# Socioeconomic inequalities in lung and upper aero-digestive tract cancer incidence in Scotland

Katharine H Sharpe

BA (hons), MBA

Submitted in fulfilment of the requirements for the

Degree of Doctor of Philosophy


School of Medicine, Dentistry and Nursing

University of Glasgow

February 2018

# Abstract

Socioeconomic inequality in cancer risk and incidence burden has received limited attention compared to genetic and behavioural risk factors. Where they have been studied, the temporal relationship between socioeconomic factors and cancer risk has been under explored due to the mainly cross-sectional nature of most research. Moreover, the inter-relationships of the multiple measures of socioeconomic status and, in particular, area and individual measures and their interaction with risk behaviours have also had limited attention. The overarching aim of this thesis was to investigate socioeconomic inequalities in the risk of lung and upper aero-digestive tract cancers and the relationship between this risk and socioeconomic status, area and individual based measures of socioeconomic circumstances, and behaviours such as smoking, alcohol consumption, diet and exercise.

To understand and quantify the relative contribution by age, sex and tumour subtype to the socioeconomic inequalities of all cancer risk, a descriptive epidemiological study of cancer incidence in Scotland (2000-07) was undertaken. Age standardised rates per 100,000 population were calculated by direct standardisation to the European standard. A linear regression model was used to calculate the Slope Index of Inequality (SII) and Relative Index of Inequality (RII) which were employed to rank tumour and subtype contribution to all cancer risk socioeconomic inequalities by age for each sex for lung and upper aero-digestive tract (UADT) cancers separately. There were 216,305 cases excluding non-melanoma skin cancer (all cancer) comprising 37,274 lung, 8,216 head and neck and 6,534 oesophageal cancers classified into anatomical or morphological subtypes. Socioeconomic circumstances were measured using the Scottish Index of Multiple Deprivation (SIMD). Analyses were partitioned by five-year age group and sex. RII was adapted to rank the contribution of each tumour type to all cancer socioeconomic inequalities and to examine subtype by age and sex simultaneously. The rank was defined as the proportion of all cancer socioeconomic inequality.

All cancer socioeconomic inequality was greater for males than females (RII=0.366; female RII=0.279). The combination of lung and UADT socioeconomic inequalities

contributed 91% and 81% respectively to all cancer socioeconomic inequality. For both sexes lung and UADT subtypes showed significant socioeconomic inequalities (*P*<0.001) except oesophageal adenocarcinoma in males (P=0.193); for females, socioeconomic inequality was borderline significant (P=0.048). Although RII rank differed by sex, all lung and larynx subtypes contributed the most to all cancer socioeconomic inequality with RII rank for oral cavity, oesophagus–squamous cell and oropharynx following. For males 40-44 years old, socioeconomic inequalities increased abruptly peaking at 55-59 years. For females, socioeconomic inequalities gradually peaked 10 years later. In both sexes, the socioeconomic inequalities peak age preceded age of peak incidence. This study showed that socioeconomic inequalities in lung and UADT cancers vary greatly by age, tumour subtype and sex; these variations were likely to largely reflect differences between the sexes in risk behaviours which vary by birth cohort and are socioeconomically patterned.

Longitudinal data enabled exploration of the temporal relationship between socioeconomic status and cancer incidence. An investigation of several individual and a single area-based measure of socioeconomic circumstances was undertaken in the second study of this thesis. The effect of country of birth, marital status, one area socioeconomic circumstances measure (Carstairs) and five individual socioeconomic variables (economic activity, education, occupational social class, car ownership, household tenure) on the risk associated with lung, UADT and all cancer combined (excluding non-melanoma skin cancer) were explored. A linked dataset using the Scottish Longitudinal Study and Scottish Cancer Registry was created to follow 203,658 cohort members aged 15+ years from 1991–2006. Relative risks (RR) were calculated using Poisson regression models by sex offset for person-years of follow-up. There were 21,832 first primary tumours (including 3,505 lung and 1,206 UADT cancers). Regardless of cancer, economic inactivity (versus activity) was associated with increased risk (male: RR 1.14 95% CI 1.10, 1.18; female: RR 1.06 95% CI 1.02, 1.11). For lung cancer, area deprivation remained significant after full adjustment suggesting that the area deprivation cannot be fully explained by individual variables. Not having a qualification (versus degree) was associated with increased lung cancer risk; likewise for UADT cancer risk (females only). Occupational social class associations were most pronounced

and elevated for UADT risk. No car access (versus ownership) was associated with increased risk (excluding all cancer risk for males). Renting accommodation (versus home ownership) was associated with increased lung cancer risk, UADT cancer risk for males only and all cancer risk for females only. Regardless of cancer group, elevated risk was associated with no education and living in deprived areas. This study demonstrated that different and independent socioeconomic variables were inversely associated (greater incidence with lower socioeconomic circumstances) with different cancer risks in both sexes; no one socioeconomic variable had a dominant risk association or captured all aspects of socioeconomic circumstances or the full life-course. The association of multiple socioeconomic variables was likely to reflect the complexity and multifaceted nature of low socioeconomic circumstances as well as the various roles of these dimensions over the life-course.

A final study investigated the role of behaviours (smoking, alcohol, diet and exercise) on the association of low socioeconomic circumstances with all cancer risk and lung and upper aero-digestive tract cancers combined (LUADT). The Scottish Cancer Registry and Scottish Health Survey data were linked to create a population study (1995-2011). There were 42,983 adults over 16 years old who were followed for 3,750,611 person-years. There were 2,130 first primary cancers diagnosed including 453 LUADT cancers. Poisson regression models, minimally adjusted by age and sex, were developed to estimate the risk association between five individual socioeconomic variables (economic activity, highest qualification, occupational social class, car ownership and housing tenure), one area-based socioeconomic indicator (SIMD) and all cancer and LUADT cancer. A further socioeconomic indicator was developed to reflect multiple low socioeconomic circumstances. This was defined as the count, at the individual participant level, of socioeconomic variables in the highest risk category. A similar multiple high risk behaviour derived variable, defined as the count of highest risk category for the following variables: current smoking status, units of alcohol consumed in a week, daily fruit and vegetable consumption and exercise sessions per week, was also calculated at the individual participant level. The minimally adjusted Poisson models were successively adjusted for behaviours (smoking, alcohol, diet and exercise) to establish any remaining contribution to cancer risk not explained by

behaviour. Multiple low socioeconomic circumstances were very strongly associated with increased risk for both cancer groups. For all cancer risk, the elevated risk was nearly fully attenuated for all categories of multiple low socioeconomic circumstances when adjusted for smoking only. For LUADT cancer and in the minimally adjusted model, the risk increased in a dose-response manner. The risk associated with LUADT cancer for study participants in the highest category of multiple low socioeconomic circumstances was more than three-times greater when compared to their affluent counterparts (RR 3.35 95% CI 2.26, 4.97); this elevated risk remained at 86% compared to those with no socioeconomic disadvantage, even after full adjustment for smoking, alcohol, diet and exercise behaviours. When looking at single socioeconomic status (SES) indicators, only those who rented accommodation from a local authority remained at a 50% increased risk of LUADT cancer even after adjustment for all the behaviours (RR 1.50 95% CI 1.05, 2.16). This study demonstrated that smoking is a major inequality issue and a significant cancer risk which is socially patterned.

Further analytical research is required to fully understand the pathways and mechanisms between socioeconomic circumstance and lung and upper aero-digestive cancer risk. This thesis suggests that when monitoring socioeconomic inequalities and cancer risk, it is less effective to focus on all cancer as a group given the mix of diseases resulting from very different aetiological processes, some associated with high SES and others with low SES. It also suggests that both individual and area measures of SES are valid measures and are required to capture the multi-dimensional nature of SES as well as the life-course and intergenerational implications of SES. In addition to this “multi-dimensional” attribute to SES, it is essential to consider multiple low social circumstances occurring simultaneously and therefore compounding vulnerability to cancer risk. Behaviours, particularly smoking and alcohol, explained much of the elevated lung and upper aero-digestive tract cancer risk for individual SES indicators. Clearly, in this context, smoking is a major inequality issue and a significant cancer risk.

This thesis provides useful insights for raising the issue of inequalities in cancer, for advocacy and for building policy and interventions to tackle inequalities in cancer incidence. Policies need to focus on more broadly upstream causes.

Traditionally, these policies have been focused on downstream behaviours (e.g. public space smoking ban and alcohol minimum pricing), but upstream policies that take on the fundamental political decisions regarding the distribution of income, wealth and power are required at both Westminster and Holyrood and beyond.

# Table of Contents

# List of Tables

# List of Figures and Equations

# List of accompanying material

# Preface

I am not the same person I was. The events of the 15th July 2014 were a watershed — both literally and figuratively.

It was the day that I was diagnosed with invasive lobular carcinoma (ILC) after a regular screening mammogram performed the month before was deemed "all clear". This is not unheard of for lobular carcinoma. I learned, through application of my PhD skills, that the (Indian file) cellular structure of ILC (Oliveira et al 2014) can be undetectable via a mammogram (McCart Reed et al 2015); that is, I had been diagnosed with an "occult" tumour. ILC can mimic normal breast cell structure (McCart Reed et al 2015), making it only visible via an MRI (Oliveira et al, 2014). And because of loss of the E-caderin protein (McCart Reed et al, 2015), cell-to-cell adhesion is also lost; therefore, ILC does not form a lump but an "Indian file" infiltrating normal breast cells. This accounts for the low sensitivity of mammography and underestimation of tumour size (Oliveira et al, 2014). Tumours as large as 5cm (in my case 8cm) can be missed in mammography if they have similar density to the normal parenchyma breast tissue (Oliveira et al, 2014). As a result, it is often undetected until it is quite extensive (McCart Reed et al, 2015) and it is often not diagnosed until it has developed to more advanced stages (Oliveira et al, 2014).

Needless to say, I wasn't expecting that. I was "healthy" having led my life carefully with respect to smoking, alcohol consumption, diet, exercise, etc. I had envisioned myself at 85 years old, largely unscathed by the chronic illnesses others would face due to my "healthy" lifestyle. I saw myself as much younger than my chronological age. A breast cancer diagnosis at 55 years old was a devastating shock.

After a three year enforced break and given my age, why did I want to complete my PhD? I was five and a half years into a six year (part-time) programme, but it was not because I am a "completer-finisher" and not because of the "self-esteem" a PhD would bestow, but because of the way that I saw this phase of my life had ended. My career as an Information Consultant in the NHS had been brutally

interrupted and its end was not of my choosing; I had not been “in control”. I am clear that I have already benefited from the PhD. The skills I learned at the University of Glasgow have been literally priceless to me as I went through three years of diagnosis, treatment and recovery. I know, without a doubt, that the analytical and critical skills developed and fine-tuned by the PhD training enabled me to understand my diagnosis; to respond to it empowered with a knowledge most patients would not have had; to interact with my medical and nursing team with confidence; and, most importantly, to participate in and shape my treatment and recovery most effectively. I am in no doubt that that the knowledge gained via my PhD studies, added to my experience in health information and management, made a difference.

I have always “pushed the envelope out” in terms of the boundary of my personal comfort zone — despite the anticipatory anxiety — and been able to (eventually) overcome that anxiety. Now, I discover that the position of that envelope has moved; it is closer than it was before, limiting my capacity but not removing that capacity, reflecting the effect of those events since 15th July 2014.

Nevertheless, I have managed, to this point, well — not only because of my own tangible and intangible resources, but because of the resilient safety net that I had around me. My husband in particular was (and continues to be) an unbelievable, unrelenting source of strength – I was very lucky. My supervisors’ flexibility, empathy and patience allowing me to take three years, the time I personally needed, to be ready to return to and complete my PhD was essential to reach this point today. And with the support of Scotland’s National Health Service — I am very fortunate here too. I know this because I have a direct comparison. Having grown up in America, not only do I have experience of using the US health care system, but I began my career in healthcare in the American system. Furthermore and more importantly, my sister, four years younger and living in Phoenix, Arizona was diagnosed just two months after me. Hearing my story of an all clear mammogram and with my urging, she had an MRI examination, and unfortunately was diagnosed with practically the same stage of ILC as me. Our diagnoses and treatments were virtually identical: chemotherapy, bi-lateral mastectomy, full axial removal, radiotherapy and on-going hormone therapy.

How that was delivered and our resulting experiences, however, could not be more contrasting. My sister had to work throughout her treatment because she would not have otherwise had health insurance to pay part of the $120,000 fee for her bi-lateral mastectomy. Her chemotherapy treatment included 10 sessions delivered bi-weekly exposing her to greater costs and more side effects than the six cycles delivered three-weekly that I had here in Scotland with the knowledge that evidence-based SIGN guidelines (SIGN 2013) had established cancer outcomes were equivalent for both treatment regimes.

So you may ask…what does all this have to do with a PhD in socioeconomic inequalities in the risk of lung and aero-digestive tract cancers in Scotland? A lot actually, as will be disclosed through the next chapters, the interaction between society and the individual is critical to the inequalities in health outcomes, including cancer.

One famous UK Prime Minister, Mrs. Margaret Thatcher, stated in an interview for a women's magazine, "There is no such thing as society" (Women's Own 1987). As a country, I believe we are on the edge of a precipice and at the very regrettable risk of "throwing the baby out with the bath water" (—realising Thatcher's vision). Recently Macmillan Cancer Support reported that, in the UK, a cancer diagnosis was more common than getting married or having your first baby (Macmillan Cancer Support 2017). What their report didn't say was that diagnosis was more likely to occur in those who are socioeconomically disadvantaged — the incidence of most cancers is socially patterned. The question remains — why?

Although, breast cancer is one of a few cancers where the socioeconomic pattern for diagnosis does not follow the typical pattern of increased incidence among those with low socioeconomic circumstances (Faggiano F 1997), it does revert to type in survival (Kogevinas et al 1997b). As I reflect on my own experience, my education empowered me to handle and cope with my cancer journey most effectively. Because I had a solid education, my financial resources, knowledge and skills enabled me to optimise my situation (however adverse or privileged) throughout my life (the "life-course" in the literature). My husband and our combined resources (mostly his) made it possible for me to stop work once I was

diagnosed so that we could focus on optimising my recovery. I reflect on my very fortunate (socioeconomic) circumstances and know that I am indeed lucky; I have managed to regain some control over my life. As will be discussed through this PhD, "loss of control" is recognised as one of the pivotal elements in the manifestation of stress and its role in cancer is being recognised (Behrens et al 2016). What is clear to me is that as society is eroded, division is widened and positions become entrenched; life will not get better for any of us, but particularly for those who are struggling with everyday issues while facing socioeconomic disadvantage.

Fortunately, there is a whole section of those in society who is aware of the increased risk association of poorer health with poverty. Many are working to raise the profile of health inequalities such that not only is there an understanding among experts of why they occur (WHO 2011), but an appreciation among the general public (von dem Knesebeck et al 2017) and among politicians — for socioeconomic circumstances and inequality are ultimately a political decision — of the importance of understanding and addressing the underlying causes of those inequalities to the benefit of everyone (Peres et al 2017).

It — socioeconomic inequalities — are all relative: within oneself, in terms of the consistency between actual socioeconomic position and one's expectations for oneself (described as "status inconsistency" in the literature (Behrens et al 2016)), and one's relative position compared to others (Uphoff et al 2013). The challenge is to recognise this and focus on a better understanding of the problem; i.e., cancer incidence inequalities and how the current flow of direction can be arrested, or even better, reversed. I hope that this thesis is able to make a contribution, however small, to that objective.

# Acknowledgement

Grateful thanks are due to a number of organisations and individuals. In particular I would like to thank the following for their help, time and efforts.

I must first acknowledge the significant contributions of Professor David Conway and Dr Alex McMahon at the University of Glasgow for their guidance, advice and direction throughout. It is actually quite difficult to enumerate all facets of their support during what has resulted in a ten year period, some phases of which have been quite challenging. Both David and Alex were there to help particularly when needed most and we have enjoyed some healthy political debates during these quite turbulent times.

I am grateful to my third supervisor, Dr David Brewster (until my early retirement and then his recent retirement, just before I returned to my studies), from ISD, who first introduced me to the possibility of doing a PhD and provided helpful direction given his knowledge of cancer in general, epidemiology and the Scottish Cancer Registry.

I would also like to specifically thank Pauline Daniel at the University of Glasgow, Dental School for her help and persona — her assistance was always delivered with a welcome smile and with genuine warmth. I must also acknowledge Professor Lorna Macpherson, my advisor, for her valuable recommendations, advice and assistance in her role of annual reviews. She provided a constant, calm, considered "arms-length" perspective.

Chapter 3 and 4 – I am grateful to Paula McClements, Chris Watling, Chris Povey, Cath Storey, Scottish Cancer Registry and ISD Scotland. All of these individuals played a significant role in this thesis for they diligently provided the Scottish Cancer data necessary for analysis and linkage. Without their essential help, none of this thesis, in terms of the data, would have been possible.

For Chapter 4, which was dependent on using the Scottish Longitudinal Study (SLS) linkage study, I must thank Professor Paul Boyle, Professor Gillian Raab (my SLS

advisor), Claire Boag (the SLS administrator) and Dr Z Feng who also supported my visits when Professor Raab was not available. I enjoyed their friendship, consistent support and guidance throughout the period I was working on the SLS study and beyond.

In respect of Chapter 5, which was focused on the Scottish Health Survey and Scottish Cancer Registry linkage, I am very grateful for the contributions of Dr Janet Murray, Carole Morris and Leanne Hopkins at eDRIS, ISD Scotland; all of whom assisted in making the final study included in this thesis feasible in terms of approvals and data. Once again, their support and assistance was significant and much appreciated.

In light of the events of July 2014 and their continuing influence on how I proceed with life, I also would like to acknowledge the contributions of Doctors Wright and Porteous, both of whom have gone beyond the normal medical remit and have provided me with insight to the process of completing a PhD which has been very valuable to me.

Finally, I must mention my husband, Alan Sharpe, who has helped me throughout this decade long experience. Without his support, provided through thick and thin, it is very conceivable that I would not have had the determination to return to the PhD or the stamina to see the thesis to its final stage. For his encouragement and belief in me, I am eternally grateful.

# Author's Declaration

Parts of the research work included in this thesis have been published with co-authors including:

Chapter 3

Sharpe, K. H., McMahon, A. D., McClements, P., Watling, C., Brewster, D. H., & Conway, D. I. 2012, Socioeconomic inequalities in incidence of lung and upper aero-digestive tract cancer by age, tumour subtype and sex: A population-based study in Scotland (2000-2007). Cancer Epidemiology; 36: E164-E170.

Chapter 4

Sharpe, K. H., McMahon, A. D., Raab, G. M., Brewster, D. H., & Conway, D. I. 2014, Association between socioeconomic factors and cancer risk: a population cohort study in Scotland (1991-2006). PLoS ONE; 9: e89513.

I declare that this thesis is my own composition and has not been submitted in part or in whole for any other degree.

Katharine H Sharpe

Glasgow, February 2018

# Abbreviations

| | |
|---|---|
| BMI | Body Mass Index |
| CHI | Community Health Index |
| CI | Confidence Interval |
| CSDH | Commission on Social Determinants of Health |
| EDRC | Economic Development and Review Committee of the OECD |
| EMBASE | Excerpta Medica dataBASE |
| EOR | Excess Odds Ratio |
| EU | European Union |
| GERD | Gastro (O)esophageal Reflux Disease |
| HDI | Human Development Index |
| HNC | Head and Neck Cancer |
| HPV | Human Papillomavirus |
| IARC | International Agency for Research on Cancer |
| ICOHIRP | International Centre for Oral Health Inequalities and Research Policy |
| I-E | Income-Employment |
| IGF-1 | Insulin-like Growth Factor |
| IMD | Index of Multiple Deprivation |
| IMF | International Monetary Fund |
| INHANCE | International Head and Neck Cancer Epidemiology Consortium |
| IRR | Incidence Rate Ratio |
| ISD | Information Services Division |
| LUADT | Lung and upper aero-digestive tract |
| NRS | National Records Scotland |
| OECD | Organisation for Economic Cooperation and Development |
| OR | Odds Ratio |
| RII | Relative Index of Inequality |
| RR | Relative Risk |
| SCR | Scottish Cancer Registry |
| SES | Socioeconomic Status |
| SHeS | Scottish Health Survey |
| SII | Slope Index of Inequality |
| SIMD | Scottish Index of Multiple Deprivation |
| SLS | Scottish Longitudinal Study |
| SLS-DSU | Scottish Longitudinal Study Development & Support Unit |
| UADT | Upper aero-digestive tract |
| UK | United Kingdom |
| US(A) | United States (of America) |
| WCRF/AIRC | World Cancer Research Fund/American Institute for Cancer Research |
| WHO | World Health Organisation |

# 1 Introduction and literature review — Part I: background and context

## 1.1 Thesis structure

This thesis investigates the association of socioeconomic status (SES) with the incidence of cancer in Scotland. Chapter 1 Part I sets out the context, the background to SES and its definition, describes the key indicators of SES, including different approaches to measuring socioeconomic inequalities and defines the cancers of focus and their behavioural risk factors. Chapter 1 Part II provides a detailed narrative literature review of the evidence of inequalities in cancer incidence, indentifies the debates in the literature, provides the rationale for the PhD studies and sets out the aims of this thesis.

Three studies were conducted to investigate different angles of the SES and cancer incidence relationship. Chapter 3 assesses the association of cancer incidence by sex, age, cancer site and morphology using the Scottish Cancer Registry and an area-based measure of SES. Chapter 4 explores the differential association of several individual and area measures of SES with cancer incidence using a prospective cohort created through record linkage between the Scottish Cancer Registry and the Scottish Longitudinal Study. Chapter 5 examines the extent that behaviour factors may explain socioeconomic inequalities via another prospective data linkage cohort including behaviour and socioeconomic factors through record linkage of the Scottish Cancer Registry and the Scottish Health Survey.

Chapter 6 gathers together the findings of this thesis; examines possible causes of the socioeconomic inequalities in cancer incidence in Scotland; and discusses the thesis strengths and limitations. Chapter 6 also draws conclusions by making recommendations for further research and for approaches to tackling inequalities in cancer incidence in Scotland and beyond. Supporting information regarding the datasets used, the necessary ethics approvals attained and the data management required are provided in the Appendices.

## 1.2 Socioeconomic inequalities in health

The association of SES and health is well established and shows a consistent pattern of poorer health with poorer socioeconomic circumstances (Marmot 2005; Mackenbach et al 2008). This socioeconomic gradient reflects the social pattern of disease across all groups in society and in the social strata. This relationship exists in lower- and middle-income (World Bank Group 2016) countries (Bangal R et al 2014) and in high and middle-income countries (Arnold et al 2016). It also persists within and between countries (Mackenbach et al 2008) suggesting that there is not an absolute level of poverty associated with poor health but a linear relationship — a "gradient" between socioeconomic circumstances and health (Watt 2002; Kawachi et al 2006)

Given the consistent and pervasive nature of this stepwise socioeconomic gradient, a wide range of diseases including cancer (Marmot 2005) have a far larger burden of incidence among the lower socioeconomic groups relative to the higher socioeconomic groups (Watt et al 2012). The relationship between SES and ill health is so well established that epidemiologists would almost always adjust by SES in the same way they adjust for age and sex when exploring the effect of other risk factors for a disease (Kawachi et al 2006).

The World Health Organisation (Solar et al 2010) developed a framework for understanding the pathways and mechanisms that socioeconomic circumstances affect health; these different theories are not mutually exclusive but assist in providing explanations for socioeconomic inequalities in health. The persistent gradient is often the basis of the "social mobility" explanation of health inequalities. Sweeting et al (2015) explored causes of socioeconomic inequalities and discussed two possible explanations. Health selection may create this gradient as poorer health is associated with downward social mobility. Alternatively or in addition, occupational, educational and power create "social causation" of health inequalities by influencing health via material or cultural processes (Sweeting et al 2015). Solar (2010) pointed out that in general, the literature on health and social mobility suggested that health status influenced

subsequent social mobility; however, it is inconsistent across different life stages. “Social causation” is often identified as an explanation of socioeconomic inequalities in health and as Solar (2010) explained, reflects underlying differences in distribution of health determinants; such as behaviour, material, environmental, psychosocial and biological factors, across socioeconomic groups. The “life-course theory” recognises the temporal nature of the causal link between exposures and outcomes and reflects the role of social determinants on health throughout life where there are “critical periods” of susceptibility in life, such as periods of fast development or significant change. Exposure to low socioeconomic circumstances during these “critical” periods has long term or latent detrimental effects on biological functions; often referred to as “biological programming”. A further aspect of the “life-course theory” in the literature discussed as the “accumulation of risk” where the ill effects of exposure to health determinants (e.g. high risk behaviours, poor material, environmental and psychosocial factors) over time accumulate gradually. In this theory, increasing intensity, frequency and duration of exposure was logically assumed to lead to increased biological system damage. An associated concept is the “chain of risk” theory where an earlier exposure to one type of low socioeconomic circumstances leads to further and potentially different types of exposures later in life (Solar et al 2010).

A review of health inequalities in Western Europe, where health and public social services were considered developed and relatively progressive, identified that socioeconomic inequalities persisted and were frequently substantial despite these relatively ”liberal” welfare policies (Mackenbach 2012). Using an index of health and social problems which included factors such as: life expectancy, maths & literacy, infant mortality, homicide rate, imprisonment, teenage births, trust, obesity, mental illness and social mobility; Wilkinson and Pickett (2009) compared the health and social position relative to income equality across several European and non-European countries. They discovered that these health and social problems were worse in more unequal countries, with the UK consistently identified as having greater inequality and greater health and social problems (Wilkinson et al 2009).

In the United States (US), one study indicated that income inequality was associated with lack of social trust, which in turn was associated with higher age adjusted mortality rates from various chronic illnesses including cancer (Kawachi et al 1997). Generally, despite major shifts in the cause of death over many decades, the socioeconomic gradient in health has remained stagnant (Watt et al 2012). The concept of social capital is introduced briefly here as social trust; social capital is fundamental to understanding socioeconomic inequalities and its association with cancer risk. Section 1.3.1.2 explores this concept in more depth.

Due to these socioeconomic inequalities, overall life expectancy and disease-free life expectancy are considerably shorter among more socioeconomically deprived groups relative to more affluent groups (Marmot 2005; Mackenbach 2006; Mackenbach 2012). Moreover, despite advances in understanding behavioural risk factors, earlier detection of cancer and improving treatments, socioeconomic inequalities in cancer are observed across the cancer continuum. Incidence, morbidity, treatment, mortality and survival persist, and in some cases are widening with improvements in disadvantaged groups falling behind the more affluent groups (Faggiano F 1997; Kogevinas et al 1997a).

Socioeconomic factors are recognised as profound contributors to health inequalities in and of themselves even after adjustment for behaviours.

## 1.3 Socioeconomic status

### 1.3.1 Definitions

Socioeconomic status incorporates concepts developed by Karl Marx and Max Weber. Marx identified social class as the result of processes of production that bring together occupations that are unequal in status. Marmot (2017) summarised the struggle between classes: the bourgeoisie, those in society that own the key to production, and the proletariat, those who do not (Marmot 2017). Weber enhanced this definition with the addition of political power and prestige (Kogevinas et al 1997a). And Krieger et al (1997) highlighted that social

class reflected the concept of relative economic interest *and* relative prestige (Krieger et al 1997). Building a picture of its pervasive nature, Kawachi (2006) pointed out SES existed in every society, whether low-, middle- or high-income; there was no threshold or cut-off of socioeconomic position where SES did not exist (Kawachi et al 2006). This implies that SES and the associated psychological and physical impact exist even for the most affluent members of society; there is something in human nature – and in our primate cousins — the observation of the position of others relative to ourselves that causes distress (Behrens et al 2016).

Various terms are often used in the literature interchangeably to describe these concepts such as socioeconomic position or socioeconomic circumstances and status. In this thesis the term socioeconomic status (SES) and socioeconomic circumstances are adopted. By SES or socioeconomic circumstances, both absolute and relative levels of income, wealth alongside aspects of power and prestige are encapsulated and reflect the dimensions of socioeconomic status. Measures of educational attainment, employment status, occupational status, income, accumulated economic assets (e.g. home and car ownership) and social participation all of which reflect general "control over life" and "power" are included. In this thesis, the inequalities focus is on low socioeconomic circumstances and not other factors such as race, ethnicity, gender, religion, sexuality or age. Marmot (2017) has moved away from status to socioeconomic position but this seems like a judgment call to avoid some of the status-power aspects in socioeconomic classification (Marmot 2017).

In summary,

Socioeconomic status is recognised as a *complex relationship* of *multidimensional* (Kogevinas et al 1997a) factors capturing both the *material* and *psychosocial* (Krieger et al 1997) aspects of an individual's social circumstances which are *dynamic and cumulative* (Kawachi et al 2006) as well as *synergistic and compounded over the life-course* (Mackenbach 2012; Marmot et al 2012) and are relative to others in society (Mackenbach 2012). The SES milieu *reflects and influences health profoundly* through both *upstream and downstream pathways* (Watt 2007; Braveman et al 2011).

#### 1.3.1.1 The complex and multidimensional nature of SES

SES results not only from an individual's economic position within society in relation to work as an employer, employee, self-employed, or being unemployed or in relation to wealth or assets as an owner, or not, of capital, land, or other forms of economic investments (Krieger et al 1997), but, it can also reflect and be influenced by the "place" or "context" of where the individual lives and works (MacIntyre et al 2002).

The health status of members of a community can also be influenced by the presence or absence of community infrastructure such as (public) libraries, transport, health centres, social services, schools, public health centres, healthy eating establishments, recreational space such as gardens and parks (Kamphuis et al 2008). Further it can be affected by attributes of the community's members such as their income, education, ethnicity, religion, age (Cagney 2006), sex, social class and presence or absence of gangs and vandalism (Bryden et al 2013). The influence of "*where you are"* in defining "*who you are"* is recognised as being critical to understanding the role that SES plays in health outcomes including cancer risk (MacIntyre et al 2002; Kawachi et al 2017).

The concept of "who you are" being influenced by "where you are" was debated in the literature in 2009 at the time that this PhD was started. A major focus at

that time was exploring the *relative* importance of area-based indicators and individual measures of SES (MacIntyre et al 2002; Costa et al 2003; Caiazzo et al 2004; Shohaimi et al 2004; Islam et al 2006; Do et al 2008; Harenstam 2009; Spadea et al 2010; Conway et al 2010b; Eriksson et al 2011; Eriksson et al 2013; Lewin et al 2014; Kawachi et al 2017).Two contrasting views were considered: i) Individual level SES would express more accurately the association between SES and health risk; or, ii) Area and individual measures were both relevant for showing different aspects of socioeconomic circumstances.

MacIntyre (2002) suggested that simple aggregation of individual attributes (e.g. unemployment and occupational social class) to constitute "area" measures of SES are limited as they do not capture local social and physical environmental attributes of a neighbourhood. The health effects of the "place" or "context" (neighbourhood, workplace, or region) also contribute to the causal pathway (MacIntyre et al 2002).

Much of the research on SES and cancer to date, particularly in the UK (Lamont et al 1997; Brewster et al 2000; Lancaster et al 2006; Cooper et al 2007; Shack et al 2008; Cooper et al 2009; Coupland et al 2012; Caygill et al 2014a; Caygill et al 2014b) has focused on area measures of SES alone because of its availability and accessibility in datasets. However, this omits the individual measures of SES which is an important gap in describing additional aspects of SES and its influence on health outcomes. Area measures have almost become a euphemism for actual individual socioeconomic circumstances, i.e., a substitute for individual SES as it is seen as a milder, more vague term considered less offensive and easier to obtain while asking about individual income or education is considered more obtrusive.

#### 1.3.1.2 The psychosocial aspect of SES

The psychosocial element of SES is described in terms of the extent of social cohesion, integration or solidarity in a community (Kawachi et al 1997) and is often described as "social capital". Social capital is the tangible and psychosocial resources available to individuals and society through social

relationships (Kawachi et al 2017). These social capital resources include civic participation, social trust in others and norms of reciprocity which engender social cooperation for mutual benefit (Watt 2011).

Social capital was viewed as not only an instrument for the privileged but a public asset (Uphoff et al 2013). This social network consisted of strong relationships between family and friends, ties between neighbours, club members, or colleagues and links between employer – employee or citizen – governments (Szreter et al 2002). In his 2006 review of social capital and its relationship with socioeconomic inequalities in health, Islam et al (2006) concluded that regardless of the study design, or the country and its level of egalitarianism, stronger or greater social capital was associated with better health outcomes (Islam et al 2006). Furthermore, he found health inequalities that did exist tended to be lower in more egalitarian societies. Uphoff et al (2013) described two potential pathways between social capital and socioeconomic inequalities in health: limited availability of social capital among the more socioeconomically deprived groups and the stress for the individual that arose from comparing his/her position relative to other SES groups in his/her society (Uphoff et al 2013). Behrens et al (2016) described this comparison of one's own SES with that of others as "status inconsistency" which may reflect loss of status control as well as the clash between expected and actual SES (Behrens et al 2016). Intrapersonal factors and shared psychosocial factors such as stress, perceived control in addition to social environmental influences are also recognised as being fundamental to creating inequalities in health outcome (Sheiham et al 2000; Watt 2002).

#### 1.3.1.3 The ever changing character of SES: dynamic, synergistic, compounded and cumulative over the life-course

The influence of the life-course pathway has been described as "dormant" or "latent" i.e. causing illness later in adult life. It can act through either *i*) "A pathway effect" such that early experiences affect decisions at future stages in life, which in turn cause illness in later life due to lasting effects potentially interacting with some modifying and triggering effect; or *ii*) A "cumulative"

effect where the intensity and duration of exposure led to illness each with independent and correlated risks (Hertzman 1999; Galobardes et al 2004).

SES is also considered dynamic (Kawachi et al 2006) reflecting that an individual's SES can vary up and down through life depending on the vagaries of life's journey. Achieving a job promotion may lead to greater income and greater social mobility reflected in movement to a better neighbourhood with an enhanced health improving environment. However, this upward social mobility may be followed by unexpected illness necessitating fewer work hours and therefore lower pay, movement to a lower paying job or even stopping work all together. All of which, depending on that individual's social capital and economic circumstances, are likely to result in an unfavourable change in that individual's SES (Marshall et al 1999; Schmeisser et al 2010; Robertson et al 2012; Behrens et al 2016).

In his review of theories of social determinants of health, Watt (2002) built on the work of Marmot, Blane and others to develop the concept of "clustering" of disadvantage over the life-course (Watt 2002). An individual who is long term unemployed may also live in poor accommodation, be unable to afford a healthy diet and smoke and consume alcohol in order to cope with life's stress. This situation described how disadvantage cumulated cross-sectionally or clustered or is "compounded" during the life-course. Individuals facing multiple low socioeconomic circumstances have "brittle" coping systems unable to withstand multiple events going wrong all at the same time. In contrast, an individual who was born to a privileged family is likely to have had the opportunity to attend a well respected university, attain a secure well-paying position and retire with a dependable and well provided pension. In this case, favourable socioeconomic experiences accumulated longitudinally or created a "chain of advantage" over the life-course (Watt 2002). This "advantaged" individual experienced a "resilient" coping system preventing escalation when faced by a trigger or risk of change in socioeconomic circumstances.

The literature explored the importance of key life stages on future health status

(Watt 2002; Galobardes et al 2004). These stages include: primary to secondary school transition, school examinations, attaining first job, leaving the parental home, establishing your own home, becoming a parent, the loss or change of employment and finally, leaving employment. These stages of life have been described as critical periods of susceptibility where a window of exposure may initiate chains of risk with additive or trigger effects. However, this compounded nature or effect of multiple low SES factors has not been fully explored and remains a gap in the literature.

Watt (2002) explored the influences of SES over each stage of the life-course, explaining that SES determines opportunities for formal education, qualifications, employment opportunities and type of employment, job security, salary, income and pensions, working and living circumstances as well as social networks, material environment of home, neighbourhood and workplace (Watt 2002). An individual's personal attributes interact with and are shaped by each of these influencing factors. If a person is 'vulnerable' and currently or previously experienced multiple social disadvantage, health damaging behaviours are likely to be encouraged and adopted. These behaviours reflect increased exposure to occupational and health hazards, chronic and acute stress, prejudice and injustice and ultimately lower self-esteem and sense of hopelessness for now and the future (Marmot 2010; Marmot et al 2012).

The influence of exposures acting during critical periods of susceptibility may be modified by later life exposures (Galobardes et al 2004). The definition of life-course may be required to include parental SES experience too. In the context of cancer, Galobardes' systematic review of childhood socioeconomic circumstances effect on all cause mortality reported one study found that men whose fathers had manual jobs or who were from large families experienced higher stomach cancer mortality independent of adult SES (Galobardes et al 2004). Three further studies reviewed found that there was no association between childhood socioeconomic characteristics and later death from non-smoking related cancers. In the context of smoking related cancers, poor childhood and adult SES could independently influence for example, lung cancer

risk via increased chance of starting to smoke, earlier age of initiation, lower probability of quitting and higher age of quitting smoking (Galobardes et al 2004).

### 1.3.2 Indicators of SES: education, occupation, income, area and individual measures

Krieger's 1997 review explored how to measure social class (Krieger et al 1997). She clarified that the resource-based measures implicit in socioeconomic status or position refer to material and social resources and assets, including income, wealth and educational credentials. Inadequate resources are often described as "poverty" and "deprivation" (Krieger et al 1997). Maxwell of the Overseas Development Institute (Maxwell 1999) identified nine terms used to describe poverty which comprehensively encapsulate the range of domains and aspects of poverty or deprivation. These are summarised here:

- ***"Income or consumption poverty"***, income was only valuable if it enabled the capabilities of individuals and supported or allowed functioning in society; otherwise, Maxwell described it as "consumption poverty".

- ***"Human (under)development"*** described as the denial of opportunities and choices to lead a long, healthy, creative life and to enjoy a decent standard of living, freedom, dignity, self-esteem and respect of others.

- ***"Social exclusion"*** reflected a feeling of powerlessness, isolation and as a result, the inability to participate in society due to the design of systems such as democratic systems, legal systems, markets and welfare systems.

- ***"(Lack of) capability and functioning"*** reflecting poorer educational attainment and lower life expectancies.

- ***"Vulnerability"*** reflected loss of (social and capital) assets as buffers and subsequent susceptibility to shocks. Today's Westminster government uses the term "just about managing" (Citizens Advice 2017) to describe individuals and families that are not rich but also not the poorest in society. Nevertheless, these "just about managing" people find day-to-day life a struggle, despite being mostly in work. As a result, these families and individuals were living very close to the edge, were susceptible to being knocked over by one of life's unexpected events and

were trying to cope with a sense of total lack of control over their everyday lives.

- ***"Livelihood unsustainability"*** reflected the importance of social capital (not just income and wealth) and coping strategies.

- ***"Lack of basic needs"*** such as the recent Scottish parliament debates which discussed "period poverty" for women and girls (Freeman T 2017) but also including those provided socially such as education, health care and other services such as (public) transport and housing.

- ***"Relative deprivation"*** recognized poverty in terms of minimum standards of nutrition and subsistence, but also the impact of inability to "keep up with the Jones", i.e. reach for and attain the normal standards of society.

Maxwell's (1999) list of domains is valuable as a construct to understanding multiple low socioeconomic circumstances. Each of the nine attributes he described reflects different facets of SES which ideally would be incorporated either separately or in a composite measure.

Prestige-based measures refer to an individual's rank or status in a social hierarchy. As they can be measured via an individual's access to and consumption of goods, services and knowledge, as linked to their occupational prestige, income and education level (Krieger et al 1997). Implicitly, "rank" suggests hierarchy, i.e. identification of gradients of SES rather than the simple approach of comparison of two categories of SES: the poor and the affluent.

No single indicator captures all aspects of SES. As a result, it is relevant to use different socioeconomic indicators as they include different dimensions of SES, which are established at different phases in the life-course, are often related and may be more or less relevant to a specific study outcome and the pathways SES may influence that outcome (Galobardes et al 2006b).

The three main indicators used in SES research, based on data availability and incorporation in routine surveys, are: education, occupation and income. Economists tend to focus on income and usually differentiate income from wealth or assets (Section 1.4 and Gini Coefficient) while sociologists focus on

occupational status and education (Reeves R 2017). Epidemiology has borrowed from both disciplines adopting a multidimensional perspective; however, the decision on which indicator to include and how to weight a component relative to others, remains complex and open to debate and often reflects the traditions and cultures of the area in which the study is performed. For example, in the UK, occupational social class as measured by the Registrar General's Social Classes which was established in 1913, and has been revised several times since then with the most recent versions being the 2000 Standard Occupation Classification (SOC2000) and Standard Occupational Classification (SOC2010; CeLSIUS 2017). This definition of occupational social class is traditionally used as an indicator of SES in the UK. In Europe, education is commonly used to measure SES as evidenced by studies performed by Mackenbach (2008) assessing socioeconomic inequalities in health in 22 European countries and several Scandinavian studies exploring the association of SES and cancer (Menvielle et al 2010a; Menvielle et al 2010b; Leuven et al 2016). Finally, in the US, income is frequently used for measuring SES (Minkler et al 2006; Boscoe et al 2014).

#### 1.3.2.1 Education

Education is generally (but not always) acquired at an early phase of the life-course. It therefore may reflect the skills and knowledge acquired to protect health (Spadea et al 2010; Dalton et al 2008c) by, for example, understanding and acting on public health messages thereby influencing health attitude and behaviour. More educated individuals are better informed (read and assimilate more medical/health material), and therefore are better equipped to make choices that benefit their health. Educational attainment also supports career choice and opportunities, income, working, living conditions and accessibility to healthcare (Sidorchuk et al 2009). Education as an SES indicator is considered more inclusive compared to occupation and income given that it captures those not working, does not reflect regional differences in cost of living and is not influenced by census household definitions (Mitra et al 2015). Given that the highest level of education is usually attained in young adulthood and therefore reflects parental characteristics, Galobardes et al (2006) suggested, in the life-

course context, education level reflects early life SES and importantly the transition from parental SES received at birth and self-acquired SES in adulthood (Galobardes et al 2006b).

It can be measured as both a continuous variable and a categorical variable. As a continuous variable reflecting the number of years in formal education it focuses on the importance of time spent in education, or as a categorical variable based on education level, it reflects accomplishment as well as prestige (Berkman et al 1997; Conway et al 2010a).

As a result, and depending on the education indicator(s) employed, it is possible to focus on its quantity (number of years) as well as its quality (level of attainment) although these two attributes of education are inter-related (Conway et al 2010a; Berkman et al 1997). Interpretation is facilitated by the fact that education level is generally constant over adult life and therefore generally avoids reverse causation bias (Mouw et al 2008). Conway et al (2015) explored this theory more thoroughly explaining that while low educational attainment caused by childhood illness could be considered inverse causation, this was unlikely to be relevant to head and neck cancers in particular. It is also unlikely to apply to the other cancers under consideration in this thesis, i.e. lung and oesophageal cancers. Conway et al (2015) did however raise the possibility of other unmeasured variables influencing head and neck cancer risk through education such as IQ and the individual's focus on well-being today or well-being in the future (Kawachi et al 2010; Conway et al 2015).

Education may also be influenced by societal beliefs and the norms of the time that the education level is attained. This can change substantially over time. For example, at a country level, educational levels have increased over time in many countries while at an individual level (Galobardes et al 2006b), it is generally constant over the life-course (Mouw et al 2008). In today's economy there is much focus on "life-long" learning as a response to changing skill requirements of jobs and the need for older employees for financial reasons to continue to work beyond what was once the normal retirement age. As a result,

reflecting education level accurately may require multiple measurements over time in the future. Finally, unbundling the range of meanings and interpretations of education continues to be complex.

#### 1.3.2.2 Occupational social class

Occupational measures such as occupational social class reflect material resources or rewards such as income while being employed and via a pension after retirement, social standing and working conditions in Weber's theory of social stratification as well as specific occupation associated risk exposures to carcinogens (Spadea et al 2010; Nkosi et al 2012). Pukkala et al (2009) evaluated the occupational risk association with cancer incidence by site and in some cases morphology for the Nordic countries. Several different occupations were evaluated — just one example of risk being associated with occupation presented was of miners and quarry workers who may be exposed to radon, silicon dust, diesel exhaust and asbestos, as part of their work, all of which are associated with increased lung cancer risk (Pukkala et al 2009). Occupational social class, income and education are all interdependent with education influencing occupational class which in turn influences income; as a result, interpreting the implications is more complex.

In the UK, the Registrar General's social class dating from 1911 (Rose 1995) or its successor, developed in 2000, the UK National Statistics Socioeconomic Classification (NS-SEC), are often used to capture social class. Both systems, because they are widely employed in the UK and have been adapted for use elsewhere, are relatively easy for researchers to apply in designing questionnaires, coding and modelling for case-control or cohort studies. Craig et al (2005) conducted a study to compare the systems and implications for interpretation. They demonstrated that both classifications systems (and a third less commonly applied, the Cambridge Social Interaction and Stratification Scale) were strongly associated with self-assessed health, the health outcome measured, although the associations were heavily attenuated by adjustment for one another and for other measures of social position. Craig et al (2005)

concluded, despite their differing theoretical bases, the three systems were closely related and that the availability of the UK National Statistics socioeconomic classification was unlikely to transform our understanding of the extent or the causes of socioeconomic inequality in health, but provided useful opportunities for sensitivity analysis.

Low occupational social class may encompass a work environment that is more associated with harmful psychological or social environments with "work stresses" which in turn may affect health. Increased risk to health may result from poorer terms and conditions, increased short-term employment, unreliable contracted hours, or increased periods of unemployment (Conway et al 2008). In Britain's economy today, austerity, zero hours contracts and lack of guaranteed hours are, unfortunately, examples of deteriorated terms and conditions.

Alternatively, high occupational class may reflect factors such as access to influential social networks, the influence of colleagues on health behaviours, or fewer occupational exposures to carcinogens as well as income or material reward (Galobardes et al 2006a).

A single occupation's prestige may also change over time. Profound societal changes, such as (de-)industrialization or change of the political system may have implications for the social standing of a particular occupation. Looking to the future, forecasts of the expansion of automation will no doubt also have significant impact on occupational prestige. As a consequence, interpretation differs depending on birth cohort, country, gender and ethnicity (Behrens et al 2016).

Occupational social classification measures prestige and status, but because of the need to summarise a large volume of occupations, the strata are often heterogeneous depending on the scheme adopted. Nevertheless, because occupation is frequently dependent on gained knowledge and experience, occupational social class is considered a relatively stable indicator of SES as it is established after the relevant educational attainment has been achieved, usually

at an earlier point in the life-course. Studies across countries at different stages of industrialisation, with very different political systems and societies have found that ranking of jobs according to social prestige was independent of country and time of survey (Behrens et al 2016). Compared to other measures of social status such as income and education, occupational social class appears to be less affected by temporal changes (Behrens et al 2016). Nevertheless, unbundling the effect of occupational social class in order to understand the pathway from SES to health inequalities remains complicated given the interrelationship between occupational social class, education and income (Sidorchuk et al 2009). Furthermore, a clear limitation of occupational social class systems in general is their inability to capture the complication of those not in a recognised occupation which will include those unemployed, as well as students, those caring for family members or looking after the home and others. To mitigate this important limitation, employment status, or relationship with the employment market is also required (Galobardes et al 2006a).

#### 1.3.2.3 Income

Finally, income reflects financial and material circumstances which can have a strong behaviour influence, acting directly or indirectly via interplay with the effect of educational and occupational social position (Sidorchuk et al 2009).

Income is likely to influence health mainly by a direct effect on material resources, and the proposed mechanisms include greater access to better-quality resources, such as food and housing and better access to services that may improve health directly (health services, leisure activities) or indirectly (e.g. education) (Dalton et al 2008c).

In the UK and Europe, income data are less commonly available and therefore not frequently used in health epidemiology studies. In 2010, the Scottish Parliament considered inclusion of gross annual income of the household (Scottish Government 2010b).This additional information presented the first possibility of considering both individual and area-based measures of wealth measured by income. In the context of this thesis, two limitations remained.

Firstly, the General Register's Office (GRO) did not anticipate 2011 census information to be available to researchers until 2014, after the studies for this thesis were concluded. Furthermore and ideally, disposable household income as opposed to gross household income would provide a more informed indication of wealth. However, collection of disposable income data would require detailed information unlikely to be willingly provided or accurately obtained. In the end, the question on household income was removed from the 2011 census as during testing, the question was completed by only 48% of those who undertook the survey and was identified by 17% of respondents as inappropriate (National Records Scotland 2015d). Income measures in the UK remain a sensitive, but important gap in the ability to fully capture this facet of socioeconomic circumstances. In the UK, area measures of socioeconomic circumstances do reflect, in aggregate, those in an area receiving financial benefits. This is described more fully in the next section (Section 1.3.2.4).

#### 1.3.2.4 Area versus individual SES indicators

Indicators of SES have been based on the characteristics of the individual as well as on the characteristics of the environment, or more ecologically based measures each reflecting different aspects of social class (Kogevinas et al 1997a). Area measures are more frequently applied as a measure of SES given ease of access. In Scotland there are two area-based measures: the older Carstairs Index and the more recently developed Scottish Index of Multiple Deprivation (SIMD).

***Carstairs Index***. The Carstairs index (Carstairs 1995) was developed in the 1980s using the 1981 census and was designed to reflect material resources and was structured similarly to the Townsend Index used in England. It is measured at postcode sector level and is based on four variables: male unemployment, households with no car, overcrowded households and the percentage of people in social classes IV (partly skilled) and V (unskilled). Scotland's 1,011 postcode sectors contained an average population of 5,012. The index is standardised such that each variable has a variance of one; therefore, each variable has an equal

influence on the resultant SES score. Dependent on census information, the index is updated every ten years and is available for 1971, 1981, 1991, 2001 and 2011. Two major changes over that period have occurred; i) The overcrowding variable was changed in 1991 to include kitchens at least two meters wide; and ii) The classification system adopted by the Registrar General changed from Social Class to National Statistics Socioeconomic Classification (NS-SEC) (Section 1.3.2.2). Despite these modifications, the basis of calculation for Carstairs has been relatively constant over time; however, dependence on census data limits the updates to a ten-yearly cycle or the cycle of censuses in the future. Furthermore, the four variables that were selected to measure material wealth are considered now to be out of date in today's society; for example, car ownership is more common now than in 1971 and female unemployment, in today's labour market, is just as important a factor affecting material wealth as male unemployment. The Carstairs Index is considered less effective in evaluating rural area deprivation given that a car may be a requirement regardless of your socioeconomic circumstances where public transport is limited or unavailable and in the context of drivers of socioeconomic inequalities in health, may provide access to health services (Berkman et al 1997). However, from a theoretical perspective, Carstairs may be considered to be a more relevant index of socioeconomic circumstances when evaluating health outcomes as it does not include any health indicators unlike the more recently developed indicators of multiple deprivation (Carstairs 1995).

***Scottish Index of Multiple Deprivation (SIMD)***. To address the limitations presented by Carstairs, the SIMD was developed. SIMD was first available in 2004 and has been updated in 2006, 2009, 2012 and 2016 (Scottish Government 2012c). During the period of conducting the analyses for this thesis, and unless otherwise stated in the relevant study methods, SIMD2009 and SIMD2012 were the most recently available and up to date versions.

SIMD covers multiple drivers of deprivation described through seven domains (income, employment, education, housing, health, crime and geographic access) covering 36 variables and is measured at datazone level. Table 1.1 summarises

the domains, indicators and data sources used for SIMD2012 (Scottish Government 2013d). The crime domain focuses on crimes of violence, sexual offences, domestic housebreaking, vandalism, drug offences and common assault while the geographic access domain provides an indicator of access to services (GP practice, post office, retail centre and primary and secondary school) in an area. As a result, both domains begin to capture attributes of the area or neighbourhood as opposed to summarising individual attributes at area level (Scottish Government 2013d).

In the SIMD2016 version, there are 6,976 datazones with 760 individuals on average. As a result and regardless of version, SIMD covers smaller populations compared to postcode sectors (Bishop J et al 2004). Given the smaller geographic area, the area is more likely to be more homogenous with respect to socioeconomic characteristics than postcode sector. The overall SIMD index is used to identify area concentrations of multiple deprivation. SIMD is sourced from administrative data as opposed to census data, e.g. Department of Work and Pensions. As a result, it can be more regularly updated than census based indices such as Carstairs. More recently, the SIMD2016 version has two substantive changes. Firstly, datazones were changed to reflect the 2011 census; previously datazones were based on the 2001 census. Secondly, the income domain was revised to reflect the new Universal Credit system (Scottish Government 2017g).

A criticism of SIMD is the fact that it includes a health domain and if used to analyse health data (GPD Team 2017), independence of the SIMD and the health indicator is jeopardised. However, the health domain is weighted to account for a relatively small part of the overall SIMD (14% of SIMD 2009, 2012 and 2016) and analyses of health inequalities using SIMD 2004 were found to give similar results whether the health domain was included or excluded, because that domain was so highly correlated with the overall index (GPD Team 2017).

**Table 1.1 Scottish Index of Multiple Deprivation 2012 Domains, Indicators and (Data Sources)[1]**

| Employment Domain | Income Domain | Crime Domain | Housing Domain | Health Domain | Education Doman | Access Domain |
|---|---|---|---|---|---|---|
| The count of the number of employment deprived people in a datazone is equal to the number of men aged 16-64 and women aged 16-60 who are on the claimant count, receive Incapacity Benefit, Employment and Support Allowance, or Severe Disablement Allowance. | The count of the number of income deprived people in a datazone is equivalent to the count of adults and their dependants in receipt of Income Support, Employment and Support Allowance, Job Seekers Allowance, Guaranteed Pension Credits and Child and Working Tax Credits (UK Department for Work and Pensions (DWP) and HM Revenues and Customs). | SIMD Crimes per 10,000 total population | Percentage of household population living in households without central heating (Census, 2001) | Standardised mortality ratio (Information Services Division (ISD), 2007-2010) | Working age people with no qualifications (2001) | Drive time to GP |
| | | | Percentage of household population living in households that are overcrowded (Census, 2001) | Comparative illness factor: standardised ratio (DWP)[2] | People aged 16-19 not in full time education, employment or training rate (School Leavers 2009/10, 2010-11, DWP 2010 | Drive time to Petrol Station |
| | | | | Hospital stays related to alcohol misuse: standardised ratio (ISD, 2007-2010) | Proportion of 17- 21 year olds entering higher education (HESA[3]2008/09, 2010/11) | Drive time to Post Office |
| | | | | Hospital stays related to drug misuse: standardised ratio (ISD, 2007-2010) | Pupil Performance on Scottish Qualifications Authority (SQA) at Stage 4 (SQA, 2008/09, 2010/11) | Drive time to Primary School |
| | | | | Emergency stays in hospital: standardised ratio (ISD, 2007-2010) | School Pupil Absences (Scottish Government, 2009/10, 2010/11) | Drive time to Secondary School |
| | | | | Estimated proportion of population being prescribed drugs for anxiety, depression or | | Drive time to retail centre |
| | | | | Proportion of live singleton births of low birth weight (ISD, 2006-09) | | Public transport travel time to GP |
| | | | | | | Public transport travel time to Post Office |
| | | | | | | Public transport travel time to retail centre |

[1] (Scottish Government 2013d)

[2]The Comparative Illness Factor is based on benefits data, counting people claiming Disability Living Allowance (DLA); Employment and Support Allowance (not also receiving DLA); Attendance Allowance; Incapacity Benefit) (not also receiving DLA); and Severe Disablement Allowance).

***Income and employment domains of SIMD.*** The income and employment combined domain (I-E) index can be combined as the I-E index which is based on the eight variables in the SIMD2009/2012 income domain and the four variables in the SIMD employment domain (Table 1.2) (Scottish Government 2008d) and is similar to the Index of Multiple Deprivation (IMD) used in England. Calculated by the Health Department's Analytical Services Division of the Scottish Government, the two domains were combined with equal weight after exponential transformation which gave greater weight to the most socioeconomically disadvantaged. I-E tenths were population weighted ensuring that each tenth contained equally sized populations. This was in contrast to SIMD deciles which are defined and ranked by datazone not population (Scottish Government 2013b). Currently, I-E index is available for each year from 1996 to 2016. The I-E domain has been considered for targeting individuals for anticipatory care (Fischbacher C 2017), identifying deprivation in rural areas (Scottish Government 2011), for review of long term monitoring of inequalities in Scotland (Scottish Government 2017c) and to support deprivation comparison across countries in the United Kingdom (Abel et al 2016).

In terms of identifying rural deprivation, the argument was that rural areas are more dispersed given a larger area and because rural areas are larger areas compared to cities, they contain a greater mix of people with different socioeconomic states in one area. As a result, it was argued that individual measures used in the I-E domain would better identify socioeconomically disadvantaged areas/ individuals in rural areas (Scottish Government 2011).This decision implied that other area attributes that are currently captured by SIMD such as transport and access to services were considered relatively less important than individual measures of SES such as I-E which may or may not be the case.

To invite individuals for anticipatory care screening, the I-E combined index was proposed because it was based on individual measures which may identify individuals suitable for anticipatory care more accurately than SIMD. The Short Life Technical Group supporting the long term monitoring of inequalities in

Scotland believed that ideally, individually linked health records and individual socioeconomic indicators were preferred but unavailable; the I-E, given it is based on individual measures, was considered the best alternative. As a result, the Scottish Government's review of long term health inequalities published in March 2017 changed from SIMD to I-E as the underlying measure of socioeconomic circumstance (Scottish Government 2017c). With respect to cross country comparison, the I-E domain is more consistent regardless of country, enabling more appropriate comparison of relative socioeconomic disadvantage within the United Kingdom, subject to potential future modifications already made in Scotland to either mitigate Westminster welfare policies in Scotland such as the bedroom tax as well as proposals to be developed and implemented related to the newly devolved social security powers for Scotland.

In the context of this thesis, the I-E was considered inferior to the full SIMD given the multidimensional nature of SES (Kogevinas et al 1997). While income and employment is a fundamental aspect of SES and may capture the material dimension of SES, it was unlikely to capture all the facets of SES and its complex nature. Furthermore, the very strength of I-E for identifying *individuals* for anticipatory care attendance, i.e. that it was based on individual level data, would be considered a weakness in the context of the objectives of this thesis where the intention was to explore not only individual measures of SES but also the area measures of SES and in the case of the latter, ideally the attributes of the area, not an aggregation of data for individuals living in that area. Finally, the I-E domain may be argued to be more appropriate for analysing health outcomes because unlike the full SIMD, it would not include health data. However as stated above, the health domain is only 14% of total SIMD (GPD Team 2017). SIMD was found to provide similar rankings whether the health domain was included or not (GPD Team 2017). Thus, the full SIMD was the preferred indicator for comparing area-based socioeconomic circumstances and has been adopted in this thesis in preference to the I-E domain.

**Table 1.2 Income and employment (I-E) domain variables**

| Income for Adults & Children (8 domains) | Employment Variables |
|---|---|
| ▪ Income support<br>▪ Income based on job seekers allowance<br>▪ Working families tax credit<br>▪ Disability tax credit | ▪ Unemployment claimant count<br>▪ Incapacity benefits recipients<br>▪ Severe disablement allowance recipients<br>▪ New Deal recipients |

***Interpretation of area SES measures***. Geographic area-based socioeconomic indices result in all people living in a particular area being allocated the same SES. Individual indicators of SES may thus prove to be better at identifying individual socioeconomic circumstances and equating this with disease risk including cancer incidence. At the SES gradient extremes, area-based indices will classify fairly socially homogenous areas; however SES categories in the middle are likely to contain individuals with a more mixed range of SES (McLaren et al 1998). An area-based deprivation score may also be a less accurate measure for comparing the most and least socioeconomically disadvantaged groups and may impact on measures of inequalities which reflect the gradient across the full population. Nevertheless, for routine monitoring purposes, area measures are remarkably consistent (Boyce 2008).

Interpretation of disadvantage measured by area-based indicators is complex. If used as a surrogate individual measure, it may be inferred that a person living in a high-income area has high-income. However, this interpretation is subject to a phenomenon known as the "ecological fallacy"; the population may be heterogeneous such that the population's attributes do not necessarily equate to the individual's attributes living in the area (Boscoe et al 2014). The larger the geographic area used, the greater the chance of misclassifying individuals. As an example, if interpreted as an area measure, it may be implied that a person who may happen to be a member of a lower socioeconomic group but lived in a high socioeconomic area would therefore have access to health promoting local resources (Berkman et al 1997).

The resources available within a community such as eating or retail establishments, recreational areas such as parks and absence or presence of transport infrastructure, as well as environmental factors, such as pollution,

interact with individual characteristics (MacIntyre et al 2002). Personal and local circumstances together amplify disadvantage and health risk. Given this interactive effect and the presence of genuine socioeconomic area effects associated with, for example, levels of crime, drug use, gang activity, accessibility of healthy food and good transport links, it is recommended that both individual and area indicators should be considered (Pickett et al 2001).

The literature uses "context" and "composition" effects to distinguish between attributes of the individual (composition effect) and attributes of the area, place or neighbourhood in which the individual lives (context effect) (Pickett et al 2001; MacIntyre et al 2002; Leyland et al 2005; Riva et al 2007). Based on a review of multi-level analyses, Riva et al (2007) discussed the conceptual and methodological challenges for future research on area effects on health including: articulating the causal pathway, recognising differences between administrative area boundaries and neighbourhood boundaries; defining ecological exposures to create meaningful area variables as opposed to aggregating data from individuals to measure area effects; and adopting longitudinal study designs as opposed to cross-sectional designs to ascertain exposure timing and duration, address selection bias and assign causality (Riva et al 2007).

#### 1.3.2.5 Timing of measurement

The measurement of SES, in relation to cancer, has occurred most frequently at diagnosis given the ease of capture at the interface with the health service. However, the known long lead-time between cancer initiation and diagnosis as well as the complex and dynamic nature of SES, and its role over the life-course, means that measurement of SES at diagnosis may under estimate, omit, or mask the effect of SES. Ideally multiple measurements over the life-course including potentially parental SES to reflect childhood circumstances are relevant (Ben Shlomo 2007).

For both area indices used in Scotland, the area that is used to define relative disadvantage has changed over time which presents interpretation issues when

reviewing trends over time. Datazones, which are the basis for calculating the SIMD, were introduced in 2004 to replace postcode sectors as the key small area geography for Scotland (Boyce 2008). Based on 2001 Census Output Areas, datazones were intended to be a stable geography over time with a reasonably consistent population size and boundaries set to respect physical boundaries and natural communities as far as possible. These attributes were intended to overcome the postcode sector limitations observed for Carstairs given that postcodes are owned by the Royal Mail. As a result, they are geographically unstable as boundaries change reflecting buildings that have been demolished or constructed. Therefore, the population in postcode sectors varies widely. In contrast, datazones are population-based, but they too can vary hugely, in particular, by geographical size. For example, in towns and cities where people live close together, datazones can contain only a few streets, while in rural areas that are sparsely populated, they can cover many square miles. In November 2014 (after the completion of the studies included in this thesis) datazones were updated to include population information from the 2011 Census, as a result, datazone boundaries have been redrawn to deliver more consistent population size (GPD Team 2017).

Nevertheless, the use of SES measurement at the time of diagnosis is less than ideal. Accessing datasets where SES is measured prior to diagnosis to unpick the issue of temporal relationship is an important priority for research in this field.

#### 1.3.2.6 Summary of literature on SES

This review of the literature has identified gaps in the approaches used to capture SES in previous studies. These include: i) Few studies have explored both individual and area measures of SES simultaneously to understand their relative importance and contribution to understanding of the pathway between SES and cancer incidence; ii) Given the multidimensional nature of SES and the compounding effects of multiple low socioeconomic circumstances over the life-course, there is a need to develop analytical approaches that can investigate this compounded effect and reflect multiple SES measures over time; iii) The

long gestation time of cancer and the dynamic nature of SES over the life-course mean that measurement of SES well before diagnosis is not only justified but necessary; iv) Analytical approaches to minimise change in underlying aspects of measuring SES such as postcode or datazone definition or geography should be adopted to support the need to focus on circumstances over the life-course. This would support minimising SES change that is a function of the administration of the underlying components of the SES indicator which may mask or mitigate change in health outcome that is due to the true SES change; and v) Finally, there is an over reliance of only using a single area-based SES indicator in reporting/monitoring SES on health and cancer in Scotland (SIMD). Further, deeper analysis is warranted to help better understand these relationships.

## 1.4 Measuring socioeconomic health inequalities

Many different measures have been developed to monitor socioeconomic associated health inequalities and have been reviewed by Harper et al (2008), Harper et al (2009), Mackenbach et al (1997) and Blair et al (2013) (Table 1.3). In summary, these measures have adopted concepts from the disciplines of economics, sociology and epidemiology. Important modifications are required to reflect the needs for monitoring cancer incidence inequalities in particular and health in general. As an example, measures used in economics such as the Gini Coefficient and the Concentration Index focus on inequality between individuals as opposed to between group inequalities and are therefore less relevant to measuring social group inequalities. Nevertheless, these two measures have been included in this review for completeness (Mackenbach et al 1997; Harper et al 2008; Harper et al 2009). These limitations are discussed more fully in Section 1.4.3.

### 1.4.1 Refinement of the definition of health inequality

A major refinement of SES concepts acknowledged a distinction between equality and equity (Mackenbach et al 1997; Harper et al 2009). Inequality is an objective concept which can be measured in terms of dissimilarity or differences. Some, like Marmot (2017), proposed that inequity requires

subjective assessment of whether those differences are fair (Marmot 2017). A further refinement was a focus on the "avoidable" inequalities or that part of the inequalities which can be influenced (Woodward et al 2000). Marmot went even further and stated that "Health inequalities that are avoidable and are not avoided are unjust. Putting them right is a matter of social justice" (Marmot 2017). The distinction made is important for policy development, implementation and monitoring and may therefore be relevant for political decisions.

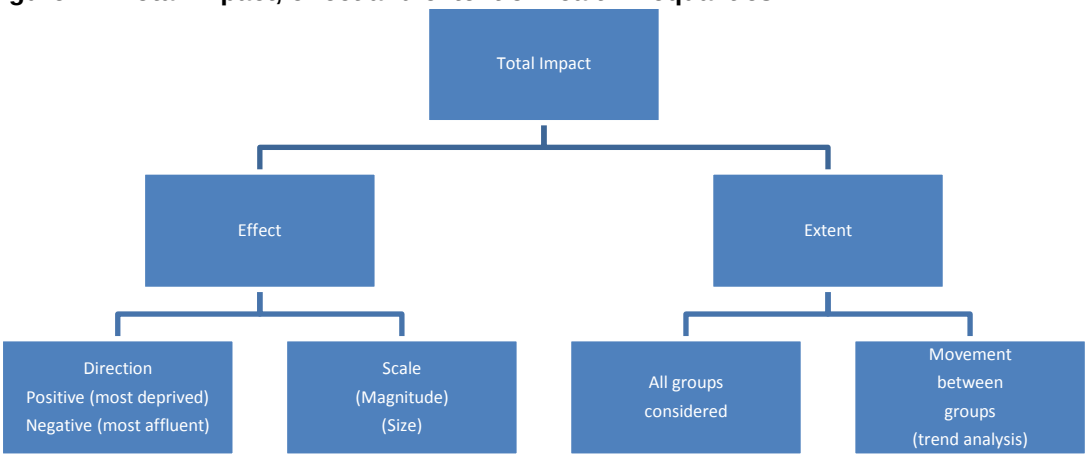
The inequality measurement debate in the context of cancer incidence or health status in general acknowledged that "re-distribution" (of incidence) from one social group to another was not comprehensible, although re-distribution of health resource (or income) was more readily understood. This distinction was implicit in the objective stated in Scottish Government policies such as *Improving Health in Scotland: the Challenge* (Scottish Government 2003) and *Equally Well* (Scottish Government 2008c). The concern was not that there are cancer incidence inequalities per se, but that these exist consistently within specific groups within society; the challenge is to improve health for all while also addressing these inequalities and to do so such that the poorest enjoy the health status of the richest (Marmot 2017); that is, to level-*up* the health gradient, not to level-*down* the health gradient.

### 1.4.2 Study design attributes

In terms of the measuring of those health status inequalities that are associated with SES, the concepts of total impact, effect and extent have been introduced by Mackenback and Kunst (1997). *Effect* measures the change in health by SES. Mackenbach and Kunst (1997) provided an example of a rate ratio for the most socioeconomically disadvantaged compared to the least socioeconomically disadvantaged. In the context of cancer incidence, by educational attainment, for example, an incidence rate ratio of 2.0 indicates *the effect* of low educational attainment on health is to double the risk compared to those with high educational attainment. *Extent* measures the inequalities in the population

as implied by the *population distribution* among the SES categories, i.e., "the gradient". *Total impact* takes into account both of these concepts. *Total impact* therefore, will increase not only because the *effect* of one further year of education on cancer incidence is larger but the *effect* is also greater if the difference in level of education between the upper and lower SES levels is larger (Figure 1.1). The Relative Index of Inequality (RII), Slope Index of Inequality (SII), Gini Coefficient, Population Attributable Risk and Index of Dissimilarity all measure total impact (Mackenbach et al 1997) (Table 1.3).

**Figure 1.1 Total impact, effect and extent of health inequalities**




Regardless of the measure used, the reference or comparison group must be defined. Options for a reference group include the best (group with most desirable outcome measurement), average or largest segment of the full population under review. Statistically, the largest group is the best reference (Harper et al 2008); however, when there are more than two groups, the referent category may be chosen to establish relative risks that are easiest to interpret, usually the higher SES strata. Other determinants of the selected measure are the timeframe covered (e.g. period in time or over time); the point in time and the nature of the social index being used (ordinal or non-ordinal such as ethnic categories) (Harper et al 2008).

Comparing extreme groups such as the best and worst off groups (pair-wise comparisons) are limited in that the two extremes may represent small populations. In addition, the measures may mask heterogeneous outcomes of the

intermediate groups. These "simple" approaches may be absolute or relative measures (Mackenbach et al 1997; Harper et al 2008).

To consider change over time, it is essential that measures reflect changes in the population distribution across the selected social group over time as well as the change in health status (Mackenbach et al 1997) (Figure 1.1).

Different social groupings may be used to evaluate SES inequality. If the groups can be ordered (e.g. income or SIMD), Slope Index of Inequality (SII) or Relative Index of Inequality (RII) can be applied. Otherwise, social groups such as sex or ethnic group or binary SES measures such as manual versus non-manual occupational groups, which are not ordered, require alternative approaches such as the rate ratio or rate difference (Harper et al 2008).

### 1.4.3 Measurement of health inequalities

As comprehensively reviewed by Blair et al (2013), various measures of inequality have been developed and can be described as measuring absolute or relative inequality using straightforward or complex approaches. Simple methods use two populations, most often defined as the least socioeconomically disadvantaged and most socioeconomically disadvantaged, or in some cases the median group where the median or least socioeconomically disadvantaged are frequently used as the reference group (Blair et al 2013). More complex methods capture the full range and distribution of the population groups and are complementary to the principle of proportional universalism (Marmot 2010) defined in the Scottish context as:

> "Proportionate universalism is the resourcing and delivering of universal services at a scale and intensity proportionate to the degree of need. Services are therefore universally available, not only for the most disadvantaged, and are able to respond to the level of presenting need" (NHS Health Scotland 2014).

The literature was clear that no one measure is likely to provide a

comprehensive understanding of the magnitude, direction and effect of SES inequalities (Mackenbach et al 1997; Harper et al 2008; Blair et al 2013). Examples from the literature demonstrate the importance of selecting the measure(s) to best reflect the objective under scrutiny. For example, when exploring the socioeconomic inequality that was associated with child dental health in Scotland, Blair et al (2013) established that the relative SES measures were not suitable given such a low prevalence of dental disease in the comparatively affluent groups.

Finally, summary measures may mask underlying factors; as a result, complementary measures may be required to gain as comprehensive an understanding of inequalities and how they are changing over time. Regardless, the number of measures used should be the lowest number required to enable complete and accurate interpretation (Blair et al 2013).

Mackenbach introduced the concept of simple versus sophisticated or complex measurements of socioeconomic inequality and provided the disadvantages and benefits of these two approaches (Mackenbach et al 1997). Because simple measures compare two groups, usually the health outcome of the low and high SES groups only, they are easier to calculate, do not impose many data restrictions and support straightforward interpretation. However, they omit available information such as the health outcome of SES groups in between i.e. the SES gradient. In contrast, the more complex methods do take into consideration the full SES spectrum, but because they are based on regression methods, the SES variable must be measured on an interval scale which is not always feasible, for example when measuring SES using occupation social class defined categorically rather than ordinally. The attributes of different socioeconomic inequality measures, including complex versus simple is provided in Table 1.3.

The list of tools to measure health inequality reflects two different approaches to describing those inequalities (Table 1.3). The most common approach focuses on measuring the social group differences assuming that these socioeconomic

groupings reflect the unequal and often unjust distribution of life resources and opportunities across a population. Only the Gini coefficient adopts the alternative approach of describing health status at individual level (Kawachi et al 2017). This avoids an a priori selection of population groups which may or may not meaningfully reflect the underlying inequality. However, in the context of health, the interpretation is awkward if not impossible because in the unlikely scenario where one person has all the cancer incidence, using the Gini Coefficient would be interpreted as positive in that incidence is less dispersed and less prevalent (Blair et al 2013). However, the Gini Coefficient value of 1.0 would indicate the highest inequality. This metric by design, is counter intuitive in this context. Furthermore, by definition, focus at the individual level completely removes that individual's social relations preventing any inquiry into the cause of (cancer risk) inequalities in society and presenting only a material or tangible cause of those inequalities (Table 1.3).

**Table 1.3 Measures of inequality**[1]

| Measure | Description |
|---|---|
| Rate Difference | ▪ Population measure<br>▪ Absolute disparity<br>▪ Arithmetic difference between two groups; one is the reference group<br>▪ The absolute 'gap'<br>▪ Measure of association<br>▪ Not weighted by population size in the SES domain<br>▪ Simple method |
| Rate Ratio | ▪ Population measure<br>▪ Relative disparity<br>▪ Generally compares the extremes but can select other groups (e.g. median) for comparison<br>▪ Measure of association<br>▪ Not weighted by population size in the SES domain<br>▪ Simple method |
| Slope index of inequality | ▪ Summary absolute measure covering full population and reflecting changes in distribution among the social groups over time.<br>▪ Each social group is given a score based on the midpoint of its range in the cumulative distribution of the population; a weighted index based on the size of the groups (population share)<br>▪ Complex method<br>▪ Total impact |
| Relative Index of Inequality | ▪ Population Measure<br>▪ Calculated as the Slope Index of Inequality/ Mean population health status<br>▪ Summary relative measure of disparity<br>▪ Complex method<br>▪ Total impact |
| Population-Attributable Risk | ▪ Population measure<br>▪ Difference between the overall rate and the rate for the best, expressed as a % of the overall rate can also be presented as absolute measure.<br>▪ Indication of the proportion of disease that could be eliminated if SES was eradicated<br>▪ Complex method<br>▪ Total impact |
| Index of dissimilarity | ▪ Individual measure<br>▪ Summary measure of inequality between social groups<br>▪ Simple method<br>▪ Total impact |
| Index of disparity | ▪ Summarises the difference between social groups rates and a reference rate<br>▪ Expresses summed differences as a proportion of the referenced rate<br>▪ Measure of disproportionality<br>▪ Complex method |
| Gini Coefficient | ▪ A summary measure describing social group difference for the entire population, at individual level.<br>▪ Measure of association between each individual's health and his/her share of health<br>▪ Not based on SES<br>▪ Complex method<br>▪ Total impact |
| Concentration Index | ▪ Can be absolute (ACI) or relative (RCI) ACI=RCI * mean of the health variable<br>Population ordered by social group status and cumulative percent of population is plotted against the groups share of total ill health<br>▪ Uses relative rank which indicates the cumulative share of the population up to the midpoint of each group interval<br>▪ Complex method<br>▪ Total impact |
| Theil Index | ▪ Measure of disproportionality<br>▪ Summary measure<br>▪ Sum of the product of each group's health status share of the whole population's total health status (within group inequality) and the natural log of each group's health status share (between group inequality)<br>▪ Applies when population of individuals is arranged into groups<br>▪ Complex method<br>▪ Total Impact |
| Between Group Variance | ▪ Summary Measure<br>▪ Sum of all squared deviations from a population average, weighted by population size<br>▪ Complex method<br>▪ Total Impact |

[1] Adapted from (Mackenbach et al 1997; Harper et al 2008)

Likewise the absolute measures of inequality identified above focus on a group's own socioeconomic circumstances independent of the circumstances of those around them. The inability to attain the normal level of consumption of their

community may lead to stress affecting health status. Relative measures reflect this and consider not only the socioeconomic circumstances of the individual group but also the socioeconomic circumstances of all groups in the population.

Most studies focusing on the association of low socioeconomic status and cancer incidence have used a range of metrics to quantify socioeconomic inequality including: i) Odds ratios (Marshall et al 1999; Conway et al 2010b); ii) Incidence rate differences (Anderson et al 2008); iii) Incidence rate ratios (Brown et al 1997; Weiderpass et al 2006; Anderson et al 2008; Baastrup et al 2008; Dalton et al 2008b); iv) Relative risk ratios (Mouw et al 2008; Clegg et al 2009; Sidorchuk et al 2009); v) Hazard ratios (Melchior et al 2005); vi) Attributed fraction (Hemminki et al 2003; Spadea et al 2009) and; vii) (European) age standardised rates (Kunst et al 2008). Very few studies have used the more complex measures identified here. Spadea and colleagues (2010) explored RII to study the cancer risk relationship to different indicators of adult socioeconomic circumstances in Turin, Italy (Spadea et al 2010). While Menvielle et al (2009) used the RII to explore lung cancer incidence association with education level across 10 European countries and Harper et al (2008) used lung cancer as an example to compare the full range of simple and complex measures of the association with low socioeconomic circumstances. By contrast to the simpler approaches, RII and SII are not often adopted and have not been widely used to measure inequalities for the cancer incidence and risk.

### 1.4.4 Definitions of burden of disease

Incidence and prevalence are the two measures of disease occurrence. Incidence risk is the proportion of people in a population that is initially free of disease who develop the disease within a specified time interval. It may be interpreted as the average probability, or risk, that an individual in a population will develop a disease during a specified period of time (Hennekens et al 1987; dos Santos Silva 1999) Equation 1.1.

**Equation 1.1 Incidence Risk**

$$\text{Incidence Risk} = \frac{\text{Number of new cases of disease arising in a defined population over a given period of time}}{\text{Number of disease-free people in that population at the beginning of that time period}}$$

Prevalence is the number of cases present in a population at a point in time and depends not only on the frequency with which new cases occur and are diagnosed, but also on the average duration of the illness reflecting recovery or death. Prevalence is the only measure of disease occurrence which can be established by cross-sectional studies and is valuable for establishing resource requirements in a population (Hennekens et al 1987; dos Santos Silva 1999) Equation 1.2.

**Equation 1.2 Point Prevalence**

$$\text{Point Prevalence} = \frac{\text{Number of existing cases of disease in a defined population at a point in time}}{\text{Number of people in that population at the same point in time}}$$

However, prevalence reflects not only incidence of disease, but also duration. It therefore is not effective for establishing and quantifying determinants of disease.

Because incidence risk assumes that those at the beginning of the time period in question are available throughout the study period, it assumes a stable population. However, in reality, populations are dynamic. Study participants may enter at different points, not just the beginning of the study and/or may be lost during follow-up for a number of reasons. In this more common case, incidence rate is used where the denominator reflects sum of the varying periods of follow-up for each person (Hennekens et al 1987; dos Santos Silva 1999) Equation 1.3.

**Equation 1.3 Incidence Rate**

$$\text{Incidence Rate} = \frac{\text{Number of new cases of disease arising in a defined population over a given time period}}{\text{Total person-time at risk during that period}}$$

Given the benefit of using all available information as the dynamic nature of populations, in this thesis incidence rate, summarised as incidence, is adopted as the basis of investigation of socioeconomic determinants of cancer. In this context, relative risk is used to estimate the magnitude of an association between exposure and disease and indicates the likelihood of developing the disease in the exposed group relative to those who are not exposed. As such, it is the ratio of the incidence of disease in the exposed group divided by the incidence of disease in the non-exposed group (Hennekens et al 1987).

The primary source for cancer incidence data is a country's population-based cancer registry.

### 1.4.5 Potential of data linkage in Scotland

Information on individuals from birth to death is available in the records of many institutions and agencies. These records can be merged or linked into a single comprehensive record using personal identifiers. This process is called record or data linkage. Linkage of the Scottish Cancer Registry with other health and administrative datasets, such as the Scottish Census, presents opportunities to evaluate more fully the factors associated with cancer risk.

In Scotland, data have been collected by the National Health Service (NHS) at national level for more than 40 years. The Information Services Division (ISD) of National Services Scotland is responsible for ensuring completion, quality and comparability of the registry data across Scotland and where relevant the UK (ISD 2010).

The Community Health Index (CHI) is a register of all patients who use the Scottish NHS. Patients have a unique 10 digit identification number with 96.5% to 99.9% of the Scottish population being covered (Pavis et al 2015). This unique number can be linked across time and location and ultimately across different datasets while maintaining data privacy. The CHI database has information on surname, forename, date of birth, sex and full postcode which can be used to link data items from two or more datasets, including the Scottish Cancer Registry (ISD 2017c), Scottish Health Surveys (ADLS 2017a) and the Scottish Longitudinal Studies (Boyle et al 2009), each of which have been used in the studies undertaken for this thesis.

In Scotland, information governance is overseen by the NHS Caldicott Guardian system which operates at local and regional levels and the Public Benefits and Privacy Panel (at the time of conducting the studies in this thesis, this body was known as the Privacy Advisory Committee) operates at the national level to ensure compliance with legislation, transparency and to maintain public trust. These structures and a network of safe havens (secure data access portals) support research by approved researchers using bespoke project specific subsets of data which are no longer identifiable (Pavis et al 2015).

#### 1.4.5.1 The Scottish Cancer Registry

The Scottish Cancer Registry held within NHS Scotland Information Services Division is a population-based database recording all new (incident) cancer cases that occur in Scotland. It therefore measures the incident occurrence of cancer in Scotland. In Scotland, approximately 55,000 cancer registrations are identified annually. The Scottish Cancer Registry database holds over 1,500,000 records dating back to 1958, when the registry began. The Scottish Cancer Registry provides historical trend and population-based data to monitor changes in cancer incidence and survival over long periods of time (ISD, 2017).

The Scottish Cancer Registry uses an electronic registration system to bring together information from hospital patient administration systems including patient discharges from hospital, radiotherapy, oncology, haematology and

pathology records; screening datasets; death records from National Records Scotland (NRS); private hospitals; and community prescribing. The European Network of Cancer Registries (ENCR) and UK Association of Cancer Registries guidelines support the quality and integrity of the data (ISD, 2017). The Scottish Cancer Registry is recognised as a high data quality dataset with less than one percent of cases identified through death certification only (Brewster et al 2002).

#### 1.4.5.2 Scottish Longitudinal Study

The Scottish Longitudinal Study (SLS), a continuous, multi-cohort study, is similar in design to the Office for National Statistics Longitudinal Study which was established in 1974 and is based on four semi-randomly selected birthdates as recorded in the relevant Census to extract a one percent sample of the 1971, 1981, 1991, 2001 and 2011 Censuses. The sample for Scotland however was deemed too small to support research, so it was discontinued in 1981 (Boyle et al 2009). The SLS was re-established in 2000 based on 20 birthdates (including the four dates used in the Longitudinal Study) resulting in a larger five point three percent proportion of the Scottish population, and although data was not available for the 1971 and 1981 Censuses, it did begin with the 1991 Census and has continued since. Following the 1991 Census, individuals born on one of the 20 birthdates are included whether or not they were born in Scotland. Census data (cultural, demographic, health, housing, employment and social variables) are updated regularly, vital events are continuously updated and health data, provided by Information Services Division, are linked on a project by project basis (SLS 2017).

