidence by histologic subtype in the United States." *Cancer* 60: 2841-2847.

Glaser SL, Swartz WG (1990) "Time trends in Hodgkin's disease incidence: the role of diagnostic accuracy." *Cancer* 66: 2196-2204.

Glaser SL (1994) "Reproductive factors in Hodgkin's Disease in women." *Am J Epidemiol* 139: 237-246.

Glaser SL, Lin RJ, Steuart SL et al (1997) "Epstein-Barr virus associated Hodgkin's disease: epidemiological characteristics in international data." *Int J Cancer* 70: 375-382.

Gledhill S, Gallagher A, Jones DB et al (1991) "Viral involvement in Hodgkin's Disease: detection of type A Epstein-Barr virus in tumour samples." *Br J Cancer* 64: 227-232.

Gledovic Z, Radovanovic Z (1991) "History of tonsillectomy and appendectomy in Hodgkin's disease." *Eur J Epidemiol* 7: 612-615.

Gleeson M, McDonald WA, Cripps AW et al (1995) "The effect on immunity of long-term intensive training in elite swimmers." *Clin Exp Immunol* 102: 210-216.

Goyal RK, McEvoy L, Wilson DB (1996) "Hodgkin's disease after renal transplantation in childhood." *J Pediatr Hemat Oncol* 18: 392-395.

Green RL, Kaplan SS, Rabin BS et al (1981) "Immune function in marathon runners." *Ann Allergy* 47: 73-75.

Greene MH, Brinton LA, Fraumeni JF et al (1978) "Familial and sporadic Hodgkin's Disease associated with occupational wood exposure." *Lancet* 2: 626-627.

Groves RM, Lyberg LE (1988) "An overview of nonresponse issues in telephone surveys." In: Groves RM, Biemer PP, Lyberg LE et al (eds) "Telephone survey methodology." New York: Wiley p191-211.

Grufferman S, Duong T, Cole P (1976) "Occupation and Hodgkin's Disease." *JNCI* 57: 1193-1195.

Grufferman S, Cole P, Levitan TR (1979) "Evidence against transmission of Hodgkin's Disease in high schools." *N Engl J Med* 300: 1006-1011.

Grufferman S. (1982) "Hodgkin's Disease." In: Schottenfeld D, Fraumeni JF Jr (eds) "*Cancer epidemiology and prevention.*" Philadelphia: WB Saunders & Co pp789.

Grufferman S, Delzell E (1984) "Epidemiology of Hodgkin's Disease." *Epidemiologic Review* 6: 76-106.

Gruhlich A, Wan X, Law M et al (1997) "Rates of non-AIDS defining cancers in people with AIDS." *J Acquired Immune Deficiency Syndromes & Human Retrovirology* 14: A18.

Gruss HJ, Pinto A, Duyster J et al (1997) "Hodgkin's disease: a tumour with disturbed immunological pathways." *Immunol Today* 18: 156-163.

Gui XE, Charles R, Rinaldo JR et al (1983) "Natural killer cell activity in renal transplant patients receiving cyclosporin." *Infect Immun* 41: 965-970.

Gulledge TP, Hackney TC (1996) "Reproducibility of low resting testosterone concentrations in endurance trained men." *Eur J Appl Physiol Occ Physiol* 73: 582-583.

Gulley ML, Raab-Traub N (1993) "Detection of Epstein-Barr virus in human tissues by molecular genetic techniques." *Archives of Pathol Lab Med* 117: 1115-1120.

Gutensohn N, Li FP, Johnson RE et al (1975) "Hodgkin's Disease, tonsillectomy and family size." *N Engl J Med* 292: 22-25.

Gutensohn N, Cole P (1977) "Epidemiology of Hodgkin's Disease in the young." *Int J Cancer* 19: 595-604.

Gutensohn N, Cole P (1981) "Childhood social environment and Hodgkin's Disease." *N Engl J Med* 304: 135-140.

Gutensohn N, Shapiro D (1982) "Social class risk factors among children with Hodgkin's disease." *Int J Cancer* 30: 433-435.

Hack V, Strobel G, Rau J-P et al (1992) "The effect of maximal exercise on the activity of neutrophil granulocytes in highly trained athletes in a moderate training period." *Eur J Appl Physiol* 65: 520-524.

Hack V, Strobel G, Weiss M et al (1994) "PMN cell counts and phagocytic activity of highly trained athletes depend on training period." *J Appl Physiol* 77: 1731-1735.

Hackney AC, Fahrner CL, Gulledge TP (1998) "Basal reproductive hormonal profiles are altered in endurance trained men." *J Sports Med Physical Fitness* 38: 138-141.

Hakkinen et al (1990) "Neuromuscular adaptation and serum hormones in females during prolonged power training." *Int J Sports Med* 11: 91-98.

Hakkinen K, Pakarinen A (1995) "Acute hormonal responses to heavy resistance exercise in men and women at different ages." *Int J Sports Med* 16: 507-513.

Hanto DW, Sakamoto K, Purtilo DT et al (1981) "The Epstein-Barr virus in the pathogenesis of post-transplant lymphoproliferative disorders." *Surgery* 40: 204-213.

Hardell L, Bengtsson NO (1983) "Epidemiologic study of socio-economic factors and clinical findings in Hodgkin's Disease, and reanalysis of previous data regarding clinical exposure." *Br J Cancer* 48: 217-225.

Harris NL, Jaffe ES, Stern H et al (1994) "A Revised European-American Classification of lymphoid neoplasms: a proposal from the International Lymphoma Stugy Group." *Blood* 84: 1361-1392.

Harris V, Sandridge AL, Black RJ et al (1998) "*Cancer registration statistics Scotland 1986-1995.*" Edinburgh: ISD Scotland Publications.

Hartge P, Brinton LA, Rosenthal JF et al (1984) "Random digit dialling in selecting a population-based control group." Am J Epidemiol 120: 825-833.

Hartley LH, Mason JW, Hogan RP et al (1972) "Multiple hormonal responses to graded exercise in relation to physical training." *J Appl Physiol* 33: 602-606.

Hartman TJ, Albanes D, Rautalahti M et al (1998) "Physical activity and prostate cancer in the Alpha-Tocopherol Beta-Carotene (ATBC) Cancer Prevention Study (Finland)." *Cancer Causes Control* 9: 11-18.

Haskell WL, Taylor HL, Wood PD et al (1980) "Strenuous physical activity, treadmill exercise test performance and plasma high density lipoprotein cholesterol." *Circulation* 62 (Suppl IV): 53-61.

Haughey BP, Graham S, Brasure J et al (1989) "The epidemiology of testicular cancer in upstate New York." *Am J Epidemiol* 130: 25-36.

Henle W, Henle G (1973) "Epstein-Barr virus-related serology in Hodgkin's disease." *Natl J Cancer (Natl Cancer Inst Monograph)* 36: 79-84.

Herbst H, Steinbrecher E, Niedobitek G et al (1992) "Distribution and phenotype of Epstein-Barr virus harbouring cells in Hodgkin's Disease." *Blood* 80: 484-491.

Hesse J, Levine PH, Ebbesen P et al (1977) "A case-control study on immunity to two Epstein-Barr virus-associated antigens and to herpes simplex virus and adenovirus in a population-based group of patients with Hodgkin's disease in Denmark, 1971-73." *Int J Cancer* 19: 49-58.

Hessol NA, Katz MH, Liu JY et al (1992) "Increased incidence of Hodgkin's disease in homosexual men with HIV-infection." *Annals of Internal Medicine* 117: 309-311.

Higginson J, Muir CS, Munoz N (1992) "*Human cancer: epidemiology and environmental causes.*" Cambridge: Cambridge university press.

Hirose K, Tajima K, Hamajima N et al (1995) "A large-scale, hospital-based case-control study of risk factors for breast cancer according to menopausal status." *Jpn J Cancer Res* 86: 146-154.

Hodgkin T. On some morbid appearances of the absorbent glands and spleen. *Med Chir Trans* 1832:17;68-114.

Hoffman-Goetz L, Simpson JR, Cipp N et al (1990) "Lymphocyte subset responses to repeated submaximal exercise in men." *J Appl Physiol* 68: 1069-1074.

Hoffman-Goetz L, Husted J (1995) "Exercise and cancer: do the biology and epidemiology correspond?" *Exerc Immunol Rev* 1: 81-96.

Holbrook NJ, Cod WI, Horner WC (1983) "Direct suppression of natural killer cell activity in human peripheral blood leucocytic cultures by glucocorticoids and its modulation by Interferon." *Cancer Res* 43: 4019-4023.

Holdstock DJ, Misiemicz JJ, Smith T et al (1970) "Propulsion (mass movement) in the human colon and its relationship to meals and somantic activity." *Gut* 11: 91-99.

Holt VL, Daling JR, Stergachis A et al (1991) "Results and effects of refusal recontact in a case-control study of ectopic pregnancy." *Epidemiology* 2: 375-379.

Hood IM, Mehendra P, Marcus RE (1996) "Hodgkin's disease after cardiac transplant: a report of 2 cases." *Clinical & Laboratory Haematology* 18: 115-116.

Hoover R (1974) "Hodgkin's Disease in schoolteachers (letter)." *N Engl J Med* 290: 1329.

Hoover R, Fraumeni JF Jr (1973) "Risk of cancer in renal transplant recipients." *Lancet* 2: 55-57.

Horn JW, Devesa SS, Burhansstipanov L (1996) "Cancer, incidence, mortality and survival among racial and ethnic minority groups in the United States." In: Schottenfeld D, Fraumeni JF Jr (eds) "*Cancer Epidemiology and Prevention. 2nd Edition.*" New York: Oxford University Press pp192.

Hsing AW, McLaughlin JK, Zheng W et al (1994) "Occupation, physical activity, and risk of prostate cancer in Shanghai, People's Republic of China." *Cancer Causes Control* 5: 136-140.

Hu YH, Nagata C, Shimuzu N et al (1997) "Association of body mass index, physical activity, and reproductive histories with breast cancer: a case-control study in Gifu, Japan." *Breast Cancer Res Treat* 43: 65-72.