The SLS is a nationally representative database and large (approximately 274,385 persons) (Hattersley et al 2007) compared to other cohorts supporting the study of relatively rare events (Boyle et al 2009), of which, cancer diagnosis is an example. The Census and many of the vital events registries supporting the SLS are compulsory, as a result, individuals living in communal establishments are included (Boyle et al 2009); this is not the case for other datasets, such as

the Scottish Health Survey also proposed for this thesis. This group of people is important given the focus on SES.

The National Health Service Central Register (NHSCR) is considered a high quality database based on the population registered with a doctor within the Scottish NHS system and is used to support linkage between the SLS and other datasets, such as the Scottish Cancer Registry (SLS 2017). However, the Census data are updated at the frequency of the Census itself, currently every 10 years; as a result, changes in socioeconomic status, for instance, are not known other than the 10-yearly points at which the Census is performed. Furthermore, no behaviour data are collected.

The SLS data were linked to other datasets through the NHSCR using name, sex and date of birth as minimum criteria and if an exact match was not possible, further information including address and postcode, name and birthdates of spouse or other household members were required ( Hattersley et al 2007; ADLS 2017b). The overall linked rate was 98.13%. The highest not traced rate at 2.52% were found among men who were aged 20 to 24 years old and among women aged 65-69 years old (3.34%). Using multivariate models including the effects of age, social class or economic status, country of birth, establishment type, marital status and local government region, Hattersey et al (2007) discovered that for both men and women and compared to those in unskilled manual positions, those in the armed forces (OR 3.51 95% CI 2.47, 4.97) were more likely to be untraced. Those born abroad, compared to those born in Scotland, also had greater odds of being untraced (OR 4.47 to 21.39). Finally, the odds of being untraced were also greater for the unemployed (OR 1.90 95% CI 1.40, 2.57) (Hattersley et al 2007).

#### 1.4.5.3 Scottish Health Survey

Similar to the Health Survey for England, the Scottish Health Survey (SHeS) provides a detailed picture of the health of the Scottish population in private households and supports monitoring health in Scotland (ADLS 2017a).

The Scottish Health Survey was first run in 1995 to capture information via personal interview with respect to cardiovascular and respiratory disease, self-assessed health and disability, common mental health problems and health service use. Important behavioural risk factors include: smoking, drinking, eating patterns, physical activity, use of prescribed medicines, anthropometric and biomedical measurements (including blood pressure, waist and hip circumference and lung function) and various biological samples (blood, urine and saliva) were also collected (Scottish Government 2017f).

The survey has been conducted for 1995, 1998, 2003 and yearly from 2008 to 2014. It is based on a stratified, clustered random probability sample of individuals living in private households across mainland Scotland and the larger islands. One in three postcode sectors with an average population of 5000 were selected for each survey (ADLS 2017a). The age range of survey participants has changed over the years. Initially, only adults aged 16-64 years were interviewed, then children two to 15 years old were included in 1998 and the age range for adults was extended to 74 years old. From 2003, all age ranges were included. From 2008, Health Boards were offered the opportunity to boost the number of adults (aged 16 years old and older) included in the survey. Various Health Boards have done so over the survey years since 2008, with the additional number of participants ranging from 475 in 2010 to 996 in 2013. The total number of adults participating in the survey fell from 7,932 in 1995 to 3,671 in 2014 (Scottish Government 2017f) and this drop in adult participation was first noticed in the 2012 survey (Scottish Government 2017f).

## 1.5 Cancer sites of focus in this thesis

The first comprehensive review of socioeconomic inequalities in cancer was published by the International Agency on Research in Cancer (IARC) in 1997 (Kogevinas et al 1997a). The focus of the IARC review was examining and explaining inequalities in cancer incidence, mortality and survival. The IARC review reported that the risk of lung, stomach, upper aero-digestive tract and cervical cancers was significantly greater for the lower socioeconomic groups.

For colon, breast, bone, ovarian and melanoma cancers, the relationship was reverse with greater risk associated with among the higher socioeconomic groups (Kogevinas et al 1997a).

Several behaviours such as smoking and alcohol consumption were strongly associated with increased cancer risk of lung and UADT cancers (Kogevinas et al 1997a). Smoking and alcohol were both recognised among the most important factors responsible for the SES gradient in cancer (Kawachi et al 2006). Furthermore, smoking and alcohol consumption, particularly heavy use, are more prevalent among lower SES (Brown et al 2016). Alcohol consumption was considered an established cause of mouth, pharynx, larynx, oesophagus, lung and breast cancer (Kawachi et al 2006). There was some evidence that malnutrition and heavy alcohol or smoking may interact to further increase cancer risk (Kawachi et al 2006). Finally, lung and UADT were selected *a priori* for this investigation given the cancer epidemiological focus of the University of Glasgow Community Oral Health Group.

### 1.5.1 Pathogenesis, ICD-10 and ICD-O-3 definitions

Lung and upper aero-digestive tract (UADT) including head and neck (larynx, oral cavity and oropharynx) and oesophageal cancers together comprise approximately 21% of the global cases diagnosed in Europe in 2012 (Ervik et al 2016).

The following outlines the ICD-10 codes, site grouping and morphology codes adopted in the analyses used in this thesis. The morphology of a cancer refers to the histological classification of the cancer tissue (histopathological type) and a description of the course of development that a tumour is likely to take: benign or malignant (behaviour).

#### 1.5.1.1 All cancer

First primary incident cancers excluding non-melanoma skin cancer (here after referred to as “all cancer” were defined by 3-digit ICD-10 codes C00-C96,

excluding C44 (non-melanoma skin cancer).

#### 1.5.1.2 Lung cancer

***Lung cancer pathogenesis***. The lungs consist of hundreds of lobules with each containing a bronchiole, the branches of the bronchiole, and ultimately cluster of alveoli (WCRF/AICR 2007). The bronchioles and alveoli (the parenchyma of the lung) are involved in gas transfer, e.g. taking in oxygen and exhaling carbon dioxide.

Four main histological subtypes of lung cancer constitute approximately 90% of all cases (Spitz et al 2006): i) Squamous cell carcinoma (30-35% of all lung cancers); ii) Adenocarcinoma (30-45%); iii) Large-cell carcinoma (9%); and iv) Small-cell carcinoma (10-15%) (WCRF/AICR 2007). The mix (of histological subtypes) has shifted over several decades which most likely reflects changes in the type of cigarettes smoked and the association of cigarette type with an increasing proportion of lung adenocarcinoma (Fehringer et al 2017). The average nicotine and tar content of cigarettes has decreased from the range 2.7mg to 37mg (high yield cigarettes) in the 1950's to the range 1.0mg to 13.5mg (low yield cigarettes) in the 1990s (Spitz et al 2006). Due to smokers' tendency to smoke more intensely and to inhale more deeply, high yield cigarettes are hypothesized to be associated with squamous cell carcinoma, while low yield cigarettes are associated with adenocarcinoma (Spitz et al 2006).

There are three commonly acknowledged pre-cancer conditions for lung cancer: i) Squamous cell dysplasia and carcinoma in-situ, a precursor of central bronchial carcinoma, ii) Adenomatous hyperplasia, considered a precursor for peripheral parenchymal adenocarcinoma of the bronchioles and alveoli, and iii) Diffuse idiopathic pulmonary neuroendocrine cell hyperplasia, considered to be rare and to be associated with development of neuroendocrine tumours of the lung (Spitz et al 2006).

Lung cancers are generally heterogeneous and consisting of cells of different histological subtypes. Pathological classification emphasizes the most common

histological subtypes. This common heterogeneity has led to the hypothesis that lung carcinomas arise from stem cell-like component or stem cell of the bronchial epithelium (Heighway D et al 2004).

***Lung cancer ICD-10 and ICD-O-3 code definition.*** First primary incident lung cancers were defined by 3-digit ICD-10 codes C33 and C34. The summarisation of the International Agency for Research on Cancer (IARC) defined morphology code groups was used for lung cancer morphology definitions (IARC 2009a) (Table 1.4).

**Table 1.4 Lung ICD-10 site and ICD-O-3 morphology codes**

| Site — morphology | ICD10 code (and ICD-O-3 morphology code) |
|---|---|
| Lung-adenocarcinoma | C33, C34 (M-8140, 8211, 8230-8231, 8250-8260, 8323, 8480-8490, 8550-8551, 8570-8574, 8576) |
| Lung-small cell carcinoma | C33, C34 (M-8041-8045, 8246) |
| Lung-squamous cell carcinoma | C33, C34 (M-8050-8078, 8083-8084) |
| Lung-other | C33, C34 (M-8010-8576, 8800-8811, 8830, 8840- 8921, 8990-8991, 9040-9044, 9120-9133, 9150, 9540-9581, 8000-8005) |

#### 1.5.1.3 Upper aero-digestive tract cancers

***UADT cancers.*** First primary incident upper aero-digestive tract (UADT) cancers consisting of head and neck cancers (HNC) and oesophageal cancers were defined by 3-digit ICD-10 codes C00–C14, C30-C32 and C15. They have been grouped mainly due to common aetiological factors in studies and reports by the International Agency for Research on Cancer (IARC) (Lagiou et al 2009).

***Head and neck cancers pathogenesis.*** Ninety percent of head and neck cancers are squamous cell carcinomas (WCRF/AICR 2007). Squamous cells are the flat skin like cells lining the mouth, nose, larynx and pharynx. Nasopharyngeal squamous cell carcinoma may be keratinising (the squamous cells include keratin, a protein forming nails and hair), non-keratinising and undifferentiated (CRUK 2017).

Most head and neck squamous cell carcinoma begin in the hypopharynx, larynx and in the oral cavity and oropharynx. Cancers arising in the sinuses and nasal cavity are rare worldwide and are strongly associated with number of occupational exposures to carcinogenic dusts most common in furniture, leather and shoe manufacturing and nickel refining industries (Littman et al 2006).

Mouth, pharynx (muscular cavity leading from mouth to larynx i.e. throat) and larynx cancers, like other types of cancer, are the result of genetic alterations that lead to small, localised lesions in the mucous membranes (e.g. leukoplakia or erythroplaki (Mayne et al 2006). These lesions may grow in an abnormal way (dysplasia). Erythroplakia is at higher risk of progressing to malignancy as it is more likely to include dysplasia (Mayne et al 2006). Carcinoma in-situ may develop from these lesions which may in turn become invasive cancers.

Head and neck cancers frequently present multiple, independent malignant foci; and as a result, second primary cancers are relatively common (WCRF/AICR 2007). The majority of laryngeal squamous cell carcinomas originate from the vocal apparatus of the larynx consisting of the vocal cords and the opening between them and the area above and below (Rousseau et al 2011). The most common oropharyngeal site of involvement is the base of the tongue (Rousseau et al 2011). Within the oral cavity, most tumours begin in the floor of the mouth, the front or bottom of the tongue to the side of the tongue or the soft palate (Rousseau et al 2011).

***Head and neck cancer ICD-10 code definition***. Cases of HNCs were defined by 3-digit ICD-10 category codes (without further sub-classification) and classified by anatomic site: lip (C00), oral cavity (C02, C03, C04, palate C05, C06), salivary glands (C07 and C08), oropharynx (C01, C09 and C10), nasopharynx (C11), hypopharynx (C12 and C13) and larynx (C32). Subtype groups adopted to analyse subsites of HNC reflected the anatomical relationship of these sites (IARC 2009; Junor et al 2010) (Table 1.5).

**Table 1.5 Head and neck cancer subsite ICD-10**

| Head and neck cancer subsite | ICD10 code |
|---|---|
| Larynx (including hypopharynx and piriform sinus) | C12, C13, C32 |
| Oral cavity (including lip) | C00, C02, C03, C04, C06 |
| Oropharynx (including base of tongue, palate and tonsil) | C01, C05, C09, C10 |
| Other (including parotid glands, other unspecified salivary glands, nasopharynx, other ill-defined sites in lip, oral cavity and pharynx, nasal cavity and middle ear and accessory sinus) | C07, C08, C11, C14, C30, C31 |

***Oesophageal cancers pathogenesis***. The oesophagus is the muscular tube connecting the pharynx (throat) to the stomach through which food is passed (WCRF/AICR 2007). Squamous cells line most of the oesophagus with the exception of the area where the oesophagus joins the stomach (gastric junction) which is lined by columnar epithelia cells. Oesophageal squamous cell carcinoma therefore arises from the upper areas of the oesophagus while oesophageal adenocarcinoma arises from the columnar epithelia cells at the gastric junction and is overwhelmingly found in the lower third of the oesophagus (Blot et al 2006). Oesophageal squamous cell carcinoma most often occurs in the middle third of the oesophagus, followed by the lower third and then the top third of the organ (Blot et al 2006).

The oesophageal epithelial squamous cells are exposed to carcinogens contained in food (WCRF/AICR 2007), alcohol (Blot et al 2006) and tobacco (Blot et al 2006). Repeated exposures are likely to irritate the lining, cause inflammation, progress to dysplasia, then carcinoma in-situ which ultimately can lead to malignancy (WCRF/AICR 2007).

Gastro-oesophageal reflux disease, where stomach acid repeatedly is regurgitated from the stomach up to the oesophagus increases risk of oesophageal adenocarcinoma. The resulting Barrett's oesophagus disease is characterised by replacement of the lower oesophageal squamous cells with columnar epithelial cells as part of the healing process (Blot et al 2006). It has

been hypothesized that the reflux generates reactive oxygen species causing oxidative stress and as a result DNA damage which in turn may cause mutations that may ultimately accumulate resulting in tumour formation (Peng et al 2009). *Helicobacter pylori* infection may also play a role in oesophageal cancer. The most common strain, CagA+, may protect against the development of oesophageal adenocarcinoma. The proposed mechanism was that the infection causes achlorhydria (the absence or reduction of hydrochloric acid in the gastric juices) which in turns reduces gastric acid reflux. Helicobacter pylori are a known risk factor of gastric cancer which is decreasing in prevalence in developed countries (Coupland et al 2012).

An exceptionally strong sphincter at the lower oesophagus which prevents food from moving to the stomach, can also increase the risk of squamous cell carcinoma 15-fold caused by chronic irritation or exposure to food borne carcinogens (WCRF/AICR 2007).

***Oesophageal cancer ICD-10 and ICD-O-3 code definition***. First primary incident oesophageal cancers were defined by 3-digit ICD-10 codes C15. The summarisation of the International Agency for Research on Cancer (IARC) defined morphology code groups was used for oesophageal cancer morphology definitions (IARC 2009a) (Table 1.6).

**Table 1.6 Oesophagus ICD-10 site and ICD-O-3 morphology codes**

| Site — morphology | ICD10 code (and ICD-O-3 morphology code) |
|---|---|
| Oesophagus-squamous cell | C15 (M-8050-8076,M-8083-8084) |
| Oesophagus-adenocarcinoma | C15 (M-8140-8141, 8143-8145,8190-8231, |
| Oesophagus-other | All remaining C15 morphologies |

## 1.6 Cancer incidence – disease burden

This section will briefly review historical trends and projections of cancer incidence at the global, European (EU), United Kingdom (UK) and Scotland levels for the anatomical sites and groupings which are the proposed subject of this thesis.

### 1.6.1 Global trends in cancer by country income level and Human Development Index

Torre et al (2016) recently undertook a detailed analysis of cancer incidence rates and trends (through to 2007) in the International Agency for Cancer (IARC) "CancerMondial" datasets. They found all cancer and lung cancer incidence to be higher in high-income countries, while incidence of oesophageal was highest in low- and middle-income countries. The authors report that trends in the incidence of all cancer and lung cancer in high-income countries are generally plateauing or starting to decrease, which they link to decreases in known risk factors; these findings contrast to the pattern they found in low- and middle-income countries, where rates of all cancer and lung cancer is increasing due, they suggested, to increases in smoking (Torre et al 2016).

These findings chime with the analysis of the global cancer burden by country level Human Development Index (HDI) undertaken by Fidler et al (2017) – which observed a higher cancer incidence burden in higher HDI countries (the most developed countries), there was a shifting trend to a greater proportion of the cancer burden projected to disproportionately affect less developed regions.

Global trends in head and neck cancer have been examined in detail by Maura Gillison's research group in the US and IARC. Chaturvedi and colleagues (2011 and 2013) from Gillison's group analysed Cancer Incidence in Five Continents databases Volumes VI to IX (1983 to 2002). They reported that among men in high-income (Western) countries oropharyngeal cancer incidence was increasing significantly, while oral cavity incidence was decreasing. These high-income country trends were accompanied by decreases in lung cancer incidence. In contrast, they found that among women both oropharyngeal and oral cavity cancer incidence rates were increasing. They also observed that oropharyngeal cancer increases were greater in developed countries, while oral cavity cancer increases were greater in less developed countries. The author's conclude that Oral human Papillomavirus (HPV) infection was emerging as the potential main driver for the rapidly rising incidence of oropharyngeal cancer. The risk association of oropharyngeal cancer (and not oral cavity cancer) first gained

prominence with a groundbreaking publication in the New England Journal of Medicine – again from Gillison's group (D'Souza et al 2007). This global phenomenon of a changing trend of flat-lining oral cavity and oropharyngeal cancer was related to changing population risk factors – Hashibe and Sturgis (2013) described it as "controlling a tobacco epidemic while a human papillomavirus epidemic emerges". The IARC analyses focused on the global burden of human papillomavirus related diseases. They reported that the greater global burden of HPV diseases and cancer falls on less developed regions of the world, except for oropharyngeal cancer where the greatest burden in more developed regions (Forman et al 2012).

### 1.6.2 All cancer trends

There were 14.1 million new cancer cases in 2012 worldwide. Estimated age – standard rates (world) per 100,000 populations indicate the United States (USA) (males: 347.0, females: 297.4) followed by the EU (males: 311.3, females: 241.3) had the highest incidence rates (Ervik et al 2016).

In Scotland, the number of new cases of cancer (excluding non-melanoma skin cancer) is predicted to rise by 33% between 2008-2012 and 2023-2027, mainly as a result of the population growing older (ISD 2015).

Projections for 2030 indicate that these figures will double (WCRF/AICR 2007). Cancer is increasing at rates faster than the increase in global population. It was becoming more common in high-income but also — and most of all — in middle- and low-income countries, absolutely and also relative to other diseases (WCRF/AICR 2007).

### 1.6.3 Lung cancer trends

In a review by Erik (2016), lung cancer was found to have been diagnosed as the most common cancer globally overall for several decades. In 2012, worldwide, there were estimated to have been 1.8 million new cases (12.9% of all incident cancer cases), 58% of which occurred in the less developed regions. The highest

estimated 2012 age-standardized incidence rates in men (per 100 000 men) were in central and eastern Europe (53.5 cases) and eastern Asia (50.4 cases). The highest estimated 2012 incidence rates in women (per 100,000 women) were in North America (33.8 cases) and northern Europe (23.7 cases) (Ervik et al 2016).

In Scotland, the projection of the number of new cases of lung cancer based on historical trends in cancer incidence and population estimates is expected to increase by 20% between 2008-12 and 2023-27 (ISD 2015). For females, the percentage increase is predicted to be 29% compared to 12% for males. While the number of cases will increase, the proportion of lung cancer relative to total cancers is expected to fall from 16.7% in 2008-12 to 15.0% in 2023-27. The number of cases of lung cancer in females was predicted to be more than in males for the first time in 2013-17 reflecting the historical patterns of smoking in the population (ISD 2015).

In Scotland, the age-standardised incidence rate (European) per 100,000 population is projected to decrease between 2008-12 and 2023-27 for men from 129 to 106. For females, the rate was projected to increase slightly between 2008-2012 and 2013-17 (from 94 to 98 respectively) before levelling off at around 96 in 2023-27. Lung cancer is predicted to continue to be the most common cancer in 2023-2027, although its proportion of all cancer will potentially fall slightly (ISD 2015).

## 1.6.4 Head and neck cancer trends

### 1.6.4.1 Head and neck cancers trends overall

In 2012 there were more than 686,000 HNC cases diagnosed worldwide by IARC, 4.9% of the total new cancer cases (Ervik et al 2016). This proportion is slightly less for EU and UK (3.9% and 3.3% respectively) (Ervik et al 2016). Estimated age-standard rates (world) per 100,000 population for the world, EU and UK are 9.2, 11.6 and 9.1 respectively (Ervik et al 2016).

#### 1.6.4.2 Head and neck cancer trends in the UK

Head and neck cancer in the UK is increasing and projected to increase at a rapid rate – this increase is made up almost entirely of oropharyngeal cancer, which is considered to be driven by HPV infection. The role of HPV in the aetiology of other sites of the head and neck is much more limited. In England, a robust Cancer Registry analysis found head and neck cancer incidence rates increased by 58% from 1995 to 2011, most rapidly for oropharyngeal cancer (Louie et al 2015). In the same analyses, incidence rates for lung cancer (strongly associated with smoking) were found to remain stable over the same period. This analysis of oropharyngeal cancer incidence trends paralleled increased rates for genital warts and genital herpes in England from 1995 to 2011, which were also reported in this paper (Louie et al 2015). These trends were also replicated in Scotland (Purkayastha et al 2016).

#### 1.6.4.3 Head and neck cancer trends in Scotland

In 2008-12, cancer of the head and neck was the fifth most common type of cancer in Scotland diagnosed in males (ISD 2015); this position is projected to remain in 2023-27 and the number of new cases is projected to increase by 37% between 2008-12 and 2023-27 (28% for males and 57% for females) (ISD 2015). The age-standardised incidence rate is also projected to increase for both males and females over this period, but the increase in the rate for females is expected to be larger than for males (6% increase for males and 31% for females) (ISD 2015). Despite these projected increases for females, head and neck cancer is expected to remain in a similar position of rank of total cancer cases compared to 2008-12 (13th in 2008-12 and 12th in 2023-27) (ISD 2015).

#### 1.6.4.4 Subsites of head and neck cancer

More recently, Louie et al (2015) reviewed trends in and projected incidence of oropharyngeal cancer for England. They noted a 58.9% increase in incidence over the 16 year period to 2011 with rates increasing for both sexes, but particularly for males. Over the period, nasopharyngeal and laryngeal cancer incidence was

found to decrease for both sexes while the oropharyngeal cancer incidence was the anatomic site that increased the most in both sexes but again, at a greater rate for males. Projected incidence analysis found that head and neck cancers would increase by 34.5% for men and 48.9% for women. As a result, head and neck cancers will move from the 15$^{th}$ largest number of cases in 2011 to 6$^{th}$ in 2025. Oropharyngeal cancers were ranked as the most frequent head and neck cancer in 2025, representing 35% of head and neck cases, whereas laryngeal cancer previously had constituted the largest proportion of head and neck cancers in 1995. The authors explained that the increase and the shift to oropharyngeal cancer resulted from reducing levels of smoking and increasing Human papillomavirus (HPV) exposure, due to a change in sexual behaviour and an increasing oncogenic role of HPV in oropharyngeal cancer.

In Scotland, Purkayastha et al (2016) recently undertook a similar analysis reporting similar trends to Louie et al (2015) for England with head and neck cancer incidence rising rapidly over recent decades. This was reported to be due to a steady increase in oral cavity cancer but a rapid rise in oropharyngeal cancer in the most recent decade. This was also found by Junor and colleagues (2010) who showed that oropharyngeal cancer was the most rapidly rising cancer in Scotland in both men and women, with increases outstripping (relatively) those observed for malignant melanoma and cervical carcinoma.

Shield et al (2017) used data from the International Agency for Research on Cancer and GLOBOCAN to project the global incidence of lip, oral cavity and pharyngeal cancers to 2035. In 2012, approximately 71% of cases occurred in men, 29% in women. The number of global cases was predicted to rise by 62% from 529,500 to 856,000 by 2035. The greatest volume of global cases were diagnosed in the oral cavity, followed by oropharynx, nasopharynx, hypopharynx and finally lip followed by ill-defined cases of these sites. The distribution of subsites differed significantly depending on the region, reflecting the different aetiologies of the cancers and demographics of the regions. The rates of oropharyngeal cancers were elevated in Europe and were associated with alcohol consumption, tobacco smoking and human papillomavirus.

## 1.6.5 Oesophageal cancer trends

### 1.6.5.1 Oesophageal cancer trends overall

Oesophageal cancer was reported as the eighth most common cancer overall (Ervik et al 2016). In 2012 worldwide, there were estimated to be 456,000 new cases (3.2% of all incidence cancer cases); 80% of which occurred in less developed countries (Arnold et al 2017). Oesophageal cancer incidence rates worldwide in men are more than double those in women (with a male-to-female ratio of 2.4:1) (Ervik et al 2016). Age-standardised rates (world) compared for 2008 incidence indicate that the UK has the highest rate compared to the world and EU (6.5, 3.4 and 5.9 respectively) (Ervik et al 2016). In 2030 the age-standardised incidence rate (world) is projected to increase to 10.01 for men with cumulative risk of oesophageal cancer increasing from 1.05 in 2005 to 1.18 in 2030 for men 75 years old or younger (Arnold et al 2017).

In a cancer projection report by ISD (2015), in Scotland, the number of oesophageal cancer cases is expected to increase by 16.1% from 2008-12 to 2023-27 with males increasing at more than twice the rate compared to females (20.1% versus 8.7%) (ISD 2015). The age-standardised rate (European) per 100,000 for males will be twice that for females over the projection period 2008-12 to 2023-27 which increased slightly in both sexes (22 to 24 for males and 11 to 12 for females) (ISD 2015).

### 1.6.5.2 Oesophageal cancer trends by histology

The two main histological types of oesophageal cancer are adenocarcinoma and the dominant histology, squamous cell carcinoma; 87% of all new cases were squamous cell carcinoma globally (Arnold et al 2017). The incidence of squamous cell oesophageal carcinoma was declining in North America and Northern Europe, but this decline was compensated for by rapid increases in the incidence of adenocarcinoma (Arnold et al 2017).

In the United Kingdom the incidence of oesophageal adenocarcinoma diagnosed among men overtook oesophageal squamous cell carcinoma incidence before 1985 (Arnold et al 2017). Of the 12 countries reviewed, the UK was the first to demonstrate this trend. Despite declining squamous cell carcinoma incidence rates in the UK, incidence of oesophageal cancer is expected to increase due to an ageing population. A doubling of annual incidence by 2010 was expected in the UK (Arnold et al 2017). Declines in squamous cell carcinoma incidence were attributed to reduced prevalence of both smoking and heavy alcohol consumption while increases in adenocarcinoma incidence were attributed to an increased prevalence of obesity which was estimated to cause 43% of adenocarcinoma cases in the most developed societies (Arnold et al 2017). Obesity was considered an independent risk factor of adenocarcinoma as well as a cause of gastro-oesophageal reflux disease which itself was recognised as the primary risk factor of oesophageal adenocarcinoma.

### 1.6.6 Cancer incidence – disease burden summary

Taken together, lung and upper aero-digestive tract (LUADT) cancer comprises 21.1%, 17.1% and 18.0% of all cancer incidence in the world, EU and UK in 2012, respectively (Ervik et al 2016). In Scotland, these cancers constitute an even higher proportion of all cancer incidence at 23.3% for the 2008-12 period (ISD 2015). This disease burden in 2008-12 was greater in males compared to females (26.9% versus 20.0%) (ISD 2015) and although these proportions are expected to fall in 2023-27, the absolute number of cases will increase over the period by 16.5% for males and 30.3% for females (ISD 2015). For the HNC, oropharyngeal cancer was the main site with a striking increasing trend in incidence globally, in the UK and in Scotland and this has been associated with human papillomavirus (HPV) (Anantharaman et al 2013).

Moreover, the incidence risk was polarising disproportionately among the lower socioeconomic groups and this looks set to continue and worsen both globally and locally.

# 1.7 Cancer incidence and behaviours: smoking, alcohol, diet and nutrition, obesity, exercise and HPV

## 1.7.1 All cancer

Parkin et al (2011) conducted several analyses to estimate the fraction of cancer attributable to lifestyle in the UK in 2010. For both men and women, tobacco smoking was the most important risk factor for cancer causing 19.4% of all new cancer cases in 2010 (23.0% men and 15.6% women). This was followed by diet (9.2%), then overweight and obesity (5.5%), alcohol consumption (4.0%) and ultimately physical exercise (1.0%). These five behaviours accounted for 35% of the cancers incident in the UK in 2010. For men, tobacco smoking and deficient fruit and vegetable consumption was responsible for 6.1% new cancers along with alcohol attributing 4.6% while for women, being overweight or obese attributed 6.9% (due to association with breast cancer).

### 1.7.1.1 Smoking

In most industrialised countries, tobacco smoking is more prevalent among low socioeconomic classes than higher socioeconomic classes. However, tobacco smoking has become the main contributor to total mortality in developing countries as well (Stellman et al 1997). Thun et al (2009) describes the characteristics of the tobacco smoking epidemic in terms of four stages: i) Stage 1 is described as low prevalence of male smoking (20%) with no evidence of female smoking prevalence; ii) Stage 2 is typified by increased smoking prevalence among men of 50% and increasing smoking prevalence among women along with earlier age of initiation of smoking. At this stage, knowledge of the associated health risks and tobacco control policies are limited; iii) Stage 3 is described as reduction in the prevalence of smoking by men, a more gradual decline in smoking by women with the prevalence among men and women converging. Nevertheless, mortality and morbidity associated with smoking continue to rise. Tobacco control activities influence the perception of smoking acceptability among the more educated; and iv) Stage 4 is described as

continued decline in smoking prevalence by both men and women with tobacco related deaths for men peaking, then declining while for women, deaths rise.

#### 1.7.1.2 Alcohol

Møller et al (1997) evaluated the role of socioeconomic inequalities in alcohol related cancers. They conclude that alcohol drinking in the social class gradients for alcohol related cancers such as UADT cancers is very likely in France, Italy, New Zealand and probably in other countries as well. The effects of alcohol intake and its association with cancer incidence reflects several factors such as tobacco use, dietary practice, physical activity, occupation and environmental exposure (Marshall et al 2009). It is proposed that the greatest risk association of alcohol with cancer incidence is likely to be concentrated among individuals who have an exceedingly high intake of alcohol through heavy intake, binge drinking and alcohol abuse (Marshall et al 2009).

#### 1.7.1.3 Diet

Willett et al (2009) discuss the history of the awareness of diet as a key cause of cancer incidence and suggest that 35% of cancer deaths in the US might be due to dietary factors. Fruits and vegetables have been identified as important in reducing the risk of many specific cancers. Willett et al (2009) warn that the benefits may be overstated given that most of the studies have been case-control studies and therefore susceptible to bias. Furthermore, the benefits of just one fruit or vegetable type may be interpreted as the benefit of fruits and vegetables in general, but may in fact be the result of chance due to multiple testing. Large prospective studies will assist in avoiding these biases (Willett 2009).

Potter (1997) identifies four types of changes in eating patterns that could produce cancer: i) Imbalance between energy intake and output; ii) Changes to the intake of either micro or macro nutrients or both; iii) Deficiencies in nutrients or bioactive compounds; and iv) The presence of substances that are not part of the normal diet which when an individual is exposed to that

substance, causes a metabolic response which may lead to cancer. He summarises that the human "original diet" provided regular exposure to a variety of substances required for human metabolism. The "original dietary pattern" resulted in variable intake and limited risk of obesity and involved very little or no intake of alcohol. Abandonment of this "original dietary pattern" has resulted in lower fruit and vegetable consumption, high intake of fat, grains and alcohol leading to increased obesity and greater cancer incidence. Potter (1997) goes on to explore the possible association of diet with social class citing that the consumption of fat, meat, alcohol, fruit and vegetable intake are socially distributed. He reports that living alone, a lower income, reduced expenditure on food and unemployment are significant predictors of poorer-quality diet among those 55 years old and older.

#### 1.7.1.4 Exercise

Lee et al (2009) reported that the global estimate for prevalence of physical activity among those 15 years old and older was 17% with a range of 11% in Africa to a high of 24% for Europe. Mechanisms for lower rates of cancer incidence among those who are more active are summarised as: i) Change in sex hormone levels; ii) Alteration of body fat; iii) Change in intestinal transit time; and iv) Change in immune function. Parkin et al (2011) discussed the importance of a history of physical exercise; i.e., levels of physical exercise over an extended period such as 20 years.

#### 1.7.1.5 Human papillomavirus

There is substantial evidence that infectious agents play a causal role in many human cancers. Human papillomavirus has been associated with cervical cancer (Mueller et al 2009) and potentially, oropharyngeal cancer (Gillison 2007). HPV infections are generally benign and ubiquitous (Mueller et al 2009). However, persistent infection with an oncogenic genotype causes most cervical cancers and a smaller proportion of other cancers, totalling 500,000 cases per year worldwide. HPV 16 is very prevalent and carcinogenic type; it is responsible for half of the HPV 16 associated cancer burden (Mueller et al 2009). According to

population-based surveys in industrialised countries, Sanjosé et al (1997) reported that men of low socioeconomic status indicated fewer sexual partners than men of high SES but the same could not be said for women.

## 1.7.2 Lung cancer

### 1.7.2.1 Smoking

Smoking was found by Fehringer et al (2017) in their comprehensive review to be the primary cause of lung cancer accounting for more than 80% of all lung cancer diagnoses. This has been shown to rise up to 90% in countries with a history of high tobacco consumption (Pesch et al 2012). Cigarette smoking was most strongly associated with lung cancer than any other anatomical site and was associated with all histological types. The SYNERGY project comprising of pooled data from eight European and one Canadian case-control studies included 13,169 cases and 16,010 controls. Using the project, Pesch et al (2012) established that males smoking 30 or more cigarettes per day had an OR of 103.5 (95% CI 74.8, 143.2) for squamous cell carcinoma, an OR of 111.3 (95% CI 69.8, 177.5) for small cell lung cancer and an OR of 21.9 (95% CI 16.6, 29.0) for adenocarcinoma. For women the ORs were similar and demonstrated a very high risk association with an OR of 62.7 (95% CI 31.5, 124.6) for squamous cell carcinoma, an OR of 108.6 (95% CI 50.7, 232.8) for small cell carcinoma and an OR of 16.8 (95% CI 9.2, 30.6) for adenocarcinoma. Adenocarcinoma was the most common subtype for women and never smokers and incidence increased over time. The authors offered possible explanations for the increase in adenocarcinoma observed including the improvement in chest imaging and consequent detection of peripheral pulmonary nodules, changes in the WHO classification and improved staining techniques resulting in fewer large cell carcinoma diagnoses and more adenocarcinoma diagnoses. The authors also proposed that cigarette modifications including tar and nicotine content and filter introduction may have resulted in greater inhalation of smaller particles which then penetrate to the distal airways. The study found that higher lung cancer risks were associated with younger starting age. The investigators hypothesized that this may reflect greater susceptibility at a younger age and/or longer smoking duration, but

concluded that the data suggested smoking habits were the more likely explanation. This was consistent with the IARC review which identified smoking duration as the strongest determinant of lung cancer diagnosis (IARC 2004; Pesch et al 2012).

Using the International Lung Cancer Consortium which consisted of 18 international case-control studies including 2,504 cases and 7,276 controls who were never smokers and 10,184 cases and 7,176 controls who were ever smokers, Kim et al (2014) studied second hand smoke and lung cancer risk by histological type. In this large collaboration supporting subtype analysis, the investigators identified that the risk of lung cancer was increased by 30% among ever smokers compared to never smokers (OR 1.31 95% CI 1.17, 1.45). The increased lung cancer risks were least elevated for adenocarcinoma with an OR of 1.26 (95% CI 1.10, 1.44), followed by squamous cell carcinoma with an OR of 1.41 (95% CI 0.99, 1.99), then an OR of 1.48 (95% CI 0.89, 2.45) for large cell lung cancer and finally an OR of 3.09 (95% C: 1.62, 5.89) for small cell lung cancer. The lower risks associated with adenocarcinoma and large cell carcinoma were hypothesized to reflect that these cancers arise from more peripheral sites of the lung while small cell lung cancer and squamous cell carcinoma which mainly occur in the larger central bronchi and are therefore more exposed to the carcinogens in smoke. The aerodynamic features of the carcinogenic particles determine where in the lung they are deposited. Larger particles are more likely to be deposited in the central bronchial regions. As a case-control study, the results are subject to recall bias, different potential influences on the mix of hospital and population-based studies, variation in definition of never smokers across studies and misclassification of ever smokers as never smokers.

#### 1.7.2.2 Alcohol

Fehringer et al (2017) pooled 22 case-control and cohort studies to create the largest international case-control study to date to study lung cancer risk and alcohol. The study was composed of 2,548 never-smoking lung cancer patients

and 9,362 never-smoking controls and was part of the International Lung Cancer Consortium and SYNERGY project. Confounding by smoking, the dominant risk factor for lung cancer was addressed by focusing only on never-smokers while SES was considered by adjusting for education. Results for overall alcohol consumption indicated that compared to non-drinkers, drinking up to 10 to 19.9 grams of alcohol per day reduced lung cancer risk (OR 0.79 95% CI 0.65, 0.96). Reduced lung cancer risk was also associated with both lung adenocarcinoma and squamous cell carcinoma but risk was increased for small cell carcinoma of the lung (OR range 1.2 to 1.7 for all alcohol consumption categories). Low and moderate levels of wine drinking (up to 29 grams per day) were associated with reduced lung cancer risk (OR 0.62 95% CI 0.43, 0.89) while low levels of spirits consumption up to 4.9 grams were associated with lower lung cancer risk (OR 0.77 95% CI 0.66, 0.91). Beer was associated with insignificant but modest increased lung cancer risk for most consumption categories compared to non-drinkers. The authors proposed that the flavonoid concentrations in wine may explain the reduced risk effect, alternatively confounding by other lifestyle behaviours such as increased exercise levels by wine drinkers compared to beer drinkers who are perceived as more healthy and may have contributed to the observed outcome. They discussed SES and the “protective” effects reported, proposing that non-drinkers may be a unique group of the population with either lower SES or medical conditions that could confound associations with lung cancer risk; along these lines, they suggested that adjustment by education may not have captured all aspects of SES.

The lungs are not directly exposed to alcohol unlike the oral cavity, hypopharynx and oesophagus. Furthermore, alcohol is absorbed into the blood stream from the stomach and small intestine and transported to the liver where it is fully metabolised. Although metabolites may come in contact with lung tissue via the blood stream and these may act as lung cancer carcinogens, the concentrations are likely to be too low to have an effect. Marshall et al (2009) in a comprehensive review of the literature (albeit not systematic) identified a number of studies suggesting alcohol consumption and lung cancer association, but also reported a number indicating no increased lung cancer risk with

increased alcohol consumption. Using hospital admission for diseases associated with alcohol consumption, Grant (2015) in his PhD thesis on the relationship of alcohol and cancer could not find an association of lung cancer risk with total alcohol intake or by histological type of lung cancer with the exception of adenocarcinoma of the lung which he concluded merited further investigation. He suggested however that there was evidence of a protective effect of wine compared to an increased risk observed from spirit and beer consumption but attributed this finding to possible confounding from SES.

Given the overriding role and dominant role of tobacco smoking in lung cancer carcinogenesis, the complications of capturing all aspects of the smoking habit (e.g. brand, filter, depth of inhalation and length of time smoke is held in the lungs) and reviews published in the early 2000s, both Grant (2015) and Marshall (2009) concluded that alcohol is not likely to be a significant risk factor for lung cancer (Marshall et al 2009; Grant I 2015).

#### 1.7.2.3 Diet and nutrition

The World Cancer Research Fund and American Institute for Cancer Research evaluated food, nutrition, physical activity and the prevention of cancer publishing their findings in 2007 (WCRF/AICR 2007). Smokers tend to have less healthy diets, are more inactive and leaner than non-smokers; as a result these features may confound the results of nutrition's association with lung cancer risk. Various classes and components of food were evaluated including fruits, non-starchy vegetables, selenium containing foods, quercetin containing foods, red meat, processed meat, fat, grains, pulses, fish eggs, plant oils, coffees, teas and various vitamins. The scientific panel included 561 publications for the lung cancer risk assessment. They concluded that for non-starchy vegetables, there was limited evidence of protection against lung cancer. For fruits, 64 studies were reviewed and several different meta-analyses were performed. The protective effect of increased fruit consumption on decreased lung cancer incidence ranged from a 6% decrease for every 80g consumed daily to 23% decrease for those who ate the most fruit. Vitamin C, carotenoid, phenol,

flavonoid and other photochemical content of fruit and their antioxidant action trapping free radicals were identified as providing a possible biological benefit leading to reduced lung cancer risk. They also noted that flavonoids inhibit expression of cytochrome P450 which was important in metabolising toxins associated with increased lung cancer risk.

The panel concluded that there was limited evidenced that red meat was associated with increased lung cancer risk, although processed meat was associated with increased risk of a lung cancer diagnosis (WCRF/AICR 2007).

#### 1.7.2.4 Obesity

Hidayat et al (2016) conducted a systematic review and meta-analysis of six identified prospective studies comprising of 5,827 lung cancer cases and 831,535 participants to investigate abdominal obesity and lung cancer risk (Hidayat et al 2016). Five of these cohort studies enabled analysis of waist circumference and lung cancer risk association after adjustment for body mass index (BMI). Each 10 cm increase in waist circumference resulted in a 10% increase in lung cancer risk association (RR 1.10 95% CI 1.04, 1.17) and each 0.1 unit increase in the waist to hip ratio resulted in a 5% increase in lung cancer risk association (RR 1.05 95% CI 1.00, 1.11). Six cohort studies were used to evaluated waist hip ratio and lung cancer risk where results indicated that greater waist to hip ratio was only associated with greater lung cancer risk for former smokers (RR 1.11 95% CI 1.00, 1.23). The authors concluded that abdominal obesity may play a role in the development of lung cancer and may be a better predictor of lung cancer risk than BMI. The authors discussed possible biological pathways to explain the results including hyperinsulinemia stimulation of insulin-like growth factor 1, cell proliferation and deregulation of apoptosis (Sartorius et al 2016), reduced sex hormone binding globulin levels and increased levels of unbound androgens and oestrogens which are more strongly associated with abdominal fatness as opposed to overall body fatness. To support these possible biological pathways, they reported that lung cancer cells have receptors for oestrogen and androgens while both small cell lung cancer and non-small cell lung cancer respond to the

presence of insulin like growth factors (IGF-1) in vitro. The apparent inconsistency of reduced lung risk association with BMI was explained via residual confounding by smoking as it has been established that lower BMI was associated with increased lung cancer risk (Dewi et al 2016). The authors noted that abdominal fatness and cancer both were both associated with reduced physical exercise, smoking and poor diet so the results reported may reflect confounding by these and other unknown risk factors. Given all the studies included were large prospective cohorts, there was reduced chance of recall or selection bias.

#### 1.7.2.5 Exercise

The World Cancer Research Fund/Americian Institue for Cancer Research (2007) studied the effect of all kinds of physical activity (defined as occupational, household, transport and recreational) on cancer prevention. Twenty-eight studies most of which were cohort design were reviewed by the panel. Generally, increased physical activity was associated with reduced lung cancer risk although the evidence was weak and mired by the possibility of reverse causation due to chronic lung disease. The mechanism that physical activity may provide a benefit was hypothesised as the increase in metabolic rate and therefore increased maximal oxygen intake which occurs with sustained moderate physical activity and as the body becomes more efficient, blood pressure is reduced along with insulin resistance.

Further theories on the protective effect of physical activity on cancer risk were provided by Shi et al (2015) who conducted a meta-analysis of 30 studies to evaluate the association between household physical activity and cancer risk. They established that total cancer risks were reduced by 16% for those with the highest category of activity compared to the lowest category of activity (RR 0.84 95% CI 0.76, 0.93). A dose-response relationship was evident and calculated at 0.99 decrease risk for each additional hour per week increase (95% CI 0.98, 0.99). The investigators proposed a number of mechanisms explaining the preventative role of physical activity and cancer incidence. Hyperinsulinemia

and insulin resistance are associated with increased cancer risk. They proposed that hyperglycaemia up regulates insulin/IGF-1 and inflammatory cytokines circulating in the blood. These may indirectly affect cancer cell development and proliferation. Physical activity may reduce insulin resistance and lower fasting insulin levels and therefore inhibit cancer cell proliferation and cellular transformation. Inflammatory cytokine markers were also considered to be associated with cancer risk and physical exercise may reduce the concentration of adipocytokines or increase anti-inflammatory levels or reduce adiposity generally. It was also proposed that there was a possibility that physical exercise increases the immune system surveillance of cancer cells by increasing the number and activity of immune cells such as macrophages, natural killer cells, lymphokine–active killer cells and cytokines. The study adds to the understanding of physical activity and cancer risk; however measurement of activity and reporting of activity was variable across studies which may lead to biased results. In many of the included studies, physical activity was self-reported so misclassification may have occurred (Shi et al 2015).

### 1.7.3 Head and neck cancer

Winn et al (2015) summarised several studies exploring the causes and mechanisms of head and neck cancer (oral cavity, larynx and pharynx) via the International Head and Neck Cancer Epidemiology Consortium (INHANCE). Thirty five international case-control studies have been pooled to provide a very large dataset of relatively rare cancers composed of 25,500 patients with head and neck cancer and 37,100 controls. Data inconsistencies and selection bias were minimised through requirement of structured questionnaires and recruitment protocols. Large sample and increased numbers in the referent categories enabled more reliable risk estimates; evaluation by cancer subsite, time and geography; and ability to assess relative role of known risk factors such as tobacco and alcohol as well as other possible aetiologic factors such as height, exercise, weight, diet and sexual behaviour.

#### 1.7.3.1 Smoking

Key findings from the INHANCE studies were that cigarette smoking among never drinkers increased risk of head and neck cancer two-fold (OR 2.13 95% CI 1.53, 2.98). Quitting smoking resulted in reducing risks to those of never smokers one to four years after stopping. With respect to tobacco type, increased risk of head and neck cancer was greatest for cigarettes (OR 3.46 95% CI 3.24, 3.70) followed by cigars (OR 2.54 95% CI 1.93, 3.34) and finally pipes (OR 2.08 95% CI 1.55, 2.81). It was highlighted in the INHANCE tobacco analysis study was the finding that a two-fold increased risk occurred for even the lowest category of smoking where one to 10 cigarettes were smoked daily. Finally, for more than 15 cigarettes smoked per day, the head and neck cancer risks associated with smoking fewer cigarettes over a longer period were greater compared to smoking more cigarettes over a shorter period of time (Peto 2012). By subsite, smoking risks were greater for laryngeal cancer compared to oral cavity and pharyngeal cancer risks (Winn et al 2015).

INHANCE focused on the oral cavity, larynx and pharynx subsites of the head and neck as these sites had the greater numbers of tumours and higher risk association with alcohol and smoking risk factors. In 2002, IARC identified cancers of the nasal cavity and paranasal sinuses as causally related to smoking. Thun et al (2009) reported that several European, Asian and US studies had found relative risks estimates of squamous cell carcinoma of these sites equal to approximately 2.0 in current smokers compared to lifelong non-smokers. They highlighted that the lower rate compared to say lung cancer risk reflects the exposure of these sites only during exhalation (Thun et al 2009). Conway (2010) commented on an INHANCE study exploring the excess odds ratio by pack-years (exposure duration) modified by cigarettes smoked/day (exposure rate). Their results suggested that more cigarettes/day for a shorter period of time was less harmful than fewer cigarettes/day for a longer time period. The results also suggested that for laryngeal cancer, the greater risk associated with smoking was due to the number of cigarettes smoked/day, not the number pack-years; however there was no safe minimum limit where smoking was protective.

***Conway et al (2010)*** used data from the INHANCE Consortium of case-control studies of head and neck cancer to explore the effect of total exposure and exposure rate for alcohol and smoking on the risk of head and neck cancer. Fifteen studies with detailed information on cigarette and alcohol consumption behaviour were used to calculate the excess odds ratio (EOR) by total exposure (pack-years and drink-years) and modification by exposure rate (cigarettes/day and drinks/day). Cases of laryngeal cancer totalled 1,761 while 2,453 pharyngeal and 1,990 oral cavity cancers were used in the alcohol analysis. For the smoking analysis, 2,551 laryngeal, 3,693 pharyngeal and 3,116 oral cavity cancers were identified and 8,000 were included in the analysis. Again, for alcohol assumption up to 10 drinks daily, EOR results suggested that more drinks/day for a shorter period was more harmful than fewer drinks/day for a longer period. EOR drink-year estimates varied by site, with pharyngeal/oral cancer risk the greatest. This suggested that the variable drink-years was the cause, not the variable drink-days. For both behaviours, there was no safe minimal limit where consumption was protective for these cancers.

#### 1.7.3.2 Alcohol

Winn et al (2015) reported alcohol drinking among never smokers was also associated with two-fold increased risk but only for those drinking three or more drinks daily (OR 2.04 95% CI 1.29, 3.21). Quitting alcohol took 20 years to achieve reduction of a never drinker's head and neck cancer risk and took much longer than achieving benefits from smoking cessation (1 to 4 years). Recognising ethanol is a precursor to acetaldehyde and was classified by IARC as a Group 1 carcinogen in 2009 (IARC 2009b). One INHANCE study computed the amount of ethanol in 30 drinks of beer, wine, or spirits consumed in a week, to determine the odds ratio of head and neck cancer associated with spirit only, beer only and wine only drinkers. They found that compared to never drinkers, the risk of head and neck cancer increased over three-fold for spirit drinkers (OR 3.6 95% CI 2.2, 5.8), over five-fold for beer drinkers (OR 5.4 95% CI 3.1, 9.2) and more than six-fold for wine drinkers (OR 6.3 95% CI 2.2, 18.6). Greater harm occurred when drinking more alcohol per week over a shorter time compared to less alcohol per

week over a longer period. This finding is notable for being the converse of the smoking risk association. Alcohol consumption risks were greater for oral cavity and pharyngeal cancers compared to laryngeal cancer.

In their review of alcohol and cancers, Marshall et al (2009) reported that the association of alcohol intake with increased oral cancer risk was one of the more consistent findings in the epidemiology of oral cancer. Citing both cohort and case-control studies, Marshall reported that those in the highest alcohol consumption categories or with history of alcoholic cirrhosis or pancreatitis had increased risk of oral cancer four to 12 times that of the general population. Likewise pharyngeal cancer risk was multiplied 10 to 12 times for those with high alcohol intake. Interestingly, two studies reported by Marshall (2009) identified that alcohol consumption separate from eating a full meal was associated with an even higher risk of oral and pharyngeal cancer compared to consumption only at mealtime. However, Marshall highlighted that poor oral health itself was associated with increased oral cancer incidence which in turn is often accompanied by alcohol intake; apportioning the role of alcohol consumption to increased oral cancer risk is confounded by other behaviours such as poor diet and tobacco use as well as oral health.

Relevant to Scotland, given whisky's role in Scotland's identity and culture, research exploring type of alcohol and risk of oral cancer reported that in some cases, but not others, whisky has been associated with the greatest increase in relative risk of oral cancer. Other studies, however, identified that wine may even reduce risk of oral cancer while further studies reported that the greatest risk of oral cancer is amongst those consuming 30% or more their alcohol in the form of spirits (Marshall et al 2009). The weight of evidence suggested that alcohol itself and not the type of alcohol beverage was the most important determinant of laryngeal cancer. A further factor identified was the fact that with respect to alcohol exposure, there were two laryngeal regions: the hypopharynx (located at the laryngeal-pharyngeal junction) which is directly exposed and the endolarynx, the main body of the larynx which is not exposed to alcoholic beverages, only air (Marshall et al 2009).

Conway (2010) commented on an INHANCE study exploring the excess odds ratio by drink-years (exposure duration) modified by drinks/day (exposure rate). For alcohol assumption up to 10 drinks daily, excess odds ratio (EOR) results suggested that more drinks/day for a shorter period was more harmful than fewer drinks/day for a longer period. EOR drink-year estimates varied by site, with pharyngeal/oral cancer risk the greatest. This suggested that the number of drink-years was the cause, not the number of drinks consumed per day. As with cigarettes, the authors found there was no safe minimal limit where consumption was protective for these cancers.

It appears to require decades of elevated exposure for alcohol intake to affect cancer risk (Marshall et al 2009). Various mechanisms have been proposed as the pathway that alcohol stimulates carcinogenesis. These include carcinogen metabolism, effects of acetaldehyde, interactions of alcohol with nutritional factors, effects of alcohol on hormone levels and physical effects of alcohol on tissues.

#### 1.7.3.3 Smoking and alcohol interaction

Taken together, the effects on head and neck cancer risk associated with both smoking and alcohol consumption were greater than the additive effects of the risks of the two behaviours alone. Among the older population and males, smoking and drinking behaviours together accounted for 64% of oral cavity cancers, 72% of pharyngeal cancers and 89% of laryngeal cancers. For women and younger adults, the proportions were less (Winn et al 2015).

Marshall et al (2009) also reported that increased laryngeal cancer risk was associated with increased alcohol consumption as well as association with increased tobacco smoking. This was evidenced by a trend of increasing relative risks of laryngeal cancer among women smokers reflecting greater smoking rates among women over time. Similar to oral cancer, the ability to attribute increased laryngeal cancer risk to alcohol versus tobacco was challenging. However, studies exploring the associated risk amongst non-smokers who consume alcohol and non-drinkers who smoke and the relative risk of laryngeal

cancer of those who stop either smoking or alcohol attempted to tease out the relative effects of both behaviours. The results were that the relative risk did not change after alcohol cessation. The interpretation was equivocal: either alcohol alone does not affect risk or alcohol induced risk cannot be reversed.

Given that the oral cavity and the hypopharynx are directly exposed to the undigested form of alcohol during consumption, perhaps the most relevant hypothesis for alcohol consumption pathway to these cancers is that the solvent properties of alcohol may enhance the effects of exposure to tobacco carcinogens (Marshall et al 2009); alcohol and tobacco are known to have synergistic effects in increasing the risk of these cancers (Winn et al 2015).

#### 1.7.3.4 Diet and nutrition

Winn et al (2015) also reported the INHANCE pooled analysis findings on diet risk association with head and neck cancers. They found that a high intake of fruits (OR 0.52 95% CI 0.43, 0.63) and vegetables (OR 0.66 95% CI 0.49, 0.90) substantially reduced risk of head and neck cancers compared to low intake. High B-carotene, lycopene and carotenoid intake in general, all found in high concentrations in fruits and vegetables were all associated with a protective role resulting in a reduced head and neck cancer risk. Likewise, ever use of calcium and vitamin C supplements were associated with lower head and neck cancer risk (Winn et al 2015).

Similarly, Schwingshackl et al (2015) conducted a systematic review and meta-analysis of observational studies comprising of an overall population of 1,784,404 subjects to investigate the effects of following a Mediterranean diet. They established that head and neck cancer risk was more than halved for those who did follow the diet compared to those who did not (RR 0.40, 95% CI 0.24, 0.66). Schwingshackl et al (2015) adopted the same definition of head and neck cancer as INHANCE, i.e. oral cavity, pharynx and larynx. The authors discussed how a Mediterranean diet might reduce the risk of a cancer diagnosis and cited a study with 7,447 subjects that showed that the highest category of nut consumption (more than 3 servings weekly) compared to the lowest was associated with 40%

reduction in cancer mortality. Interestingly, differences in consumption of extra virgin olive oil did not have any association. Both cohort and case-control studies were included in the meta-analysis; the case-control studies are subject to recall and measurement bias whereas the self-reported nutritional assessment validity and reliability for both case-control studies and cohort studies must be considered a weakness.

Little or no evidence has been collected to show any risk between carbohydrates and oral cavity or pharyngeal cancers (Sartorius et al 2016). Similarly there was limited evidence of risk associations beyond the preventative effect associated with fruit and vegetable consumption (WCRF/AICR 2007).

The World Cancer Research Fund review of food, nutrition and physical activity reported that it has been estimated that up to half of head and neck cancers could be prevented by appropriate diets and associated factors citing non-starchy vegetables, fruits and foods containing carotenoids as probably protective. Of the 238 studies included in their assessment, the head and neck cancer panel reported that a meta-analysis showed a 18% decrease risk per 100g fruits consumed daily or 24% reduction per 50g portion of citrus fruits consumed daily; the greatest effect occurred for the first increment of consumption suggesting that some fruit consumption was better than no fruit consumption and continued to show a dose-response relationship, although this could not be described as linear (WCRF/AICR 2007).

#### 1.7.3.5 Obesity

Winn et al (2015) reviewed via INHANCE obesity and head and neck cancer risk, reporting that lean BMI was associated with increased head and neck cancer risk (RR 2.13 95% CI 1.75, 2.58) compared to those with high self-reported BMI (RR range 0.43 to 0.52) at time of diagnosis regardless of smoking or drinking status (Gaudet et al 2010). The investigators propose that high BMI mitigated the weight reducing effects of heavy tobacco and alcohol consumption. Low BMI and alcohol consumption were stronger risks for oral cavity and pharyngeal cancers (Winn et al 2015). This finding goes against the large body of evidence for other

major cancers which shows a strong relationship between obesity and cancer risk (Working Group 21 2016).

#### 1.7.3.6 Exercise

Nicolotti et al (2011), also via the INHANCE consortium, analysed data from four case-control studies including 2,289 head and neck cancer cases and 5,580 controls to explore recreational physical activity and head and neck cancer risk. Moderate or high levels of exercise was associated with 22% and 28% reduced risk of head and neck cancer respectively (OR 0.78 95% CI 0.66, 0.91; OR 0.72 95% CI 0.46, 1.16) (Nicolotti et al 2011; Winn et al 2015).

#### 1.7.3.7 Sexual behaviours

No association of sexual behaviours was found with oral cavity and hypopharynx cancers; however, elevated risks of oropharynx and tonsils were associated with more than three oral sex partners and six or more sex partners (OR 1.25, 95% CI 1.01, 1.54). These behaviours might increase the risk of HPV type 16 infection, a carcinogen recognised by IARC in 2012 (IARC 2012a; Winn et al 2015).

#### 1.7.3.8 Human papillomavirus

There is much debate about the HPV aetiological fraction of oropharyngeal cancer, although there is general agreement that the HPV is mainly involved with oropharyngeal cancer as opposed to other sites of the head and neck (Gillison 2007). World-wide the HPV attributable fraction has been estimated at between 18% to 28% for oropharyngeal cancer. However, recent estimates approaching 70% have been reported in the US in a recent systematic review and meta-analysis, which included 5,396 oropharyngeal cancer cases (Mehanna et al 2013). This analysis also observed increases from 40.5% before 2000 to 72.2% after 2005, with significant increases observed in North America and Europe. The estimates that have been employed in the health economics models are around 30% (Mehanna et al 2013).

## 1.7.4 Oesophageal cancer

### 1.7.4.1 Smoking

Kamangar et al (2009) reviewed the environmental risk factors for the two main histological types of oesophageal cancer. Tobacco was identified as a significant causal factor in oesophageal cancer carcinogenesis in 1979. Arnold et al (2017) identified that smoking and heavy alcohol consumption in high-income countries accounted for 75% of all newly diagnosed oesophageal squamous cell carcinomas (Arnold et al 2017). The increased risk of oesophageal squamous cell carcinoma was similar for cigarettes, cigar and pipe smoking (Kamangar et al 2009).

Oesophageal adenocarcinoma was also associated with smoking, with a two-fold increased risk compared to non-smokers and a dose-response relationship, but the association was much weaker compared to that of squamous cell carcinoma (Kamangar et al 2009).

### 1.7.4.2 Alcohol

The World Cancer Research Fund evaluated alcohol consumption and risk of oesophageal cancer (WCRF/AICR 2007). Reviewing 74 studies, predominantly of a case-control design, they concluded that most studies demonstrated a relationship of increased alcohol consumption and increased oesophageal cancer incidence. A meta-analysis of the case-control studies showed a 4% increase in risk per drink consumed in a week and a clear dose-response relationship. They hypothesized that DNA damaged that occurred because of smoking may be less effectively repaired in the presence of alcohol, particularly acetaldehyde. They suggested that the alcohol may act as a solvent enabling the carcinogenic molecules to reach the mucosa more effectively. Other proposed effects of alcohol in the carcinogenic pathway included production of prostaglandins, lipid peroxidation and production of free radical oxygen species. Finally, the panel proposed that heavy alcohol consumption may be associated with poor diets with limited essential nutrients increasing tissue susceptibility to carcinogenic attack (WCRF/AICR 2007).

The oesophagus is also directly exposed to the undigested form of alcohol. Marshall et al (2009) reported that many studies showed relative risk increases of oesophageal cancer of eight to 10 times the normal population for heavy consumption of alcohol.

#### 1.7.4.3 Smoking and alcohol interaction

Smoking and heavy alcohol consumption in high-income countries accounted for 75% of all newly diagnosed oesophageal squamous cell carcinomas (Arnold et al 2017).

Marshall et al (2009) also reported several studies indicating that again there was a synergistic effect between alcohol and tobacco exposure and increased oesophageal cancer risk. One study indicated that heavy drinkers who were also heavy smokers increased their risk of oesophageal cancer 51 times compared to those who totally abstained from either smoking or drinking alcohol. Recent studies demonstrated that this association of alcohol and increased oesophageal cancer risk held for squamous cell carcinoma, the previously dominant form of oesophageal cancer in the UK (Section 1.6.5), but it did not hold for oesophageal adenocarcinoma (Marshall et al 2009).

#### 1.7.4.4 Diet and nutrition

The World Cancer Research Fund report investigating nutrition and cancer prevention reviewed 262 publications on the association with oesophageal cancer. Forty-seven studies, most of which followed the case-control design, evaluated the relationship with fruit consumption. With few exceptions, the studies showed a decrease in oesophageal cancer risk with increased fruit consumption which demonstrated a dose-response relationship. Meta-analysis of the case-control studies indicated a 22% decrease in risk for 50g fruit consumed daily and a 30% decrease for 50g of citrus fruit consumption per day (WCRF/AICR 2007).

As described and critiqued in Section 1.7.3.4, Schwingshackl et al (2015) investigated the effects of adhering to a Mediterranean diet using a meta-analysis of observational studies with 1,784,404 persons. They established no significant association for oesophageal cancer.

#### 1.7.4.5 Obesity

Arnold et al (2015) reviewed obesity and cancer to assess the global impact by examining data from seven countries and reported that the trends in oesophageal cancer have changed such that oesophageal adenocarcinoma is overtaking oesophageal squamous cell carcinoma. It was proposed that this trend reflected that obesity increased the risk of oesophageal adenocarcinoma more than two-fold (Arnold et al 2015) while Behrens et al (2014) reported obesity was associated with a 30% to 50% risk reduction of squamous cell oesophageal cancer (Behrens et al 2014). Possible explanations provided were earlier exposure to excess weight and exposure accumulated over the life-course may lead to insulin resistance, chronic inflammation, oxidative DNA stress and changes to endogenous hormone metabolism which in turn may lead to carcinogenesis (Behrens et al 2014; Arnold et al 2015). Obesity was also associated with increased gastro-oesophageal reflux disease, a known risk determinant for oesophageal adenocarcinoma; however, Behrens et al (2014) reported that a five unit increase in BMI is associated with a 52% increase in risk of oesophageal adenocarcinoma, independent of gastro-oesophageal reflux disease. Increasing waist circumference was also reported as associated with increased risk of oesophageal adenocarcinoma (Behrens et al 2014).

Inverse relationships between oesophageal squamous cell carcinoma where decreased BMI and waist circumference were associated with increased oesophageal squamous cell carcinoma independent of smoking status were also noted (Behrens et al 2014). So, unlike head and neck sites of the upper aero-digestive tract, oesophageal cancer was observed to have a risk association with being overweight/obese (Behrens et al 2014).

#### 1.7.4.6 Exercise

For non-smokers, diet and physical exercise are the most important modifiable behavioural risk factors that are associated with oesophageal cancer risk (Singh et al 2014).

Singh et al (2014) conducted a meta-analysis to explore physical exercise and oesophageal cancer risk. Nine (four cohort, five case-control) studies comprised of 1,871 oesophageal cancer cases among 1,281,844 patients were included in the review. Comparing the most to the least active groups, risk of oesophageal cancer was reduced by 29% (OR 0.71 95% CI 0.57, 0.89). Risk of oesophageal adenocarcinoma was reduced slightly more at 32% (OR 0.68 95% CI 0.55, 0.85) as this cancer was associated with obesity and associated chronic inflammation. Only three studies reported on the association of physical exercise with risk of oesophageal squamous cell carcinoma and the results were equivocal. As a result, no association between physical exercise and squamous cell oesophageal cancer risk (OR 1.10, 95% CI 0.21, 5.64) could be identified. The authors explored possible explanations for their findings including "healthy user bias" in more physically active people and the tendency for other unhealthy behaviours to be adopted by physically inactive people. They recognized that there may be residual confounding by SES despite adjustment by most studies. They also noted that none of the studies adjusted for presence of gastro-oesophageal reflux or erosive oesophagitis; moderate but not intense exercise had previously been found to be associated with reduced reflux disease in obese patients but not those who were not obese.

Behrens et al (2014) in their systematic review investigated physical exercise and gastro-oesophageal cancer risk by anatomical site and histology via a meta-analysis based on 24 studies comprised of 15,745 cases. A 21% risk reduction was evident for oesophageal adenocarcinoma (RR 0.79 95% CI 0.66, 0.94). Oesophageal squamous cell carcinoma risk was also reduced by 34% after exclusion of one study and including terms for study design and sex (RR 0.66 95% CI 0.46, 0.96). The investigators explored whether adiposity mediated the

inverse relationship between physical exercise and gastro-oesophageal cancer by comparing RRs for studies adjusted for adiposity and those that were not; they established some attenuation, but not full attenuation suggesting physical exercise has a protective effect of its own. Interestingly, distant past physical exercise and consistent physical exercise over time were more protective than recent past physical exercise.

### 1.7.5 Brief summary of the behavioural risk factor association with lung and UADT cancer

Upper aero-digestive tract cancers collectively share similar behavioural risk factors associations. Notably, the major risk factors of smoking (tobacco) and alcohol consumption are dominant across all sites. The evidence in relation to diet is more limited but consistent in the protective benefits of fresh fruit and vegetables. Physical activity is an emerging area of research and again there is a tendency for a protective effect associated with increased physical activity. Obesity, while not being a direct behaviour has a different effect across subsites with a risk association with adenocarcinoma of the oesophagus, while oesophageal squamous cell carcinoma, along with other sites of the head and neck show an inverse relationship. These risk behaviours are all determined and strongly associated with socioeconomic status, both individually (Kogevinas et al 1997a) and collectively (Lawder et al 2010) and therefore it is important to attempt to include these factors in any investigation of socioeconomic status where possible.

# 2 Literature review – Part II: cancer incidence disease burden by socioeconomic status

## 2.1 Approach to literature search

Several searches in PubMed were undertaken to identify relevant articles that had been published between January 2007 and June 2017 and had within in the article title the desired focus on anatomical site, incidence, socioeconomic status and cancer. The objective was not to conduct a full systematic literature review of all articles on the subject, but to establish recent publications, mainly systematic reviews, meta-analyses or significant cohort studies, focused on cancer incidence disease burden by SES which had already been undertaken. Given that this was the primary focus of the PhD, it was considered important to endeavour to capture as complete a picture as possible of the up-to-date international literature in this area.

Note that this literature review was commenced prior to the PhD research studies being undertaken (in 2009). In compiling the final version of the thesis, the literature searches described in Sections 2.1.1 – 2.1.4 were performed to provide a more complete and contemporary picture, rather than presenting a somewhat out-of-date literature review. How the PhD research studies fit in, compare, contrast and add to the body of this literature is provided in Section 6.2 of the Discussion (Chapter 6).

### 2.1.1 All cancer

The first search using the string ‘socioeconomic [title] AND cancer [title] AND risk [title]’ generated 45 papers, 15 of which were selected for further review. Articles that did not focus on the incidence, socioeconomic status or the desired cancer sites were excluded. The exclusions consisted of articles which may have included these attributes but were focused on the patient’s perception of risk, or looked at other aspects of cancer incidence, diagnosis and treatment, such as screening, ethnic or racial inequalities or other co-morbidities (e.g. diabetes), rather than, SES and risk of cancer.

### 2.1.2 Lung cancer

The second search focused on lung cancer using the following string:

> (((lung*[Title]) AND (cancer[Title] OR malignancy[Title] OR neoplasm[Title])) AND (incidence[Title] OR risk[Title] OR determinant[Title]))) AND (socio*[Title] OR social[Title] OR economic[Title] OR education*[Title] OR income[Title] OR poverty[title] OR poor[title] OR depriv*[title] OR inequal*[title] OR dispara*[title])))))

Of the 37 articles returned, nine were identified for further review. One of the papers that was identified was a systematic review and a meta-analysis conducted in 2009 (Sidorchuk et al 2009). This paper and eight others became the focus of review. Those papers that were excluded were focused on cancer treatment, cancer mortality, the patient's perception of cancer risk, or cancer prevention strategies.

While title searches identified potential articles, there was a need to search bibliographies; therefore, Web of Science was used to perform a citation search for Sidorchuk's systematic review of socioeconomic differences in lung cancer referred to above (Sidorchuk et al 2009). As of 29th June 2017, 41 articles were retrieved. After review and confirming those articles already identified through previous searches, six further articles were retrieved. In addition to the articles already mentioned, also excluded were those articles focusing on: i) Treatment, survival, or mortality; ii) Illnesses other than lung cancer or lung cancer in addition to other co-morbidities; iii) Non-socioeconomic associated causes of lung cancer; iv) Ethnic or racial disparities rather than socioeconomic inequalities; v) Describing lung cancer in prison patients; vi) Genetic causes of lung cancer; and vii) One paper that was only available in Portuguese.

### 2.1.3 UADT cancers

The third search used the following string and focused on UADT cancers specifically.

> '((((((oesophag*[Title] OR esophag*[Title] OR UADT[Title] OR "Upper Aerodigestive Tract"[Title] OR "Upper Aero-digestive Tract"[Title])) AND (cancer[Title] OR malignancy[Title] OR neoplasm[Title])) AND (incidence[Title] OR risk[Title] OR determinant[Title]))) AND (socio*[Title] OR social[Title] OR economic[Title] OR education*[Title] OR income[Title] OR poverty[title] OR poor[title] OR depriv*[title] OR inequal*[title] OR disparat*[title]))

Seventeen articles were returned, nine of which were selected for further review. The articles that were excluded focused on clinical prognostic markers for primary incidence or metastasis, or clinical treatment of these cancers.

### 2.1.4 Additional searches performed and sources reviewed

To ensure a comprehensive coverage of relevant articles, a further search of EMBASE using the following string was performed and returned 96 entries.

> (incidence or risk or determinant).tw. and (socio* or social or economic or education or income or poverty or poor or depriv* or inequal* or disparat*).m_titl. and ((oesoph* or esophag* or UADT or "Upper aerodigestive tract" or "upper aero-digestive tract") adj3 (cancer* or malignan* or neoplasm* or tumo?r*)).tw.

After abstract review and comparison to the other searches already performed to exclude duplicates, nine papers were retained for review. Excluded articles were: i) Those published before 2008; ii) Papers that did not focus on the cancers in question; iii) Those papers that were primarily focused on racial disparity or genetics, treatment, mortality or survival of cancer; or iv) Papers that used gross national income, a very high level area measure of SES.

The special supplement edition of the *European Journal of Cancer* that was published in September 2008 featured social inequalities in cancer incidence, survival and mortality in Denmark by anatomical site. The six relevant articles

focusing on lung and UADT and all cancer were considered and all adopted the same methodology so comparison across the cancer sites was facilitated.

Schottenfeld et al (2009) online book titled *Cancer Epidemiology and Prevention* included six relevant chapters by anatomical site each comprehensively reviewing the epidemiology of lung and UADT cancers. These chapters were also reviewed.

The results from the assessed studies are presented for all cancer, lung, head and neck cancers, oesophageal cancers and finally UADT cancers where head and neck and oesophageal cancers were not evaluated separately.

## 2.2 All cancer

### 2.2.1 Introduction

Five studies which included all cancer risk association with low socioeconomic circumstances are described below. Of these studies, three were cohorts (Dalton et al 2008a; Mouw et al 2008; Spadea et al 2008), one study was based on the pooled data of several cancer registries (Boscoe et al 2014) and the last was a case-control study (Leuven et al 2016). Only one of the studies included behavioural data (Mouw et al 2008) and two studies were the only studies to consider both individual and area indicators of SES (Dalton et al 2008a; Leuven et al 2016).

### 2.2.2 Publications

***Dalton et al (2008a)***. At the time that the thesis studies were commenced, the studies conducted in Denmark were instrumental in the development of a special supplement of the *European Journal of Cancer*. This 2008 supplement investigated social inequalities of cancer incidence, mortality and survival by anatomical site in the population of Denmark using six individual socioeconomic

indicators (education level, disposable income, work market affiliation[1] (employment status), social class, housing tenure and size of dwelling). The series of papers utilised the strengths of Danish routine administrative and health databases, cancer registry and linkage potential. Incidence rate ratios (IRR) were adjusted for age, education level and disposable income. Large inequalities in incidence were identified for lung (IRR male 1.53, female 1.85), oesophagus (IRR male 1.30, female 0. 87) stomach (IRR male 1.37, female 1.23), mouth and pharynx, larynx (IRR male 1.67, female 3.23) and cervix (IRR 1.33) where higher incidence rates were identified for those in lower social groups. The strengths of the study included the large cohort study design of the Danish population covering 3.22 million people and the consequent statistical power. Five individual SES measures (education level, disposable income, work market affiliation (employment status), social class and housing tenure) were analyzed. However, no data were available for adjustment for known risk behaviours such as smoking and alcohol consumption. The area indicator focused on rural versus metropolitan location as opposed to capturing aspects of area deprivation, thereby omitting the potential role of neighbourhood characteristics in SES. The individual SES measures adopted usefully captured various aspects of SES over the life-course. In addition, the investigators considered relevant demographic variables such as type of district, cohabiting status, ethnicity, Charlson co-morbidity index, depression and psychosis. However, no adjustment could be made for known risk behaviours and area SES, other than rural versus metropolitan location, was not considered. The SES variables were measured two years before diagnosis and therefore did not fully reflect the temporal relationship between exposure and cancer incidence; however, the SES variables were updated annually to reflect most up to date information. It is also worth noting that the study adjusted for education and income level; therefore the remaining SES variables (work market affiliation (employment status), social class and housing tenure) had an effect on cancer incidence independent of education and income (Dalton et al 2008a).

1 Work market affiliation (categorised as working, unemployed or early retirement pension) were defined as unemployed in November that year. Early retirement pension (formerly known as disability pension) was granted if a person was unable to work permanently due to mental or physical disability and this disability reduced the ability to work by at least 50%. Pensioners due to age (in Denmark in the study period, age 67) were categorised on the basis of their affiliation to the work market before their age-related retirement.

This study identified that inequalities were greater for housing ownership variables than for income, for a number of cancer sites, suggesting that accumulated wealth was an important predictor for cancer incidence. This finding may potentially reflect the long lag-time for cancer development (Kawachi et al 2006). Although it was associated with higher cancer risk, lower occupational social class was found to be less important than education and income. The role of education potentially reflected childhood socioeconomic circumstances as the foundation for future knowledge: i) To make healthy decisions; ii) To provide opportunities for employment; iii) To improve the future level of remuneration and influence the ability to select healthy neighbourhood and home environments; and may iv) Explain the relatively less important role of occupational social class (Dalton et al 2008a).

***Mouw et al (2008)*** explored education and risk of all and site-specific cancer in a relatively large prospective cohort study of nearly 500,000 Americans taking into consideration behavioural factors (Mouw et al 2008). The cohort used was the National Institute of Health–American Association of Retired Person Diet and Health Study. Models to calculate relative risk association with educational attainment were developed for men and women separately. A whole raft of covariates were used including: age, years of education, smoking status, time since quitting, smoking dose, race/ ethnicity, energy intake, alcohol consumption, BMI, physical activity frequency, marital status and family history of cancer. Further variables used, which were sex and cancer site-specific, were: hormonal use, age at first birth and number of births. Although comprehensive in coverage, this information was obtained via a questionnaire and therefore was subject to recall bias. For men, compared to postgraduate educational attainment, all cancer risk was not associated with the lack of education after adjustment for smoking (RR 1.05 95% CI 1.00, 1.09) or other behavioural factors (RR 1.03 95% CI 0.99, 1.07) but was protective for women (RR 0.86 95% CI 0.80, 0.92 and RR 0.84 95% CI 0.79, 0.90). Education level was generally established early in life, potentially reflecting early SES, and therefore particularly relevant to cancer incidence given its long gestation period. As a result of attainment at an earlier point in life, education may avoid the criticism

of capturing reverse causation, particularly for the cancers of focus in this thesis (Section 1.3.2.1). Nevertheless, SES at any point in the life-course could be influenced by area SES factors as well as multiple individual factors, with education being just one. As a result a multidimensional approach to measuring SES may provide further insights.

***Spadea et al (2010)*** conducted their study using a cohort of 1,407,164 residents in Turin (Italy) in the period 1985 to 1999 to explore the relative importance and independent effect of three individual socioeconomic indicators (education, occupational social class and housing characteristics) and one area-based deprivation indicator (proportion of manual workers, those with low education tenants, those in accommodation without a bath, families with children and one parent and a crowding index) that was measured during adulthood to establish the magnitude of cancer incidence inequalities. They found that for all cancer for men, all four indicators contributed to inequalities in a fully adjusted model with housing characteristics (RII 1.26 95% CI 1.18, 1.34) the strongest association, followed by education (RII 1.17 95% CI 1.09, 1.27). Occupational social class association (RII 1.10 95% CI 1.02, 1.18) and area deprivation (RII 1.09 95% CI 1.03, 1.16) were weakly associated. For females, lower educational attainment (RII 0.78 95% CI 0.72, 0.85) was protective and poor housing characteristics were weekly associated with a greater risk (RII 1.12 95% CI 1.04, 1.19). The confidence intervals for RII, of both occupational social class and area deprivation SES indicators included zero and therefore were not significant (Spadea et al 2010). For older men, while education and occupational social class indicated no association with all cancer risk, housing characteristics and area deprivation indicated a nine percent and 12% greater risk (RII 1.09 95% CI 1.02, 1.18 and RR 1.12 95% CI 1.04, 1.20 respectively). No behaviour risk factors were considered and the cancers were evaluated only at anatomical site level (lung) or grouped together (UADT) thereby masking potential differences at morphology (lung) or individual anatomical site level (UADT). However, the study conducted by Spadea et al (2010) was one of the few where more complex measures of socioeconomic inequalities were used (Spadea et al 2010).

In a cancer registry study, ***Boscoe et al (2014)*** investigated the area SES association with cancer incidence for cancer by anatomical site using data for 16 USA states sourced from the North American Association of Central Cancer Registries. Their analysis included 2.9 million malignant tumours and population data from the National Cancer Institute's Surveillance, Epidemiology and End Results programme (Boscoe et al 2014). The area SES indicator used was the census track poverty level described as the proportion of population living below poverty level based on income. Risk ratios of cancer incidence between the highest and lowest poverty category were calculated, adjusted for ethnicity. For all sites and both sexes combined, the difference in risk between the greatest and lowest poverty category was less than two percent. Explanations for the findings focused on risk factors such as tobacco, alcohol, intravenous drug use, sexual transmission and poor diet which the authors suggest are associated with higher poverty. However, no individual indicators of SES were available or analysed, a simple measure of socioeconomic inequality was adopted, no risk behaviour data was available and measurement of SES occurred at diagnosis (Boscoe et al 2014).

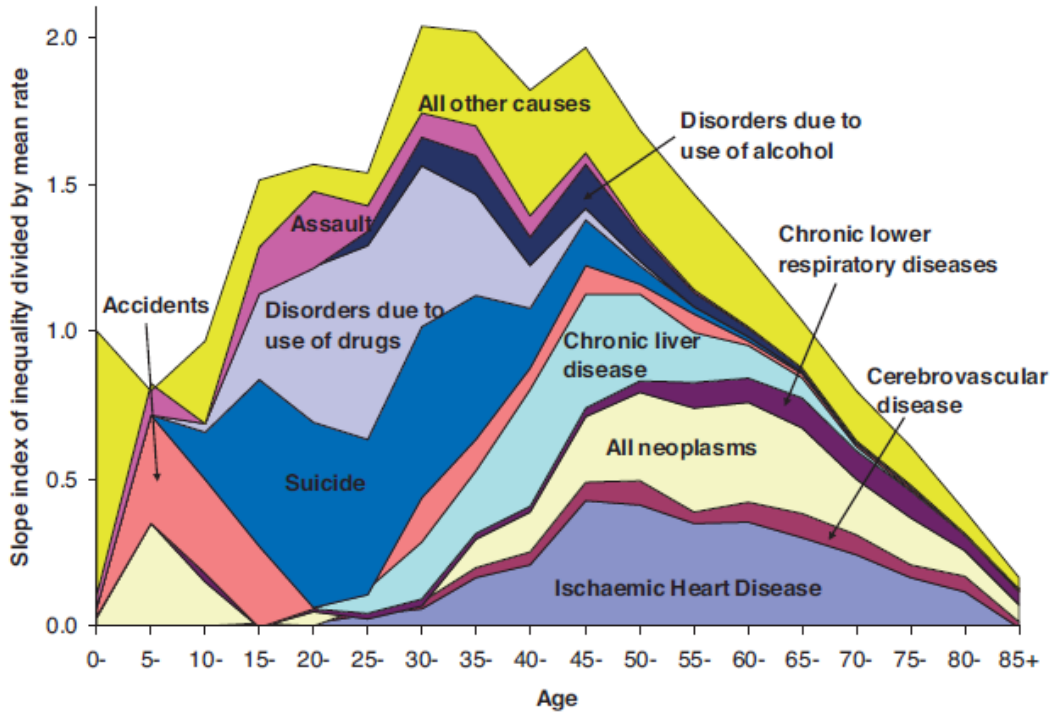
In a natural experiment analysis, ***Leuven et al (2016)*** took advantage of a two year increase in, and standardisation (reform) of, Norway's compulsory schooling. This reform occurred from 1960 to 1972. Through record linkage of the population, cancer and education registries they conducted a population-based cohort study comparing those with and without the additional education of two years to explore the association with all cancer and common cancers risk (including lung cancer) (Leuven et al 2016). Hazard ratios (HRs) for males and females indicated a very small and statistically significant decreased risk of all cancer associated with a two year increase of school education. The authors also explored the association of education reform and all cancer risk and concluded that the estimates were not statistically significant for either males or females. They also concluded that, with the exception of lung cancer (Section 2.2.2), education had no effect on either all cancer or the most common sites in isolation. The study is an interesting addition to the socioeconomic inequalities in cancer risk literature; however, individuals were followed only until the age

of 67, a relatively young age where cancer may not yet be diagnosed. In addition, the study assessed the addition of two years of compulsory education, but did not capture the effect of further or higher post school education on cancer risk. This suggests that it is with higher post school education that the widening of inequalities in health outcomes in general and cancer risk in particular are aggravated.

### 2.2.3 Summary of all cancer literature findings

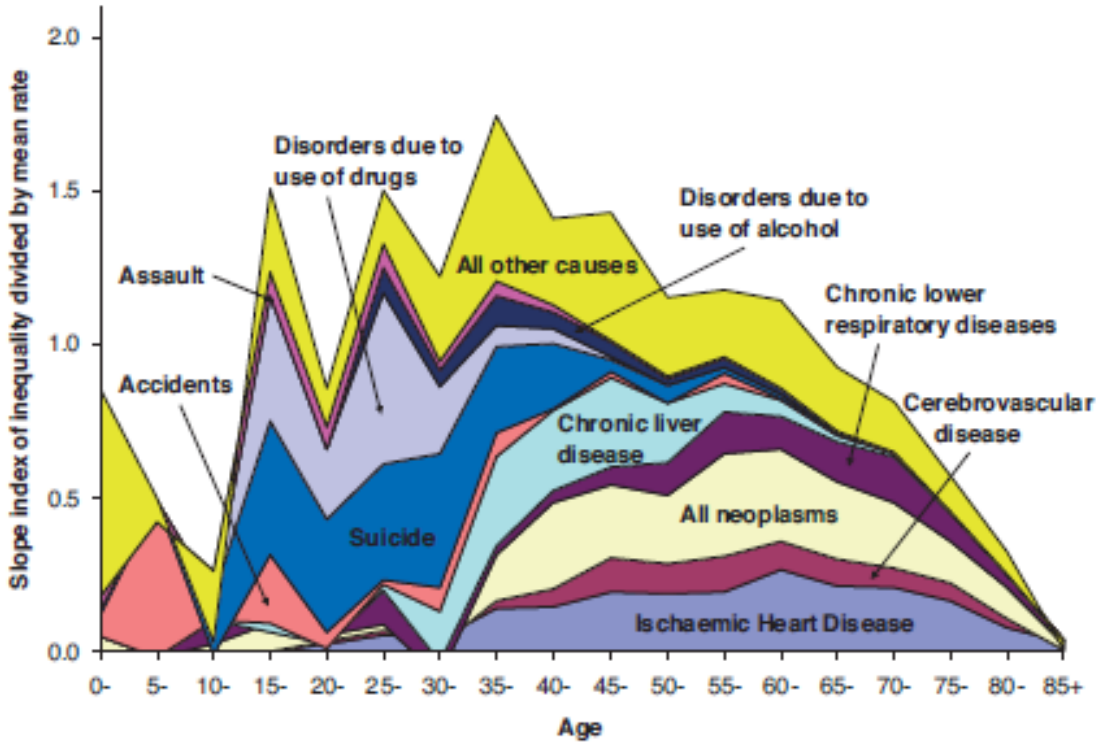
In Scotland, little or no change in the inequality gap for cancer incidence (Scottish Government 2017c) has occurred despite investment in public services including the NHS (Scottish Government 2016) and a number of policies focused on addressing inequalities (Scottish Government 2003; Scottish Government 2007; Scottish Government 2008a; Scottish Government 2008b; Scottish Government 2008c; Scottish Government 2008e; Scottish Government 2010a). Using the Slope Index of Inequality and SIMD, Leyland et al (2007a; 2007b) explored age and sex cause specific socioeconomic inequality in mortality in Scotland (Leyland et al 2007b; Leyland et al 2007a). They found that while a general reduction in mortality in ischemic heart disease and malignant neoplasms had occurred over the period 1981-2001, the reductions were socially patterned. These publications provide a helpful way of presenting the picture of socioeconomic inequalities (Figures 2.1, 2.2). Using similar approaches for cancer incidence may reveal information that is pertinent to understanding the pattern of inequalities in cancer incidence in Scotland, as well providing clues to the causes that contribute to these inequalities. These analyses tend to be focused on the Scottish Cancer Registry using traditional Carstairs SES area measures, with more recent analyses using SIMD or the I-E domain of SIMD (Section 1.3.2.4)

**Figure 2.1 Age specific contribution to inequalities of specific causes of death across SIMD income quintiles, men, Scotland 2000-02[1]**




[1](Leyland et al 2007b)

**Figure 2.2 Age specific contribution to inequalities of specific causes of death across SIMD income quintiles, women, Scotland 2000-02[1]**




[1](Leyland et al 2007b)

## 2.3 Lung cancer

### 2.3.1 Introduction

Seventeen studies which included lung cancer risk association with low socioeconomic circumstances are described below. Of these studies, nine were cohort studies (Dalton et al 2008a; Mouw et al 2008; Spadea et al 2010; Meijer et al 2013; Sondergaard et al 2013; Garcia-Gil et al 2014; Mitra et al 2015; Li et al 2015; and Vohra et al 2016), one study (Sidorchuk et al 2009) was a systematic review and meta-analysis, four studies were case-control design (Nkosis et al 2012; Hystad et al 2013; Behren et al 2016; and Leuven et al 2010) and the remaining three were either cancer registry (Boscoe et al 2014; Kuznetsov et al 2011) or descriptive (hospital cases) studies (Denton et al 2017). Nine of the studies included behavioural data (Mouw et al 2008; Sidorchuk et al 2009; Nkosi et al 2012; Meijer et al 2013; Hystad et al 2013; Garcia-Gil et al 2014; Behren et al 2016; Vohra et al 2016; and Denton et al 2017), while six studies considered both individual and area indicators of SES (Dalton et al 2008a; Spadea et al 2010; Nkosi et al 2012; Hystad et al 2013; and Li et al 2015). Only one study (Spadea et al 2010) used a more complex method (Relative Index of Inequality) to quantify the SES inequalities while all other studies used more simple measures comparing the two extremes or each SES category relative to a selected reference. Finally recognising the temporal relationship between SES exposure and diagnosis, eight of the studies (Spadea et al 2010; Hystad et al 2013; Sondergaad et al 2013; Garcia-Gil et al 2014; Behren et al 2016; Leuven et al 2016; Vohra et al 2016) measured SES well before diagnosis, while seven studies measured SES at the time of diagnosis (Sidorchuk et al 2009; Kuznetkov et al 2011; Nkosi et al 2012; Boscoe et al 2014; Mitra et al 2015; Li et al 2015; Denton et al 2017) and two studies measured SES two years before diagnosis (Dalton et al 2008b; Meijer et al 2013).

### 2.3.2 Publications

***Dalton et al (2008a)*** evaluated SES association with cancer risk by anatomical site in a large cohort in Denmark (Dalton et al 2008b) (Section 2.2.2). Their

study on lung cancer which adjusted IRRs for age, period, education and disposable income, indicated for both males and females a decrease in lung cancer incidence was associated with greater social advantage for longer period of education, higher income, affiliation to work market (employment status), housing tenure (ownership) and larger dwelling size. For both males and females, adjusted IRR for the Bohemian social class[2] compared to the manual class was the largest of all the adjusted IRRs (IRR 7.2 95% CI 5.2, 9.4 and 18.4 95% CI 10.1, 33.4), respectively) observed. The explanations for these results, focused on smoking behaviour but a wider discussion of pathways for SES to lead to cancer diagnosis was outlined in a lessons to be learned article that covered all sites (Dalton et al 2008a) (Section 2.2.2) (Dalton et al 2008b).

As described earlier, ***Mouw et al (2008)*** performed a large cohort study in the US which investigated the association between education and cancer incidence (Section 2.2.2). Lung cancer incidence in men was strongly associated with education level even after adjustment for smoking and other behaviour factors. Compared to those with a postgraduate education, the RRs for men with less than a high school education were 3.67 after adjustment for age, reducing to 2.02 after additional adjustment for smoking and attenuating further to 1.95 after additional adjustment for other behaviour factors. These models repeated for women provided RRs of 2.14, 1.43 and 1.43 respectively (in all cases the confidence intervals above 1.0). Notably, adjustment for smoking did not attenuate completely all the effect of low education level suggesting possible residual confounding by smoking, or other factors at play. These may include other SES dimensions and more complex pathways between SES and cancer incidence for which there were no data available (Mouw et al 2008).

***Sidorchuk et al (2009)*** published a systematic review and meta-analysis exploring socioeconomic differences in lung cancer incidence. They reviewed 64

[2] Social Class definition employed here was based on theory of creative class: the creative class (e.g. researchers, designers, and architects), creative professionals (e.g. managers, business and finance, lawyers, doctors), bohemians (e.g. artists, models), the service class (e.g. nurses, hairdressers, and caterers), the manual class (e.g. construction workers, transport and production workers), and the agricultural class (e.g. farmers, fishermen).

studies that were comprised of 44 case-control studies, six case-control studies nested in a cohort and 14 cohort studies. All reviewed studies used three or more SES variables (educational attainment, occupational categories and income). Their main analysis was restricted to studies that had made an adjustment for smoking. They also performed separate analyses for studies using individual SES indicators and studies using area SES indicators. When adjusted for smoking, they found that greater lung cancer risk was associated with lower educational attainment and occupational categories (RR 1.33 95% CI 1.14, 1.55), with educational attainment the strongest (RR 1.65 95% CI 1.19, 2.28). The income indicator was associated with elevated lung cancer risk, but was not significant (RR 1.25 95% CI 0.93, 1.70) (Sidorchuk et al 2009).

While the Sidorchuk et al (2009) review identified a strong relationship of greater incidence with lower educational attainment or occupation, the analysis did not mutually adjust the SES indicators for each other or differentiate between area and individual SES factors (only one study used an area SES indicator). However, mutually adjusting could be considered over-adjustment or may lead to co-linearity. Furthermore, while adjustment was made for smoking behaviour (the most important risk factor) other behaviours that are known to be associated with lung cancer risk (e.g. diet) were not considered. Consequently, there remain opportunities to further refine estimates of SES contribution to lung cancer incidence in future research.

The Turin (Italy) study conducted by ***Spadea et al 2010*** (previously described in Section 2.2.2) identified that for men, education (RII 1.72 95% CI 1.45, 2.04) and housing characteristics (RII 1.72 95% CI 1.51, 1.95) were very strongly associated with lung cancer incidence, however occupational social class (RII 1.10 95% CI 0.94, 1.27) was found not to be associated after adjusting for smoking behaviour. Area deprivation (RII 1.24 95% CI 1.09, 1.41) was associated with an increased lung cancer incidence but to a lesser extent. For females, no education was strongly protective (RII 0.54 95% CI 0.37, 0.77) for lung cancer incidence while poor housing characteristics were less strongly associated with increased lung cancer risk (RII 1.45 95% CI 1.06, 1.98). For women, the

association of lung cancer risk with both area deprivation and occupational social class were not significant. Increased risks were explained by the social patterning of smoking which was more prevalent among men in the lower social classes. Education's larger gradient suggested that cultural and material aspects of SES during adolescence may have led to starting and continuing with smoking at an earlier age while housing characteristics' greater gradient represented material aspects of SES and the chances of stopping smoking. The lack of association for women was interpreted as reflecting differences between the sexes in social stratification of smoking which was more prevalent among the less disadvantaged, but had been reported to be reversing in recent years. Spadea and colleagues (2010) proposed that material indicators of SES may identify inequalities more quickly because smoking may be a way of facing economic stress. No behaviour risk factors were considered, although the more complex measure of socioeconomic inequality, the Relative Index of Inequality, was used and therefore reflected the full social gradient (Spadea et al 2010).

***Kuznetsov et al (2011)*** used the population-based cancer registry data for Bavaria, Germany, an area measure of SES and a multilevel study design to evaluate socioeconomic inequalities in lung cancer incidence and mortality. The index of multiple deprivation used was similar to the UK and Scottish Indices of Multiple Deprivation and included: income, education, employment, environment, social capital and security. Unlike the SIMD, for the Bavarian index, the community population size employed varied widely. No individual social variables or behaviour data were available for this analysis. Age adjusted RRs demonstrated a monotonic relationship of socioeconomic deprivation with increased lung cancer incidence for those from higher area deprivation for men (RR 1.41 95% CI 1.28, 1.54) but not women (RR 1.09 95% CI 0.96, 1.24). The authors discussed socially patterned behaviours and situations (smoking, diet, physical activity and environmental and occupational carcinogens) to explain the results and to explain that there was no association with SES and lung cancer for women (Kuznetsov et al 2011).

***Nkosi et al (2012)*** used data from a population-based case-control study focused on lung cancer and performed in Montreal, Canada, in 1996 – 2002 to explore the effect of different aspects of smoking behaviour on SES association with the risk of lung cancer. There were 1,203 subjects with incident lung cancer and 1,513 population controls. One area and four individual SES variables were used. The area-based measure was median-census tract household income assessed at time of interview. The individual SES measures were: self-reported household income (at interview), residential property value (from a publically available tax assessment 1995 database), education level (at interview) and occupational class (using employment history obtained at interview). They considered possible confounders such as: country of birth, diet (weekly fruit and vegetable consumption) and smoking (status, lifetime number of cigarette years and time since smoking cessation). Comparisons were made between the highest and lowest SES categories and the results were adjusted for age, country of birth and diet. For each SES variable, they found that incremental addition of aspects of smoking behaviour progressively reduced the contribution of most of the SES variables (e.g. area indicator: OR 0.97 95% CI 0.51, 1.02) but the lung cancer incidence inequalities that were associated with self-reported income (OR 0.72 95% CI 0.38, 1.39) and education (OR 0.57 95% CI 0.57, 1.02) were strongest, although they were fully attenuated by smoking. When all three smoking variables were included as well as the SES variables, all of the SES variables were fully attenuated (property value: OR 0.81 95% CI 0.55, 1.20; occupational social class: OR 1.00 95% CI 0.68, 1.47).

The authors concluded that if adequately modelled, smoking behaviour fully explained socioeconomic inequalities in lung cancer risk. They also discussed the pathways that smoking and SES may be implicated in lung cancer genesis and diagnosis. Two hypotheses were discussed: smoking as a mediator of SES acting as an intermediate risk factor or smoking as a confounder of SES and its association with lung cancer risk, related, but on a different pathway.

Nkosi and colleagues' (2012) study is a comprehensive assessment of smoking and its role in explaining SES, however, as a case-control study it is subject to

selection and participation bias as well as recall bias (dos Santos Silva 1999). Furthermore, the variables adopted to capture smoking behaviour include cigarette-years. In the same publishing year, Peto advised that the variable cigarette-years (referred to as pack-years) was not appropriate for epidemiological research. He explained that: "This (using pack-years) is a serious error, as the excess incidence for 20 pack-years is much greater after 40 years of smoking 0.5 packs per day than for 10 years at 2 packs per day. The effect of smoking is trivial for the first decade but substantial after 40 years" (Peto 2012). The message stated by Nkosi et al (2012) is that thorough modelling of smoking behaviour is relevant to understanding the extent that SES effect is independent of smoking on lung cancer incidence, but the modelling approach (use of pack-years) may be refined in future studies. Furthermore, full attenuation of SES effect by smoking does not imply that there is no SES effect at all (Nkosi et al 2012).

In Denmark, ***Meijer et al (2013)*** conducted a population cohort study that evaluated the role of neighbourhood SES via an area indicator defined as the proportion unemployed and population density along with individual SES indicators (education level, disposable income, occupational social class) on cancer incidence including lung. Multilevel analysis was used and taking into account cancer's long latency period, both area indicators were measured in 1995 while the timing of individual SES measures was assessed two years prior to diagnosis, acknowledging that diagnosis or the run up to definitive diagnosis may potentially result in SES change (downward most likely). Other relevant factors that were considered were sex, marital status and a refined classification of those not working (students, pensioners, disability pensioners and the unemployed). Fully adjusted hazard ratios indicated that the incidence of lung cancer was greater for the low socioeconomic strata for each individual SES indicator and the risk increased with increased area unemployment and increased population density. The authors concluded that both neighbourhood and individual SES indicators were associated with lung cancer risk. They hypothesised that greater density of convenience stores, greater air pollution and social influences may explain the higher lung cancer incidence association

with more population dense areas and lower SES neighbourhoods. However, no information on behaviours was available which may explain some of the association identified. While the authors reported from other studies that smoking attenuated or explained SES inequalities in lung cancer, they did somewhat ignore or fail to investigate the interrelationship or interaction between SES and smoking.

***Hystad et al (2013)*** used a Canadian population-based case-control study consisting of 1,224 lung cancer cases and 1,802 controls to explore long-term area SES exposure and lung cancer risk and how that changes over time. They also explored the extent that smoking, environmental and occupational factors mediate the relationship between area SES and lung cancer risk. They discussed the challenges of study participants moving residence and length of residence and the potential importance of these variables given cancer's long latency period. The study design enabled multiple individual level and behavioural variables to be included: age, sex, educational attainment (years), household income, life-time cigarette smoking measured in pack-years, years since quitting, person-years of residential second hand smoke exposure, average weekly alcohol consumption, weekly meat consumption, weekly vegetable consumption, average monthly physical activity, person-years of occupational second hand smoke exposure, years working with daily/weekly carcinogen exposure, industrial odours or dusts, exposure to nitrogen dioxide, years living within 100 meters of a major road and average ecological-level radon estimates. Hystad et al (2013) defined area SES as mean household income, percentage of adults without a  diploma, percentage of adult unemployment, percentage of rental dwellings and percentage of residents that moved in the last five years. These area SES variables were collated from five censuses to establish a single composite area SES for study entry (1994) and two latency periods (1975 and 1975-1985) separately. They also measured long-term neighbourhood socioeconomic status based on the five residential histories occurring within the 20 year period from 1975 to 1994. Incremental logistic regression models were developed to compute odds ratios and assess the degree that area SES and lung cancer incidence association changed. The effects were restricted to the most

socioeconomically deprived area SES group compared to the least socioeconomically deprived. The unadjusted OR for lung cancer risk was 1.66 (95% CI 1.31, 2.09). After adjustment for individual SES variables, the OR was attenuated to 1.46 (95% CI 1.13, 1.89); after full adjustment, the OR was further attenuated, but remained significant (OR 1.38 95% CI 1.01, 1.88). These long-term area SES ORs were greater than the ORs for the area SES at study entry (1994, point-in- time) and the ORs for area SES captured earlier (1975, 1975-1985). Focusing on the long-term area SES index only and successively adjusting the model for smoking variables, other individual health behaviours, occupational exposures and environmental exposures; the authors found that the addition of smoking attenuated the long-term area SES effect by 20%; all other additional variables had little effect. This study was well designed and executed; however, some case-control studies are more likely to experience selection, participation and recall bias (dos Santos Silva 1999) and the number of cases and controls to support the long-term area SES measure was relatively small. Furthermore, and as mentioned previously, Peto clearly identified modelling the effect of smoking using pack-years as being undesirable. It is more effective and accurate from a lung cancer disease aetiology perspective to use the components of pack-years: number of cigarettes smoked daily and particularly duration of smoking (Peto 2012). Finally, the study considered vegetable consumption only; however, fruit consumption has also been identified as a potential risk factor for lung cancer (WCRF/AICR 2007). Interestingly, and albeit small compared to either the SES or smoking variables, weekly meat consumption was associated with a 17% greater risk of lung cancer after full adjustment, but this was not associated with neighbourhood SES suggesting that it was not a mediator in the neighbourhood SES – lung cancer association (OR 1.17 95% CI 1.06, 1.29) (Hystad et al 2013).

***Sondergaard et al (2013)*** conducted a large population cohort study in Denmark using data linkage. They evaluated the family environment in childhood and the genes shared by siblings to investigate whether these factors explained the inverse association of education level and lung cancer such that higher lung cancer risk was associated with lower education level. The large cohort was

comprised of 1,381,369 individuals; 1,415 of whom were diagnosed with lung cancer. Individuals were followed from age 28 until the year 2009. Education level was defined as primary school, high school, vocational, short and middle length higher education and advanced higher education. The covariates the researchers included were: sex, psychiatric hospitalisations in young adulthood and disability pension at age 28. The last two variables were used as an indicator of serious health conditions, both of which may affect level of education attained. The researchers performed both cohort (the unadjusted model) and inter-sibling analyses (the adjusted model) to estimate hazard ratios (HR) using Cox regression models. The authors used likelihood ratio tests to assess if a linear trend existed over the five education levels (primary school, high school, vocational education (reference), short and middle-length higher education and advanced higher education. The researchers found that, in the *cohort* analysis (considered the unadjusted model), compared to a vocational education, those with the lowest education (primary school) had the greatest risk of lung cancer (HR 1.64 95% CI 1.45, 1.84). Notably, the lung cancer risk reduced for each step up in the education ladder. In the cohort analysis, a trend estimate over the five education levels was significant (HR 0.76 95% CI 0.73, 0.79). However, the risk of lung cancer for *siblings* with only primary school educational attainment (compared to those with a vocational education) was not significant (HR 1.24 95% CI 1.00, 1.54). Similarly, in the *sibling* analysis all the HRs for each of the education levels was fully attenuated, however the trend estimate for every step up the education ladder (HR 0.89 95% CI 0.82, 0.96) remained significant. The authors concluded that factors shared by siblings explained a part of the association between education level and lung cancer risk suggesting that shared sibling exposure to known lung cancer behaviour risk factors such as smoking, diet and physical activity may explain the findings. They pointed out that genetic and non-genetic factors had not been collected, so could not be considered. This study provided helpful insight into the association of educational attainment and lung cancer risk by contributing to the discussion on the role of SES over the life-course and in the pathway to lung cancer induction. In particular, through the sibling analysis, the study suggested that low family

circumstances in childhood had an effect on educational attainment and ultimately risk of lung cancer (Sondergaard et al 2013).

***Boscoe et al (2014),*** in a cancer registry study described in more detail in Section 2.2.2, used an area poverty indicator based on income to study cancer incidence association with SES by site for males and females separately in the United States. Those with low socioeconomic circumstances were compared to those with high socioeconomic circumstances. They reported that lung cancer incidence rate ratios were 1.6 and 1.2 for males and females respectively where confidence intervals excluded 1.0. No individual indicators of SES were available or analysed, a simple measure of socioeconomic inequality was adopted, no risk behaviour data was available and measurement of SES occurred at diagnosis (Boscoe et al 2014).

***Garcia-Gil et al (2014)*** linked an area SES indicator to a longitudinal database of medical records for a representative population of Catalonia, Spain to explore area SES association with incidence of cancers during 2009-2012, including lung cancer. The area SES measure that was adopted considered the proportion of the census tract population in 2001 that was described as: unemployed, a manual worker, a temporary worker, attained basic education only and had dropped out of school before 16 years old. The Incidence rate ratios (IRR) adjusted for sex, age, and behaviours such as smoking, alcohol consumption and obesity were considered and co-morbidities such as hypertension and diabetes were collected at the 2009 baseline. They identified that fully adjusted IRRs for lung cancer incidence of the most socioeconomically disadvantaged compared to least socioeconomically disadvantaged remained elevated (IRR 1.16 95% CI 1.08, 1.25). However, when comparing the most to the least socioeconomically disadvantaged, the effect of low socioeconomic circumstances on the risk of lung cancer was stronger for men (IRR 1.47 95% CI 1.35, 1.59) and reversed for women (IRR 0.79 95% CI 0.66, 0.93). They explained these findings via the differences in age-sex composition of the study population and prevalence of cancer risk behaviours, exposure to occupational carcinogens and diet; but did not fully explain why, for women, the risk of lung cancer was lower for those in

the lower socioeconomic strata. The study recognised the long latency period for cancer in using an area SES indicator based on 2001 data for a study period 2009-2012, however lung cancer in particular is recognised to have an induction period that is even longer (Kawachi et al 2006). The area SES indicator used was dependent on summarised attributes of the population and did not include any attributes of the area itself. Furthermore, no individual SES measures (individual or multiple individual at person level) of the study participants were available for analysis (Garcia-Gil et al 2014).

***Mitra et al (2015)*** evaluated social determinants of lung cancer incidence using a Canadian population-based prospective cohort created by linking the Canadian census holding individual SES indicators with the Canadian cancer registry and comprised of 2.7 million individuals. Using age-standardised incidence rates, rate ratios and rate differences between the least and most socioeconomic disadvantaged, they quantified the risk for all lung cancer and for each histologic subtype (adenocarcinoma, squamous cell carcinoma and small cell carcinoma) for each of the three SES indicators used (individual educational attainment, occupation and pre-tax income). Analysis was performed by age group and sex; no behaviour data were available for adjustment and only individual level SES indicators were used, omitting any area SES indicator. They established that lung cancer risk was greater for those in the lower socioeconomic strata for all three SES indicators and for both sexes, but the associations for females were weaker. Rate differences indicated that if all of the cohort members had experienced the rate of those with a university degree, lung cancer incidence would have been reduced by 56% in men and 55% in women. With respect to income and using the experience of those in the highest income quintile, the incidence would have been 33% and 25% lower in men and women, respectively. Finally, if all cohort members had experienced the rate of those in managerial occupations, the incidence would have been 54% lower in men and 44% lower in women. Squamous cell and small cell carcinoma were also distinctly associated with all three SES indicators such that the incidence was greater with increasingly low socioeconomic circumstances; and the risk of squamous cell lung cancer was greatest for those with less than a secondary

education (RR 3.3 95% CI 2.9, 3.9). However, adenocarcinoma was associated only with education and income. As in many of the studies reviewed, education was identified as having the strongest association with lung cancer risk (Mitra et al 2015).

***Li et al (2015)*** conducted a population cohort study that was based in Sweden that was comprised of 3.2 million people aged 50 years old or older; 33,704 of whom were diagnosed with lung cancer. Their objective was to explore the association of neighbourhood deprivation with lung cancer risk after adjusting for individual SES measures. A neighbourhood deprivation index was constructed using education, income, unemployment and welfare assistance data. The investigators also used two individual SES indicators: family income and education. Using multilevel logistic regression, the authors computed adjusted ORs for the individual SES variables, age, sex and the co-morbidities (including chronic obstructive pulmonary disease) and tobacco abuse which were also applied as a surrogate for smoking behaviour. Other covariates used for further adjustment included marital status, immigrant status, urban/rural status, geographic mobility, alcoholism and related liver disease. A consistent pattern of higher incidence rates with each increasing level of neighbourhood level deprivation was observed across all individual level socio-demographic categories and co-morbidities. In addition, all categories showed a gradient effect across the levels of neighbourhood deprivation. The fully adjusted results indicated high neighbourhood deprivation was associated with increased lung cancer incidence (OR 1.27 95% CI 1.22, 1.32). The greatest odds of lung cancer incidence were among men (OR 1.44 95% CI 1.41, 1.47), immigrants (OR 1.18 95% CI 1.14, 1.22), those with the lowest educational attainment level (OR 1.59 95% CI 1.53, 1.66), or those who were also affected by co-morbidities (OR range 1.30 to 3.69). Again, however, smoking behaviour data were not available to fully assess the impact of this potential confounder (Li et al 2015).

***Behren et al (2016)*** conducted a large case-control study to explore occupational prestige (Treiman's Standard International Occupational Prestige Scale) over a trajectory in order to understand the development of occupational

prestige over the working life and its association with lung cancer risk by histological subtype in men. Lung cancer cases and controls were sourced from an international pooled dataset including 12 studies from 13 countries with detailed job histories and smoking behaviour data (smoking status and pack-years). Models were adjusted for centre, age, any employment in occupations with established lung cancer risk and education level. Fully adjusted ORs showed a monotonic direct relationship of increased incidence with increased low socioeconomic circumstances for all lung cancers, squamous cell carcinoma and small cell carcinoma but were less clear for adenocarcinoma. Evaluating fully adjusted ORs for change in social mobility between first and last occupation did not show a clear pattern; only medium to low prestige (downward mobility) and low to medium prestige (upward mobility) was associated with increased lung cancer risk (OR 1.24 95% CI 1.08, 1.41 and OR 1.19 95% CI 1.04, 1.36 respectively). For change in social prestige from first occupation to longest occupation, only downward mobility for medium prestige to low prestige was associated with increased lung cancer risk (OR 1.16 95% CI 1.01, 1.32). The authors concluded that low occupational prestige in men is associated with lung cancer independent of smoking behaviour and occupational exposures as smoking behaviour only partly attenuated the elevated ORs between lung cancer and occupational social prestige.

Behren's (2016) study was well designed and comprehensively evaluated the association of occupational prestige, social mobility and lung cancer risk in men including discussion of the pathways through which occupational prestige may influence lung cancer development and incidence. Despite being a large pooled case-control study, it was subject to potential bias such as retrospective recall of smoking behaviour and potential differences in recall accuracy between those who had high versus those who had low occupational prestige. Furthermore, because of its international construct, participating countries were likely to be in different stages of the smoking epidemic with different associations with the different social classes. A population based cohort study design may overcome some of these limitations (Behrens et al 2016).

As described in Section 2.2.2, ***Leuven et al (2016)*** conducted a population-based case-control study in Norway to explore the association of education with cancer and also included a separate lung cancer analysis (Leuven et al 2016). Hazard ratios (HR) for males and females respectively indicated a relatively large lung cancer protective effect of increased education (males: HR 0.8 SL 99%, SE 0.003; females: HR 0.88 SL 99% SE 0.004). The implication was that one extra year of education was associated with a 12% reduction for women and 20% reduction for men in lung cancer risk. The HR estimates for education reform were only significant for men (males: HR 0.89 SL 5%, SE 0.05; females: HR 0.96 SL 5% 0.06). As stated previously (Section 2.2.2) two observations to note on study design that can be made are the relatively short follow-up age (67 years old) and the focus on compulsory education excluding further education (Leuven et al 2016).

***Vohra and colleagues (2016)*** conducted a rapid-review of the literature to evaluate the relationship between SES in childhood and cancer in adulthood. Twenty-two publications from 13 studies were identified from the North American and European countries which focused on individual SES measures during childhood and cancer outcomes (incidence and mortality). Most studies were cohorts with retrospective data collection on childhood circumstances, were focused on both men and women and participant follow-up was achieved via linkage to cancer registries. Childhood socioeconomic circumstances were in general established via participant surveys at the start of the study and were most commonly measured via father's occupation. Adult socioeconomic circumstances were most frequently measured by occupation, education or deprivation level. Studies also included covariates (where available) such as tobacco use, alcohol consumption, physical activity, BMI, height, weight, blood pressure, lung function, psychosocial measures of stress and blood lipid measurement. With respect to lung cancer in particular, the rapid-review concluded that childhood socioeconomic circumstances were most likely to contribute, along with adult socioeconomic circumstances, to lung cancer risk through cumulative exposure to smoking; however, the stronger effect was in the adult SES. Nevertheless, the authors discussed that a residual influence of

childhood SES remained, implying that early establishment of smoking behaviour in families with low socioeconomic circumstances may be a critical mechanism for lung cancer causation. Although unable to provide an aggregated quantification of the risk association between childhood SES and lung cancer due to the small number of cases and differences in the approaches adopted, this rapid-review did provide a useful assessment of the current status of the literature in this area and provided useful direction for further research (Vohra et al 2016).

***Denton et al (2017)*** conducted a relatively simple descriptive study of 2,369 lung cancer patients in Australia to review area SES and lung cancer patient attributes (histology, geographic area of residence and smoking behaviour) and survival. A postcode based SES indicator incorporating multiple social and economic variables was used; the survival analysis was adjusted for smoking status. Focusing on the findings related to incidence, they found that there was no difference between socioeconomic groups in the proportion of non-small cell lung cancer (93%) to small cell lung cancer (7%). However, among the low SES group with non-small cell lung cancer, there were higher rates of squamous cell carcinoma (27% versus 22%, low and high SES respectively). With respect to smoking behaviour, a statistically higher proportion of low SES patients were smokers compared to the high SES group (92% versus 82% respectively $P<0.01$) (Denton et al 2017). The authors provided no further explanation for the observed socioeconomic inequalities or lack of inequalities (Denton et al 2017).

### 2.3.3 Summary of lung cancer literature findings

In summary, regardless of the SES measure used, and with the exception of one study (Denton et al 2017) which was a small descriptive analysis of hospital cases, all 16 studies reported the consistent finding of greater lung cancer risk with lower socioeconomic status with the risk greater for men than women. The only exceptions were two studies that reported either no SES association (Kutnetkov et al 2011) or a protective effect of low area deprivation for women (Garcia-Gil 2014). This latter study performed in Catalonia, Spain may, in

particular, reflect the differences in stage of the smoking epidemic between men and women and between northern and southern European countries (Karim-Kos et al 2008; Lortet-Tieulent et al 2014). Of the 11 studies that included educational attainment in the SES variables, six studies (Mouw et al 2008; Sidorchuk et al 2009; Nkosi et al 2012; Mitra et al 2015; Li et al 2015) reported that higher lung cancer risk was associated with low education; and that this risk was the greatest of all the SES variables investigated. In the study performed by Dalton et al (2008a), the important role of education was reflected in the fact that the minimally adjusted models included education level while Leuven et al (2016) in their natural study of the education reform in Norway, found that one extra year of school was associated with a 12% risk reduction in women and a 20% reduction in men. Finally, Sondergaard et al (2013) through their education and sibling analysis suggested that family circumstances in childhood had an effect on educational attainment and ultimately lung cancer risk. Of the six studies that considered area-based SES variables, one study (Hystad et al 2013) concluded that long-term area SES (compared to short-term SES) was associated with greater risk after adjustment for individual behaviours and individual SES variables. All six studies (Kuznetkov et al 2011; Meijer et al 2013; Hystad et al 2013; Boscoe et al 2014; Garcia-Gil et al 2014; Li et al 2015) investigating area-based SES, found that greater area deprivation was associated with greater lung cancer risk for both genders with the exception of two studies (Kuznetkov et al 2011; Garcia-Gil et al 2014;) which found either no association or a protective effect for women respectively. With respect to histological subtype, these two studies reviewed identified greater risk association with low SES for all lung, squamous cell and non-small cell carcinoma subtypes, but not adenocarcinoma of the lung (Mitra et al 2015; Behren et al 2016).

## 2.4 Head and neck cancers

### 2.4.1 Introduction

Eight studies which included head and neck cancer risk association with low socioeconomic circumstances are described below. Of these studies three were cohort studies (Anderson et al 2008; Mouw et al 2008; Purkayastha et al 2016), one study was a systematic review and meta-analysis of case-control studies (Conway et al 2008), two further studies were case-control (Conway et al 2010b, Conway et al 2015) and the remaining two studies were descriptive cancer registry studies (Conway et al 2007; Boscoe et al 2014). Four studies adjusted by risk behaviours including, at minimum, alcohol and smoking (Conway et al 2008; Mouw et al 2008; Conway et al 2010b; Conway et al 2015) while the remaining studies were not able to adjust for behavioural risk factors (Anderson et al 2008) due to availability (Conway et al 2007; Anderson et al 2008; Boscoe et al 2014; Purkayastha et al 2016). With respect to SES measurement, three studies considered individual variables only ((Mouw et al 2008; Conway et al 2015), three studies considered area SES variables only (Conway et al 2007; Boscoe et al 2014; Purkayastha et al 2016), while the remaining two studies considered both individual and area SES variables (Anderson et al 2008; Conway et al 2010b). None of the studies considered presentation of the SES inequalities using a complex method such as the RII or SII; in all cases, rate ratios, odds ratios or age and sex standardised rates were used. Six of the studies measured SES exposure at time of diagnosis thereby omitting the temporal relationship between exposure and diagnosis (Conway et al 2007; Conway et al 2008; Conway et al 2010b; Boscoe et al 2014; Purkayastha et al 2016), while Mouw et al 2008 did consider temporal relationship by using education as the only SES variable. Anderson (2008) measured SES two years before diagnosis; however, this is unlikely to reflect the significantly longer lag time between exposure and diagnosis identified for cancer in general (Anderson et al 2008).

### 2.4.2 Publications

***Conway et al (2007)*** conducted a descriptive epidemiological analysis of the Scottish Cancer Registry of oral cancer by Carstairs area-based deprivation index (1976 to 2002). The authors observed a widening of socioeconomic inequality in the burden of oral cancer. This inequality emerged in the late 1970s in men and in the 1980s in women. By 2002, there was a dose-like response with both men and women having more than doubled the incidence rates (Conway et al 2007).

***Anderson and colleagues (2008)*** used a large Danish population cohort study to focus on the association of SES with cancer incidence of the mouth and pharynx together and larynx separately. This study was part of a series that focused on socioeconomic inequalities in cancer risk, mortality and survival in Denmark which is described in greater detail in Section 2.2.2. Little discussion was offered to explore SES pathways to diagnoses other than risk behaviours such as tobacco and alcohol consumption, diet, oral hygiene and infection (Anderson et al 2008).

For all three cancers, they established in both men and women decreasing incidence with increasing social advantage (longer education, more income, closer work market affiliation (employment status), better housing tenure and larger dwelling). For men, higher social class (creative core and professional) as well as agricultural class were associated with lower mouth and pharynx incidence compared to manual workers. For females, agricultural class had the lowest mouth and larynx incidence risk compared to manual workers; the other social classes were not significant. For both males and females and for all three cancers, early retirement pensioners had a much higher risk of all three cancers with the IRRs. IRRs for housing variables were greater than those for the other SES variables. It may be suggested that this observation regarding housing SES variables may reflect the housing conditions in Denmark or it may suggest that accumulated wealth over a long period of time which may be related to the SES-level exposure over that period and may be particularly relevant to cancer diagnosis which is known to have a long latency period (Anderson et al 2008).

***Conway et al (2008)*** conducted a systematic review and meta-analysis of case-control studies to assess socioeconomic inequality and oral cancer risk. Forty-one studies provided 15,344 cases and 33,852 controls and three individual SES indicators were used (educational attainment, occupational social class, monthly household income); low SES in each indicator was strongly associated with increased oral cancer risk. Pooled OR 1.85 95% CI 1.60, 2.15 for low relative to high educational attainment; 1.84 95% CI 1.47, 2.31 for low relative to high occupational social class; and 2.41 95% CI 1.59, 3.65 for low relative to high income were calculated. Not all the studies adjusted for any or all of the confounding variables: age, sex, tobacco use, alcohol consumption, but when comparing OR results for studies which had adjusted for confounding variables compared to those that had not, no significant differences in the results were identified. The main limitation was related to the underlying case-control nature of the studies included (Conway et al 2008).

***Mouw et al (2008)*** conducted an education and cancer risk cohort study in the US that is described in more detail in Section 2.2.2. For head and neck cancer risk in men who had not completed  compared to men with a postgraduate education, after adjustment for smoking, alcohol consumption and other behaviour factors, the RRs were just fully attenuated but remained elevated (RR=1.29 CI 95% 0.99, 1.67) and were not significant for women (RR=1.21 CI 95% 0.69, 2.13). SES measured by one factor only (education) may omit effects associated with other SES dimensions both individual and area-based. Given education is generally established in early adulthood, it is likely to capture SES early in the life-course which is commensurate with the long lead-time between cancer initiation and diagnosis. However, other SES dimensions such as occupational status may also contribute to head and neck cancer risk which could be explored through further research. Furthermore, other known oropharyngeal cancer risk factors such as human papillomavirus were not considered (Mouw et al 2008).

To investigate the SES association with oral cancers in more depth, ***Conway et al (2010b)*** conducted a further case-control study in Scotland (nested within a

larger multi-centre European study) to explore the oral cancer incidence that was associated with two area SES indicators (Carstairs and SIMD), eight individual SES indicators (educational attainment, years in education, first occupation, last occupation, longest occupation, every manual, social mobility and period of unemployment), along with various behaviours (vegetables consumed per week, fruit consumed per week, mean lifetime alcohol units consumed per week and smoking status). The study had a small number of case-control pairs (n=~100). Their results showed that those living in the most socioeconomically disadvantaged areas compared to the most advantaged (OR 4.66 95% CI 1.79, 12.18) and those who were unemployed compared to those employed (OR 2.27, 95% CI 1.21, 4.26) had higher risk of cancer than those with high educational attainment compared to those with a secondary school education (OR 0.17, 95% CI 0.05, 0.58). After adjustment for smoking and alcohol consumption, all SES indicators were not significant with smoking being such a dominant risk factor with nearly all case participants reporting a smoking history (Conway et al 2010b).

***Boscoe et al (2014)*** used an area poverty indicator that was based on income to study cancer incidence by site for males and females separately in the United States; this study is described in greater detail in Section 2.2.2. They reported that oral cavity and pharyngeal cancer incidence rate ratios comparing the highest poverty category to the referent lowest poverty category were 1.42 for oral cancer and 1.21 for pharyngeal cancer. Laryngeal cancer had the highest rate ratios of 1.85 for men and 2.08 for females while nasopharyngeal cancer rate ratios were 1.8 (male) and 1.1 (female). For all head and neck cancer sites, confidence intervals excluded 1.0 and rate ratios across the four area poverty categories were monotonic such that the risk of diagnosis increased with increasing area poverty.

***Conway et al (2015)*** through the global INHANCE consortium estimated the association of head and neck cancer risk with education and household income by age, site, sex and geographic location. Thirty-one case-control studies from 27 countries contributed 23,934 cases and 31,954 controls creating a large study

population and overcoming the limitation of study size and, as a consequence, study power. Education was standardized and stratified into low education (including no education, primary education or first stage of basic education), intermediate education (lower secondary or second stage of basic education or completed upper secondary education) and high education (including further education, vocational education and higher education). Household income, available in only the seven US case-control studies, was standardised and stratified into five groups. The authors identified that the odds of being diagnosed with head and neck cancer was more than two-fold for those with low education compared to those with a high education (OR 2.50 95% CI 2.02, 3.09). While smoking and alcohol consumption explained much of the additional risk, 31% was not explained and remained elevated with a 61% increase in risk even among never smokers and never drinkers (OR 1.61 95% CI 1.13, 2.31). Low household income relative to high household income was associated with over two-fold extra risk of head and neck cancer (OR 2.44 95% CI 1.62, 3.67) with 39% not explained by smoking and alcohol. There were no differences by age, sex, or head and neck cancer subsite. Taking into consideration smoking and alcohol behaviour, the risk of head and neck cancer was greatest (65% increase in risk) for those with low education living in higher income inequality countries (OR 1.65 95% CI 1.27, 2.15). This study also uniquely was able to remove the question of residual confounding by smoking and alcohol with an analysis of the risk association among never tobacco/alcohol users. The risk associations' odds ratios were comparable to the adjusted estimates for both education and income. The authors fully reviewed the pathways in which SES may confer head and neck risk beyond behaviour factors suggesting psychosocial, material and life-course pathways may explain the proportion estimated as not related to behaviour, particularly tobacco and alcohol consumption (Conway et al 2015).

More recently, ***Purkayastha et al (2016)*** updated the historical incidence trends analysis, in the Scottish Cancer Registry using the SIMD index from 1975 to 2012 and projected incidence from 2012 to 2025 with better refinement of the head and neck cancer subsites: oral cavity, oropharyngeal cancer and laryngeal cancer. The study identified that 28,217 diagnoses were made over the historical

period. Age-sex standardised rates were calculated and SES was measured using primarily the Carstairs 1991 Index and SIMD for the period 2000 to 2012. Peak incidence occurred for the 61-65 age group for oropharyngeal cancer (RR 2.34 95% CI 2.08, 2.63) but in the older 76-80 age group for oral cavity (RR 3.54 95% CI 3.20, 3.91) and 71-75 age group for laryngeal cancers (RR 4.74 95% CI 4.30, 5.23). Compared to the least deprived group, the most deprived group had a more than two-fold increased risk of head and neck cancer (RR 2.59 95% CI 2.45, 2.74) with laryngeal cancer having the highest risk with more than a three-fold increased risk (RR 3.34 95% CI 3.02, 3.69). For head and neck cancers together and using European age-standardised rates a clear SES gradient, where incidence was greater for each subsequent SIMD decile, was observed for the period 1975 to 2012. The relative risks of the most deprived compared to the least deprived broadly correlated across subsites and there was no different relationship for oropharyngeal cancer which was a reported clinical finding. Incidence projections per 100,000 population indicated a striking increase from 17 in 2012 to around 25 in 2025 for head and neck cancer comprised mainly from rapid increase in oropharyngeal cancer while rates remained stable at around five to seven for oral cavity cancer and began to decrease from around six to four for laryngeal cancer. As a population cohort design, the study provided a very thorough and robust assessment of head and neck cancer incidence trends utilising cancer registry data. However, due to lack of data availability at population level and the datasets used which did not collect behavioural/HPV data; this study did not consider behaviour factors or information on HPV status or individual SES measures which may provide insight into the factors underlying the trends identified. Moreover, the Carstairs or SIMD index was recorded on date of diagnosis so the temporal relationship could not be ascertained (Purkayastha et al 2016).

### 2.4.3 Summary of all cancer literature findings

In Denmark, Anderson et al (2008) identified that regardless of the SES variable used, head and neck cancer incidence decreased with greater social advantage (Anderson et al 2008) with early retirement pensioners at a much higher risk.

Similarly, in Scotland, Conway et al (2008) found low socioeconomic circumstances were strongly associated with increased risk of head and neck cancer (Conway et al 2008) with this inequality initiating in the 1970's for men and 1980's women (Conway et al 2007). Focusing on educational attainment only, Mouw et al (2008) found that compared to a postgraduate education, men with high school education only were at a greater risk of head and neck cancer but that this risk was attenuated after adjustment for smoking and alcohol consumption and insignificant for women (Mouw et al 2008). Conway et al (2010b) observed similar findings, recognising that smoking was such a dominant behaviour risk factor (Conway et al 2010b). When using a large study comprised of multiple international case-control studies providing greater power, Conway et al (2015) was able to quantify the smoking and alcohol contribution to the two-fold elevated risk of head and neck cancer associated with low education at 61%; leaving the balance of elevated risk unexplained (Conway et al 2015). Finally, laryngeal cancer was identified as the head and neck site with the greatest risk association with low SES (Boscoe et al 2014; Purkayastha et al 2016)

## 2.5 Oesophageal cancer

### 2.5.1 Introduction

There were 13 studies focusing on oesophageal cancers. Eight of these were cancer registry studies providing a description of the current position or trends over time (Brewster 2000; Baastrup et al 2008; Cooper et al 2009; Gossage et al 2009; Coupland et al 2012; Boscoe et al 2014; Kiadaliri 2014; Bodek et al 2016). Two of the studies were case-control design (Giri et al 2014; Caygill et al 2014b), two were cohort studies (Mouw et al 2008; Lagergren et al 2016), while the final study was a report covering several different study types (Kogevinas et al 1997b). Most of the studies did not adjust for behaviours leaving two that adjusted for at least smoking and alcohol behaviours (Kogevinas et al 1997b; Mouw et al 2008). With respect to the SES variables employed, five of the studies included individual SES variables (Kogevinas et al 1997b; Baastrup et al 2008; Mouw et al 2008; Cooper et al 2009; Giri et al 2014; Lagergren et al 2016)

while six did not (Brewster et al 2000; Gossage et al 2009; Boscoe et al 2014; Kiadaliri 2014; Caygill et al 2014a) and the two remaining studies included race or ethnicity (Bodek et al 2016; Coupland et al 2012). Ten of the studies included an area measure of deprivation (Kogevinas et al 1997b; Brewster et al 2000; Cooper et al 2009; Gossage et al 2009; Coupland et al 2012; Boscoe et al 2014; Kiadaliri 2014; Giri et al 2014; Caygill et al 2014a; Bodek et al 2016) while three studies used only the individual SES variable(s) (Baastrup et al 2008; Mouw et al 2008; Lagergren et al 2016). Only one study considered the temporal relationship between SES exposure and incidence (Mouw et al 2008); similarly, only two studies applied a more complex measure of SES that reflected the full spectrum of SES groups (Kiadaliri 2014; Bodek et al 2016).

### 2.5.2 Publications

***Kogevinas et al (1997)*** reported that oesophageal cancer risk was also socially patterned for both men and women and associated with the high risk behaviours of smoking and alcohol consumption, reflecting the synergistic effect of these behaviours as well as their individual effects. The IARC report also indicated that squamous cell carcinoma was more likely to be associated with these behaviours compared to oesophageal adenocarcinoma which occurs at the junction of the stomach and oesophagus (Kogevinas et al 1997a).

***Brewster et al (2000)*** in Scotland analysed the Cancer Registry data for incident cases of the oesophagus by histological type from 1977 to 1996 using sex and age standardised incidence rates by deprivation category. SES was measured using the Carstairs deprivation categories, which is a census based area measure of SES. Incidence of oesophageal adenocarcinoma increased strikingly over the period for both men and women (139.5 and 124.6 estimated percent change 1977 to 1996). No association of oesophageal adenocarcinoma with SES was identified for either men or women. However, the SES analysis was limited to a historic (large) area-based measure, with no individual measures available. And this measure was linked to the patient's postcode at diagnosis limiting the ability to determine a temporal relationship (Brewster et al 2000).

***Baastrup et al (2008)*** reviewed the risk association of SES for oesophageal cancer in a large population study in Denmark. Age-period standardised incidence rates of oesophageal cancer decreased in a stepwise manner with increasing education level; but were stable for women. Adjusted IRRs showed decreasing incidence with increasing social advantage for work market affiliation (employment status), social class, housing tenure and dwelling size for men. For women, only those who were early retirement pensioners or rented a home were associated with elevated oesophageal cancer risk. Behaviour risks that are associated with oesophageal cancer (smoking, alcohol and obesity) were not considered in the models used due to data not being available. The authors referred to studies estimating that 50% and 40% of all incident oesophageal cancers were caused by tobacco smoking and alcohol consumption respectively, both of which were known to be associated with low SES (Section1.7.4). The extent that the calculated SES association could be attenuated was not reviewed. In addition, area indicators of SES were similarly not available for analysis (Baastrup et al 2008).

***Mouw's et al (2008)*** cohort study which is described more fully in Section 2.2.2 found that for men without a diploma, compared to those with a postgraduate education, the relative risk of oesophageal cancer was doubled even after full adjustment for smoking and other behaviour factors (RR= 2.00 CI 95% 1.39, 2.86). Potential explanations for this "stark" finding considered by the authors were residual confounding by smoking or other psychosocial or biological factors. The author focused only on education and perhaps other, additional SES measures may shed further light on the aetiology of this disease (Mouw et al 2008).

***Gossage et al (2009)*** evaluated the effect of economic deprivation from 1993 to 2002 on oesophageal cancer incidence in the London area. Using the income domain of an area IMD at time of diagnosis, they established that from 1993-95 to 2000-02 the incidence of oesophageal cancer amongst affluent males increased by 51% while it increased only two percent amongst the most socioeconomically deprived males. A higher proportion of low SES vs. high SES

patients (24% v. 17%) diagnosed with oesophageal cancer were under 60 years old (p=0.04) and 40% of all cases were squamous cell carcinoma among the low income group compared to the 31% among the high income group (p=0.03). The authors proposed that increasing adenocarcinoma amongst the affluent was likely to be associated with increased obesity and gastro-oesophageal reflux disease (GERD), but raised that there was evidence that obesity was prevalent among the more socioeconomically deprived too. Higher squamous cell carcinoma among the low income group was likely to be explained by higher prevalence of smoking among this group. This study clearly defined the direction of travel for histological types of oesophageal cancer; however, measurement of SES at diagnosis captures a point in time when SES may be the result of diagnosis rather than the cause of diagnosis. Furthermore, the lack of behavioural data presents an opportunity for refinement through further research (Gossage et al 2009).

***Cooper et al (2009)*** studied the influence of age, sex, deprivation and ethnicity on oesophageal cancer in the West Midlands, England. They reported that directly standardised incidence rates had increased for oesophageal cancer for both men and women from 1977-1981 to 2000-04. While oesophageal squamous cell carcinoma incidence had not changed. The overall increase reflected large increases of adenocarcinoma in both sexes, but particularly men. Two area deprivation measures were used. The first was the Townsend Index which is based on unemployment, overcrowding non-car ownership and non-home ownership at the postcode level and the second was the income domain of the Index of Multiple Deprivation which consisted of the percentage of the area population claiming various income-based benefits and tax credits. Both analyses by both measures showed an inverse relationship with incidence for squamous cell carcinoma which was lost in later years for both sexes, but this loss occurred at an earlier point for women. The study contributed to a better understanding of the change in oesophageal cancer incidence over time and how that has changed for men and women in England. The assessment of SES was limited by the use of area measures which were more assessable but subject to ecological fallacy (Boscoe et al 2014), reflected only one facet of SES and moreover, the

postcode at diagnosis omitted the temporal inference (Cooper et al 2009).

***Coupland et al (2012)*** conducted a population-wide study in England describing incidence and survival of oesophageal cancer by anatomical region (upper and middle, lower, not otherwise specified and gastric cardia) and area deprivation using the National Cancer Data Repository which contained information from the eight English cancer registries on all patients diagnosed with cancer in their catchment area. The study reported that incidence was greater for men than for women and in those from more socioeconomically deprived areas using the income domain of the Index of Multiple Deprivation. Most tumours were located in the lower oesophagus and among the more deprived. The difference between the sexes in incidence rates at four-times for men compared to that for women was greatest for lower oesophageal cancers. Risk factors such as reducing H pylori infection, increasing obesity and increasing gastro-oesophageal reflux disease were discussed as possible explanations for the dominant and increasing incidence of cancer in the lower oesophagus. The authors suggest that these risk factors are likely to be more common among lower SES groups. However, no individual measures of SES and no adjustment for risk behaviours was performed and the postcode was recorded at diagnosis omitting the temporal inference (Coupland et al 2012).

***Boscoe et al (2014)*** undertook a study using an area poverty indicator study described in detail earlier (Section 2.2.2) that analysed cancer incidence by site for males and females separately. They reported that oesophageal cancer incidence rate ratios were 1.33 and 1.19 for males and females respectively where confidence intervals excluded 1.0 for those living in poor areas compared with those living in more affluent areas. Rate ratios were also increased with increasing deprivation over the four area poverty categories used. Boscoe's analysis by area poverty did not consider the role of individual SES indicators nor did it incorporate behavioural confounders such as smoking or consider subtypes of oesophageal cancer which were likely to demonstrate different characteristics (Boscoe et al 2014).

***Kiadaliri et al (2014)*** conducted a descriptive study in Iran, a lower/middle-income country using the National Cancer Registry and focusing on gender and social inequalities in oesophageal cancer incidence over the period 2003 to 2009. Using the human development index to measure SES and RII to measure socioeconomic inequality, they found an inverse relationship between SES and oesophageal cancer incidence where incidence increased with decreasing SES. In their review of the literature, they found that unlike other countries (USA (Brown et al 2001), Finland (Weiderpass et al 2006), Puerto Rico (Torres-Cintron et al 2012)), risk was similar for both males and females (female to male rate ratio by year hovered around 1.0 with CI including 1.0). The authors explained that in a high incidence area like Iran, this observation was not unexpected. Explanations for their findings included smoking, low consumption of fruit and vegetables and obesity given these behaviours were more prevalent among low SES areas, although these factors were not controlled for in the study (Kiadaliri 2014).

***Caygill et al (2014)*** evaluated social deprivation in Barrett's oesophagus as a precursor to oesophageal adenocarcinoma in Rotherham, England using 1,076 diagnosed Barrett's oesophagus from 1978 to 2012. The area SES index of multiple deprivation (IMD) was used. The study cases were divided before and after 2001 based on date of diagnosis. Case distribution amongst the SES strata was similar to the Rotherham population before 2001, but the two most affluent groups had a 37% increase in cases after 2001 indicating a quantitative link between Barrett's oesophagus as a precursor to oesophageal adenocarcinoma and SES (Caygill et al 2014a). As discussed in Section 1.7.4.5 gastro-oesophageal reflux disease is associated with adiposity as well as being a risk factor for Barrett's oesophagus, a known precursor of adenocarcinoma of the oesophagus. The exact mechanism for the adiposity, as it relates to affluence, was not thoroughly discussed as being overweight or obese was most common among the more disadvantaged in the United Kingdom (Loring et al 2014)

***Giri et al (2014)*** conducted a retrospective case-control study in 2014 at a tertiary hospital in India. They described the characteristics of 207 oesophageal

cancer cases indicating 30.9% were illiterate, 73.9% were in the lowest SES group and 28.0% were farmers living in rural areas. Although no behaviour data was available, the authors reported that alcohol and tobacco consumption (cigarette, bidi or both) were prevalent in the study area, amongst farmers and the lower SES groups. This study included a relatively small number of cases and presented the first steps of understanding oesophageal cancer incidence in an area of India. However, it demonstrated that oesophageal cancer risk inequalities existed in developing countries as well as the developed world reinforcing the picture that socioeconomic inequality is pervasive and exists regardless of the 'wealth' of a country (Giri et al 2014).

***Bodek et al (2016)*** presented preliminary findings at a 2016 American Gastroenterological Association conference on trends in incidence and survival of oesophageal cancer (adenocarcinoma and squamous cell separately) in the United States from 1992-2007 using the Surveillance, Epidemiology and End Results (SEERS) cancer registry database. The area SES measure used was the proportion of the population below poverty line (>15%); both absolute and relative socioeconomic inequality measures were used. They focused on racial disparities; however, they found that in poorer areas oesophageal adenocarcinoma in the non-Hispanic white population was disproportionally increasing while squamous cell carcinoma was decreasing most heavily among the non-Hispanic black population living in poorer areas. The findings for adenocarcinoma were not explained by adiposity as non-Hispanic whites were less disposed to obesity; the authors proposed other genetic factors may play a role. For squamous cell carcinoma, the authors proposed that smoking cessation efforts may be acting on the higher absolute numbers of smokers in poorer areas (Bodek et al 2016).

***Lagergren et al (2016)*** recently conducted a study focusing on marital status, education and income level in relation to oesophageal cancer diagnosis by histological type. This large Swedish population cohort from 1991-2010 found that, compared to those who were married, an increased relative risk of oesophageal cancer for individuals who had been divorced, had never been

married, or were widowed. Those who had the greatest number of years of education or the highest income also had the lowest risk of oesophageal cancer. The associations were in the same direction (reduced risk associated with increased education, increased income. or those who were married) for both histologies of oesophageal cancer but the risks were greatest for oesophageal squamous cell carcinoma. As a large cohort with full follow-up, the study had strong statistical power; however, no information was available on behaviours including tobacco smoking, alcohol consumption, obesity and dietary factors (e.g. high fat, processed and red meat consumption and low fruit and vegetable consumption (Rustgi et al 2014) and exercise (Singh et al 2014)) which may assist in explaining the outcomes observed (excluding alcohol for oesophageal adenocarcinoma, Section 1.7.4.3).

### 2.5.3 Summary of oesophageal cancer literature findings

Greater oesophageal cancer incidence observed among both males and females from low socioeconomic groups was reported by five of the studies (Coupland et al 2012; Boscoe et al 2014; Kiadaliri 2014; Giri et al 2014; Caygill et al 2014a). In the studies investigating educational attainment, a strong association was identified between lower educational attainment and greater oesophageal cancer incidence (Baastrup et al 2008; Mouw et al 2008; Lagergren et al 2016) and was stronger for men. For females, early retirement pensioners and those renting accommodation were at a greater risk of oesophageal cancer (Baastrup et al 2008). A general trend of increasing incidence of adenocarcinoma among the more affluent was observed in one of the studies (Gossage et al 2009). Some studies identified no association of oesophageal adenocarcinoma with SES (Brewster et al 2000) while others identified a trend of disproportionately increasing incidence among those from lower SES groups (Bodek et al 2016). Consistent with the former finding of greater oesophageal adenocarcinoma incidence among those who were more affluent was the 37% increase in the number of affluent Barrett's oesophagus patients compared to the expected number (Caygill et al 2014a). Finally, Coupland et al (2012) identified that most oesophageal cancers occurred in the lower anatomical region of the oesophagus

and nearly 75% of these occurred in males and among the most deprived (Coupland et al 2012). All of the studies offered behavioural explanations for their findings with the social patterning of behaviours offered as an explanation for their findings by eight of the studies (Baastrup et al 2008; Cooper et al 2009; Coupland et al 2012; Boscoe et al 2014; Kiadaliri 2014; Giri et al 2014; Caygill et al 2014a; Bodek et al 2016). Only one study discussed possible pathways including the psychosocial pathway or other biological factors not considered in the study (Mouw et al 2008).

## 2.6 Upper aero-digestive tract (UADT) cancer

### 2.6.1 Introduction

Three studies focused on UADT cancers as a group (Spadea et al 2010; Schmeisser et al 2010; Conway et al 2010a). Two were case control studies (Schmeisser et al 2010; Conway et al 2010a) while the third was a cohort study (Spadea et al 2010). All three studies reviewed individual measures of SES, but only one considered an area measure of deprivation (Spadea et al 2010). Estimates of risk association were based on simple comparison in two of the studies (Schmeisser et al 2010; Conway et al 2010a) while the cohort study provided both relative risks and the complex measure of inequality called the Relative Index of Inequality (Spadea et al 2010). One of the case-control studies measured the SES variables at diagnosis (Conway et al 2010a) while the other case-control study took a life-course view and measured changes in occupational social class over time (Schmeisser et al 2010). The third study measured the SES variables at study entry (Spadea et al 2010).

### 2.6.2 Publications

***Spadea et al (2010)*** via the Turin, Italy study described previously, estimated the association with social inequalities of head and neck and oesophageal cancer together (Section 2.2.2). They established that all four SES indicators were strongly associated with increased UADT incidence for men. After mutual adjustment, housing characteristics (RI I 1.92 95% CI 1.57, 2.35) was the

socioeconomic indicator most strongly associated with the incidence of UADT cancer followed by education (RII 1.82 95% CI 1.39, 2.36), then occupational social class (RII 1.60 95% CI 1.27, 2.02) and finally area deprivation (RII 1.38 95% CI 1.13, 1.68). For women, only housing characteristics (RII 1.87 95% CI 1.17, 3.00) remained associated with UADT cancer incidence after mutual adjustment, but it was strongly associated. Explanations for these results were similar to those discussed under Section 2.3.2 for lung cancer with smoking as well as alcohol being the primary risk factors for UADT cancer. As with smoking, alcohol consumption was socially patterned with the greatest use among those in the lower social groups. Again, no behavioural risk factors were available for this analysis, although the more complex measure of socioeconomic inequality, the Relative Index of Inequality was used and therefore reflected the full social gradient (Spadea et al 2010).

***Conway et al (2010a)*** in a European 14 centre case-control study analysed the association of components of socioeconomic risk individually after adjusting for known behaviours (smoking, alcohol consumption and diet) with UADT cancer risk. Various aspects of occupational social class were evaluated including first, last, longest and current occupation and experience of unemployment. They found that after adjustment for age, sex, centre and behavioural factors that low relative to high education remained strongly associated with UADT cancer risk while low occupational social class variables were fully attenuated. Their analysis suggested that 67% of UADT risk associated with education variables was explained by behaviours of smoking, alcohol and diet; however 33% of SES risk remained. Direct and indirect pathways for how low education increases UADT risk were discussed. Behavioural risks were proposed as an intermediate step in the carcinogenetic pathway stemming from social factors (material, psychosocial, eco-social or life-course). They went on to speculate that the process may result in biological ageing caused by poor social circumstances. This study was limited in its case-control design and hospital-based controls in many European countries (Conway et al 2010a).

***Schmeisser et al (2010)*** conducted a case-control study to investigate the life-course social mobility and risk of UADT cancer in men in a follow-up study to Conway et al (2010a). The full occupational histories were used to assess changes in Standard International Occupational Prestige. SES risk was adjusted by known behavioural confounders (smoking, alcohol consumption and diet) as well as centre and age. They found that, after full adjustment, the OR for the lowest versus highest of social prestige categories was 1.28 (95% CI 1.04, 1.56). When compared to the highest category of social prestige, those with no social mobility for the middle and low prestige categories showed elevated ORs. Fully adjusted site ORs demonstrated that low social prestige was greatest for oesophageal cancer risk (OR 2.02 95% CI 1.26, 3.23). Relative to those who were continuously in the high social prestige group, those who were downwardly mobile had an OR of 1.71 (95% CI 0.75, 3.87). Finally, the gap between controls versus cases of social prestige widened during working life. While this study did consider life-course factors, it was not able to consider parental SES influence on childhood SES (e.g. education) which could affect adult SES. This study evaluated pathways from SES to disease, but could not fully explain the phenomenon. Various theories were discussed including biological ageing, stress induced neuro-endocrine responses leading to chronic inflammation and impaired immune systems and disease susceptibility, along with mental health status as evidenced by self-rated hopelessness correlating with low SES and higher cardiovascular disease risk. The study points out that fewer studies explore the SES pathway influence on cancer risk than for cardiovascular disease. Case-control limitations as described for Conway et al (2010a) hold here (Schmeisser et al 2010).

### 2.6.3 Summary of UADT cancer literature findings

In each case, an increased risk association of SES with UADT cancer incidence was found. Spadea et al (2010) identified that for men this association existed for education, occupational social class, housing characteristics and area deprivation, while only housing characteristics were associated with elevated UADT risk for women. Conway et al (2010a) calculated that education explained

67% of the elevated risk association, but left 33% unexplained, while Schmeisser et al (2010) identified that the risk association of low social prestige was greatest for oesophageal cancer and the gap in social prestige between the controls and cases widened over working-life. Behavioural explanations for the findings were offered by Spadea et al (2010), while Conway et al (2010a) and Schmeisser et al (2010) explored the possible pathways between SES exposure and cancer incidence with behaviour considered an intermediary step.

## 2.7 Gaps identified in the literature

The studies undertaken to-date largely measure SES using a limited number of individual variables or a single area variable. Few studies reviewed included a more comprehensive list of individual SES variables as well as area-based indicators – it was usually one or the other. In order to appreciate their relative importance and contribution to understanding the pathway between SES and cancer incidence, both individual and area-based SES variables are required. There is perhaps an over reliance on area-based measures for routine monitoring of health or cancer inequality which has given the impression that the SES cancer risk relationship is well known and fully understood. However, there is limited use/availability of individual measures of SES and their inter-relationship with area.

Interpretation of those studies relying on area-based SES variables was subject to ecological fallacy (Boscoe et al 2014) and the underlying changes in the definition of area used, both of which may mask or mitigate the true socioeconomic inequalities (Section 1.3.2.4). Analytical approaches to minimise change in underlying aspects of measuring SES such as postcode or datazone definition or geography should be adopted to support the need to focus on socioeconomic circumstances over the life-course. This would support minimising SES change that is a function of the administration of the underlying components of the SES indicator that may mask or mitigate change in the health outcome that is due to the true SES change.

No study considered a combination of measures to try to capture compounded socioeconomic disadvantage. Given the multidimensional nature of SES and the compounding effects of multiple low socioeconomic circumstances over the life-course, there is a need to develop analytical approaches that can capture this compounded effect.

Many of the studies were case-control design with smaller case numbers and associated risks of bias; less frequent were studies using the prospective cohort design. Furthermore, fewer studies have exploited the power of linking administratively collected databases enabling a population-based study design. Those studies based on cancer registry data alone are often fairly simple providing a descriptive epidemiology only and these are further limited by the over-reliance on area-based SES measures at the time of diagnosis thus limiting temporal inference. The long gestation time of cancer and the dynamic nature of SES over the life-course mean that the measurement of SES well before diagnosis is not only justified but mostly necessary in order to reflect the temporal relationship between SES exposure and cancer diagnosis.

With respect to covariates, most studies reported that risk behaviour data were not available to assess confounding; many of those that did incorporate smoking, the major behavioural risk factor for the cancers in question, adopted pack-year variables that are considered to be misleading in epidemiological research (Peto 2012). In addition, many studies performed analysis at anatomical or site group level, often a requirement due to the small number of cases at subsite level. Given that there are different aetiologies for different morphologies, aggregation was likely to mask true SES effects. Although many of the studies discussed and in some cases have analyzed the role of behaviours, fewer studies have explored the complex multidimensional nature of SES in order to understand more fully the role of SES in the causal pathway of cancer incidence. One aspect of this picture is the timing of capture of SES; most studies rely on SES measured at point of diagnosis which is subject to reverse causation bias and neglects the role of SES over the life-course.

Few studies have utilised the complex measurements of inequality, which take into account the SES gradient, such as the Relative Index of Inequality and the Slope Index of Inequality, most studies relied on the more simple approaches of rate ratios of the most and least socioeconomically deprived which fail to acknowledge the full extent of SES inequality.

Finally, the linkage potential of routine administrative health and resource databases in Scotland to investigate inequalities in cancer incidence has yet to be undertaken.

To study and monitor socioeconomic inequalities in health in ways that are useful to informing policies that result in reducing the inequality gap, it is necessary to seek clarity about how to measure and interpret socioeconomic status.

## 2.8 Aim

It is known that socioeconomic inequalities in cancer exist (Kogevinas et al 1997a). It is also known they are important for lung, head and neck, oesophageal and UADT cancers. However, the relative degree, extent and relationship for subsites by age, sex, SES measure or over time are not well known. Moreover, the pathways and explanations are not well understood.

The overall aim of this thesis is to better understand the burden that is associated with socioeconomic inequalities in the cancer incidence examining both individual and area measures of SES.

## 2.9 Hypotheses

The following hypotheses describe the focus of this thesis.

- The distribution of the burden of cancer incidence is unequal among SES groups and this varies by SES measure, cancer site and over time and is increasing over time. Different approaches to measuring and presenting

SES inequalities will be required to summarise inequalities.

- Certain individual SES factors have a greater role in determining cancer risk than others (e.g. education). Area-based SES factors play a role, but may be less significant than individual SES factors.

- Multiple or compounded low SES factors will confer an increased risk association with cancer incidence.

- Risk behaviours associated with SES will explain a proportion of the SES gradients observed for the selected cancer sites incidence.

- The temporal relationship between SES exposure and cancer diagnosis requires to be reflected in the timing of SES exposure measurement.

These will be tested via the following studies and associated objectives:

Chapter 3

- To undertake a detailed analysis of the Scottish Cancer Registry to investigate socioeconomic inequality by age, sex and tumour subtype/site
- To quantify the relative contribution to all cancer socioeconomic inequalities by tumour subtype/site and differences by sex and age in order to assist in providing explanations for socioeconomic inequalities.
- To rank tumour and subtype contribution to all cancer socioeconomic inequalities by age for each sex for lung and UADT cancers using complex metrics of inequality (Slope Index of Inequality and Relative Index of Inequality).

Chapter 4

- To explore the association of cancer incidence with demographic, social and five individual socioeconomic variables (economic activity, occupational social class,

educational attainment level, car ownership and household tenure) variables through novel data linkage between the Scottish Cancer Registry and the Scottish Longitudinal Study.

- To assess more finely the socioeconomic factors associated with cancer incidence through: i) Examining the consistency of the relationship between area and individual SES measures associated with cancer incidence; ii) Explaining whether any single measure was particularly associated with cancer incidence; iii) Assessing whether the area measure was fully explained by the individual measures; and iv) Exploring whether there were any synergistic effects between the area deprivation measure and each individual SES variable; and v) assessing temporal relationship between the SES measure and cancer incidence.

Chapter 5

- To undertake data linkage between the Scottish Cancer Registry and the Scottish Health Survey to create cohort study designed to investigate multiple SES and behavioural risk factors and their association with cancer risk (all cancer and lung and upper aero-digestive tract cancers together).
- To assess whether behaviour risk factors explain the previously identified socioeconomic magnitude as measured by individual and area SES (and in combination) measures in all cancer (excluding non-melanoma skin cancer) and lung and upper aero-digestive tract incident cancers taken together.

Chapter 6

- Through discussion of the thesis findings collectively and in relation to the existing literature, to contribute to explanations of inequality in cancer incidence and to the evidence-base for developing public health policies aiming to reduce inequalities in cancer incidence and draw conclusions.

# 3 Socioeconomic inequality in Lung and UADT cancer incidence in Scotland: quantification of contribution to all cancer risk and examination by tumour subtype, five-year age group and sex

## 3.1 Introduction

In 2012, worldwide, there were 14.1 million new cancer cases. Estimated age – standard rates (world) per 100,000 populations indicate the United States (USA) (males: 347.0, females: 297.4) followed by the EU (males: 311.3, females: 241.3) had the highest incidence rates (Ervik et al 2016). In the UK, a cancer diagnosis is more common than getting married or having a first baby (Knapton S 2017).

Projections for 2030 indicate that these figures will double. Cancer is increasing at rates faster than the increase in global population. It is becoming more common in high-income but also — and most of all — in middle and low-income countries, absolutely and also relative to other diseases (WCRF/AICR 2007).

In Scotland, the number of new cases of cancer (excluding non-melanoma skin cancer) is predicted to rise by 33% between 2008-12 and 2023-27, mainly as a result of the population growing older (ISD 2015).

Lung and upper aero-digestive tract (UADT) comprising head and neck (larynx, oral cavity and oropharynx) and oesophageal cancers together are the most common worldwide; 21% of global cases were diagnosed in Europe in 2012 (IARC 2008). These cancers show socioeconomic inequalities with greater incidence among lower socioeconomic groups (Hemminki et al 2003; Anderson et al 2008; Conway et al 2008; Conway et al 2010a). Previous research on oesophageal cancer and socioeconomic status (SES) identified increased risk of squamous cell carcinoma (Morgan et al 2007) in lower socioeconomic groups while adenocarcinoma showed no clear association (Brewster et al 2000). Others who studied lung cancer histological subtypes found increasing incidence among lower socioeconomic groups for all subtypes, although the association was less

strong for adenocarcinoma in both sexes (Bennett et al 2008). Case-control and population cohort studies have found increased incidence in lower socioeconomic groups for larynx (Anderson et al 2008), oral cavity (Conway et al 2007; Anderson et al 2008; Conway et al 2008) and oropharynx (Anderson et al 2008) cancers.

Area-based indices of SES are increasingly used worldwide to measure effects of SES on health outcomes (Kogevinas et al 1997a). Based on income, employment, education, housing, health, crime and geographic access data, the Scottish Index of Multiple Deprivation (SIMD) is a small area measure of SES regularly used in Scotland (Leyland et al 2007a). Small area SES indices are more likely to be homogenous with respect to socioeconomic characteristics and more closely describe individual SES (MacIntyre et al 2002; Leyland et al 2007a). Given its area basis, SIMD also provides a surrogate measure of physical environmental SES, another important and recognised deprivation factor associated with health and disease (MacIntyre et al 2002).

Several inequality measures are used to monitor socioeconomic associated health inequalities (Harper et al 2009). The Slope Index of Inequality (SII) and Relative Index of Inequality (RII) capture the effect (direction and magnitude) of the inequality gradient as well as the extent (population deprivation distribution) of absolute and relative SES inequality (Harper et al 2008).

## 3.2 Study aims and objectives

The objective of this study was to undertake a detailed analysis of socioeconomic inequality by age, sex and tumour subtype as this had yet to be fully undertaken. Furthermore, this study quantified the relative contribution to all cancer socioeconomic inequalities by tumour subtypes and differences by sex and age in order to assist in providing explanations for socioeconomic inequalities. Finally, this study explored SII and RII to rank tumour and subtype contribution to socioeconomic inequalities by age for each sex for lung and UADT cancers.

## 3.3 Methods

Cancer incidence data were sourced from the Scottish Cancer Registry (ISD 2010) for all cancer excluding non-melanoma skin cancer, lung and UADT sites for the period 2000 to 2007. Data extracted were: age, sex, cancer diagnosis (ICD-10), year of diagnosis, postcode at diagnosis and morphology (ICD-O-2 from 1999 to 2005 or ICD-O-3 from 2006). Subtype groups reflect anatomical relationship (Junor et al 2010) for head and neck sites and summarisation of the International Agency for Research on Cancer (IARC) defined morphology code groups for lung and oesophageal cancers (IARC, 2009) (Sections 1.5.1.1, 1.5.1.2 and 1.5.1.3). Unspecified or non-specific cancers were grouped together in the 'other' category by anatomical site. Patient's residential postcode was linked to SIMD decile via datazone, a measure of geography with mean population size of 778 and composed of census output areas (Bishop J et al 2004). At the time of this study, two versions of SIMD were available. SIMD 2006 was selected as it was based on the National Records of Scotland (NRS) 2004 mid year population estimates coinciding most closely with the midpoint of the study period (Section 1.3.2.4).

The Scottish population at the start of the period (2000) was sourced from the General Registrar Office (Scotland) to establish for each sex, decile and cancer; age standardised rates per 100,000 population calculated by direct standardisation to the European standard population (IARC 2002) and age-specific rates per 100,000 population.

A linear regression model was used to calculate SII (Harper et al 2009), SII confidence intervals and *P*-values of the age-specific and age standardised rates. The study used statistically significant SII results defined as $P<0.05$ for the all ages analysis and $P<0.01$ for the age-specific analysis to establish RII based on the mean incidence rate of all cancer (Leyland et al 2007b). Where SII was not statistically significant, RII was set at 0.000. All analyses were conducted using SAS version 9.1, (SAS Institute Inc. Cary, NC, USA).

## 3.4 Results

Incident cancers totalled 216,305 and were comprised of 105,040 cases in males and 111,265 cases in females over the period 2000 to 2007. Age-specific rates per 100,000 population increased from 12.5 (five to nine years) to 3737.9 (85+ years) for males and from 9.7 (five to nine years) to 2195.4 (85+ years) for females. This comprised for males: 20,427 lung, 5,746 head and neck and 4,078 oesophageal cancers and for females: 16,847 lung, 2,470 head and neck and 2,456 oesophageal cancers (Figure 3.1 and Tables 3.1, 3.2).

**Figure 3.1 All cancer (excluding non-melanoma skin) age-specific incidence rate per 100,000 population by five-year age and sex, Scotland 2000-07**

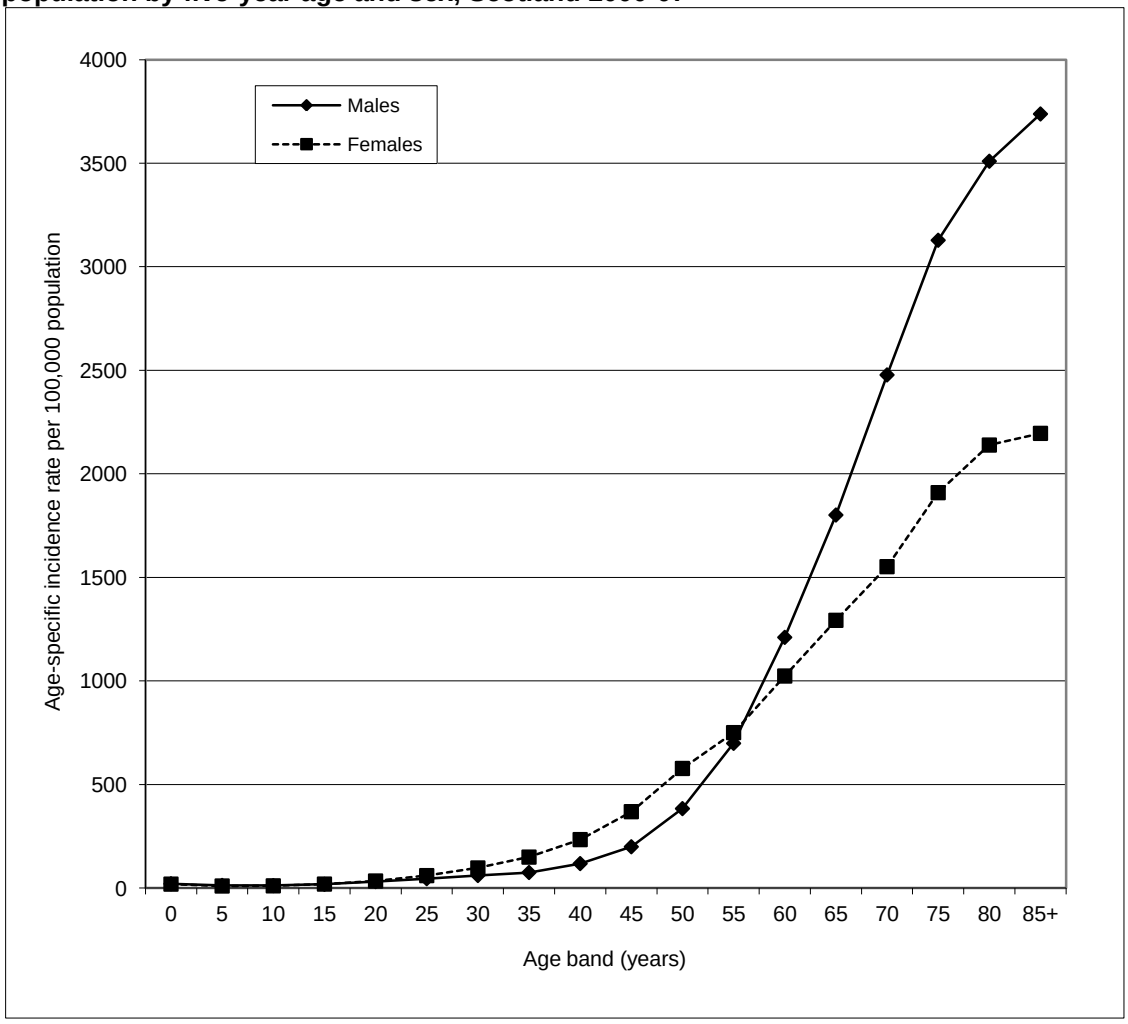


**Table 3.1 Cases, incidence[a], Slope Index of Inequality (95% confidence intervals), *P*-value and Relative Index of Inequality (RII)[b] for lung and upper aero-digestive tract cancer subtypes for all ages for males ranked[c] by RII, Scotland 2000-07**

Males

| SITE | Cases | | Incidence[a] | Slope Index of Inequality | *P*-value | RII | Rank[c] of all cancer[d] RII |
|---|---|---|---|---|---|---|---|
| | n | (%) | | | | | |
| All cancer (exc. NMSC)[d] | 105040 | 100.0% | 362.0 | 132.5 (159.6, 105.4) | <0.001 | 0.366 | 100% |
| Lung | 20427 | 19.4% | 69.1 | 86.1 (94.5, 77.7) | <0.001 | 0.238 | 65% |
| Head and neck | 5746 | 5.5% | 20.7 | 25.5 (30.0, 21.0) | <0.001 | 0.070 | 19% |
| Oesophagus | 4078 | 3.9% | 14.1 | 8.7 (11.4, 5.9) | <0.001 | 0.024 | 7% |
| Lung-other | 6652 | 6.3% | 22.1 | 28.1 (31.5, 24.6) | <0.001 | 0.078 | 21% |
| Lung-squamous cell carcinoma | 5397 | 5.1% | 18.3 | 26.6 (29.2, 24.0) | <0.001 | 0.074 | 20% |
| Lung-small cell carcinoma | 5271 | 5.0% | 18.0 | 21.7 (25, 18.3) | <0.001 | 0.060 | 16% |
| Head and neck-larynx (inc. hypopharynx and piriform sinus) | 2303 | 2.2% | 8.2 | 12.4 (15.1, 9.7) | <0.001 | 0.034 | 9% |
| Lung-adenocarcinoma | 3107 | 3.0% | 10.7 | 9.8 (11.9, 7.6) | <0.001 | 0.027 | 7% |
| Head and neck-oral cavity (including lip) | 1697 | 1.6% | 6.1 | 6.8 (7.9, 5.6) | <0.001 | 0.019 | 5% |
| Oesophagus-squamous cell carcinoma | 1212 | 1.2% | 4.3 | 5.2 (6.4, 4.0) | <0.001 | 0.014 | 4% |
| Head and neck-oropharynx (inc. base of tongue, palate and tonsil) | 1124 | 1.1% | 3.7 | 3.5 (4.4, 2.6) | <0.001 | 0.010 | 3% |
| Head and neck-other | 622 | 0.6% | 2.2 | 1.9 (2.6, 1.2) | <0.001 | 0.005 | 1% |
| Oesophagus-other | 410 | 0.4% | 1.4 | 1.3 (1.6, 1.0) | <0.001 | 0.004 | 1% |
| Oesophagus-adenocarcinoma | 2456 | 2.3% | 8.5 | 2.2 (5.7, -1.4) | 0.193 | 0.000 | 0% |

[a] Age Standardised Incidence Rate per 100,000 population (standardised to the European standard population)

[b] RII defined as Slope Index of Inequality/ incidence rate for all cancer

[c] Rank defined as proportion of all cancer RII

[d] All cancer excluding non-melanoma skin cancer

**Table 3.2 Cases, incidence[a], Slope Index of Inequality (95% confidence intervals), *P*-value and Relative Index of Inequality (RII)[b] for lung and upper aero-digestive tract cancer subtypes for all ages for females ranked[c] by RII, Scotland 2000-07**

Females

| SITE | Cases | | Incidence[a] | Slope Index of Inequality | *P*-value | RII | Rank[c] of all cancer[d] RII |
|---|---|---|---|---|---|---|---|
| | n | (%) | | | | | |
| All cancer (exc. NMSC)[d] | 111265 | 100.0% | 309.3 | 86.4 (103.6, 69.2) | <0.001 | 0.279 | 100% |
| Lung | 16847 | 15.1% | 43.4 | 59.1 (67.2, 51.0) | <0.001 | 0.191 | 68% |
| Head and neck | 2470 | 2.2% | 7.2 | 7.6 (9.0, 6.3) | <0.001 | 0.025 | 9% |
| Oesophagus | 2456 | 2.2% | 5.7 | 3.1 (4.3, 1.9) | <0.001 | 0.010 | 4% |
| Lung-small cell carcinoma | 4904 | 4.4% | 13.7 | 21.3 (24.9, 17.7) | <0.001 | 0.069 | 25% |
| Lung-other | 6205 | 5.6% | 13.9 | 19 (22.8, 15.1) | <0.001 | 0.061 | 22% |
| Lung-squamous cell carcinoma | 2822 | 2.5% | 7.4 | 11.4 (12.5, 10.2) | <0.001 | 0.037 | 13% |
| Lung-adenocarcinoma | 2916 | 2.6% | 8.3 | 7.5 (8.6, 6.3) | <0.001 | 0.024 | 9% |
| Head and neck-larynx (inc. hypopharynx and piriform sinus) | 632 | 0.6% | 1.9 | 3.7 (4.7, 2.7) | <0.001 | 0.012 | 4% |
| Head and neck-oral cavity (including lip) | 982 | 0.9% | 2.7 | 2 (2.4, 1.7) | <0.001 | 0.007 | 2% |
| Oesophagus-squamous cell carcinoma | 1249 | 1.1% | 3.0 | 1.9 (2.7, 1.2) | <0.001 | 0.006 | 2% |
| Head and neck-oropharynx (inc. base of tongue, palate and tonsil) | 461 | 0.4% | 1.2 | 1.0 (1.4, 0.5) | <0.001 | 0.003 | 1% |
| Oesophagus-adenocarcinoma | 395 | 0.4% | 1.2 | 0.6 (1.1, 0.0) | 0.048 | 0.002 | 1% |
| Head and neck-other | 342 | 0.3% | 0.7 | 0.5 (0.7, 0.2) | <0.001 | 0.002 | 1% |
| Oesophagus-other | 342 | 0.3% | 2.0 | 0.7 (1.5, -0.2) | 0.105 | 0.000 | 0% |

[a] Age Standardised Incidence Rate per 100,000 population (standardised to the European standard population)

[b] RII defined as Slope Index of Inequality/ incidence rate for all cancer

[c] Rank defined as proportion of all cancer RII

[d] All cancer excluding non-melanoma skin cancer

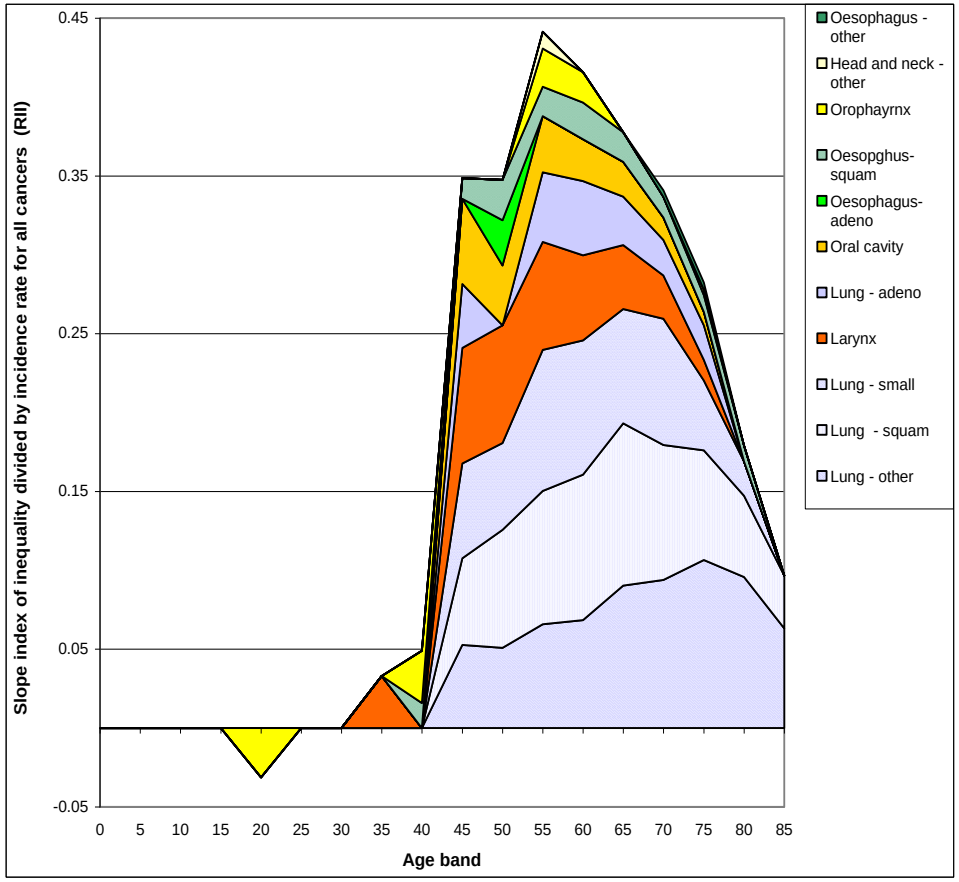
All cancer socioeconomic inequality was greatest among males (RII=0.366) compared to females (RII=0.279) with the combination of lung and UADT contributing to all cancer socioeconomic inequality 91% and 81% respectively. For males, all cancer socioeconomic inequality contribution by site was lung (65%), head and neck (19%) and oesophagus (7%); for females, all cancer socioeconomic inequality contribution was lung (68%), head and neck (9%) and oesophagus (4%) (Tables 3.1, 3.2).