Huh J, Park C, Juhng S et al (1996) "A pathologic study of Hodgkin's Disease in Korea and its association with Epstein-Barr virus infection." *Cancer* 75: 949-955.

Hyams L, Wynder EL (1968) "Appendectomy and cancer risk: an epidemiological evaluation." *J Chron Dis* 21: 391-415.

IARC Working Group on the Evaluation of Carcinogenic Risks in Humans (1996) "*IARC Monograph on the evaluation of carcinogenic risks for humans from human immunodeficiency virus and human T-cell lymphotropic virus.*" Lyon: IARC.

IARC (1997) "*IARC monographs on the evaluation of carcinogenic risks in humans. Epstein-Barr virus and Kaposi's Sarcoma Herpesvirus/Human Herpesvirus 8.*" 70: 157.

Ilic M, Vlajinas H, Marinkovic J (1996) "Case-control study of risk factors for prostate cancer." *Br J Cancer* 74: 1682-1685.

Innes J, Newall J (1961) "Seasonal incidence of clinical onset of Hodgkin's Disease." *BMJ* 2: 765.

Ioachim HL, Dorsett B, Cronin W et al (1991) "Acquired immunodeficiency syndrome associated lymphomas: clinical, pathologic, imunologic and viral characteristics of 111 cases." *Human Pathology* 22: 659-673.

ISD (1999) "*Scottish Health Statistics.*" Edinburgh: Information & Statistics Division Scottish NHS pp86-87.

Jackson H Jr, Parker F Jr (1944) "Hodgkin's Disease II. Pathology." *N Engl J Med* 231: 35-44.

Jacobs DR Jr, Ainsworth BE, Hartman TO et al (1993) "A simultaneous evaluation of 10 commonly used physical activity questionnaires." *Med Sci Sports Exerc* 25: 81-91.

Jaffe HW, Bregman DJ, Selick RM (1983) "Aquired immunodeficiency syndrome in the United States: the first 1000 cases." *J Infect Dis* 148: 339-345.

Janerich DT, Glebatis D, Flink E et al (1979) "Case-control studies on the effect of sex steroids on women and their offspring." J Chron Dis 32: 83-88.

Jarcard A, Emile J-F, Metivier F et al (1994) "Epstein-Barr virus associated Hodgkin's Disease following renal transplantation." *Am J Hematol* 46: 256-257.

Jarman B, Townsend P, Carstairs V (1991) "Deprivation indices." *BMJ* 303: 523.

Jarret RF, Gallagher A, Jones DB et al (1991) "Detection of Epstein-Barr virus genomes in Hodgkin's Disease: relation to age." *J Clin Pathol* 44: 844-848.

Jarrett RF (1992) "Hodgkin's Disease." *Ballieres Clinical Haematology* 5: 57-79.

Jarrett AF, Armstrong AA, Alexander FE (1996) "Epidemiology of EBV and Hodgkin's lymphoma." *Annals of Oncology* 7(Suppl 4): S5-S10.

Jenson J, Oftebro H, Breigan B et al (1991) "Comparison of changes in testosterone concentration after strength and endurance exercise in well-trained men." *Eur J Appl Physiol* 63: 467-471.

Jobe JB, Tourangeau R, Smith AF (1993) "Contribution of survey research to the understanding of memory." *Appl Cognit Psychol* 7: 567-584.

Johnson ES, Fischman HR, Matanoski GM et al (1986) "Cancer mortality among white males in the meat industry." *J Occup Med* 28: 23-32.

Johnson SK, Johnson RE (1972) "Tonsillectomy history in Hodgkin's Disease." *N Engl J Med* 287: 1122-1125.

Jurkowski JE, Jones NL, Walker WC et al (1981) "Ovarian response to exercise." *Med Sci Sports Exerc* 13: 109-114.

Kalandidi A, Tzonov A, Lipworth L et al (1996) "A case-control study of endometrial cancer in relation to reproductive, somatometric and life style variables." *Oncology* 53: 354-359.

Kanavaros P, Sakalidou A, Tzardi M et al (1994) "Frequent detection of Epstein-Barr virus (EBV) EBER transcripts and latent membrane protein-1 (LMP-1) in tumour cells in Hodgkin's disease arising in childhood." *Pathology Research and Practice* 190: 1026-1030.

Kannel WB, Sorlie P (1979) "Some health benefits of physical activity: the Framingham Study." *Arch Intern Med* 139: 857-61.

Khan G, Coates PJ, Gupta RK et al (1992) "Presence of Epstein-Barr virus in Hodgkin's disease is not exclusive to Reed-Sternberg cells." *Am J Pathology* 140: 757-762.

Khan G, Norton AJ, Slavin G (1993) "Epstein-Barr virus in Hodgkin's disease. relation to subtype and age." *Cancer* 71: 3124-3129.

Kinlen LJ, Sheil AGR, Peto J et al (1979) "Collaborative United Kingdom-Australasian study of cancer in patients treated with immunosuppressive drugs." *BMJ* 2: 1461-1466.

Kinlen LJ (1996) "Immunologic factors, including AIDS." In: Schottenfeld D, Fraumeni JF (eds) "*Cancer Epidemiology and Prevention, 2nd edition.*" New York: Oxford University Press pp532-545.

Kirchoff LV, Evans AS, McClelland KE et al (1980) "A case-control study of Hodgkin's disease in Brazil: epidemiological aspects." *Am J Epidemiol* 112: 595-608.

Klitz W, Aldrich CL, Fildes N et al (1994) "Localisation of predisposition to Hodgkin's Disease in the HLA class II region." *Am J Hum Genet* 54: 497-505.

Klurfeld DM, Weber MM, Kritchevsky D (1987) "Inhibition of chemically induced mammary and colon tumour promotion by caloric restriction in rats fed increased dietary fat." *Cancer Res* 47: 2759-2762.

Knopf KB, Locker GY (1995) "Hodgkin's Disease and HIV infection (letter)." *N Engl J Med* 333: 65-66.

Koblin BA, Hessol NA, Zauber AG et al (1996) "Increased incidence of cancer among homosexual men, New York city and San Francisco, 1978-90." *Am J Epidemiol* 144: 916-923.

Koffler KH, Menken A, Redmond RA et al (1992) "Strength training accelerates gastro-intestinal transit time in middle aged and older men." *Med Sci Sports Exerc* 24: 415-419.

Kordeck R, Jesionek-Kupniecka D, Biernat W et al (1996) "Expression of the latent membrane protein of Epstein-Barr virus in Hodgkin's disease: age and subtype distribution in Polish patients." *Acta Haematologica Polonica* 27: 15-20.

Kraemer WJ (1988) "Endocrine responses to resistance exercise." *Med Sci Sports Exerc* 20: S152-S157.

Kraemer et al (1991) "Endogenous anabolic hormonal and growth factor responses to heavy resistance exercise in males and females." *Int J Sports Med* 12: 228-235.

Kraemer WJ, Fleck SJ, Dziados JE et al (1993) "Changes in hormonal concentrations after different heavy resistance exercise protocols in women." *J Appl Physiol* 75: 594-604.

Kramer MM, Wells CL (1996) "Does physical activity reduce risk of estrogen-dependent cancer in women?" *Med Sci Sports Exerc* 28: 322-334.

Kraemer WJ, Hakkinen K, Newton RU et al (1998) "Acute hormonal responses to heavy resistance exercise in younger and older men." *Eur J Appl Physiol Occ Physiol* 77: 206-211.

Kraemer WJ, Staron RS, Hagerman FC et al (1998a) "The effects of short-term resistance training on endocrine function in men and women." *Eur J Appl Physiol Occ Physiol* 78: 64-76.

Kravdal O, Hansen S (1993) "Hodgkin's Disease: the protective effect of childbearing." *Int J Cancer* 55: 909-914.

Kravdal O, Hansen S (1996) "The importance of childbearing for Hodgkin's Disease: new evidence from incidence and mortality models." *Int J Epidemiol* 25: 737-743.

Kriska AM, Caspersen CJ (1997) "A collection of physical activity questionnaires for health-related research." *Med Sci Sports Exerc* 29: S1-S205.

Kryscio RJ, Myers MH, Prusine ST et al (1969) "The space-time distribution of Hodgkin's Disease in Connecticut 1940-1969." *JNCI* 50: 1107.

Kuoppasalmi K, Naveri H, Harkonen M et al (1980) "Plasma cortisol, androstenedione, testosterone and luteinising hormone in running exercise of different intensities." *Scand J Clin Lab Invest* 40: 403-409.

Kuoppasalmi K, Naveri H, Rehiunen S et al (1976) "Effect of strenuous anaerobic running exercise on plasma growth hormone, cortisol, luteinising hormone, testosterone, androstenedione, estrone, and estradiol." *J Steroid Biochem* 7: 823-829.

Kusuda M, Tariyama K, Kamidigo NO et al (1998) "A comparison of the epidemiologic, histologic and virologic studies on Hodgkin's disease in Western Kenya and Nagasaki, Japan." *Am J Tropical Med Hygiene* 59: 801-807.

Kvale G, Hoiby EA, Pedersen E (1979) "Hodgkin's Disease in patients with previous infectious mononucleosis." *Int J Cancer* 23: 593-597.

La Vecchia C, Negri E, D'Avonzo B et al (1989) "Occupation and lymphid neoplasms." *Br J Cancer* 60: 385-388.

La Vecchia C, Levi F, Lucchini F et al (1991) "Hodgkin's disease mortality in Europe." *Br J Cancer* 64: 723-734.

LaCoste D, Dabis F, Boulogne N et al (1998) "Cancer incidence in a cohort of human immunodeficiency virus seroconverters." *Cancer* 82: 996-997.

Lambe M, Hsieh C-C, Tsaih SW et al (1998) "Childbearing and the risk of Hodgkin's disease." *Cancer Epidemiology Biomarkers Prev* 7: 831-834.

Lampe JW, Slavin JL, Apple FS (1991) "Iron status of active women and the effect of running a marathon on bowel function and gastrointestinal blood loss." *Int J Sports Med* 12: 173-179.