For both sexes lung and UADT subtypes showed significant socioeconomic inequality gradients ($P$<0.001) except oesophageal adenocarcinoma in males (P=0.193); for females, socioeconomic inequality was very small and only borderline significant (RII=0.002, P=0.048) (Tables 3.1, 3.2).

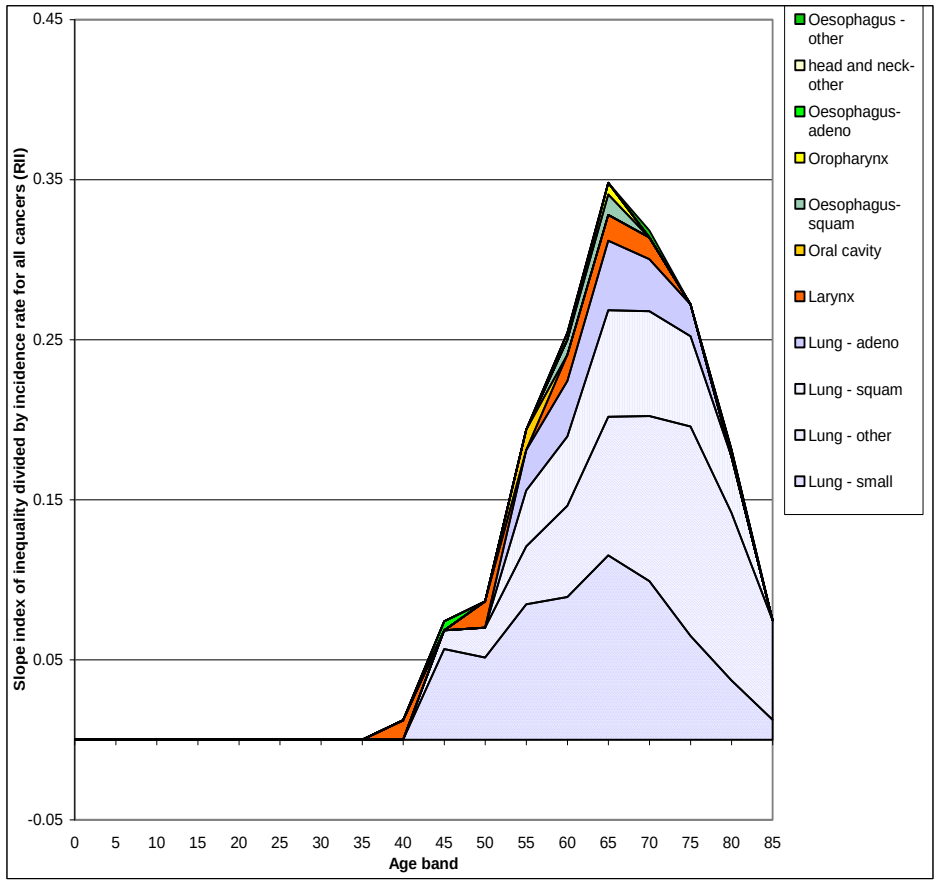
For males, 32.8% of the age-sex-deprivation decile strata were statistically significant at the 99% confidence level (females, 21.7%). Age-specific analysis confirmed that peak lung and UADT socioeconomic inequality for males (RII=0.441) was greater than for females (RII=0.348) and occurs 10 years earlier (55-59 years). For both sexes, the 40-85 age range showed the widest socioeconomic inequality reflecting that RII was never equal to 0 at these ages. For males, oropharynx and larynx cancers show socioeconomic inequality at even earlier ages (20-24 and 35-39 years respectively). For males, socioeconomic inequality abruptly widened at 40-44 years reaching the peak and remaining elevated until 75-79 years, then falling. For females, the socioeconomic inequality peak (65-69 years) was reached more gradually; thereafter socioeconomic inequality rate of decline was similar to males (Figure 3.2). For both sexes, peak incidence occurred at 80-84 years (Figure 3.3).

**Figure 3.2 Socioeconomic inequalities (RII[a]) in lung and aero-digestive tract cancers by age and subtype for each sex**

(a) Males

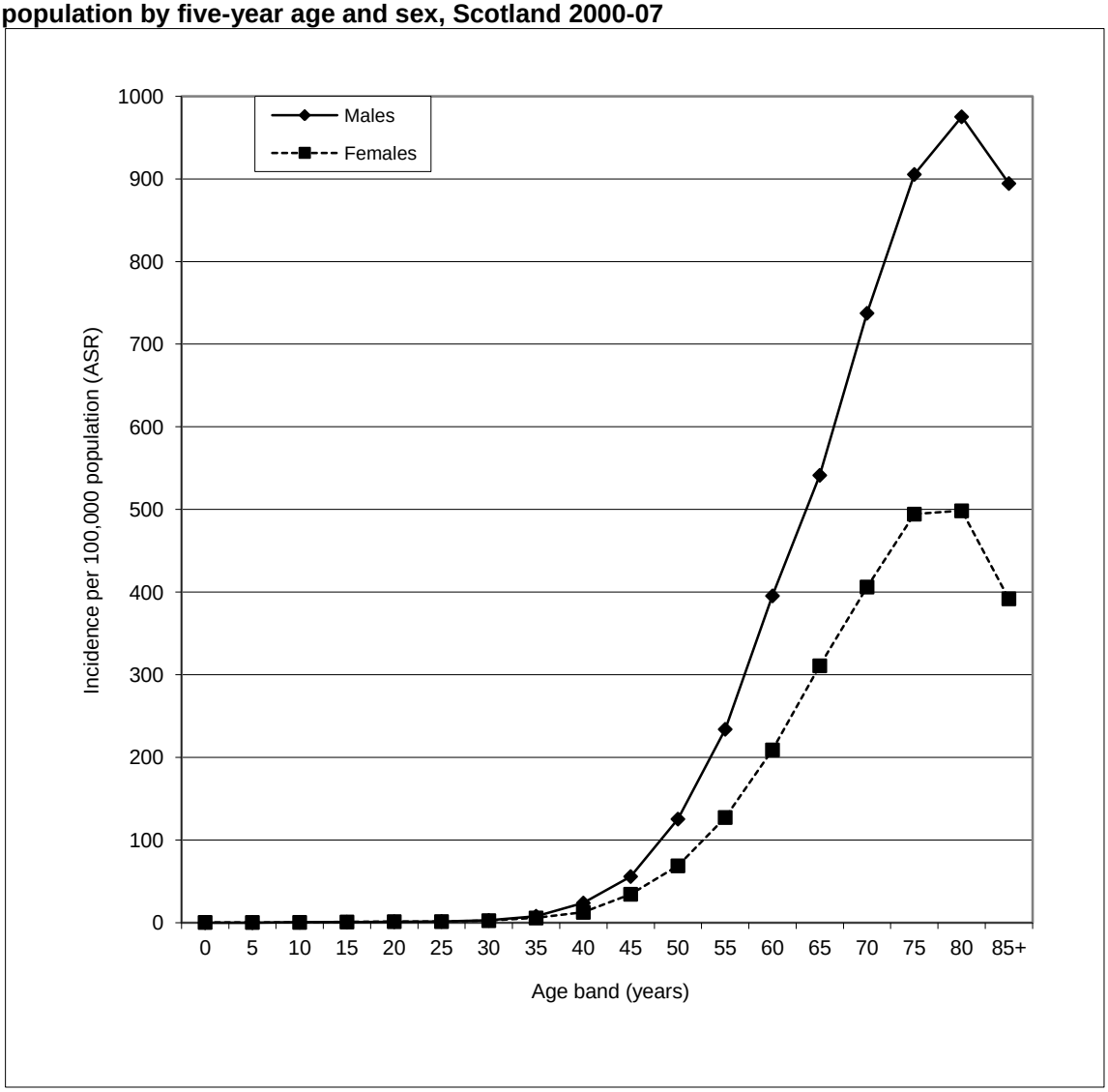



(b) Females




[a] RII or Relative Index of Inequality defined as the Slope Index of Inequality / incidence rate for all cancers

**Figure 3.3 Lung and upper aero-digestive tract age-specific incidence rate per 100,000 population by five-year age and sex, Scotland 2000-07**




The age with widest socioeconomic inequality varied depending on the sex and cancer subtype. With the exception of other lung morphologies, lung cancer inequalities by subtype peaked for females 65-69 years. For males, lung socioeconomic inequality peaked at earlier and differing ages for adenocarcinoma (60-64 years) and small cell carcinoma (55-59 years) but occurred at the same age as females for squamous cell carcinoma (65-69 years) (Figure 3.2).

For head and neck cancers in males; larynx dominated socioeconomic inequality in the 35-79 age range with greatest contribution in males aged 50-54 years old.

Oral cavity showed socioeconomic inequality in males 45-79 years old; within this age range, socioeconomic inequality contribution was greatest for males 45-49 years old. Oropharynx cancer in males was the only cancer demonstrating socioeconomic inequality for the socioeconomically advantaged (20-24 years old) as well as those in low socioeconomic circumstances (40-44 and 55-74 years old). Overall, the picture was similar in females with reduced magnitude and older ages affected; however, all head and neck socioeconomic inequality affected those with low socioeconomic circumstances. In females, larynx dominated socioeconomic inequality at ages 40-74 years old followed by lip and oral cavity at ages 50-59 and 65-69 years old and oropharynx at ages 55-59 and 65-69 years old (Figure 3.2).

For oesophageal cancer in males, adenocarcinoma socioeconomic inequality contribution was the greatest for any age, but was only evident at ages 50-54 years. Although oesophageal squamous cell carcinoma contributed less at any single age, its contribution was spread over much wider age range (40-84 years old). Oesophageal cancer socioeconomic inequality was less evident in females with squamous cell carcinoma dominant and peaking at ages 65-69 years; small but significant socioeconomic inequality for adenocarcinoma featured for females 45-49 years old (Figure 3.2).

For males age 50-54 years, laryngeal cancer and squamous cell lung cancer socioeconomic inequality were equivalent. Thereafter, lung cancer socioeconomic inequality dominated; however, squamous cell carcinoma was greatest from ages 55-69 years while other lung morphologies dominated the oldest age groups. For females, laryngeal cancer was the only cancer presenting socioeconomic inequality at age 40-44 years. Thereafter small cell lung carcinoma dominated until age 69. Like males, other lung morphologies dominated at the older age range (Figure 3.2).

## 3.5 Discussion

All cancer socioeconomic inequality was established as greater for males than females. Lung cancer contributed the majority to all cancer socioeconomic inequality (males 65%, females 68 %), followed by head and neck (males 19%, females 9%) with oesophagus contributing the least (males 7%, females 4%). LUADT cancers together contributed 91% and 81% to all cancer socioeconomic inequality for males and females respectively. Although RII rank differed by sex, lung and larynx subtypes contributed most to all cancer socioeconomic inequality with RII rank for oral cavity, oesophagus–squamous cell and oropharynx following for both sexes. Finally, for males 40-44 years old, lung and UADT socioeconomic inequality increased abruptly and peaked at 55-59 years. For females, lung and UADT socioeconomic inequality increased more gradually, peaking 10 years later. For both sexes, lung and UADT socioeconomic inequality peak occurred at an earlier age than peak incidence.

These findings were consistent with others who reported greater socioeconomic inequalities with greater incidence among lower socioeconomic groups for lung (Hemminki et al 2003; Harper et al 2008; Shack et al 2008; Dalton et al 2008b), head and neck (Hemminki et al 2003; Conway et al 2007; Anderson et al 2008; Conway et al 2008; Conway et al 2010a) and oesophagus (Brewster et al 2000; Hemminki et al 2003; Baastrup et al 2008).

Behavioural factors including alcohol (excluding lung cancer (Spitz MR et al 2006)) and tobacco consumption are recognised as playing an important role in the risk of these cancers. These findings reflected the strong gradient in smoking prevalence across SIMD deciles in Scotland. In most communities facing low socioeconomic circumstances, 2006 smoking prevalence rates were similar to 1970 rates (ScotPHO 2008). Recent self-reported survey data showed association of deprivation and alcohol consumption in Scotland was less clear. However, adults living in areas of low socioeconomic circumstances consumed more alcohol on their heaviest drinking day and were more likely to exceed binge

drinking thresholds than those living in more advantaged areas (Beetson C et al 2011).

Deprivation was associated with all lung cancer subtypes for males and females with incidence greatest for lower socioeconomic groups and a weaker association for adenocarcinoma (Bennett et al 2008). Previous studies reported that smoking was most strongly associated with small cell lung cancer, then squamous cell carcinoma and finally adenocarcinoma (Menvielle et al 2009). This study's findings were consistent with all subtypes showing statistically significant socioeconomic inequality; in females this inequality followed the same smoking association ranking. This reinforced the conclusion that past smoking behaviour differences among socioeconomic groups largely explain socioeconomic inequality in lung cancer risk. However, others have demonstrated although smoking behaviour is a major factor; unexplained socioeconomic inequality risk remained (Menvielle et al 2009).

Lung cancer's dominant contribution to all cancer socioeconomic inequality reflects two factors: the volume of cases and the distribution of those cases amongst the deprivation deciles. However, had inequality of case distribution amongst the deciles not been present, lung cancer would not have ranked in terms of all cancer socioeconomic inequality contribution. The volume of cases is secondary; lung cancer's proportion of all cancer cases is 19.4% for males (15.1% females) while the lung cancer socioeconomic inequality proportion of all cancer socioeconomic inequality is 65% for males (68% females). This point was also demonstrated in laryngeal cancer contribution to all cancer socioeconomic inequality. Despite a smaller volume of cases, laryngeal cancer ranked higher than adenocarcinoma of the lung in terms of contribution to all cancer socioeconomic inequality.

The peak of lung cancer inequalities occurred at a younger age for males than females (males: 55-59 years old; females: 65-69 years old) and is likely to reflect differences between birth cohorts and the sexes in terms of smoking behaviour (e.g. duration of smoking, number and type of cigarettes smoked and initiation

of smoking among women in large numbers occurring years later than men) as well as the long latency period between exposure to tobacco smoke and lung cancer development (Harkness et al 2002). Differences by sex in rank order of lung cancer subtypes (Tables 3.1 and 3.2 male and female respectively) may reflect later uptake by women coinciding with changes in cigarette type which may result in change in histologic distribution of lung cancers. A body of research is building suggesting that adenocarcinoma may be associated with low-tar filtered cigarettes while squamous cell carcinoma is more likely to be associated with high-tar, unfiltered cigarettes (Ito et al 2011). For both males and females other lung cancer morphologies dominate the 70 – 85 year and older ages most likely due to poorer patient health status preventing more invasive diagnostic procedures required to determine tumour histopathology.

This study confirmed that head and neck cancer incidence was associated with low socioeconomic circumstances. Smoking and alcohol behaviours probably explain most but not all of the socioeconomic inequality (Boing et al 2010; Conway et al 2010a). At the anatomical site level, differences in tobacco and alcohol use remain the predominant explanation for socioeconomic inequality in cancers of the mouth, pharynx and larynx (Conway et al 2010b) with occupational exposure also contributing after adjustment for smoking and alcohol consumption (Menvielle et al 2004). The relationship of occupational exposure to socioeconomic inequalities is yet to be fully explored.

The small but significant socioeconomic inequality for oropharyngeal cancer in males affecting the more advantaged (20-24 years old) was most likely a chance finding reflecting small sample size or misclassification of deprivation. Nevertheless, oropharyngeal cancer is the fastest increasing cancer in Scotland (Mehanna H et al 2010) with other countries reporting similar increases in tonsillar cancers in particular (Syrjanen 2004; Ryerson et al 2008). Human papillomavirus (HPV) has been proposed as a possible explanation for the aetiology of head and neck cancers. A recent review reported tonsillar cancer had a significantly higher HPV detection rate than any other head and neck

cancer and HPV 16, a high risk form of the virus, was the most prevalent type (Mammas et al 2011). However detection does not confirm causation.

Consistent with others (Brewster et al 2000; Weiderpass et al 2006) this study demonstrated a small but significant socioeconomic inequality for oesophageal adenocarcinoma in both sexes with wider inequalities in males, evident only after examination by age. The primary aetiological factors for oesophageal cancer are gastro-oesophageal reflux and high body mass index (BMI). Smoking is a weaker factor while use of non-steroidal anti-inflammatory drugs may be protective; Helicobacter pylori bacterial infection is associated with lower risk as are vegetable and fruit intake (Rutegard et al 2010). A case-control study reviewing socioeconomic factors and oesophageal cancer risk identified socioeconomic inequality among low socioeconomic groups for both adenocarcinoma and squamous cell carcinoma even after adjustment for dominant risk factors; although BMI, smoking and reflux symptoms attenuated the excess risk, adjustment for Helicobacter pylori infection did not influence results (Jansson et al 2005). These findings suggest that while BMI and reflux symptoms provide a partial explanation, other factors must explain the remaining observed socioeconomic inequality. In Scotland, a 2003 survey showed an association of increasing prevalence of obesity with increasing low socioeconomic circumstances that was stronger in females than males (Grant I et al 2007). A further study in Scotland demonstrated that dyspepsia and oesophagitis admissions increased seven-fold in the last twenty-five years with the increased dyspepsia rate associated with increasing deprivation (Baron et al 2008). Further exploration of these risk factors in the context of their association with SES and diagnosis of oesophageal cancer is required.

It might be expected that socioeconomic inequality and incidence peaks would coincide at the same age given greater incidence would maximise the opportunity for inequality; however this study's results identified age at diagnosis decreased with increasing deprivation. Previous researchers have hypothesized that low socioeconomic groups may develop cancer earlier than advantaged groups due to an increased rate of biological ageing reflecting

socioeconomic patterning of risk factor exposure (Adams et al 2004). This study identified socioeconomic inequality declined at older ages and at similar rates for both sexes. Scottish low socioeconomic circumstances are distributed evenly across age groups; nevertheless, nearly 60% of deaths among more affluent males occur at ages 75+ years while just 33% occur at the same age among the more socioeconomically disadvantaged (females: 76% and 55% respectively) (Leyland et al 2007b). In Scotland socioeconomically disadvantaged populations are more likely to be diagnosed with cancer, diagnosis is more likely to occur at an earlier age and this diagnosis is more likely to lead to death at an earlier age. Risk behaviour prevalence is more equally distributed among the social groups in the older population; at younger ages, public health messages have more readily resulted in changed behaviour among the more advantaged.

To capture the contribution of a specific site (and single age) to all cancer socioeconomic inequality in terms of gradient direction and magnitude, the traditional RII definition was adapted by using the mean incidence rate for all cancer. The resulting measure could be summed to provide an overall socioeconomic inequality. This was only feasible when the denominator used was consistent across the summed RIIs. This adaptation supported evaluating multiple cancer subtypes and patient features by providing the aggregate and individual subtype contribution simultaneously. However, all cancer socioeconomic inequality was a net measure of both negative and positive inequalities affecting the more advantaged and more socioeconomically disadvantaged respectively. Breast, prostate and cutaneous skin cancers are more likely to be diagnosed among the higher socioeconomic groups. They are also very common cancers. As a result, the overall effect on all cancer inequality RII may have been attenuation. Furthermore, age-specific analysis revealed socioeconomic inequality previously masked. These dynamics may have reduced the level of all cancer socioeconomic inequality, particularly if evaluating inequalities affecting either the more socioeconomically disadvantaged or the more advantaged separately.

Strengths of this study include a deprivation measure based on small geographic areas with small populations. Area measures of deprivation potentially include the important contribution of neighbourhood on influencing deprivation of an individual. This was particularly relevant when deprivation was the exposure variable and not a confounder as in this study. However, SIMD measured low socioeconomic circumstances at one point in time; for cancer the latent period between exposure and development of the disease is often years if not decades. This analysis measured SES at diagnosis, the closest point to exposure available.

To reduce the possibility of chance findings the more rigorous 99% confidence interval definition of statistical significance was applied to the age specific analysis. A potential weakness of this approach is the use of SII and RII as these measures implicitly require a linear relationship between incidence and social group, however previous work with Scottish Cancer Registry data has demonstrated linear relationships for the tumour types reviewed (ISD 2009).

As a population study all registered cases were included; the Scottish Cancer Registry has high case ascertainment with only 0.4% of registrations based on death certificate only. However, aggregate data does not support exploring other characteristics of the cancer or patient attributes including individual SES (education, income and occupation) and behaviours which may further explain the socioeconomic inequality described.

## 3.6 Conclusion

In summary, peak socioeconomic inequality in lung and UADT cancer was greater for males than females, occurred decades earlier and for both sexes preceded peak incidence. As a proportion of all cancer socioeconomic inequality, lung and UADT cancers contributed 91% of socioeconomic inequality in males and 81% in females. Differences in socioeconomic inequality cancer subtype ranked by sex were likely to largely reflect differences between the sexes in risk behaviours which varied by birth cohort and were socioeconomically patterned. These

findings reinforced established preventive strategies targeting younger socioeconomically deprived populations.

# 4 Association between socioeconomic factors and cancer risk: a population cohort study in Scotland (1991-2006)

## 4.1 Introduction

As discussed in more depth in Section 1.2, the association of socioeconomic status (SES) and health is well established and shows a mostly consistent pattern of poorer health with lower SES (Kogevinas et al 1997a; Mackenbach et al 2008). SES is usually measured in routine statistics using an area indicator or in epidemiological studies with a single individual indicator such as educational attainment.

Lung and upper aero-digestive tract (UADT) cancers taken together are the most common cancers in the world compared to the other individual sites; 21% of global cases were diagnosed in Europe in 2008 (IARC, 2008). As reviewed in detail (Sections 2.3 – 2.6), these cancers show stark socioeconomic inequalities with greater incidence among lower socioeconomic groups (Hemminki et al 2003; Anderson et al 2008; Conway et al 2008; Conway et al 2010a). The United Kingdom (UK) has the second highest age standardised incidence rate (ASR) for these cancers among Northern European countries with Scotland ranking the highest in the UK (IARC 2008; National Cancer Intelligence Network 2013). In Scotland, cancer incidence is higher in more deprived areas with the level of inequality stubbornly remaining stable over time (Scottish Government 2013c). Furthermore, as shown in Chapter 3, lung and UADT cancers contributed 90% (males) and 81% (females) to total social inequality in cancer risk in Scotland when measured using the recently developed Scottish Index of Multiple Deprivation, an area measure of social circumstances (Sharpe et al 2012).

While the relative importance of area and individual SES association with cancer mortality has previously been explored (Singh et al 2002; Galobardes et al 2004; Puigpinos et al 2009; Stringhini et al 2017), recently work has begun to focus on cancer incidence – including cervical cancer in South East England (Weiderpass

et al 2006; Currin et al 2009). Such studies have more frequently focused on single SES factors such as occupational social class (Brown et al 1997; Marshall et al 1999; Melchior et al 2005), educational attainment (Baastrup et al 2008), disposable income (Baastrup et al 2008), or area-based SES indicators alone (Currin et al 2009; Menvielle et al 2009)

Only recently, researchers have begun to explore relative impact of area and individual SES measures, such as education level and urban/rural areas in the US National Longitudinal Mortality Study (NLMS) linked to SEER cancer registry data in an analysis (Clegg et al 2009), or education level by European region cancer risk analysis in the European Prospective Investigation into Cancer (EPIC) cohort study (Menvielle et al 2009).

Other social indicators have been investigated, including marital status which has been associated with increased cancer risk (Dalton et al 2008c) and ethnicity (often a proxy for SES in the US) (Ward et al 2004). Moreover, all cancer, lung, colorectal, breast and prostate cancer mortality by country of birth showed higher mortality for all cancer and lung cancer among people born in Scotland (Wild et al 2006). Very few studies have assessed the association with cancer incidence of both area and (multiple) individual SES variables along with marital status and country of birth — Spadea et al (2010) linked the Turin Longitudinal Study and the Piedmont Cancer Registry (1985-1999) (Spadea et al 2010) while Li et al (2015), for lung cancer alone, evaluated area and individual SES variables, marital and immigration status.

## 4.2 Study aims and objectives

This study investigated the association of cancer incidence with one demographic variable (country of birth), one social variable (marital status), one area SES variable through Carstairs deprivation index (McLoone 2000) and five individual socioeconomic variables (economic activity, occupational social class, educational attainment level, car ownership and household tenure). This study aimed to assess more finely the socioeconomic factors associated with cancer

incidence through: i) Examining the consistency of relationship between an area and several individual SES measures associated with cancer incidence; ii) Establishing if any single measure was particularly associated with cancer incidence; iii) Assessing if the area measure was fully explained by the individual measures; and iv) Exploring if there were any synergistic effects between the area deprivation measure and each individual SES variable.

## 4.3 Methods

The 1991 Census data and mortality data from the Scottish Longitudinal Study (SLS) (Boyle et al 2009) managed by National Records of Scotland (NRS) was linked via the National Health Service Central Register (Section 1.4.5.2) to data from the Scottish Cancer Registry (SCR) managed by NHS National Services Scotland (NSS) to develop a cohort. The SLS links data from the Censuses and other administrative sources for a semi-random 5.3% representative sample of the Scottish population. It is the only administrative source of self-reported individual SES factors in Scotland. The five individual categorical socioeconomic variables from the 1991 Census based on the variable's ability to capture SES at various stages of life and the variable's focus on established and different determinants of SES (Krieger et al 2003; Ellaway et al 1998; Galobardes et al 2006a; Galobardes et al 2006b). Country of birth (Scotland, rest of UK and rest of world) and legal marriage status (single, married, widowed and divorced) variables were also included.

Economic activity was grouped into active (full time and part time employees, self-employed, on a government scheme) and inactive (waiting to start a job, unemployed, student status, permanently sick, retired, looking after home or family, or other inactive). Occupational social class was grouped using the Registrar General defined categories: Social Class I (professional, managerial, technical), Social Class II (intermediate), Social Class, IIINM (skilled non-manual), Social Class IIIM (skilled manual), Social Class, IV (partly skilled) and Social Class V (unskilled) (Rose 1995). Education qualifications reflected highest attained degree (first degree and higher, other non-degree, none or missing or under 18

years old). Car ownership was grouped into one or more cars or no car, while household tenure was grouped into owned (owner occupier) or rented (with job, farm or other business, local authority or council, new town corporation, housing association or charitable trust, or private landlord). All variables were measured at 1991 Census, the start of the follow-up period.

Carstairs decile was used as the area-based deprivation measure providing the socioeconomic environmental dimension. Carstairs is measured for Scotland's 1,011 postcode sectors with average population 5,012 and is based on the area level measure of four decennial census variables here taken from the 1991 Census: male unemployment, households with no car, overcrowded households and the percentage of people in higher occupational socioeconomic classes. Unlike other more recent area measures, Carstairs was available for 1991, the start of the study cohort (Bishop J et al 2004).

The study population consisted of 206,830 SLS participants who were 15+ years old present at the 1991 Census and who had been traced at the NHS Central Register so that follow-up data were available. These records linked to individual SCR records recording date of diagnosis and diagnosis code for first primary cancers. There were 2,950 individuals diagnosed with cancer prior to 1 April 1991 and 222 individuals with a missing Carstairs score which were excluded leaving 203,658 cohort members who were followed for up to 16 years from the study start (the 1991 Census date) to the study end date defined as the earliest date of incident cancer, death or the 31 December 2006.

The analysis focused on first primary incident cancers excluding non-melanoma skin cancer (hereafter referred to as all cancer (C00-C96, excluding C44) lung cancer (C33, C34) and upper aero-digestive tract (UADT) cancers (C00–C14, C30-C32 and C15).

Relative risks (RR) and 95% confidence intervals (CI) were computed using Poisson regression models by sex, corrected for under dispersion and offset by person-years of follow-up adjusted for age at start of the cohort in 10 year

categories beginning with 45-54 years (minimally adjusted model). RRs (and 95%CI) for cancer for each variable category were computed by mutually adjusting all the variables for each other (fully adjusted model). Reference categories selected for each variable were: country of birth (Scotland), marital status (married), area SES (least deprived), economic activity (active), educational attainment level (first degree and higher), occupational social class (professional, managerial, technical), car ownership (1 or more car(s)) and household tenure (owned). RRs with 95% CI that did not include the value of 1.0 were regarded as statistically significant. The relationship between area deprivation and educational attainment level was tested in a stratification analysis. Finally, using the multivariate Poisson models, interactions were assessed between area deprivation and each individual socioeconomic variable as well as the difference in RRs between the sexes (females as reference) — with significance established at $P<0.0001$. Age-adjusted sub group analyses were conducted to explore further statistically significant area and individual socioeconomic variable interactions. All analyses were performed using SAS version 9.2 (SAS Institute Inc. USA).

The University of Glasgow Medical Ethics Committee, NSS Privacy Advisory Committee and SLS Research Board approved this study (Appendices 4.1). Analysis was conducted on a secure standalone computer, following strict disclosure protocols. Outputs leaving the safe setting (including this paper) were screened for disclosure by SLS prior to release. Data are publically available to researchers through a similar process of approvals and access.

At the time of conducting this study, the structures and systems described in Section 1.4.5 regarding secure data access portals in Scotland were not entirely in place. For example, researchers using the Scottish Longitudinal Study (SLS) were required to use the physical safe haven facility at Ladywell House, Edinburgh for all of their analysis. The computers used were isolated from outside networks as was the analysis room. Once vetted by the SLS via their application procedures for access and following attendance of the SLS training programme, researchers using the facilities were supervised by SLS staff;

analysis could only take place while an SLS staff member was present and while using the safe haven, communication with others outside the safe haven was restricted. All analysis sessions therefore were required to be organised in advance. Access to analytical programme code and outputs were strictly reviewed by SLS staff and supplied electronically by secure email to researchers only after they had been reviewed for possible disclosure issues.

## 4.4 Results

The cohort consisted of 203,658 individuals (106,819 females and 96,839 males) present in the 1991 Census with an average age of 45.2 and 42.8 years for females and males respectively (Tables 4.1, 4.2). 21,832 first primary cancers were diagnosed during 3.05 million person-years of follow-up (52.3% male, 47.7% female). 3,505 lung cancer cases were diagnosed during 3.12 million person-years of follow-up (52.6% female, 47.4% male) and 1,206 UADT cancer cases during 3.12 million person-years of follow-up (52.6%. female, 47.4% male) (Tables 4.3, 4.4).

When compared to the relevant referent categories and regardless of sex or cancer group, the minimally adjusted models showed elevated cancer risk association for individuals born in Scotland, divorced or widowed, living in more deprived areas, unemployed, with no education, employed in skilled manual, partly skilled or unskilled jobs, with no access to a car or renting a home (Tables 4.3, 4.4). In the fully adjusted models, RRs for each variable were attenuated (some fully) depending on the sex and cancer group; these differences are detailed by each variable below. With the exception of country of birth and single marital status, all statistically significant RRs were greater for males compared to females ($P<0.0001$)

**Table 4.1 Cohort number, proportion, average age and standard deviation (SD) by variable for females, April 1991, Scotland**

| | | Number | (%) | Average Age (SD) |
|---|---|---|---|---|
| Total | | 106,819 | (100.0) | 45.2 (19.4) |
| Country of birth | Scotland | 95,057 | (88.9) | 45.3 (19.5) |
| | Rest UK | 8,710 | (8.2) | 44.8 (18.9) |
| | Rest of World | 3,052 | (2.9) | 44.6 (18.6) |
| Marital Status | Married | 60,425 | (56.6) | 46.0 (14.7) |
| | Divorced | 5,832 | (5.5) | 44.2 (13.0) |
| | Single | 26,610 | (24.9) | 30.2 (18.4) |
| | Widowed | 13,952 | (13.1) | 71.3 (11.1) |
| Carstairs area | Least Deprived 1 | 8,698 | (8.1) | 44.2 (18.1) |
| | 2 | 10,007 | (9.4) | 44.6 (18.9) |
| | 3 | 12,897 | (12.1) | 45.1 (19.0) |
| | 4 | 13,131 | (12.3) | 45.5 (19.3) |
| | 5 | 11,995 | (11.2) | 45.7 (19.7) |
| | 6 | 11,487 | (10.8) | 46.5 (20.0) |
| | 7 | 9,963 | (9.3) | 46.2 (19.7) |
| | 8 | 9,988 | (9.4) | 45.3 (19.9) |
| | 9 | 9,216 | (8.6) | 45.3 (19.7) |
| | Most deprived 10 | 9,437 | (8.8) | 43.4 (19.5) |
| Economic activity | Economically active | 53,249 | (50.6) | 36.8 (12.6) |
| | Economically inactive | 51,958 | (49.4) | 54.9 (20.4) |
| | Under 16 years old | 1,612 | (1.5) | 15.0 (0.0) |
| Education level | First degree and higher | 4,823 | (5) | 38.7 (14.6) |
| | Other non-degree | 8,653 | (8.9) | 43.2 (15.2) |
| | None | 83,421 | (86.1) | 47.2 (19.0) |
| | Under 18 years old or missing[1] | 9,922 | (10.2) | 33.5 (22.5) |
| Occupational social class | I, II Professional, managerial, technical | 18,454 | (17.3) | 40.2 (13.1) |
| | III N Skilled non manual | 25,462 | (5.1) | 37.2 (14.5) |
| | III M Skilled manual | 5,481 | (23.8) | 37.7 (15.0) |
| | IV Partly skilled | 11,579 | (10.8) | 37.6 (14.5) |
| | VI Unskilled | 7,252 | (6.8) | 46.6 (13.7) |
| | No job in last 10 years, under 16 years old or missing[2] | 38,591 | (36.1) | 56.1 (22.3) |
| Car ownership | 1 or more car(s) | 66,422 | (62.2) | 41.2 (16.6) |
| | No cars | 40,397 | (37.8) | 51.8 (21.7) |
| Household tenure | Owned | 59,032 | (55.3) | 43.7 (18.2) |
| | Rented | 47,787 | (44.7) | 47.1 (20.7) |

[1]5.04% of total population was under 18 years old therefore education not recorded; 4.1% of total population education level not stated

[2]For 0.4% of total population occupational social class was not adequately described or not stated, 27.5% of total population was less than 16 years old or held no job in last 10 years.

Source: Scottish Longitudinal Study

**Table.4.2 Cohort number, proportion, average age and standard deviation (SD) by variable for males, April 1991, Scotland**

| | | Number | (%) | Average Age (SD) |
|---|---|---|---|---|
| Total | | 96,839 | (100.0) | 42.8 (17.9) |
| Country of birth | Scotland | 85,802 | (88.6) | 42.7 (18.0) |
| | Rest UK | 8,259 | (8.5) | 43.4 (17.3) |
| | Rest of World | 2,778 | (2.9) | 44.6 (18.0) |
| Marital Status | Married | 59,920 | (61.9) | 48.2 (14.9) |
| | Divorced | 4,163 | (4.3) | 45.1 (12.1) |
| | Single | 29,120 | (30.1) | 27.8 (13.9) |
| | Widowed | 3,636 | (3.8) | 70.6 (11.4) |
| Carstairs area | Least Deprived 1 | 8,411 | (8.7) | 42.8 (17.1) |
| | 2 | 9,504 | (9.8) | 42.3 (17.1) |
| | 3 | 11,906 | (12.3) | 42.9 (17.5) |
| | 4 | 12,344 | (12.7) | 42.8 (18.0) |
| | 5 | 10,854 | (11.2) | 43.0 (18.1) |
| | 6 | 10,068 | (10.4) | 43.4 (18.5) |
| | 7 | 8,872 | (9.2) | 43.5 (18.4) |
| | 8 | 8,964 | (9.3) | 43.1 (18.1) |
| | 9 | 7,995 | (8.3) | 42.6 (18.4) |
| | Most deprived 10 | 7,921 | (8.2) | 41.4 (18.0) |
| Economic activity | Economically active | 70,719 | (73) | 38.1 (13.1) |
| | Economically inactive | 24,452 | (25.3) | 58.5 (20.3) |
| | Under 16 years old | 1,668 | (1.7) | 15.0 ( 0.0) |
| Education level | First degree and higher | 7,066 | ( 8.0) | 41.1 (14.2) |
| | Other non-degree | 6,404 | (7.3) | 43.4 (14.8) |
| | None | 74,757 | (84.7) | 44.6 (17.5) |
| | Under 18 years old or missing[1] | 8,612 | (9.8) | 28.2 (19.6) |
| Occupational social class | I, II Professional, managerial, technical | 23,434 | (24.2) | 43.2 (13.7) |
| | III N Skilled non manual | 9,347 | (9.7) | 38.0 (15.6) |
| | III M Skilled manual | 26,577 | (27.4) | 40.8 (14.9) |
| | IV Partly skilled | 14,359 | (14.8) | 40.6 (16.2) |
| | VI Unskilled | 4,609 | (4.8) | 39.8 (16.1) |
| | No job in last 10 years, under 16 years old or missing[2] | 18,513 | (19.1) | 50.1 (25.7) |
| Car ownership | 1 or more car(s) | 68,702 | (70.9) | 41.2 (16.6) |
| | No cars | 28,137 | (29.1) | 46.7 (20.2) |
| Household tenure | Owned | 56,760 | (58.6) | 42.0 (17.0) |
| | Rented | 40,079 | (41.4) | 43.9 (19.1) |

[1]5.04% of total population was under 18 years old therefore education not recorded; 4.1% of total population education level not stated

[2]For 0.4% of total population occupational social class was not adequately described or not stated, 27.5% of total population was less than 16 years old or held no job in last 10 years.

Source: Scottish Longitudinal Study

**Table 4.3 Minimally adjusted[1] relative risks (RR) and 95% confidence intervals (CI) by cancer, demographic and socioeconomic variable for females, Scotland 1991-2006**

| Female | | All cancer | | | | Lung | | | | UADT | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Level | Number | RR | 95% CI | | Number | RR | 95% CI | | Number | RR | 95% CI | |
| Country of birth | Scotland | 10946 | Reference | | | 1344 | Reference | | | 416 | Reference | | |
| | Rest UK | 869 | **0.87** | **0.83** | **0.92** | 66 | **0.55** | **0.51** | **0.60** | 30 | **0.78** | **0.72** | **0.85** |
| | Rest of World | 254 | **0.73** | **0.66** | **0.81** | 20 | **0.47** | **0.41** | **0.55** | 7 | **0.53** | **0.45** | **0.62** |
| Marital status | Married | 6721 | Reference | | | 767 | Reference | | | 235 | Reference | | |
| | Divorced | 696 | **1.19** | **1.12** | **1.26** | 92 | **1.50** | **1.39** | **1.61** | 34 | **1.83** | **1.70** | **1.98** |
| | Single | 1979 | **0.92** | **0.89** | **0.96** | 94 | **0.64** | **0.59** | **0.69** | 51 | **0.91** | **0.85** | **0.98** |
| | Widowed | 2673 | **1.11** | **1.06** | **1.15** | 477 | **1.50** | **1.44** | **1.57** | 133 | **1.18** | **1.12** | **1.24** |
| Area deprivation | Least Deprived 1 | 899 | Reference | | | 65 | Reference | | | 31 | Reference | | |
| | 2 | 993 | 0.95 | 0.88 | 1.02 | 88 | **1.16** | **1.04** | **1.30** | 34 | **0.88** | **0.79** | **0.97** |
| | 3 | 1359 | 0.99 | 0.92 | 1.05 | 119 | **1.17** | **1.06** | **1.30** | 49 | 0.97 | 0.88 | 1.07 |
| | 4 | 1502 | 1.07 | 1.00 | 1.14 | 155 | **1.49** | **1.35** | **1.64** | 46 | 0.90 | 0.82 | 1.00 |
| | 5 | 1324 | 1.02 | 0.95 | 1.09 | 158 | **1.63** | **1.47** | **1.79** | 37 | **0.75** | **0.67** | **0.83** |
| | 6 | 1347 | 1.05 | 0.98 | 1.12 | 151 | **1.54** | **1.40** | **1.71** | 56 | **1.13** | **1.03** | **1.24** |
| | 7 | 1217 | **1.09** | **1.02** | **1.17** | 147 | **1.72** | **1.55** | **1.90** | 48 | **1.18** | **1.07** | **1.30** |
| | 8 | 1204 | **1.11** | **1.04** | **1.19** | 155 | **1.91** | **1.73** | **2.11** | 49 | **1.21** | **1.10** | **1.34** |
| | 9 | 1100 | **1.10** | **1.02** | **1.18** | 195 | **2.59** | **2.36** | **2.85** | 52 | **1.37** | **1.24** | **1.50** |
| | Most deprived 10 | 1124 | **1.15** | **1.07** | **1.23** | 197 | **2.80** | **2.55** | **3.09** | 51 | **1.45** | **1.32** | **1.60** |
| Economic activity | Economically active | 4126 | Reference | | | 310 | Reference | | | 109 | Reference | | |
| | Economically inactive | 7864 | **1.07** | **1.03** | **1.11** | 1120 | **1.47** | **1.40** | **1.54** | 344 | **1.31** | **1.24** | **1.38** |
| | Under 16 years old | 79 | **0.82** | **0.69** | **0.98** | 0 | 0.00 | 0.00 | 0.00 | 0 | 0.00 | 0.00 | 0.00 |
| Education level | First degree and higher | 386 | Reference | | | 12 | Reference | | | 7 | Reference | | |
| | Other non-degree | 806 | 0.99 | 0.90 | 1.09 | 56 | **1.78** | **1.44** | **2.20** | 25 | **1.52** | **1.26** | **1.82** |
| | None | 9979 | **1.10** | **1.02** | **1.19** | 1280 | **3.24** | **2.67** | **3.94** | 394 | **1.88** | **1.59** | **2.22** |
| | Under 18 years old or missing[2] | 898 | 1.02 | 0.93 | 1.12 | 82 | **2.69** | **2.18** | **3.31** | 27 | **1.53** | **1.27** | **1.84** |
| Occupational social class | I, II Professional, managerial, technical | 1638 | Reference | | | 120 | Reference | | | 41 | Reference | | |
| | IIIa N Skilled non manual | 2250 | 1.03 | 0.98 | 1.09 | 157 | **1.04** | **0.95** | **1.12** | 63 | **1.16** | **1.07** | **1.26** |
| | IIIb M Skilled manual | 525 | **1.09** | **1.01** | **1.18** | 65 | **1.88** | **1.69** | **2.08** | 21 | **1.68** | **1.50** | **1.89** |
| | IV Partly skilled | 1137 | **1.14** | **1.07** | **1.21** | 117 | **1.66** | **1.52** | **1.81** | 38 | **1.57** | **1.43** | **1.72** |
| | V Unskilled | 926 | **1.14** | **1.07** | **1.22** | 173 | **2.28** | **2.11** | **2.47** | 37 | **1.48** | **1.35** | **1.63** |
| | No job in last 10 years, under 16 years old or missing[3] | 5593 | **1.07** | **1.02** | **1.12** | 798 | **1.66** | **1.55** | **1.78** | 253 | **1.43** | **1.32** | **1.54** |
| Car ownership | 1 or more car(s) | 6427 | Reference | | | 560 | Reference | | | 195 | Reference | | |
| | No car | 5642 | **1.15** | **1.11** | **1.18** | 870 | **1.82** | **1.75** | **1.89** | 258 | **1.45** | **1.39** | **1.52** |
| Housing tenure | Owned | 5987 | Reference | | | 515 | Reference | | | 200 | Reference | | |
| | Rented | 6082 | **1.14** | **1.11** | **1.18** | 915 | **1.90** | **1.83** | **1.97** | 253 | **1.31** | **1.25** | **1.36** |

[1]Minimally adjusted model is adjusted for age only.
[2]4.6% of total population was under 18 years old therefore education not recorded; 4.6% of total population education level not stated.
[3]0.4% of total population occupational social class was not adequately described or not stated, 35.7% of total population was less than 16 years old or held no job in last 10 years.
Relative Risk (RR), 95% CI, estimated using age at diagnosis adjusted Poisson regression models corrected for under dispersion and offset by person-years of follow-up.
**Bold and Red** indicates CI did not include 1.0 and increased risk; **Bold and Green** indicates CI level did not include 1.0 and reduced risk.
Source: Scottish Longitudinal Study

**Table 4.4 Minimally adjusted[1] relative risks (RR) and 95% confidence intervals (CI) by cancer, demographic and socioeconomic variable for males, Scotland 1991-2006**

| Male | | All cancer | | | | Lung | | | | UADT | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Level | Number | RR | 95% CI | | Number | RR | 95% CI | | Number | RR | 95% CI | |
| Country of birth | Scotland | 8765 | Reference | | | 1894 | Reference | | | 698 | Reference | | |
| | Rest UK | 731 | **0.87** | **0.82** | **0.91** | 125 | **0.68** | **0.63** | **0.73** | 47 | **0.67** | **0.62** | **0.73** |
| | Rest of World | 267 | **0.86** | **0.79** | **0.94** | 56 | **0.84** | **0.75** | **0.93** | 8 | **0.36** | **0.30** | **0.44** |
| Marital status | Married | 7452 | Reference | | | 1558 | Reference | | | 552 | Reference | | |
| | Divorced | 449 | **1.11** | **1.04** | **1.19** | 123 | **1.52** | **1.41** | **1.63** | 55 | **1.65** | **1.52** | **1.78** |
| | Single | 941 | **0.71** | **0.67** | **0.74** | 140 | **0.67** | **0.62** | **0.72** | 92 | **0.87** | **0.81** | **0.93** |
| | Widowed | 921 | **1.01** | **0.96** | **1.07** | 254 | **1.30** | **1.23** | **1.37** | 54 | 1.14 | 1.05 | 1.23 |
| Area deprivation | Least Deprived 1 | 790 | Reference | | | 92 | Reference | | | 46 | Reference | | |
| | 2 | 819 | 0.95 | 0.89 | 1.02 | 130 | **1.31** | **1.18** | **1.46** | 46 | 0.93 | 0.83 | 1.04 |
| | 3 | 1124 | 1.01 | 0.94 | 1.07 | 187 | **1.44** | **1.30** | **1.59** | 90 | **1.45** | **1.31** | **1.60** |
| | 4 | 1183 | 0.99 | 0.93 | 1.06 | 208 | **1.50** | **1.36** | **1.66** | 93 | **1.45** | **1.31** | **1.60** |
| | 5 | 1116 | 1.05 | 0.98 | 1.12 | 229 | **1.84** | **1.67** | **2.03** | 87 | **1.58** | **1.43** | **1.75** |
| | 6 | 1043 | 1.02 | 0.96 | 1.09 | 226 | **1.89** | **1.71** | **2.08** | 71 | **1.29** | **1.16** | **1.43** |
| | 7 | 1002 | **1.12** | **1.05** | **1.20** | 237 | **2.26** | **2.05** | **2.49** | 88 | **1.88** | **1.70** | **2.08** |
| | 8 | 974 | **1.09** | **1.02** | **1.17** | 238 | **2.27** | **2.06** | **2.51** | 77 | **1.63** | **1.47** | **1.80** |
| | 9 | 874 | **1.11** | **1.04** | **1.19** | 249 | **2.71** | **2.46** | **2.98** | 86 | **2.05** | **1.85** | **2.27** |
| | Most deprived 10 | 838 | **1.16** | **1.08** | **1.24** | 279 | **3.35** | **3.05** | **3.68** | 69 | **1.72** | **1.55** | **1.91** |
| Economic activity | Economically active | 4069 | Reference | | | 671 | Reference | | | 335 | Reference | | |
| | Economically inactive | 5685 | **1.14** | **1.10** | **1.18** | 1404 | **1.68** | **1.61** | **1.76** | 418 | **1.67** | **1.59** | **1.75** |
| | Under 16 years old | 9 | **0.27** | **0.17** | **0.42** | 0 | **0.00** | **0.00** | **0.00** | 0 | **0.00** | **0.00** | **0.00** |
| Education level | First degree and higher | 481 | Reference | | | 39 | Reference | | | 26 | Reference | | |
| | Other non-degree | 534 | 1.01 | 0.92 | 1.10 | 59 | **1.32** | **1.12** | **1.55** | 36 | **1.17** | **1.02** | **1.35** |
| | None | 8298 | **1.17** | **1.09** | **1.24** | 1861 | **3.05** | **2.68** | **3.47** | 660 | **1.82** | **1.63** | **2.03** |
| | Under 18 years old or missing[2] | 450 | 0.95 | 0.87 | 1.04 | 116 | **3.24** | **2.80** | **3.75** | 31 | **1.40** | **1.21** | **1.62** |
| Occupational social class | I, II Professional, managerial, technical | 2001 | Reference | | | 287 | Reference | | | 120 | Reference | | |
| | IIIa N Skilled non manual | 696 | 1.06 | 0.99 | 1.12 | 109 | **1.19** | **1.09** | **1.30** | 52 | **1.42** | **1.30** | **1.56** |
| | IIIb M Skilled manual | 2384 | **1.13** | **1.08** | **1.17** | 541 | **1.81** | **1.71** | **1.92** | 227 | **1.82** | **1.71** | **1.93** |
| | IV Partly skilled | 1394 | **1.14** | **1.09** | **1.20** | 291 | **1.67** | **1.56** | **1.78** | 119 | **1.68** | **1.56** | **1.80** |
| | V Unskilled | 421 | **1.14** | **1.06** | **1.23** | 97 | **1.84** | **1.68** | **2.02** | 49 | **2.27** | **2.07** | **2.49** |
| | No job in last 10 years, under 16 years old or missing[3] | 2867 | **1.11** | **1.06** | **1.16** | 750 | **2.20** | **2.07** | **2.33** | 186 | **1.84** | **1.72** | **1.98** |
| Car ownership | 1 or more car(s) | 6150 | Reference | | | 1073 | Reference | | | 430 | Reference | | |
| | No car | 3613 | **1.06** | **1.03** | **1.09** | 1002 | **1.68** | **1.62** | **1.74** | 323 | **1.67** | **1.60** | **1.74** |
| Housing tenure | Owned | 5199 | Reference | | | 847 | Reference | | | 360 | Reference | | |
| | Rented | 4564 | **1.08** | **1.05** | **1.11** | 1228 | **1.76** | **1.70** | **1.82** | 393 | **1.50** | **1.44** | **1.56** |

[1]Minimally adjusted model is adjusted for age only.
[2]5.4% of total population was under 18 years old therefore education not recorded; 3.5% of total population education level not recorded; 3.5% of total population education level not stated.
[3]0.5% of total population occupational social class was not adequately described or not stated, 18.5% of total population was less than 16 years old or held no job in last 10 years.
Relative Risk (RR), 95% CI, estimated using age at diagnosis adjusted Poisson regression models corrected for under dispersion and offset by person-years of follow-up.
**Bold and Red** indicates CI did not include 1.0 and increased risk; **Bold and Green** indicates CI level did not include 1.0 and reduced risk.
Source: Scottish Longitudinal Study

For both sexes and each cancer group, being born outwith Scotland was associated with reduced risk of cancer compared to being born in Scotland. The only exception was lung cancer risk for males (RR 0.90, 95% CI 0.81, 1.00) (Tables 4.5, 4.6). Regardless of cancer group or sex, being single was associated with reduced cancer risk compared to being married. For females, being divorced or widowed was associated with increased cancer risk compared to the reference regardless of cancer group. For males being divorced was associated with increased risk for lung and UADT cancer while being widowed was associated with increased lung cancer risk only (Tables 4.5, 4.6).

Regardless of sex, all cancer risk was not associated with area deprivation. For females, lung cancer RRs were more variable among those from more affluent area deprivation deciles, but showed clear increased risk association for the three most deprived deciles. For males and compared to females, lung cancer RRs for area deprivation were more pronounced showing clear increasing gradient of elevated risk for all area deprivation deciles. For females, area deprivation was associated with reduced UADT cancer for the more affluent deciles while the 95% CI for more deprived deciles included 1.0. For males and UADT cancer, RRs 95% CIs were generally greater than 1.0 suggesting association with stronger increased risk compared to females, but were more variable for the more deprived area deciles (Tables 4.5, 4.6).

Regardless of sex or cancer group, increased cancer risk was associated with inactive economic status. For males, UADT cancer risk (RR 1.45, 95% CI 1.37, 1.53) was strongest followed by lung and then all cancer. For females the cancer group order starting with the highest risk was lung cancer (RR 1.29, 95% CI 1.22, 1.36), UADT then all cancer. For both males and females, education level was not associated with all cancer risk. Regardless of sex, no education or holding a non-degree qualification was associated with increased lung cancer risk compared to holding a degree. For females, elevated UADT cancer risk was also associated with these categories, but only associated with no education for males (Tables 4.5, 4.6).

**Table 4.5 Fully adjusted[1] relative risks (RR) and 95% confidence intervals (CI) by cancer, demographic and socioeconomic variable for females, Scotland 1991-2006**

| Female | | All cancer | | | | Lung | | | | UADT | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Level | Number | RR | 95% CI | | Number | RR | 95% CI | | Number | RR | 95% CI | |
| Country of birth | Scotland | 10946 | | Reference | | 1344 | | Reference | | 416 | | Reference | |
| | Rest UK | 869 | **0.90** | **0.85** | **0.95** | 66 | **0.66** | **0.61** | **0.72** | 30 | **0.86** | **0.79** | **0.93** |
| | Rest of World | 254 | **0.74** | **0.67** | **0.82** | 20 | **0.52** | **0.45** | **0.61** | 7 | **0.54** | **0.46** | **0.64** |
| Marital status | Married | 6721 | | Reference | | 767 | | Reference | | 235 | | Reference | |
| | Divorced | 696 | **1.12** | **1.05** | **1.19** | 92 | **1.17** | **1.09** | **1.26** | 34 | **1.55** | **1.43** | **1.68** |
| | Single | 1979 | **0.91** | **0.87** | **0.95** | 94 | **0.60** | **0.55** | **0.64** | 51 | **0.88** | **0.82** | **0.95** |
| | Widowed | 2673 | **1.07** | **1.03** | **1.12** | 477 | **1.29** | **1.23** | **1.35** | 133 | **1.08** | **1.03** | **1.14** |
| Area deprivation | Least Deprived 1 | 899 | | Reference | | 65 | | Reference | | 31 | | Reference | |
| | 2 | 993 | 0.93 | 0.87 | 1.00 | 88 | 1.06 | 0.95 | 1.18 | 34 | **0.84** | **0.75** | **0.93** |
| | 3 | 1359 | 0.95 | 0.89 | 1.02 | 119 | 0.97 | 0.87 | 1.07 | 49 | **0.89** | **0.81** | **0.99** |
| | 4 | 1502 | 1.01 | 0.95 | 1.08 | 155 | **1.15** | **1.04** | **1.27** | 46 | **0.81** | **0.73** | **0.89** |
| | 5 | 1324 | 0.95 | 0.89 | 1.02 | 158 | **1.17** | **1.06** | **1.29** | 37 | **0.64** | **0.58** | **0.72** |
| | 6 | 1347 | 0.97 | 0.91 | 1.04 | 151 | 1.05 | 0.95 | 1.17 | 56 | 0.95 | 0.86 | 1.05 |
| | 7 | 1217 | 0.99 | 0.93 | 1.07 | 147 | 1.11 | 1.00 | 1.23 | 48 | 0.97 | 0.88 | 1.07 |
| | 8 | 1204 | 1.00 | 0.93 | 1.07 | 155 | **1.17** | **1.05** | **1.29** | 49 | 0.97 | 0.87 | 1.07 |
| | 9 | 1100 | 0.97 | 0.90 | 1.05 | 195 | **1.52** | **1.37** | **1.68** | 52 | 1.07 | 0.96 | 1.18 |
| | Most deprived 10 | 1124 | 1.00 | 0.93 | 1.08 | 197 | **1.53** | **1.38** | **1.69** | 51 | 1.09 | 0.98 | 1.21 |
| Economic Activity | Economically active | 4126 | | Reference | | 310 | | Reference | | 109 | | Reference | |
| | Economically inactive | 7864 | **1.06** | **1.02** | **1.11** | 1120 | **1.29** | **1.22** | **1.36** | 344 | **1.20** | **1.12** | **1.28** |
| | Under 16 years old | 79 | 0.93 | 0.77 | 1.12 | 0 | 0.00 | 0.00 | 0.00 | 0 | 0.00 | 0.00 | 0.00 |
| Education Level | First degree and higher | 386 | | Reference | | 12 | | Reference | | 7 | | Reference | |
| | Other non-degree | 806 | 0.96 | 0.88 | 1.06 | 56 | **1.57** | **1.27** | **1.94** | 25 | **1.46** | **1.21** | **1.75** |
| | None | 9979 | 0.99 | 0.91 | 1.08 | 1280 | **1.94** | **1.60** | **2.37** | 394 | **1.42** | **1.20** | **1.69** |
| | Under 18 years old or missing[2] | 898 | 0.96 | 0.87 | 1.06 | 82 | **1.66** | **1.35** | **2.05** | 27 | 1.20 | 1.00 | 1.46 |
| Occupational social class | I, II Professional, managerial, technical | 1638 | | Reference | | 120 | | Reference | | 41 | | Reference | |
| | IIIa N Skilled non manual | 2250 | 1.00 | 0.95 | 1.06 | 157 | **0.83** | **0.76** | **0.90** | 63 | 1.07 | 0.98 | 1.18 |
| | IIIb M Skilled manual | 525 | 1.03 | 0.95 | 1.12 | 65 | **1.27** | **1.14** | **1.41** | 21 | **1.45** | **1.29** | **1.64** |
| | IV Partly skilled | 1137 | 1.06 | 0.99 | 1.13 | 117 | 1.08 | 0.98 | 1.18 | 38 | **1.33** | **1.20** | **1.48** |
| | V Unskilled | 926 | 1.04 | 0.97 | 1.11 | 173 | **1.36** | **1.25** | **1.48** | 37 | **1.22** | **1.10** | **1.36** |
| | No job in last 10 years, under 16 years old or missing[3] | 5593 | 0.98 | 0.92 | 1.04 | 798 | 0.99 | 0.91 | 1.07 | 253 | **1.15** | **1.05** | **1.26** |
| Car ownership | 1 or more car(s) | 6427 | | Reference | | 560 | | Reference | | 195 | | Reference | |
| | No car | 5642 | **1.07** | **1.03** | **1.11** | 870 | **1.27** | **1.21** | **1.33** | 258 | **1.23** | **1.16** | **1.29** |
| Housing tenure | Owned | 5987 | | Reference | | 515 | | Reference | | 200 | | Reference | |
| | Rented | 6082 | **1.08** | **1.04** | **1.11** | 915 | **1.34** | **1.28** | **1.40** | 253 | 1.02 | 0.97 | 1.07 |

[1]Fully adjusted model is adjusted for age and mutually adjusting all the variables for each other.
[2]4.6% of total population was under 18 years old therefore education not recorded; 4.6% of total population education level not stated.
[3]0.4% of total population occupational social class was not adequately described or not stated, 35.7% of total population was less than 16 years old or held no job in last 10 years.
Relative Risk (RR), 95% CI, estimated using age at diagnosis adjusted Poisson regression models corrected for under dispersion and offset by person-years of follow-up.
**Bold and Red** indicates CI did not include 1.0 and increased risk; **Bold and Green** indicates CI level did not include 1.0 and reduced risk.
Source: Scottish Longitudinal Study

**Table 4.6 Fully adjusted[1] relative risks (RR) and confidence intervals (CI) by cancer, demographic and socioeconomic variable for males, Scotland 1991-2006**

| Male | | All cancer | | | | Lung | | | | UADT | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Level | Number | RR | 95% CI | | Number | RR | 95% CI | | Number | RR | 95% CI | |
| Country of birth | Scotland | 8765 | | Reference | | 1894 | | Reference | | 698 | | Reference | |
| | Rest UK | 731 | **0.88** | **0.84** | **0.93** | 125 | **0.85** | **0.79** | **0.92** | 47 | **0.78** | **0.72** | **0.85** |
| | Rest of World | 267 | **0.87** | **0.80** | **0.94** | 56 | 0.90 | 0.81 | 1.00 | 8 | **0.39** | **0.32** | **0.47** |
| Marital status | Married | 7452 | | Reference | | 1558 | | Reference | | 552 | | Reference | |
| | Divorced | 449 | 1.05 | 0.98 | 1.13 | 123 | **1.15** | **1.06** | **1.24** | 55 | **1.28** | **1.18** | **1.39** |
| | Single | 941 | **0.69** | **0.66** | **0.73** | 140 | **0.54** | **0.50** | **0.58** | 92 | **0.72** | **0.67** | **0.77** |
| | Widowed | 921 | 0.98 | 0.94 | 1.04 | 254 | **1.09** | **1.03** | **1.16** | 54 | 0.99 | 0.91 | 1.07 |
| Area deprivation | Least Deprived 1 | 790 | | Reference | | 92 | | Reference | | 46 | | Reference | |
| | 2 | 819 | 0.93 | 0.87 | 1.00 | 130 | **1.17** | **1.05** | **1.31** | 46 | **0.85** | **0.76** | **0.95** |
| | 3 | 1124 | 0.98 | 0.92 | 1.04 | 187 | **1.18** | **1.06** | **1.30** | 90 | **1.22** | **1.10** | **1.35** |
| | 4 | 1183 | 0.96 | 0.90 | 1.02 | 208 | **1.16** | **1.05** | **1.28** | 93 | **1.15** | **1.04** | **1.28** |
| | 5 | 1116 | 1.00 | 0.94 | 1.07 | 229 | **1.35** | **1.22** | **1.49** | 87 | **1.18** | **1.07** | **1.31** |
| | 6 | 1043 | 0.97 | 0.91 | 1.04 | 226 | **1.31** | **1.18** | **1.45** | 71 | 0.90 | 0.81 | 1.01 |
| | 7 | 1002 | 1.05 | 0.98 | 1.13 | 237 | **1.48** | **1.34** | **1.63** | 88 | **1.24** | **1.12** | **1.38** |
| | 8 | 974 | 1.02 | 0.95 | 1.10 | 238 | **1.45** | **1.31** | **1.60** | 77 | 1.05 | 0.94 | 1.17 |
| | 9 | 874 | 1.04 | 0.97 | 1.12 | 249 | **1.65** | **1.49** | **1.83** | 86 | **1.26** | **1.13** | **1.40** |
| | Most deprived 10 | 838 | 1.08 | 1.00 | 1.16 | 279 | **1.89** | **1.71** | **2.10** | 69 | 0.97 | 0.86 | 1.08 |
| Economic Activity | Economically active | 4069 | | Reference | | 671 | | Reference | | 335 | | Reference | |
| | Economically inactive | 5685 | **1.14** | **1.10** | **1.19** | 1404 | **1.28** | **1.22** | **1.35** | 418 | **1.45** | **1.37** | **1.53** |
| | Under 16 years old | 9 | **0.38** | **0.24** | **0.60** | 0 | 0.00 | 0.00 | 0.00 | 0 | 0.00 | 0.00 | 0.00 |
| Education Level | First degree and higher | 481 | | Reference | | 39 | | Reference | | 26 | | Reference | |
| | Other non-degree | 534 | 0.98 | 0.90 | 1.06 | 59 | **1.24** | **1.05** | **1.45** | 36 | 1.10 | 0.95 | 1.26 |
| | None | 8298 | 1.06 | 0.99 | 1.13 | 1861 | **1.95** | **1.70** | **2.22** | 660 | **1.14** | **1.01** | **1.28** |
| | Under 18 years old or missing[2] | 450 | 0.91 | 0.83 | 1.00 | 116 | **1.95** | **1.68** | **2.27** | 31 | 0.92 | 0.79 | 1.07 |
| Occupational social class | I, II Professional, managerial, technical | 2001 | | Reference | | 287 | | Reference | | 120 | | Reference | |
| | IIIa N Skilled non manual | 696 | 1.03 | 0.97 | 1.09 | 109 | 0.93 | 0.85 | 1.02 | 52 | **1.28** | **1.17** | **1.40** |
| | IIIb M Skilled manual | 2384 | **1.06** | **1.01** | **1.11** | 541 | **1.19** | **1.12** | **1.27** | 227 | **1.47** | **1.37** | **1.58** |
| | IV Partly skilled | 1394 | **1.09** | **1.03** | **1.14** | 291 | 1.07 | 1.00 | 1.15 | 119 | **1.30** | **1.21** | **1.41** |
| | V Unskilled | 421 | 1.08 | 1.00 | 1.17 | 97 | 1.10 | 1.00 | 1.21 | 49 | **1.68** | **1.52** | **1.85** |
| | No job in last 10 years, under 16 years old or missing[3] | 2867 | 1.04 | 0.99 | 1.10 | 750 | **1.21** | **1.13** | **1.29** | 186 | 1.23 | 1.14 | 1.34 |
| Car ownership | 1 or more car(s) | 6150 | | Reference | | 1073 | | Reference | | 430 | | Reference | |
| | No car | 3613 | 1.01 | 0.97 | 1.04 | 1002 | **1.17** | **1.12** | **1.22** | 323 | **1.34** | **1.27** | **1.41** |
| Housing tenure | Owned | 5199 | | Reference | | 847 | | Reference | | 360 | | Reference | |
| | Rented | 4564 | 1.01 | 0.97 | 1.04 | 1228 | **1.23** | **1.17** | **1.28** | 393 | **1.10** | **1.05** | **1.15** |

[1]Fully adjusted model is adjusted for age and mutually adjusting all the variables for each other.
[2]5.4% of total population was under 18 years old therefore education not recorded; 3.5% of total population education level not stated.
[3]0.5% of total population occupational social class was not adequately described or not stated, 18.5% of total population was less than 16 years old or held no job in last 10 years.
Relative Risk (RR), 95% CI, estimated using age at diagnosis adjusted Poisson regression models corrected for under dispersion and offset by person-years of follow-up.
**Bold and Red** indicates CI did not include 1.0 and increased risk; **Bold and Green** indicates CI level did not include 1.0 and reduced risk.
Source: Scottish Longitudinal Study

For UADT cancer risk and compared to the professional, managerial and technical reference, most occupational social class categories were associated with increased RRs for both males and females. Occupational social class associations with lung cancer risk were very limited (males) or variable (females) while associations with all cancer risk were limited (males) or did not exist (females). Having no access to a car was associated with increased risk compared to owning a car regardless of cancer group and sex with the exception of all cancer risk in males. Renting a home was associated with increased lung cancer risk compared to owning a home for both sexes. Likewise elevated UADT cancer risk was associated with home rental for males, but not females while elevated all cancer risk was associated with home rental for females but not males (Tables 4.5, 4.6).

For males, highest qualification (lung), social class (all cancer, lung), car ownership (lung, UADT) and housing tenure (lung, UADT) presented statistically significant interactions with area, while for females, social class (lung), housing tenure (lung, UADT) and car ownership (UADT) interactions with area were statistically significant ($P<0.0001$). Exploratory sub group analysis of the statistically significant interactions uncovered no discernable trends as even a single cross-product category can trigger significance.

Regardless of sex and cancer group, elevated risk was associated with no education and living in deprived areas. RRs for males exceeded those for females and risk order was consistent for both sexes (lung followed by UADT with all cancer the lowest elevated risk). For males, elevated risk was associated with all area-education level combinations regardless of cancer group excluding the all cancer risk among males with a degree living in deprived areas. Elevated lung cancer risk in females was also associated with no education living in more affluent areas (RR 1.77, 95% CI 1.22, 2.36) (Table 4.7).

**Table 4.7 Area deprivation and education interrelationship: age adjusted relative risks (RR) and 95% confidence intervals (CI) by cancer and sex, Scotland 1991-2006**

| | | | Female | | | Males | | |
|---|---|---|---|---|---|---|---|---|
| Cancer | Area | Education level | RR | 95% CI | | RR | 95% CI | |
| All | Deprived | No education | **1.13** | **1.06** | **1.22** | **1.21** | **1.12** | **1.30** |
| | Deprived | Diploma or higher education | 0.93 | 0.81 | 1.07 | 1.01 | 0.85 | 1.19 |
| | Affluent | No education | 1.05 | 0.98 | 1.12 | **1.12** | **1.04** | **1.21** |
| | Affluent | Diploma or higher education | | | | Reference | | |
| Lung | Deprived | No education | **2.62** | **1.97** | **3.49** | **3.65** | **2.87** | **4.63** |
| | Deprived | Diploma or higher education | 1.27 | 0.74 | 2.20 | **2.04** | **1.31** | **3.20** |
| | Affluent | No education | **1.77** | **1.33** | **2.36** | **2.36** | **1.85** | **3.00** |
| | Affluent | Diploma or higher education | | | | Reference | | |
| UADT | Deprived | No education | **1.64** | **1.09** | **2.49** | **2.10** | **1.55** | **2.84** |
| | Deprived | Diploma or higher education | 1.18 | 0.53 | 2.61 | **1.80** | **1.02** | **3.18** |
| | Affluent | No education | 1.21 | 0.80 | 1.85 | **1.75** | **1.29** | **2.38** |
| | Affluent | Diploma or higher education | | | | Reference | | |

Source: Scottish Longitudinal Study

## 4.5 Discussion

This study found a complex and different pattern of socioeconomic factors associated with risk in different cancer groups in both sexes with no single individual or area-based socioeconomic factor predominant.

### 4.5.1 Place of birth

Being born in Scotland was associated with increased risk regardless of the cancer group and sex, which was in keeping with a previous study in Scotland that had focused on place of birth and cancer mortality (Wild et al 2006). The observed lack of any difference for lung cancer risk in males born in Scotland compared to the rest of the world may reflect the different stage in the smoking epidemic in Scotland for males and females relative to each other (with males reducing more rapidly than females), as well as the shifting of the global epidemic with reductions in high-income countries to increases in low- and middle-income countries (Mackay et al 2003; ScotPHO 2008; Mackay 2012).

### 4.5.2 Marital status

Relative to being married and in contrast to females, this study found no risks associated with all cancer or UADT cancers for widowed males but not for widowed females. This may reflect financial implications of widowhood for a

cohort of older women where marriage imparted greater financial security and little or no change in financial security for their male counter parts. Being divorced or widowed was also associated with increased cancer risk for females, while being single was associated with reduced risk for both sexes. These findings are broadly consistent with Danish studies which identified increased lung (Dalton et al 2008b), mouth and pharyngeal (Anderson et al 2008) and laryngeal (Anderson et al 2008) cancer risk associated with being divorced or widowed for both sexes. In contrast to this study's results for UADT cancer, being single was associated with elevated head and neck cancer risk in two Danish studies and one Italian (Randi et al 2004; Anderson et al 2008; Dalton et al 2008b). The Danish studies separately identified cohabiting and single individuals while this study was limited to legal marriage categories only (Anderson et al 2008; Baastrup et al 2008; Dalton et al 2008a; Dalton et al 2008b; Dalton et al 2008c). Further, the Italian study found compared to married individuals; those who had never married faced an increased risk of oral cavity and pharyngeal cancer diagnosis (Randi et al 2004). Reduced risk levels for single individuals seen in this study may have reflected the risk of individuals who were cohabiting but legally single as well as the risk of single individuals living alone. Many have suggested cohabiting or married individuals experience improved health status due to stronger social relationships and potentially healthier behaviours reflecting greater psychological reinforcement provided by partner support (Shouls et al 1996; Johnson et al 2008; Dalton et al 2008b; Islami et al 2009; Hashibe et al 2011; Yen et al 2013; Lagergren et al 2016), while being divorced or widowed may increase unhealthy behaviour due to reduced income and increased stress (Harvei et al 1997; Randi et al 2004; Bailey 2009). Poverty and social exclusion also has been shown to increase risk of divorce and separation as well as disability, illness, addiction and social isolation (World Health Organisation 2003).

### 4.5.3 Area-based socioeconomic deprivation

The finding that area deprivation remained significant for lung cancer risk even after adjustment for the individual SES factors is consistent with an analysis of

deprivation in small areas in Denmark, which reported increased neighbourhood population density and unemployment were associated with increased lung cancer risk (Meijer et al 2013). This neighbourhood effect of increased risk may reflect physical and social environments, including exposure to traffic or industrial related air pollution, reduced access to shops and services promoting healthier lifestyles and increased stressful environments and general sense of hopelessness associated with lack of supportive social networks, resources and opportunity (Frohlich et al 2001; MacIntyre S et al 2003). In the context of area air quality, a recent review of several European and US studies focusing on air pollution and the respiratory system found between 7–30% of lung cancer incidence was attributed to chronic exposure to air pollution (Beelen et al 2008). Consistent with other parts of the UK, in Scotland, Pye et al (2006) found greater air pollution concentrations in the more deprived deciles reflecting heavier road traffic in cities and higher proportion of deprived populations in urban locations. When compared to England and Northern Ireland, however, the inequality gradient associated with air pollution concentration was less steep in Scotland (Pye S 2006). Relative to the rest of the UK, higher lung cancer incidence rates in Scotland in general and among the more deprived areas does not appear to reflect current higher air pollution levels. Nevertheless and despite being below WHO guidelines (World Health Organisation 2006), air pollution in Scotland is greatest in more deprived areas (Pye S 2006). This may contribute to an already 'unhealthy' neighbourhood environment in deprived areas adding to stress and exacerbating already unhealthy lifestyles which potentially lead to lung cancer diagnosis or diagnosis at an even earlier age among the more deprived (Adams et al 2004).

### 4.5.4 Occupational social class

After full adjustment, the finding of increased UADT cancer risk for most occupational social class categories compared to the professional, managerial and technical group in both males and females was consistent with Anderson et al (2008) who studied occupational factors in relation to mouth, pharyngeal and laryngeal cancer risk. However, oesophageal cancer risk for females has

previously not been associated with social class (Baastrup et al 2008). Although the number of cases in the study presented here did not allow disaggregation of UADT cancers into subsites, the results were consistent with the findings of other studies reported earlier (Section 2.5), which showed differences in SES association with oesophageal cancer risk between the sexes (females weaker than males) as well as differences in SES association with different oesophageal cancer morphologies (increased risk association for squamous cell carcinoma and no association for adenocarcinoma). However the first study presented in this thesis did not explore any individual socioeconomic variables, including occupational social class. Furthermore, oropharyngeal cancer, ranked relatively low in terms of contribution to socioeconomic inequalities of all cancer risk for both males and females is one of the fastest increasing cancers in Scotland (Chapter 3) (Sharpe et al 2012).

In contrast to others, the present study did not find a strong association with lung cancer risk in either sex (Dalton et al 2008b). This may reflect the higher proportion of individuals who were economically inactive or had not held a job in the last 10 years (Tables 4.5, 4.6). The findings of occupational social class association with increased cancer risk is likely to reflect not only employment status but also prestige, qualifications, rewards and job characteristics (e.g. reporting relationship, locus of control and autonomy) all of which have been associated with social status differences in health, sickness absence and premature death (World Health Organisation 2003). The stronger increased UADT cancer risk association of occupational social class for males compared to females was consistent with the theory that the socioeconomic roles performed by males and females differ (Galobardes et al 2006a). For women, health has been shown to be more negatively affected by the psychosocial stress over the life-course of balancing caring, paid work and managing a household while work conditions alone more frequently negatively affect men’s health (UCL Institute of Health Equity 2013).

### 4.5.5 Employment status

Similar to this study's results, several studies have reported that not working versus working was associated with elevated risk in both sexes of all cancer (Dalton et al 2008c), lung (Dalton et al 2008b) mouth and pharyngeal (Anderson et al 2008), laryngeal (Anderson et al 2008), oesophageal cancers (Baastrup et al 2008) and oral cancer (Conway et al 2010a). Having a job is better for health outcomes than being unemployed, but the nature of the social relationships and their implication for stress at work can negatively contribute to illness (Adler N.E. et al 2002). Unemployment and negative health consequences are well established with health effects felt at the first signs of job insecurity leading to psychological stress and anxiety as well as financial impact (World Health Organisation 2003; Marmot 2005) and may lead to increased use of tobacco and alcohol in part as coping mechanisms (Baum et al 1999).

### 4.5.6 Educational attainment

The findings in relation to the elevated cancer risk associated with no attainment of educational qualifications are largely consistent with others who found reduced mouth and pharyngeal cancer risk for males with higher educational attainment and no risk difference for females for these cancers (Anderson et al 2008). No risk differences were also previously reported for educational attainment and oesophageal cancer for both sexes (Baastrup et al 2008), while reduced lung cancer risk associated with higher educational attainment was found for both sexes. Education may impact on the life-course – reflecting the early years home/ family socioeconomic circumstances and is recognised as a key factor in establishing a foundation for adult life (Galobardes et al 2006b). Many studies also suggest that education inequalities may have an underpinning role in health and social inequalities influencing the occupation attained and income earned in later life (Kogevinas et al 1997a; Galobardes et al 2006a). While this study has not been able to establish education as the most important factor influencing health outcome, others studying the impact of socioeconomic circumstances on health (including cancer incidence) over the life-course concluded that education level is the primary determinant (Smith

2007). This study's result that education was not the most important factor influencing health outcome may be explained by the theory that the relative importance of education may be dependent on levels of other SES measures suggesting that education was less important to health status among individuals who reside in households below poverty thresholds (Shavers 2007).

### 4.5.7 Car ownership

Compared to owning a car the analysis found that no car access was associated with increased risk of all cancer groups for females, but only for lung and UADT cancer for males. This observation that no car access was not associated with increased risk for all cancer in males is likely to reflect the mix of cancer sites included in this cancer group, some of which are more likely to be diagnosed among more affluent individuals (e.g. prostate cancer and melanoma) who are more likely to be car owners while other cancers are more likely to be diagnosed among the more deprived (lung and UADT cancer) who are less likely to own a car. Consistent with these results where lack of car access is associated with elevated lung cancer risk, Lancaster et al (2006) established elevated risk association regardless of sex in North England. Car ownership is a strong marker of material socioeconomic circumstances: the 2011 Scottish Household Survey (SHS) indicated car availability was strongly associated with income and that car access differed by sex with 76% of males and only 60% of females holding a license (Scottish Government 2012b). In this study, the proportion of car owners by sex for the full cohort was consistent with the SHS results (Table 4.1). The higher lung cancer RR for women without a car compared to men may have reflected differences in the smoking epidemic stage between men and women as well as the general shift in prevalence of the smoking habit from the more affluent to the more deprived as the more affluent adopted healthier non-smoking behaviour more quickly. The lower UADT RRs for women without a car compared to men was likely to reflect the weaker association of deprivation with UADT cancer risk among women. These results suggested for both sexes, to a lesser or greater degree depending on sex and cancer, car ownership was a marker of material wealth and as a resource enabling access to work, schools,

shops, leisure activities, friends and family, was an important socioeconomic dimension associated with cancer risk (Pevalin et al 2008).

### 4.5.8 Housing tenure

Several earlier Danish studies established increased risk associated with rented compared to owner occupied accommodation for all cancer (Dalton et al 2008c), lung (Dalton et al 2008b), mouth and pharynx (Anderson et al 2008), laryngeal (Anderson et al 2008) and oesophageal (Baastrup et al 2008) cancer regardless of sex. In contrast, this study found that the association was also not consistent for women and UADT cancer risk and for men and all cancer risk. With respect to women diagnosed with UADT cancer, it may be expected that renting would be associated with higher risk compared to the home owner category as housing condition is independently associated with deterioration of health, especially in women. Furthermore, renters are more likely to report more housing problems than owner occupiers (Pevalin et al 2008). The differences may reflect that household tenure is a material wealth indicator and the finding that deteriorating health applies to women home owners in poverty as well as renters (Ellaway et al 1998). Finally, these results may reflect the weaker association of UADT cancer with socioeconomic status for women compared to men observed in Chapter 3. Like the results for no car access, no difference in all cancer risk for males is likely to reflect the mix of cancer sites included in the all cancer group some of which are more likely to be diagnosed in the more affluent while other cancers are more likely to be diagnosed among the more deprived.

### 4.5.9 Inter-relationship of individual and area-based factors

The findings on the inter-relationship between area deprivation and education show the synergistic effect of area and individual SES measured by education and are consistent with others focusing on cancer (Spadea et al 2010) and lung function (Meijer et al 2013). Consistent with others this study established, low education level and high deprivation was associated with increased lung and UADT cancer risk in males and the risk order implied greater influence of education (Spadea et al 2010). For females, being educated to some extent

mitigated the effects of living in a deprived area, likewise living in an affluent area mitigated the effect of no education. Given that these cancers are largely driven by smoking and alcohol behaviours, which are both more prevalent among more deprived groups (Chapter 2; (Lawder et al 2010)), the implication is that social and cultural aspects of SES are important in uptake and continuation of smoking and alcohol consumption (Spadea et al 2010). Education level captures the impact of socioeconomic and cultural circumstances at an early age when adopting these habits. In addition, the differences between the sexes in the smoking epidemic are likely to explain the mitigating effects identified.

### 4.5.10 Relative influence of area vs individual SES measures

As all variables used to establish the area deprivation measure were included in the model (excluding accommodation overcrowding due to no discernible differences in the cohort population), the finding that that these variables remained statistically significant in the fully adjusted model further supports the need for a separate and independent role of individual socioeconomic factors in addition to the area measure. However, these results may reflect confounding by other unavailable and unmeasured factors including geographic attributes such as environmental pollution, individual risk behaviours and other individual SES variables such as income and house value — a potentially important individual SES measure given the cancers under investigation are most likely to be diagnosed among the old who are also more likely to have access to accumulated wealth.

### 4.5.11 Potential explanations of relationship with low SES and cancer risk

It has been suggested that low SES, regardless of measure, potentially implies some form of ‘stress’ which may come from a range of sources such as insecurity of work, unemployment, fear of crime, debt, low material resources and low social capital and community cohesion (Marmot 2010; Kawachi et al 2017). Lifelong adverse experiences have strong and long lasting impacts on health and occur most often among the most socioeconomically deprived (Cambois et al

2011). Furthermore, these impacts occur across the life-course, with disadvantage at critical life transition points being identified as contributing to deteriorating health — such as early childhood, moving from primary to secondary school, starting work, leaving or moving home, starting a family, job change, facing redundancy and retirement (World Health Organisation 2003). Recent studies, seeking a "biological" explanation report telomere lengths which vary by age, sex and ethnicity are associated with biological ageing and cancer (MacIntyre et al 2002). Various studies, including systematic reviews, have explored the predictive potential of (shortened) telomere length for cancer risk (Prescott et al 2012; Wentzensen et al 2011; Ma, X et al 2011). Moreover, shortened telomere length has also been found to be associated with lower SES, measured by certain socioeconomic variables (Shiels et al 2011; Robertson et al 2012) such as low relative household income and renting a home and adverse early life experiences (Kelly-Irving et al 2013) as well as lifestyle factors such as smoking and poor diet (Shiels et al 2011). Cancers strongly associated with smoking such as lung cancer display most consistent results showing shorter telomere length association with incidence (Wu et al 2003). Behaviours such as smoking (Wu et al 2003), alcohol consumption abuse (Wu et al 2003) and obesity (Wu et al 2003) are also associated with accelerated telomere attrition as well as recognised as risk factors for lung and UADT cancer.

### 4.5.12 Study strengths

To date, many studies have focused on cancer mortality; here for the first time in Scotland, this study examined multiple individual SES metrics as well as an area measure to explore cancer risk. Area-based rather than individual measures of SES, created for the smallest available administrative unit, out of necessity, are increasingly used worldwide to measure effects of SES on health outcomes and to plan services (Hemminki et al 2003) and may be used as surrogates for individual social indicators (Berkman et al 1997; Pickett et al 2001; MacIntyre et al 2002; Leyland 2005; Riva et al 2007). However, while convenient, these area-based measures are not without their limitations. This study recognised that individual SES classification based on area SES measures may not reflect

individual SES accurately ('ecological fallacy') (McLaren et al 1998; Shaw et al 2009) as well as the importance of investigating the relative influence of individual as well as area socioeconomic circumstances when considering SES as the exposure (Galobardes et al 2006a). This study also undertook for the first time, linkage of SCR incidence data with the SLS to provide a large cohort with a number of primary tumours followed for several years. Finally, the SCR is a population-based cancer registry with evidence of high data quality and less than 1% of cases identified through death certification only (Brewster et al 2002).

The analysis excluded any diagnosis of cancer prior to the April 1991 Census and cohort start; this coupled with measurement of area and individual SES variables at the 1991 Census provided measurement at the earliest time possible prior to diagnosis. This gave the advantage of knowing individual SES before cancer diagnosis rather than the traditional area measurement at time of diagnosis. Measurement at time of diagnosis, the typical time when SES measures are captured in relation to cancer, may reflect the reverse impact of diagnosis on socioeconomic circumstances (Kawachi et al 2006). Therefore, the finding suggesting a strong role of low SES provides stronger evidence than the typical cross-sectional nature of area-based SES captured in routine cancer registry analyses.

Finally, the Scottish Cancer Registry has high case ascertainment rate with only 0.4% of registrations based on death certificate only (Brewster et al 2002).

### 4.5.13 Study limitations

To capture socioeconomic circumstances at the earliest point in the study the Registrar General's occupational social class was used given availability of the more recently developed National Statistics Socioeconomic Classification (NS-SEC) system at study start (Craig et al 2005). However, this measure focused on manual versus non-manual distinction between occupations and is only applicable to those in paid employment, omitting important segments of society such as the unemployed, retired and permanently sick (Black D 1980; Jones et al

2004; Craig et al 2005). Finally, as indicated previously, there was no access to any risk behaviour data.

The analysis employed person-years models which estimate the risk of cancer incidence in the absence of competing risks, even those competing risks that may be correlated (for example, a smoking related cause of death other than cancer). Because individuals succumbing to a non-cancer smoking-related death may be at greater risk of cancer had they lived, the estimated risks may underestimate the effects of the variables under investigation. However, due to the desire to measure the association of SES exposure with cancer incidence, in-effect performing a prognostic marker effect test, this approach was preferred to alternatives such as the cumulative incidence function (Dignam et al 2012). It may be suggested that multi-level modelling would have been a more suitable analytical approach given that the study was exploring one area and five individual SES indicators. The only area deprivation indicator (Carstairs) was measured at postcode sector level of which there are 1,011 in Scotland. Given the small number of cases by cancer group and sex, there were many postcode sectors with either no or only a very few cases and therefore no individual measurements available. As a result, multi-level modelling was not appropriate for the data available. Finally, the approach adopted (fully adjusted model) recognised the *a priori* hypothesis (and conscious SES variable selection) that different individual SES variables capture different SES dimensions at different points in the life-course. Area measures of socioeconomic inequality, including the one used in this study are frequently composite measures reflecting a number of different aspects of socioeconomic circumstances. For area deprivation measures, a composite index is often used to capture as much of the multi-dimensional nature of deprivation as possible. In this study, depending on the cancer and sex, both the area measure and the included individual variables were associated with cancer risk to various magnitudes. This complex picture is likely to be further complicated by other unavailable demographic or socioeconomic dimensions such as ethnicity (Adler et al 2002), long term income (Benzeval et al 2001) and wealth (Pollack et al 2007) or power (Glasgow Centre for Population Health 2017). Despite this emerging understanding for cancer risk,

few if any, composite individual measures tailored to the specific population and outcomes have been considered.

## 4.6 Conclusion

This study recognised the strengths and limitations of relying on area measures of deprivation alone and begins to reassess more finely the socioeconomic factors associated with cancer risk. This study also showed that there was generally a strong temporal relationship between low SES measured (up to 16 years) before cancer diagnosis. In part this refutes the arguments of reverse causation.

This association of multiple socioeconomic and demographic variables with cancer risk is likely to reflect not only the complex, multifaceted nature of deprivation, but also the various and cumulative effects of different socioeconomic determinants over the life-course (Braveman et al 2011). This may reflect the fact that an individual's socioeconomic circumstances may change over the course of their life and the impact of which can accumulate over time. This complexity was also likely to reflect the longer lag-time between exposure and diagnosis for cancer incidence; for example, lung cancer lag period is estimated at several decades (Bilello et al 2002).

This study identified that different socioeconomic variables are not proxies of each other, but are independently associated with different cancer risks in both sexes. No single measure of socioeconomic circumstances dominated the risk profile or comprehensively reflected all aspects of socioeconomic stratification or captured the full effect of low socioeconomic circumstances at different stages in the life-course or transmitted over generations. The different components of SES not only suggest different cohort subgroups, but point to different pathways such as different behaviours or to critical periods of the life-course. The results emphasize the importance of using multiple SES measures in epidemiological studies.

In conclusion, different and independent socioeconomic variables were inversely associated with different cancer risks in both sexes; no one socioeconomic variable on its own captured all aspects of socioeconomic circumstances over the life-course. Association of multiple socioeconomic variables was likely to reflect the complexity and multifaceted nature of deprivation as well as the various roles of these dimensions over the life-course which in turn reflected the longer gestation period for cancer.

# 5 Investigating the role of behaviours in socioeconomic risk association with cancer incidence in Scotland: a Scottish Cancer Registry and Scottish Health Survey Data Linkage Population study (1995 – 2011)

## 5.1 Introduction

### 5.1.1 Role of area and individual socioeconomic factors and behaviour risk factors

Differences in cancer risk that are associated with socioeconomic factors have previously been explained by known risk factors such as behaviours, for example, smoking (Thun et al 2009), alcohol (Marshall et al 2009); obesity and body composition (Ballard-Barbash et al 2009), lack of physical exercise (Lee et al 2009), diet and nutrition (Willett 2009), sexual behaviour (Kawachi et al 2006; Mueller et al 2009), occupational exposure (Siemiatycki et al 2009), and various biological agents, including, human papillomavirus, Helicobacter pylori, and hepatitis B and C (Mueller et al 2009). Studies have also indicated that these behaviours do not occur in isolation — they cluster. The prevalence of multiple behavioural risk factors was high in the Scottish population and this was strongly associated with poorer socioeconomic circumstances (Lawder et al 2010).

The Scottish Health Surveys (SHeS) are cross-sectional cluster-sampled surveys designed to provide data, at both the national and regional level, about the health of the population living in private households in Scotland (Gray et al 2010). Longitudinal information can be created through linkage with the Scottish Cancer Registry (SCR), maintained by the Public Health Intelligence Unit of National Services Scotland (Section 1.4.5).

At the time of undertaking the analysis for this study in 2014, only four studies had been previously published in a peer reviewed journal which had linked the SCR and SHeS data (Hamer et al 2009a; Hamer et al 2009c; Evans et al 2011; Atherton et al 2012). In three of the four studies, SES, measured through

occupational social class, was employed as a confounder rather than as the exposure. In all cases, only one measure of SES, occupational social class, was used, omitting the wealth of SES measures held in the SHeS. In addition, in all four studies, only two or three surveys were used. As a result, in these studies, a wider understanding of the impact of SES as well may not have been fully tapped. Furthermore, the opportunity to increase the number of study participants may have been missed. For example, where it may have been possible to use six of the SHeS studies, only two were used (Atherton 2012) either because of the time of the study relative to the completion dates of more recent SHeS surveys or potentially because the excluded (earlier) surveys did not capture consistent or the same information as the selected surveys. Finally, only two of the studies were designed as a cohort (Hamer 2009a, 2009c) and due to the timing of the study relative to availability of surveys, this will have resulted in a more restricted follow-up period. In this thesis, there were opportunities to more fully exploit the SHeS by using more of the available surveys thereby creating a larger cohort with a longer follow-up period. This was possible not only because this SHeS and SCR linkage study was conducted in 2013–14, but also because all available surveys were used. A brief summary of the key methodological aspects and approaches of these SCR – SHeS linkage studies are discussed here:

***Atherton et al (2012)*** explored SES and self-assessed health of those with and without a cancer history. They used pooled data from the 2003 and 2008 SHeS data linked to SCR data to create a cross-sectional dataset that was comprised of 17,505 participants, 432 of whom had been previously diagnosed with cancer. Cancer survivors were defined as those with a diagnosis prior to their participation in either SHeS survey. SES was established via the three categories of the National Statistics — Socioeconomic Classification scheme (managerial and professional, intermediate occupations and routine and manual occupations). Logistic regression models for those with and without a cancer diagnosis by SES group were created to estimate odds ratios of poor self-assessed health adjusted for age, sex, time since cancer diagnosis and co-morbidities. The authors found that cancer survivors in the lowest SES group were almost three-times more

likely than those in the highest group to report poor general health (OR 2.96 95% CI 1.82, 4.80). This difference between the lowest and highest SES groups widened for those participants who were four or more years after a cancer diagnosis. This study used SES as the exposure, although only one measure of SES was employed. A cross-sectional study design was undertaken using only two SHeS surveys.

***Evans et al (2011)*** used the 1995, 1998 and 2003 SHeS surveys that were linked to the SCR to create a cross-sectional dataset of 25,631 people with no cancer diagnosis and 507 cancer survivors. They compared self-assessed health and well-being for cancer survivors with those who never had a cancer diagnosis. The outcomes were measured separately via reduced activity and psychological morbidity, using logistic regression models to calculate odds ratios adjusted for sex, age and occupational social class and time from diagnosis. In this study, SES was again, considered a confounder and not an exposure.

***Hamer et al (2009c)*** examined the association between different types of physical activity and mortality in 293 participants who had been diagnosed with cancer before they entered the SHeS (1995, 1998 and 2003); 78 of these participants died during the follow-up period. The authors used Cox proportional hazard models to estimate hazard ratios adjusted for age, sex, SES via occupational social class, marital status, BMI, smoking status, alcohol consumption as well as co-morbidities (diabetes, hypertension, admissions for cardiovascular events, type of cancer, number of cancer recurrences) and survey year. Compared to no physical activity, the authors concluded that vigorous exercise (HR 0.47 95% CI 0.23, 0.96) as opposed to domestic or walking (HR 0.95 95% CI 0.57, 1.56) reduced all-cause mortality. In this study, again, SES was used as a confounder, not the exposure; the study design was a cohort, but only three surveys were used, most likely due to availability, and the number of study participants was small.

***Hammer et al (2009a)*** used the 1995, 1998 and 2003 SHeS incorporating 15,453 participants (295 diagnosed with cancer) linked to the Scottish Cancer Registry

to explore psychological stress and cancer mortality. The risk of mortality in relation to psychological distress was estimated using Cox proportional hazards models adjusted for age, sex, Registrar General Classification of occupation, marital status, BMI, smoking status, alcohol consumption, physical activity and survey year. In the fully adjusted models, cancer survivors with distress compared to survivors with no distress were nearly two-times as likely to die from cancer (HR 1.97 95% CI 1.05, 3.71). The authors suggested that this elevation may reflect the inclusion of individuals with a cancer history. When these individuals were removed, higher levels of distress were associated with increased lung cancer mortality only. This study again, considered SES as a confounder and not the exposure. The study design was a cohort, but only three surveys were used, most likely due to availability.

Since 2014, four further papers have been published (Wang et al 2015; Leung et al 2016; Batty et al 2017; Stewart et al 2017). Two of these papers focused on different aspects of psychological stress and cancer survival or mortality (Leung et al 2016; Batty et al 2017). One study focused on health behaviour and well-being of cancer survivors compared to the general population (Wang et al 2015). Only one paper was related to the examination of socioeconomic circumstances (Stewart et al 2017). In this case, the researchers focused on the cancer related health behaviours of young people not in education, employment or training.

These studies demonstrated the possibility of using the SHeS to investigate other public health issues beyond cardiovascular illness (the original objective for development of the SHeS). Using linked SCR and SHeS datasets to investigate cancer related public health issues, at the time of conducting this study was relatively untapped. Furthermore using these datasets to explore SES and cancer incidence had not previously been attempted.

The overall percentage of SHeS participants consenting to linkage with other datasets is high and has remained above 83%, despite declining overall survey response levels (Morris 2017; ADLS 2017a) (Table 5.1). Those SHeS respondents agreeing to linkage to the CHI database which supports linkage to other datasets

such as the Scottish Cancer Registry is very high at 97.8% in 2012 or higher (Table 5.1).

**Table 5.1 Scottish Health Survey linkage metrics by survey year, Scotland 1995 to 2014**[1]

| Year | Total Respondents | Respondents consenting to linkage | | Respondents linked to CHI | | Respondents not consenting |
|---|---|---|---|---|---|---|
| | | N | % | N | % (of consenting) | N |
| 1995 | 7,932 | 7,363 | 93% | 7,259 | 98.6% | 569 |
| 1998 | 12,939 | 8,296 | 64% | 8,275 | 99.7% | 4,643 |
| 2003 | 11,420 | 10,470 | 92% | 10,407 | 99.4% | 950 |
| 2008 | 8,170 | 7,028 | 86% | 6,994 | 99.5% | 1,142 |
| 2009 | 10,116 | 8,667 | 86% | 8,615 | 99.4% | 1,449 |
| 2010 | 9,015 | 7,765 | 86% | 7,739 | 99.7% | 1,250 |
| 2011 | 9,510 | 8,155 | 86% | 8,119 | 99.6% | 1,355 |
| 2012 | 6,602 | 5,617 | 85% | 5,496 | 97.8% | 985 |
| 2013 | 6,733 | 5,731 | 85% | 5,699 | 99.4% | 1,002 |
| 2014 | 6,327 | 5,323 | 84% | 5,296 | 99.5% | 1,004 |

[1] **(Morris 2017)**

However, it is known that certain subgroups of the population such as men, younger individuals and those from more disadvantaged areas are more likely to not participate or not consent to linkage of their data making the dataset less representative of the full Scottish population (Gray et al 2013). An example of this issue was the inconsistency between SHeS data and alcohol- related mortality and hospital admissions (Gray et al 2012; Gray et al 2013). These analyses suggested that there was however no association of alcohol intake with area deprivation while alcohol related admissions and mortality was much higher for those living in the most disadvantaged areas of Scotland. Furthermore, the survey identified participants living in private households and therefore excluded those living in residential care homes and prisons as well as those in the armed forces (Gray et al 2010).

The SHeS, commissioned by the Scottish Government Health Directorate, was based on face-to-face interviews and collected a wide range of demographic, socioeconomic and behavioural variables. Their purpose was to monitor health in order to shape policy and develop new health initiatives. Early and current socioeconomic information was available, with a range of measures of socioeconomic status taken at the time of the survey. Linkage to the SCR

provided prospective outcome information.

Records from SHeS conducted in 1995, 1998, 2003, 2008, 2009, 2010 and 2011 were linked to SCR records from 1995 to 2011 to provide a maximum follow-up of approximately 16 years. Linkage of these two datasets provided the opportunity to assess the risk associations of socioeconomic circumstances and behaviours on cancer risk in a prospective manner.

There were 46,368 adults over 16 years old that were included in the survey. Over time, the range of ages included in the surveys was widened. The survey in 1995 only included adults up to 65 years old; in 1998, children over two years old and adults up to the age of 75 years were sampled; in 2003, the full age range was surveyed (ADLS 2017a).

The SHeS sample was drawn based on postcode sectors and household addresses. Stratification was based on geographical areas and not on individual characteristics of the population. Different samples were drawn for each survey (Hamer et al 2009b; Gray et al 2010).

Despite its favourable attributes (e.g. response rates, number of years, behaviours captured), the SHeS has not been used to its full potential and presented the opportunity to evaluate relative effects of area and individual socioeconomic circumstances on various health endpoints. Like the English Health Survey, the principle focus of the SHeS has been to address questions related to cardiovascular disease and associated risk factors (Gray et al 2010).

As found in previous chapters (Sections 1.2, 2.0, 3.0, and 4.0), the association between SES and health was well established and showed a mostly consistent pattern of poorer health with lower SES. Lung and upper aero-digestive tract (LUADT) cancers taken together were among the most common cancers in the world compared to the other individual sites with a strong smoking aetiology; with 21% of global cases diagnosed in Europe in 2012 (Ervik et al 2016). These cancers showed stark socioeconomic inequalities with greater incidence among lower socioeconomic groups (Gray et al 2009a; Hamer et al 2009b; Landy et al

2012).

The study that was described in Chapter 3 investigated the incidence of lung, UADT and all cancer associated with area and individual measures of demographic and socioeconomic circumstances (marriage status, economic activity, occupational social class, education level, car ownership and household tenure) and found that different and independent socioeconomic variables were inversely associated with different cancer risks in different sexes; no one socioeconomic variable on its own captured all aspects of socioeconomic circumstances or the life-course. However, no behaviour risk data were available. The investigation of the relationship between SES and behaviour may explain some or all of the socioeconomic inequalities identified. Behavioural factors such as alcohol (WCRF/AICR 2007; Winn et al 2015; Fehringer et al 2017) and smoking (Kamangar et al 2009; Pesch et al 2012; Winn et al 2015) are strong risk factors for these cancers with diet (fruit and vegetable consumption) (WCRF/AICR 2007; Winn et al 2015; Schwingshackl et al 2015), body mass index and obesity (Winn et al 2015; Hidayat et al 2016), physical activity (Nicolotti et al 2011; Behrens et al 2014; Singh et al 2014; Winn et al 2015) and sexual behaviour (Winn et al 2015) reported to be relevant depending on the cancer site.

In addition, and as discussed in Section 1.3, factors including the long lag-time between initiation of cancer and diagnosis, the complex and dynamic nature of SES over the life-course, and the desire to avoid reverse causation bias made it desirable to measure SES at the earliest point in time prior to cancer diagnosis as possible. Linkage of the SCR to the SHeS provided the opportunity to measure SES at the time the study participant entered the study rather than at diagnosis, the only option possible with traditional cancer registry analyses.

### 5.1.2 Study aims and objectives

The study aim was to assess whether behaviour risk factors explained the previously identified socioeconomic inequality as measured by individual and area SES (and in combination) measures in all cancer (excluding non-melanoma

skin cancer) and LUADT cancer incidence.

This aim of was to be achieved through developing a data linkage cohort study to analyze incidence of multiple SES (including SIMD, economic activity highest qualification, occupational social class, access to cars, housing tenure and multiple low SES) and behavioural (smoking status, cigarettes smoked per day, age started smoking, duration smoked, units of alcohol consumed per week, binge units consumed per day, fruit and vegetable consumption per day, exercise sessions per week and multiple high risk behaviours) risk factors and their association with cancer risk for all cancer and for lung and upper aero-digestive tract cancers together, hereafter described as LUADT.

## 5.2 Methods

### 5.2.1 Cohort definition

Data were obtained from the seven SHeS available at the time of analysis which were conducted in 1995, 1998, 2003, 2008, 2009, 2010 and 2011 and were linked to the SCR data from 1975-2011 (latest incidence year available at the time of performing the study in 2013-14) to create a cohort (Gray et al 2010). For individuals who participated in more than one SHeS, only their first SHeS interview data were included. As a result, any individual who was included in a subsequent SHeS was described by the data provided at the first survey in which they were involved. All individuals diagnosed with cancer prior to survey interview were excluded. The Scottish Government reviews progress on addressing socioeconomic inequalities in all cancer incidence; therefore all cancer combined (excluding non-melanoma skin cancer) were also evaluated (ICD-10 C00-C96 excluding C44) (Scottish Government 2008d; Scottish Government 2009; Scottish Government 2012a; Scottish Government 2013c; Scottish Government 2015; Scottish Government 2017c).

Due to small case numbers, lung and upper aero-digestive tract cancers (LUADT: C32, C33-C34, C00-C14, C15, M-8050-8076) and both sexes were combined. Oesophageal adenocarcinoma was excluded due to a previous Scottish study

identifying no association of this histological type with SES (Brewster et al 2000) (Section 2.5.2). The combined LUADT cancer group also reflected that smoking was identified as a high cancer risk behaviour for all of these cancers (Bilancia et al 2009; Thun et al 2009; Kamangar et al 2009; Al Dakkak et al 2011; Winn et al 2015; Fehringer et al 2017; Arnold et al 2017) (Chapter 2.0).

The cohort was defined by the study start date equal to the first SHeS interview date and the study end date defined as the earliest of date of diagnosis, date of death, or 31st December 2011. Individuals were followed for up to 16 years.

## 5.2.2 Socioeconomic variables

One area-based and six individual socioeconomic measures were used to capture different known aspects of socioeconomic circumstances. All SES variables were based on the individual's first participation in the SHeS thereby potentially avoiding reverse causation bias and recognising the long lead-time between cancer initiation and diagnosis. All SES variables were based on the variable list and definitions as stated in the SHeS documentation relevant to that survey year (Joint Health Surveys Unit of Social and Community Planning Research and University College London 1999; Joint Health Surveys Unit of Social and Community Planning Research and University College London 2016; ScotCen for Social Research 2016; ScotCen for Social Research et al 2016; Joint Health Surveys Unit 2016a; Scottish Centre for Social Research 2016a; Scottish Centre for Social Research 2016b).

For consistency over the study period and multiple SHeS years, variables were re-defined for this study. The variable definitions, categories for this study and reference category are described below. The category with the lowest risk was identified as the reference category in order to establish relative risks that were greater than 1.0.

### 5.2.2.1 Area SES indicator: SIMD

Area deprivation was assessed using the national Scottish Index of Multiple

Deprivation (SIMD) quintile which is a measure of area-based deprivation using 31 indicators in six individual domains of: current income, employment, housing, health, education, skills and training and geographic access to services and telecommunication. SIMD is calculated at datazone level (500-1000 population), enabling small pockets of deprivation to be identified. The datazones are ranked from most deprived (1) to least deprived (6505) on the overall SIMD index and grouped into deciles or quintiles for purposes of epidemiological analysis (Scottish Government 2012c). Different SIMD years were employed depending on the survey year that the individual first participated in the SHeS (Appendix 5.3 and Section 1.3.2.4).

Generally, earlier survey (1995, 1998 and 2003) variable and category definitions differed from later survey (2008, 2009, 2010 and 2011) definitions necessitating creation of consistent variables and categories for all seven surveys (Appendix 5.3).

SIMD (SCSIMD2012) codes for surveys conducted in 1995 and 1998 differed from subsequent years such that 1=most deprived and 5=least deprived. In all subsequent survey years the SIMD variable (SIMD5 for 2003 and SIMD5_RP thereafter) reversed this convention such that 5=most deprived and 1=least deprived (Bishop J et al 2004; ISD Geography Analysis Support 2012; Jackson 2017). As a result the SIMD coding for SHeS years 1995 and 1998 were adjusted accordingly. For modelling purposes, the reference category for SIMD was defined as “least deprived”.

#### 5.2.2.2 Economic activity

Economic activity was defined by the variable named ECONACT and used in both the 1998 and 2003 SHeS. ECONACT described economic activity as employed, unemployed, retired and other economically inactive. The level of documentation for the 1998 SHeS survey did not clearly state if the “unemployed” category excluded those who are permanently unable to work or not (Appendix 5.3) (Joint Health Surveys Unit of Social and Community Planning Research and University College London 2016). However, the enhanced 2003

SHeS survey documentation did clearly indicate that the "unemployed" category adopted the International Labour Organisation (ILO) definition of unemployment which "covers people who are: out of work, want a job, have actively sought work in the previous four weeks and are available to start work within the next fortnight; or out of work and have accepted a job that they are waiting to start in the next fortnight" (Office of National Statistics 2017). Therefore, the "other economically inactive" category in the 2003 SHeS did include those survey participants that were permanently unable to work (Appendix 5.3).

The economic activity variables used by the other SHeS surveys (1995, 2008, 2009, 2010, 2011) differed and were required to be collapsed into the desired categories. The 1995 SHeS economic activity variable (ECSTA) described economic activity more finely. Therefore the "employed" category included "full time", "part time" and "unspecified hours work" groups. While the "other economically active" category included: "keeping house", "full time student with no job", "permanently sick" and the "other inactive" groups. The 2008, 2009, 2010 and 2011 SHeS variables for economic activity (ECONAC_08) also described economic activity more finely but differently from the 1995 SHeS survey. For these surveys, "employed" included individuals in "paid employment" and "self-employed or government training", while "unemployed" included those individuals "looking for or intending to look for paid work", "retired" included the "retired" group and "other economically inactive" included those "permanently unable to work", "looking after home or family" and "doing something else". With the exception of the 1998 SHeS, the approach for including economic activity adopted for this study recognised that the economically inactive group was a heterogeneous group that was not easily categorised as employed or unemployed while retirees in theory have a source of (albeit reduced) pension income and therefore were separately identified (Brown et al 2012). For modelling purposes, the reference category for economic activity was defined as "employed" (Appendix 5.3).

#### 5.2.2.3 Occupational social class

Occupational social class for the household's chief income earner was grouped

using the Registrar General defined categories: Social Class I (professional, managerial, technical), Social Class II (intermediate), Social Class, IIINM (skilled non-manual), Social Class IIIM (skilled manual), Social Class IV (partly skilled) and Social Class V (unskilled) (Rose 1995; Berkman et al 1997). Although the variable names changed from survey to survey, all SHeS surveys adopted the same code structure for occupational social class, thus no regrouping or assumptions were required. For modelling purposes, the reference category for occupational social class was defined as "professional, managerial and technical" (Appendix 5.3, Section 1.3.2.2).

#### 5.2.2.4 Highest educational qualification

Education qualifications reflected the highest attained degree of the participant on entry of the SHeS (first degree and higher, other non-degree, none or missing) (Nordahl et al 2014) and were based on the variable (HEDQUL08) for the 2008, 2009, 2010 and 2011 Scottish Health Surveys. The category "First degree or higher" included "Degree or higher"; while the category "Other non-degree" included "HNC or equivalent", "Higher grade or equivalent" and "Other school leaver"; and finally the category "None" was defined by "No qualifications". The variables for 1995, 1998 and 2003 were more refined and differed by year, but collapsed into these same categories. For modelling purposes, the reference category for highest qualification was defined as "first degree and higher" (Appendix 5.3).

#### 5.2.2.5 Car ownership

Car ownership was recorded in each of the surveys consistently as the number of cars owned in the household and was grouped into no car or one or more cars. For modelling purposes, the reference category for car ownership was defined as "one or more car(s)" (Appendix 5.3).

#### 5.2.2.6 Housing tenure

The study variable used for housing tenure was defined by the named variable TENURE used in the 1995 SHeS which established categories of "owner

occupier", "rented from a local authority or housing association" and "privately rented". In the 1998 survey, the variable was renamed (OWNORENT) and refined such that the "rent privately" and "rent from local authority or housing association" variables were expanded; collapsing these categories re-created the desired housing tenure variable structure. As the 2003 SHeS survey retained the same variable name and structure as for the 1998 survey, it was possible to combine the "rent privately" category (including "rent privately", "unfurnished, rents privately", "furnished, rents from employer" and "rents other with payment" and "rent free") and the "rent from local authority/ *housing* association" category (including "rents from local authority/ *new* town" and "rents from housing authority"). From 2008, the SHeS variable for housing tenure was simplified. As a result, it was not possible to distinguish between renting privately from renting from a local authority or housing association. Therefore, for the 2008, 2009, 2010 and 2011 SHeS surveys, the "owner occupier" category included: "buying it with the help of a mortgage or loan", "own outright" and "pay part rent and part mortgage (shared ownership)"; "rented from local authority or housing association" included "rent it" and "rent privately" included "tied accommodation (e.g. where the accommodation goes with your job)" and "live here rent free (including rent-free relative's/friend's property) (Appendix 5.3).

As far as possible, this approach reflected the different socioeconomic implications of renting from a private landlord versus a local authority or housing association (MacIntyre et al 1998; Ellaway et al 1998; MacIntyre et al 2001). For modelling purposes, the reference category for housing tenure was defined as "owner occupier".

#### 5.2.2.7 Multiple low socioeconomic circumstances

To reflect multiple low socioeconomic circumstances, a derived variable, defined as the count, at individual level, of socioeconomic variables in the high risk category was created where a minimum score of zero and a maximum score of six was possible. As a result, an individual who did not have any socioeconomic risk factor falling into the high risk category for any SES variable

would have a score of zero for multiple low socioeconomic circumstances. This was in contrast to an individual who fell in the high risk category for each of the socioeconomic variables, in which case, his or her multiple socioeconomic circumstances score equated to six. As a consequence, each socioeconomic risk factor was equally weighted. Values between greater than zero and up to five indicated that some, but not all of the six socioeconomic risk factors were in the high risk category. As an example, an individual who was unemployed, lived in an area designated as most deprived and had no education would have a multiple low socioeconomic circumstances score value of three. For modelling purposes, the reference category for multiple low socioeconomic circumstances was defined as "no low socioeconomic circumstances" (Table 5.2).

**Table 5.2 Definition of multiple low socioeconomic circumstances**

| Socioeconomic risk factor | High risk category |
|---|---|
| SIMD | 5 Most deprived |
| Economic activity | Unemployed |
| Highest educational qualification | None |
| Occupational social class | V Unskilled or manual |
| Car ownership | None |
| Housing tenure | Renting (Local Authority or private) |

### 5.2.3 Behaviour variables

Behavioural variables were selected to capture known risk factors (Sections 1.7, 5.1.1) for LUADT cancers and reflected the data available in the SHeS (Joint Health Surveys Unit of Social and Community Planning Research and University College London 1999; Joint Health Surveys Unit of Social and Community Planning Research and University College London 2016; ScotCen for Social Research 2016; ScotCen for Social Research et al 2016; Joint Health Surveys Unit 2016a; Scottish Centre for Social Research 2016a; Scottish Centre for Social Research 2016b). Body mass index and obesity were not included as they are not strictly behavioural factors and are determined largely by, i.e., physical exercise and diet (Goodarzi 2017).

For consistency over the study period involving multiple SHeS years, the

behaviour variables were re-defined for this study. The study variable definitions and reference categories are described below while Appendix 5.3 provides this for each of the SHeS surveys. For modelling purposes and for each variable, the category which was most likely to have the lowest risk was identified as the reference category in order to establish relative risks (RRs) that were greater than 1.0.

#### 5.2.3.1 Smoking

Smoking behaviour was captured by smoking status (ex-smoker, current smoker), number of cigarettes smoked per day (ex-smoker, <20, >=20), age started smoking (<20, 20-39, 40 or older) and duration smoked in years (1-20, 21-40, 41-50, over 50 years). Four smoking variables were used to address possible self-reporting bias or "social desirability" (Connor et al 2009) bias of smoking status resulting in under estimating prevalence of smoking in the cohort. A systematic review of 54 studies comparing self-reported status versus biological measurement of cotinine demonstrated under estimation of prevalence, but was unable to estimate the extent (Connor et al 2009). A further reason to use four smoking variables was to provide a definition of high risk reflecting what was recognised in epidemiological literature as a more accurate method for modelling the clinical risk of smoking (Peto 2012) (Section 2.3.2).

Different smoking variables were collected in 1995 requiring creation of new smoking variables which were consistent with the other survey years (1998, 2003, 2008, 2009, 2010 and 2011). For the 1995 survey, this was achieved by deriving smoking status and number of cigarettes smoked per day from the supplied variables (NUMSMOK, SMOKECIG, SMOKENOW, SMOKEREG and STARTSMK) and categories. Study participants who responded "No" to the question "Have you ever smoked cigarettes?" captured by the variable SMOKECIG were classified as "Non-smokers". Those who responded "Yes" to the question "Do you smoke cigarettes at all nowadays" captured by SMOKENOW were classified as "Current" smokers. Those who responded "No" to SMOKENOW and responded "Smoke them only occasionally" or "Never really smoked cigarettes, just tried them once or twice" to a question about how regularly or occasionally

they had smoked (SMOKEREG) were classified as “Ex-smokers”.

The variable “Age started smoking” was defined as being equal to the named variable STARTSMK which recorded this information while “Duration smoked (years)” was derived from the respondent’s age less age started smoking (Joint Health Surveys Unit of Social and Community Planning Research and University College London 1999). To quantify number of cigarettes smoked by a current smoker, the named variable NUMSMOK was used and grouped into non-smoker, <20 cigarettes per day and => 20 cigarettes per day. For all subsequent surveys (1998, 2003, 2008, 2009, 2010 and 2011) the variable for smoking status (CIGST1) and the variable banding current smokers (CIGST2) was consistent. For this study, smoking status was described as non-smoker (including “never smoked cigarettes at all”, “used to smoke cigarettes occasionally” and “used to smoke cigarettes regularly”) or current smoker including “current cigarette smoker”. Number of cigarettes smoked per day was based on the named variable CIGST2 categories summarised as non-smoker (including non-smoker), <20 per day (including “light smoker, under 10 per day” and “moderate smoker 10 to under 20 a day”) and =>20 per day (including “heavy smoker, 20 or more a day”) (Appendix 5.3).

For modelling purposes, the reference category for each smoking variable was: smoking status (non-smoker), cigarettes smoked per day (non-smoker), age started smoking (>=40 years old) and duration smoked (1-20 years).

#### 5.2.3.2 Alcohol consumption

Consumption of alcohol was captured by number of units of alcohol consumed per week (never drink, <1, males 2-10 and females 2-7, males 11-21 and females 8-14, males 22-35 and females 15-21, males >35 and females >21 and ex-drinker). The categories used reflected the survey categorisation for the sex specific alcohol consumption variables (ALCBASMT and ALCBASWT) in the SHeS years 1998, 2003, 2008, 2009, 2010 and 2011 (Appendix 5.3). These in turn reflected the Scottish Government public health policy at the time which described “hazardous drinking”, or a drinking level that may cause harm in the

future, as between 21-50 units a week for men and 14-35 units for women (Scottish Government 2013a). Hazardous drinking included "binge drinking" defined as excessive alcohol consumption on any one occasion and was set at more than 8 units for men and more than 6 units for women reflecting the then current Scottish Government public health policy (Scottish Government 2013a). The hazardous effect of binge drinking was considered to occur whether or not the weekly limits were breached (Scottish Government 2013a). "Harmful drinking", was defined as drinking behaviour that currently caused evidence of health damage and was defined as consuming more than 50 units per week for men and over 35 units for women (Scottish Government 2013a).

The categories for the 1995 SHeS male and female specific variables for alcohol consumption (ALCOHOLW and ALCOHOLM) differed from all other subsequent surveys. In 1995, the categories were consistent with the UK Royal College of Physicians definition of sensible drinking (21 units/ week for men and 14 units/ week for women; one unit of alcohol was defined as 10 ml (eight grams) of ethanol) (Working Party of the Royal College of Physicians 2001). However to capture the definitions of "hazardous" and "harmful" drinking described above, subsequent survey (1998, 2003, 2008, 2009, 2010 and 2011) alcohol consumption variables were modified. As a result, it was necessary to align the 1995 categories into the most relevant categories used by the alcohol consumption variables (ALCBASMT and ALCASWT) adopted by all the future surveys (1998, 2003, 2008, 2009, 2010 and 2011). A further modification was made by the survey administrators in 2003 to correct how normal and strong beer weekly alcohol units were handled which was reflected thereafter (Joint Health Surveys Unit 2016b).

For the 1995 SHeS, it was relatively easier to align categories for male alcohol consumption to the desired categories compared to that for females. Males who indicated they were an ex-drinker were described as an "Ex-drinker" in the male alcohol consumption variable ALCBASMT, men who indicated they were a non or occasional drinker were described as "Never drink", men who indicated they consumed one to 10 units per week were described as consuming "2-10 units per

week"; and men who reported consuming 11-21 units per week were described as consuming "11-21 units per week". However, those who reported consuming more than 21 units per week were classified as "consuming "22 to 35 units per week" in this study as no further information supported differentiating between those who consumed 22-35 units and those who consumed more than 35 units per week (Appendix 5.3).

For females, the consistency between the 1995 alcohol consumption variable for women (ALCOHOLW) and the future survey variable capturing alcohol consumption (ALCBASWT) was more tenuous. Women who indicated that they were ex-drinkers or non/occasional drinkers or consumed more than 21 units of alcohol per week were categorised as described. However, women who reportedly drank 1-10 units per week were described as consuming "2-7 units per week" in this study and women who indicated they drank between 11-21 units weekly were categorised as drinking "15-21 units per week" (Appendix 5.3).

This study captured harmful binge drinking via the maximum alcohol units consumed in a single day over the last seven days (binge rate/day males <=8 and females <=6, males >8 and females >6) and reflected the Scottish Government public health policy at the time (Scottish Government 2013a). In 1995 no suitable variable was available to assess binge drinking rate; as a result, no variable could be derived. In 1998 and 2003, named variable D7GROUP was added to reflect "units drunk on the heaviest day in past week". In all subsequent surveys (2008, 2009, 2010 and 2011) the named variable D7UT08_2 was available. Like D7GROUP, D7UT08_2 identified "Units drunk on heaviest day in past week" but made revisions associated with alcopops and wine (Scottish Centre for Social Research 2016a).

For modelling purposes, the reference category for each alcohol consumption variable was: weekly alcohol units consumed (never drink) and binge units per day (<= binge rate/day M: 8, F: 6).

#### 5.2.3.3 Diet

Diet was measured via fruit and vegetable portions consumed per day. The categories reflected the WHO and national recommendations to consume five portions or more of fruit and vegetables daily (>=4, 2 or 3, 1 or never) (Williams 1995; WHO 2003).

The Scottish Health Surveys in 1995 and 1998 had no variable to measure total daily consumption of fruits and vegetables together. Furthermore, the time period for fruit or vegetable consumption varied (daily, weekly, or monthly). As a result, daily fruit (FRUIT) or vegetable (GREENVEG) consumption was calculated separately by dividing by the relevant number of days (one, seven, or 28 respectively) then combining the two. For all subsequent surveys (2003, 2008, 2009, 2010 and 2011), daily fruit and vegetable consumption combined was available (PORFV) (Appendix 5.3).

For modelling purposes, the reference category for daily fruit and vegetable portions consumed was defined as >=4.

#### 5.2.3.4 Exercise

The study variable capturing level of exercise carried out in a week reflected the number of 30 minute sessions per week. In 1995, the variable (EXNUM) recorded the number of times per week the participant exercised; however, no assessment of time was indicated. In 1998, the physical exercise variable (ADTOT30C) recorded the number of 30 minute sessions or more in a week that a participant exercised while in 2003, 2008, 2009, 2010 and 2011 the same variable (ADTOT30C) included multiples of activities lasting 15-29 minutes as well as activities performed for 30 or more minutes. The study variable was grouped into less than five 30 minute sessions weekly and five or more 30 minute sessions weekly reflecting the Physical Activity Guidelines in the UK: Review and Recommendations (Bull et al 2010) (Appendix 5.3).

For modelling purposes, the reference category for weekly exercise sessions was

defined as >=5.

#### 5.2.3.5 Multiple high risk behaviours

In an approach similar to that adopted by Lawder et al (2010), to capture multiple high risk behaviours, a derived variable was defined as the count of the highest risk category for each of the following variables; current smoking status, units of alcohol consumed in a week, daily fruit and vegetable consumption and exercise sessions per week. A minimum score of zero and a maximum score of four were possible (Lawder et al 2010). The multiple high risk behaviours variable also took into consideration order of risk contribution for the behaviours as indicated in the literature (Section 1.7, 2.0), that is: smoking; smoking and alcohol; smoking, alcohol and diet; and finally smoking, alcohol, diet and exercise. As an example, an individual who was a current smoker, consumed more units of alcohol in a week than the recommended limit and never ate fruit or vegetables would have a multiple high risk behaviours score value of 3 (Table 5.3).

**Table 5.3 Definition of multiple high risk behaviours**

| Risk Behaviour | High risk category |
|---|---|
| Smoking status | Current smoker |
| Units of alcohol consumed per week | Male: > 35 units Female: > 21 units |
| Binge units consumed in a day | >binge rate guidance Male: =>8 units Female: => 6 Units |
| Daily fruit and vegetable consumption | <1 or never |
| Exercise sessions per week | <5 episodes per week |

For modelling purposes, the referent category for multiple high risk behaviours was defined as zero.

### 5.2.4 Missing values

Records with missing data for key variables were deleted: SIMD (n=1,247), economic activity (n=999), car ownership (n=5) and housing tenure (n=43). Records with a smoking status indicating that the individual had been a smoker or was a current smoker but also had an illogical smoking duration of “never

smoked" implied incorrect data and were therefore deleted (n=17).

The resulting file (minimum dataset) was used for all further analysis. Remaining observations with missing values for the explanatory behaviour variables were deleted, the number of which depended on the modelling performed and are reported in Tables 5.12-5.14 and 5.18-5.20. This approach supported examination of attenuation by the behaviour variables of the socioeconomic variable association with cancer risk.

## 5.2.5 Ethics approval

Ethics approval for the SHeS was awarded by the NHS Multi-Centre Research Ethics Committee (MREC03/0/19 for 2003; 07/MRE09/55 for 2008; 08/MRE09/62 for 2009-11; SHS 1998 was approved by the Research Ethics Committees for All Health Boards for Scotland; SHS 2003 was approved by the Multi Research Ethics Committee for Scotland). The supply and use of linked data was provided by the Privacy Advisory Committee of National Services Scotland (PAC 100/12 — IR XRB13040). Approval of the study design and approach was provided by the University of Glasgow, College of Medical, Veterinary & Life Sciences Ethics Committee for Non-Clinical Research Involving Human Subjects. Participants gave full informed consent to participate in SHeS and for linkage of SHeS data with other administrative datasets to support secondary studies (Appendices 5.1, 5.2).

## 5.2.6 Modelling

### 5.2.6.1 Minimally adjusted models

The RRs and 95% CI were computed for each cancer group using minimally adjusted Poisson regression models offset by person-years of follow-up adjusted by sex (female as reference) and age group at start of the cohort (16-44 (referent group), 45-59, 60-74 and 75+ years old). Models were developed for each of the five individual socioeconomic status variables, one area socioeconomic status variable, the eight behaviour variables and the two derived

variables capturing multiple circumstances (high risk behaviours and low socioeconomic circumstances).

All significant socioeconomic and behavioural variables where 95% CI excluded 1.0 were ranked from high to low based on the RRs for each cancer group.

#### 5.2.6.2 Identification of variables with greatest predictive value

In order to identify which of the four smoking and two alcohol variables had the greatest predictive value for each of the cancer groups, logistic regression models adjusted for age and sex were developed. Logistic regression was used given that forward selection is easier using this method compared to Poisson regression in the SAS version 9.1, the statistical package used. Furthermore, the logistic regression function also provides the C-index (Wilford et al 2008)

Using forward selection, the relevant smoking and alcohol variables were identified by applying a 0.2 significance level for entry and a 0.25 significance level for the variable to remain in the model. These significance levels are commonly applied for selection of potential confounders in exposure or disease models (Greenland et al 1989).

#### 5.2.6.3 Evaluation of interactions

Two factors are interactive when the effect of one factor varies depending on the level of the other factor. That is, the two factors do not act independently on the dependent variable (Petrie et al 2009). For this study, the dependent or response variable is all cancer or LUADT cancer risk.

Previous studies have established an interaction between sex and socioeconomic circumstances influencing cancer risk (Caracta 2003; Campos-Matos et al 2016). The study described in Chapter 3 also identified differences between socioeconomic circumstances and males and females in Scotland (Section 3.4); however, due to the number of cases in the cohort, the sexes were combined (Section 5.2.1). As a result, data were reviewed to establish any evidence of

interactions between sex and socioeconomic circumstances.

Other studies demonstrated an interactive effect between smoking and alcohol (Hashibe et al 2009; Prabhu et al 2014), therefore, *a priori* evidence of such an effect between the selected alcohol and smoking variables identified via the forward selection logistic regression modelling for each cancer group were also reviewed.

Interactions between sex and socioeconomic circumstances and the most important smoking and alcohol variables that were identified via forward selection modelling were evaluated for each cancer group using Poisson regression models and the likelihood ratio using a 0.05 significance level. The Poisson model that was developed to test for smoking and alcohol interactions for LUADT cancer risk was corrected for under dispersion; no correction was required for the all cancer group.

#### 5.2.6.4 Identification of socioeconomic risk attenuation by behaviour variables

Due to the interest in each of the socioeconomic variables, models of each socioeconomic variable adjusted for age and sex for each cancer group were used. This minimally adjusted model for each of the seven socioeconomic variables was then further adjusted successively for smoking; smoking and alcohol; smoking, alcohol and diet; and smoking, alcohol, diet and exercise to establish to what extent the associated RRs were attenuated. The variables used for smoking and alcohol behaviour for each cancer group were determined by forward selection modelling as described in Section 5.2.6.2. Identified interactions specific to each cancer group were also included in the Poisson regression models developed (Section 5.2.6.3).

Relative risks (RR) with 95% CI that did not include the value of 1.0 were regarded as statistically significant. All statistical analyses were undertaken using SAS version 9.1(SAS Institute Inc. USA).

## 5.3 Results

### 5.3.1 Cohort description

There were 42,983 individuals aged 16 years old or older who had participated in one of the SHeS that were conducted in 1995, 1998, 2003, 2008, 2009, 2010, or 2011 in the cohort. First primary cancers were identified via the linkage with the SCR from 1975 to 2011. The linkage created the following results:

- 2130 cancers (5.0%),
- 453 LUADT cancers (1.1%),
- 40853 non-cancers (95.0%).

The study population consisted of fewer males (44%) compared to the 2011 Scottish population as a whole (51.5%) (National Records Scotland 2015e). The study cohort age distribution also differed slightly with only 44% in the 16 to 44 age group compared to 47% in the 2011 Scottish population; however, the proportion of those under 60 years old (72%) and the proportion of those 60 or more years old (28%) was the same for both the study cohort and 2011 Scottish population (National Records Scotland 2015b). The proportion "employed" in the study population was only 56% compared to the 69% in the 2011 census (National Records Scotland 2015a); this may reflect greater agreement to participate in the SHeS as well as greater availability of time by retirees. The proportion holding a first degree or higher degree was only 20% in the study cohort, but was 26% in the 2011 Scottish population aged 16 years old and older (National Records Scotland 2015c). However, the proportion with no degree was only 20% in the study cohort compared to 27% for the comparably aged 2011 Scottish population (National Records Scotland 2015c). Overall, 80% of the study population held some level of education compared to only 73% in the 2011 Scottish census of comparable age (National Records Scotland 2015c). With respect to the area deprivation levels, the study population appeared less affluent with only 17.9% in the least deprived category and 21.3% in the most deprived category compared to 20.3% least deprived and 19.5% most deprived

categories in the 2011 Scottish population of comparable age (National Records Scotland 2015b). Seventy-five percent of the study population had access to one or more car(s) while the proportion was only 69.5% of the 2011 Scottish census (National Records Scotland 2014b). According to the 2011 census, 17% of all employed people were in “professional occupations”, 13% in “associate professional and technical occupations” and 13% in “skilled trade occupations”. Although not exactly the same classification as used in the study population, this compared less favourably to the study population (35% professional, managerial and technical, 15% skilled non manual and 24% skilled manual) (National Records Scotland 2014c). Likewise with respect to housing tenure, a greater proportion of the study population lived in their own home (67%) compared to the 2011 Scottish population (62%) (National Records Scotland 2014a). On the whole, the study cohort appeared to be more affluent than the 2011 Scottish population in terms of educational attainment, social class, car and house ownership, but less affluent in terms of area deprivation. However, the age distribution overall was broadly similar and differences in economic activity appeared neutral overall. Nevertheless, the proportion of women, who were less likely to be diagnosed with the cancers in question, was greater (ISD 2015) (Table 5.4).

## 5.3.2 All cancer risk

### 5.3.2.1 Behavioural risks (minimally adjusted models)

Current smokers were at a 43% greater risk of a cancer diagnosis compared to non-smokers (RR 1.43 95% CI 1.29, 1.59). The more cigarettes smoked daily, the greater the risk of a cancer diagnosis such that those who smoked less than 20 cigarettes per day had a reduced risk (RR 1.26 95% CI 1.10, 1.45) compared to those who smoked 20 or more cigarettes daily (RR 1.50 95% CI 1.30, 1.74). The age that an individual started smoking was not significant with confidence intervals including 1.0. However, the duration smoked was strongly associated with a greater risk of cancer diagnosis associated with the longer the period smoked; the relationship presented a very clear increasing gradient. Compared to those who smoked 20 years or less, those who smoked more than 50 years had a greater than four-fold risk of a cancer diagnosis (RR 4.42 95% CI 3.10, 6.30).

**Table 5.4 Cohort number (N) and percent (%) by demographic and socioeconomic factors by cancer group, Scotland 1995-2011**

| | | All cancer[1] | | LUADT[2] | | No Cancer | | Full cohort | |
|---|---|---|---|---|---|---|---|---|---|
| Variable | Category | N | % | N | % | N | % | N | % |
| Sex | Male | 940 | 5.0 | 229 | 1.2 | 17960 | 95.0 | 18900 | 100.0 |
| | Female | 1190 | 4.9 | 224 | 0.9 | 22893 | 95.1 | 24083 | 100.0 |
| | Total | 2130 | 5.0 | 453 | 1.1 | 40853 | 95.0 | 42983 | 100.0 |
| Age at survey | 16-44 | 261 | 1.4 | 28 | 0.1 | 18664 | 98.6 | 18925 | 100.0 |
| | 45-59 | 673 | 5.6 | 140 | 1.2 | 11301 | 94.4 | 11974 | 100.0 |
| | 60-74 | 990 | 10.6 | 233 | 2.5 | 8306 | 89.4 | 9296 | 100.0 |
| | 75+ | 206 | 7.4 | 52 | 1.9 | 2582 | 92.6 | 2788 | 100.0 |
| SIMD | 1 Least deprived | 372 | 4.8 | 73 | 0.9 | 7315 | 95.2 | 7687 | 100.0 |
| | 2 | 329 | 4.2 | 56 | 0.7 | 7590 | 95.8 | 7919 | 100.0 |
| | 3 | 484 | 5.3 | 87 | 1.0 | 8661 | 94.7 | 9145 | 100.0 |
| | 4 | 487 | 5.4 | 106 | 1.2 | 8587 | 94.6 | 9074 | 100.0 |
| | 5 Most deprived | 458 | 5.0 | 131 | 1.4 | 8700 | 95.0 | 9158 | 100.0 |
| Economic activity | Employed | 758 | 3.1 | 115 | 0.5 | 23497 | 96.9 | 24255 | 100.0 |
| | Unemployed | 81 | 2.8 | 24 | 0.8 | 2804 | 97.2 | 2885 | 100.0 |
| | Retired | 918 | 9.8 | 222 | 2.4 | 8491 | 90.2 | 9409 | 100.0 |
| | Other economically inactive | 373 | 5.8 | 92 | 1.4 | 6061 | 94.2 | 6434 | 100.0 |
| Highest qualification | First degree and higher | 213 | 2.5 | 25 | 0.3 | 8244 | 97.5 | 8457 | 100.0 |
| | Other non-degree | 1263 | 4.9 | 271 | 1.0 | 24712 | 95.1 | 25975 | 100.0 |
| | None | 651 | 7.7 | 157 | 1.8 | 7856 | 92.3 | 8507 | 100.0 |
| | Missing | 3 | 6.8 | 0 | 0.0 | 41 | 93.2 | 44 | 100.0 |
| Social class | I,II Professional, managerial, technical | 657 | 4.4 | 110 | 0.7 | 14425 | 95.6 | 15082 | 100.0 |
| | III N Skilled non manual | 342 | 5.0 | 58 | 0.8 | 6500 | 95.0 | 6842 | 100.0 |
| | III M Skilled manual | 542 | 5.2 | 130 | 1.3 | 9838 | 94.8 | 10380 | 100.0 |
| | IV Partly skilled | 360 | 5.4 | 88 | 1.3 | 6299 | 94.6 | 6659 | 100.0 |
| | VI Unskilled | 176 | 6.7 | 54 | 2.1 | 2454 | 93.3 | 2630 | 100.0 |
| | Missing | 53 | 3.8 | 13 | 0.9 | 1337 | 96.2 | 1390 | 100.0 |
| Number of cars | 1 or more car(s) | 1436 | 4.5 | 231 | 0.7 | 30732 | 95.5 | 32168 | 100.0 |
| | No car | 694 | 6.4 | 222 | 2.1 | 10121 | 93.6 | 10815 | 100.0 |
| Housing tenure | Owner occupier | 1345 | 4.7 | 213 | 0.7 | 27345 | 95.3 | 28690 | 100.0 |
| | Rent LA | 577 | 9.1 | 194 | 3.0 | 5789 | 90.9 | 6366 | 100.0 |
| | Rent privately | 208 | 2.6 | 46 | 0.6 | 7719 | 97.4 | 7927 | 100.0 |
| Multiple SES | 0 No deprivation | 708 | 3.8 | 104 | 0.6 | 17832 | 96.2 | 18540 | 100.0 |
| | 1 | 624 | 5.5 | 102 | 0.9 | 10752 | 94.5 | 11376 | 100.0 |
| | 2 | 442 | 6.3 | 116 | 1.6 | 6619 | 93.7 | 7061 | 100.0 |
| | 3 | 265 | 6.5 | 95 | 2.3 | 3839 | 93.5 | 4104 | 100.0 |
| | 4-6 Low socioeconomic | 91 | 4.8 | 36 | 1.9 | 1811 | 95.2 | 1902 | 100.0 |

[1] All cancer excluding non-melanoma skin cancer

[2] LUADT (Lung and upper aero-digestive tract cancers)

**Table 5.5 Cohort number (N) and percent (%) by behavioural factor by cancer group, Scotland 1995-2011**

| | | All cancer[1] | | LUADT[2] | | No Cancer | | Full cohort | |
|---|---|---|---|---|---|---|---|---|---|
| Variable | Category | N | % | N | % | N | % | N | % |
| Smoking status | Never smoked | 549 | 3.4 | 39 | 0.2 | 1580 | 96.6 | 16355 | 100.0 |
| | Ex-smoker | 712 | 5.9 | 119 | 1.0 | 1141 | 94.1 | 12131 | 100.0 |
| | Current | 766 | 6.1 | 290 | 2.3 | 1173 | 93.9 | 12497 | 100.0 |
| | Missing | 103 | 5.2 | 5 | 0.3 | 1897 | 94.9 | 2000 | 100.0 |
| Number of cigarettes smoked | Never Smoked | 1116 | 6.6 | 39 | 0.2 | 1580 | 93.4 | 16922 | 100.0 |
| | Non-smoker | 567 | 5.3 | 104 | 1.0 | 1008 | 94.7 | 10654 | 100.0 |
| | <21 | 1451 | 18.2 | 98 | 1.2 | 6513 | 81.8 | 7964 | 100.0 |
| | >=21 | 1766 | 31.6 | 105 | 1.9 | 3831 | 68.4 | 5597 | 100.0 |
| | Don't know | 2027 | 42.7 | 102 | 2.1 | 2719 | 57.3 | 4746 | 100.0 |
| | Missing | 2130 | 52.9 | 5 | 0.1 | 1897 | 47.1 | 4027 | 100.0 |
| Age started smoking | 0-9 | 34 | 11.0 | 11 | 3.5 | 276 | 89.0 | 310 | 100.0 |
| | 10-19 | 1045 | 6.0 | 323 | 1.8 | 1646 | 94.0 | 17514 | 100.0 |
| | 20-29 | 250 | 6.6 | 61 | 1.6 | 3537 | 93.4 | 3787 | 100.0 |
| | 30-39 | 45 | 9.3 | 5 | 1.0 | 441 | 90.7 | 486 | 100.0 |
| | >=40 | 24 | 9.2 | 2 | 0.8 | 237 | 90.8 | 261 | 100.0 |
| | Never smoker | 549 | 3.4 | 39 | 0.2 | 1580 | 96.6 | 16355 | 100.0 |
| | Missing | 183 | 4.3 | 12 | 0.3 | 4087 | 95.7 | 4270 | 100.0 |
| Years | 1-20 | 84 | 1.4 | 27 | 0.4 | 6025 | 98.6 | 6109 | 100.0 |
| | 21-30 | 149 | 3.5 | 88 | 2.1 | 4102 | 96.5 | 4251 | 100.0 |
| | 31-40 | 332 | 7.3 | 145 | 3.2 | 4231 | 92.7 | 4563 | 100.0 |
| | 41-50 | 437 | 10.8 | 136 | 3.4 | 3604 | 89.2 | 4041 | 100.0 |
| | >50 | 394 | 11.9 | 39 | 1.2 | 2924 | 88.1 | 3318 | 100.0 |
| | Never Smoker | 549 | 3.4 | 12 | 0.1 | 1580 | 96.6 | 16355 | 100.0 |
| | Missing | 185 | 4.3 | 6 | 0.1 | 4161 | 95.7 | 4346 | 100.0 |
| | Never drink | 125 | 6.1 | 24 | 1.2 | 1936 | 93.9 | 2061 | 100.0 |
| Alcohol units/ week | Ex-drinker | 126 | 5.3 | 35 | 1.5 | 2249 | 94.7 | 2375 | 100.0 |
| | =<1 per week | 280 | 5.9 | 50 | 1.0 | 4504 | 94.1 | 4784 | 100.0 |
| | M:2 - 10 F:2-7 | 527 | 4.2 | 97 | 0.8 | 1198 | 95.8 | 12511 | 100.0 |
| | M:11 - 21 F:8-14 | 283 | 3.9 | 55 | 0.8 | 6933 | 96.1 | 7216 | 100.0 |
| | M:22 - 35 F:15-21 | 154 | 3.7 | 38 | 0.9 | 3954 | 96.3 | 4108 | 100.0 |
| | M:>35 F:>21 | 77 | 3.4 | 19 | 0.8 | 2195 | 96.6 | 2272 | 100.0 |
| | M:>35 F:>21 | 71 | 4.6 | 20 | 1.3 | 1456 | 95.4 | 1527 | 100.0 |
| | Missing | 487 | 7.9 | 115 | 1.9 | 5642 | 92.1 | 6129 | 100.0 |
| Binge units/ day | <=binge rate/day M: 8 F:6 | 907 | 3.7 | 178 | 0.7 | 2357 3 | 96.3 | 24480 | 100.0 |
| | >binge rate/day M: 8 F:6 | 243 | 3.3 | 61 | 0.8 | 7151 | 96.7 | 7394 | 100.0 |
| | Missing | 980 | 8.8 | 214 | 1.9 | 1012 | 91.2 | 11109 | 100.0 |
| Fruit and vegetable consumption | >=4 | 286 | 3.1 | 37 | 0.4 | 8840 | 96.9 | 9126 | 100.0 |
| | 2 or 3 | 612 | 5.6 | 104 | 0.9 | 1040 | 94.4 | 11014 | 100.0 |
| | 1 | 519 | 7.2 | 112 | 1.5 | 6739 | 92.8 | 7258 | 100.0 |
| | <1 or never | 496 | 5.8 | 148 | 1.7 | 7993 | 94.2 | 8489 | 100.0 |
| | Missing | 217 | 3.1 | 52 | 0.7 | 6879 | 96.9 | 7096 | 100.0 |
| Exercise episodes/ week | <5 episodes/ | 1351 | 5.5 | 302 | 1.2 | 2343 | 94.5 | 24787 | 100.0 |
| | >=5 episodes/ | 322 | 2.5 | 40 | 0.3 | 1278 | 97.5 | 13104 | 100.0 |
| | Missing | 457 | 9.0 | 111 | 2.2 | 4635 | 91.0 | 5092 | 100.0 |

[1] All cancer excluding non-melanoma skin cancer

[2] LUADT (Lung and upper aero-digestive tract cancers)

**Table 5.6 Cohort number (N) and percent (%) for multiple risk behaviours by cancer group, Scotland 1995-2011**

| | | All cancer | | LUADT | | No cancer | | Full cohort | |
|---|---|---|---|---|---|---|---|---|---|
| Variable | Category | N | % | N | % | N | % | N | % |
| Multiple risk behaviours | 0 No high risk behaviour | 1364 | 4.5 | 163 | 0.5 | 29122 | 95.5 | 30486 | 100 |
| | 1 Smoking[1] | 637 | 6.6 | 238 | 2.5 | 9031 | 93.4 | 9668 | 100 |
| | 2 Smoking[1] Alcohol[2] | 82 | 4.4 | 33 | 1.8 | 1798 | 95.6 | 1880 | 100 |
| | 3 or 4 Smoking[1] Alcohol[2] Diet[3] Exercise[4] | 47 | 5.0 | 19 | 2.0 | 902 | 95.0 | 949 | 100 |

High risk behaviour defined as:
[1] Smoking status= current smoker.
[2]Units of alcohol consumed in a day= Male: >35 units, Female:> 21 units
[3]Daily fruit and vegetable consumption <1 or never
[4]Exercise sessions per week <5

While the number of alcohol units consumed per week was not associated with a greater risk of cancer diagnosis, those who drank more than eight or six units of alcohol daily (males or females respectively) were at a 20% greater risk of a cancer diagnosis (RR 1.20 95% CI 1.01, 1.43). Compared to those who consumed four or more portions of fruits and vegetables daily, those who ate only one portion or less increased their risk of a cancer diagnosis by 30-36% (1 portion daily: RR 1.30 95% CI 1.08, 1.56; less than 1 or never: RR 1.36 95% CI 1.13, 1.63). Individuals who did not manage to exercise the recommended threshold compared to those that did were at a 39% greater risk of a cancer diagnosis (RR 1.39 95% CI 1.19, 1.62). Finally and compared to those with no high risk behaviours, risk of a cancer diagnosis was elevated by up to 75% for individuals with one or more behaviours in the high risk category with the highest risk among those who were in the highest risk categories for smoking and alcohol (RR 1.75 95% CI 1.39, 2.20) (Table 5.6).

**Table 5.7 Age and sex adjusted behavioural relative risks (RR)[1], 95% confidence intervals (CI) by cancer group, Scotland 1995-2011**

| | | All cancer[2] | | | | LUADT[2] | | | |
|---|---|---|---|---|---|---|---|---|---|
| Variable | Category | *P* - Value | RR | CI | | *P* - Value | RR | CI | |
| Smoking status | Non-smoker | | *Reference* | | | | *Reference* | | |
| | Current | **<0.001** | 1.43 | 1.29 | 1.59 | **<0.001** | 3.81 | 3.05 | 4.75 |
| Cigarettes smoked/ day | Non-smoker | | *Reference* | | | | *Reference* | | |
| | <20 | **0.001** | 1.26 | 1.10 | 1.45 | **<0.001** | 2.47 | 1.86 | 3.28 |
| | >=20 | **<0.001** | 1.50 | 1.30 | 1.74 | **<0.001** | 3.37 | 2.54 | 4.47 |
| Age started smoking | >=40 | | *Reference* | | | | *Reference* | | |
| | <20 | 0.76 | 1.06 | 0.71 | 1.60 | **0.041** | 4.26 | 1.06 | 17.15 |
| | 20-39 | 0.55 | 0.88 | 0.58 | 1.34 | 0.215 | 2.44 | 0.60 | 9.95 |
| Duration smoked (years) | 1-20 | | *Reference* | | | | *Reference* | | |
| | 21-40 | **<0.001** | 2.13 | 1.59 | 2.85 | **<0.001** | 7.11 | 2.88 | 17.54 |
| | 41-50 | **<0.001** | 3.15 | 2.26 | 4.39 | **<0.001** | 16.98 | 6.55 | 44.01 |
| | >50 | **<0.001** | 4.42 | 3.10 | 6.30 | **<0.001** | 26.80 | 10.06 | 71.42 |
| Units Alcohol/ week | Never drink | | *Reference* | | | | *Reference* | | |
| | =<1 per week | 0.582 | 0.92 | 0.68 | 1.24 | **0.048** | 0.59 | 0.35 | 1.00 |
| | M:2-10 F:2-7 | 0.795 | 0.96 | 0.73 | 1.28 | 0.092 | 0.66 | 0.41 | 1.07 |
| | M:11-21 F:8-14 | 0.487 | 0.90 | 0.66 | 1.22 | 0.059 | 0.60 | 0.36 | 1.02 |
| | M:22-35 F:15-21 | 0.852 | 0.97 | 0.70 | 1.34 | 0.519 | 0.83 | 0.48 | 1.45 |
| | M:>35 F:>21 | 0.857 | 1.03 | 0.74 | 1.43 | 0.884 | 0.96 | 0.55 | 1.67 |
| | Ex-drinker | 0.865 | 0.97 | 0.70 | 1.35 | 0.577 | 0.86 | 0.49 | 1.48 |
| Binge units/ day | <=binge rate/day M:8 F:6 | | *Reference* | | | | *Reference* | | |
| | >binge rate/day M:8 F:6 | **0.039** | 1.20 | 1.01 | 1.43 | **0.011** | 1.53 | 1.10 | 2.12 |
| Fruit and vegetable consumption/ day | >=4 | | *Reference* | | | | *Reference* | | |
| | 2 or 3 | 0.26 | 1.11 | 0.93 | 1.33 | 0.369 | 1.21 | 0.80 | 1.82 |
| | 1 | **0.01** | 1.30 | 1.08 | 1.56 | **0.002** | 1.90 | 1.27 | 2.84 |
| | <1 or never | **0.00** | 1.36 | 1.13 | 1.63 | **<0.001** | 2.66 | 1.80 | 3.93 |
| Exercise sessions/ week | >= 5 episodes/ week | | *Reference* | | | | *Reference* | | |
| | <5 episodes/ week | **<0.001** | 1.39 | 1.19 | 1.62 | **<0.001** | 2.26 | 1.58 | 3.25 |
| Multiple behaviours | 0 No high risk behaviour | | *Reference* | | | | *Reference* | | |
| | 1 Smoking | **<0.001** | 1.46 | 1.31 | 1.62 | **<0.001** | 4.06 | 3.25 | 5.06 |
| | 2 Smoking Alcohol | **<0.001** | 1.75 | 1.39 | 2.20 | **<0.001** | 5.24 | 3.56 | 7.73 |
| | 3 or 4 Smoking Alcohol Diet | **0.003** | 1.56 | 1.16 | 2.10 | **<0.001** | 4.87 | 2.98 | 7.95 |

[1]Relative Risk (RR), 95% CI estimated using age at diagnosis Poisson regression models offset by person-years of follow-up adjusted by sex (female as reference) and age group at start of the cohort (16-44 (reference), 45-59, 60-74, 75+ years old)**, Bold** indicates CI did not include 1.0

[2] All cancer excluding non-melanoma skin cancer

[3]LUADT (Lung and upper aero-digestive tract cancers)

#### 5.3.2.2 Socioeconomic risks (minimally adjusted models)

Compared to those in work, individuals who were retired were 43% more likely to be diagnosed with cancer (RR 1.43 95% CI 1.22, 1.68) while those who were otherwise economically inactive were at 17% greater risk of a cancer diagnosis (RR 1.17 95% CI 1.02, 1.35). Those who had no education were 38% more likely to be diagnosed with cancer compared to those with a first or higher degree (RR 1.38 95% CI 1.13, 1.69). Likewise, those living in accommodation rented from the local authority or housing association had an 18% greater risk of a cancer diagnosis compared to those who owned their own home (RR 1.18 95% CI 1.06, 1.32) (Table 5.7).

Finally and compared to those with no socioeconomic factors in the highest risk category, those with any socioeconomic factors in the highest risk category had at least a 16% greater risk of a cancer diagnosis (RR 1.16 95% CI 1.02, 1.32). This greater risk increased with each additional high risk socioeconomic factor; those with four to six highest socioeconomic factors had a 47% greater risk of cancer incidence (RR 1.47 95% CI 1.16, 1.86). The area measure of deprivation, SIMD was not statistically significant for any category with RRs hovering around 1.0 (RR range 1.08 to 1.01) (Table 5.7).

#### 5.3.2.3 Relative risks associated with behaviours and SES (minimally adjusted model) ranking

Ranking the highest category for each variable where the relative risk confidence interval excluded 1.0, identified 11 variables with increased all cancer risk association where seven of those variables are behaviour and four socioeconomic with 4-6 multiple low socioeconomic circumstances ranking 5$^{th}$ – a higher ranking than occurred for the LUADT cancer group (6$^{th}$, Table 5.14). Duration smoked ranked first and with a RR of 4.42 dominated the other risks, although all cancer RRs were much lower than that for LUADT cancer risk (26.8 versus 4.42) (Tables 5.13, 5.8).

**Table 5.8 Age and sex adjusted socioeconomic factors relative risks (RR)[1] and 95% confidence intervals (CI) by cancer group, Scotland 1995-2011**

| | | All cancer[2] | | | | LUADT[3] | | | |
|---|---|---|---|---|---|---|---|---|---|
| Variable | Category | *P* - Value | RR | CI | | *P* - Value | RR | CI | |
| Economic activity | Employed | | *Reference* | | | | *Reference* | | |
| | Unemployed | 0.092 | 1.24 | 0.97 | 1.58 | **0.003** | 1.97 | 1.25 | 3.10 |
| | Retired | **<0.001** | 1.43 | 1.22 | 1.68 | **<0.001** | 1.96 | 1.42 | 2.68 |
| | Other economically inactive | **0.030** | 1.17 | 1.02 | 1.35 | **0.001** | 1.64 | 1.22 | 2.20 |
| SIMD | 1 Least deprived | | *Reference* | | | | *Reference* | | |
| | 2 | 0.398 | 1.08 | 0.90 | 1.29 | 0.879 | 1.03 | 0.71 | 1.48 |
| | 3 | 0.203 | 1.11 | 0.95 | 1.30 | 0.906 | 1.02 | 0.74 | 1.41 |
| | 4 | 0.711 | 1.03 | 0.88 | 1.21 | 0.475 | 1.12 | 0.82 | 1.54 |
| | 5 Most deprived | 0.893 | 1.01 | 0.86 | 1.18 | **0.023** | 1.42 | 1.05 | 1.93 |
| Highest qualification | First degree and higher | | *Reference* | | | | *Reference* | | |
| | Other non-degree | 0.080 | 1.18 | 0.98 | 1.43 | **0.013** | 1.78 | 1.13 | 2.80 |
| | None | **0.002** | 1.38 | 1.13 | 1.69 | **0.000** | 2.33 | 1.46 | 3.74 |
| Social class | I,II Professional, managerial, technical | | *Reference* | | | | *Reference* | | |
| | III M Skilled manual | 0.778 | 1.02 | 0.89 | 1.17 | **0.051** | 1.31 | 1.00 | 1.71 |
| | IV Partly skilled | 0.820 | 1.02 | 0.88 | 1.18 | **0.028** | 1.39 | 1.04 | 1.87 |
| | VI Unskilled | 0.529 | 1.06 | 0.88 | 1.28 | **0.001** | 1.74 | 1.24 | 2.43 |
| Number of cars | 1 or more car(s) | | *Reference* | | | | *Reference* | | |
| | No car | 0.052 | 1.11 | 1.00 | 1.23 | **<0.001** | 2.05 | 1.69 | 2.50 |
| Housing tenure | Owner occupier | | *Reference* | | | | *Reference* | | |
| | Rent LA | **0.004** | 1.18 | 1.06 | 1.32 | **<0.001** | 2.37 | 1.93 | 2.90 |
| | Rent privately | 0.243 | 1.10 | 0.94 | 1.30 | 0.056 | 1.38 | 0.99 | 1.93 |
| Multiple SES | 0 No deprivation | | *Reference* | | | | *Reference* | | |
| | 1 | **0.023** | 1.16 | 1.02 | 1.32 | 0.104 | 1.27 | 0.95 | 1.70 |
| | 2 | **0.022** | 1.18 | 1.02 | 1.35 | **<0.001** | 2.01 | 1.52 | 2.66 |
| | 3 | **0.005** | 1.26 | 1.07 | 1.48 | **<0.001** | 2.81 | 2.09 | 3.76 |
| | 4-6 High multiple deprivation | **0.002** | 1.47 | 1.16 | 1.86 | **<0.001** | 3.35 | 2.26 | 4.97 |

[1]Relative Risk (RR), 95% CI estimated using age at diagnosis Poisson regression models offset by person-years of follow-up adjusted by sex (female as reference) and age group at start of the cohort (16-44 (referent group), 45-59, 60-74 and 75+ years old)**, Bold** indicates CI did not include 1.0
[2] All cancer excluding non-melanoma skin cancer
[3]LUADT (Lung and upper aero-digestive tract cancers)

The variables duration smoked and multiple high risk behaviour were ranked 1st and 2nd for both cancer endpoints. Although position varied between the two cancer groups, cigarettes smoked, economic activity, multiple low socioeconomic circumstances, smoking status, exercise sessions per week, highest qualification, fruit and vegetable consumption, binge units and housing tenure were statistically significant in both cancer groups. Confidence intervals for age started smoking, car ownership and occupational social class excluded 1.0 for LUADT cancer only (Tables 5.8, 5.13).

**Table.5.9 Ranking of greatest significant relative risks (RR) for all cancer group, Scotland 1995-2011**

| | Characteristics | Category | RR | Focus |
|---|---|---|---|---|
| 1 | Duration smoked (years) | >50 | 4.42 | Behaviour |
| 2 | Multiple high risk behaviours | 2 Smoking Alcohol | 1.75 | Behaviour |
| 3 | Cigarettes smoked/day | >=20 | 1.50 | Behaviour |
| 4 | Economic activity | Retired | 1.43 | Socioeconomic circumstances |
| 5 | Multiple low socioeconomic circumstances | 4-6 Low socioeconomic circumstances | 1.47 | Socioeconomic circumstances |
| 6 | Smoking status | Current | 1.43 | Behaviour |
| 7 | Exercise sessions /week | <5 episodes/week | 1.39 | Behaviour |
| 8 | Highest qualification | None | 1.38 | Socioeconomic circumstances |
| 9 | Fruit and vegetable consumption/day | <1 or never | 1.36 | Behaviour |
| 10 | Binge units/day | >binge rate/day M:8 | 1.20 | Behaviour |
| 11 | Housing tenure | Rent Local Authority accommodation | 1.18 | Socioeconomic circumstances |

### 5.3.2.4 Identification of variables with greatest predictive value

The forward stepwise logistic regression model for all cancer risk included all four smoking variables and the two alcohol variables based on the likelihood ratio test where *P*<0.20 suggesting that these variables contributed to the model (Table 5.10).

**Table 5.10 Likelihood ratio test statistics[1] for smoking and alcohol variables for all cancer risk**

| Focus | Step | Variable | DF | Chi-square | Pr >ChiSq |
|---|---|---|---|---|---|
| Alcohol variables | 1 | Units alcohol/ week | 6 | 63.052 | <.0001 |
| | 2 | Binge units/ day | 1 | 11.1767 | 0.0008 |
| Smoking variables | 1 | Cigarettes smoked / day | 2 | 91.8587 | <.0001 |
| | 2 | Smoking status | 1 | 62.7926 | <.0001 |
| | 3 | Smoking duration | 4 | 33.5798 | <.0001 |
| | 4 | Age started smoking | 2 | 14.2316 | 0.0008 |

[1]Logistic regression forward selection at 0.02 selection level

### 5.3.2.5 Smoking/alcohol and socioeconomic/sex interactions

The interactions between the four smoking variables, all of which were identified as contributing predictive value for all cancer risk and both alcohol variables (alcohol units per week and binge units per day) were tested to establish if the effect of either variable on all cancer risk varied according to the level of the other variable; that is, the smoking and alcohol explanatory variables did not act independently on all cancer risk, the dependent variable. Sex and socioeconomic interactions for each of the seven socioeconomic variables were also tested using Poisson regression models.

Using Chi-square probability < 0.05, results indicated that none of the interactions between any smoking variable and either of the two alcohol variables were significant (*P*-values ranging from 0.146 to 0.972). Likewise, none of the interactions tested between sex and the seven socioeconomic variables proved significant (*P*-values ranging from 0.121 to 0.804). This suggested that these variables did not have a modifying effect on the all cancer risk and that they acted independently (Table 5.11).

**Table 5.11 Likelihood ratio test for smoking and alcohol and sex and socioeconomic variable interactions for all cancer risk**

| Interaction | Chi Square | *P*-value[1] |
|---|---|---|
| Cigarettes smoked/ day * alcohol units/ day | 10.08 | 0.523 |
| Cigarettes smoked/ day * binge units/ day | 0.44 | 0.801 |
| Smoking duration * alcohol units/ day | 29.99 | 0.185 |
| Smoking duration * binge units/ day | 6.82 | 0.146 |
| Smoking status * alcohol units/ day | 1.59 | 0.902 |
| Smoking status * binge units/ day | 0 | 0.972 |
| Age started smoking * binge units / day | 3.62 | 0.164 |
| | | |
| Economic activity * sex | 6.53 | 0.088 |
| SIMD * sex | 1.63 | 0.804 |
| Highest qualification * sex | 4.23 | 0.121 |
| Occupational social class * sex | 5.57 | 0.234 |
| Car ownership * sex | 0.21 | 0.648 |
| Housing tenure * sex | 1.5 | 0.472 |
| Multiple high deprivation * sex | 4.76 | 0.191 |
| Multiple high risk behaviour * sex | 4.44 | 0.350 |

#### 5.3.2.6 Socioeconomic risk attenuation for all cancer risk successively adjusted for behaviour

In the minimally adjusted models, those who were economically inactive or retired had a 17% to 43% elevated risk of all cancer respectively compared to those who were employed (RR 1.17 95% CI 1.02, 1.35 and RR 1.43 95% CI 1.22, 1.68 respectively). The elevated risk for those participants who were economically inactive was attenuated after adjustment for smoking behaviour with confidence intervals including 1.0 (RR 1.17 95% CI 0.97, 1.40). The elevated risks for those who were retired remained significant with confidence intervals excluding 1.0 until adjustment for all behaviours when the risk was just fully attenuated (RR 1.31 95% CI 1.00, 1.73) (Table 5.12).