LaPorte RE, Montoye HJ, Caspersen CJ (1985) "Assessment of physiologic activity in epidemiologic research: problems and prospects." *Public Health Reports* 100: 131-146.

Le Marchand L, Kolonel LN, Yoshizawa CN (1991) "Lifetime occupational physical activity and prostate cancer risk." *Am J Epidemiol* 133: 103-111.

Lehtinen T, Lumio J, Dillner J et al (1993) "Increased risk of malignant lymphomas indicated by elevated Epstein-Barr virus antibodies – a prospective study." *Cancer Causes Control* 4: 187-193.

Lehtinen T, Lehtinen M (1998) "Common and emerging infectious causes of hematological malignancies in the young." *APMIS* 106: 585-597.

Leocini L, Spina D, Nyong'O A et al (1996) "Neoplastic cells of Hodgkin's disease show differences in EBV expression between Kenya and Italy." *Int J Cancer* 65: 781-784.

Leon AS, Jacobs DR Jr, DeBacker G et al (1981) "Relationship of physical characteristics and life habits to treadmill exercise capacity." *Am J Epidemiol* 113: 653-660.

Levi F, Canon C, Touitou Y, Reinberg A, Mathe G. (1988) "Seasonal modulation of the circadian structure of circulating T and natural killer lymphocyte subsets from healthy subjects." *J Clin Invest* 81; 407-413.

Levi F, La Vecchia C, Negri E et al (1993) "Selected physical activities and risk of endometrial cancer." *Br J Cancer* 67: 846-851.

Levine PH, Lin A, Tucker MA (1995) "What can we learn about the aetiology of Hodgkin's Disease from family studies?" In: Jarrett RF (ed) "*The Etiology of Hodgkin's Disease.*" New York: Plenum Press pp27-32.

Lewicki R, Tchorzewski H, Denys A et al (1987) "Effect of physical exercise on some parameters of immunity in conditioned sportsmen." *Int J Sports Med* 8: 309-314.

Li DX, Oda K, Mikata A et al (1995) "Epstein-Barr virus genomes in Hodgkin's disease and non-Hodgkin's lymphoma." *Pathology International* 45: 735-741.

Liaw KL, Adami J, Gridley G et al (1997) "Risk of Hodgkin's disease subsequent to tonsillectomy: a population-based cohort study in Sweden." *Int J Cancer* 72: 711.

Linsted KD, Tonstad S, Kuzmo JW (1991) "Self-report of physical activity and patterns of mortality in Seventh Day Adventist men." *J Clin Epidemiol* 44: 355-364.

Liu S-M, Chow K-C, Chiu C-F et al (1998) "Expression of Epstein-Barr virus in patients with Hodgkin's disease in Japan." *Cancer* 83: 367-371.

Longnecker MP, Gerharrdsson de Verdier M, Frumkin M et al (1995) "A case control study of physical activity in relation to risk of cancer of the right colon and rectum in men." *Int J Epidemiol* 24: 42-50.

Lucia A, Chicharro JL, Perez M et al (1996) "Reproductive function in male endurance athletes: sperm analysis and hormonal profile." *J Appl Physiol* 81: 2627-2636.

Luger A, Deuster PA, Kyle SB et al (1987) "Acute hypothalamic-pituitary-adrenal response to the stress of treadmill exercise: physiologic adaptation to endurance training." *N Engl J Med* 316: 1309-1315.

Lukes RJ, Butler JJ (1966) "The pathology and nomenclature of Hodgkin's Disease." *Cancer Res* 26: 1063.

Lukes RJ, Butler JJ, Hicks EB (1966a) "Natural history of Hodgkin's Disease as related to its pathologic picture." *Cancer* 19: 317.

Lukes RJ, Craver LF, Hall TC et al (1966b) "Report of the nomenclature committee." *Cancer Res* 26: 1311.

Lumio J, Karjalainen S (1993) "Patients treated in hospital for infectious mononucleosis and risk of cancer." *Scand J Infect Dis* 25: 283-288.

Lyter DW, Bryant J, Thackeray R et al (1995) "Incidence of human immunodeficiency virus-related and non-related malignancies in a large cohort of homosexual men." *J Clin Oncol* 13: 2540-2546.

MacFarlane GJ, Evstifeeva T, Boyle P et al (1995) "International patterns in the occurrence of Hodgkin's Disease in children and young males." *Int J Cancer* 61: 165-169.

Mack TM, Cozen W, Shibata DK et al (1995) "Concordance for Hodgkin's Disease in identical twins suggesting genetic susceptibility to the young-adult form of the disease." *N Engl J Med* 332: 413-418.

Mackinnon LT (1992) "*Exercise and immunology.*" Champagne, Illinois: Human Kinetics Publishing pp36-39, 69-75, 89-90.

MacMahon B (1957) "Epidemiologic evidence on the nature of Hodgkin's Disease." *Cancer* 10: 1045-1054.

MacMahon B (1966) "Epidemiology of Hodgkin's Disease." *Cancer Res* 26: 1189-1200

Marcus PM, Newcomb PA, Storer BE (1994) "Early adulthood physical activity and colon cancer risk among Wisconsin women." *Cancer Epidemiol Biomarkers Prev* 3: 641-644.

Mason TJ, Fraumeni JF (1974) "Hodgkin's Disease among Japanese Americans (letter)." *Lancet* 1: 215.

Matanoaski GM, Sartwell PE, Elliott EA (1975) "Hodgkin's Disease mortality among physicians." *Lancet* 1: 926-927.

McKinney PA, Alexander FE, Ricketts TJ et al (1989) "A specialist leukaemia/ lymphoma registry in the UK. Part I: Incidence and geographical distribution of Hodgkin's disease." *Br J Cancer* 60: 942-947.

McKinney PA, Alexander FE, Roberts BE et al (1990) "Yorkshire case control study of leukaemias and lymphomas: parallel multivariate analyses of seven disease categories." *Leuk Lymph* 2: 67-80.

McTiernan A, Stanford JL, Weiss NS et al (1996) "Occurrence of breast cancer in relation to recreational exercise in women age 50-64 years." *Epidemiology* 7: 598-604.

Medeiros LJ, Greiner TC (1995) "Hodgkin's Disease." *Cancer* 75(Suppl 1): 357-369.

Meittenen OS (1985) "The case-control study: valid selection of subjects." J Chron Dis 38: 543-548.

Mezzetti M, La Vecchia C, De Carli A et al (1998) "Population attributable risk for breast cancer; diet, nutrition, and physical exercise." *JNCI* 90: 389-394.

Milham S Jr, Hesser JE (1967) "Hodgkin's Disease in woodworkers." *Lancet* 2: 136-137.

Milham S (1974a) "*Mortality experience of the AFL-C10 United Brotherhood of Carpenters and Joiners of America 1969-1970.*" Cincinnati, Ohio: National Institute for Occupational Safety and Health pp74-129.

Milham S Jr (1974b) "Hodgkin's disease as an occupational disease of school teachers." *N Engl J Med* 290: 1329.

Milham S (1983) "*Occupational mortality in Washington State, 1959-79.*" Washington DC: Government Printing Office [DHHS(NIOSH) Publication #83-116].

Miller AB, Barclay THC, Choi NW et al (1980) "A study of cancer, parity and age at first pregnancy." *J Chron Dis* 33: 595-605.

Miller BA, Blair AE, Raynor HL et al (1989) "Cancer and other mortality patterns among US furniture workers." *Br J Ind Med* 46: 508-515.

Miller RW, Beebe GW (1973) "Infectious mononucleosis and the empirical risk of cancer." *JNCI* 50: 315-321.

Mink PJ, Folsom AR, Sellers TA et al (1996) "Physical activity, waist-to-hip ratio, and other risk factors for ovarian cancer: a follow-up study of older women." *Epidemiology* 7: 38-45.

Min-Lee IM, Paffenbarger RS, Hseih CC (1991) "Physical activity and risk of developing colo-rectal cancer among college alumni." *JNCI* 83: 1324-1329.

Min-Lee I, Paffenbarger RS, Hsieh CC (1992) "Time trends in physical activity among college alumni." *Am J Epidemiol* 135: 915-925.

Min-Lee I, Manson JE, Ajani U et al (1997) "Physical activity and risk of colon cancer: the Physicians' Health Study (United States)." *Cancer Causes Control* 8: 568-574.

Mittendorf R, Longnecker MP, Newcomb PA et al (1995) "Strenuous physical activity in young adulthood and risk of breast cancer (US)." *Cancer Causes Control* 6: 347-353.

Modan B, Goldman B, Shani M et al (1969) "Epidemiological aspects of neoplastic disorders in Israeli migrant populations." *JNCI* 42: 375-381.

Monterosso V, Zhou Y, Koo S et al (1998) "Hodgkin's disease in Costa Rica: a recent report of 40 cases analysed for Epstein-Barr virus." *Am J Clin Pathol* 109: 618-624.

Montoye HJ (1975) "*Physical activity and health: an epidemiologic study of a total community.*" Englewood Cliffs, NJ: Prentice Hall.

Moreau A, Dantal J, Heymann MF et al (1996) "Post-transplantation Hodgkin's disease: a clinico-pathologic study of 3 cases." *Annales de Pathologie* 16: 124-127.

Mourad W, Alsohaibani MO, Saddick M et al (1998) "Epstein-barr virus expression in Hodgkin's disease: correlation with histologic subtypes and T and B lymphocyte distributions." *Annals Saudi Medicine* 18: 296-300.

Mueller BU, Pizzo PA (1995) "Cancer in children with primary or secondary immuno-deficiencies." *J Pediatrics* 126: 1-10.

Mueller N, Marie Swanson G, Hsieh C-C et al (1987) "Tonsillectomy and Hodgkin's disease: results from companion population-based studies." *JNCI* 78:1-5.

Mueller N, Evans A, Harris NL et al (1989) "Hodgkin's Disease and Epstein-Barr Virus." *N Engl J Med* 320: 689-695.

Muir CS, Waterhouse J, Mack T et al (1987) "*Cancer incidence in 5 continents.*" Lyon: IARC Scientific Publications No 88.