Compared to those owning their home, the minimally adjusted model demonstrated that the risk of all cancer diagnosis was elevated for those who rented a local authority property (RR 1.18 95% CI 1.06, 1.32). For these participants, the risk association remained elevated after successively adjusting for smoking and alcohol (RR 1.21 95% CI 1.01, 1.45) and was only fully attenuated with confidence intervals including 1.0 after the addition of fruit and vegetable consumption behaviour (RR 1.10 95% CI 0.90, 1.33) (Table 5.12).

Compared to those with no low socioeconomic circumstances in the minimally adjusted model, all participants with at least one situation of the greatest low SES circumstances had an elevated risk association of all cancer (RR range 1.16 to 1.47). For all participants, the elevated risk was fully attenuated after adjustment for smoking or after adjustment for smoking and alcohol (Table 5.13).

For area socioeconomic circumstances as measured by SIMD, highest qualification, occupational social class and car ownership, there were no associations with all cancer risk; all confidence intervals included 1.0 for the minimally adjusted models (Table 5.11, 5.12).

**Table 5.12 Socioeconomic factors relative risks (RR) and 95% confidence intervals (CI) successively adjusted for behaviour for all cancer[1], Scotland 1995-2011**

| | | Age, Sex | | | | + Smoking[2] | | | | + Smoking[2], alcohol[3] | | | | + Smoking[2], alcohol[3], diet | | | | + Smoking[2], alcohol[3], diet, exercise | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Characteristic | Category | *P* | RR | CI | | *P* | RR | CI | | *P* | RR | CI | | *P* | RR | CI | | *P* | RR | CI | |
| Economic | Employed | | *Reference* | | | | *Reference* | | | | *Reference* | | | | *Reference* | | | | *Reference* | | |
| Activity | Unemployed | 0.092 | 1.24 | 0.97 | 1.58 | 0.629 | 1.08 | 0.79 | 1.48 | 0.759 | 1.06 | 0.73 | 1.53 | 0.722 | 1.07 | 0.72 | 1.60 | 0.821 | 1.05 | 0.70 | 1.56 |
| | Retired | **<0.001** | 1.43 | 1.22 | 1.68 | **0.001** | 1.40 | 1.14 | 1.71 | **0.011** | 1.39 | 1.08 | 1.78 | **0.030** | 1.35 | 1.03 | 1.78 | 0.053 | 1.31 | 1.00 | 1.73 |
| | Other economically inactive | **0.030** | 1.17 | 1.02 | 1.35 | 0.094 | 1.17 | 0.97 | 1.40 | 0.360 | 1.12 | 0.88 | 1.44 | 0.516 | 1.09 | 0.84 | 1.42 | 0.742 | 1.05 | 0.80 | 1.37 |
| | Cohort | 26610 | | | | 19756 | | | | 16059 | | | | 12936 | | | | 12924 | | | |
| | Missing | **0** | | | | 6854 | | | | 10551 | | | | 13674 | | | | 13686 | | | |
| SIMD | 1 Least deprived | | *Reference* | | | | *Reference* | | | | *Reference* | | | | *Reference* | | | | *Reference* | | |
| | 2 | 0.398 | 1.08 | 0.90 | 1.29 | 0.618 | 1.05 | 0.86 | 1.29 | 0.966 | 0.99 | 0.77 | 1.28 | 0.873 | 1.02 | 0.78 | 1.34 | 0.772 | 1.04 | 0.79 | 1.37 |
| | 3 | 0.203 | 1.11 | 0.95 | 1.30 | 0.358 | 1.09 | 0.90 | 1.32 | 0.337 | 1.12 | 0.89 | 1.43 | 0.318 | 1.14 | 0.88 | 1.48 | 0.258 | 1.16 | 0.90 | 1.51 |
| | 4 | 0.711 | 1.03 | 0.88 | 1.21 | 0.770 | 0.97 | 0.80 | 1.17 | 0.721 | 0.96 | 0.75 | 1.22 | 0.697 | 0.95 | 0.73 | 1.24 | 0.730 | 0.95 | 0.73 | 1.25 |
| | 5 Most deprived | 0.893 | 1.01 | 0.86 | 1.18 | 0.413 | 0.92 | 0.76 | 1.12 | 0.573 | 0.93 | 0.73 | 1.19 | 0.607 | 0.93 | 0.71 | 1.22 | 0.641 | 0.94 | 0.72 | 1.23 |
| | Cohort | 26610 | | | | 19756 | | | | 16059 | | | | 12936 | | | | 12924 | | | |
| | Missing | 0 | | | | 6854 | | | | 10551 | | | | 13674 | | | | 13686 | | | |
| Highest | First degree and higher | | *Reference* | | | | *Reference* | | | | *Reference* | | | | *Reference* | | | | *Reference* | | |
| Qualification | Other non-degree | 0.207 | 1.16 | 0.92 | 1.45 | 0.207 | 1.16 | 0.92 | 1.45 | 0.784 | 0.97 | 0.75 | 1.24 | 0.652 | 0.94 | 0.71 | 1.24 | 0.647 | 0.94 | 0.71 | 1.24 |
| | None | 0.143 | 1.19 | 0.94 | 1.50 | 0.143 | 1.19 | 0.94 | 1.50 | 0.639 | 1.06 | 0.82 | 1.37 | 0.966 | 0.99 | 0.75 | 1.32 | 0.934 | 0.99 | 0.74 | 1.32 |
| | Cohort | 26581 | | | | 19734 | | | | 16042 | | | | 12920 | | | | 12909 | | | |
| | Missing | 29 | | | | 6876 | | | | 10568 | | | | 13690 | | | | 13701 | | | |

Relative Risk (RR), 95% CI estimated using age at diagnosis Poisson regression models offset by person-years of follow-up adjusted by sex (female as reference) and age group at start of the cohort (16-44 (reference), 45-59, 60-74, 75+ years old)

[1]All cancer excluding non-melanoma skin cancer

[2]Adjusted for sex, age group at start of cohort, cigarettes smoked/ per day, smoking status, smoking duration and age started smoking

[3]Adjusted for sex, age group at start of cohort, cigarettes smoked/ per day, smoking status, smoking duration, age started smoking, units of alcohol/ week and binge units/ day

**Bold** indicates CI did not include 1.0

**Table 5.13 Socioeconomic factors relative risks (RR) and 95% confidence intervals (CI) successively adjusted for behaviour for all cancer[1], Scotland 1995-2011 continued**

| | | Age, Sex | | | | + Smoking[2] | | | | + Smoking[2], alcohol[3] | | | | + Smoking[2], alcohol[3], diet | | | | + Smoking[2], alcohol[3], diet, exercise | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Characteristic | Category | *P* | RR | CI | | *P* | RR | CI | | *P* | RR | CI | | *P* | RR | CI | | *P* | RR | CI | |
| Occupational | I,II Professional, managerial, technical | | *Reference* | | | | *Reference* | | | | *Reference* | | | | *Reference* | | | | *Reference* | | |
| Social Class | III N Skilled non manual | 0.576 | 0.96 | 0.81 | 1.12 | 0.921 | 0.99 | 0.82 | 1.20 | 0.666 | 0.95 | 0.75 | 1.20 | 0.623 | 1.07 | 0.83 | 1.37 | 0.693 | 1.05 | 0.82 | 1.36 |
| | III M Skilled manual | 0.778 | 1.02 | 0.89 | 1.17 | 0.986 | 1.00 | 0.85 | 1.17 | 0.619 | 0.95 | 0.78 | 1.16 | 0.567 | 0.94 | 0.76 | 1.16 | 0.575 | 0.94 | 0.76 | 1.17 |
| | IV Partly skilled | 0.820 | 1.02 | 0.88 | 1.18 | 0.847 | 1.02 | 0.85 | 1.21 | 0.475 | 1.08 | 0.87 | 1.34 | 0.443 | 1.10 | 0.87 | 1.39 | 0.444 | 1.10 | 0.87 | 1.39 |
| | VI Unskilled | 0.529 | 1.06 | 0.88 | 1.28 | 0.793 | 0.97 | 0.77 | 1.22 | 0.873 | 1.02 | 0.76 | 1.38 | 0.441 | 1.13 | 0.83 | 1.54 | 0.510 | 1.11 | 0.81 | 1.51 |
| | Cohort | 25611 | | | | 19205 | | | | 15634 | | | | 12582 | | | | 12570 | | | |
| | Missing | 999 | | | | 7405 | | | | 10976 | | | | 14028 | | | | 14040 | | | |
| Car Ownership | 1 or more car(s) | | *Reference* | | | | *Reference* | | | | *Reference* | | | | *Reference* | | | | *Reference* | | |
| | No car | 0.052 | 1.11 | 1.00 | 1.23 | 0.911 | 1.01 | 0.89 | 1.14 | 0.410 | 1.07 | 0.91 | 1.26 | 0.788 | 1.02 | 0.86 | 1.22 | 0.824 | 1.02 | 0.86 | 1.22 |
| | Cohort | 26610 | | | | 19756 | | | | 16059 | | | | 12936 | | | | 12924 | | | |
| | Missing | 0 | | | | 6854 | | | | 10551 | | | | 13674 | | | | 13686 | | | |
| Housing Tenure | Owner occupier | | *Reference* | | | | *Reference* | | | | *Reference* | | | | *Reference* | | | | *Reference* | | |
| | Rent LA | **0.004** | 1.18 | 1.06 | 1.32 | **0.044** | 1.15 | 1.00 | 1.31 | **0.041** | 1.21 | 1.01 | 1.45 | 0.357 | 1.10 | 0.90 | 1.33 | 0.438 | 1.08 | 0.89 | 1.32 |
| | Rent privately | 0.243 | 1.10 | 0.94 | 1.30 | 0.842 | 1.02 | 0.84 | 1.24 | 0.587 | 0.94 | 0.74 | 1.18 | 0.171 | 0.83 | 0.63 | 1.08 | 0.167 | 0.83 | 0.63 | 1.08 |
| | Cohort | 26610 | | | | 19756 | | | | 16059 | | | | 12936 | | | | 12924 | | | |
| | Missing | 0 | | | | 6854 | | | | 10551 | | | | 13674 | | | | 13686 | | | |

Relative Risk (RR), 95% CI estimated using age at diagnosis Poisson regression models offset by person-years of follow-up adjusted by sex (female as reference) and age group at start of the cohort (16-44 (referent group), 45-59, 60-74, 75+ years old)

[1]All cancer excluding non-melanoma skin cancer

[2]Adjusted for sex, age group at start of cohort, cigarettes smoked/ per day, smoking status, smoking duration and age started smoking

[3]Adjusted for sex, age group at start of cohort, cigarettes smoked/ per day, smoking status, smoking duration, age started smoking, units of alcohol/ week and binge units/ day

**Bold** indicates CI did not include 1.0

**Table 5.14 Socioeconomic factors relative risks (RR) and 95% confidence intervals (CI) successively adjusted for behaviour for all cancer[1], Scotland 1995-2011 continued**

| | | Age, Sex | | | + Smoking[2] | | | + Smoking[2], alcohol[3] | | | + Smoking[2], alcohol[3], diet | | | + Smoking[2], alcohol[3], diet, exercise | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Characteristic | Category | *P* | RR | CI | *P* | RR | CI | *P* | RR | CI | *P* | RR | CI | *P* | RR | CI |
| Multiple Low | 0 No deprivation | *Reference* | | | *Reference* | | | *Reference* | | | *Reference* | | | *Reference* | | |
| Socioeconomic | 1 | **0.023** | 1.16 | 1.02 1.32 | 0.036 | 1.18 | 1.01 1.37 | **0.015** | 1.26 | 1.05 1.53 | 0.091 | 1.20 | 0.97 1.48 | 0.089 | 1.20 | 0.97 1.49 |
| | 2 | **0.022** | 1.18 | 1.02 1.35 | 0.475 | 1.07 | 0.90 1.27 | 0.305 | 1.12 | 0.90 1.40 | 0.661 | 1.06 | 0.83 1.35 | 0.660 | 1.06 | 0.83 1.35 |
| | 3 | **0.005** | 1.26 | 1.07 1.48 | 0.355 | 1.09 | 0.90 1.33 | 0.177 | 1.18 | 0.93 1.50 | 0.560 | 1.08 | 0.83 1.41 | 0.603 | 1.07 | 0.82 1.40 |
| | 4-6 High multiple deprivation | **0.002** | 1.47 | 1.16 1.86 | 0.161 | 1.21 | 0.93 1.58 | 0.100 | 1.31 | 0.95 1.80 | 0.208 | 1.25 | 0.88 1.76 | 0.279 | 1.21 | 0.86 1.71 |
| | Cohort | 26610 | | | 19756 | | | 16059 | | | 12936 | | | 12924 | | |
| | Missing | 0 | | | 6854 | | | 10551 | | | 13674 | | | 13686 | | |

Relative Risk (RR), 95% CI estimated using age at diagnosis Poisson regression models offset by person-years of follow-up adjusted by sex (female as reference) and age group at start of the cohort (16-44 (reference), 45-59, 60-74, 75+ years old)

[1]All cancer excluding non-melanoma skin cancer

[2]Adjusted for sex, age group at start of cohort, cigarettes smoked/ per day, smoking status, smoking duration and age started smoking

[3]Adjusted for sex, age group at start of cohort, cigarettes smoked/ per day, smoking status, smoking duration, age started smoking, units of alcohol/ week and binge units/ day

**Bold** indicates CI did not include 1.0

## 5.3.3 LUADT cancer risk

### 5.3.3.1 Behavioural risks (minimally adjusted models)

For LUADT cancer, the minimally adjusted behaviour risk factor RRs were largely significant with confidence intervals excluding 1.0 with the exception of alcohol units consumed per week, two to three fruit or vegetables consumed daily and age started smoking at 20-39 years old (Table 5.6).

The risk association of LUADT cancer incidence was nearly four-fold higher for those who were current smokers compared to non-smokers (RR 3.81 95% CI 3.05, 4.75); it was more than three-fold greater for those who smoked a greater number of cigarettes daily compared to non-smokers (RR 3.37 95% CI 2.54, 4.47), over four-times greater for those who started smoking at a younger age compared to those who starting the habit at age 40 or older (RR 4.26 95% CI 1.06, 17.15), and most profoundly, almost 27-times greater for those who smoked the greatest number of years relative to non-smokers (RR 26.80 95% CI 10.06, 71.42). The elevated LUADT risks associated with the duration smoked were the highest for all SES factors and behaviours; at multiples of the other calculated relative risks they were dominant (Table 5.6).

Risk association of LUADT cancer was 53% greater for those who drank more than eight or six units of alcohol a day (males or females respectively) compared to those who drank fewer than or equal to the daily limit (RR 1.53 95% CI 1.10, 2.12). Likewise, LUADT cancer risk association was nearly double for those who ate only one portion of fruit or vegetables daily (RR 1.90 95% CI 1.27, 2.84) and more than two point five-times for those who ate even less portions of fruits and vegetables per day (RR 2.66 95% CI 1.80, 3.93) compared to those who ate four or more portions. Those who exercised below the five-times per week threshold relative to those who exercised five or more times per week, had a two-fold higher risk association with LUADT cancer (RR 2.27 95% CI 1.80, 3.93) (Table 5.6).

Compared to those with no high risk behaviours, any high risk behaviour

increased risk by at least four-fold (RR range 4.06 to 5.24). The RRs for those who both smoked and consumed alcohol in the highest risk categories had the highest risk of LUADT cancer over five-times that for those who had no high risk behaviours (RR 5.24 95% CI 3.56, 7.73). At nearly five-fold greater LUADT cancer risk, those individuals who were in the highest risk category for at least three behaviours had almost as high a risk of LUADT cancer as those who both smoked and consumed alcohol excessively (RR 4.87 95% CI 2.98, 7.95). Notably, for these individuals, the estimated risk range included the risk range for those who smoked and consumed alcohol excessively (Table 5.6).

#### 5.3.3.2 Socioeconomic risks (minimally adjusted models)

Compared to those in work, those unemployed and those retired had almost a two-fold elevated risk association with LUADT cancer (RR 1.97 95% CI 1.25, 3.10 and RR 1.96 95% CI 1.22, 2.20 respectively). LUADT cancer risk association was also elevated more than two-fold for those who had no education compared to those with a first degree or higher (RR 2.33 95% CI 1.46, 3.74). Similar elevated LUADT cancer risk (nearly two-fold) was also present for those who were educated but not at degree level (RR 1.78 95% CI 1.13, 2.80) when compared to those with a degree. Compared to the areas with more affluent socioeconomic circumstances as measured by SIMD, only those living in low socioeconomic areas had a greater risk association of LUADT cancer, but this risk was elevated 42% (RR 1.42 95% CI 1.05, 1.93). The risk association with LUADT cancer was greater for each subsequent lower occupational social class. Compared to professionals, managers and technicians, those in all other occupations were at a 31% to 74% greater risk of LUADT cancer with those who were employed in unskilled jobs at the greatest elevated risk (RR 1.74 95% CI 1.24, 2.43). Those who had no car were at a two-fold increase risk in LUADT cancer compared to those with at least one car (RR 2.05 95% CI 1.69, 2.50). Likewise, those who rented their home from a local authority experienced more than two-fold greater risk of LUADT cancer compared to those who owned their own home (RR 2.37 95% CI 1.93, 2.90). Finally and compared to those with no multiple SES highest risk factors, those with two or more SES highest risk factors were at least twice as likely to

be diagnosed with LUADT; risk of LUADT cancer was over three-times greater for those with four to six highest SES factors (RR 3.35 95% CI 2.26, 4.97) (Table 5.7).

For all risks, both behavioural and socioeconomic, the highest LUADT cancer relative risk was for smoking more than 50 years with a very elevated RR of 26.80 (95% CI 10.06, 71.42).

#### 5.3.3.3 Relative risks associated with behaviours and SES (minimally adjusted model) ranking

The 15 highest risk categories which were also significant with the confidence interval excluding 1.0 were ranked in order of decreasing size of relative risk association of LUADT cancer. Compared to the relevant lowest risk population, the high risk population had elevated risk associated with LUADT cancer that varied widely from 42% to 2,680%. Compared to those who smoked less than 20 years, those who smoked the longest had the highest associated risk of LUADT cancer which dominated all other relative risks. Those who were both current smokers and consumed more than the threshold number of alcohol units weekly followed with a five-fold increase in risk. Of all the socioeconomic factors featured, those with multiple low socioeconomic circumstances had more than three-times the associated risk of LUADT cancer and ranked 6$^{th}$ of the 15 risk categories; experiencing multiple low socioeconomic circumstances was the first and highest risk socioeconomic factor. Generally, those who exhibited high risk behaviours ranked higher than those with low socioeconomic circumstances (Table 5.15).

**Table 5.15 Ranking in decreasing size of relative risks (RR) for LUADT cancer group**

| | Characteristic | Category | RR | Focus |
|---|---|---|---|---|
| 1 | Duration smoked (years) | >50 | 26.80 | Behaviour |
| 2 | Multiple high risk behaviour | 2 Smoking Alcohol | 5.24 | Behaviour |
| 3 | Age started smoking | <20 | 4.26 | Behaviour |
| 4 | Smoking status | Current | 3.81 | Behaviour |
| 5 | Cigarettes smoked/ day | >=20 | 3.37 | Behaviour |
| 6 | Multiple low socioeconomic circumstances | 4-6 Low socioeconomic circumstances | 3.35 | SE circumstances |
| 7 | Fruit and vegetable consumption/ day | <1 or never | 2.66 | Behaviour |
| 8 | Housing tenure | Rent LA | 2.37 | SE circumstances |
| 9 | Highest qualification | None | 2.33 | SE circumstances |
| 10 | Exercise sessions/ week | <5 episodes/ week | 2.26 | Behaviour |
| 11 | Car ownership | No car | 2.05 | SE circumstances |
| 12 | Economic activity | Unemployed | 1.97 | SE circumstances |
| 13 | Occupational social class | V! Unskilled | 1.74 | SE circumstances |
| 14 | Binge units/ day | >binge rate/ day M:8 F:6 | 1.53 | Behaviour |
| 15 | SIMD | 5 Most deprived | 1.42 | SE circumstances |

[1]Minimally adjusted by age and sex

### 5.3.3.4 Identification of variables with greatest predictive value

Of the four smoking variables, the forward stepwise logistic regression model for LUADT cancer risk included only cigarettes smoked/ day and smoking duration (years) while both alcohol variables were selected via the procedure based on the likelihood ratio test where $P$<0.20 (Table 5.16).

**Table 5.16 Likelihood ratio test statistics[1] for smoking and alcohol variables for LUADT cancer risk**

| Focus | Step | Variable | DF | Chi-square | Pr >Chi Sq |
|---|---|---|---|---|---|
| Alcohol variables | 1 | Binge units/day | 1 | 28.4175 | <0.001 |
| | 2 | Units alcohol /week | 6 | 19.5204 | 0.0034 |
| Smoking variables | 1 | Cigarettes smoked/ day | 2 | 108.6558 | <0.001 |
| | 2 | Smoking duration | 4 | 31.5495 | <0.001 |

[1]Logistic regression forward selection at 0. 20 selection level

### 5.3.3.5 LUADT cancer risk and smoking/ alcohol and socioeconomic/ sex interactions

The interactions between the two smoking variables (cigarettes smoked per day and smoking duration) that were identified as providing the greatest predictive value for LUADT cancer risk and both alcohol variables (alcohol units per week and binge units per day) were tested to establish if the effect of either variable on LUADT cancer risk varied according to the level of the other variable; that is, the smoking and alcohol explanatory variables did not act independently on

LUADT cancer risk, the dependent variable. Sex and socioeconomic interactions for each of the seven socioeconomic variables were also tested using Poisson regression models.

Using Chi-square probability < 0.05, results indicated that only the interactions between binge drinking and both smoking variables (cigarettes smoked per day , smoking duration) were significant (*P*<0.0001) while none of the interactions tested between sex and the seven socioeconomic variables proved significant (*P*-values ranging from 0.115 to 0.737). This suggested that the combination of cigarettes smoked per day and binge drinking as well as the combination of years smoked and binge drinking synergistically increased the risk of LUADT cancer (Table 5.17).

**Table 5.17 Likelihood ratio test for smoking and alcohol and sex and socioeconomic variable interactions for LUADT cancer risk**

| Interaction | Chi-Sq | *P*-Value[1] |
|---|---|---|
| Cigarettes smoked per day * binge drinking | 40.23 | **<0.001** |
| Smoking duration * binge drinking | 76.42 | **<0.001** |
| Economic activity * sex | 1.90 | 0.593 |
| SIMD * sex | 7.44 | 0.115 |
| Highest qualification * sex | 2.51 | 0.285 |
| Occupational social class * sex | 2.92 | 0.572 |
| Car ownership * sex | 1.10 | 0.294 |
| Housing tenure * sex | 0.61 | 0.737 |
| Multiple high deprivation * sex | 3.92 | 0.417 |

[1]**Bold** indicates significance at Chi-square probability <0.05

### 5.3.3.6 Socioeconomic risk attenuation for LUADT cancer risk successively adjusted for behaviour

Of the seven socioeconomic indicators used, elevated LUADT risks were fully attenuated after adjustment for smoking behaviour in three of the indicators: low area socioeconomic circumstances as measured by SIMD, highest qualifications and occupational social class. In contrast, elevated LUADT risks for those renting a local authority home and participants in the highest category of multiple low socioeconomic circumstances remained even after fully adjusting for all the risk behaviours (smoking, alcohol, diet and exercise) (Tables 5.18, 5.19, 5.20).

For those who rented a local authority home, compared to individuals who owned their own home, the more than two-fold elevated risk of LUADT cancer (RR 2.37 95% CI 1.93 2.90) remained even after adjustment for their smoking, alcohol, diet and exercise behaviours such that their risk remained 54% greater than that of home owners (RR 1.54 95% CI 1.07, 2.21) (Table 5.19).

Likewise, the elevated risk of LUADT cancer of more than three-times (RR 3.35 95% CI 2.26, 4.97) compared to the risk of those with no low socioeconomic factors remained elevated with 86% greater risk for those in the highest category of multiple low socioeconomic circumstances even after full adjustment for all behaviours (RR 1.86 95% CI 1.04, 3.31) (Table 5.20).

Both of the two remaining socioeconomic factors (economic activity and car ownership) were attenuated at intermediate points of adjustment for risk behaviours. For those who were retired who had close to twice the risk of LUADT cancer (RR 1.96 95% CI 1.42, 2.68) compared to participants in work, their risk reduced with each successive additional adjustment (excluding diet) and was only fully attenuated after adjustment for all the risk behaviours (RR 0.97 95% CI 0.55, 1.71) (Table 5.18).

The LUADT cancer risks for those participants who had no access to a car were initially two-times the risk of those with a car in the minimally adjusted model (RR 2.05 95% CI 1.69, 2.50). Although still elevated, their risks reduced to 36% greater than those with a car after adjustment for smoking behaviour (RR 1.36 95% CI 1.07, 1.73) but did not fall further after adjustment for alcohol behaviours (RR 1.36 95% CI 1.01, 1.83). Compared to car owners, the risk of LUADT cancer for those with no car only became attenuated after adjusting for their smoking, alcohol and diet behaviours (RR 1.25 95% CI 0.90, 1.73) (Table 5.19).

**Table 5.18 Socioeconomic factors relative risks (RR) and 95% confidence intervals (CI) for LUADT[1] successively adjusted for behaviour, Scotland 1995-2011**

| | | Age, Sex | | | + Smoking[2] | | | + Smoking[2], alcohol[3] | | | + Smoking[2], alcohol[3], diet | | | + Smoking[2], alcohol[3], diet, exercise | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Characteristic | Category | *P* | RR | CI | *P* | RR | CI | *P* | RR | CI | *P* | RR | CI | *P* | RR | CI |
| Economic | Employed | | *Reference* | | | *Reference* | | | *Reference* | | | *Reference* | | | *Reference* | |
| Activity | Unemployed | **0.003** | 1.97 | 1.25 3.10 | 0.226 | 1.45 | 0.79 2.66 | 0.648 | 1.18 | 0.57 2.44 | 0.347 | 1.43 | 0.68 3.00 | 0.480 | 1.31 | 0.62 2.76 |
| | Retired | **<0.001** | 1.96 | 1.42 2.68 | **0.011** | 1.70 | 1.13 2.55 | **0.029** | 1.74 | 1.06 2.84 | **0.029** | 1.86 | 1.06 3.24 | 0.067 | 1.69 | 0.96 2.96 |
| | Other economically inactive | **0.001** | 1.64 | 1.22 2.20 | 0.224 | 1.28 | 0.86 1.90 | 0.683 | 1.12 | 0.66 1.89 | 0.753 | 1.10 | 0.62 1.93 | 0.904 | 0.97 | 0.55 1.71 |
| | Cohort | 26610 | | | 19756 | | | 16059 | | | 12936 | | | 12924 | | |
| | Missing | 0 | | | 6854 | | | 10551 | | | 13674 | | | 13636 | | |
| SIMD | 1 Least deprived | | *Reference* | | | *Reference* | | | *Reference* | | | *Reference* | | | *Reference* | |
| | 2 | 0.879 | 1.03 | 0.71 1.48 | 0.840 | 1.04 | 0.69 1.58 | 0.459 | 1.22 | 0.72 2.05 | 0.882 | 1.05 | 0.58 1.87 | 0.852 | 1.06 | 0.59 1.89 |
| | 3 | 0.906 | 1.02 | 0.74 1.41 | 0.866 | 0.97 | 0.65 1.44 | 0.469 | 1.21 | 0.73 2.00 | 0.365 | 1.28 | 0.75 2.21 | 0.334 | 1.31 | 0.76 2.24 |
| | 4 | 0.475 | 1.12 | 0.82 1.54 | 0.526 | 1.13 | 0.77 1.66 | 0.473 | 1.20 | 0.73 1.99 | 0.678 | 1.12 | 0.65 1.93 | 0.690 | 1.12 | 0.65 1.92 |
| | 5 Most deprived | **0.023** | 1.42 | 1.05 1.93 | 0.326 | 1.21 | 0.83 1.76 | 0.190 | 1.39 | 0.85 2.26 | 0.269 | 1.34 | 0.80 2.27 | 0.313 | 1.31 | 0.78 2.21 |
| | Cohort | 26610 | | | 19756 | | | 16059 | | | 12936 | | | 12924 | | |
| | Missing | 0 | | | 6854 | | | 10551 | | | 13674 | | | 13686 | | |
| Highest | First degree and higher | | *Reference* | | | *Reference* | | | *Reference* | | | *Reference* | | | *Reference* | |
| Qualification | Other non-degree | **0.013** | 1.78 | 1.13 2.80 | 0.221 | 1.40 | 0.82 2.38 | 0.679 | 1.12 | 0.65 1.95 | 0.924 | 1.03 | 0.55 1.92 | 0.924 | 1.03 | 0.55 1.92 |
| | None | **0.000** | 2.33 | 1.46 3.74 | 0.121 | 1.53 | 0.89 2.63 | 0.404 | 1.27 | 0.73 2.21 | 0.729 | 1.12 | 0.59 2.10 | 0.729 | 1.12 | 0.59 2.10 |
| | Cohort | 26581 | | | 19734 | | | 16042 | | | 12920 | | | 12909 | | |
| | Missing | 29 | | | 6876 | | | 10568 | | | 13690 | | | 13701 | | |

Relative Risk (RR), 95% CI estimated using age at diagnosis Poisson regression models offset by person-years of follow-up adjusted by sex (female as reference) and age group at start of the cohort (16-44 (reference), 45-59, 60-74, 75+ years old)
[1]LUADT includes lung and upper aero-digestive cancers
[2]Adjusted for sex, age group at start of cohort, cigarettes smoked/ per day, smoking status, smoking duration and age started smoking
[3]Adjusted for sex, age group at start of cohort, cigarettes smoked/ per day, smoking status, smoking duration, age started smoking, units of alcohol/ week and binge units/ day
**Bold** indicates CI did not include 1.0

**Table 5.19 Socioeconomic factors relative risks (RR) and 95% confidence intervals for LUADT[1] successively adjusted for behaviour, Scotland 1995-2011 continued**

| | | Age, Sex | | | | + Smoking[2] | | | | + Smoking[2], alcohol[3] | | | | + Smoking[2], alcohol[3], diet | | | | + Smoking[2], alcohol[3], diet, exercise | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Characteristic | Category | *P* | RR | CI | | *P* | RR | CI | | *P* | RR | CI | | *P* | RR | CI | | *P* | RR | CI | |
| Occupational Social Class | I,II Professional, managerial, technical | | *Reference* | | | | *Reference* | | | | *Reference* | | | | *Reference* | | | | *Reference* | | |
| | III N Skilled non manual | 0.925 | 0.98 | 0.70 | 1.39 | 0.506 | 0.87 | 0.59 | 1.30 | 0.381 | 0.81 | 0.50 | 1.30 | 0.645 | 0.88 | 0.52 | 1.50 | 0.638 | 0.88 | 0.52 | 1.49 |
| | III M Skilled manual | 0.051 | 1.31 | 1.00 | 1.71 | 0.914 | 1.02 | 0.75 | 1.38 | 0.483 | 0.87 | 0.60 | 1.27 | 0.556 | 0.88 | 0.57 | 1.35 | 0.555 | 0.88 | 0.57 | 1.35 |
| | IV Partly skilled | **0.028** | 1.39 | 1.04 | 1.87 | 0.697 | 0.93 | 0.66 | 1.33 | 0.846 | 0.96 | 0.63 | 1.46 | 0.756 | 1.08 | 0.68 | 1.71 | 0.824 | 1.05 | 0.66 | 1.68 |
| | VI Unskilled | **0.001** | 1.74 | 1.24 | 2.43 | 0.355 | 1.21 | 0.81 | 1.82 | 0.232 | 1.35 | 0.83 | 2.22 | 0.098 | 1.55 | 0.92 | 2.62 | 0.118 | 1.52 | 0.90 | 2.56 |
| | Cohort | 25611 | | | | 19205 | | | | 15634 | | | | 12582 | | | | 12570 | | | |
| | Missing | 999 | | | | 7405 | | | | 10976 | | | | 14028 | | | | 14040 | | | |
| Car Ownership | 1 or more car(s) | | *Reference* | | | | *Reference* | | | | *Reference* | | | | *Reference* | | | | *Reference* | | |
| | No car | **<0.001** | 2.05 | 1.69 | 2.50 | **0.011** | 1.36 | 1.07 | 1.73 | **0.041** | 1.36 | 1.01 | 1.83 | 0.182 | 1.25 | 0.90 | 1.73 | 0.213 | 1.23 | 0.89 | 1.70 |
| | Cohort | 26610 | | | | 19756 | | | | 16059 | | | | 12936 | | | | 12924 | | | |
| | Missing | 0 | | | | 6854 | | | | 10551 | | | | 13674 | | | | 13686 | | | |
| Housing | Owner occupier | | *Reference* | | | | *Reference* | | | | *Reference* | | | | *Reference* | | | | *Reference* | | |
| Tenure | Rent LA | **<0.001** | 2.37 | 1.93 | 2.90 | **<.0001** | 1.76 | 1.37 | 2.26 | **0.001** | 1.75 | 1.26 | 2.43 | **0.015** | 1.57 | 1.09 | 2.25 | **0.028** | 1.50 | 1.05 | 2.16 |
| | Rent privately | 0.056 | 1.38 | 0.99 | 1.93 | 0.130 | 1.33 | 0.92 | 1.91 | 0.225 | 1.30 | 0.85 | 1.98 | 0.483 | 1.19 | 0.73 | 1.93 | 0.539 | 1.16 | 0.72 | 1.89 |
| | Cohort | 26610 | | | | 19756 | | | | 16059 | | | | 12936 | | | | 12924 | | | |
| | Missing | 0 | | | | 6854 | | | | 10551 | | | | 13674 | | | | 13686 | | | |

Relative Risk (RR), 95% CI estimated using age at diagnosis Poisson regression models offset by person-years of follow-up adjusted by sex (female as reference) and age group at start of the cohort (16-44 (reference), 45-59, 60-74, 75+ years old)
[1]LUADT includes lung and upper aero-digestive cancers
[2]Adjusted for sex, age group at start of cohort, cigarettes smoked/ per day, smoking status, smoking duration and age started smoking
[3]Adjusted for sex, age group at start of cohort, cigarettes smoked/ per day, smoking status, smoking duration, age started smoking, units of alcohol/ week and binge units/ day
**Bold** indicates CI did not include 1.0

**Table 5.20 Socioeconomic factors relative risks (RR) and 95% confidence intervals (CI) for LUADT[1] successively adjusted for behaviour, Scotland 1995-2011 continued**

| | | Age, Sex | | | | + Smoking[2] | | | | + Smoking[2], alcohol[3] | | | | + Smoking[2], alcohol[3], diet | | | | + Smoking[2], alcohol[3], diet, exercise | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Characteristic | Category | *P* | RR | CI | | *P* | RR | CI | | *P* | RR | CI | | *P* | RR | CI | | *P* | RR | CI | |
| Multiple High Deprivation | 0 No deprivation | *Reference* | | | | *Reference* | | | | *Reference* | | | | *Reference* | | | | *Reference* | | | |
| | 1 | 0.104 | 1.27 | 0.95 | 1.70 | 0.897 | 1.02 | 0.72 | 1.45 | 0.828 | 1.05 | 0.70 | 1.57 | 0.823 | 0.95 | 0.60 | 1.51 | 0.810 | 0.94 | 0.59 | 1.50 |
| | 2 | **<0.001** | 2.01 | 1.52 | 2.66 | 0.267 | 1.22 | 0.86 | 1.74 | 0.286 | 1.26 | 0.82 | 1.93 | 0.629 | 1.13 | 0.69 | 1.83 | 0.705 | 1.10 | 0.68 | 1.78 |
| | 3 | **<0.001** | 2.81 | 2.09 | 3.76 | **0.007** | 1.63 | 1.15 | 2.33 | 0.087 | 1.46 | 0.95 | 2.26 | 0.282 | 1.31 | 0.80 | 2.15 | 0.352 | 1.27 | 0.77 | 2.08 |
| | 4-6 Low socioeconomic circumstances | **<0.001** | 3.35 | 2.26 | 4.97 | **0.005** | 1.91 | 1.22 | 2.98 | **0.007** | 2.05 | 1.22 | 3.46 | **0.024** | 1.95 | 1.09 | 3.47 | **0.036** | 1.86 | 1.04 | 3.31 |
| | Cohort | 26610 | | | | 19756 | | | | 16059 | | | | 12936 | | | | 12924 | | | |
| | Missing | 0 | | | | 6854 | | | | 10551 | | | | 13674 | | | | 13686 | | | |

Relative Risk (RR), 95% CI estimated using age at diagnosis Poisson regression models offset by person-years of follow-up adjusted by sex (female as reference) and age group at start of the cohort (16-44 (reference), 45-59, 60-74, 75+ years old)
[1]LUADT includes lung and upper aero-digestive cancers
[2]Adjusted for sex, age group at start of cohort, cigarettes smoked/ per day, smoking status, smoking duration and age started smoking
[3]Adjusted for sex, age group at start of cohort, cigarettes smoked/ per day, smoking status, smoking duration, age started smoking, units of alcohol/ week and binge units/ day
**Bold** indicates CI did not include 1.0

## 5.4 Discussion

### 5.4.1 Principle findings

Those with multiple low socioeconomic circumstances had very strong association for increased risk of both all cancer and LUADT cancer. For all cancer risk, the elevated risk associated with all categories of multiple low socioeconomic circumstances was nearly fully attenuated when adjusted for smoking. For LUADT cancer, the risk associated with multiple low social circumstances increased in a dose-response manner such that each additional socioeconomic disadvantage level resulted in increased risk association. Those in the highest category of multiple low socioeconomic circumstances (with four to six SES indicators in the lowest category of SES) were more than three-times more likely to be diagnosed with LUADT cancer compared to their affluent counterparts. Furthermore, for those facing the greatest disadvantage, this elevated risk association remained at 86% greater than those with no socioeconomic disadvantage, after full adjustment for smoking, alcohol, diet and exercise behaviours.

When looking at single SES indicators, only two situations of low socioeconomic circumstances resulted in elevated risk associations of all cancer: being retired (43%) and renting a local authority home (18%). In both cases, these modest elevated risk associations were ultimately fully attenuated with the successive addition of smoking and alcohol behaviours (renting a local authority home) or smoking, alcohol and diet behaviours (retired). For LUADT cancer, the unadjusted risk associated with each of the six SES indicators for those in the lowest socioeconomic circumstances was elevated by a minimum of 42% (SIMD-most deprived) to a maximum of 237% (housing tenure – renting from a local authority). However, the elevated risks associated with the lowest SES categories were fully attenuated after the addition of smoking behaviour in the model; only those who rented accommodation from a local authority remained with a 50% elevated risk of LUADT cancer even after adjustment for all the behaviours. The risk of those who were retired was fully attenuated after adjustment for smoking, alcohol and diet behaviours. What was not captured

through this analysis was why an individual had retired. If retirement was "involuntary" due to ill health reasons or redundancy, it is possible that their health deteriorated due to stress and anxiety associated with lack of control, financial concerns and health limitations. These may ultimately initiate or accelerate the higher risk association with LUADT cancer identified. Alternatively, if retirement was "voluntary" no or even improved mental and physical health may result (Van de Heide et al 2013). A greater understanding of the circumstances surrounding retirement is required.

For the all cancer group, this study found that SIMD was not statistically significant for any SIMD category even before adjustment for risk behaviours.

### 5.4.2 What is already known on this topic

Consistent with this study's finding that smoking was a fundamental risk factor dominating the risks of a cancer diagnosis among the socioeconomically disadvantaged, Hiscock et al (2012) found that over time (2001-08), smoking prevalence has been concentrated among those facing multiple facets of socioeconomic disadvantage in England. In their study, while the affluent group experienced increased prevalence of never smokers, those from very disadvantaged backgrounds had not experienced a decline in smoking prevalence and had not experienced improved socioeconomic circumstances.

The lack of association of SES and cancer risk after adjustment for behaviours, is likely to reflect the mix of cancer sites included in the all cancer group some of which were more likely to be diagnosed among more affluent groups (e.g. breast cancer, melanoma and prostate cancer) while other cancers (e.g. lung cancer) were more likely to be diagnosed among the more socioeconomically disadvantaged and have greater behaviour association as well. Given the mix of cancer sites in the all cancer group, it is not surprising that SIMD was not statistically significant for any SIMD category. Lack of a relationship with individual measures of SES was also observed. This was consistent with Leuven's natural experiment study in Norway focusing on the association of education reform in Norway, who noted little evidence that education was associated with

all cancer risk (Leuven et al 2016) (Section 2.2).

This study also established for both cancer incidence end points, compared to being employed, being retired was associated with a nearly two-fold increased risk association for LUADT cancer and a 43% increased risk for all cancer. For both cancer groups, these risk associations were only fully attenuated with the addition of all the risk behaviours. This most likely reflects that cancer in general and LUADT cancer in particular are diseases of the older population with a long lag-time between initiation and diagnosis. Although confidence intervals included 1.0, in both cases, compared to being employed, the elevated risk that was associated with being retired remained at 69% and 31% for LUADT and all cancer, respectively. This may also reflect differences by birth cohort in uptake and cessation of smoking which vary by socioeconomic status such that those who are older and more affluent are more likely to have started smoking and to have ceased smoking, while those who are older but more disadvantaged are less likely to have stopped smoking (Sidorchuk et al 2009; Hiscock et al 2012).

Previous studies have reported results that are consistent with this study's results for occupational social class where an elevated risk of LUADT cancer for partly skilled and unskilled employees was attenuated fully after addition of smoking behaviours. Melchior et al (2005) used a cohort to study occupational social class, cancer incidence and behaviours including smoking, alcohol consumption, diet and other behaviours (excluding exercise). They found that an observed occupational gradient in smoking and alcohol related cancers was greatest for manual workers, but was fully attenuated after adjustment for behaviours. Occupational social class captures the prestige associated with a particular job and the consequential benefits of greater social standing as well as material reward and resources (Galobardes et al 2006a). Lower occupational social classes may experience less control, greater stress and greater exposure to high risk behaviours which may be pursued or retained as a means of coping with life's circumstances (Marmot et al 1991).

This study also established that renting a local authority home was associated

with an increased risk of all cancer and LUADT cancer. In the case of all cancer, the risk was fully attenuated after inclusion of smoking, alcohol and diet behaviours; while for LUADT cancer, the risk remained elevated even after full behaviour adjustment. Although no data on risk behaviours were available, in the study by Spadea et al (2010), they reported similar findings. They too evaluated multiple individual measures of SES (education, occupational social class and housing characteristics) and one area measure. They found that housing characteristics had the strongest association with increased risk for men and was weakly associated for women (Spadea et al 2010). Housing tenure and housing characteristics are indicators of material wealth which also capture potential exposure to stressful or unhealthy environments. As a material wealth measure (Galobardes et al 2006), it captures the major financial outlay for both owners and renters starting from young adulthood; and for home owner-occupiers, it can indicate significant accumulation of wealth in later adult years. Housing tenure also captured a longer time horizon which mirrored the life-course attribute of SES as well as the longer gestation period from cancer initiation to diagnosis. This attribute may also make housing tenure or housing characteristics a particularly sensitive measure of SES when evaluating cancer risk.

This study's results of fully attenuated LUADT cancer risk association with low educational attainment and low occupational social class after adjustment of smoking were not consistent with Sidorchuk et al (2009) who identified that, even after adjustment of smoking behaviour, greater lung cancer risk was associated with lower educational attainment and lower occupational social class. In this study, the LUADT cancer group included head and neck and squamous cell oesophageal carcinoma as well as lung cancer. As these cancers are also strongly associated with smoking behaviour, it may be expected that similar results would be identified. However, the SHeS has been shown not to be fully representative of the Scottish population with certain groups such as men (who are more at risk of LUADT cancer than women) and the more socioeconomic disadvantaged less likely to participate and/or consent to linkage of their data (Gray et al 2012; Gray et al 2013). Furthermore, with the

behaviours being self-reported, the responses are subject to bias. Additionally, these findings were not consistent with the INHANCE study which identified a third unexplained associated risk for head and neck cancer associated with low educational attainment although the INHANCE data only adjusted for smoking and alcohol and was based on case-control data so could be subject to residual confounding (Conway et al 2015).

### 5.4.3 What this study adds

This study identified that those who faced multiple low socioeconomic circumstances had more than three-times greater risk association for LUADT cancer compared to those with no socioeconomic disadvantage. Furthermore, for those facing the greatest disadvantage, this elevated risk association remained at 86% compared to their affluent counterparts, even after full adjustment for smoking, alcohol, diet and exercise behaviours. Multiple indicators of socioeconomic circumstances and their risk association with cancer have been evaluated separately along with behaviours or they have been evaluated together, but without adjustment for relevant behaviours. At the date of this thesis and to the best of current knowledge, this study was the first to explore the association with cancer incidence of multiple measures of disadvantage with a compound variable at individual level alongside individual behaviour factors. Capturing multiple exposures of low socioeconomic circumstances, as in this study, is more likely to reflect the multi-dimensional nature of socioeconomic status and its compounding effects as well as potentially the effect of socioeconomic status at various points over the life-course. It was, however, possible that further confounding by behaviours which were not measured or not fully measured could have been responsible for the remaining SES effect. Moreover, several variables were used to reflect as completely as possible the most important known risk behaviours. For example, this study employed several smoking variables: smoking status, smoking duration, age started smoking and cigarettes smoked per day, to consider as completely as possible smoking behaviour association with risk of all cancer and LUADT cancer. This approach supported the focus on duration and Peto's

observation that the variable pack-years was less important than number of cigarettes smoked daily and duration of smoking (Peto 2012).

### 5.4.4 Strengths of this study

Due to high response rates, the high level of consent to linkage of 84% (Morris 2017; Table 5.1) across all available surveys, an overall 99.4% successful linkage rate of respondents agreeing to linkage to the CHI database (Morris 2017; Table 5.1) and through the combination of all available SHeS surveys, this study was based on a prospective cohort of 46,368 persons available for follow-up for up to 16 years with a total number of person-years of follow-up available of 3,737,854 years. As a result, this enabled measurement of SES before diagnosis, reflecting the extended temporal relationship between SES exposure and cancer initiation. The larger cohort provided by using all available SHeS surveys also facilitated making comparisons with increased power to detect differences where they exist. Furthermore, there have been a limited number of cancer incidence/ SHeS linkage studies undertaken to date.

This study utilised individual level socioeconomic circumstances described by multiple and diverse detailed variables, albeit self-reported. At the same time, it used an area-based measure of socioeconomic circumstances, as well as several individual major risk behaviours known to be associated with the incidence of the cancers in question. These data enabled analysis of the extent these behaviours explained the association between low socioeconomic circumstances and cancer incidence.

Reflecting the finding that no single SES indicator captured all aspects of SES presented in Chapter 3 and because of the breadth of individual socioeconomic measures also held by the SHeS, it was possible to create an *individual* indicator of multiple low socioeconomic circumstances; recognising that low socioeconomic circumstances frequently "cluster" (Watt 2002) thereby compounding the vulnerability to cancer risk (Watt 2002; Galobardes et al 2004).

A further strength of this study was the potential to begin to infer a life-course

interpretation of SES on health. Housing tenure and its association with accumulated wealth attained during active professional life and retirement, the later periods of the life-course, coupled with highest qualification which may reflect early life and parental SES (Galobardes et al 2006b; Dalton et al 2008a) combined, supported this longer life-course view. Together, these variables covered an extended period of time commensurate with the long gestation period of cancer initiation to diagnosis and reflected facets of SES over the full life-course.

Housing tenure, one of the indicators which were used to measure material aspects of SES, was found to have a 50% greater LUADT cancer risk association for those renting a local authority or housing association accommodation relative to home ownership. This finding was interesting in itself, although there was no data on housing condition, (e.g. overcrowding, dampness and building materials). Housing condition (compared to housing amenities) was the SES attribute considered most relevant to advanced industrialised societies such as Scotland (Galobardes et al 2006b). Finally, markers for unfavourable social circumstances, poor housing conditions, early life socioeconomic conditions and material lifetime well-being along with exposure to environmental tobacco smoking and occupational hazards, have all been previously considered to contribute to susceptibility to lung cancer (Sidorchuk et al 2009).

A major strength of this study is the prospective cohort study design which enabled the measurement of socioeconomic circumstances before diagnosis (at least at one point in time) thereby assessing the temporal relationship and minimising the possibility of reverse causation (where a cancer diagnosis may, lead to lower socioeconomic circumstances). Finally, the cohort design provided a longer period of follow-up allowing the consideration of the potentially long lead-time between cancer gestation and diagnosis.

The minimum dataset used (Section 5.2.4) removed records where no information was available for the socioeconomic indicators thereby ensuring that full information for assigning SES was available for the exposure variables. This

also enabled one to observe and compare patterns of attenuation across the SES indicators and behaviour adjustment models. Others adopted a different approach to missing values, e.g. allocating records with missing values for SES variables to the not disadvantaged category (Hiscock et al 2012). However, to address potential misclassification, SES measures were averaged over two successive three year periods and, as a result, variation between years was smoothed. This option was not available in this study; nevertheless, the approach adopted was a thorough, conservative method of data management and analysis.

## 5.4.5 Limitations of this study

### 5.4.5.1 Scottish Health Survey population

The Scottish Health Surveys are representative of individuals living in private households and thus exclude those living in communal establishments, such as residential care and prisons or those in the armed forces. There are potential sources of bias as well that may arise from the agreement to participate in the original interview and agreement to linkage of records (Gray et al 2013). These potential biases are likely to be important when exploring socioeconomic factors as in this study. The bias of the SHeS sample towards a more affluent population which was not representative of all of Scotland is potentially a significant limitation for this study, particularly in capturing the socioeconomic and behaviour circumstances of the less advantaged population, the main focus of this study. Nevertheless, the results found a clear 86% increased risk association with multiple low socioeconomic circumstances and LUADT cancer risk even after adjustment for all behaviours. These RRs (and others reported) may in fact be under-estimated as a result of the recognised SHeS sample bias. In addition, due to the bias associated with self-reported behaviour, residual confounding by smoking and the other behaviours cannot be ruled out. There was a substantial number of missing behavioural variables. However, to maximise the dataset, it was not possible to create a minimum analysis dataset that excluded all missing variables. This would have resulted a significant reduction in the number of cases (which given n=453) would have reduced the statistical power of the

modelling.

The Scottish population is relatively stable with low immigration supporting follow-up for the majority of consenting SHeS participants. Previous studies identified a small number (four percent to seven percent) of total 1995 and 1998 participants who were immigrants for whom follow-up morbidity records may be incomplete in the linked dataset (Gray et al 2010). This may have resulted in incomplete information on pre-existing cancers identified through SCR linkage diagnosed prior to survey for those immigrating to Scotland. Given the trend over the last decade of greater migration to UK in general and Scotland in particular, it is possible that the omission of pre-existing cancers has increased leading to an over estimate of the relative risks. Between the two census years 2001 and 2011, the population in Scotland from non UK countries or the Republic of Ireland increased by around 315,000, approximately 5.9%, a relatively small proportion of the total 2011 Scottish population and therefore not likely to be a significant factor for this study (National Records Scotland 2017). Nevertheless, it was possible that the number of cancer diagnoses identified was under-estimated as cohort participants may die before cancer diagnosis of other competing morbidities (e.g. stroke, heart attack etc.) and this was more likely to occur in the more disadvantaged populations. As a result, the relative risks presented may be under-estimated. As discussed in the longitudinal survey study (Chapter 3), alternative study designs could have been considered, however, the focus of this study (as in the longitudinal study) was to consider SES as the exposure, and in this context, alternative analytical methods such as the cumulative incidence function would have been less appropriate in this case (Dignam et al 2012).

The younger age restriction applied in the 1995 and 1998 surveys of those under 65 years old and 75 years old respectively limited the population most likely to be diagnosed with cancer from these surveys. And the inclusion all adults 16 years old and older for the surveys from 2003 onwards, because of the closer proximity to the 2011 study end date, limited the follow-up period for cancer diagnosis for these individuals, a disease with a long gestation period.

Nevertheless, total number of person-years of follow-up available was 3,737,854 years.

#### 5.4.5.2 Behaviour association

***Self-completed behaviour data and question design***. The SHeS behaviour data were self-completed without objective or external validation. The reliability of some measures, such as diet, was particularly questionable as respondents have been known to provide answers that convey more favourable nutritional profiles than objective data suggest (Gray et al 2009a). Although this pattern is likely to be no different across both groups (i.e. those with and without cancer). Furthermore, and as described previously (Sections 5.2.2, 5.2.3), the earlier survey years (1995, 1998 and in some cases 2003) in particular did not use the same questions to capture behaviour and socioeconomic circumstances as the later surveys. As a result, some important refinements available for the later surveys were missing altogether (e.g. binge drinking in 1995) or were not captured in the same level of detail (e.g. physical exercise). In addition, the socioeconomic circumstances and behaviours of cohort participants were only included at the start of the cohort. Given the SHeS cross-sectional design, it was not possible to obtain the same individual's measures of SES and behaviours throughout the cohort follow-up period so causal attributions cannot be fully established (Hiscock et al 2012). Previous research focusing on mortality established that multiple measures of behaviours more completely measured the explanatory effect of SES association on health outcome (Stringhini et al 2010). Finally, information on human papillomavirus infection was not available for inclusion; however it is a recognised risk for oropharyngeal cancer incidence (Conway et al 2016). Nevertheless, the survey data remain useful for carrying out comparisons across population groups within similar periods of time as carried out in the present study (dos Santos Silva 1999) and as previously stated, provide information on the temporal relationship between SES exposure and cancer diagnosis.

***Alcohol consumption***. With respect to self-reported alcohol consumption it is known that the SHeS sample was not representative of the full Scottish

population and existing demographic and socioeconomic adjusted weighting do not fully compensate (Gray et al 2013). Other researchers identified this issue (Gray et al 2013) and have explored ways to address this limitation. At the time of conducting this study, their work had only just begun. As a result, this study used two sex specific complementary variables (number of units consumed per week and binge drinking behaviour) to address this point in the absence of alternatives.

Individuals in the lower socioeconomic group were most likely to report either "do not drink alcohol at all" or "drink very little alcohol"; however, they were also the group that were most likely to drink at harmful levels (Gillan 2010; Brown et al 2016). Compared to the most affluent population, 7% of whom reported consuming no alcohol, 23% of those in the lowest income group reported that they abstain. This may reflect stretched financial resources and priority of alcohol purchase for those living in low socioeconomic circumstances. Those with the highest incomes were more likely to drink hazardously, but harmful drinkers in the lowest income group drank significantly more than harmful drinkers in the highest income group (Gillan 2010; Brown et al 2016). To reflect these behavioural and socioeconomic differences a "binge drinking" variable based on Scottish Government guidance and specific to males and females separately was created; however, and as already identified (Section 5.2.3), the 1995 survey design did not include this information, although it was possible in all subsequent surveys.

Adoption of the four smoking status variables and evaluating them separately rather than deriving a single measure of lifetime cumulative dose recognised that cancer risk at a given cumulative dose sometimes varied substantially with the duration of exposure (Peto 2012).

#### 5.4.5.3 Socioeconomic status association

***Area-based SIMD***. The area-based SIMD measure has the disadvantage (Chapter 1) of considering everyone in the area as having the same SES. This is described as the "ecological fallacy". Lower socioeconomic area or place effects are also

considered important and can include elements of the socio-physical environment that can play a role in health and disease (Ellaway et al 2012). In addition, the change in convention of SIMD during the period of this study where the definition of the least and most deprived categories was reversed was likely to create misclassification of deprivation, unless researchers were conscious of the change and the need to correct for the convention reversal (Section 5.2.2.1). This may be less of an issue for those focused on socioeconomic circumstances as the exposure, but may be overlooked by those viewing SES as a confounder and was not an issue in this analysis.

***Highest educational qualification***. There remain some limitations to the measures of SES used. The individual SES measures each have their own limitations (and strengths). Highest qualification (as opposed to number of years of formal education) was used to reflect achievement implying that length of time in formal education was less important than educational achievement on the assumption that achievement itself was required for enhanced job opportunities and income leading to higher SES at later stages in life. However, the meaning of education varies by birth cohort, which was not possible to consider in this study either by re-classification of education into low, medium and high, depending on participant age, or by stratifying by age. This was due to the number of SES and behaviour variables under consideration and the size of the cohort. Even if this limitation had been possible to overcome, the inability to measure the quality of the education (whether measured in years or in attainment) could not be assessed with the data available (Galobardes et al 2006b).

***Housing tenure***. Similar to education, the interpretation of housing tenure is dependent on the context at the time of collection as well as geographic implications (Galobardes et al 2006b). In addition tenure does not include important aspects related to the quality or condition of housing such as overcrowding or dampness.

***Occupational social class***. Occupational social class was also recorded and

while it importantly reflected access to social networks, work based stress, level of autonomy and control, as well as social standing and/or occupationally related carcinogenic exposures, it did not include information for those who were not currently employed either as retirees or for other reasons (e.g. students, home makers and the unemployed). It also may struggle to effectively categorise those who are self employed. Furthermore, the interpretation of occupational social class, like education and housing tenure may change over time (Galobardes et al 2006a). However, this study did not rely on occupational social class only to record SES, but used five other relevant individual and an area SES attribute which were likely to compensate for this limitation. In particular, economic activity was included and did separately classify retirees from those not currently employed.

***SES indicators not measured***. Further limitations include not directly considering other indicators of SES such as income, which may have independent effects on health inequalities. Income is a variable which many UK/European researchers (unlike US counterparts) seem reluctant to collect. However, the study did measure economic activity, occupational social class and education level, all of which indirectly reflect income. In addition, the study was not able to consider direct measurements of early life other than educational attainment or other contextual influences (such as family's socio-economic position or neighbourhood characteristics).

Like level of educational attainment, income has a dose-response relationship with health outcomes, such as cancer incidence, with increasing income providing a reduction in risk association and is described as having cumulative effect over the life-course. Of the various SES indicators, income is the one indicator that can change most dramatically over a short space of time (Galobardes et al 2006b). This last attribute could not be captured using the methods adopted in this study. In addition, income is subject to reverse causation bias as it could *reflect* health status or may *influence* health status through ability to access better quality circumstances such as neighbourhood, housing and food. More recent attempts to include income in the Scottish census

failed due to a general reluctance to supply this information (National Records Scotland 2015d), demonstrating sensitivity in Scotland to questions relating to income likely to be shared by SHeS participants. However, interest in enhancing the nature and completeness of SES information remains. In response to the Scottish Government's request for written comment on Scotland's economic data and how effective these data were for scrutiny of policy, the Scottish Public Health Observatory indicated the importance of income data as a key determinant of population health and health inequalities (Scottish Public Health Observatory 2017). Despite this limitation, adult occupation, which was measured in this study, is strongly linked to income. The association with health is likely to directly reflect material resources (both monetary and other tangible benefits) which in turn determine material living standards and ultimately health (Galobardes et al 2006b).

Another aspect of change over time not measured in this study was early life circumstances. Others (e.g. Giesinger et al 2014) have used parental SES measures such as father's occupational social class and mother's education level to reflect childhood SES and to investigate the intergenerational influence on the life-course, smoking behaviour and ultimately mortality (Giesinger et al 2014). However for this study, information on parental socioeconomic circumstances was not available but is a potential avenue for future research. Nevertheless, an aspect of childhood SES was indirectly measured via participant educational attainment as this has been described to be related to childhood /parental circumstances (Dubow et al 2009).

***Multiple low socioeconomic circumstance.***This somewhat novel development of an individual measure of multiple deprivation was a derived variable from multiple low socioeconomic circumstances. It treated each of the six individual SES variables equally with no weighting to each variable applied. The results suggest that renting a local authority or housing association accommodation was an important factor in determining cancer risk and may suggest that equal weighting was not necessarily appropriate.

### 5.4.6 Conclusion

This study has developed a novel individual measure capturing multiple exposures of low socioeconomic circumstances. However, application of weights to the different domains of socioeconomic status, which may be desirable, would require to have been hypothetical.

The study also confirmed the strong relationship between SES with behaviours, particularly smoking. Smoking has been described as a major inequality issue and a significant cancer risk. Nordahl et al (2014) defined the association of greater risk of morbidity with low socioeconomic circumstances in terms of the "differential exposure" and "differential vulnerability" to behaviour risk factors for those in lower social strata. Not only is smoking prevalence greater among those living in lower socioeconomic circumstances ("differential exposure"), but also these individuals are more vulnerable to that exposure ("differential vulnerability") (Nordahl 2014). This concept was identified in a recent Scottish alcohol morbidity and mortality analysis which found that while those from lower SES groups consumed similar levels of alcohol to the higher SES groups, the authors found and described the impact of poverty as compounding the impact of alcohol on health (Katikireddi et al 2017).

This study, through the individual SES indicator of multiple low socioeconomic circumstances, demonstrated a stronger SES risk association with cancer, particularly LUADT cancer, which itself was stronger than individual measures and not fully attenuated by behaviours. The study also confirmed the strong temporal association identified earlier.

# 6 Overall Discussion

## 6.1 Overall thesis findings

The first study of the thesis (Chapter 3) found that socioeconomic inequalities in the risk of lung and UADT cancers were among the most unequally distributed cancers in Scotland — with the steepest socioeconomic gradients and those from the poorest socioeconomic backgrounds bearing the greatest burden. These inequalities were greater for males than females. An adapted version of the complex measure of SES, the Relative Index of Inequality which reflected the full gradient of inequality was used. This quantified that across all SES groups, lung and UADT cancers, recognised as the major smoking related cancers, together contributed 91% and 81% to all cancer inequality for males and females respectively. Lung cancer was the primary contributor to all cancer inequalities (males 65%, females 68%) followed by head and neck cancer (males 19%, females 9%) and finally oesophageal cancer (males 7%, females 4%). While lung cancer dominated the contribution to all cancer inequalities for both sexes, laryngeal cancer ranked higher than lung adenocarcinoma for males; both the volume of cases and the distribution of those cases amongst the SES groups were relevant to the impact on health of socioeconomic inequalities, but given the lower case volume of laryngeal cancer, the SES gradient was relatively more important.

The second thesis study (Chapter 4) identified that to measure these inequalities, area-based measures of socioeconomic circumstances (routinely used in cancer registry analysis as in Chapter 3) alone did not reflect the full range of dimensions of socioeconomic status and the relationship of low SES on cancer risk. However, no single individual measure dominated the low SES cancer risk either. Socioeconomic circumstances risk associations seemed to be multifaceted, dynamic over time and could accumulate over the life-course. As a result, multiple measures at a point in time and over time are essential in order to reflect SES at critical stages over life's journey. Furthermore, Chapter 4 confirmed a temporal relationship between when low socioeconomic status was

experienced and the timing of cancer diagnosis and therefore SES should where possible be measured some years ahead of cancer diagnosis.

Chapter 5 found that given the mixed aetiologies and the variable associations with socioeconomic inequalities of the cancers making up the all cancer group, no association with socioeconomic inequalities was identified after adjustment for behaviours (Section 2.2, 4.4 and 5.3.2). Taking into account the significant contribution of lung and UADT cancers to all cancer socioeconomic inequalities quantified in Chapter 3, it was sensible to focus on this cancer group to understand to what extent behaviours explained these inequalities. The study linking the Scottish Cancer Registry to the Scottish Health Survey (Chapter 5) demonstrated that most of the socioeconomic variables were fully attenuated after adjustment by smoking (SIMD, education level and occupational social class) or the combination of smoking and alcohol (car ownership). Elevated risk of LUADT cancer only persisted for those renting a local authority home even after adjustment for all behaviours, while the elevated risk for those who had retired was only fully attenuated after adjustment for smoking, alcohol and diet behaviours. Smoking is a major inequality issue and a significant cancer risk.

Chapter 5, via the novel development of the derived individual SES indicator of multiple low socioeconomic circumstances, demonstrated a stronger SES risk association with cancer, particularly with LUADT cancer, which was stronger than individual measures and not fully attenuated by behaviours. It also identified the importance of reflecting the *compounded* effect of multiple socioeconomic disadvantage on health via the derived individual indicator of multiple SES exposure.

## 6.2 Comparison with previous research

The literature review (Chapter 2 Part II) had previously indicated that low socioeconomic status was associated with greater risk of lung and UADT cancer (Kogevinas et al 1997a; Conway et al 2007; Conway et al 2008; Conway et al 2010a; Conway et al 2010b; Conway et al 2015).

The findings of this thesis confirm these observations. What this thesis adds is a level of detail which was previously less described, including the quantification by histology or subsite, age and sex, of the relative contribution of these cancers to all cancer inequalities. Chapter 3 estimated that, taken together, these cancers contributed 91% and 81% of the total socioeconomic inequality gradient/burden for males and females respectively in Scotland.

Many studies agreed with the findings of this thesis that all cancer risk was marginally or not associated with low SES (Boscoe et al 2014; Leuven et al 2016). Mouw et al (2008) concurred for men, but not women where the authors identified a protective effect of lower educational attainment (Mouw et al 2008). This is likely due to the inclusion of and weight of the burden of breast cancer which is greater in more affluent women. The study performed in Chapter 3 shed some light on these findings indicating that a more detailed age and sex presentation using a complex measure of inequalities (Relative Index of Inequality) unveiled peak all cancer inequalities occurring at age 55 and 65 for males and females respectively. The socioeconomic burden of cancer which leads to lower socioeconomic groups having greater cancer incidence risk seem to suggest that the incidence/risk is brought forward 10 to 20 years earlier than the age-specific distribution of cancer incidence (Chapter 3).

At site and morphology level, the results of this thesis presented in Chapter 3, identified that laryngeal (of all UADT sites) cancer contributed significantly to the socioeconomic inequalities of all cancer and more so than lung adenocarcinoma. These observations are supported by (Boscoe et al 2014; Purkayastha et al 2016), where laryngeal cancer was identified as the head and neck site with the greatest risk association with low SES.

In Chapter 4, the thesis results identified that no single SES measure (area or individual) dominated the SES risk association with elevated risk of lung and UADT cancers. Similarly, Spadea et al (2010) also identified that for men, an increased risk association for all individual and area SES measures with UADT cancer incidence; while for women, only poorer housing characteristics were

associated with elevated risk (Spadea et al 2010). Meijer et al (2013) also identified increased risk associations for lung cancer with specific SES measures including among those with low education, low disposable income, low occupational social class and those living in high area deprivation; the authors concluded that both area and individual SES measures were implicated (Meijer et al 2013).

This contrasts with others who have identified that education and income were particularly prominent SES risk factors, followed by occupational social class (Dalton et al 2008a). However, Hystad et al (2013) after conducting a study reviewing long-term SES adjusted for several individual SES measures and behavioural risk factors concluded that long-term area SES was the dominant factor. This analysis may importantly reflect the very long gestation period for lung cancer which was addressed in the studies conducted here by measuring SES at cohort entry, well before cancer diagnosis, and may reflect the compounding effects of physical area environment effects of pollution (Laurent et al 2007; Raaschou-Nielsen et al 2013)

In Chapter 5, the results indicated that renting local authority accommodation (relative to private home ownership) or being retired (relative to being employed) were associated with increased risk of lung and UADT cancer. Likewise, Dalton et al (2008), Anderson et al (2008) and Spadea et al (2010) observed similar results. These findings were also consistent with Baastrup et al (2008) who found for females, early retirement pensioners and those renting accommodation were at a greater risk of oesophageal cancer (Baastrup et al 2008).

The findings in Chapter 5 also concur with others who identified that behaviours, particularly smoking and alcohol consumption, largely attenuated the risk associated with lung and UADT cancers (Mouw et al 2008; Nkosi et al 2012). However, in Chapter 5, the novel addition of a multiple low SES indicator identified that the potential compounded effect of multiple low socioeconomic circumstances was an important risk association — particularly for LUADT cancer

risk where risk was elevated even after adjustment for all behaviours. Conway et al (2015) supported this finding with the quantification of the contribution of smoking and alcohol consumption at 61% to the elevated risk of head and neck cancer associated with low educational attainment, leaving the remainder unexplained (and likely socioeconomically driven). Similarly, when quantified in an earlier study for UADT cancer, an estimated 67% elevated UADT risk associated with low education was explained by behaviours leaving 33% of SES risk unexplained (Conway et al 2010a).

The results of Chapters 4 and 5 confirmed the temporal nature of SES exposure and cancer diagnosis and were supported by (Spadea et al 2010; Garcia-Gil et al 2014; Vohra et al 2016; Leuven et al 2016). These studies focused on different aspects of low socioeconomic circumstances measured over the life-course with differing findings. Hystad et al (2013) focused on long-term area SES and lung cancer risk and found that the addition of smoking attenuated the long-term area SES effect by 20%; all other additional variables had little effect. Sondergaard et al (2013) explored the family environment in childhood and lung cancer risk noting that, through the sibling analysis, their study suggested that low family circumstances in childhood had an effect on educational attainment and ultimately risk of lung cancer. Behren et al (2016) focused on the association of occupational prestige and social mobility over time with lung cancer risk showing low occupational prestige in men was associated with lung cancer independent of smoking behaviour and occupational exposures as smoking behaviour only partly attenuated the elevated ORs between lung cancer and occupational social prestige. Leuven et al (2016) evaluated a change in school education to explore its association with lung cancer risk and confirmed one extra year of education was associated with a 12% reduction for women and a 20% reduction for men. Finally, Vohra's rapid review of SES in childhood and cancer in adulthood found that both childhood and adult SES contributed to lung cancer risk, but adult SES was more powerful (Vohra et al 2016).

## 6.3 Explanations for socioeconomic inequalities in cancer incidence observed

The complex interdependent interactions of SES have been proposed to affect health through a number of causal theories to better understand the social determinants of health and inequalities in health (Marmot 2005). These include: i) **"Context vs composition"** ("place versus people") which is described as being about neighbourhood factors — both physical and social environments (Macintyre et al 2002) can encapsulate social capital (Kawachi et al 1997) and can also be described as eco-social theory (Solar et al 2010); ii) **Life-course** perspective which is proposed to include SES from early years and throughout life – considering both social mobility and cumulative effects (Sweeting et al 2015); and iii) **Selection or reverse causation** — whereby health or disease states drive SES risk associations with diseases (Kawachi et al 2000). Within these broad theories there are several proposed and studied explanatory pathways including: a) Access to healthcare, b) Exposure to behavioural risk factors (Galobardes et al 2006b), c) Psychosocial "stress" factors (Solar et al 2010), and d) Material factors (Krieger et al 1997). However, there is a fundamental driver of health inequalities which is only now receiving attention — political decisions — that impact on the distribution of income, wealth and power (Solar et al 2010; Beeston et al 2014b)

In the next sections, each of these causal theories and pathways will be briefly discussed in relation to the thesis findings on cancer risk.

### 6.3.1 Context versus composition

Context versus composition is often referred to as "people versus place" (MacIntyre et al 2002). Context focuses on the environment, both physical and social, to which individuals are exposed while composition focuses on the individual attributes or rather the collective attributes of the community. The social aspect of context or place is often referred to as "social capital" or "social cohesion" and refers to the social networks, trust and support from which individuals can mutually benefit (Marmot 2010; Kawachi et al 2017).

Section 1.3.1.2 describes the concept in more detail. Physical attributes of "context" or "place" refer to, e.g., the availability of parks, recreational facilities, access to transport, healthy food establishments and absence of crime, industrial related pollution and air pollution in general. However, researchers in this area, propose that the influence on health is not an either/or scenario but a "mutually reinforcing and reciprocal relationship between people and place" (Cummins et al 2007); i.e., both context and composition are relevant and interdependent. Individuals influence the place in which they live and vice versa; and both influence health. Chapter 3, as with most analyses of routine administrative health datasets (including cancer registry analyses), was limited to use of the area-based socioeconomic status variable Carstairs. In Chapter 4, SIMD was similarly used to reflect the area deprivation (Section 1.3.2.4). The findings of this thesis were consistent with the perspective that *both* people and place influence health. Chapter 4 identified that area deprivation remained significant for lung cancer risk even after adjustment for individual SES factors; neither area nor individual SES factors seemed to dominate. It could also be perceived that all factors, i.e., both area and individual attributes were important. Chapter 5 built on these findings showing that the compounded effects of multiple low socioeconomic circumstances (including area and multiple individual SES measures) were strongly associated with increased cancer risk. However, it was not possible to assess the relative weights. Moreover, data on the social networks of social capital was not directly measured in the studies performed for this thesis; it remains an important area for further research.

#### 6.3.1.1 Social capital

As discussed in greater depth in Section 1.3.1.2, social capital is also described as "social cohesion" or "community" and is perceived to contribute to socioeconomic inequalities in health through the psychosocial pathway where social capital is either limited or unavailable or stress arises from comparison of the social capital of different SES (Uphoff et al 2013).

In the strictest sense, social capital was not measured in this thesis. Nevertheless, the area measures of SES used in this thesis, the Carstairs and the SIMD indices, were used as a "surrogate" measure of neighbourhood socioeconomic circumstances which reflect to some degree the social environment which fosters social capital. The SIMD in particular reflects attributes of the neighbourhood including crime levels and transport availability. From these attributes of SIMD it is possible to infer attributes of "social capital". For example, those living in neighbourhoods with high crime levels or poor transport may refrain from healthy behaviours such as going for a walk, attending community events, or visiting friends. As such and for lung and LUADT cancer, the studies in this thesis (Chapter 4 and 5 respectively) found that area measures of SES alone were not adequate to identify all of the complex dimensions of SES suggesting that a neighbourhood or place contributed to SES exposure, but did not capture all aspects of its multifaceted nature. Social capital is a plausible explanation of the thesis findings.

### 6.3.2 Life-course

Life-course perspectives were discussed in 1.3.1.3 when defining socioeconomic inequalities in health. Events over key transition points during the life-course and the associated accumulated risk potentially assist in explaining how SES could impact on biological processes associated with disease. The interaction between SES, parental behaviours and developmental processes in the early years, a key transition point in the life-course, and the change in SES from childhood through to adulthood and on into older years along with cumulative SES exposures are both important transitions that affect health inequalities (Galobardes et al 2004).

In the context of this thesis, in Chapter 5, the combination of education level (reflecting earlier SES), occupational social class (reflecting mid-life SES) and housing tenure (potentially reflecting SES at later life) supported inferring a life-course perspective. While it was not possible to examine cumulative risks in terms of temporal accumulation, the study in Chapter 5 attempted to create a

novel measure of multiple SES at a cross-sectional time point. This included SES measures that were reached at an earlier time point (e.g. educational attainment). The compounded SES measure most closely reflected the multifaceted, dynamic, life-course nature of SES exposure and as such, in Chapter 5, identified that an elevated risk of LUADT remained even after adjustment for behaviours (smoking, alcohol, diet and exercise). Investigating social mobility, while previously implicated in both lung (Behrens et al 2010) and UADT cancer risk associations (Schmeisser et al 2010), was beyond the scope of this thesis; however it is an area that warrants further investigation.

#### 6.3.2.1 Social mobility

Social mobility over the life-course reflects the mobility of a person's social standing (Behrens et al 2016). Movement is not static, however, changing over the life-course (Galobardes et al 2004) and over critical stages of life's journey (Watt 2002; Galobardes et al 2004). Social mobility over the life-course can be considered either "vertical", that is, across socioeconomic status strata (upwards or downwards) within a country or "lateral" between countries (Susser et al 1997).

Susser (1997) discussed social mobility occurring" laterally" as well as "vertically", explaining that migration either within or between countries was a form of social mobility. Along with movement of people, cultural norms, behaviour and disease also migrate (Susser 1997). As an example, Susser (1997) summarised the smoking epidemic development, starting with the upper classes among men in the 19th century, encouraged during World War I through distribution of cigarettes among working class men and upper class women and ultimately, a common habit predominantly of the lower social classes. The decline of smoking in the UK followed a similar pattern, first the upper class ceased smoking; the reduction of smoking among the lower classes followed.

In the context of this thesis, social mobility (either lateral or vertical) was not within the scope of investigation. However, and in terms of "lateral" social mobility, as described in Chapter 5, Scotland's population is relatively stable and

although migration has increased in the last decade, the proportion of the population that are not from Scotland or other UK nations, is relatively small and unlikely to significantly impact on socioeconomic inequalities in lung and UADT cancer risk. However, the uptake and decline of the smoking epidemic as described by Susser (1997) is a very feasible explanation for why smoking remains prevalent among the more disadvantaged in Scotland. Hiscock et al (2012) identified that over time, there were many more never smokers among the affluent group (around 50%) and many fewer never smokers among the more disadvantaged group (around 40%); reflecting the differential between the social groups in shunning the habit.

In the context of this thesis and "vertical" social mobility, in Chapter 5, the addition of the novel multiple indicator of socioeconomic circumstances did capture the cumulative nature of social mobility over the life-course. Educational attainment is normally achieved at an earlier stage of the life-course while occupational social class indicates socioeconomic circumstances during the working life of adulthood and finally, housing tenure is a measure of accumulated wealth and socioeconomic circumstances at a later stage of the life-course. Together these socioeconomic measures enabled a cumulative perspective over the life-course and may be interpreted to infer social mobility.

#### 6.3.2.2 Biological programming

Barker (1991) theorised that poor health outcomes reflected socioeconomic circumstances and accumulated disadvantage or advantage throughout life (Barker 1991). They occurred as a result of 'biological programming' and accumulated health and social disadvantage experienced earlier in life or even in earlier generations. As a result, poor health outcomes such as a illness (including cancer diagnosis) were proposed to be more likely (Barker 1991; Marmot et al 2012).

A related theory was proposed by Hystad et al (2013) where the wear and tear on the body that accumulates as an individual was exposed to chronic stress that

resulted in physiological consequences such as fluctuating or heightened neural or neuroendocrine responses (Hystad et al 2013).

Although not directly measured in the studies performed for this thesis, these explanations are also plausible and support the findings of Chapter 5 where multiple low socioeconomic circumstances were associated with elevated LUADT cancer risk before and after adjustment for behaviours.

In the context of this thesis and its aims and objectives, as an explanation for the findings identified, the life-course and accumulated risk theories are compelling. The multiple measures of area and individual SES employed in Chapters 4 and 5 as well as the multiple low socioeconomic circumstances indicator developed for Chapter 5 reflected different aspects of socioeconomic circumstances at different stages of the life-course and demonstrated elevated lung, UADT cancer risk in Chapter 4 and LUADT cancer risk in Chapter 5. However, as mentioned, data availability, or lack thereof, somewhat limited full exploration of these areas.

### 6.3.3 Selection

Theories of selection focus on reverse causation, i.e. that poor health causes lower SES (Kawachi et al 2000). Selection has been described such that individuals sort themselves into neighbourhoods and social groups – and that for example those who value physical activity or smoking may respectively select to live in neighbourhoods with parks, or among social groups who smoke (Kawachi et al 2000). Kawachi et al (2000) goes further and argues for example that good health and high IQ are genetically related thus explaining why individuals with high educational attainment are generally in higher socioeconomic groups – although there seems to be limited empirical data to support this. A raft of literature consistently suggests that SES exposures do influence health directly (Kawachi et al 2000).

The results of this thesis with the consistent strong findings of low SES and increased cancer risk across multiple area and individual measures of SES make

it difficult to conclude that reverse causation is the nature of the relationship, particularly in Chapters 4 and 5 where a temporal relationship can be determined with SES measured well before cancer diagnosis. Nevertheless, the data could not entirely rule out elements of reverse causation. For example, low educational attainment could itself be caused by underlying childhood health that could also be involved in the aetiological pathway of the disease. In terms of cancer risk, and the cancers of focus in this thesis, this seems unlikely to be a dominant pathway. However, the studies in this thesis did not have unobserved third variables such as IQ or time preference (whether one places emphasis on their present or future wellbeing), which have been proposed as explaining some of the SES attributes such as educational attainment (Kawachi et al 2010).

### 6.3.4 Access to healthcare pathway

Mclaren et al (1998) used Scotland's health data to explore deprivation and health. They identified issues that were related to the health service, such as the provision and quality of health services, access to health services and utilisation of health services. They concluded that the distribution of all these influencing factors on health was not equal across all people in society (McLaren GL et al 1998). Car ownership, a variable incorporated in area measures used in Scotland may infer convenient access to health care, particularly in rural environments (Berkman et al 1997). In the context of this thesis, access to health care was not directly measured but could also be tenuously inferred from the car ownership SES variable. Chapter 4 established that lack of car ownership was associated with elevated risk of lung and UADT cancer particularly for males and as such was a marker of material wealth and a resource for enabling access to work, school, shops, leisure activities, doctor's visits, friends and family (Pevalin et al 2008).

Related to access to health care services is the uptake of screening, both of which have been reported to be related to socioeconomic circumstances and follow the consistent pattern of lower access associated with lower socioeconomic circumstances (McLaren GL et al 1998; Boscoe et al 2014)

(Netuveli et al 2006). Only a fraction of all cancers can be detected by screening programmes (and some not reliably (Oliveira et al, 2014)); the vast majority of cancers are diagnosed when individuals with symptoms visit their GP or dentist but this is also subject to unequal access (Netuveli et al 2006). Lung and UADT cancers fall into the group of cancers with no current screening programmes. Netuveli et al (2006) concluded that opportunistic screening for oral cancer in general dental practices in the UK was unlikely to be effective as a prevention strategy given their study identified that the probability of regular dental attendance was low in all groups with a higher risk of oral cancer (Netuveli et al 2006).

Given the high incidence rates for lung cancer in Scotland (Section 1.6.3) and as part of the Scottish Government's "Detect Cancer Early" programme, a pilot lung cancer screening programme was developed in 2012 with the objective of increasing early detection by 25% (Scottish Cancer Prevention Network 2012). The target population for the pilot were those most at risk, that is, individuals who have smoked 20 cigarettes or more per day over 20 years or longer. As a result of this definition, those individuals were more likely to be in lower socioeconomic groups. In the context of this thesis, screening in general while fundamental to better chances of survival due to earlier detection, does not explain the greater *risk* of lung and UADT cancers association for the more deprived. Alternative explanations must be pursued.

### 6.3.5 Behavioural pathway

Several theories exist to explain the association between SES exposure and cancer incidence; the most frequently offered as an explanation are behaviours, i.e. socioeconomic factors influence behaviours such as those explored in this thesis (smoking, alcohol consumption, diet and exercise) which in turn lead to cancer (Marmot 2010; Marmot et al 2012). Many (Gray et al 2009b; Lawder et al 2010; Eberth et al 2014; Gupta et al 2015) have explored further this relationship and have identified that while unhealthy behaviour is certainly a cause of increased cancer risk, these behaviours can be socially determined.

Individuals do make decisions on how to behave; however, these choices are heavily influenced and shaped by the socioeconomic circumstances experienced through out life and are ultimately, the consequence of those circumstances (CSDH 2008; Watt 2002; Watt 2007).

It is clear that in Scotland, prevalence of the key risk behaviours associated with lung and UADT cancer, i.e., smoking, diets low in fresh fruit and vegetables and lack of exercise were socially patterned with the greatest prevalence among the most deprived (Lawder et al 2010). The clustering of multiple unhealthy behaviours was also socially determined (Lawder et al 2010). The most recent available data from the 2015 Scottish Health Survey among adults 16 years old or older in Scotland presented a strong gradient in smoking prevalence across SIMD quintiles where smoking prevalence increased from 11% in the least deprived quintile to 35% in the most deprived quintile despite substantial decrease in smoking rates in the Scottish population overall since the introduction of the public space ban (ScotPHO 2017c).

A similar gradient existed for the consumption of five or more portions of fruits and vegetables a day by adults 16 years old and older. In the least deprived SIMD quintile 26% of adults achieved the target while only 15% of those in the most deprived quintile met the five-a-day goal in 2015 (ScotPHO 2017a).