Mujika I, Chatard JC, Padilla S et al (1996) "Hormonal responses to training and its tapering off in competitive swimmers: relationship with performance." *Eur J Appl Physiol Occ Physiol* 74: 361-366.

Munoz N, Davidson RJL, Witthoff B et al (1978) "Infectious mononucleosis and Hodgkin's Disease." *Int J Cancer* 22: 10-13.

Nagata C, Kaneda N, Kabuto M et al (1997) "Factors associated with serum levels of estradiol and sex hormone binding globulin among pre-menopausal Japanese women." *Environ Health Perspec* 105: 994-997.

National Center for Health Statistics, Wilmore JH (1989) "Design issues and alternatives in assessing physical fitness among apparently healthy adults in a health examination survey of the general population." In: US Department of Health and Human Services, Public Health Service, National Center for Health Statistics "*Assessing physical fitness and physical activity in population based surveys.*" DHHS Publication No (PHS) 89-1253.

Nehlsen-Cannarella SL, Nieman DC, Balk-Lamberton AJ et al (1991a) "The effect of moderate exercise training on immune response." *Med Sci Sports Exerc* 23: 64-70.

Nehlsen-Cannarella SL, Nieman DC, Jessen J et al (1991b) "The effects of acute moderate exercise on lymphocyte function and serum immunoglobin levels." *Int J Sports Med* 12: 391-398.

Neilly IJ, Dawson AA, Bennett BB, Douglas AS. (1995a) "Evidence for a seasonal variation in presentation of Hodgkin's Disease." *Blood* 86;1343.

Neilly IJ, Dawson AA, Bennett B et al (1995b) "Evidence for seasonal variation in the presentation of Hodgkin's disease." *Leuk Lymph* 19: 325-328.

Nelson ME, Meredith CN, Dawson-Hughes B et al (1988) "Hormones and bone mineral status in endocrine trained and sedentary menopausal women." *J Clin Endocronol* 66: 927-933.

Newcomb PA, Klein R, Klein BEK et al (1995) "Association of dietry and lifestyle factors with sex-hormones in post-menopausal women." *Epidemiology* 6: 318-321.

Newell G (1972) "Seasonal onset of Hodgkin's Disease." *Lancet* 1: 1024-1025.

Newell GR, Rawlings W, Kinnear BK et al (1973) "Case-control study of Hodgkin's Disease. 1 Results of the interview questionnaire." *JNCI* 51: 1437-1441.

Newell GR, Lynch HK, Gibeau JM et al (1985) "Seasonal diagnosis of Hodgkin's disease." *JNCI* 74: 53-56.

Nieman DC, Johanson LM, Lee JW et al (1990a) "Infectious episodes in runners before and after the LA marathon." *J Sports Med Physical Fitness* 30: 316-328.

Nieman DC, Nehlsen-Cannarella SL, Markhoff PA et al (1990b) "The effect of moderate exercise training on natural killer cells and acute upper respiratory tact infection." *Int J Sports Med* 11: 467-473.

Nieman DC, Nehlsen-Cannarella SL, Donohue KM et al (1991) "The effect of acute moderate exercise on leukocyte and lymphocyte sub-populations." *Med Sci Sports Exerc* 23: 578-585.

Nieman DC, Henson DA, Johnson R et al (1992) "Effects of brief, heavy exertion on circulating lymphocyte sub-populations and proliferative response." *Med Sci Sports Exerc* 24: 1339-1345.

Nieman DC, Miller AR, Henson DA et al (1993a) "Effects of high vs. moderate intensity exercise on Natural Killer cell activity." *Med Sci Sports Exerc* 25: 1126-1134.

Nieman DC, Henson DA, Gusewitch G et al (1993b) "Physical activity and immune function in elderly women." *Med Sci Sports Exerc* 25: 823-831.

Nieman DC, Buckley KS, Henson DA et al (1995) "Immune function in marathon runners versus sedentary controls." *Med Sci Sports Exerc* 27: 986-992.

Nieman DC (1997) "Exercise immunology: practical applications." *Int J Sports Med* 18: S91-S100.

O'Grady J, Stewart S, Elton RA, et al (1994) "Epstein-Barr virus in Hodgkin's Disease and the site of origin of tumour." *Lancet* 343: 265-266.

O'Neil MJ (1979) "Estimating non-response bias due to refusals in telephone surveys." *Public Opin Q* 43: 218-232.

O'Shea J, Ortaldo JR (1992) "The biology of natural killer cells: insights into the molecular basis of function." In: Lewis CE, McGee JO (eds) "*The natural killer cell.*" Oxford: Oxford university Press pp 1-40.

Obminski Z, Aniol Stryzyzewska K, Burkhard Jagodzinska K et al (1997) "Endocrine and metabolic responses to an indoor rowing competition in male and female elite athletes." *Biology of Sport* 14: 145-153.

Oettle GJ (1991) "Effect of moderate exercise on bowel habit." *Gut* 32: 941-944.

Oldhafer KJ, Bunzendahl H, Frei U et al (1989) "Primary Hodgkin's lymphoma: an unusual cause of graft dysfunction after kidney transplantation." *Am J Med* 87: 218-220.

Oliveria SA, Blair SN (1993) "Association between physical activity and prostate cancer." *Am J Epidemiol* 138: 643-644.

Oliveria SA, Kohl HW, Trichopoulos D et al (1996) "The association between cardiorespiratory fitness and prostate cancer." *Med Sci Sports Exerc* 28: 97-104.

Oliveria SA, Christos PJ (1997) "The epidemiology of physical activity and cancer." *Annals NY Acad Sci* 833: 79-90.

Oliveria SA, Min-Lee I (1997) "Is exercise beneficial in the prevention of prostate cancer?" *Sports Med* 23: 271-278.

Olsen J, Sabroe S (1979) "A follow-up study of non-retired and retired workers of the Danish Carpenters/Cabinet Makers Union." *Int J Epidemiol* 8: 379-382.

Olsen SH, Kelsey JL, Pearson TA et al (1992) "Evaluation of random digit dialling as a method of selection in case-control studies." *Am J Epidemiol* 135: 210-222.

Olson SH, Vena JE, Dorn JP et al (1997) "Exercise, occupational activity, and risk of endometrial cancer." *Annals of Epidemiology* 7: 46-53.

Olsson H, Olssen ML, Ranstam J (1990) "Late age at first full-term pregnancy as a risk factor for women with malignant lymphomas." *Neoplasia* 37: 185-190.

Oza AM, Tonks S, Lim J et al (1994) "A clinical epidemiological study of human leukocyte antigen DPB alleles in Hodgkin's disease." *Cancer Res* 54: 5101-5105.

Paffenbarger RS Jr, Wing AL, Hyde RT (1977) "Characteristics of youth indicative of adult-onset Hodgkin's Disease." *JNCI* 58: 1489-1491.

Paffenbarger RS Jr, Wing AL, Hyde RT (1978) "Physical activity as an index of heart attack risk in college alumni." *Am J Epidemiol* 108: 161-175.

Paffenbarger RS Jr, Hyde RT, Wing AL et al (1986) "Physical activity, all-cause mortality and longevity of college alumni." *N Engl J Med* 314: 605-613.

Paffenbarger RS Jr, Hyde RT, Wing AL (1987) "Physical activity and risk of cancer in diverse populations: a preliminary report." *Am J Clin Nutr* 45: 312-317.

Paffenbarger RS Jr, Blair SN, Lee I-M et al (1993) "Measurement of physical activity to assess health effects in free-living populations." *Med Sci Sports Exerc* 25: 60-70.

Pallesen G, Hamilton-Dutoit SJ, Rowe M et al (1991) "Expression of Epstein-Barr virus latent gene products in tumour cells of Hodgkin's Disease." *Lancet* 307: 320-322.

Park CS, Juhng SW, Brigate DJ et al (1994) "Analysis of Epstein-Barr virus in Hodgkin's disease – experience of a single university hospital in Korea." *J Clin Lab Analysis* 8: 412-417.

Parkin DM, Muir CS, Whelan SL et al (eds) (1992) "*Cancer incidence in 5 continents*" (Volume VI). Lyons: IARC Scientific Publications No 120.

Paulino AFG, Paulino-Cabrera E, Weiss LM et al (1996) "Hodgkin's disease in the Philippines." *Modern Pathology* 9: 115-119.

Pearce NE, Smith AH, Fisher DO (1985) "Malignant lymphomas and multiple myeloma linked with agricultural occupation in a New Zealand cancer-registry based study." *Am J Epidemiol* 121: 225-237.

Pearce NE, Smith AH, Reif JS (1988) "Increased risks of soft tissue sarcoma, malignant lymphoma and acute myeloid leukaemia in abbatoir workers." *Am J Indust Med* 14: 63-72.

Pedersen BK, Tvede N, Hansen FR et al (1988) "Modulation of natural killer cell activity in peripheral blood by physical exercise." *Scand J Immunol* 27: 673-678.

Pedersen BK, Tvede N, Christensen LD et al (1989) "Natural killer cell activity in peripheral blood of highly trained and untrained persons." *Int J Sports Med* 10: 129-131.

Pedersen BK, Tvede N, Klarlund K et al (1990) "Indomethacin in vitro and in vivo abolishes post-exercise suppression of natural killer cell activity in peripheral blood." *Int J Sports Med* 11: 127-131.

Pedersen BK, Ullum H (1994) "NK cell response to physical activity: possible mechanisms of action." *Med Sci Sports Exerc* 26: 140-146.

Pedersen, Bruunsgaard (1995) "How physical exercise influences the establishment of infection." *Sports Med* 19: 393-400.

Peh SC, Looi LM, Pallesen G (1997) "Epstein-Barr virus (EBV) and Hodgkin's disease in a multi-ethnic population in Malaysia." *Histopathology* 30: 227-233.

Penn I (1978) "Malignancies associated with immunosuppressive or cytotoxic therapy." *Surgery* 88: 492-502.

Penn I (1981) "Malignant lymphomas in organ transplant recipients." *Transplantation* 13: 736-738.

Peters EM, Bateman ED (1983) "Ultramarathon running and upper respiratory tract infections – an epidemiologic study." *S Afr Med J* 64: 582-584.

Peters EM, Goetzsche JM, Grobbelaar B et al (1993) "Vitamin C supplementation reduces the incidence of post-race symptoms of URTI in ultra-marathon runners." *Am J Clin Nutr* 57: 170-176.

Peters RK, Garabrant DH, Yu MC et al (1989) "A case control study of occupational and dietary factors in colorectal cancer in young men by subsite." *Cancer Res* 49: 5459-5468.

Petersen GR, Milham S (1974) "Hodgkin's Disease mortality and occupational exposure to wood." *JNCI* 53: 957-958.

Pinkus GS, Lones M, Shintaku IP et al (1994) "Immunohistochemical detection of Epstein-Barr virus encoded latent membrane protein in Reed-Sternberg cells and variants of Hodgkin's disease." *Modern Pathology* 7: 454-461.

Poppema S, Visser L (1994) "Epstein-Barr virus positivity in Hodgkin's Disease does not correlate with an HL-A A2 negative phenotype." *Cancer* 73: 3059-3063.

Preciado MV, De Matteo E, Diez B et al (1995) "Epstein-Barr virus (EBV) latent membrane protein (LMP) in tumor cells of Hodgkin's disease in pediatric patients." *Med Pediatr Oncol* 24: 1-5.

Pross HF, Bainen MG (1982) "Studies of human natural killer cells I. In vivo parameters affecting normal cytotoxic function." *Int J Cancer* 29: 383-390.

Pukkala E, Poskiparta M, Apter D et al (1993) "Lifelong physical activity and cancer risk among Finnish female teachers." *Eur J Cancer Prev* 2: 369-376.

Pukkala E, Notkola V (1997) "Cancer incidence among Finnish farmers, 1979-93." *Cancer Causes Control* 8: 25-33.

Pyne DB, Baker MS, Fricker PA et al (1995) "Effects of an intensive 12-week training programme by elite swimmers on neutrophil oxidative activity." *Med Sci Sports Exerc* 27: 536-542.

Quintanilla-Martinez L, Gamboa-Momnquez A et al (1995) "Association of Epstein-Barr virus, latent mambrane protein and Hodgkin's Disease in Mexico." *Modern Pathology* 8: 675-679.

Rabkin CS, Hilgartner MW, Hedberg KW et al (1992) "Incidence of lymphomas and other cancers in HIV-infected and HIV-uninfected patients with hemophilia." *JAMA* 267: 1090-1094.

Rabkin CS, Biggar RJ, Baptiste MS et al (1993) "Cancer incidence trends in women at high risk of human immunodeficiency virus (HIV) infection." *Int J Cancer* 55: 208-212.

Rabkin CS, Yellin F (1994) "Cancer incidence in a population with a high prevalence of infection with human immunodeficiency virus type 1." *JNCI* 86: 1711-1716.

Ragni MV, Belle SH, Jaffe RA, et al (1993) "Acquired immunodeficiency syndrome-associated non-Hodgkin's lymphomas and other malignancies in patients with haemophilia." *Blood* 81: 1889-1897.

Rang EH, Acheson ED (1981) "Cancer in furniture workers." *Int J Epidemiol* 10: 253-261.

Raymond E, Tricollet V, Samuel D et al (1995) "Epstein-Barr virus-related localised hepatic lymphoproliferative disorders after liver transplantation." *Cancer* 76: 1344-1351.

Razzouk BJ, Gan YJ, Mendonca C et al (1997) "Epstein-Barr virus in pediatric Hodgkin's disease: age and histiotype are more predictive than geographic region." *Med Pediatr Oncol* 28: 248-254.

Reddy BS, Sugie S, Lowenfels A (1988) "Effect of voluntary exercise on azooxymethane-induced colon carcinogenesis in male F3 rats." *Cancer Res* 48: 7079-7081.

Reif J, Pearce N, Fraser J (1989) "Cancer risk in New Zealand farmers." *Int J Epidemiol* 18: 768-774.

Reynolds P, Saunders LD, Layefsky ME et al (1993) "The spectrum of Acquired Immunodeficiency Syndrome (AIDS)-associated malignancies in San Francisco 1980-1987." *Am J Epidemiol* 137: 19-30.

Richter EA, Sutton JR (1994) "Hormonal adaptations to physical activity." In: Bouchard C, Shepherd RJ, Stephens T (eds) "*Physical activity, fitness and health.*" Champaign: Human Kinetics pp331-342.

Rickinson AB (1994) "Epstein-Barr virus infection and Epstein-Barr virus associated tumours." In: Minson A, Niel J, McCrae M (eds) "*Viruses and cancer. 51st symposium of the society for general microbiology.*" Cambridge university press.

Robert NJ, Schneidermann H (1984) "Hodgkin's disease and the Acquired Immunodeficiency Syndrome (letter)." *Ann Intern Med* 101: 142-143.

Robinson LL, Stoker V, Frizzera G et al (1987) "Hodgkin's disease in pediatric patients with naturally occurring immunodeficiency." *Am J Pediatr Hematol/Oncol* 9: 189-192.

Rockhill B, Willett WC, Hunter DJ et al (1998) "Physical activity and breast cancer risk in a cohort of young women." *JNCI* 90: 1155-1160.

Roder JC, Halitosis T, Klein M et al (1980) "A new immunodeficiency disorder in humans involving natural killer cells." *Nature* 284: 553-555.

Roitt IM, Brostoff J, Male DK (1991) "*Kurzes Lehrbuch der Immunologie.*" New York: Thiemi, Stuttgart.

Rosdahl N, Olesen Larsen S, Clemmesen J (1974) "Hodgkin's Disease in patients with previous infectious mononucleosis: 30 years experience." *BMJ* 2: 253-256.

Ross PK, Paganin-Hill A, Henderson E (1988) "Epidemiology of prostatic cancer." In: Skinner DG, Lieskovsky G (eds) "*Diagnosis and management of genitourinary cancer.*" Philadelphia: WB Saunders pp 40-45.

Ruuskanen O, Vanha-Pertilla T, Kouvalainen K (1971) "Tonsillectomy, appendectomy and Hodgkin's Disease." *Lancet* 1: 1127-1128.

Sandler RS, Holland KL (1990) "Fate of incorrectly addressed mailed questionnaires." *J Clin Epidemiol* 43: 45-47.

Sandler RS, Pritchard ML, Bangdiwala SI (1995) "Physical activity and the risk of colo-rectal adenomas." *Epidemiology* 6: 602-606.

Schechtman KB, Barzilai B, Rost K et al (1991) "Measuring physical activity with a single question." *Am J Public Health* 81: 770-3.

Scheib RG, Siegel RS (1985) "Atypical Hodgkin's disease and the Acquired Immunodeficiency Syndrome." *Ann Intern Med* 102: 554.

Schwartz B, Cummings DC, Riordan E et al (1981) "Exercise associated amenorrhea: a distinct entity." *Am J Obstet Gynecol* 141: 662-670.

Serraino D, Franceschi S, Talamini R et al (1991) "Socio-economic indicators, infectious diseases, and Hodgkin's disease." *Int J Cancer* 47: 352-357.

Serraino D, Pezzotti P, Dorrucci M et al (1997) "Cancer incidence in a cohort of human immunodeficiency virus seroconverters." *Cancer* 79: 1004-1008.

Severson RK, Nomurs AMY, Grove JS et al (1989) "A prospective analysis of physical activity and cancer." *Am J Epidemiol* 130: 522-529.

Shangold MM, Gatz ML, Thysen B (1981) "Acute effects of exercise on plasma concentrations of prolactin and testosterone in recreational women runners." *Fertil Steril* 35: 699-702.

Shapiro S, Weinblatt E, Frank CW et al (1965) "The HIP Study of incidence and prognosis of coronary heart disease." *J Chron Dis* 18: 527-58.

Shephard, Shek (1994) "Potential impact of physical activity and sport on the immune system: a brief review." *Br J Sports Med* 28: 247-255.

Shepherd RJ, Shek PN (1995) "Cancer, immune function, and physical activity." *Can J Appl Physiol* 20: 1-25.

Shinkai S, Share S, Shek PN et al (1992) "Acute exercise and immune function." *Int J Sports Med* 13: 452-461.

Siconolfi SF, Lasater TM, Snow RCK et al (1985) "Self-reported physical activity compared with maximal oxygen uptake." *Am J Epidemiol* 122: 101-105.

Silingardi V, Venezia L, Tampieri A et al (1982) "Tonsillectomy, appendectomy, and malignant lymphomas." *Scand J Haematol* 28: 59-64.

Slattery ML, Schumacher MC, Smith KR et al (1988) "Physical activity, diet and risk of colon cancer in Utah." *Am J Epidemiol* 128: 989-999.

Slattery ML, Edwards SL, Caan BJ et al (1995) "Response rates among control subjects in case-control studies." *Ann Epidemiol* 5: 245-249.

Slattery ML, Potter J, Caan B et al (1997) "Energy balance and colon cancer – beyond physical activity." *Cancer Res* 57: 75-80.

Sleckman BG, Mauch PM, Ambinder RF et al (1988) "Correlation between EBV status and risk factor profile in Hodgkin's Disease." *Can Epidemiol Biomark Prev* 7: 1117-21.

Slivnick DJ, Ellis TM, Nawrocki JF et al (1990) *Semin Oncol* 17: 673-682.

Smith PG, Kinlen LJ, Doll R (1974) "Hodgkin's Disease mortality among physicians." *Lancet* 2: 525.

Spiers PS (1969) "Hodgkin's disease in workers in the wood industry." *US Public Health Reports* 84: 385-389.

Staal GP, Ambinder R, Bescharner NE et al (1989) "A survey of Epstein-Barr virus DNA in lymphoid tissues: frequent detection in Hodgkin's Disease." *Am J Clin Pathol* 1: 1-5.

Stech CG (1981) "Trends in non-response rates 1952-79." *Public Opinion Q* 45: 45-57.

Steenland K, Nowlin S, Palu S (1995) "Cancer incidence in the National Health and Nutrition Survey I follow-up data – diabetes, cholesterol, pulse, and physical activity." *Cancer Epidemiology Biomarkers Prev* 4: 807-811.