Also based on the 2015 Scottish Health Survey, the percentage of adults aged 16 years and over who met the physical activity guideline of 30 minutes activity five-times a week demonstrated a very clear gradient with decreasing proportion of men and women meeting the physical activity recommendations with each increase in deprivation (ScotPHO 2017b).

Using household income as the SES measure, the most recent Scottish Health Survey report (2015) indicated that the mean weekly unit consumption of alcohol for both men and women who were hazardous/harmful drinkers (consuming more than 14 units per week) demonstrated a clear gradient increasing from 28.3 units (men) or 22.9 units (women) for the least deprived to

54.8 units (men) or 49.1 units (women) for the most deprived (Brown et al 2016). These clear existing and persistent SES gradients in unhealthy behaviours were consistent with the three-fold elevated LUADT cancer risk for those in highest category of multiple low socioeconomic circumstances identified in Chapter 5.

### 6.3.6 Psychosocial pathways

The WHO conceptual framework for action on social determinants of health summarised by Solar et al (2010) describes the psychosocial cause of disease as where an individual experiences or perceives differences between his/her social circumstances compared to others, causing stress which in turn leads to poor health (Solar et al 2010). This perceived inconsistency also contributes to a sense of lack of control undermining health outcomes (ICOHIRP 2015). Psychosocial factors, along with behaviours, biological factors and finally material circumstances, are considered intermediary determinants of the social determinants of health (Solar et al 2010).

#### 6.3.6.1 Status inconsistencies

As reported by Behrens et al (2016), status inconsistencies are defined as loss of status control or the incongruity between an individual's expected and actual SES which affects that individual's psychosocial response potentially resulting in chronic stress, mental health issues such as depression, loss of job control and reduced social support (Behrens et al 2016). These factors in turn may impact on material circumstances (Behrens et al 2016). Although perception and experience of status inconsistencies were not evaluated in this thesis, this explanation was conceivable as through several decades, the income inequality in the UK, as measured by the Gini Coefficient (1.0 is most unequal and 0 is equal), widened most dramatically from 0.24 in 1979 to 0.34 in 1991, during the 1980s, making the differences between social strata more stark. A new high of 0.36 was reached in 2009-10; income inequality remained around this level until 2015-16 (Wilkinson et al 2009). Given the long lead-time, measured in terms of decades, between cancer initiation and diagnosis, it is notable that the rate of increase in inequality in the UK was greatest nearly 40 years ago. In a more

global context, the UK's relative position today remains highly unequal as reported in a House of Commons Library Briefing Paper No 7484 on Income Inequality in the UK (2016). OECD countries were ranked based on the Gini Coefficient, the UK was ranked sixth most unequal (0.385) out of 36 countries while Chile ranked first (0.465) and the USA ranked third (0.394). Those with the lowest levels of income inequality were Iceland (0.244), Norway (0.252) and Denmark (0.254) (Dorling 2017). For Scotland, the effect of the policies driving the increased UK income inequality was rapid de-industrialisation which, within the UK, particularly impacted Scotland (Walsh et al 2010).

These observations were consistent with the greater incidence and earlier diagnosis of cancer among the more deprived populations identified in Chapter 3. Likewise, in Chapter 4 increased risk of lung, UADT and all cancer was identified for those who were economically inactive relative to those who were active, while in Chapter 5, LUADT cancer risk remained elevated by 50% after adjusting for all behaviours for those who rented a local authority accommodation compared to home owners. In both Chapters 4 and 5, the temporal relationship between SES exposure and diagnosis of cancer was confirmed and is consistent with the widest income inequalities occurring in the 1980s, as measured by the Gini Coefficient, and the possible initiation of "status inconsistencies" during this period.

### 6.3.7 Material pathways

Material circumstances include income, wealth, living conditions, such as housing and neighbourhood quality, as well as the ability to buy healthy food or warm clothing (Solar et al 2010; ICOHIRP 2015). They also include physical working conditions, e.g., occupational hazards such as exposure to asbestos (IARC 2012b). The WHO conceptual model for social determinants of health considered material circumstances as one of the intermediary determinants of health (Solar et al 2010). In this thesis, material pathways can be inferred through three of the SES variables used: housing tenure, car ownership and economic activity (although no income data were available). In Chapter 5,

housing tenure was one of the SES variables that remained associated with elevated LUADT cancer risk even after adjustment for all behaviours (smoking, alcohol, diet and exercise). Those who rented accommodation from a local authority experienced a 50% greater risk of LUADT cancer compared to those who owned their own home. Likewise, those who had no access to a car were twice as likely to be at risk of LUADT cancer compared to those who had access to one or more cars; only when adjusted for smoking, alcohol and diet behaviours was the risk attenuated. Finally, those who were retired had a 96% elevated risk of LUADT cancer which was only attenuated after the addition of all behaviours. Through these three SES variables, the material pathways theory of inequalities does appear to be upheld.

### 6.3.8 Social production of disease/political economy of health

The WHO conceptual framework for action on the social determinants of health identified social production of disease/political economy of health as one of the three fundamental theories of social determinants of health summarised by Solar et al (2010). This theory focuses on the political and economic determinants of health. Social production of disease/political economy of health is not just the perception of inequality (as reflected in the psychosocial pathways), but also the structural causes of socioeconomic inequalities which result in unequal resources and systematic lack of investment in community infrastructure (Solar et al 2010). Health inequalities are widely regarded as political decisions or rather the impact of political decisions, and conversely if such decisions create them – therein lies the solution (Woodward et al 2000) as NHS Health Scotland point out “health inequalities are avoidable because they are rooted in political and social decisions” (Scottish Government 2014).

#### 6.3.8.1 Political causes

Health inequalities are political. Political decisions fundamentally create the socioeconomic environment of societies, and by extension political decisions are at the route of resolving them.

The core drivers of health inequalities, as recognised in the NHS Health Scotland (2015) policy review, are inequalities in income, wealth and power.

Income inequalities within and between countries have been associated with a wide range of health and social outcomes including ultimately life expectancy (CSDH 2008; Wilkinson et al 2009). Inequalities in wealth (assets) have been brought into focus by the work of the economist (Piketty 2013), who described widening inequalities in wealth as almost being built into the current capitalist society model with return on wealth assets outstripping achievable income gains from labour/productivity in the long-run. The NHS Health Scotland report focuses on the role of wealth (rather than only income — which has been the traditional financial consideration/metric of SES used) as a fundamental cause of health inequalities. They report wealth inequality in Scotland (2012-14) as being high — with the wealthiest 10% of households owning nine point four-times the wealth of the bottom 40%, and that this inequality continues to widen while mobility between wealth bands continues to slow down.

The World Health Organisation (2010) defines four types of power, all of which are based on an individual's relationship with themselves or others: i) *Power over* is the most sinister power where a more powerful group of society forces another, weaker group to act in a certain way. It is associated with coercion, domination and oppression, ii) *Power to* is the ability to change an existing structure or chain of events, iii) *Power with* is the ability to collectively work with others in order to influence an outcome and finally, iv) *Power within* is the ability, on an individual level, to be in control, to make decisions yourself of your own will (Solar et al 2010). Realisation of inequality of power is an important aspect of health inequalities (alongside wealth and income inequalities). It is reflected in NHS Health Scotland's recent policy review (Beeston et al 2014b).

In the context of political causes of socioeconomic inequalities, Ottersen et al (2014) described these concepts in terms of: i) Democratic deficit, ii) Weak accountability mechanisms (the inability to influence), iii) Institution

"stickiness" or worse, changing rules (laws) to address the needs of the powerful alone and to the detriment of everyone else, but to the greatest extent, the most vulnerable, and iv) Inadequate time focused on the impact on health policy of policies in other non-health areas (Ottersen et al 2014).

The premise is that this lack of equity in distribution of power in society, distorts the distribution of health such that "social norms, policies and practices that tolerate or actually promote unfair distribution of, and access to, power, wealth, and other necessary social resources create systematic inequalities in daily living conditions" (UCL Institute of Health Equity 2013).

William Farr, who lived in the late 19$^{th}$ century was a British epidemiologist and has been regarded as one of the founders of medical statistics. He identified that socioeconomic circumstances were a direct result of physical and political conditions (Kawachi et al 2006). This is interesting, given the focus of this thesis and also the methods of this thesis, as William Farr lends his name to the federation of data linkage research hubs across the UK including Scotland – the Farr Institute (Farr Institute 2017).The appreciation of the impact of policy on socioeconomic conditions is therefore not a new concept or unique to only certain populations. Thomas Jefferson once stated in 1809 that "the care of human life and happiness…is the only legitimate object of good government". The policies followed by a government are fundamentally linked to the socioeconomic circumstances of the population. WHO stated quite categorically that the primary responsibility for health equity was that of government (Solar et al 2010) and Dahlgren et al (2007) stated that economic growth should be seen as a resource for human development and not an end in itself (Dahlgren et al 2007).

In an attempt to move away from GDP as the single measure of national progress, the United Nations agreed in 2012 that "happiness" was the proper measure of social progress and should be a goal of public policy (Helliwell et al 2017). Likewise, in June 2016, the OECD sought to continue to redefine "growth" to include the population's well-being and to establish this wider definition as a

focus of governments' efforts (OECD 2016). Against these objectives, the Sustainable Development Solutions of the United Nations has been publishing a World Happiness Report annually for six years. The most recent, the World Happiness Report 2017 extended the focus on social drivers of happiness and as in the past reports, did not focus on a single health outcome but on "happiness". One hundred fifty five countries were compared using an index based on the interviews of 1,000 individuals in each country and available administrative data covering the period 2014-16 to establish a comparable index composed of six variables; GDP, social support, healthy life expectancy, social freedom, generosity and absence of corruption in government and business (Helliwell et al 2017).

The report presented some intriguing findings. The United Kingdom with a score of 6.714 (out of 10) falls to 19$^{th}$ in the ranking of the 155 countries with social support, followed by GDP, healthy life expectancy, freedom to make life choices, generosity and finally perception of corruption as the relative contributory factors explaining that score and rank. This represented a decrease of 0.172 compared to the 2005-07 score. The top ten countries (Norway, Denmark, Iceland, Switzerland, Finland, Netherland, Canada, New Zealand, Australia and Sweden) remained the same over the two periods, although the order was slightly rearranged. Very powerfully, the report quantified specifically for the UK the extent that misery could be reduced by eliminating key causes of misery — poverty (defined as below 60% of median income), lack of education, unemployment, being not partnered, physical illness (falling in the lowest 20%) and emotional health issues (falling in the lowest 20%). Based on 2014-16 data, elimination of emotional health issues was identified as potentially making the greatest contribution to improving happiness in the United Kingdom. This was also true of the USA and Australia. Furthermore, the report explored the effect of childhood on adult life-satisfaction using the British Cohort Study which followed children born in 1970 through to today. Considering intellectual development (highest qualification), behavioural factors, emotional health and family background, the strongest predictor of adult life-satisfaction was the child's emotional and behavioural health with mother's mental health being the

most important factor that influenced these outcomes. Based on data from Germany, the UK and Australia, the report identified that lack of income in and of itself does not create misery, but comparison of your income to others if your income was less (regardless of the level of income), does. This was also true of education level (Helliwell et al 2017). Finally, using data from the German Socioeconomic Panel, the report established that for both men and women life-satisfaction does not recover after any period of unemployment and the negative effect was greater for men than for women. This lasting negative impact of unemployment remains even after regaining employment, potentially because of the fear that it may happen again. On a macro level, rising unemployment negatively affects those in a job as well as those out of employment (Helliwell et al 2017), potentially due to fear for their jobs of those employed "at the moment".

In general, these potential explanations complement and reinforce the literature and research on socioeconomic inequality and cancer incidence presented in Chapter 2. Socioeconomic inequality can be profoundly reinforced or ameliorated by political policy; focus on income (GDP) alone is ineffective and even destructive. Opportunities to affect happiness and improved health outcomes, including reduced cancer incidence, begins with the socioeconomic circumstances of the child and the generation before but do not end there as the SES experience over the life-course must also be supported. This clearly includes the period of time in work as an adult (including employment, job type and job characteristics). *Any* employment does not lead to greater security, empowerment and enhanced control over life's circumstances with the consequent benefit of the opportunity to improve health outcomes, including cancer risk, but *good* employment can support this objective.

The World Happiness Report (2017) identified important lessons for socioeconomic inequalities in the UK, but it was measured before the most significant changes faced by the UK in 50 years, i.e. the credit crunch, austerity and Brexit as currently pursued by the UK government. The OECD Economic Survey for the United Kingdom (2017) however, is timelier, focusing on the

economic (and social) performance just before and subsequent to the June 2016 EU referendum. Their report warned that 45% of zero-hours contracts were held by individuals with low literacy, numeracy or both skills combined. Furthermore, over 25% of those employed have only low skills limiting job quality (EDRC 2017). Self-employment is also a large proportion of new job creation in the last quarter of 2017 in the UK, but as a non-standard form of employment, this was considered by the OECD as potentially detrimental to skill creation and job quality (EDRC 2017). Couple this starting point with the potential for economic shock as Brexit approaches and ultimately arrives, presents a very concerning outlook for socioeconomic inequalities in general, let alone cancer risk. Not surprisingly, the quality of life in the UK, as measured by the OECD, was close to or above the OECD average. In particular, social connections were significantly stronger, personal security was higher and environmental quality was better. However, well-being indicators showed significant inequalities between the high achievers and low achievers in terms of health status, jobs and earnings, education and skills, civic engagement and governance.

Unfortunately, the limited definition of social support applied in the World Happiness Report did not cover the full definition of social capital. The World Happiness Report focused only on family support, excluding the wider community networks considered part of social capital (Section 1.3.1.2). As a consequence, it is possible to argue that financial welfare should be reduced in order to force more of the deprived population to take any employment given that more than 50% of happiness is driven by non GDP factors (social support, generosity and life expectancy) and the finding that negative experience of unemployment is long lasting with a general impact on happiness. As a result, this interpretation could be used to support austerity policies.

Key requirements to support happiness, wellbeing and implicitly reduced cancer risk are the full range of dimensions of socioeconomic status and circumstances including education, income, occupational social class, housing tenure, employment status and area-based socioeconomic status all of which are reflected in the studies performed for this thesis. Issues not captured such as

work-life balance, autonomy, variety, job security, social capital and health and safety risks must also be reflected.

## 6.4 Overall thesis strengths

This thesis has been structured to evaluate both area and individual measures of socioeconomic inequalities in lung and UADT cancer risk from several perspectives. The studies involved epidemiological analysis, the population-wide Scottish Cancer Registry and cohorts developed via data linkage with the Scottish Longitudinal Study and Scottish Health Survey.

### 6.4.1 Temporal relationship

Given the importance of SES over the life-course and reflecting the long lead time from cancer initiation to diagnosis, the studies that were performed have measured socioeconomic inequalities at the earliest point possible and well before diagnosis, the most commonly used measurement point.

Reflecting references in the literature to the “compounding” effects of low socioeconomic circumstances, this thesis additionally and uniquely created an individual “index” of multiple low socioeconomic circumstances. This was an important contribution, despite being fairly crudely defined (Chapter 5), as it identified that facing multiple low socioeconomic circumstances was associated with elevated lung and UADT cancer risk, even after adjustment for all behaviours measured.

### 6.4.2 Multiple measures of SES

Recognising the potential role of individual and area socioeconomic circumstances, this thesis investigated both of these categories of SES indicators using multiple measures of individual circumstances to reflect as comprehensively as possible the three main SES indicators (education, income and occupational social class). Berkman et al (1997) reviewed these three indicators as measures of social class in health studies and although there are

limitations as well as strengths, they concluded that the three measures reflect Weber's three domains of social class: ownership and economic resources (class), prestige and community ranking (status) and political power (Berkman et al 1997). Furthermore, as outlined in Chapter 1 and remarked by Berkman et al (1997), virtually all reviews on SES gradients and health as measured by these three indicators report consistent and strong associations between SES morbidity and mortality (Berkman et al 1997).

### 6.4.3 Sophisticated measures of health inequality

This thesis also adopted for the initial descriptive population study (Chapter 3), the Relative Index of Inequality (RII), the measure considered most effective in capturing total impact, effect and extent of socioeconomic inequalities as it is not limited to only the two extreme social groups, but reflects the full social gradient and captures the direction and magnitude of those inequalities. Applying approaches used in Scotland to quantify the contribution of cause specific mortality to all cause mortality in Scotland, this thesis modified the RII to establish the relative contribution to all cancer risk, by cancer site and histology of the cancers of interest.

### 6.4.4 Behavioural data

As explored in detail in Chapters 1 and 2, individual behaviours are often used to explain the socioeconomic inequalities in lung and UADT cancer risk observed. In recognition of this, this thesis explored area and multiple individual measures of SES as well as individual behaviours (smoking, alcohol, diet and exercise) to quantify to what extent the observed socioeconomic inequalities could be explained by behaviours alone.

### 6.4.5 Scotland's epidemiology and research environment

Scotland provides a unique research environment fundamental to supporting the most desirable study design: i) A well-established and developing history of data linkage; ii) High quality administrative datasets such as the Scottish Cancer

Registry; iii) Availability of representative cohort research studies such as the Scottish Longitudinal Study; iv) Multiple cross-sectional representative surveys such as the Scottish Health Survey; v) Systems to support researcher access to these datasets; and vi) Robust comprehensive epidemiology, statistical and analytical approaches. Each of the studies performed for this thesis has taken advantage of these strengths in order to most accurately measure socioeconomic inequalities and their association with lung and UADT cancer risk. As a result, the ability to draw evidence based conclusions is optimised (Brewster et al 2002; Pavis et al 2015; ISD 2017a).

## 6.5 Overall thesis limitations

The thesis limitations are largely related to the limitations inherent in and to availability of the study data. This section first covers the main data interpretation issues, ecological fallacy and residual confounding, before detailing study population and data limitations.

### 6.5.1 Ecological fallacy

Ecological fallacy may remain as a limitation of this thesis given that area measures of socioeconomic circumstances are at risk of this source of bias. As discussed in Section 1.3.2.4 ecological fallacy occurs when individual level relationships are inferred from summary or aggregate relationships observed at area level. However, research has demonstrated that area SES effects exist independently of individual SES measures suggesting aggregation of individual attributes; although inadequate on their own, these do nevertheless contribute to the "neighbourhood" or "place" effect referred to by MacIntyre and Ellaway (2003). The characteristics of the neighbourhood including transport links, healthy food establishments, parks, lack of crime, availability of social infrastructures such as libraries and clubs are perfectly relevant aspects of the "compositional" attributes of a neighbourhood. This aspect has been reflected throughout this thesis. However, the approach taken in considering several individual SES variables in the cohort studies conducted for Chapters 4 and 5 (economic activity, educational attainment, occupational social class, car

ownership and housing tenure) did ensure that individual attributes were fully considered at individual level. Furthermore, areas identified as the most deprived all have the same SIMD and are likely to be very homogeneous (Bishop et al 2004).

### 6.5.2 Residual confounding

Remaining confounding of the behaviours measured, or unmeasured at all, on the identified socioeconomic exposure continues to be a limitation of this thesis. Many of the behaviours associated with lung and UADT cancer risk are indeed socially patterned themselves, so it is very possible that adjustment may not fully remove this confounding if all aspects of the behaviour are not reflected. This is particularly relevant in Chapter 5 where behaviours were considered as explanatory variables. An attempt was made to incorporate multiple variables and dimensions of behaviours, including interactions, to reflect as comprehensively as possible, each of the behaviours applied (smoking, alcohol, diet and exercise).

Residual confounding is also a limitation of the study undertaken in Chapter 4 where the relative importance of area and individual SES indicators was explored. It is possible that the individual socioeconomic indicators used (economic activity, education level, occupational social class, car ownership and housing ownership) and supplemented by the socio-demographic variables, marriage status and country of birth, did not fully capture all aspects of individual SES. As a result, the area deprivation measure used, Carstairs, may have remained significant only because individual SES measures did not comprehensively capture all aspects of individual SES. However, the associations identified were strong, suggesting that if this was an issue the effect was not strong enough to fully attenuate the relative risks. Nevertheless, an omission in the individual SES indicators measured would be income. As discussed in Chapter 1 (Section 1.3.2.3), this thesis could not measure income directly because the information is not reflected in any of the datasets employed. Income can, however, be inferred from occupational social class and education level. Ideally,

and to be as comprehensive as possible, income should be incorporated. This would require inclusion of questions regarding income in the relevant surveys. Most recently, this was unsuccessfully attempted in the 2011 Census.

### 6.5.3 Study population limitations

The Scottish Cancer Registry (Chapter 3) was a population study covering all of Scotland while the Scottish Longitudinal Study (Chapter 4) and the Scottish Health Survey (Chapter 5) studies used samples of the Scottish population. In terms of the population sample, this is less of an issue with the Scottish Longitudinal Study as it is based on randomly selected birthdates, limiting the prospect of sample bias. It should be noted that the primary outcome data used for the linkage of these studies was the Scottish Cancer Registry which is a robust dataset with high quality and completeness (Brewster et al 2002). Representativeness has been raised as a concern of the Scottish Health Survey (Gray et al 2013), however, the adjustment approach adopted (Chapter 5) is likely to address this concern.

#### 6.5.3.1 Selection and recall bias

As discussed in detail in Section 5.4.5.1, population selection bias is a more relevant issue for the Scottish Health Survey which by design does not include segments of the Scottish population more likely to be more deprived, for example, those living in residential care. Furthermore, because the Scottish Health Survey was the source of behaviour data and because that data is self-reported, it is at risk of “recall bias”; that is, more favourable (and socially acceptable) behaviours may be reported compared to the reality (Gorman et al 2017). This explanation for the results identified in Chapter 5 cannot be ruled out; however, when defining the cohort, records that had conflicting responses to different questions were omitted. Furthermore, multiple behaviour variables were used, where available and finally, even with a potentially more favourable behaviour response, SES was not fully attenuated in all cases; this suggests that the reported elevated risks may in fact be an under estimate.

#### 6.5.3.2 Summary of data interpretation limitations

While the data interpretation issues raised — ecological fallacy and residual confounding — may have persisted, nevertheless the approach to the number and type of variables selected were designed to minimise these issues. Furthermore, the associations identified were strong and in many cases, the relative risks were not fully attenuated after adjustment. The population limitations of selection and recall bias were also unlikely to undermine the conclusions made. To minimise these possible issues, multiple behaviour variables were used and where illogical or missing, omitted.

### 6.5.4 Data availability limitations

#### 6.5.4.1 Social mobility

Social mobility was beyond the scope of the studies performed for this thesis. Measurements of the relevant socioeconomic indicators taken at multiple points in time would be required but this was not feasible using the datasets employed in this thesis. However, to capture an aspect of the dynamic nature of SES, this thesis did capture a "surrogate" measure of early SES via education which reflects parental influence as well as transition to a more independent status as a young adult.

#### 6.5.4.2 Childhood socioeconomic circumstances

Likewise, the studies performed in this thesis were not able (Chapter 5 Scottish Health Survey linkage study) or were not designed (Chapter 4 Scottish Longitudinal Study linkage study) to capture childhood circumstances directly. Nevertheless, and as explained above, education does reflect, in part, childhood socioeconomic circumstances as parental educational attainment is known to influence the next generation's education level and attainment (Galobardes et al 2006b). School education begins at an early age and continues to young adulthood and therefore can be used to infer the socioeconomic circumstances of that period. Alternatively, parental SES could be used as a proxy for childhood SES.

#### 6.5.4.3 Unavailable risk behaviours

Finally, in the study performed for Chapter 5, there were some risk behaviours such as sexual behaviour associated with HPV infection which could not be considered given the datasets used and accessibility.

Lack of HPV infection data was identified as a limitation of the study performed in Chapter 5 of this thesis (Section 5.4.5.2). Approximately 60% of oropharyngeal cancer incidence is associated with HPV infection (Conway et al 2016). Consequently, omission of these data is an opportunity for improvement; however, there are no routine data collections of these data in Scotland (Conway et al 2016). Conway et al (2016) conducted a study in dental practices to assess the feasibility of undertaking a full population study in Scotland of the incidence, prevalence and persistence of HPV (Conway et al 2016). They concluded that through dental practice patients and using the approach piloted, such a study was possible. Their focus was on assessing the effect of Scotland's HPV vaccination programme targeting young women, but their approach, through linkage with other routine datasets may also be valuable in filling the HPV infection data gap identified here.

#### 6.5.4.4 Income data

Throughout this thesis, it has been acknowledged that income is an important missing dimension of socioeconomic circumstances not available in Scotland or the United Kingdom. The studies performed here did not incorporate important SES variables such as income, benefits or Department of Work and Pensions data which could not be considered given the datasets used and accessibility.

As described in Section 1.3.2.3, income data are an important measure of socioeconomic circumstances that studies performed in Scotland and the UK are unable to reflect except indirectly through occupational social class and educational attainment or through receipt of welfare benefits. The studies performed in this thesis (Chapters 3, 4, 5) used:

- SIMD (Chapter 3);

- Carstairs area deprivation measure with five individual SES indicators (economic activity, occupational social class, educational attainment, car ownership and housing tenure) and two socio-demographic variables (country of birth and marriage status) (Chapter 4); or

- SIMD with the same five individual SES indicators (Chapter 5).

Individual SES indicators retained an association with cancer risk in Chapter 4 after mutually adjusting each SES variable for the other SES variables or in Chapter 5 after full adjustment for the four behaviours (smoking, alcohol, diet and exercise); it is thus possible that residual confounding had occurred. Inclusion of income would enhance the understanding of the relationship between socioeconomic inequalities and cancer risk. Income along with education, wealth and occupational social class are key determinants of population health and as confirmed through this thesis, cancer risk. Better and more comprehensive SES indicators, including income would support a more accurate description of current circumstances and assist in shaping policies to promote better health and address the role of socioeconomic inequalities as a determinant of that health status. Possible avenues for obtaining this information are to revisit the introduction of household income in the 2021 census and to use, through linkage, the Scottish Household Survey data on household income. Both of these options, however, would not capture the effect of income volatility because of the frequency of implementation. Income is the primary SES indicator which reflects the dynamic nature of SES (Galobardes et al 2006b), an important attribute of SES. Opportunities to capture this aspect should be considered; the new social security powers recently devolved to Scotland may present a complementary window of opportunity (Section 6.5.4.4).

In addition, the devolution of income tax and limited social security powers to the Scottish Parliament gives new opportunities for data linkage. Many of the social security devolved benefits (i.e. Attendance Allowance, Carer's Allowance,

Disability Living Allowance, Personal Independence Payment, Industrial Injuries Benefit, Severe Disablement Allowance, Cold Weather Payment, Funeral Payment, Sure Start Maternity Grant, Winter Fuel Payment and Discretionary Housing Payments (Wane et al 2016)) have been paid to people in poor health and in low-income households. This new Scottish based social security system and developing dataset presents an opportunity to use Scotland's data management capabilities to link social security data to administrative and survey health data. This would support future research to further understand the relationship between socioeconomic inequalities and health as well as the effectiveness of health and economic policies designed to improve it.

## 6.6 Recommendations/further research

The recommendations from the thesis findings include those relating to: i) Further research including suggestions for work that could address some of the limitations identified in 6.5; ii) Monitoring inequalities in cancer incidence; and finally iii) Informing approaches to tackle inequalities in cancer incidence.

### 6.6.1 Recommendations for research to address identified limitations

#### 6.6.1.1 Research support and data linkage infrastructure

***Research support***. As discussed in Section 6.4.5, Scotland's research environment is supportive and uniquely provides opportunities to investigate important public health issues such as socioeconomic inequalities and cancer risk. However, fully understanding the linkage context is required in order to interpret, refine and develop further research studies.

In the context of this thesis, linkage was performed for the Scottish Longitudinal Study (SLS) — Scottish Cancer Registry study (Chapter 4) with support from the SLS Development & Support Unit (SLS-DSU) in a safe-haven setting at the National Records of Scotland in Edinburgh. By contrast, the Scottish Health Survey — Scottish Cancer Registry study (Chapter 5) was performed using a "virtual" safe haven with support from the electronic Data Research and

Innovation Service (eDRIS). Two different approaches to supporting the researcher were evident with benefits and limitations experienced in both situations. For the Scottish Longitudinal Study – Scottish Cancer Registry study, the support provided by SLS-DSU was very hands on and beneficial for a PhD student working directly with a large unfamiliar dataset for the first time. As a result of this physical interaction, a supporting research officer who was an expert in the Scottish Longitudinal Study dataset and that officer's availability during analysis sessions, it was possible to address questions comprehensively and timeously. However, the physical safe haven and the requirement to always organise analysis time which was mutually agreeable was limiting.

This contrasted significantly with the experience of the Scottish Health Survey – Scottish Cancer Registry study where support was as "virtual" as the "virtual" safe haven. Interaction with the supporting research officer was never in person and always via email and/or telephone. This coupled with staff change-over during the study development, initiation and implementation phases resulted in a less cohesive service. On the other hand, the "virtual" safe haven, where researchers could access the data for analysis remotely, was far more convenient, a more efficient use of time and potentially much less expensive in terms of travel costs.

In the context of both study experiences, it would be desirable in both cases, for "linkage metrics" to be provided as part of the package of supplying the data such that the researcher was aware of the success rate of linkage for their specific study. Such a report could provide an analysis of the number and the proportion of records for the datasets, the numbers of records linked and the basis of that linkage.

***Maintenance of data linkages***. The linkages between the Scottish Cancer Registry and the Scottish Health Surveys and the Scottish Cancer Registry and Scottish Longitudinal Study performed for this thesis should be updated and performed on a regular basis enabling greater statistical power, longer follow-up and the additional fields proposed under Section 6.6.1 to be captured.

#### 6.6.1.2 Scottish Health Survey recommendations

Based on the study performed using the Scottish Health Survey, there are opportunities for improvement that would assist researchers and the value of the research performed in the future. These are described as follows.

***Response rate***. The Scottish Government, which manages the survey, should consider means of improving the Scottish Health Survey response rate which currently is close to or at 83% — this does represent a fall from 94% achieved at the first Scottish Health Survey in 1995 (Section 5.1). Although the response rate has been stable for the six surveys conducted since 2008, further degradation should be avoided and strategies to maintain and improve the response rate should be considered. The proportion of those who do agree to participate who also consent to linkage is exceptional at over 99% and presents a significant asset for research dependent on record linkage.

***Variable definition consistency***. As highlighted in the Sections 5.2.2 and 5.2.3 and discussed as a limitation (Sections 5.4.5.2, 5.4.5.3) of the study conducted and described in Chapter 5, consistency of variables over time is critical for development of cohort studies, the most desired epidemiological study design. Cohort studies overcome design limitations which are associated with cross-sectional studies (e.g. measurement error assessment, residual confounding from unmeasured variables, inability to incorporate suitable time-lag period and ecological fallacy) and case-control studies (e.g. ability to select a suitable controls and to obtain unbiased measures of past exposure) (dos Santos Silva 1999).

The earlier survey years (1995, 1998 and in some cases 2003) in particular did not use the same questions to record behaviour and socioeconomic circumstances as the later surveys. As a result, some important refinements available for the later surveys are missing all together (e.g. binge drinking in 1995) or were not captured in the same level of detail (e.g. physical exercise). It is recommended that variable definition (as adopted for the later 2008, 2009,

2010 and 2011) is more consistent and that this is maintained for all future Scottish Health Surveys.

***Sample selection***. As discussed in Section 5.4.5.1, the Scottish Health Survey targets individuals living in private accommodation. This creates a selection bias when exploring socioeconomic inequalities and when drawing conclusions, based on the Scottish Health Survey sample, about the Scottish population in general. In order to overcome the issue of selection, it is recommended that the sample incorporate individuals living in communal establishments as well. This may however, lead to poorer participation rates and poorer levels of consent to linkage which, as a consequence, may require larger samples.

#### 6.6.1.3 Further SES measures and variable attributes

***Economic activity definitions***. As reported in the recent economic survey for the UK conducted on behalf of the Organisation of Economic Cooperation and Development (EDRC 2017), a third of new jobs developed in the UK were identified as "self-employed" while a further third were described as "zero-hours contracts". These two categories of economic activity would both be summarised as "employed"; however, the implications for the quality of the jobs the individual holds as well as the consequent cost on health, including cancer incidence, is likely to be significant. Therefore, it may be valuable to consider capturing these two additional subcategories of "employed" in future surveys including the Scottish Health Survey, while supporting consistency of definitions between surveys.

***Parental SES measures/childhood socioeconomic circumstances***. As discussed in Section 5.4.5.3, parental SES measures to more fully capture intergenerational transfer of SES and associated health outcome risks are an important aspect reflecting SES during the early stages of the life-course. Future research using the Scottish Longitudinal Survey linked to the full census may support both identifying the parents of SLS participants and capturing the socioeconomic circumstances of the parents through measures such as father's occupation and mother's educational attainment (Galobardes et al 2006b).

Alternatively, this may be feasible through NHS birth records; however, this is likely to only offer parental SIMD which may be only focused on the mother and exclude the father. In any case, the principle would require linkage not between two administrative or survey datasets as was the case in this thesis (Chapters 4 and 5), but between multiple datasets. This type of complex cross-sectoral linkage has begun to be undertaken, including recently as a pilot study focusing on looked after children and their dental health needs in Scotland (McMahon et al 2017; Clark et al 2017) and is therefore feasible.

School examination attainment data are held by ScotXed, a unit which is part of the Education Analytical Services Division of the Scottish Government. School examination data and attendance could potentially provide information on childhood circumstances and may have important relationships with SES. Linkage between health and education datasets has been demonstrated by a recent study focused on looked after children and their dental care experience (McMahon et al 2017; Clark et al 2017).

***Stress measures***. The Scottish Health Survey collects biannually stress at work questions focusing on the experience of stress at work, as well as work/life balance and working conditions, beginning with the 2011 survey (the last survey used for this thesis). A recent consultation on the survey content recommended that these questions are included each year (APS Group Scotland 2017). In the context of measuring the outcome of low socioeconomic circumstances, this move is strongly supported. In addition, proposals suggested in the consultation document included extending the questions to reflect the participant's contract status (e.g. self-employed, zero-hours contract), whether the participant had multiple jobs and whether the respondent had experience of benefit sanctions (APS Group Scotland 2017). The Labour Force Survey (Office for National Statistics 2016), is conducted quarterly for all of the UK and is the basis of the official statistics for unemployment and employment. The Labour Survey does reflect this type of contract question; however, the survey samples private households and is based on telephoning or face-to-face interviews of 38,000 people in all of the UK. Furthermore, response rate was around 45-50%.

Alternative approaches for capturing this important information should be considered. As a regular survey, the Scottish Health Survey may be a potential source. As discussed in Sections 5.4.5.1 and in 6.5.4, because the Scottish Health Survey targets private households, those living in communal accommodation and more likely to experience low socioeconomic circumstances may be omitted; options to include these individuals in the sample should be pursued.

Alternatively, possible linkage with the Labour Force Survey could be considered, recognizing the relatively poor response rates identified and adjusting via weighting accordingly. However, as indicated and in the context of investigating socioeconomic inequalities, the Labour Force Survey omits an important segment of the population, those not living in private households.

Other measures of stress that could be considered in the future include anxiety prescribing. Through the ePharmacy Programme which records the unique patient identified CHI on prescriptions made by GPs and dispensed by pharmacies it is possible to identify which prescriptions have been dispensed for individual patients. These data are becoming more complete in recent years and the CHI capture rates currently vary by prescriber type, geography and type of drug; however the CHI capture rate for GP prescribing is over 95% (ISD 2017b).

#### 6.6.1.4 Multiple low socioeconomic circumstances index and multiple measures of SES over time

This thesis identified that, for those facing multiple low socioeconomic circumstances, the incidence risk association of lung cancer was more than three-times the risk for those with no low socioeconomic circumstances. This risk remained 86% elevated even after adjustment for smoking, alcohol, diet and exercise. This measure captured the clustering of low socioeconomic circumstances, an important feature of SES with compounding negative effect on health. To further enhance this indicator, it is proposed that the scope of individual SES indicators should be extended as further SES indicators are developed or become feasible to include. These new SES variables may include those explored in Section 6.6.1.3, that is, household income level, refined

employed status categories (self-employed and zero-hours contracts), parental SES measures and stress measures. Consideration should also be given to a more refined calculation method that reflects the relative importance of retirement (economic activity) and renting accommodation from a housing or local authority (housing tenure) given that these two individual indicators were associated with elevated lung and UADT cancer risk, even after adjustment for behaviours.

Multiple measures of SES circumstances over the life-course was not possible in the studies conducted for this thesis, however, this is an important aspect of SES; further research focusing on this aspect of cancer risk's association with socioeconomic circumstances should be pursued.

## 6.6.2 Recommendations for further research

A number of options are open to further wider research in relation to updating current work, methodological research, exploration of wider factors and research into developing and evaluating specific interventions to tackling inequalities in cancer incidence.

### 6.6.2.1 Extension of the definition to include a wider definition of equality

This thesis defined socioeconomic inequalities in terms of area and individual SES indicators of inequality. However, the literature also examines racial or ethnic inequalities as well as inequalities in age, gender, disability, sexual orientation, religion and other vulnerable groups — to embrace these dimensions of diversity (SCOTPHO 2010; Gordon et al 2010). These dimensions have had relatively limited attention in relation to cancer incidence inequalities and would be worthy of further research.

### 6.6.2.2 Wider areas of research

Section 6.5.4 identified a number of data limitations of this thesis. These included: social mobility, social capital, childhood socioeconomic circumstances, parental SES, unavailable risk behaviours (HPV infection) and other income related data (e.g. benefits or Department of Work and Pensions data). This

thesis, given the cross-sectional design of the Scottish Health Survey was not able to examine cumulative risks in terms of temporal accumulation – and to fully assess inequalities across the life-course. These limitations present opportunities for further research into cancer incidence inequalities. In addition, there are opportunities to link social epidemiology perspectives with other fields of research from genetic, biomarkers and behavioural research. The concept of more fully defining these environmental exposures has been described as the "exposome project" to complement the "genome project" which aimed to map the whole genomic code (Wild 2012).

### 6.6.3 Monitoring cancer incidence inequalities

The Scottish Government established a technical advisory group in early 2008 to counsel a Ministerial Task Force on health inequalities which was led by the Minister for Public Health. This group explored further ways to monitor progress in reducing inequalities over the long term (Scottish Government 2008d). They identified not only cancer mortality as a key measure, but also cancer incidence, focusing on all malignancies excluding non melanoma skin cancer (Scottish Government 2008d). Alternative measures such as Slope Index of Inequality (SII) and Relative Index of Inequality (RII) were used to demonstrate no change over time in the inequality gradient for the incidence of all malignancies (excluding non melanoma skin cancer). As explored in Section 1.4.3, this approach is desirable as it measures the full SES gradient.

However, as demonstrated through this thesis (Sections 4.4. and 5.3.2) and as established by other research studies discussed previously (Section 2.2), monitoring of all cancer is unlikely to identify socioeconomic inequalities given the breadth of cancer sites included with each of these sites reflecting complex and different aetiologies and different SES relationships. As an example, for lung cancer, the cancer with the greatest number of incident cases in Scotland, greater incidence is associated with lower socioeconomic circumstances, while breast cancer, which is also among the top four incident cancer cases in Scotland, is more likely to be diagnosed among the least disadvantaged. Given

evidence of different SES gradients depending on the cancer site, opportunities remain to apply these alternative measures (RII and SII) at the level of cancer site, or a grouping of cancer sites known to make significant contribution to cancer risk inequalities.

The technical advisory group also recommended the combination of income and employment domains of the Scottish Index of Multiple Deprivation (SIMD) as a more sensitive measure of deprivation in comparison to the full SIMD index. This recommendation was first adopted in the most recently published inequalities monitoring document (Scottish Government 2017c). However, based on the evidence presented in this thesis (Section 1.3.2.4), this approach does not recognise or reflect the multidimensional aspect of SES, but instead emphasizes income and employment status alone.

As demonstrated in Chapter 5, the novel indicator of multiple low socioeconomic circumstances was a powerful tool for identifying SES inequalities. Given that the Scottish Health Survey is an annual survey, it may be more relevant to adopt multiple individual SES measures as well as the area measure SIMD and to monitor a cancer group known to contribute significantly to all cancer inequalities, i.e. lung and upper aero-digestive tract cancers (Chapter 3). However, multiple years of data are likely to be required, given the Scottish Health Survey sample size. As a result, this is likely to preclude adopting such an approach as expanding the survey is unlikely to be suitable given cost constraints. Nevertheless, a rolling average of the most recent five years may provide an adequate volume of cases and support, at minimum, an appreciation of the current position. Given the long lead-time between cancer initiation and diagnosis is decades, it may not be expected that immediate or even medium term reduction in cancer incidence could be achieved with suitably individually targeted interventions that recognised the underlying socioeconomic pattern of behavioural risk factors although repeating the outcome analysis would certainly give a bigger cohort (with more cancer cases developing).

## 6.6.4 Tackling cancer incidence inequalities

### 6.6.4.1 Cancer policy

Epidemiological data are critical to highlighting the scale of this public health problem. The thesis findings contribute to this large and growing body of work in relation to health inequalities, which is relatively more limited for cancer-specific research.

A report from Macmillan recently highlighted widening inequalities in cancer survival in Scotland, likely to be more related to access to care (Dhanda 2014), but cancer incidence inequalities, which are likely associated with more fundamental issues associated with socioeconomic determinants (and even more outside the control of health services) were not included and are often left unaddressed in policy discourse.

In 2016, the Scottish Government published *Beating Cancer: Ambition and Action*, the strategy for cancer services. The document recognised that cancer incidence was more common in the more deprived areas of Scotland with 30% to 50% higher rates compared to the more advantaged areas. Behaviours and screening uptake were provided as explanations while the gap between least and most deprived was expected to continue to increase. The document recognised the social pattern of the key risk behaviours: smoking, alcohol, diet and exercise and mentioned policies focused on reducing exposure by the more disadvantaged. For example, the report identified that smoking prevalence among the most disadvantaged, while still higher than the most affluent, had decreased the most (from 39% in 2013 to 34% in 2014). In general, though, the document itself did not focus on socioeconomic inequalities in cancer incidence or provide a concrete plan for addressing these inequalities and was evidence of a silo approach to policy evidence, despite the stated aim of the Scottish Government to underpin all policy with equality. As outlined in 6.6.4.1, 6.6.4.2 and 6.6.4.5 there are several policies that the Scottish Government has developed that are focused on the structural determinants of socioeconomic inequalities. At minimum, reference to these other Scottish Government policies

is suggested, particularly since many of the interventions to prevent ill-health in general are relevant to preventing cancer incidence (Watt et al 2012).

#### 6.6.4.2 Behaviour-related policies

***Smoking***. The Scottish Parliament legislated for introducing a smoking ban on 26 March 2006 and increased the age of sale for tobacco from 16 to 18 years on 1 October 2007. The smoking ban established it would be an offence to smoke in any wholly or substantially enclosed public space in Scotland, with a small number of exceptions, such as prisons, care homes and police interview rooms. Building on this, *Scotland's Future is Smoke-Free* focused on dissuading children and young people from smoking. Actions were focused on a holistic approach to health and well-being in Scottish Schools and using Curriculum for Excellence, the Scottish Government's education action to support this objective. Actions were adopted to reduce the attractiveness (e.g. restrict tobacco product displays), availability and affordability of tobacco products, (Scottish Government 2008e)

***Alcohol***. The *Alcohol (Minimum Pricing) (Scotland) Act 2012* was passed in June 2012. The legislation has not yet been implemented due to a long legal challenge led by the Scotch Whisky Association. The UK Supreme Court ruled on the 15 November 2017 that the legislation that allows minimum unit pricing was lawful. As a result, the Scottish Government intends to introduce the new pricing approach on 1 May 2018 (Alcohol Team 2017). Minimum unit pricing will set a floor price for a unit of alcohol. The more alcohol a drink contains, the stronger it is and therefore the more expensive it will be. The most recent research estimated that the proposed minimum price of 50p per unit would result in a reduction in alcohol related deaths of around 120 per year (full effect) and hospital admissions would fall by 2,000 per year (full effect) when 20 years of implementation were achieved (Angus et al 2016).

***Diet***. In his last budget, the then UK Chancellor the Exchequer, George Osborne announced a *UK Soft Drinks Industry Levy* to tackle childhood obesity. The levy will be introduced in April 2018 and will have a lower rate which will apply to

added sugar drinks with a total sugar content of 5gr or more per 100 millilitres and a higher rate for drinks with 8gr or more per 100 millilitres. Alcoholic drinks with an Alcohol by volume of up to 1.2% are also included in the levy (HMRC 2016). The new levy will increase the cost of a can of soft drink by £0.06 to £0.08. While this new levy is a step in the right direction, Foods Standards Scotland, established in 2015, identified a number of limitations to the tax: i) The levy focuses on only one type of high sugar content foods and should be applied to other calorie-dense foods in order to address a wider target group beyond children; ii) Concern that the levy may not achieve re-formulation of soft drinks such that sugar content is reduced and may instead result in smoothing of costs across the high and low sugar product range; iii) Concern that as a result, the levy will have little effect on consumer behaviour; iv) The possible unintended consequence of the new levy restricting Scottish Government's options for wider sugar fiscal measures, particularly given that both Food and Health are devolved policy areas while taxation remains a reserved power; v) To be effective in changing consumer behaviour, Food Standards Scotland and Cancer Research UK believe the levy should be 20% (Food Standards Scotland 2017; Collinson 2017); and finally vi) Traffic light nutritional labelling should be made mandatory rather than voluntary (Food Standards Scotland 2017).

Peres et al (2017) identify a number of ***diet related policies*** which could also be considered. These include school food policy, vending machine content and hospital patient meals, all of which should reflect national diet guidelines. The presentation and placement of food in supermarkets, self-serve cafes and multiple product offers all affect choice and could be re-designed to improve the population's diet. A further opportunity to influence sugar consumption is to reduce sugar production through changes in agricultural subsidies (Peres et al 2017).

In October 2017, the Scottish Government announced a consultation, *A Healthier Future – Action and Ambitions on Diet, Activity and Healthy Weight*, which included: i) Proposals to restrict promotion of unhealthy food including banning unhealthy adverts before 21:00 (requiring reserved powers to be devolved to

Scotland); ii) Restricting price promotions and advertising of foods high in fat, sugar and salt near schools, visitor attractions, and on buses, trains, and transport hubs; and iii) Investment in weight management interventions for those at risk of type-two diabetes. The proposals also targeted "out of home" food providers including large and small business, whether in the public, private or voluntary sector. A targeted "out of home" provider strategy will include calories labelling, portion size, promotions, advice on healthier cooking methods and nutritional standards for public sector procurement. The strategy included funding to assist Scottish small and medium enterprises to reformulate their products with support from the Food Standards Scotland and the Food and Drink Federation Scotland. Nutritional labelling will be reviewed with the objective of improving communication of important information to consumers. In consultation with the UK government there is also an objective to extend the Soft Drinks Industry Levy to include sugary milk-based drinks and dissolvable powder drinks containing less than 95% milk (Scottish Government 2017a). The strategy seeks to reduce the current situation where 70% of children's excess weight gains occurred by age five. Focusing on children, women and families at greatest risk, the proposed actions include: i) Better integration of services with women before their first pregnancy; ii) Using the health visitor pathway and early years workforce to promote healthy eating, portion control and mealtime behaviours; and iii) Using social marketing to improve the way target groups and those leaving home for the first time shop, cook and eat; and finally iv) To train front-line staff to discuss diet behaviour. The proposals are comprehensive and bold and welcomed by Cancer Research UK and the British Medical Association Scotland (BBC 2017a).

***Exercise***. Physical activity is also addressed by the Scottish Government's *Healthier Future – Action and Ambitions on Diet, Activity and Healthy Weight* consultation. The Commonwealth Games Legacy programme targeted interventions for inactive groups such as the elderly, disabled, teenage girls and those with lower socioeconomic status. This consultation re-stated the commitment to being the first "Daily Mile" nation with a roll out to nurseries, colleges, universities and workplaces and to making Scotland's active travel

infrastructure safe for pedestrians and cyclists. The document also recognises the importance of all communities having access to active places and commits to exploring how this can be achieved (Scottish Government 2017a).

To monitor progress on Scotland becoming an active nation, the Scottish Government developed an *Active Scotland Outcomes Framework* which describes the key outcomes for sport and physical activity in Scotland over the next 10 years. Outcomes include: encouragement of the inactive to be active, enable the active to stay active, development of physical confidence early, improving Scotland's active infrastructure in terms of people *and* places, support community wellbeing and resilience through physical activity and improving opportunities to participate, progress and achieve in sport. The entire framework is underpinned and assessed in terms of equality — a baseline report was published in 2015 for this purpose (Cruickshank et al 2015).

#### 6.6.4.3 Further strategies to address health inequalities

***Advocacy***. Peres et al (2017) recently published a plan for solutions to (oral) health inequalities (Peres et al 2017). They discuss the important role that organisations such as the Royal Colleges and Dental Associations, can make to promote a healthy environment that supports healthy decisions by influencing politicians and the public. Policy statements on, for example, sugar free drinks conveyed through social media, journalists and television can change both environment and behaviour even in the face of industry lobbyists deploying "individual rights" arguments (Peres et al 2017). In the context of globalisation and neoliberal policies empowering corporations at the expense of governments and their public health objectives, advocacy is a challenging but essential strategy to address socioeconomic inequalities in oral health and health in general. Against this challenging political and economic backdrop, working across organisations is a necessary and effective tool (Peres et al 2017). To date, this approach to progress the health inequalities agenda has not been significantly used in Scotland.

***Service integration***. Barriers to integration include policy silos where different sectors work in isolation of each other both in terms of policy development as well as service delivery, reflecting the needs of the professionals delivering the service — not the community and their needs. As suggested by the findings in this thesis, socioeconomic inequalities are multifaceted and complex (Chapters 4 and 5). This finding is likely to reflect the nature of the needs of those who are experiencing low socioeconomic circumstances. As a result, services that are limited in focus are not likely to effectively address the needs of those facing low socioeconomic circumstances if provided in silos, thus integration is essential (Peres et al 2017).

***Context versus composition***. As evidenced by the results of this thesis, both area and individual characteristics are important to describing and addressing inequalities (Chapters 4 and 5). It stands to follow therefore that development of policies to address inequalities must consider both attributes (Peres et al 2017); policies focused on one or the other are likely to be ineffective given the interdependencies between "people and place" (MacIntyre et al 2002).

***Further proposed actions***. The International Agency for Research on Cancer has not updated their monograph 138 on Socioeconomic Inequalities and Cancer since publication in 1997 (Kogevinas et al 1997a). Given two decades have passed and the significant socioeconomic implications of government policy over that period, an update of the monograph is warranted. In addition, there is limited explicit attention of social epidemiology in the cancer research agenda, as an example, the Cancer Research UK funding schemes are with greater focus on areas such as biomarker research, pre-clinical research, early diagnosis, statistics and methodology, cancer biology etc. None of the research areas or the listed schemes was specifically focused on cancer from the socioeconomic perspective (Cancer Research UK 2017).

#### 6.6.4.4 General health inequalities policy

Dahlgren and Whitehead (2007) developed a framework to reflect the relative influence of the social determinants on health. At the centre resides the

individual, potentially described in terms of age, sex and constitutional factors. The next layer includes the individual behavioural factors (e.g. smoking, alcohol, diet and exercise in the context of this thesis), followed by social and community networks (social capital or cohesion), then living and working conditions are described (e.g. work environment, education, agriculture and food production, unemployment, water and sanitation, health care services and housing) and finally, general socioeconomic, cultural and environmental conditions (Dahlgren et al 2017). In the Commission on Social Determinants of Health (Solar et al 2010) the outer two Dahlgren layers are described as "structural determinants" while the inner two layers are considered as "intermediate determinants" of health inequalities.

Dahlgren et al (2007) recommended that to achieve economic growth with equity, it is essential to: i) Recognise growth as a resource for human development, especially the disadvantaged; ii) Develop efficient economic growth strategies that promote human development generally, reduce disadvantage and increase access to education and health services; iii) Conversely, avoid inefficient strategies that increase poverty and the income gap, reduce access to education and health; iv) Develop health-adjusted measures of GNP considering the total costs of poverty; and finally v) Research global factors and processes that affect health equity and constrain the power of countries to address health inequalities (Dahlgren et al 2007).

#### 6.6.4.5 Political choices

***Political context***. The Scottish Government is responsible for devolved matters (Scottish Parliament 2017), and those not explicitly reserved to the UK Parliament in Westminster, as outlined by Schedule 5 to the Scotland Act 1998. Devolved matters that were covered by the Scotland Act 1998 included healthcare provision, education, justice, policing, rural affairs, economic development and transport. Following the Scottish independence referendum in 2014, the devolved responsibility was expanded to include some elements of social security, policing of transport, the Crown Estate in Scotland, road signage and speed limits and further elements of taxation, principally

personal income tax. The Scottish Government's budget is driven by the UK block grant which is determined by the Barnett Formula with, as indicated, the ability to also increase or decrease income tax rates and cut off levels (Scottish Government 2017b). However it must be noted that the block grant provided to Scotland is driven by spending in England and thereby in part by English health policy and related spending. Relative reduction in health spending or partial health privatisation in England will likely have a consequent impact on available funding for Scotland.

In the context of addressing inequalities in cancer incidence in Scotland, in the current constitutional setting, the wider political choices therefore need to be considered at both UK Westminster and Scotland Holyrood levels. This is a challenging task in the current environment of budget erosion and austerity (Barr et al 2017). While the main levers of taxation and benefits and social security largely still lie with Westminster (Audit Scotland 2016; Scottish Government 2017b) the relatively limited devolved powers can still lead to policy innovation (such as alcohol minimum unit pricing). This is also supported by the prioritisation of health policies in the Scottish Government as health is a fully devolved power (Scottish Government 2017b). The strengths of routine administrative health, social and public health services along with data linkage potential are further enablers in developing and enhancing policy in Scotland.

Through the new social security powers devolved to Scotland under the Scotland Act 2016 (Wane et al 2016), the Scottish Government now has the power to determine the structure and value of certain benefits or replace them with new benefits in line with this legislative framework. As an example of the changes to date, the Scottish Government has recently announced plans to increase Carer’s Allowance so it is comparable to Jobseeker’s Allowance (Citizen Advice Scotland 2017) and to offer Universal Credit claimants the choice of being paid fortnightly or to have payments made directly to social landlords (Scottish Government 2017h). Although it is too early to know the full detail of these new policies, those that have been defined to date demonstrate plans to reduce or mitigate

the effect of socioeconomic stratification and to consider the wider arena of policy to begin addressing redistribution of income, wealth and power.

***Westminster policy***. In the UK, since the financial crisis in 2008, successive governments have pursued programmes of austerity which are founded on the pillars of neoliberalism, that is, budget cuts to reduce fiscal deficits, measures to balance fiscal and trade deficits, deregulation of the economy and privatisation of state enterprises (e.g. Royal Mail) (Barr et al 2017). Westminster austerity policies have had direct impacts on public services and welfare benefit cuts (Barr, 2017). Stuckler et al (2013) go a step further to demonstrate that countries, including the UK, have through harsh austerity measures and cuts to social programmes "turned recessions into veritable epidemics" (Stuckler et al 2013).

Further prescient issues that must be considered are the implications of Brexit which are difficult to fully assess given the level of uncertainty as to final agreement, but there are many who anticipate negative effects on health (some of which are already materialising such as adverse NHS staffing implications resulting from reduced net migration to the UK from the EU) and the NHS (Pfeiffer et al 2010; Iacobucci 2016; Modi 2017; Fahy et al 2017). Whether intended or not, the projections for the UK as a result of Brexit are likely to be comparable to the definition of structural adjustment programmes outlined by Breman et al (2007) which are recognised as having significant negative socioeconomic effects followed by poorer health outcomes (Pfeiffer et al 2010; Mendez-Parra et al 2016). These may compound the effects of previously pursued austerity policies.

Policy on health inequality, built on the Marmot report, *Fair Society Healthy Lives* published in 2010 (Marmot 2010), has lost some of its early momentum in terms of political buy-in with the changing UK government and over-riding Brexit challenges. Ironically, societal inequalities have been described as the key drivers of the Brexit vote in a *Joseph Rowntree Report* (Goodwin et al 2016). However, it is important to note that Scotland did not vote to leave the EU in

the 2016 referendum, but voted in fact to remain in all 32 councils with a 62% overall remain vote (BBC 2016). Interestingly, H Curtis (2016) identified that "voters in Scotland were more likely to vote remain irrespective of their social background" (Curtis 2016).

***Holyrood policy***. *Achieving Our Potential: a Framework to Tackle Poverty and Income Inequality in Scotland, 2008* (Scottish Government 2008a) identified the objective of increasing overall income and the proportion of income earned by the three lowest deciles as a group by 2017 and to increase the healthy life expectancy at birth in the most deprived areas of Scotland. The means of delivering this objective were to encourage and support cooperative working across the different sectors and to complement other policies such as *Equally Well* and the *Early Years Framework*.

The *Health Inequalities Taskforce* was implemented in an effort to raise the profile and cross-sectoral working to tackle health inequalities (Beeston et al 2014a; Beeston et al 2014b; Beeston et al 2014c). However, progressing action on key economic levers has not been without its detractors – including urging the utilisation of newly devolved income taxation powers (ScotPHO 2014; Socialist Health Association Scotland 2015). Scotland's 2017 budget, announced 14 December 2017, represents a move in this direction and to date is the most significant use of the income tax varying powers available to the Scottish Government. While the implications may be viewed as small steps relative to the task, the adoption of a five-band income tax system will mean no one earning less than £33,000 in Scotland will pay more tax than they do now (BBC 2017b) while higher earners will pay more, but marginally so. As a result, those on lower incomes are protected while additional funds are raised to invest in public services and support businesses (Mackay 2017). This may be considered a positive structural enhancement of the income tax system in Scotland. It was recently recognised as an "astute and progressive budget that has placed Scotland in the vanguard of tax reform – not just here but in the entire UK" (Macwhirter 2017).

In 2016, the independent advisor on poverty and inequality provided an assessment of Scotland's performance (Eisenstadt 2016). She reported the Scottish Government had taken a range of actions to lessen UK austerity including welfare reform mitigating, council tax reduction, providing crisis and community care grants, actively supporting social housing, funding advice services, further supporting education maintenance and promoting the Living Wage. Eisenstadt (2016) concluded that as a likely result of these policies and despite the current socioeconomic climate in the UK, relative poverty rates for children and working age adults was 14% while for pensioners it was 15%. These rates were better than for the UK as a whole as reflected in the recent Joseph Rowntree Foundation report at 30% for children and 16% for pensioners for 2015-16 (Barnard et al 2017). They were also lower than the rates for Scotland in recent prior years (Eisenstadt 2016).

Moreover, the Scottish Government has developed 55 high level *National Indicators* to monitor Scotland's progress across key government areas (Scottish Government 2017e). Fourteen of these indicators are focused on the social determinants of health defined by Solar et al (2010) as structural or intermediate or by Watt et al (2007) as upstream or downstream respectively. With exception of the three indicators that are focused on the disadvantaged populations, the other indicators are high level focusing on Scotland as a whole. These provide evidence of policies focused on reducing stratification, not just mitigation of socioeconomic hardship and address structural (upstream) determinants such as the level of education attainment, reducing the proportion of employees earning less than the Living Wage and improving the skill profile of the population (Scottish Government 2017e). Intermediary or downstream determinants are also among the National Indicators with a focus on smoking, alcohol and physical activity. Many of these indicators which are currently at "all Scotland" level could appropriately be monitored and evaluated by socioeconomic group.

A further indicator that could be added would be the reduction of health-related inequalities, specifically, socioeconomic inequalities of lung and UADT cancer

risk. Since 1997, cancer has been the primary cause of death in Scotland overtaking coronary heart disease (Scottish Government 2017d) and as such could be a tangible health inequality measure. As already presented in this thesis the lung and UADT cancer group contributes 80-90% of all cancer risk inequalities. Thus, there may be scope for the Scottish Government to review national indicators and targets and to be more focused on health inequalities as a fundamental policy outcome — across all policies.

## 6.7 Conclusions

Collectively, this thesis points to the fundamental importance – and empirical challenges — of examining the relationship between person and place socioeconomic factors in order to understand cancer incidence burden and inequalities.

It demonstrated that lung and upper aero-digestive tract cancers incidence are the most unequally (unfairly) distributed cancers in Scotland. The temporal relationship was confirmed suggesting that SES measures should be recorded as early as possible and throughout the life-course to reflect the long lead-time between exposure and incidence as well as the complex, multi-dimensional and compounded nature of socioeconomic circumstances. This finding in part refuted the reverse causation explanation for inequalities.

Age-sex specific analysis and use of the complex measure of inequalities showed that those experiencing low socioeconomic inequalities are more likely to be diagnosed with lung and upper aero-digestive tract cancers at a younger age than their more affluent counter parts; inequality was greater for men than women and occurred decades earlier than peak incidence for both sexes. Lung and UADT cancer contributed 91% to all cancer inequalities for men and 81% for women.

This thesis also confirmed that relying on either area or individual SES variables alone was inadequate; both area and individual SES variables are required as they capture different aspects of the multifaceted nature of deprivation as well

as the various and cumulative effects of SES exposure. The results of this thesis demonstrated the importance of using multiple SES measures in epidemiological studies.

The thesis studies identified that multiple measures of SES over the life-course, both area and individual, as well as, the novel development of an indicator of compounded multiple low socioeconomic circumstances contribute to explaining the complex mechanisms leading to increased cancer risk. Behavioural risk factors were also confirmed to be an important part of the socioeconomic cancer risk pathway. Behaviours, particularly smoking and alcohol, explained much of the elevated lung and upper aero-digestive tract cancer risk for individual SES indicators. Clearly, in this context, smoking was confirmed as a major inequality issue and a significant cancer risk. However, those facing multiple low socioeconomic circumstances are particularly vulnerable with elevated risk association of lung and UADT cancer, even after adjustment for behaviours. Further research is required to investigate the remaining association between multiple low socioeconomic circumstances and lung and upper aero-digestive tract cancer incidence that may be due to chronic stress, unmeasured risk factors and/or residual confounding from important aetiological factors. This information could then be used to explore specific modifiable pathways for people with low socioeconomic circumstances that may influence lung and UADT cancer risk.

Some of the thesis data linkage methodology was pioneering and involved earlier research infrastructure, which has since substantially improved building on Scotland's wealth of routine health administrative datasets (Pavis et al 2015).

This thesis provides useful insights for raising the issue of inequalities in cancer, for advocacy, and for building policy and interventions to tackle inequalities in cancer incidence. Policies need to focus more broadly on upstream causes. Traditionally, these policies have been focused on downstream behaviours (e.g. public space smoking ban and alcohol minimum pricing), but upstream policies that take on the fundamental political decisions regarding the distribution of

income, wealth and power are required at both Westminster and Holyrood and beyond.

# References


Abel, G. A., Barclay, M. E., & Payne, R. A. (2016) Adjusted indices of multiple deprivation to enable comparisons within and between constituent countries of the UK including an illustration using mortality rates. BMJ Open; 6: 11.

Adams, J., White, M., & Forman, D. (2004) Is the rate of biological aging, as measured by age at diagnosis of cancer, socioeconomically patterned? J Epidemiol Community Health.2005 Feb;59(2):146-51.

Adler N.E. & Newman K. Socioeconomic Disparities In Health: Pathways And Policies. Health Aff 21[2], 60-76. 2002.

ADLS (2017a) Scottish Health Survey Administrative Data Liaison Service. [website] http://www.adls.ac.uk/find-administrative-data/linked-administrative-data/scottish-health-survey/ [accessed 18-9-2017a].

ADLS (2017b) Scottish Longitudinal Study Administrative Data Liaison Service. [website] http://www.adls.ac.uk/find-administrative-data/linked-administrative-data/scottish-longitudinal-study/ [accessed 18-9-2017b].

Al Dakkak, I. & Khadra, M. (2011) Socio-economic status and upper aerodigestive tract cancer. Evidence-based dentistry; 12: 87-88.

Alcohol Team (30-11-2017) Minimum Unit Pricing Scottish Government Website. [website] http://www.gov.scot/Topics/Health/Services/Alcohol/minimum-pricing [accessed 6-12-2017].

Anantharaman, D., Gheit, T., Waterboer, T., Abedi-Ardekani, B., Carreira, C., McKay-Chopin, S., Gaborieau, V., Marron, M., Lagiou, P., Ahrens, W., Holcatova, I., Merletti, F., Kjaerheim, K., Talamini, R., Simonato, L., Castellsague, X., Macfarlane, T. V., Biggs, A. M., Thakker, N., Znaor, A., Thomson, P., Canova, C., Conway, D. I., Healy, C. M., Tommasino, M., Pawlita, M., & Brennan, P. (2013) Human Papillomavirus Infections and Upper Aero-Digestive Tract Cancers: The ARCAGE Study. Journal of the National Cancer Institute; 105: 536-545.

Anderson, Z. J., Lassen, C. F., & Clemmensen, I. H. (2008) Social inequality and incidence of and survival from cancers of the mouth, pharynx and larynx in a population-based study in Denmark, 1994-2003. European Journal of Cancer; 44: 1950-1961.

Angus, C., Holmes, J., Pryce, R., Meier, P., & Brennan, A. (2016). *Model-based appraisal of the comparative impact of Minimum Unit Pricing and taxation policies in Scotland*, University of Sheffield, Sheffield, Version 3.

APS Group Scotland (2017). *Questionnaire Content of the Scottish Health Survey*, Scottish Government, Edinburgh.

Arnold, M., Laversanne, M., Brown, L. M., Devesa, S. S., & Bray, F. (2017) Predicting the Future Burden of Esophageal Cancer by Histological Subtype: International Trends in Incidence up to 2030. Am J Gastroenterol. 2017 Aug; 112(8):1247-1255.

Arnold, M., Leitzmann M, Freisling, H., Bray, F., Romieu, I., Renehan A, & Soerjomataram, I. (2015) Obesity and cancer: An update of the global impact. Cancer Epidemiol; 41: 8-15.

Arnold, M., Renteria E, Conway, D., & Bray, F. (2016) Inequalities in cancer incidence and mortality across medium to highly developed countries in the twenty-first century. Cancer Causes & Control; 27: 999-1007.

Atherton, I. M., Evans, J. M., Dibben, C. J., Woods, L. M., & Hubbard, G. (2012) Differences in self-assessed health by socioeconomic group amongst people with and without a history of cancer: an analysis using representative data from Scotland. J Cancer Surviv.; 6: 458-467.

Audit Scotland (2016). *Managing new financial powers: An update,* Edinburgh.

Baastrup, R., Sørensen, M., Hansen, J., Hansen, R. D., Wurtzen, H., & Winther, J. F. (2008) Social inequality and incidence of and survival from cancers of the oesophagus, stomach and pancreas in a population-based study in Denmark, 1994-2006. European Journal of Cancer; 1962-1977.

Bailey, J. (2009) Effect of Marital Status on Cancer Incidence and Survival Rates. American Family Physician; 80: 1052-1052.

Ballard-Barbash, R., Friedenreich, C., Slattery, M., & Thune, I. (2009) Obesity and Body Composition. in *Cancer Epidemiology and Prevention*, D. Schottenfeld & J. Fraumeni, eds., Oxford University Press, New York.

Bangal R., Giri P., Bangal S., More M., & Singh K. (2014) Socio-demographic profile and associated risk factors in cancer patients attending the oncology OPD of a tertiary care teaching hospital in Western Maharashtra, India. International Journal of Medical Science and Public Health; 3: 1389-1392.

Barker, D. J. (1991) The foetal and infant origins of inequalities in health in Britain. J Public Health Med; 13: 64-68.

Barnard, H., Kumar, A., Wenham, A., Drake, B., Collingwood, A., & Leese, D. (2017). *UK Poverty 2017 Summary*, Joseph Rowntree Foundation, York.

Baron, J. H. & Sonnenberg, A. (2008) Hospital admissions and primary care attendances for nonulcer dyspepsia, reflux oesophagitis and peptic ulcer in Scotland 1981-2004. European Journal of Gastroenterology & Hepatology; 20: 180-186.

Barr, B., Higgerson, J., & Whitehead, M. (2017) Investigating the impact of the English health inequalities strategy: time trend analysis. BMJ; 358: j3310-

Batty, G. D., Russ, T. C., Stamatakis, E., & Kivimaki, M. (2017) Psychological distress in relation to site specific cancer mortality: pooling of unpublished data from 16 prospective cohort studies. BMJ; 356: j108.

Baum, A., Garofalo, J. P., & Yali, A. M. (1999) Socioeconomic status and chronic stress. Does stress account for SES effects on health? Ann N Y.Acad.Sci; 896: 131-144.

BBC (24-6-2016) EU referendum: Scotland backs Remain as UK votes Leave BBC News Online. [website] 16-12-2017].

BBC (26-10-2017a) Scottish government crackdown on unhealthy food offers. BBC News Online. [website] http://www.bbc.co.uk/news/uk-scotland-scotland-politics-41762128 [accessed 6-12-2017a].

BBC (14-12-2017b) Scottish income tax changes unveiled. BBC News Online. [website] http://www.bbc.co.uk/news/uk-scotland-42356953 [accessed 16-12-2017b].

Beelen, R., Hoek, G., van den Brandt, P. A., Goldbohm, R., Fischer, P., Schouten, L. J., Armstrong, B., & Brunekreef, B. (2008) Long-term exposure to traffic-related air pollution and lung cancer risk. Epidemiology; 19: 702-710.

Beeston, C., McCartney G., Ford, J., Wimbush, E., Beck, S., MacDonald, W., & Fraser, A. (2014a). *Erratum for 'Health Inequalities Policy Review for the Scottish Ministerial Task Force on Health Inequalities' published by NHS Health Scotland in June 2013*, NHS Health Scotland, Edinburgh.

Beeston, C., McCartney G, Ford, J., Wimbush, E., Beck, S., MacDonald, W., & Fraser, A. (2014b). *Health Inequalities Policy Review for the Scottish Ministerial Task Force on Health Inequalities' published by NHS Health Scotland in June 2013*, NHS Health Scotland, Edinburgh.

Beeston, C., McCartney G, Ford, J., Wimbush, E., Beck, S., MacDonald, W., & Fraser, A. (2014c). *Health Inequalities Policy Review for the Scottish Ministerial Task Force on Health Inequalities' published by NHS Health Scotland in June 2013 Key Messages*, NHS Health Scotland, Edinburgh.

Beetson C., Robinson M., Craig N., & Graham L. (2011). *Monitoring and evaluating Scotland's Alcohol Strategy. Setting the Scene: Theory of change and baseline picture*, NHS Health Scotland, Edinburgh.

Behrens, G., Jochem, C., Keimling, M., Ricci, C., Schmid, D., & Leitzmann, M. F. (2014) The association between physical activity and gastroesophageal cancer: systematic review and meta-analysis. Eur J Epidemiol; 29: 151-170.

Behrens, T., Gross, I. M., Siemiatycki, J., Conway, D. I., Olsson, A., Stucker, I., Guida, F., Jockel, K. H., Pohlabeln, H., Ahrens, W., Bruske, I., Wichmann, H. E., Gustavsson, P., Consonni, D., Merletti, F., Richiardi, L., Simonato, L., Fortes, C., Parent, M. E., McLaughlin, J., Demers, P., Landi, M. T., Caporaso, N., Zaridze, D., Szeszenia-Dabrowska, N., Rudnai, P., Lissowska, J., Fabianova, E., Tardon, A., Field, J. K., Dumitru, R. S., Bencko, V., Foretova, L., Janout, V., Kromhout, H., Vermeulen, R., Boffetta, P., Straif, K., Schuz, J., Hovanec, J., Kendzia, B., Pesch, B., & Bruning, T. (2016) Occupational prestige, social mobility and the association with lung cancer in men. BMC Cancer; 16: 395.

Ben Shlomo, Y. (2007) Rising to the challenges and opportunities of life course epidemiology. Int J Epidemiol; 36: 481-483.

Bennett, V. A., Davies, E. A., Jack, R. H., Mak, V., & Moller, H. (2008) Histological subtype of lung cancer in relation to socio-economic deprivation in South East England. BMC Cancer; 8: 139.

Benzeval, M. & Judge, K. (2001) Income and health: the time dimension. Social Science & Medicine; 52: 1371-1390.

Berkman, L. F. & MacIntyre, S. (1997) The measurement of social class in health studies: old measures and new formulations. In *Social Inequalities and Cancer*, M. Kogevinas et al., eds., International Agency for Research on Cancer, Lyon, 51-64.

Bilancia, M. & Fedespina, A. (2009) Geographical clustering of lung cancer in the province of Lecce, Italy: 1992-2001. International Journal of Health Geographics [Electronic Resource].8: 40.

Bilello, K. S., Murin, S., & Matthay, R. A. (2002) Epidemiology, etiology, and prevention of lung cancer. Clinics in Chest Medicine; 23: 1.

Bishop J., Clark D., Harris V., Stockton D., & Sutton M. (2004). *Deprivation and urban and rural measurements Summary repor*t. [website] http://www.isdscotlandarchive.scot.nhs.uk/isd/files//Measuring_deprivation_in_ISD_v3.pdf [accessed 01-03-2009].

Black D., Morris J., Smith C., & Townsend P. (1980). *Inequalities in Health: Report of a Research Working Group*, Department of Health and Social Services, London.

Blair, Y. I., McMahon, A. D., & Macpherson, L. M. (2013) Comparison and relative utility of inequality measurements: as applied to Scotland's child dental health. PLoS ONE; 8: 3.

Blot, W. J., Mclaughlin, J. K., & Fraumeni, J. F. (2006) Esophageal Cancer. In *Cancer Epidemiology and Prevention*, D. Schottenfeld & J. Fraumeni, eds., Oxford University Press, New York.

Bodek, D. D. & Patel, P. (2016) Sa1131 Racial and economic inequality trends in esophageal and colorectal malignancies in the US from 1992-2007. Gastroenterology 2016; 150: 4, Supplement 1, S251.

Boing, A. F., Ferreira Antunes, J. L., de Carvalho, M. B., Gois Filho, J. F., Kowalski, L. P., Michaluart, P., Jr., Eluf-Neto, J., Boffetta, P., & Wunsch-Filho, V. (2010) How much do smoking and alcohol consumption explain socioeconomic inequalities in head and neck cancer risk? J Epidemiol Community Health; 2011;65: 709-714.

Boscoe, F. P., Johnson, C. J., Sherman, R. L., Stinchcomb, D. G., Lin, G., & Henry, K. A. (2014) The relationship between area poverty rate and site-specific cancer incidence in the United States. Cancer; 120: 2191-2198.

Boyce, J. (18-4-2008) Scottish Index of Multiple Deprivation (SIMD) Scottish Government Website. [website] http://www.gov.scot/resource/doc/933/0059715.pdf [accessed 29-11-2017].

Boyle, P. J., Feijten, P., Feng, Z., Hattersley, L., Huang, Z., Nolan, J., & Raab, G. (2009) Cohort Profile: The Scottish Longitudinal Study (SLS). International Journal of Epidemiology; 38: 385-392.

Braveman, P., Egerter, S., & Williams, D. R. (2011) The Social Determinants of Health: Coming of Age. Annual Review of Public Health, Vol 32; 32: 381-398.

Breman, A. & Shelton, C. (2007) Structural adjustment programs and health. in *Globalisation and Health*, I. Kawachi & S. Wamala, eds., Oxford University Press, Oxford, pp. 219-233.

Brewster, D. H., Fraser, L. A., McKinney, P. A., & Black, R. J. (2000) Socioeconomic status and risk of adenocarcinoma of the oesophagus and cancer of the gastric cardia in Scotland. British Journal of Cancer; 83: 387-390.

Brewster, D. H., Stockton, D., Harvey, J., & Mackay, M. (2002) Reliability of cancer registration data in Scotland, 1997. Eur J Cancer; 38: 414-417.

Brown, A., Harding, S., Bethune, A., & Rosato, M. (1997) Incidence of health of the nation cancers by social class. Population Trends; 90: 40-47.

Brown, J., Demou, E., Tristram, M. A., Gilmour, H., Sanati, K. A., & Macdonald, E. B. (2012) Employment status and health: understanding the health of the economically inactive population in Scotland. BMC Public Health; 12: 327.

Brown, L., Campbell-Jack, D., Gray, L., Hovald, P., Kirkpatrick, G., Knudsen, L., Leyland, A. H., Montagu, I., & Rose, J. (2016). *Scottish Health Survey 2015 edition*, Scottish Government, Edinburgh, Volume 1 Main Report.

Brown, L. M., Hoover, R., Silverman, D., Baris, D., Hayes, R., Swanson, G. M., Schoenberg, J., Greenberg, R., Liff, J., Schwartz, A., Dosemeci, M., Pottern, L., & Fraumeni, J. (2001) Excess incidence of squamous cell esophageal cancer among US black men: Role of social class and other risk factors. American Journal of Epidemiology; 153: 114-122.

Bryden, A., Roberts, B., Petticrew, M., & Mckee, M. (2013) A systematic review of the influence of community level social factors on alcohol use. Health & Place; 21: 70-85.

Bull, F. C. & Expert Working Groups (2010). *Physical Activity Guidelines in the UK: Review and Recommendations*, School of Sport, Exercise and Health Sciences, Loughborogh University.

Cagney, K. A. (2006) Neighborhood age structure and its implications for health. Journal of Urban Health-Bulletin of the New York Academy of Medicine; 83: 827-834.

Caiazzo, A., Cardano, M., Cois, E., Costa, G., Marinacci, C., Spadea, T., Vannoni, F., & Venturini, L. (2004) Inequalities in health in Italy. Epidemiol Prev; 28(3 Suppl): i-ix, 1-161.

Cambois, E. & Jusot, F. (2011) Contribution of lifelong adverse experiences to social health inequalities: findings from a population survey in France. European Journal of Public Health; 21: 667-673.

Campos-Matos, I., Russo, G., & Perelman, J. (2016) Connecting the dots on health inequalities - a systematic review on the social determinants of health in Portugal. Int J Equity Health; 15: 26.

Cancer Research UK (1-1-2017) Our funding schemes CRUK website. [website] http://www.cancerresearchuk.org/funding-for-researchers/our-funding-schemes?f%5B0%5D=field_scheme_research_area%3A4156&f%5B1%5D=field_scheme_research_area%3A4121&f%5B2%5D=field_scheme_research_area%3A4201&f%5B3%5D=field_scheme_research_area%3A4071&items_per_page=20 [accessed 17-12-2017].

Caracta, C. F. (2003) Gender differences in pulmonary disease. Mount Sinai Journal of Medicine; 70: 215-224.

Carstairs, V. (1995) Deprivation indices: Their interpretation and use in relation to health. Journal of Epidemiology and Community Health; 49: S3-S8.

Caygill, C., Bhattacharjee, S., Charlett, A., Fox, A. J., Gatenby, P., Watson, A., Royston, C., & Bardhan, K. D. (2014a) PTU-179 Social deprivation in Barrett's oesophagus? Gut; Conference: British Society of Gastroenterology Annual General Meeting; 63(Suppl 1): 10.1136/gutjnl-2014-307263.253.

Caygill, C., Bhattacharjee, S., Charlett, A., Fox, A. J., Gatenby, P. A., Watsom, A., Royston, C., & Dev, B. K. (2014b) Sa1851 Index of social deprivation in a barrett's esophagus cohort: The influence of affluence? Gastroenterology; Conference: 146(Suppl 1): S-312.

CeLSIUS (2017) Social class (SC) UCL. [website] https://www.ucl.ac.uk/celsius/online-training/socio/se040100 [accessed 23-8-2017].

Chaturvedi, A. K., Anderson, W. F., Lortet-Tieulent, J., Curado, M. P., Ferlay, J., Franceschi, S., Rosenberg, P. S., Bray, F., & Gillison, M. L. (2013) Worldwide trends in incidence rates for oral cavity and oropharyngeal cancers. J Clin Oncol; 31: 4550-4559.

Chaturvedi, A. K., Engels, E. A., Pfeiffer, R. M., Hernandez, B. Y., Xiao, W., Kim, E., Jiang, B., Goodman, M. T., Sibug-Saber, M., Cozen, W., Liu, L., Lynch, C. F., Wentzensen, N., Jordan, R. C., Altekruse, S., Anderson, W. F., Rosenberg, P. S., & Gillison, M. L. (2011) Human papillomavirus and rising oropharyngeal cancer incidence in the United States. J Clin Oncol; 29: 4294-4301.

Citizen Advice Scotland (30-5-2017) Details and timescales for first wave of devolved Scottish benefits announced. Citizens Advice Scotland. [website] https://www.cas.org.uk/news/details-and-timescales-first-wave-devolved-scottish-benefits-announced [accessed 18-12-2017].

Citizens Advice (2017). *Just about managing. The problems that can make the difference between 'just managing' - and not*. Citizens Advice [website] https://www.citizensadvice.org.uk/Global/CitizensAdvice/Work%20Publications/JustAboutManaging-final.pdf [accessed 15-9-2017].

Clark, D., King, A., Sharpe, K., Connelly, G., Elliott, L., Macpherson, L. M. D., McMahon, A. D., Milligan, I., Wilson, P., Conway, D. I., & Wood, R. (2017) Linking routinely collected social work, education and health data to enable monitoring of the health and health care of school-aged children in state care ('looked after children') in Scotland: a national demonstration project. Public Health; 150: 101-111.

Clegg, L. X., Reichman, M. E., Miller, B. A., Hankey, B. F., Singh, G. K., Lin, Y. D., Goodman, M. T., Lynch, C. F., Schwartz, S. M., Chen, V. W., Bernstein, L., Gomez, S. L., Graff, J. J., Lin, C. C., Johnson, N. J., & Edwards, B. K. (2009) Impact of socioeconomic status on cancer incidence and stage at diagnosis: selected findings from the surveillance, epidemiology, and end results: National Longitudinal Mortality Study. Cancer Causes & Control; 20: 417-435.

Collinson, P (2017) UK leads the way with its fizzy drink tax - it should cover sweets too Guardian Online Newspaper. [website] https://www.theguardian.com/society/2017/oct/11/big-soda-small-steps-philip-hammond-must-extend-pop-tax-to-sweets [accessed 6-12-2017].

Connor, G. S., Schofield-Hurwitz, S., Hardt, J., Levasseur, G., & Tremblay, M. (2009) The accuracy of self-reported smoking: a systematic review of the relationship between self-reported and cotinine-assessed smoking status. Nicotine Tob Res; 11: 12-24.

Conway, D. I. (2010) " Everything in moderation...? ". Evid.Based Dent; 11: 89-90.

Conway, D. I., Brenner, D. R., McMahon, A. D., Macpherson, L. M., Agudo, A., Ahrens, W., Bosetti, C., Brenner, H., Castellsague, X., Chen, C., Curado, M. P., Curioni, O. A., Maso, L. D., Daudt, A. W., Gois Filho, J. F., D'Souza, G., Edefonti, V., Fabianova, E., Fernandez, L., Franceschi, S., Gillison, M., Hayes, R. B., Healy, C. M., Herrero, R., Holcatova, I., Jayaprakash, V., Kelsey, K., Kjaerheim, K., Koifman, S., La Vecchia, C., Lagiou, P., Lazarus, P., Levi, F., Lissowska, J., Luce, D., Macfarlane, T. V., Mates, D., Matos, E., McClean, M., Menezes, A. M., Menvielle, G., Merletti, F., Morgenstern, H., Moysich, K., MÃ¼ller, H., Muscat, J., Olshan, A. F., Purdue, M. P., Ramroth, H., Richiardi, L., Rudnai, P., Schantz, S., Schwartz, S. M., Shangina, O., Simonato, L., Smith, E., Stucker, I., Sturgis, E. M., Szeszenia-Dabrowska, N., Talamini, R., Thomson, P., Vaughan, T. L., Wei, Q., Winn, D. M., Wunsch-Filho, V., Yu, G. P., Zhang, Z. F., Zheng, T., Znaor, A., Boffetta, P., Chuang, S. C., Ghodrat, M., Lee, Y. C. A., Hashibe, M., & Brennan, P. (2015) Estimating and explaining the effect of education and income on head and neck cancer risk: INHANCE consortium pooled analysis of 31 case-control studies from 27 countries. Int J Cancer; 136: 1125-1139.