Sterling WA, Wu LYF, Dowling EA (1974) "Hodgkin's disease in a renal transplant recipient: a case report." *Transplantation* 17: 315-317.

Stiller CA (1998) "What causes Hodgkin's disease in children." *Eur J Cancer* 34: 523-528.
Strasner A, Davis JM, Kohut ML et al (1997) "Effects of exercise intensity on natural killer cell activity in women." *Int J Sports Med* 18: 56-62.

Sturgeon SR, Brinton CA, Berman ML et al (1993) "Past and present physical activity and endometrial cancer risk." *Br J Cancer* 68: 584-589.

Svejgaard A, Platz P, Ryder LP et al (1975) "HL-A and disease association. A survey." *Transplant Rev* 22: 3.

Sulke AN, Jones DB, Wood PJ (1985) "Variation in natural killer cell activity in peripheral blood during the menstrual cycle." *BMJ* 290: 884-886.

Sullivan JL, Bryon KS, Brewster FE et al (1980) "Deficient natural killer cell activity in X-linked lymphoproliferative disorder." *Science* 210: 543-545.

Sullivan SN, Wong C (1992) "Runners diarrhea: different patterns and associated factors." *J Clin Gastroenterol* 14: 101-104.

Sutton JR, Coleman MJ, Casey J et al (1973) "Androgen responses during physical exercise." *BMJ* 1: 520-521.

Taioli E, Barone J, Wynder EL (1995) "A case control study on breast cancer and body mass." *Eur J Cancer* 31A: 723-728.

Takvorian T, Kalnicki S, Bloomer W et al (1993) "Hodgkin's Disease." In: Holland JF, Frei E, Bast RC, Kufe DW, Morton DL, Weichselbaum RR (eds) "*Cancer Medicine*" (Third edition) Philadelphia: Lea & Febiger.

Tavani A, Pregnolato A, La Vecchia C et al (1997) "A case-control study of reproductive factors and risk of lymphomas and myelomas." *Leukemia Res* 21: 885-888.

Taylor GM, Gokhale DA, Crowther D et al (1996) "Increased frequency of HLA-DPB1*0301 in Hodgkin's Disease suggests that susceptibility is HVR-sequence and subtype associated." *Leukaemia* 10: 854-859.

Taylor HL, Jacobs DR, Shucker B et al (1978) "A questionnaire for the assessment of leisure time physical activities." *J Chron Dis* 31: 741-755.

Thomas DB, Jiminez LM, McTiernan A et al (1992) "Breast cancer in men: Risk factors with hormonal implications." *Am J Epidemiol* 135: 734-748.

Thompson HJ (1994) "Effect of exercise intensity and duration on the induction of mammary carcinogenesis." *Cancer Res* 54: 1960s-1963s.

Thun MJ, Calle EE, Namboodiri MM et al (1992) "Risk factors for fatal colon cancer in a large prospective study." *JNCI* 84: 1491-1500.

Thune I, Brenn T, Lund E et al (1997) "Physical activity and risk of breast cancer." *N Engl J Med* 336: 1269-1275.

Thune I, Lund E (1994) "Physical activity and risk of prostatic and testicular cancer – a cohort study of 53000 Norwegian men." *Cancer Causes Control* 5: 549-556.

Thune I, Lund E (1996) "Physical activity and risk of colorectal cancer in men and women." *Br J Cancer* 73: 1134-1140.

Thyss A, Caldini C, Bourier C et al (1984) "Comparison of natural killer cell activity during the first and second halves of the menstrual cycle in women (letter)." *Br J Cancer* 50: 127-128.

Tirelli U, Errante D, Dolcetti R et al (1995) "Hodgkin's disease and human immunodeficiency virus infection – clinicopathological and virological features of 114 patients from the Italian Co-operative Group on AIDS and Tumours." *J Clin Oncol* 13: 1758-1767.

Tomita Y, Ohsawa M, Kanno H et al (1996) "Epstein-Barr virus in Hodgkin's disease patients in Japan." *Cancer* 77: 186-192.

Tutton PJM, Barkla DH (1980) "Influence of prostaglandin analogues on epithelial cell proliferation and xenopgraft growth." *Br J Cancer* 41: 47-51.

Tvede N, Steensberg J, Baslund B et al (1991) "Cellular immunity in highly-trained elite racing cyclists and controls during periods of training with high and low intensity." *Scand J Sports Med* 1: 163-166.

U.S. Department of Health & Human Services (1996) "*Physical activity and health: A report of the Surgeon General.*" Atlanta, Georgia: U.S. Department of Health & Human Sciences, Centers for Disease Control & Prevention, National Center for Chronic Disease Prevention & Health Promotion.

Uccini S, Monardo F, Stoppacciaro A et al (1990) "High frequency of Epstein-Barr virus genome detection in Hodgkin's Disease of HIV positive patients." *Int J Cancer* 46: 581-585.

UKTCSG (1994) "Social, behavioural and medical factors in the aetiology of testicular cancer: results from the UK study group." *Br J Cancer* 70: 513-520.

Urhausen et al (1987) "7 week follow-up study of behaviour of testosterone and cortisol in rowers." *Eur J Appl Physiol* 56: 528.

Vasankari TJ, Kujala UM, Heinonen OJ et al (1993) "Effects of endurance training on hormonal responses to prolonged physical exercise in males." *Acta Endocrinol* 129: 109-113.

Vena JE, Graham S, Zielezney M et al (1985) "Lifetime occupational exercise and colon cancer." *Am J Epidemiol* 122: 357.

Vena JE, Graham S, Zielezny M et al (1987) "Occupational exercise and risk of cancer." *Am J Clin Nutr* 45: 318-327.

Vestlev PM, Pallesen G, Sandvej K et al (1992) "Prognosis of Hodgkin's Disease is not influenced by EBV latent membrane protein (letter)." *Int J Cancer* 50: 671-672.

Vianna NJ, Greenwald P, Davies JNP (1971) "Tonsillectomy and Hodgkin's Disease: the lymphoid tissue barrier." *Lancet* 1: 431-432.

Vianna NJ, Greenwald P, Davies JNP (1979) "Extended epidemic of Hodgkin's disease in high school students." *Lancet* 1: 1209.

Vianna NJ, Keogh MD, Polan AK et al (1974) "Hodgkin's Disease mortality among physicians." *Lancet* 2: 131-133.

Vianna NJ, Polan AK (1973) "Epidemiological evidence for the transmission of Hodgkin's Disease." *N Engl J Med* 289: 499-502.

Vihko VJ, Apter DL, Pukkala EJ et al (1992) "Risk of breast cancer among female teachers of physical education and languages." *Acta Oncologica* 31: 201-204.

Vineis P, Ciccone G, Magnino A (1993) "Asbestos exposure, physical activity and colon cancer – a case-control study." *Tumori* 79: 301-303.

Volek JS, Kraemer WJ, Bush JA et al (1997) "Testosterone and cortisol in relationship to dietary nutrients and resistance exercise." *J Appl Physiol* 82: 49-54.

Wacholder S, McLaughlin JK, Silverman DT et al (1992) "Selection of controls in case-control studies. I Principles." Am J Epidemiol 135: 1019-1028.

Wannamethee G, Gerald, Shaper A et al (1993) "Heart rate, physical activity, and mortality from cancer and other non-cardiovascular disease." *Am J Epidemiol* 137: 735-748.

Warren MP (1980) "The effects of exercise on pubertal progression and reproductive function in girls." *J Clin Endocrinol Metabol* 51: .

Washburn RA, Goldfield SRW, Smith KW et al (1990) "The validity of self-reported exercise-induced sweating as a measure of physical activity." *Am J Epidemiol* 132: 107-113.

Washburn RA, Montoye HJ (1986) "The assessment of physical activity by questionnaire." *Am J Epidemiol* 123: 563-576.

Washburn TC, Medearis DN, Childs B (1965) "Sex differences in susceptibility to infection." *Pediatrics* 35: 57-64.

Webb et al (1984) "Serum testosterone concentration during two hours of moderate intensity treadmill running in trained men and women." *Endocrine Res* 10: 27-38.

Weinpahl J (1990) "Physical activity, diet, and risk of colon cancer in Utah (letter)." *Am J Epidemiol* 131: 567.

Weinreb M, Day PJ, Niggli F et al (1996) "The role of Epstein-Barr virus in Hodgkin's disease from different geographical areas." *Archives of Disease in Childhood* 74: 27-31.

Weinreb M, Day PJR, Murray PG et al. (1992) "Epstein-Barr virus (EBV) and Hodgkin's Disease in children: incidence of EBV latent membrane protein in malignant cells." *J Pathol* 168: 365-369.

Weinrels M, Day PJR, Niggli EK et al (1996) "The consistent association between Epstein-Barr virus and Hodgkin's Disease in children in Kenya." *Blood* 87: 3828-3836.

Weintraub J, Warnke RA (1982) "Lymphomas in cardiac allograft recipients: clinical and histological features and immunological phenotype." *Transplantation* 33: 347-351.

Weiss LW, Cureton KJ, Thompson FN (1983) "Comparison of serum testosterone and androstenedione responses to weight-lifting in men and women." *Eur J Appl Physiol Occup Physiol* 50: 413-419.

Weiss LM, Movahed LA, Warnke RA et al (1989) "Detection of Epstein-Barr virus genomes in Reed-Sternberg cells of Hodgkin's disease." *N Engl J Med* 320: 502-506.

Weiss LM, Strickler SG, Hu E et al (1986) "Immunoglobin gene rearrangements in Hodgkin's disease." *Human Pathology* 17: 1009-1014.

Welsh RM, Vargas-Cortes N (19??) "Natural killer cells in viral infection." In: Lewis CE, McGee JO (eds) "*The natural killer cell.*" Oxford: Oxford university Press.

West DW, Schuman KL, Lyon JL et al (1984) "Differences in risk estimation from a hospital and a population-based case-control study." Int J Epidemiol 13: 235-239.

Westerbeek RMC, Blair V, Eden OB et al (1998) "Seasonal variation in the onset of childhood leukaemia and lymphoma." *Br J Cancer* 78: 119-124.

Westergaard T, Melbye M, Pedersen JB et al (1997) "Birth order, sibship size, and risk of Hodgkin's disease in children and young adults: a population based study of 31 million person-years." *Int J Cancer* 72: 977-981.

Wheeler GD, Wall SR, Belcastro AN et al (1984) "Reduced serum testosterone and prolactin levels in male distance runners." *JAMA* 252: 514-516.

White E, Jacobs EJ, Daling JR (1996) "Physical activity in relation to colon cancer in middle-aged men and women." *Am J Epidemiol* 144: 42-50.

Whittemore AS, Kolonel AN, Wu AH et al (1995) "Prostate cancer in relation to diet, physical activity, and body size in blacks, whites, and Asians in United States and Canada." *JNCI* 87: 652-661.

Whittemore AS, Paffenbarger RS Jr, Anderson K et al (1985) "Early precursors of site-specific cancers in college men and women." *JNCI* 74: 43-51.

Wiklund K, Lindefars B-M, Holm LE (1988) "Risk of malignant lymphomas in Swedish agricultural and forestry workers." *Br J Ind Med* 45: 19-24.

Williams CJ (1990) "Lymphatic and haematological cancers." In: Williams C (ed) "*Cancer biology and management: an introduction.*" Chichester: John Wiley & sons.

Wilson PWF, Paffenbarger RS, Morris JN et al (1986) "Assessment methods for physical activity and physical fitness in population studies: report of an NHLBI workshop." *Am Heart J* 111: 1177-1192.

Witherspoon RP, Fisher LD, Schoch G et al (1989) "Secondary cancers after bone marrow transplantation for leukaemia or aplastic anaemia." *N Engl J Med* 321: 784-789.

Wolman JJ. (1956) *Q Rev Pediat* 11: 109.

Zahm SH, Hoover RN, Fraumeni JF (1995) "Hodgkin's Disease and parity (letter)." *Int J Cancer* 62: 362-363.

Zambon P, Simomato I, Penelli N et al (1981) "Descriptive epidemiology of Hodgkin's Disease in the province of Padova 1970-74." *Tumori* 67: 95-100.

Zheng W, Shu XO, McLaughlin JK et al (1993) "Occupational physical activity and the incidence of cancer of the breast, corpus uteri and ovary in Shanghai." *Cancer* 71: 3620-4.

Zhou X-G, Hamilton-Dutoit SJ, Yan Q-H et al (1993) "The association between Epstein-Barr virus and Chinese Hodgkin's disease." *Int J Cancer* 55: 359-363.

Zorate-Osorno A, Roman LN, Kingmo DW et al (1995) "Hodgkin's Disease in Mexico: prevalence of Epstein-Barr virus sequences and correlations with histologic subtype." *Cancer* 75: 1360-1366.

Zwitter M, Zakelj MP, Kosmelj K (1996) "A case-control study of Hodgkin's Disease and pregnancy." *Br J Cancer* 73: 246-251.

Appendix A

Wood-related occupations:

The occupation that has been most frequently investigated for an association with HD is woodworking. The results of case/control and cohort studies investigating this relationship have been reasonable consistent.

Six studies have described a statistically significantly elevated risk of HD occurring in those working with wood (Milham & Hesser, 1967; Petersen & Milham, 1974; Fonte et al, 1982; Acheson et al, 1984; Brownson & Reif, 1988; Wiklund et al, 1988) with relative risks of HD ranging from 1.8-7.2. However, in the Wiklund et al (1988) study the majority of the men in the exposed group (silviculture workers) were employed in administrative positions. This result, therefore, may be more related to social class rather than exposure to wood. A further study found a statistically significant relationship with mortality of HD in white males in states of the USA with the percentage of pine forest in each state (Spiers, 1969). However, in this study the proportion of people employed in the lumber industry was not a significant predictor of HD mortality. Five other studies have found a non-statistically significantly increased risk of HD associated with woodworking (Grufferman et al, 1976; Greene et al, 1978; Abramson et al, 1978; Kirchoff et al, 1980; McKinney et al, 1990). While the results of Abramson et al (1978) were not statistically significant for total HD working with wood was associated with a significantly increased risk of MC subtype (RR=5.2) when analysed separately. In the study by McKinney et al (1990) exposure to wood dust was found to increase risk of HD if exposure took place in a hobby setting but there was no relationship with occupational exposure to wood dust. This could be due to the use of protective equipment at work. Not all studies have found an elevated risk of HD associated with woodworking (Milham, 1974a; Olsen & Sabroe, 1979; Rang & Acheson, 1981; Miller et al, 1989; La Vecchia et al, 1989; Franceshi et al, 1991).

All authors have not used the same definition of woodworking. This could account for some of the discrepancies in results. Carpentry (Greene et al, 1978; Brownson & Reif, 1988), working with wood including paper (Petersen & Milham, 1974; Greene et al, 1978), membership of furniture making unions (Milham, 1974a; Olsen & Sabroe, 1979, Miller et al, 1989), and silviculture workers (Wiklund et al, 1988) have all been considered as the exposed group. These subjects would have had heterogeneous contact with wood.

In the aggregate the available data suggests that woodworkers have an increased risk of HD. This association may be due to exposure to wood dust, preservatives or particular chemicals within these industries.

Agriculture and food processing:

Farmers have an increased risk for all lymphatic and haematopoietic cancers. This can be seen when looking at only HD. Two case/control studies (La Vecchia et al, 1989; Franceshi et al, 1991) and one cohort study (Cerhan et al, 1998) have found a statistically significant increased risk of HD for agricultural workers. However, in one case/control study the exposed group included food processing workers (La Vecchia et al, 1989) and in the other the statistically significant effect was only seen in those employed in agriculture for over 10 years (Franceshi et al, 1991). A further four case/control studies (Pearce et al, 1985; Bernard et al, 1987; Brownson & Reif, 1988; McKinney et al, 1990) and three cohort studies (Kravdal & Hansen, 1993; Wiklund et al, 1988; Pukkala & Notkola, 1997) found an increased risk of HD but the results were not statistically significant. Pukkala & Notkola (1997) found the effect of agriculture was limited to men but Kravdal & Hansen (1993) found female farmers in their cohort to be at increased risk.

Increased risk in agricultural workers may possibly be due to exposure to animal viruses e.g. bovine leukaemia virus (Brownson & Reif, 1988). Indeed, livestock farmers, as compared with other farmers, show the highest OR for HD (Reif et al, 1989) while vets and meat workers (Johnson et al, 1986; Blair & Hayes, 1982; Pearce et al, 1988) share with farmers an increased risk of lymphoid neoplasms. This suggests that exposure to animals

may be involved. However, the results of studies designed to test the zoonotic hypothesis have been equivocal (Brownson & Reif, 1988). Alternatively exposure to herbicides and pesticides, especially phenoxy acids, may be responsible.

Appendix B

General health **Page 01**

Please can I ask for some details about your general health.

1. *CASES:* Prior to your diagnosis with Hodgkin's Disease have you ever suffered from, any of the following illnesses?

   CONTROLS: Prior to your _____ birthday have you ever suffered from, any of the following illnesses? 1=Yes 2=No

1=(Previous) cancer, leukaemia, lymphoma or tumour	☐	
2=Diabetes	☐	
If yes, do you take insulin? 1=Yes, 2=No		☐
3=Thyroid disease	☐	
If yes, was it 1=underactive,2=overactive		☐
4=Rheumatoid arthritis	☐	
5=Pernicious anaemia	☐	
6=Multiple sclerosis	☐	
7=Systemic lupus	☐	
8=Sjögren's syndrome	☐	
9=Immune disorders	☐	
10=Asthma	☐	
11=Eczema	☐	
12=Glandular fever	☐	
If yes, was it confirmed by a blood test?		☐
13=Chronic or recurrent infection - specify	☐	
14=Appendicitis	☐	
If yes, was your appendix removed?		☐
15=Epilepsy	☐	
16=Have you had your tonsils or adenoids removed?	☐	
17=Have you ever had a transplant?	☐	
18=Have you ever had a blood transfusion or received blood products?	☐	
19=Have you ever had a serious or unusual illness?	☐	

If yes to any of these please complete a record for each condition, enter the total number of illnesses below and say:

Now I shall ask you for further details about this illness / some of these illnesses.

Condition ☐☐ ICD Code ☐☐☐.☐

1. ***(For codes 1, 3, 8, 9, 13, 19)*** What exactly was wrong with you?

2. When was this? Year ☐☐ or Age ☐☐ years

3. Did you have treatment from 1=Hospital 2=GP 3=Neither ☐

If hospital:

3a. ***(Code 1 only)*** What treatment did you have? ☐

3b. ***(Code 18 only)*** Did you have 1=a transfusion of plasma or whole blood / 2=concentrated blood products? ☐

Condition ☐☐ ICD Code ☐☐☐.☐

1. ***(For codes 1, 3, 8, 9, 13, 19)*** What exactly was wrong with you?

2. When was this? Year ☐☐ or Age ☐☐ years

3. Did you have treatment from 1=Hospital 2=GP 3=Neither ☐

If hospital:

3a. ***(Code 1 only)*** What treatment did you have? ☐

3b. ***(Code 18 only)*** Did you have 1=a transfusion of plasma or whole blood / 2=concentrated blood products? ☐

Condition ☐☐ ICD Code ☐☐☐.☐

1. ***(For codes 1, 3, 8, 9, 13, 19)*** What exactly was wrong with you?

2. When was this? Year ☐☐ or Age ☐☐ years

3. Did you have treatment from 1=Hospital 2=GP 3=Neither ☐

If hospital:

3a. ***(Code 1 only)*** What treatment did you have? ☐

3b. ***(Code 18 only)*** Did you have 1=a transfusion of plasma or whole blood / 2=concentrated blood products? ☐

Prior to *CASES:* your diagnosis with Hodgkin's Disease.
CONTROLS: your _____ birthday.

Have you had any of the following infections?

	1=Yes 2=No 9=NK	If Yes, when?			If NK school stage	If Yes, when?			If NK school stage
		Mth	Yr	or Age		Mth	Yr	or Age	
Measles									
Mumps									
German Measles									
Whooping Cough									
Chicken Pox									
Shingles									
Pneumonia									
Pneumonia									
Herpes Simplex I (cold sores)									
Herpes Simplex II (Genital Herpes)									
Any other serious infectious disease - specify ICD									
Any other serious infectious disease - specify ICD									

The next series of questions apply to your parents, full or half brothers and sisters, children and anyone who lived in the same household while you were young. I shall call these 'relatives'. We already have lists of these people.

INTERVIEWER: GO QUICKLY THROUGH FORMS ON PAGES ------- AND READ OUT THE NAMES OF THE CANDIDATE RELATIVES.

Have any of these relatives ever suffered from any of the following illnesses:

		Yes=1 No=2	No. of Relatives	Names
1.	Cancer, leukaemia, Hodgkin's Disease lymphoma or tumours	☐	☐☐	_____
2.	Multiple sclerosis	☐	☐☐	_____
3.	Glandular fever	☐	☐☐	_____
4.	Diabetes	☐	☐☐	_____
5.	Thyroid disease	☐	☐☐	_____
6.	Rheumatoid arthritis	☐	☐☐	_____
7.	Pernicious anaemia	☐	☐☐	_____
8.	Systemic lupus	☐	☐☐	_____
9.	Sjögren's syndrome	☐	☐☐	_____
10.	Immune disorders	☐	☐☐	_____
11.	Inherited disease	☐	☐☐	_____
12.	Unusual illness	☐	☐☐	_____
13.	Anything else which runs in the family	☐	☐☐	_____

IF YES TO ANY OF THESE PLEASE COMPLETE A RECORD FOR EACH CONDITION AND ENTER THE TOTAL NUMBER OF ILLNESSES BELOW AND SAY:
I shall now ask you details of each of these.

Total number of illness records following ☐☐

Please complete section for each relative - illness combination in Section 6a page 1.

Name of Relative: _____ Relative No: ☐☐

Condition: _____ Code: ☐☐

1. *(Omit for codes 2, 3, 6, 7, 8)*

What exactly was wrong with them?_____ ☐☐☐.☐

2. When was it first diagnosed? Year 1 9 ☐☐ or Age ☐☐ yrs

3. Were they treated in hospital? 1=Yes 2=No 9=NK ☐

Name of Relative: _____ Relative No: ☐☐

Condition: _____ Code: ☐☐

1. *(Omit for codes 2, 3, 6, 7, 8)*

What exactly was wrong with them?_____ ☐☐☐.☐

2. When was it first diagnosed? Year 1 9 ☐☐ or Age ☐☐ yrs

3. Were they treated in hospital? 1=Yes 2=No 9=NK ☐

Name of Relative: _____ Relative No: ☐☐

Condition: _____ Code: ☐☐

1. *(Omit for codes 2, 3, 6, 7, 8)*

What exactly was wrong with them?_____ ☐☐☐.☐

2. When was it first diagnosed? Year 1 9 ☐☐ or Age ☐☐ yrs

3. Were they treated in hospital? 1=Yes 2=No 9=NK ☐

Appendix C

Sport and Athletics **Page 01**

This section is about your sporting and athletic activities while you were at school.

1. Did you ever belong to a school team for any sport or athletics? ☐
   1=Yes 2=No

If yes: 1a. Which activity or sport(s)?_____

2. At that time did you represent a sports club or similar in competitions? ☐
   1=Yes 2=No

If yes: 2a. Which activity or sport(s)?_____

If there is no "yes" answer to question 1 or 2, go to next page.

3. Did you compete at regional or national level? 1=Yes 2=No ☐

4. I shall now ask for more details of your main sporting activities at different stages of schooling. [You have mentioned...but can you choose the five most important].

ACTIVITIES	AGE 7-11 years	AGE 12-14 years	AGE 15-18 years
Activity 1	Hrs:	Hrs:	Hrs:
Code:	1=H 2=S 3=O	1=H 2=S 3=O	1=H 2=S 3=O
MET Score:			
Activity 2	Hrs:	Hrs:	Hrs:
Code:	1=H 2=S 3=O	1=H 2=S 3=O	1=H 2=S 3=O
MET Score:			
Activity 3	Hrs:	Hrs:	Hrs:
Code:	1=H 2=S 3=O	1=H 2=S 3=O	1=H 2=S 3=O
MET Score:			
Activity 4	Hrs:	Hrs:	Hrs:
Code:	1=H 2=S 3=O	1=H 2=S 3=O	1=H 2=S 3=O
MET Score:			
Activity 5	Hrs:	Hrs:	Hrs:
Code:	1=H 2=S 3=O	1=H 2=S 3=O	1=H 2=S 3=O
MET Score:			

5. Did your sporting activities include touring for competitions or training camps? ☐
   1 = Yes 2 = No 9 = NK

If yes: 5a. How many nights were spent away per year?

While aged 7-11 years ☐☐☐

12-14 years ☐☐☐

15-18 years ☐☐☐

5b. How many other people stayed with you? ☐
   1=none / 2=1or2 / 3=3-5 / 4=6-10 / 5=more than that.

Current age _____

1. Since leaving school, have you ever taken part in competitive sport or athletics? 1 = Yes 2 = No ☐

If yes: 1a. What activity or sport(s)? _____
If No, go to next page.

2. Did you compete at regional or national level? 1=Yes 2=No ☐

3. I shall now ask for more details of your main sporting activities since you left school. [You have mentioned.... but can you choose the five most important].

ACTIVITIES	Leaving School-21 years	AGE 22-29 years	AGE 30-39 years
Activity 1	Hrs:	Hrs:	Hrs:
Code:	1=H 2=S 3=O	1=H 2=S 3=O	1=H 2=S 3=O
MET Score:			
Activity 2	Hrs:	Hrs:	Hrs:
Code:	1=H 2=S 3=O	1=H 2=S 3=O	1=H 2=S 3=O
MET Score:			
Activity 3	Hrs:	Hrs:	Hrs:
Code:	1=H 2=S 3=O	1=H 2=S 3=O	1=H 2=S 3=O
MET Score:			
Activity 4	Hrs:	Hrs:	Hrs:
Code:	1=H 2=S 3=O	1=H 2=S 3=O	1=H 2=S 3=O
MET Score:			
Activity 5	Hrs:	Hrs:	Hrs:
Code:	1=H 2=S 3=O	1=H 2=S 3=O	1=H 2=S 3=O
MET Score:			

4. Did your sporting activities include touring for competitions or training camps? 1=Yes 2=No 9=NK ☐

If yes: 4a. How many nights were spent away per year?

While aged Leaving school-21 years ☐☐☐

22-29 years ☐☐☐

30-39 years ☐☐☐

4b. How many other people stayed with you? ☐
1=none / 2=1or2 / 3=3-5 / 4=6-10 / 5=more than that.

1. Apart from what we have discussed have you regularly done any other strenuous physical activity, at school or later to...
   your 40th birthday? *(older respondents)* 1=Yes
   your diagnosis with Hodgkin's Disease? *(younger CASES)* 2=No ☐
   your __th birthday? *(younger CONTROLS)* 9=NK

If No go to next page.

If yes: 1a. What activities?_____

*End age:*____

2. I shall now ask for more details of these activities.
   [You have mentioned.... but can you choose the five most important].

ACTIVITIES	AGE 7-11 years	AGE 12-14 years	AGE 15-Leaving School	Leaving School - 21 years	AGE 22-29 years	AGE 30-39 years
Activity 1	Hrs:	Hrs:	Hrs:	Hrs:	Hrs:	Hrs:
Code:	1=H 2=S 3=O	1=H 2=S 3=O	1=H 2=S 3=O	1=H 2=S 3=O	1=H 2=S 3=O	1=H 2=S 3=O
MET Score:						
Activity 2	Hrs:	Hrs:	Hrs:	Hrs:	Hrs:	Hrs:
Code:	1=H 2=S 3=O	1=H 2=S 3=O	1=H 2=S 3=O	1=H 2=S 3=O	1=H 2=S 3=O	1=H 2=S 3=O
MET Score:						
Activity 3	Hrs:	Hrs:	Hrs:	Hrs:	Hrs:	Hrs:
Code:	1=H 2=S 3=O	1=H 2=S 3=O	1=H 2=S 3=O	1=H 2=S 3=O	1=H 2=S 3=O	1=H 2=S 3=O
MET Score:						
Activity 4	Hrs:	Hrs:	Hrs:	Hrs:	Hrs:	Hrs:
Code:	1=H 2=S 3=O	1=H 2=S 3=O	1=H 2=S 3=O	1=H 2=S 3=O	1=H 2=S 3=O	1=H 2=S 3=O
MET Score:						
Activity 5	Hrs:	Hrs:	Hrs:	Hrs:	Hrs:	Hrs:
Code:	1=H 2=S 3=O	1=H 2=S 3=O	1=H 2=S 3=O	1=H 2=S 3=O	1=H 2=S 3=O	1=H 2=S 3=O
MET Score:						

3. Did these activities involve residential training or touring ?
   1 = Yes 2 = No 9 = NK ☐

If yes: 3a. How many nights were spent away/year on average?

While aged
7-11 years ☐☐☐ Leaving school-21 years ☐☐☐
12-14 years ☐☐☐ 22-20 years ☐☐☐
15-Leaving school ☐☐☐ 30-39 years ☐☐☐

3b. How many other people stayed with you? ☐
1=none / 2=1or2 / 3=3-5 / 4=6-10 / 5=more than that.

This section considers your more recent physical activity. In the 5 years up to....
either your diagnosis with Hodgkin's Disease.
or your ___ birthday.

1. How much exercise did you get from your recreational activities in a typical week?

   1=a lot / 2=Moderate amount / 3=Little or none. ☐

2. Did you regularly engage in any physical activity such as walking, jogging, cycling for long enough to work up a sweat?

   1=Yes 2=No ☐

If Yes: 2a. How often?

   1=at least twice a week / 2=weekly / 3=less often. ☐