Conway, D. I., Brewster, D. H., McKinney, P. A., Stark, J., McMahon, A. D., & Macpherson, L. M. D. (2007) Widening socio-economic inequalities in oral cancer incidence in Scotland, 1976-2002. British Journal of Cancer; 96: 818-820.

Conway, D. I., McKinney, P. A., McMahon, A. D., Ahrens, W., Schmeisser, N., Benhamou, S., Bouchardy, C., MacFarlane, G. J., Macfarlane, T. V., Lagiou, P., Minaki, P., Bencko, V., Holcatova, I., Merletti, F., Richiardi, L., Kjaerheim, K., Agudo, A., Castellsague, X., Talamini, R., Barzan, L., Canova, C., Simonato, L., Lowry, R. J., Znaor, A., Healy, C. M., McCartan, B. E., Marron, M., Hashibe, M., & Brennan, P. (2010a) Socioeconomic factors associated with risk of upper aerodigestive tract cancer in Europe. European Journal of Cancer; 46: 588-598.

Conway, D. I., McMahon, A. D., Smith, K., Black, R., Robertson, G., Devine, J., & McKinney, P. A. (2010b) Components of socioeconomic risk associated with head and neck cancer: A population-based case-control study in Scotland. British Journal of Oral & Maxillofacial Surgery; 48: 11-17.

Conway, D. I., Petticrew, M., Marlborough, H., Bertbiller, J., Hashibe, M., & Macpherson, L. M. D. (2008) Socioeconomic inequalities and oral cancer risk: A systematic review and meta-analysis of case-control studies. International Journal of Cancer; 122: 2811-2819.

Conway, D. I., Robertson, C., Gray, H., Young, L., McDaid, L. M., Winter, A. J., Campbell, C., Pan, J., Kavanagh, K., Kean, S., Bhatia, R., Cubie, H., Clarkson, J. E., Bagg, J., Pollock, K. G., & Cuschieri, K. (2016) Human Papilloma Virus (HPV) Oral Prevalence in Scotland (HOPSCOTCH): A Feasibility Study in Dental Settings. PLoS ONE; 11: e0165847.

Cooper, S. C., Day, R., Brooks, C., Livings, C., Thomson, C. S., & Trudgill, N. J. (2009) The influence of deprivation and ethnicity on the incidence of esophageal cancer in England. Cancer Causes & Control; 20: 1459-1467.

Cooper, S. C., Day, R., Brooks, C., Livings, C., & Trudgill, N. J. (2007) Oesophageal carcinoma in the West Midlands, United Kingdom: Changing incidence and the influence of socio-economic status and ethnicity. Gastroenterology; 132: A418-A418.

Costa, G., Marinacci, C., Caiazzo, A., & Spadea, T. (2003) Individual and contextual determinants of inequalities in health: The Italian case. International Journal of Health Services; 33: 635-667.

Coupland, V. H., Allum, W., Blazeby, J. M., Mendall, M. A., Hardwick, R. H., Linklater, K. M., Moller, H., & Davies, E. A. (2012) Incidence and survival of oesophageal and gastric cancer in England between 1998 and 2007, a population-based study. BMC Cancer; 12: 11.

Craig, P. & Forbes, J. (2005) Social position and health: are old and new occupational classifications interchangeable? Journal of Biosocial Science; 37: 89-106.

Cruickshank, G., Geyer, J., & O'Connor, N. (2015). *Active Scotland Outcomes: Indicator Equality Analysis*, Scottish Government Social Research 2015, Edinburgh.

CRUK (2017) Nasopharyngeal Cancer Cancer Research UK. [website] http://www.cancerresearchuk.org/about-cancer/nasopharyngeal-cancer/types [accessed 28-8-2017].

CSDH (2008). *Closing the gap in a generation: health equity through action on the social determinants of health. Final Report of the Commission on Social Determinants of Health*. World Health Organisation, Geneva.

Cummins, S., Curtis, S., Diez-Roux, A. V., & MacIntyre, S. (2007) Understanding and representing 'place' in health research: a relational approach. Soc Sci Med; 65: 1825-1838.

Currin, L. G., Jack, R. H., Linklater, K. M., Mak, V., Moller, H., & Davies, E. A. (2009) Inequalities in the incidence of cervical cancer in South East England 2001-2005: an investigation of population risk factors. BMC Public Health; 9: 62.

Curtis, J (24-10-2016) Why did Scotland vote to remain? UK in a Changing Europe. [website] http://ukandeu.ac.uk/why-did-scotland-vote-to-remain/ [accessed 16-12-2017].

D'Souza, G., Kreimer, A. R., Viscidi, R., Pawlita, M., Fakhry, C., Koch, W. M., Westra, W. H., & Gillison, M. L. (2007) Case-control study of human papillomavirus and oropharyngeal cancer. N Engl J Med; 356: 1944-1956.

Dahlgren, G. & Whitehead, M. (2007). *European strategies for tackling social inequalities in health: Levelling up Part 2*, WHO, Europe.

Dahlgren, G and Whitehead, M (2017) Dahlgren-Whitehead rainbow Economic and Social Research Council. [website] http://www.esrc.ac.uk/about-us/50-years-of-esrc/50-achievements/the-dahlgren-whitehead-rainbow/ [accessed 6-12-2017].

Dalton, S. O., Schuz, J., Engholm, G., Johansen, C., Kru¨ ger Kjæra, S., Steding-Jessen, M., Storm, H. H., & Olsen, J. H. (2008a) Social inequality in incidence of and survival from cancer in a population-based study in Denmark, 1994-2003: Summary of findings. European Journal of Cancer; 44: 2074-2085.

Dalton, S. O., Steding-Jessen, M., Engholm, G., Schüz, J., & Olsen, J. H. (2008b) Social inequality and incidence of and survival from lung cancer in a population-based study in Denmark, 1994-2003. European Journal of Cancer; 44: 1989-1995.

Dalton, S. O., Steding-Jessen, M., Gislum, M., Frederiksen, K., Engholm, G., & Schuz, J. (2008c) Social inequality and incidence of and survival from cancer in a population-based study in Denmark, 1994-2003: Background, aims, material and methods. European Journal of Cancer; 44: 1938-1949.

Denton, E. J., Hart, D., Russell, P. A., Wright, G., & Conron, M. (2017) Lung cancer and socio-economic status: inextricably linked to place of residence. Internal Medicine Journal; 47: 563-569.

Dewi, N. U., Boshuizen, H. C., Johansson, M., Vineis, P., Kampman, E., Steffen, A., Tjonneland, A., Halkjaer, J., Overvad, K., Severi, G., Fagherazzi, G., Boutron-Ruault, M. C., Kaaks, R., Li, K., Boeing, H., Trichopoulou, A., Bamia, C., Klinaki, E., Tumino, R., Palli, D., Mattiello, A., Tagliabue, G., Peeters, P. H., Vermeulen, R., Weiderpass, E., Torhild, G., I, Huerta, J. M., Agudo, A., Sanchez, M. J., Ardanaz, E., Dorronsoro, M., Quiros, J. R., Sonestedt, E., Johansson, M., Grankvist, K., Key, T., Khaw, K. T., Wareham, N., Cross, A. J., Norat, T., Riboli, E., Fanidi, A., Muller, D., & Bueno-de-Mesquita, H. B. (2016) Anthropometry and the Risk of Lung Cancer in EPIC. Am J Epidemiol; 184: 129-139.

Dhanda, J. (2014). *The Dividing Line in Cancer Care for 2030; Exploring the effect of inequalities on four million lives*, Macmillan Cancer Support, London.

Dignam, J. J., Zhang, Q., & Kocherginsky, M. (2012) The Use and Interpretation of Competing Risks Regression Models. Clinical Cancer Research; 18: 2301-2308.

Do, D. P., Finch, B. K., Basurto-Davila, R., Bird, C., Escarce, J., & Lurie, N. (2008) Does place explain racial health disparities? Quantifying the contribution of residential context to the Black/white health gap in the United States. Social Science & Medicine; 67: 1258-1268.

Dorling, D (25-10-2017) Turning the Tide On Inequality Social Europe Website. [website] https://www.socialeurope.eu/turning-tide-inequality [accessed 26-10-2017].

dos Santos Silva, I. 1999, *Cancer Epidemiology: Principles and Methods* International Agency for Research on Cancer, Lyon.

Dubow, E. F., Boxer, P., & Huesmann, L. R. (2009) Long-term Effects of Parents' Education on Children's Educational and Occupational Success: Mediation by Family Interactions, Child Aggression, and Teenage Aspirations. Merrill.Palmer.Q (Wayne.State Univ Press); 55: 224-249.

Eberth, B., Olajide, D., Craig, P., & Ludbrook, A. (2014) Smoking-related disease risk, area deprivation and health behaviours. J Public Health; 36: 72-80.

EDRC (2017). *OECD Economic Surveys: United Kingdom*, Economic and Development Review Committee OECD, London.

Eisenstadt, N. (2016). *Independent Advisor on Poverty and Inequality: Shifting the Curve - a report for the First Minister*, Scottish Government, Edinburgh.

Ellaway, A., Benzeval, M., Green, M., Leyland, A., & MacIntyre, S. (2012) "Getting sicker quicker": Does living in a more deprived neighbourhood mean your health deteriorates faster? Health & Place; 18: 132-137.

Ellaway, A. & MacIntyre, S. (1998) Does housing tenure predict health in the UK because it exposes people to different levels of housing related hazards in the home or its surroundings? Health & Place; 4: 141-150.

Eriksson, M. & Emmelin, M. (2013) What constitutes a health-enabling neighborhood? A grounded theory situational analysis addressing the significance of social capital and gender. Soc Sci Med; 97: 112-123.

Eriksson, M., Ng, N., Weinehall, L., & Emmelin, M. (2011) The importance of gender and conceptualization for understanding the association between collective social capital and health: A multilevel analysis from northern Sweden. Social Science & Medicine; 73: 264-273.

Ervik, M., Lam, F., Ferlay, J., Mery, L., Soerjomataaram, I., & Bray, F. 2016, *Cancer Today.* International Agency for Research on Cancer, Lyon.

Evans, J., Atherton, I., Dibben, C., & Hubbard, G. (2011) The health and well-being of people with a previous diagnosis of cancer: a record-linkage study in Scotland. Journal of Public Health and Epidemiology; 3: 429-432.

Faggiano, F., Partane, T., Kogevinas, M., Boffeta, P. (1997). *Socioeconomic differences in cancer incidence and moratality.* International Agency for Research on Cancer, Lyon.

Fahy, N., Hervey, T., Greer, S., Jarman, H., Stuckler, D., Galsworthy, M., & Mckee, M. (2017) How will Brexit affect health and health services in the UK? Evaluating three possible scenarios. Lancet; 390: 2110-2118.

Farr Institute (2017) Farr Institute of Health Informatics Research Farr Institute Website. [website] http://www.farrinstitute.org/ [accessed 16-12-2017].

Fehringer, G., Brenner, D. R., Zhang, Z. F., Lee, Y. C. A., Matsuo, K., Ito, H., Lan, Q., Vineis, P., Johansson, M., Overvad, K., Riboli, E., Trichopoulou, A., Sacerdote, C., Stucker, I., Boffetta, P., Brennan, P., Christiani, D. C., Hong, Y. C., Landi, M. T., Morgenstern, H., Schwartz, A. G., Wenzlaff, A. S., Rennert, G., McLaughlin, J. R., Harris, C. C., Olivo-Marston, S., Orlow, I., Park, B. J., Zauderer, M., Barros Dios, J. M., RaviÃ±a, A. R., Siemiatycki, J., Koushik, A., Lazarus, P., FernÃ¡ndez-Somoano, A., Tardon, A., Le Marchand, L., Brenner, H., Saum, K. U., Duell, E. J., Andrew, A. S., Szeszenia-Dabrowska, N., Lissowska, J., Zaridze, D., Rudnai, P., Fabianova, E., Mates, D., Foretova, L., Janout, V., Bencko, V., Holcatova, I., Pesatori, A. C., Consonni, D., Olsson, A., Straif, K., & Hung, R. J. (2017) Alcohol and Lung Cancer Risk Among Never Smokers: A Pooled Analysis from the International Lung Cancer Consortium and the SYNERGY Study. Int J Cancer; 140: 1976-1984.

Fidler M., Bray, F., & Soerjomataram, I. (2017) The global cancer burden and human development: A review. Scandinavian Journal of Public Health; 1-10.

Fischbacher C (2017). *Alternatives to SIMD for targeting anticipatory care*, Information Services Division, NHS National Services Scotland, Edinburgh.

Food Standards Scotland (6-2-2017) Food Standards Scotland's response to Soft Drinks Industry Level Consultation Food Standards Scotland Website. [website] http://www.foodstandards.gov.scot/news-and-alerts/fss-response-to-soft-drinks-industry-levy [accessed 6-12-2017].

Forman, D., de Martel, C., Lacey, C. J., Soerjomataram, I., Lortet-Tieulent, J., Bruni, L., Vignat, J., Ferlay, J., Bray, F., Plummer, M., & Franceschi, S. (2012) Global burden of human papillomavirus and related diseases. Vaccine; 30(Suppl 5): F12-F23.

Freeman T. Free sanitary products for low-income women in Aberdeen pilot. Holyrood [website] https://www.holyrood.com/articles/news/free-sanitary-products-low-income-women-aberdeen-pilot. [accessed 11-7-2017].

Frohlich, K. L., Corin, E., & Potvin, L. (2001) A theoretical proposal for the relationship between context and disease. Sociology of Health & Illness; 23: 776-797.

Galobardes, B., Shaw, M., & Lawlor, D. (2006a) Indicators of socioeconomic position (part 2). Journal of Epidemiology and Community Health; 60: 95-101.

Galobardes, B., Lynch, J., & Davey Smith, G. (2004) Childhood Socioeconomic Circumstances and Cause-specific Mortality in Adulthood: Systematic Review and Interpretation. Epidemiologic Reviews; 26: 7-21.

Galobardes, B., Shaw, M., Lawlor, D. A., Lynch, J. W., & Smith, G. D. (2006b) Indicators of socioeconomic position (part 1). Journal of Epidemiology and Community Health; 60: 7-12.

Garcia-Gil, M., Elorza, J. M., Banque, M., Comas-Cufi, M., Blanch, J., Ramos, R., Mendez-Boo, L., Hermosilla, E., Bolibar, B., & Prieto-Alhambra, D. (2014) Linking of Primary Care Records to Census Data to Study the Association between Socioeconomic Status and Cancer Incidence in Southern Europe: A Nation-Wide Ecological Study. PLoS ONE; 9(10): e109706.

Gaudet, M. M., Olshan, A. F., Chuang, S. C., Berthiller, J., Zhang, Z. F., Lissowska, J., Zaridze, D., Winn, D. M., Wei, Q., Talamini, R., Szeszenia-Dabrowska, N., Sturgis, E. M., Schwartz, S. M., Rudnai, P., Eluf-Neto, J., Muscat, J., Morgenstern, H., Menezes, A., Matos, E., Bucur, A., Levi, F., Lazarus, P., La Vecchia, C., Koifman, S., Kelsey, K., Herrero, R., Hayes, R. B., Franceschi, S., Wunsch-Filho, V., Fernandez, L., Fabianova, E., Daudt, A. W., Dal Maso, L., Paula, C. M., Chen, C., Castellsague, X., Benhamou, S., Boffetta, P., Brennan, P., & Hashibe, M. (2010) Body mass index and risk of head and neck cancer in a pooled analysis of caseâ€“control studies in the International Head and Neck Cancer Epidemiology (INHANCE) Consortium. Int J Epidemiol; 39: 1091-1102.

Giesinger, I., Goldblatt, P., Howden-Chapman, P., Marmot, M., Kuh, D., & Brunner, E. (2014) Association of socioeconomic position with smoking and mortality: the contribution of early life circumstances in the 1946 birth cohort. Journal of Epidemiology and Community Health; 68: 275-279.

Gillan E (2010) Minimum Pricing in Scotland www.eurocare.org. [website] http://www.google.co.uk/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0CDEQFjAA&url=http%3A%2F%2Fwww.eurocare.org%2Fcontent%2Fdownload%2F8040%2F42862%2Fversion%2F1%2Ffile%2FGillan%2B-%2BMinimum%2BPricing%2Bin%2BScotland.ppt&ei=7dfjUb65MYaQ0QWE44HwDA&usg=AFQjCNE1d8qYhrqHNMOHxJCbrGnNMRec_A&sig2=pQ4oKPjGLYj4T0tzLWmVkw&bvm=bv.48705608,d.d2kbout:Tabs [accessed 1-10-2017].

Gillison, M. L. (2007) Current topics in the epidemiology of oral cavity and oropharyngeal cancers. Head and Neck-Journal for the Sciences and Specialties of the Head and Neck; 29: 779-792.

Giri, P., Singh, K., & Phalke, D. (2014) Study of socio-demographic determinants of esophageal cancer at a tertiary care teaching hospital of Western Maharashtra, India. South Asian Journal of Cancer; 3: 54-56.

Glasgow Centre for Population Health (16-8-2017) Power inequality. NHS Health Scotland. [website] http://www.healthscotland.scot/health-inequalities/fundamental-causes/power-inequality [accessed 25-10-2017].

Goodarzi, M. O. (2017) Genetics of obesity: what genetic association studies have taught us about the biology of obesity and its complications. Lancet Diabetes Endocrinol. [Epub]. https://doi.org/10.1016/ S2213-8587(17)30200-0.

Goodwin, M and Heath, O (31-8-2016) Brexit vote explained: poverty, low skills and lack of opportunities Joseph Rowntree Foundation Website. [website] https://www.jrf.org.uk/report/brexit-vote-explained-poverty-low-skills-and-lack-opportunities [accessed 16-12-2017].

Gordon, D., Graham, L., Robinson, M., & Taulbut, M. (2010). *Dimensions of diversity: Population differences and health imporviement opportunities*, NHS Health Scotland, Edinburgh.

Gorman, E., Leyland, A. H., McCartney, G., Katikireddi, S. V., Rutherford, L., Graham, L., Robinson, M., & Gray, L. (2017) Adjustment for survey non-representativeness using record-linkage: refined estimates of alcohol consumption by deprivation in Scotland. Addiction; 112: 1270-1280.

Gossage, J. A., Forshaw, M. J., Khan, A. A., Mak, V., Moller, H., & Mason, R. C. (2009) The effect of economic deprivation on oesophageal and gastric cancer in a UK cancer network. International Journal of Clinical Practice; 63: 859-864.

GPD Team (1-5-2017) Deprivation Advice for Analysts National Services Scotland. [website] http://www.isdscotland.org/Products-and-Services/GPD-Support/Deprivation/_docs/PHI-Deprivation-Guidance.pdf [accessed 24-8-2017].

Grant I (2015). *Alcohol related cancers in Scotland,* University of Edinburgh.

Grant I, Fischbacher C, & Whyte B (2007). *Obesity in Scoltand: an epidemiology briefing*, NHS National Services Scotland, Edinburgh.

Gray, L., McCartney, G., White, I., Given, L., Katikireddi, S., V, & Leyland, A. (2012) Exploring Impacts of Survey Non-Response Using Record-Linkage of Scottish Health Survey Data (2003 to 2008). Journal of Epidemiology and Community Health; 66: A28-A29.

Gray, L., McCartney, G., White, I. R., Katikireddi, S. V., Rutherford, L., Gorman, E., & Leyland, A. H. (2013) Use of record-linkage to handle non-response and improve alcohol consumption estimates in health survey data: a study protocol. BMJ Open; 2013: 1-11.

Gray, L., Batty, G., Craig, P., Stewart, C., Whyte, B., Finlayson, A., & Leyland, A. H. (2010) Cohort Profile: The Scottish Health Surveys Cohort: linkage of study participants to routinely collected records for mortality, hospital discharge, cancer and offspring birth characteristics in three nationwide studies. International Journal of Epidemiology; 39: 345-350.

Gray, L. & Leyland, A. H. (2009a) A multilevel analysis of diet and socio-economic status in Scotland: investigating the 'Glasgow effect'. Public Health Nutrition; 12: 1351-1358.

Gray, L. & Leyland, A. H. (2009b) Is the "Glasgow effect" of cigarette smoking explained by socio-economic status?: A multilevel analysis. BMC Public Health; 9: 245.

Greenland, S. & Mickey, R. M. (1989) Re: "The impact of confounder selection criteria on effect estimation. Am J Epidemiol; 130: 1066.

Gupta, B., Lalloo, R., & Johnson, N. W. (2015) Life course models for upper aero-digestive tract cancer. Int Dent J; 65: 111-119.

Hamer, M., Chida, Y., & Molloy, G. J. (2009a) Psychological distress and cancer mortality. J Psychosom.Res; 66: 255-258.

Hamer, M., Stamatakis, E., & Steptoe, A. (2009b) Dose-response relationship between physical activity and mental health: the Scottish Health Survey. British Journal of Sports Medicine; 43: 1111-1114.

Hamer, M., Stamatakis, E., & Saxton, J. M. (2009c) The impact of physical activity on all-cause mortality in men and women after a cancer diagnosis. Cancer Causes & Control; 20: 225-231.

Harenstam, A. (2009) Exploring gender, work and living conditions, and health - suggestions for contextual and comprehensive approaches. Scandinavian Journal of Work Environment & Health; 35: 127-133.

Harkness, E. F., Brewster, D. H., Kerr, K. M., Fergusson, R. J., & MacFarlane, G. J. (2002) Changing trends in incidence of lung cancer by histologic type in Scotland. Int J Cancer; 102: 179-183.

Harper, S. & Lynch, J. (2009). *Methods for measuring cancer disparities: using data relevant to Health People 2010 cancer-related objectives*, University of Michigan, Ann Arbor.

Harper, S., Lynch, J., Meersman, S. C., Breen, N., Davis, W. W., & Reichman, M. E. (2008) An overview of methods for monitoring social disparities in cancer with an example using trends in lung cancer incidence by area-socioeconomic position and race-ethnicity, 1992-2004. American Journal of Epidemiology; 167: 889-899.

Harvei, S. & Kravdal, O. (1997) The importance of marital and socioeconomic status in incidence and survival of prostate cancer - An analysis of complete Norwegian birth cohorts. Preventive Medicine; 26: 623-632.

Hashibe, M., Brennan, P., Chuang, S., Boccia, S., Castellsague, X., Chen, C., Curado, M. P., Maso, L. D., Daudt, A. W., Fabianova, E., Fernandez, L., WÃ¼nsch-Filho, V., Franceschi, S., Hayes, R. B., Herrero, R., Kelsey, K., Koifman, S., La Vecchia, C., Lazarus, P., Levi, F., Lence, J. J., Mates, D., Matos, E., Menezes, A., McClean, M. D., Muscat, J., Eluf-Neto, J., Olshan, A. F., Purdue, M., Rudnai, P., Schwartz, S. M., Smith, E., Sturgis, E. M., Szeszenia-Dabrowska, N., Talamini, R., Wei, Q., Winn, D. M., Shangina, O., Pilarska, A., Zhang, Z. F., Ferro, G., Berthiller, J., & Boffetta, P. (2009) Interaction between tobacco and alcohol use and the risk of head and neck cancer: pooled analysis in the INHANCE consortium. Cancer Epidemiology Biomarkers & Prevention; 18: 541-550.

Hashibe, M., Siwakoti, B., Wei, M., Thakur, B. K., Pun, C. B., Shrestha, B. M., Burningham, Z., Lee, Y. C., & Sapkota, A. (2011) Socioeconomic status and lung cancer risk in Nepal. Asian Pac.J Cancer Prev; 12: 1083-1088.

Hashibe, M. & Sturgis, E. M. (2013) Epidemiology of oral-cavity and oropharyngeal carcinomas: controlling a tobacco epidemic while a human papillomavirus epidemic emerges. Otolaryngol Clin North Am; 46: 507-520.

Hattersley, L., Raab, G., & Boyle P (2007). *THE SCOTTISH LONGITUDINAL STUDY Tracing rates and sample quality for the 1991 Census SLS sample*, LSCS, St Andrews, LSCS Working Paper 2.0.

Heighway D and Betticher D (2004) Lung tumours: an overview Atlas of Genetics and Cytogenetics in Oncology and Haematology. [website] http://atlasgeneticsoncology.org/Tumors/LungTumOverviewID5030.html [accessed 28-8-2017].

Helliwell J, Layard R, & Sachs J (2017). *World Happiness Report 2017*, Sustainable Development Solutions Network, New York.

Hemminki, K., Zhang, H., & Czene, K. (2003) Socioeconomic factors in cancer in Sweden. International Journal of Cancer; 105: 692-700.

Hennekens, C. H. & Buring, J. (1987). *Epidemiology in Medicine,* Lippincott Williams and Wilkins, Boston.

Hertzman, C. (1999) The biological embedding of early experience and its effects on health in adulthood. Ann N Y.Acad.Sci; 896: 85-95.

Hidayat, K., Du, X., Chen, G., Shi, M., & Shi, B. (2016) Abdominal Obesity and Lung Cancer Risk: Systematic Review and Meta-Analysis of Prospective Studies. Nutrients; 8: 810.

Hiscock, R., Bauld, L., Amos, A., & Platt, S. (2012) Smoking and socioeconomic status in England: the rise of the never smoker and the disadvantaged smoker. Journal of Public Health; 34: 390-396.

HMRC (5-12-2016) Policy paper Soft Drinks Industry Levy UK Government Website. [website] https://www.gov.uk/government/publications/soft-drinks-industry-levy/soft-drinks-industry-levy [accessed 6-12-2017].

Hystad, P., Carpiano, R. M., Demers, P. A., Johnson, K. C., & Brauer, M. (2013) Neighbourhood socioeconomic status and individual lung cancer risk: Evaluating long-term exposure measures and mediating mechanisms. Soc Sci Med; 97: 95-103.

Iacobucci, G. (2016) NHS funding at serious risk from Brexit, think tank warns. BMJ; 354: i3731.

IARC (2002). *Cancer Incidence in Five Continents Vol. VIII*, International Agency for Research on Cancer, Lyon, France, No.155.

IARC (2004). *Tobacco Smoke and Involuntary Smoking No 83,* International Agency for Research on Cancer. IARC Monographs on the Evaluation of Carcinogenic Risks to Humans World Health Organisation International Agency for Research on Cancer, Lyon.

IARC. GLOBOCAN 2008 FAST STATS.

IARC (2009a). *Cancer Incidence in Five Continents Vol. IX*, International Agency for Research on Cancer, Lyons, France, IARC Scientific Publication No. 160.

IARC. (2-11-2009b) IARC Strengthens its findings on several carcinogenic personal habits and household exposures. Press Release 196. Lyon France, World Health Organisation, International Agency for Research on Cancer.

IARC. (28-6-2012a) Agents Classified by the IARC Monographs, Volumes 1-119. Lyon, France, World Health Organisation, International Agency for Research on Cancer.

IARC (2012b). *Chemical Agents and Related Occupations; Volume 100 F; A Review of Human Carcinogens*, IARC, Lyon, France,100 F.

ICOHIRP (2015). *Social inequalities in oral health: from evidence to action*, International Centre for Oral Health Inequalities Research & Policy.

ISD (2009). Cancer Statistics. [accessed February 2009].

ISD (17-3-2010) Scottish Cancer Registry ISD. [website] http://www.isdscotland.org/isd/183.html [accessed 1-4-2010].

ISD (18-8-2015) Cancer Incidence Projections Scottish Cancer Registry. [website] http://www.isdscotland.org/Health-Topics/Cancer/Cancer-Statistics/Incidence-Projections/ [accessed 16-6-2017].

ISD (1-1-2017a) Data Quality Assurance ISD Website. [website] http://www.isdscotland.org/Products-and-Services/Data-Quality/ [accessed 16-12-2017a].

ISD (2017b) Prescribing and Medicines ISD Website. [website] http://www.isdscotland.org/Health-Topics/Prescribing-and-medicines/Publication-Schedule/ [accessed 11-12-2017b].

ISD (2017c) Scottish Cancer Registry http://www.isdscotland.org/Health-Topics/Cancer/Scottish-Cancer-Registry/. [website] http://www.isdscotland.org/Health-Topics/Cancer/Scottish-Cancer-Registry/ [accessed 31-5-2017c].

ISD Geography Analysis Support (2012). *Deprivation measures: Guidance for ISD Analysts*, ISD, Edinburgh.

Islam, M. K., Merlo, J., Kawachi, I., Lindstrom, M., & Gerdtham, U. G. (2006) Social capital and health: Does egalitarianism matter? A literature review. International Journal for Equity in Health; 5:3.

Islami, F., Kamangar, F., Nasrollahzadeh, D., Aghcheli, K., Sotoudeh, M., Abedi-Ardekani, B., Merat, S., Nasseri-Moghaddam, S., Semnani, S., Sepehr, A., Wakefield, J., Moller, H., Abnet, C. C., Dawsey, S. M., Boffetta, P., & Malekzadeh, R. (2009) Socio-economic status and oesophageal cancer: results from a population-based case-control study in a high-risk area. Int J Epidemiol; 38: 978-988.

Ito, H., Matsuo, K., Tanaka, H., Koestler, D. C., Ombao, H., Fulton, J., Shibata, A., Fujita, M., Sugiyama, H., Soda, M., Sobue, T., & Mor, V. (2011) Nonfilter and filter cigarette consumption and the incidence of lung cancer by histological type in Japan and the United States: analysis of 30-year data from population-based cancer registries. International Journal of Cancer; 128: 1918-1928.

Jackson, S. Scottish Health Survey (2003) SIMD Coding. [electronic communication 5-9-2017].

Jansson, C., Johansson, A. L. V., Nyren, O., & Lagergren, J. (2005) Socioeconomic factors and risk of esophageal adenocarcinoma: A nationwide Swedish case-control study. Cancer Epidemiology Biomarkers & Prevention; 14: 1754-1761.

Johnson, S., McDonald, J. T., & Corsten, M. J. (2008) Socioeconomic factors in head and neck cancer. Journal of Otolaryngology: Head and Neck Surgery; 37(4): 597-601,

Joint Health Surveys Unit of Social and Community Planning Research and University College London (1999) Scottish Health Survey 1995 [data collection] UK Data Service. [website] http://doi.org/10.5255/UKDA-SN-3807-1 [accessed 9-4-2014].

Joint Health Surveys Unit of Social and Community Planning Research and University College London (17-8-2016) Scottish Health Survey 1998 [computer file] Colchester, Essex: UK Data Archive [distributor]. [website] http://dx.doi.org/10.5255/UKDA-SN-4379-1 [accessed 30-9-2013].

Joint Health Surveys Unit, University College London (2016a) Scottish Health Survey 2003 [data collection] UK Data Service. [website] http://doi.org/10.5255/UKDA-SN-5318-2 [accessed 30-9-2013a].

Joint Health Surveys Unit, University College London (2016b) Scottish Health Survey 2003 [data collection] Revisions to 2003 Dataset UK Data Service. [website] http://doc.ukdataservice.ac.uk/doc/5318/mrdoc/pdf/5318revisions2003v2.pdf [accessed 30-9-2013b].

Jones, A. M. & Nicolas, A. L. (2004) Measurement and explanation of socioeconomic inequality in health with longitudinal data. Health Economics; 13: 1015-1030.

Junor, E. J., Kerr, G. R., & Brewster, D. H. (2010) Oropharyngeal cancer. Fastest increasing cancer in Scotland, especially in men. BMJ; 340:c2512.doi: 10.1136/bmj.c2512.:

Kamangar, F., Chow, W. H., Abnet, C. C., & Dawsey, S. M. (2009) Environmental Causes of Esophageal Cancer. Gastroenterology Clinics of North America; 38: 27-vii.

Kamphuis, C. B. M., Giskes, K., Kavanagh, A. M., Thornton, L. E., Thomas, L. R., van Lenthe, F. J., Mackenbach, J. P., & Turrell, G. (2008) Area variation in recreational cycling in Melbourne: a compositional or contextual effect? Journal of Epidemiology and Community Health; 62: 890-898.

Karim-Kos, H. E., de Vries, E., Soerjomataram, I., Lemmens, V., Siesling, S., & Coebergh, J. W. W. (2008) Recent trends of cancer in Europe: A combined approach of incidence, survival and mortality for 17 cancer sites since the 1990S. European Journal of Cancer; 44: 1345-1389.

Katikireddi, S. V., Whitley, E., Lewsey, J., Gray, L., & Leyland, A. H. (2017) Socioeconomic status as an effect modifier of alcohol consumption and harm: analysis of linked cohort data. Lancet Public Health; 2: e267-e276.

Kawachi, I., Adler, N. E., & Dow, W. H. (2010) Money, schooling, and health: Mechanisms and causal evidence. Ann N Y.Acad.Sci; 1186: 56-68.

Kawachi, I. & Berkman, L. F. 2000, *Social Cohesion, Social Capital, and Health* Oxford University, New York.

Kawachi, I. & Kennedy, B. P. (1997) Health and social cohesion: why care about income inequality? BMJ; 314: 1037-1040.

Kawachi, I., Subramanian, S., V, & Almeida-Filho (2017) A glossary for health inequalities. J Epidemiol Community Health; 56: 647-652.

Kawachi, I. & Kroenke, C. (2006) Socioeconomic Disparities in Cancer Incidence and Mortality. In *Cancer Epidemiology and Prevention*, D. Schottenfeld & J. Fraumeni,, ed., Oxford University Press, New York.

Kelly-Irving, M., Mabile, L., Grosclaude, P., Lang, T., & Delpierre, C. (2013) The embodiment of adverse childhood experiences and cancer development: potential biological mechanisms and pathways across the life course. Int J Public Health; 58: 1661-8564.

Kiadaliri, A. A. (2014) Gender and social disparities in esophagus cancer incidence in Iran, 2003-2009: a time trend province-level study. Asian Pac. J Cancer Prev; 15: 623-627.

Kim, C. H., Lee, Y. C., Hung, R. J., McNallan, S. R., Cote, M. L., Lim, W. Y., Chang, S. C., Kim, J. H., Ugolini, D., Chen, Y., Liloglou, T., Andrew, A. S., Onega, T., Duell, E. J., Field, J. K., Lazarus, P., Le Marchand, L., Neri, M., Vineis, P., Kiyohara, C., Hong, Y. C., Morgenstern, H., Matsuo, K., Tajima, K., Christiani, D. C., McLaughlin, J. R., Bencko, V., Holcatova, I., Boffetta, P., Brennan, P., Fabianova, E., Foretova, L., Janout, V., Lissowska, J., Mates, D., Rudnai, P., Szeszenia-Dabrowska, N., Mukeria, A., Zaridze, D., Seow, A., Schwartz, A. G., Yang, P., & Zhang, Z. F. (2014) Exposure to secondhand tobacco smoke and lung cancer by histological type: a pooled analysis of the International Lung Cancer Consortium (ILCCO). Int J Cancer; 135: 1918-1930.

Knapton S.(10-7-2017) Cancer now more common than getting married or having a first baby. Telegraph. [website] http://www.telegraph.co.uk/science/2017/07/09/cancer-now-common-getting-married-having-first-baby/ [accessed 10-7-2017].

Kogevinas, M., Pearce, N., Susser, M., & Boffetta, P. (1997a). *Social Inequalities and Cancer IARC Scientific Publications No 138* International Agency for Research on Cancer, Lyon.

Kogevinas, M. & Porta, M. (1997b) Socioeconomic differences in cancer survival: a review of the evidence. In *Social Inequalities and Cancer*, M. Kogevinas et al., eds., International Agency for Research on Cancer, Lyon.

Krieger, N., Chen, J. T., Waterman, P. D., Soobader, M. J., Subramanian, S. V., & Carson, R. (2003) Choosing area based socioeconomic measures to monitor social inequalities in low birth weight and childhood lead poisoning: The Public Health Disparities Geocoding Project (US). Journal of Epidemiology and Community Health; 57: 186-199.

Krieger, N., Williams, D. R., & Moss, N. E. (1997) Measuring social class in US public health research: Concepts, methodologies, and guidelines. Annual Review of Public Health; 18: 341-378.

Kunst, A., Menvielle, G., Ezendam, N., Strand, B., Spadea, T., Pukkala, E., van Agt, H., Schapp, M., Roskam, A., Coeburgh, J. W. W., & Mackenbach, J. P. (2008). *Eurocadet: Key determinants of the future incidence of cancer across Europe: impact of prevention; Final report Work package 3* SP23-CT-2005-006528.

Kuznetsov, L., Maier, W., Hunger, M., Meyer, M., & Mielck, A. (2011) Associations between regional socioeconomic deprivation and cancer risk: Analysis of population-based Cancer Registry data from Bavaria, Germany. Prev Med; 53: 328-330.

Lagergren, J., Andersson, G., Talback, M., Drefahl, S., Bihagen, E., Harkonen, J., Feychting, M., & Ljung, R. (2016) Marital status, education, and income in relation to the risk of esophageal and gastric cancer by histological type and site. Cancer; 122: 207-212.

Lagiou, P., Georgila, C., Minaki, P., Ahrens, W., Pohlabeln, H., Benhamou, S., Bouchardy, C., Slamova, A., Schejbalova, M., Merletti, F., Richiardi, L., Kjaerheim, K., Agudo, A., Castellsague, X., Macfarlane, T. V., MacFarlane, G. J., Talamini, R., Barzan, L., Canova, C., Simonato, L., Lowry, R., Conway, D. I., McKinney, P. A., Znaor, A., McCartan, B. E., Healy, C., Nelis, M., Metspalu, A., Marron, M., Hashibe, M., & Brennan, P. J. (2009) Alcohol-related cancers and genetic susceptibility in Europe: the ARCAGE project: study samples and data collection. Eur J Cancer Prev; 18: 76-84.

Lamont, D. W., Toal, F. M., & Crawford, M. (1997) Socioeconomic deprivation and health in Glasgow and the west of Scotland-a study of cancer incidence among male residents of hostels for the single homeless. J Epidemiol Community Health; 51(6): 668-71.

Lancaster, G. A., Green, M., & Lane, S. (2006) Linkage of survey data with district-level lung cancer registrations: a method of bias reduction in ecological studies. Journal of Epidemiology and Community Health; 60: 1093-1098.

Landy, R., Walsh, D., & Ramsay, J. (2012) Do socio-economic, behavioural and biological risk factors explain the poor health profile of the UK's sickest city? Journal of Public Health; 34: 591-598.

Laurent, O., Bard, D., Filleul, L., & Segala, C. (2007) Effect of socioeconomic status on the relationship between atmospheric pollution and mortality. Journal of Epidemiology and Community Health; 61: 665-675.

Lawder, R., Harding, O., Stockton, D., Fischbacher, C., Brewster, D. H., Chalmers, J., Finlayson, A., & Conway, D. I. (2010) Is the Scottish population living dangerously? Prevalence of multiple risk factors: the Scottish Health Survey 2003. BMC Public Health; 10: 330.

Lee, I. & Oguma, Y. (2009) Physical Activity. In *Cancer Epidemiology and Prevention*, D. Schottenfeld & J. Fraumeni, eds., Oxford University Press, New York.

Leung, J., Atherton, I., Kyle, R. G., Hubbard, G., & McLaughlin, D. (2016) Psychological distress, optimism and general health in breast cancer survivors: a data linkage study using the Scottish Health Survey. Supportive Care in Cancer; 24: 1755-1761.

Leuven, E., Plug, E., & Ronning, M. (2016) Education and cancer risk. Labour Economics; 43: 106-121.

Lewin, A., Pannier, B., Meline, J., Karusisi, N., Thomas, F., & Chaix, B. (2014) Residential neighborhood, geographic work environment, and work economic sector: associations with body fat measured by bioelectrical impedance in the RECORD Study. Annals of Epidemiology; 24: 180-186.

Leyland, A. H. (2005) Socioeconomic gradients in the prevalence of cardiovascular disease in Scotland: the roles of composition and context. J Epidemiol Community Health; 59: 799-803.

Leyland, A. H., Dundas, R., McLoone, P., & Boddy, F. A. (2007a) Cause-specific inequalities in mortality in Scotland: two decades of change. A population-based study. BMC Public Health; 7: 172

Leyland, A. H., Dundas, R., McLoone, P., & Boddy, F. A. (2007b). *Inequalities in mortality in Scotland 1981-2001*, MRC Social and Public Health Sciences Unit, Glasgow, Occasional Paper 16.

Li, X., Sundquist, J., Zoller, B., & Sundquist, K. (2015) Neighborhood deprivation and lung cancer incidence and mortality: a multilevel analysis from Sweden. J Thorac.Oncol; 10: 256-263.

Littman, A. J. & Vaughan, T. L. (2006) Cancers of the Nasal Cavity and Paranasal Sinuses. in *Cancer Epidemiology and Prevention*, D. Schottenfeld & J. Fraumeni, ed., Oxford University Press, New York.

Loring, B. & Robertson, A. (2014). *Obesity and inequalities*, WHO, Copenhagen.

Lortet-Tieulent, J., Soerjomataram, I., Ferlay, J., Rutherford, M., Weiderpass, E., & Bray, F. (2014) International trends in lung cancer incidence by histological subtype: Adenocarcinoma stabilizing in men but still increasing in women. Lung Cancer; 84: 13-22.

Louie, K. S., Mehanna, H., & Sasieni, P. (2015) Trends in head and neck cancers in England from 1995 to 2011 and projections up to 2025. Oral Oncol; 51: 341-348.

Ma, H., Zhou, Z., Wei, S., Liu, Z., Pooley, K., Dunning, A., Svenson, U., Roos, G., Hosgood, H., Shen, M., & Wei, Q. (2011) Shortened telomere length is associated with increased risk of cancer: a meta-analysis. PLoS ONE; 6: e20466.

MacIntyre S & Ellaway, A. (2003) Ecological approaches: rediscovering the role of the physical and social environment. In *Social epidemiology*, L. F. Berkman & I. Kawachi, eds., Oxford University Press, New York, 332-348.

MacIntyre, S., Ellaway, A., & Cummins, S. (2002) Place effects on health: how can we conceptualise, operationalise and measure them? Social Science & Medicine; 55: 125-139.

MacIntyre, S., Ellaway, A., Der, G., Ford, G., & Hunt, K. (1998) Do housing tenure and car access predict health because they are simply markers of income or self esteem? A Scottish study. Journal of Epidemiology and Community Health; 52: 657-664.

MacIntyre, S., Hiscock, R., Kearns, A., & Ellaway, A. (2001) Housing tenure and car access: further exploration of the nature of their relations with health in a UK setting. Journal of Epidemiology and Community Health; 55: 330-331.

Mackay, D (14-12-2017) Scottish budget statement by Finance Secretary Derek Mackay SNP Website. [website] https://www.snp.org/scottish_budget_statement_by_finance_secretary_derek_mackay [accessed 16-12-2017].

Mackay, J. (2012) The global epidemiology of tobacco and related chronic diseases. Public Health; 126: 199-201.

Mackay, J. & Amos, A. (2003) Invited review series: Tobacco and lung health - Women and tobacco. Respirology; 8: 123-130.

Mackenbach, J. P. (2006). *Health Inequalities: Europe in Profile*, University Medical Centre Rotterdam, The Netherlands.

Mackenbach, J. P. (2012) The persistence of health inequalities in modern welfare states: the explanation of a paradox. Soc Sci Med; 75: 761-769.

Mackenbach, J. P. & Kunst, A. E. (1997) Measuring the magnitude of socio-economic inequalities in health: An overview of available measures illustrated with two examples from Europe. Social Science & Medicine; 44: 757-771.

Mackenbach, J. P., Stirbu, I., Roskam, A. J. R., Schaap, M. M., Menvielle, G., Leinsalu, M., Kunst, A. E., & European Union Working Group on Socioeconomic Inequalities in Health (2008) Socioeconomic inequalities in health in 22 European countries. New England Journal of Medicine; 358: 2468-2481.

Macmillan Cancer Support (10-7-2017) Cancer diagnosis now as common a 'life milestone' as marriage or getting a degree. We are Macmillan Cancer Support. [website] https://www.macmillan.org.uk/aboutus/news/latest_news/cancer-diagnosis-now-as-common-a-life-milestone-as-marriage-or-getting-a-degree.aspx [accessed 10-7-2017].

Macwhirter, I (17-12-2017) Ian Macwhirter: The budget was a step in the right direction, but now we need a wealth tax. The Herald Online News. [website] http://www.heraldscotland.com/opinion/columnists/15777859.Iain_Macwhirter__The_budget_was_a_step_in_the_right_direction__but_now_we_need_a_wealth_tax/ [accessed 17-12-2017].

Mammas, I. N., Sourvinos, G., Zaravinos, A., & Spandidos, D. A. (2011) Vaccination against Human Papilloma Virus (HPV): Epidemiological Evidence of HPV in Non-genital Cancers. Pathology & Oncology Research; 17: 103-119.

Marmot, M. (2005) Social determinants of health inequalities. Lancet; 365: 1099-1104.

Marmot, M. (2010). *Fair Society, Healthy Lives; The Marmot Review*, The Marmot Review, London.

Marmot, M. (2017) Social justice, epidemiology and health inequalities. European Journal of Epidemiology 32: 537-546.

Marmot, M., Allen, J., Bell, R., Bloomer, E., & Goldblatt, P. (2012) WHO European review of social determinants of health and the health divide. Lancet; 380: 1011-1029.

Marmot, M. G., Smith, G. D., Stansfeld, S., Patel, C., North, F., Head, J., White, I., Brunner, E., & Feeney, A. (1991) Health inequalities among British civil servants: the Whitehall II study. Lancet; 337: 1387-1393.

Marshall, B., Chevalier, A., Garillon, C., Goldberg, M., & Coing, F. (1999) Socioeconomic status, social mobility and cancer occurrence during working life: a case-control study among French electricity and gas workers. Cancer Causes & Control; 10: 495-502.

Marshall, J. R. & Freudenheim, J. O. (2009) Alcohol. in *Cancer Epidemiology and Prevention*, D. Schottenfeld & J. Fraumeni, eds., Oxford University Press, New York.

Maxwell, S. (1999). *The Meaning and Measurement of Poverty*, Overseas Development Institute, London.

Mayne, S. T., Morse, D. E., & Winn, D. M. (2006) Cancers of the Oral Cavity and Pharynx. in *Cancer Epidemiology and Prevention*, D. Schottenfeld & J. Fraumeni, eds., Oxford University Press, New York.

McCart Reed, A. E., Kutasovic, J. R., Lakhani, S. R., & Simpson, P. T. (2015) Invasive lobular carcinoma of the breast: morphology, biomarkers and 'omics. Breast Cancer Res; 17: 12.

McLaren GL & Bain MRS (1998). *Deprivation and health in Scotland: insights from NHS data*. ISD, Edinburgh.

McLoone, P. (2000). *Carstairs Scores for Scottish Postcode Sectors from the 1991 Census.* Public Health Research Unit, Glasgow University.

McMahon, A. D., Elliott, L., Macpherson, L. M., Sharpe, K. H., Connelly, G., Milligan, I., Wilson, P., Clark, D., King, A., Wood, R., & Conway, D. I. (2017) Inequalities in the dental health needs and access to dental services among looked after children in Scotland: a population data linkage study. Arch.Dis.Child; 103: 39-43.

Mehanna H & Jones T.M, G. V. A. K. (2010) Oropharyngeal carcinoma related to human papillomavirus. British Medical Journal; 340: c1439.

Mehanna, H., Beech, T., Nicholson, T., El Hariry, I., McConkey, C., Paleri, V., & Roberts, S. (2013) Prevalence of human papillomavirus in oropharyngeal and nonoropharyngeal head and neck cancer--systematic review and meta-analysis of trends by time and region. Head and Neck; 35: 747-755.

Meijer, M., Bloomfield, K., & Engholm, G. (2013) Neighbourhoods matter too: the association between neighbourhood socioeconomic position, population density and breast, prostate and lung cancer incidence in Denmark between 2004 and 2008. Journal of Epidemiology and Community Health; 67: 6-13.

Melchior, M., Goldberg, M., Krieger, N., Kawachi, I., Menvielle, G., Zins, M., & Berkman, L. F. (2005) Occupational class, occupational mobility and cancer incidence among middle-aged men and women: a prospective study of the French GAZEL cohort. Cancer Causes & Control; 16: 515-524.

Mendez-Parra, M., Papadavid, P., & te Velde, D. W. (2016). *Brexit and development. How will developing countries be affected?*, Overseas Development Institute, London.

Menvielle, G., Boshuizen, H., Kunst, A. E., Dalton, S. O., Vineis, P., Bergmann, M. M., Hermann, S., Ferrari, P., Raaschou-Nielsen, O., Tjonneland, A., Kaaks, R., Linseisen, J., Kosti, M., Trichopoulou, A., Dilis, V., Palli, D., Krogh, V., Panico, S., Tumino, R., Buchner, F. L., van Gils, C. H., Peeters, P. H. M., Braaten, T., Gram, I. T., Lund, E., Rodriguez, L., Agudo, A., Sanchez, M. J., Tormo, M. J., Ardanaz, E., Manjer, J., Wirfalt, E., Hallmans, G., Rasmuson, T., Bingham, S., Khaw, K. T., Allen, N., Key, T., Boffetta, P., Duell, E. J., Slimani, N., Gallo, V., Riboli, E., & Bueno-de-Mesquita, H. B. (2009) The Role of Smoking and Diet in Explaining Educational Inequalities in Lung Cancer Incidence. Journal of the National Cancer Institute; 101: 321-330.

Menvielle, G., Boshuizen, H., Kunst, A. E., Vineis, P., Dalton, S. O., Bergmann, M. M., Hermann, S., Veglia, F., Ferrari, P., Overvad, K., Raaschou-Nielsen, O., Tjonneland, A., Kaaks, R., Linseisen, J., Palli, D., Krogh, V., Tumino, R., Rodriguez, L., Agudo, A., Sanchez, M. J., Arozena, J. M. A., Cirera, L., Ardanaz, E., Bingham, S., Khaw, K. T., Boffetta, P., Duell, E., Slimani, N., Gallo, V., Riboli, E., & Bueno-de-Mesquita, H. B. (2010a) Occupational exposures contribute to educational inequalities in lung cancer incidence among men: evidence from the EPIC prospective cohort study. International Journal of Cancer; 126: 1928-1935.

Menvielle, G., Luce, D., Goldberg, P., & Leclerc, A. (2004) Smoking, alcohol drinking, occupational exposures and social inequalities in hypopharyngeal and laryngeal cancer. International Journal of Epidemiology; 33: 799-806.

Menvielle, G., Soerjomataram, I., de Vries, E., Engholm, G., Barendregt, J. J., Coebergh, J. W., & Kunst, A. E. (2010b) Scenarios of future lung cancer incidence by educational level: Modelling study in Denmark. European Journal of Cancer; 46: 2625-2632.

Minkler, M., Fuller-Thomson, E., & Guralnik, J. M. (2006) Gradient of disability across the socioeconomic spectrum in the United States. New England Journal of Medicine; 355: 695-703.

Mitra, D., Shaw, A., Tjepkema, M., & Peters, P. (2015) Social determinants of lung cancer incidence in Canada: A 13-year prospective study. Health Rep; 26: 12-20.

Modi, N. (2017) Brexit is bad for health, and doctors should say so. BMJ; 357: j2265.

Møller, H. & Tønnesenm, H. (1997) Alcohol drinking, social class and cancer. In *Social Inequalities and Cancer*, M. Kogevinas et al., eds., IARC, Lyon.

Morgan, M. A., Lewis, W. G., Chan, D. S., Burrows, S., Stephens, M. R., Roberts, S. A., Havard, T. J., Clark, G. W., & Crosby, T. D. (2007) Influence of socio-economic deprivation on outcomes for patients diagnosed with oesophageal cancer. Scand J Gastroenterol; 42: 1230-1237.

Morris, C. Scottish Health Survey Linkage Metrics. eDRIS. [electronic communication 27-9-2017].

Mouw, T., Koster, A., Wright, M. E., Blank, M. M., Moore, S. C., Hollenbeck, A., & Schatzkin, A. (2008) Education and Risk of Cancer in a Large Cohort of Men and Women in the United States. PLoS ONE; 3: e3639.

Mueller, N., Birmann, B. , Parsonnet, J. , Schiffman, M. , & Stuver, S. (2009) Infectious Agents. In *Cancer Epidemiology and Prevention*, D. Schottenfeld & J. Fraumeni, eds., Oxford University Press, New York.

National Cancer Intelligence Network (2013). *e_Atlas*. [website] http://www.ncin.org.uk/cancer_infomration_tools/eatlas/ [accessed 23-1-2013].

National Records Scotland (13-8-2014a) Census 2011: Bulletin Figures and Tables National Records Scotland. [website] http://www.scotlandscensus.gov.uk/bulletin-figures-and-tables [accessed 6-10-2017a].

National Records Scotland (13-8-2014b) Census 2011: Households with access to one or more vehicles 2001, 2011 National Records Scotland. [website] http://www.scotlandscensus.gov.uk/documents/censusresults/release2a/rel2asbfigure22.pdf [accessed 6-10-2017b].

National Records Scotland (13-8-2014c) Census 2011: Release 3H - Detailed characteristics on Labour Market and Transport in Scotland National Records Scotland. [website] http://www.scotlandscensus.gov.uk/news/census-2011-release-3h-detailed-characteristics-labour-market-and-transport-scotland [accessed 6-10-2017c].

National Records Scotland (2015a) Labour Market Census 2011 National Records Scotland. [website] http://www.scotlandscensus.gov.uk/labour-market [accessed 6-10-2017a].

National Records Scotland (2015b) Population estimates 2011 National Records Scotland. [website] https://www.nrscotland.gov.uk/statistics-and-data/statistics/statistics-by-theme/population/population-estimates/2011-based-special-area-population-estimates/population-estimates-by-simd-2016 [accessed 6-10-2017b].

National Records Scotland (2015c) Scotland's 2011 Census: Education National Records Scotland. [website] http://www.scotlandscensus.gov.uk/education-0 [accessed 6-10-2017c].

National Records Scotland (2015d). *Scotland's Census 2011 General Report*, National Records Scotland, Edinburgh.

National Records Scotland (2015e) Scotland's Census: Population & Households Scotland Census 2011. [website] http://www.scotlandscensus.gov.uk/population-households [accessed 6-10-2017e].

National Records Scotland (2017) Scotland's Census Shaping Our Future; Standard Outputs National Records Scotland. [website] http://www.scotlandscensus.gov.uk/ods-web/standard-outputs.html [accessed 12-9-2017].

Netuveli, G., Sheiham, A., & Watt, R. G. (2006) Does the 'inverse screening law' apply to oral cancer screening and regular dental check-ups? J Med Screen; 13: 47-50.

NHS Health Scotland (2014). *Proportionate universalism and health inequalities*, NHS Health Scotland.

Nicolotti, N., Chuang, S. C., Cadoni, G., Arzani, D., Petrelli, L., Bosetti C, Brenner, H., Hosono, S., La Vecchia, C., Matsuo, K., & Mueller, H. (2011) Recreational physical activity and risk of head and neck cancer: a pooled analysis within the international head and neck cancer epidemiology (INHANCE) Consortium. European Journal of Epidemiology; 26: 619-628.

Nkosi, T. M., Parent, M. E., Siemiatycki, J., & Rousseau, M. C. (2012) Socioeconomic position and lung cancer risk: how important is the modelling of smoking? Epidemiology; 23: 377-385.

Nordahl, H. (2014) Social inequality in chronic disease outcomes. Dan.Med J; 61: B4943.

Nordahl, H., Hvidtfeldt, U. A., Diderichsen, F., Rod, N. H., Osler, M., Frederiksen, B. L., Prescott, E., Tjonneland, A., Lange, T., Keiding, N., Andersen, P. K., & Andersen, I. (2014) Cohort profile: the Social Inequality in Cancer (SIC) cohort study. Int J Epidemiol; 43: 1750-1758.

OECD (2016). *Strategic Orientations of the Secretary-General: For 2016 and beyond; Meeting of the OECD Council at Ministerial Level*, OECD, Paris.

Office for National Statistics (2016). *Labour Force Survey: User Guide Volume 1 - LFS Background and methodology 2016*, Office for National Statistics, London.

Office of National Statistics (2017) Definitions of International Labour Organisation (ILO) measures Labour Market Trends. [website] http://ww2.prospects.ac.uk/cms/ShowPage/Home_page/Main_Menu___News_and_information/Graduate_Market_Trends/Definitions_of_International_Labour_Organisation_measures/p!edXbLa [accessed 2-9-2017].

Oliveira, T. M., Elias, J., Jr., Melo, A. F., Teixeira, S. R., Filho, S. C., Goncalves, L. M., Faria, F. M., Tiezzi, D. G., Andrade, J. M., & Muglia, V. (2014) Evolving concepts in breast lobular neoplasia and invasive lobular carcinoma, and their impact on imaging methods. Insights Imaging; 5: 183-194.

Ottersen, O. P., Dasgupta, J., Blouin, C., Buss, P., Chongsuvivatwong, V., Frenk, J., Fukuda-Parr, S., Gawanas, B. P., Giacaman, R., Gyapong, J., Leaning, J., Marmot, M., McNeill, D., Mongella, G. I., Moyo, N., Mogedal, S., Ntsaluba, A., Ooms, G., Bjertness, E., Lie, A. L., Moon, S., Roalkvam, S., Sandberg, K. I., & Scheel, I. B. (2014) The political origins of health inequity: prospects for change. Lancet; 383: 630-667.

Parkin, D., Boyd, L., & Walker, L. (2011) The fraction of cancer attributable to lifestyle and environmental factors in the UK in 2010 Summary and conclusions. British Journal of Cancer; 105: S77-S81.

Parkin, D. M. (2011) Cancers attributable to inadequate physical exercise in the UK in 2010. Br J Cancer; 105 Suppl 2: S38-S41.

Pavis, S. & Morris, A. M. (2015) Unleashing the power of administrative health data: the Scottish model. Public Health Research & Practice; 25: 1-6.

Peng DF and El-Rifai W (2009) Esophagus: Barrett's esophagus, dysplasi and adenocarcinoma. Atlas of Genetics and Cytogenetics in Oncology and Haematology. [website] http://atlasgeneticsoncology.org/Tumors/BarrettsEsophagID5591.html [accessed 28-8-2017].

Peres, M. & Watt, R. G. (2017). *Policy solutions for oral health inequalities*, ICOHIRP, Adelaide, Australia.

Pesch, B., Kendzia, B., Gustavsson, P., Jockel, K. H., Johnen, G., Pohlabeln, H., Olsson, A., Ahrens, W., Gross, I. M., Bruske, I., Wichmann, H. E., Merletti, F., Richiardi, L., Simonato, L., Fortes, C., Siemiatycki, J., Parent, M. E., Consonni, D., Landi, M. T., Caporaso, N., Zaridze, D., Cassidy, A., Szeszenia-Dabrowska, N., Rudnai, P., Lissowska, J., Stucker, I., Fabianova, E., Dumitru, R. S., Bencko, V., Foretova, L., Janout, V., Rudin, C. M., Brennan, P., Boffetta, P., Straif, K., & Bruning, T. (2012) Cigarette smoking and lung cancer-relative risk estimates for the major histological types from a pooled analysis of case-control studies. Int J Cancer; 131: 1210-1219.

Peto, J. (2012) That the effects of smoking should be measured in pack-years: misconceptions. British Journal of Cancer; 107: 406-407.

Petrie A & Sabourin, J. C. 2009, *Medical Statistics at a Glance*, 3rd edn, Blackwell Publishing, Oxford.

Pevalin, D. J., Taylor, M. P., & Todd, J. (2008) The dynamics of unhealthy housing in the UK: A panel data analysis. Housing Studies; 23: 679-695.

Pfeiffer, J. & Chapman, R. (2010) Anthropological Perspectives on Structural Adjustment and Public Health. Annu.Rev Anthropol.149-165.

Pickett, K. & Pearl, M. (2001) Multilevel analyses of neighbourhood socioeconomic context and health outcomes: a critical review. Journal of Epidemiology and Community Health; 55: 111-122.

Piketty, T. (2013). *Capital in the Twenty-First Century,* Éditions du Seuil, Paris.

Pollack, C. E., Chideya, S., Cubbin, C., Williams, B., Dekker, M., & Braveman, P. (2007) Should health studies measure wealth? A systematic review. American Journal of Preventive Medicine; 33: 250-264.

Potter, J. (1997) Diet and cancer: possible explanations for the higher risk of cancer in the poor. In *Social Inequalities and Cancer*, M. Kogevinas et al., eds., International Agency for Research on Cancer, Lyon, pp. 265-283.

Prabhu, A., Obi, K. , & Rubenstein, J. (2014) The Synergistic Effects of Alcohol and Tobacco Consumption on the Risk of Esophageal Squamous Cell Carcinoma: A Meta-Analysis. American Journal of Gastroenterology; 109: 822-827.

Prescott, J., Wentzensen, I. , Savage, S. , & De Vivo, I. (2012) Epidemiologic evidence for a role of telomere dysfunction in cancer etiology. Mutation Research-Fundamental and Molecular Mechanisms of Mutagenesis; 730: 75-84.

Puigpinos, R., Borrell, C., Antunes, J., Azlor, E., Pasarin, M., Serral, G., Pons-Vigues, M., Rodriguez-Sanz, M., & Fernandez, E. (2009) Trends in socioeconomic inequalities in cancer mortality in Barcelona: 1992-2003. BMC Public Health; 9: 35.

Pukkala, E., Martinsen, J. I., Lynge, E., Gunnarsdottir, H. K., Sparen, P., Tryggvadottir, L., Weiderpass, E., & Kjaerheim, K. (2009) Occupation and cancer - follow-up of 15 million people in five Nordic countries. Acta Oncologica; 48: 646-790.

Purkayastha, M., McMahon, A. D., Gibson, J., & Conway, D. I. (2016) Trends of oral cavity, oropharyngeal and laryngeal cancer incidence in Scotland (1975-2012) - A socioeconomic perspective. Oral Oncol; 61: 70-75.

Pye S, King, K., Sturn, R. (2006). *Air Quality and Social Deprivation in the UK: an environmental inequalities analysis-final report to DEFRA,* Contract RMP/2035. AEAT/ENV/2170.

Raaschou-Nielsen, O., Andersen, Z. J., Beelen, R., Samoli, E., Stafoggia, M., Weinmayr, G., Hoffmann, B., Fischer, P., Nieuwenhuijsen, M. J., Brunekreef, B., Xun, W. W., Katsouyanni, K., Dimakopoulou, K., Sommar, J., Forsberg, B., Modig, L., Oudin, A., Oftedal, B., Schwarze, P. E., Nafstad, P., De Faire, U., Pedersen, N. L., Ostenson, C. G., Fratiglioni, L., Penell, J., Korek, M., Pershagen, G., Eriksen, K. T., Sorensen, M., Tjonneland, A., Ellermann, T., Eeftens, M., Peeters, P. H., Meliefste, K., Wang, M., Bueno-de-Mesquita, B., Key, T. J., de Hoogh, K., Concin, H., Nagel, G., Vilier, A., Grioni, S., Krogh, V., Tsai, M. Y., Ricceri, F., Sacerdote, C., Galassi, C., Migliore, E., Ranzi, A., Cesaroni, G., Badaloni, C., Forastiere, F., Tamayo, I., Amiano, P., Dorronsoro, M., Trichopoulou, A., Bamia, C., Vineis, P., & Hoek, G. (2013) Air pollution and lung cancer incidence in 17 European cohorts: prospective analyses from the European Study of Cohorts for Air Pollution Effects (ESCAPE). The lancet oncology; 14: 813-822.

Randi, G., Altieri, A., Gallus, S., Chatenoud, L., Montella, M., Franceschi, S., Negri, E., Talamini, R., & La Vecchia, C. (2004) Marital status and cancer risk in Italy. Preventive Medicine; 38: 523-528.

Reeves R. [16-7-2017] How the middle class hoards wealth and opportunity. The Guardian. [website] https://www.theguardian.com/inequality/2017/jul/15/how-us-middle-classes-hoard-opportunity-privilege [accessed 16-7-2017].

Riva, M., Gauvin, L., & Barnett, T. A. (2007) Toward the next generation of research into small area effects on health: a synthesis of multilevel investigations published since July 1998. Journal of Epidemiology and Community Health; 61: 853-861.

Robertson, T., Batty, G., Der, G., Green, M. J., McGlynn, L. M., McIntyre, A., Shiels, P. G., & Benzeval, M. (2012) Is Telomere Length Socially Patterned? Evidence from the West of Scotland Twenty-07 Study. PLoS ONE; 7: e41805.

Rose, D. (1995). *Social Research Update*, University of Surrey, Guildford, England, Issue 9.

Rousseau A and Badoual C (2011) Head and Neck: Squamous cell carcinoma: an overview Atlas of Genetics and Cytogenetics in Oncology and Haematology. [website] http://atlasgeneticsoncology.org/Tumors/HeadNeckSCCID5078.html [accessed 28-8-2017].

Rustgi, A. K. & El Serag, H. B. (2014) Esophageal carcinoma. N Engl J Med; 371: 2499-2509.

Rutegard, M., Nordenstedt, H., Lu, Y., Lagergren, J., & Lagergren, P. (2010) Sex-specific exposure prevalence of established risk factors for oesophageal adenocarcinoma. British Journal of Cancer; 103: 735-740.

Ryerson, A. B., Peters, E. S., Coughlin, S. S., Chen, V. W., Gillison, M. L., Reichman, M. E., Wu, X. C., Chaturvedi, A. K., & Kawaoka, K. (2008) Burden of Potentially Human Papillomavirus-associated Cancers of the Oropharynx and Oral Cavity in the US, 1998-2003. Cancer; 113: 2901-2909.

Sanjosé, S. d., Bosch, F., Muñoz, & Shah, K. (1997) Social differences in sexual behaviour and cervical cancer. In *Social Inequalities and Cancer*, M. Kogevinas et al., eds., International Agency for Research on Cancer, Lyon, 309-318.

Sartorius, B., Sartorius, K., Aldous, C., Madiba, T. E., Stefan, C., & Noakes, T. (2016) Carbohydrate intake, obesity, metabolic syndrome and cancer risk? A two-part systematic review and meta-analysis protocol to estimate attributability. BMJ Open; 6: e009301.

Schmeisser, N., Conway, D. I., McKinney, P. A., McMahon, A. D., Pohlabeln, H., Marron, M., Benhamou, S., Bouchardy, C., MacFarlane, G. J., Macfarlane, T. V., Lagiou, P., Lagiou, A., Bencko, V., Holcatova, I., Merletti, F., Richiardi, L., Kjaerheim, K., Agudo, A., Talamini, R., Polesel, J., Canova, C., Simonato, L., Lowry, R., Znaor, A., Healy, C., McCarten, B. E., Hashibe, M., Brennan, P., & Ahrens, W. (2010) Life course social mobility and risk of upper aerodigestive tract cancer in men. Eur J Epidemiol; 25: 173-182.

Schottenfeld, D. and Fraumeni, J. (2009) Cancer Epidemiology and Prevention. Oxford Scholarship Online. [website] http://www.oxfordscholarship.com.ezproxy.lib.gla.ac.uk/view/10.1093/acprof:oso/9780195149616.001.0001/acprof-9780195149616 [accessed 25-8-2017].

Schwingshackl, L. & Hoffmann, G. (2015) Adherence to Mediterranean diet and risk of cancer: an updated systematic review and meta-analysis of observational studies. Cancer Med; 4: 1933-1947.

ScotCen for Social Research, University College London. Department of Epidemiology and Public Health (2016) Scottish Health Survey 2010 [data collection] UK Data Service. [website] http://doi.org/10.5255/UKDA-SN-6987-3 [accessed 30-9-2013].

ScotCen for Social Research, University College London. Department of Epidemiology and Public Health and University of Glasgow. MRC/CSO Social and Public Health Science Unit (2016) Scottish Health Survey 2011 [data collection] UK Data Service. [website] http://doi.org/10.5255/UKDA-SN-7247-4 [accessed 1-7-2013].

ScotPHO (2008). *Tobacco smoking in Scotland: an epidemiological briefing*. NHS Health Scotland, Edinburgh.

ScotPHO (5-9-2014) Informing Investment to reduce health Inequalities (III) ScotPHO website. [website] http://www.scotpho.org.uk/comparative-health/health-inequalities-tools/informing-investment-to-reduce-health-inequalities-iii [accessed 16-12-2017].

ScotPHO (12-9-2017a) Diet and nutrition: adults Scottish Health Survey 2015. [website] http://www.scotpho.org.uk/behaviour/diet-and-nutrition/data/adults/ [accessed 25-10-2017a].

ScotPHO (12-9-2017b) Physical activity: adults Scottish Health Survey 2015. [website] http://www.scotpho.org.uk/behaviour/physical-activity/data/adults/ [accessed 25-10-2017b].

ScotPHO (12-9-2017c) Tobacco use: adult smoking in Scotland Scottish Health Survey 2015. [website] http://www.scotpho.org.uk/behaviour/tobacco-use/data/adult-smoking-in-scotland/ [accessed 25-10-2017c].

Scottish Cancer Prevention Network (1-4-2012) Lung Cancer Screening Trial Scottish Cancer Prevention Network. [website] https://www.cancerpreventionscotland.org.uk/newsletter/lung-cancer-screening-trial/ [accessed 25-10-2017].

Scottish Centre for Social Research, University College London. Department of Epidemiology and Public Health (2016a) Scottish Health Survey 2008 [data collection] UK Data Service. [website] http://doi.org/10.5255/UKDA-SN-6383-3 [accessed 30-9-2013a].

Scottish Centre for Social Research, University College London. Department of Epidemiology and Public Health (2016b) Scottish Health Survey 2009 [data collection] UK Data Service. [website] http://doi.org/10.5255/UKDA-SN-6713-3 [accessed 30-9-2013b].

Scottish Government (2003). *Improving Health in Scotland*, Astron, Edinburgh.

Scottish Government (2007). *Better Health, Better Care: An Action Plan*, Scottish Government, Edinburgh.

Scottish Government (2008a). *Achieving our potential: A Framework to tackle poverty and income inequality in Scotland*, RR Donnelly, Edinburgh.

Scottish Government (2008b). *Better Cancer Care, An Action Plan*, RR Donnelly, Edinburgh.

Scottish Government (2008c). *Equally Well: Report of the Ministerial Task Force on Health Inequalities*, Scottish Government, Edinburgh.

Scottish Government (2008d). *Long - Term Monitoring of Health Inequalities First Report on Headline Indicators - September 2008,*Scottish Government, Edinburgh.

Scottish Government (2008e). *Scotland's Future is Smoke-Free: A Smoking Prevention Action Plan*, Scottish Government, Edinburgh.

Scottish Government (2009). *Long-term monitoring of health inequalities: Health indicators - September 2009*, RR Donnelly, Edinburgh.

Scottish Government. (2010a).*Keep Well*, Scottish Government, Edinburgh.

Scottish Government. The Census (Scotland) Order 2010 Draft Scottish Statutory Instruments. Draft Scottish Statutory Instruments. B3429472. [accessed 14-4-2010b].

Scottish Government (2011). *Income, employment and access deprived rural datazone,* Scottish Government, Edinburgh.

Scottish Government (2012a). *Long term monitoring of health inequalities*, APS Group Scotland, Edinburgh.

Scottish Government (2012b). *Scotland's People Annual Report: Results from 2011 Scottish Household Survey*, Scottish Government, Edinburgh.

Scottish Government (2012c) The Scottish Index of Multiple Deprivation Scottish Government Website. [website] http://www.gov.scot/Topics/Statistics/SIMD [accessed 1-2-2014c].

Scottish Government (2013a). *Getting our priorities right*, Scottish Government, Edinburgh.

Scottish Government (2013b) Guidance on the definition of SIMD quintiles Scottish Government. [website] http://www.gov.scot/Resource/0043/00439496.pdf [accessed 15-9-2017b].

Scottish Government (2013c). *Long-term Monitoring of Health Inequalities: Headline Indicators - October 2013*, Scottish Government, Edinburgh.

Scottish Government (2013d) Scottish Index of Multiple Deprivation 2012 Scottish Government Website. [website] http://www.gov.scot/Topics/Statistics/SIMD/DataAnalysis/Background-Data-2012 [accessed 27-8-2017d].

Scottish Government (2014). *Poverty and income inequality in Scotland: 2012/13*, Scottish Government, Edinburgh.

Scottish Government (2015). *Long Term Monitoring of Health Inequalities: Headline Indicators – October 2015* Edinburgh.

Scottish Government (2016). *Scotland's Budget: Draft Budget 2017-18*, Scottish Government, Edinburgh.

Scottish Government (2017a). *A healthier future - action and ambitions on diet, activity and healthy weight*, Scottish Government, Edinburgh.

Scottish Government (2017b) Devolved and Reserved Matters Scottish Government Website. [website] http://www.parliament.scot/visitandlearn/Education/18642.aspx [accessed 6-12-2017b].

Scottish Government (2017c). *Long-term monitoring of health inequalities*, Scottish Government, Edinburgh.

Scottish Government (9-10-2017d) Mortality Rates Scottish Government Website. [website] http://www.gov.scot/Topics/Statistics/Browse/Health/TrendMortalityRates [accessed 26-10-2017d].

Scottish Government (3-10-2017e) National Indicators Scottish Government Website. [website] http://www.gov.scot/About/Performance/scotPerforms/indicator [accessed 26-10-2017e].

Scottish Government (2017f) Scottish Health Survey: Design and Content Scottish Government. [website] http://www.gov.scot/Topics/Statistics/Browse/Health/scottish-health-survey/SurveyDesignContent [accessed 18-9-2017f].

Scottish Government (2017g). *SIMD Scottish Index of Multiple Deprivation SIMD 16 Technical Notes*, Scottish Government, Edinburgh.

Scottish Government (4-10-2017h) Universal Credit: information for new applicants Scottish Government Website. [website] https://beta.gov.scot/publications/universal-credit-applicant-information/ [accessed 18-12-2017h].

Scottish Parliament (1-1-2017) List of Devolved Powers 1999-2016 Scottish Parliament Website. [website] http://www.parliament.scot/images/Parliament%20Publications/ListDevolvedPowers_1999-2016.pdf [accessed 18-12-2017].

Scottish Public Health Observatory (2017). *Economic Data Inquiry Scottish Public Health Observatory EDI-023*, Scottish Government, Edinburgh.

Shack, L., Jordan, C., Thomson, C. S., Mak, V., Moller, H., & UK Assoc, C. R. (2008) Variation in incidence of breast, lung and cervical cancer and malignant melanoma of skin by socioeconomic group in England. BMC Cancer; 8: 271.

Sharpe, K. H., McMahon, A. D., McClements, P., Watling, C., Brewster, D. H., & Conway, D. I. (2012) Socioeconomic inequalities in incidence of lung and upper aero-digestive tract cancer by age, tumour subtype and sex: A population-based study in Scotland (2000-2007). Cancer Epidemiology; 36: E164-E170.

Shavers, V. L. (2007) Measurement of socioeconomic status in health disparities research. Journal of the National Medical Association; 99: 1013-1023.

Shaw, D., Macpherson, L., & Conway, D. (2009) Tackling Socially Determined Dental Inequalities: Ethical Aspects of Childsmile, the National Child Oral Health Demonstration Programme in Scotland. Bioethics; 23: 131-139.

Sheiham, A. & Watt, R. G. (2000) The common risk factor approach: a rational basis for promoting oral health. Community Dent Oral Epidemiol; 28: 399-406.

Shi, Y., Li, T., Wang, Y., Zhou, L., Qin, Q., Yin, J., Wei, S., Liu, L., & Nie, S. (2015) Household physical activity and cancer risk: a systematic review and dose-response meta-analysis of epidemiological studies. Sci Rep; 5: 14901.

Shield, K. D., Ferlay, J., Jemal, A., Sankaranarayanan, R., Chaturvedi, A. K., Bray, F., & Soerjomataram, I. (2017) The global incidence of lip, oral cavity, and pharyngeal cancers by subsite in 2012. CA Cancer J Clin; 67: 51-64.

Shiels PG, McGlynn LM, MacIntyre A, Johnson PCD, & Batty GD (2011) Accelerated Telomere Attrition Is Associated with Relative Household Income, Diet and Inflammation in the pSoBid Cohort. PLoS ONE; 6: e22521.

Shohaimi, S., Welch, A., Bingham, S., Luben, R., Day, N., Wareham, N., & Khaw, K. T. (2004) Area deprivation predicts lung function independently of education and social class. European Respiratory Journal.24(1):157-61.

Shouls, S., Congdon, P., & Curtis, S. (1996) Modelling inequality in reported long term illness in the UK: Combining individual and area characteristics. Journal of Epidemiology and Community Health; 50: 366-376.

Sidorchuk, A., Agardh, E. E., Aremu, O., Hallqvist, J., Allebeck, P., & Moradi, T. (2009) Socioeconomic differences in lung cancer incidence: a systematic review and meta-analysis. Cancer Causes & Control; 20: 459-471.

Siemiatycki, J., Richardson, L., & Boffetta, P. (2009) Occupation. In *Cancer Epidemiology and Prevention*, D. Schottenfeld & J. Fraumeni, eds., Oxford University Press, New York.

SIGN (2013). *SIGN 134 Treatment of primary breast cancer*, SIGN, Edinburgh.

Singh, G. K., Miller, B. A., Hankey, B. F., Feuer, E. J., & Pickle, L. W. (2002) Changing area socioeconomic patterns in US cancer mortality, 1950-1998: Part I - All cancers among men. Journal of the National Cancer Institute; 94: 904-915.

Singh, S., Devanna, S., Edakkanambeth, V. J., Murad, M. H., & Iyer, P. G. (2014) Physical activity is associated with reduced risk of esophageal cancer, particularly esophageal adenocarcinoma: a systematic review and meta-analysis. BMC Gastroenterol; 14: 101.

SLS (2017) Introduction to the Scottish Longitudinal Study Scottish Longitudinal Study. [website] http://sls.lscs.ac.uk/about/ [accessed 18-9-2017].

Smith, J. P. (2007) The impact of socioeconomic status on health over the life-course. Journal of Human Resources; 42: 739-764.

Socialist Health Association Scotland (2015). *Commission on Health Inequalities; Report for the Scottish Labour Party*, Scottish Labour Party, Glasgow.

Solar, O. & Irwin, A. (2010). *A Conceptual Framework for Action on the Social Determinants of Health; Social Determinants of Health Discussion Paper 2*, WHO Press, Geneva.

Sondergaard, G., Mortensen, L. H., Andersen, A. M. N., Andersen, P. K., Dalton, S. O., & Osler, M. (2013) Social inequality in breast, lung and colorectal cancers: a sibling approach. BMJ Open; 3: pii: e002114.

Spadea, T., d'Errico, A., Demaria, M., Faggiano, F., Pasian, S., Zanetti, R., Rosso, S., Vicari, P., & Costa, G. (2009) Educational inequalities in cancer incidence in Turin, Italy. European Journal of Cancer Prevention; 18: 169-178.

Spadea, T., Zengarini, N., Kunst, A., Zanetti, R., Rosso, S., & Costa, G. (2010) Cancer risk in relationship to different indicators of adult socioeconomic position in Turin, Italy. Cancer Causes & Control; 21: 1117-1130.

Spitz MR, Wu X, & Wilkinson A, W. Q. (2006) Cancer Epidemiology and Prevention. 3rd edn, D. Schottenfeld & J. F. Fraumeni, eds., Oxford University Press Inc., Oxford, pp. 1-1392.

Spitz, M. R., Wu, X. I. F. E., Wilkinson, A. N. N. A., & Wei, Q. I. N. G. (2006) Cancer of the Lung. In *Cancer Epidemiology and Prevention*, D. Schottenfeld & J. Fraumeni, eds., Oxford University Press, New York.

Stellman, S. & Resnicow, K. (1997) Tobacco smoking, cancer and social class. In *Social Inequalities and Cancer*, No. 138 edn, M. Kogevinas et al., eds., IARC, Lyon, 229-247.

Stewart, C. H., Berry, P., Przulj, D., & Treanor, C. (2017) Cancer-related health behaviours of young people not in education, employment or training ('NEET'): a cross-sectional study. BMC Cancer; 17: 165.

Stringhini, S., Carmeli, C., Jokela, M., Avendano, M., Muennig, P., Guida, F., Ricceri, F., d'Errico, A., Barros, H., Bochud, M., Chadeau-Hyam, M., Clavel-Chapelon, F., Costa, G., Delpierre, C., Fraga, S., Goldberg, M., Giles, G. G., Krogh, V., Kelly-Irving, M., Layte, R., Lasserre, A. M., Marmot, M. G., Preisig, M., Shipley, M. j., Vollenweider, P., Zins, M., Kawachi, I., Steptoe, A., Mackenbach, J. P., Vineis, P., & Kivimaki, M. (2017) Socioeconomic status and the 25 x 25 risk factors as determinants of premature mortality: a multi-cohort study and meta-analysis of 1.7 million men and women. Lancet; 389: 1229-1237.

Stringhini, S., Sabia S, Shipley, M. j., Brunner, E., Nabi, H., Kivimaki, M., & Singh-Manoux A (2010) Association of socioeconomic position with health behaviours and mortality. JAMA; 30: 1159-1166.

Stuckler, D. & Basu, S. (2013) *The Body Economic: Why Austerity Kills* Basic Books, Philedelphia.

Susser, I. (1997) Social theory and social class. In *Social Inequalities and Cancer*, M. Kogevinas et al., eds., International Agency for Research on Cancer, Lyon.

Sweeting, H., Green, M., Benzeval, M., & West, P. (2015) The emergence of health inequalities in early adulthood: evidence on timing and mechanisms from a West of Scotland cohort. BMC Public Health; 16: 41.

Syrjanen, S. (2004) HPV infections and tonsillar carcinoma. Journal of Clinical Pathology; 57: 449-455.

Szreter, S. & Woolcock, M. (2002) Health by association? Social capital, social theory, and the political economy of public health. International Journal of Epidemiology; 33: 650-667.

Thun, M. J. & Henley, S. J. (2009) Tobacco. In *Cancer Epidemiology and Prevention*, D. Schottenfeld & J. Fraumeni, eds., Oxford University Press, New York.

Torre, L. A., Siegel, R. L., Ward, E. M., & Jemal, A. (2016) Global Cancer Incidence and Mortality Rates and Trends-An Update. Cancer Epidemiology Biomarkers & Prevention; 25: 16-27.

Torres-Cintron, M., Ortiz, A. P., Ortiz-Ortiz, K. J., Figueroa-Valles, N. R., Perez-Irizarry, J., Diaz-Medina, G., Torre-Feliciano, T., & Suarez-Perez, E. (2012) Using a socioeconomic position index to assess disparities in cancer incidence and mortality, Puerto Rico, 1995-2004. Preventing chronic disease; 9: E15.

UCL Institute of Health Equity (2013). *Review of social determinants and the health divide in the WHO European Region: final report*, WHO Regional Office for Europe, Copenhagen.

Uphoff, E. P., Pickett, K. E., Cabieses, B., Small, N., & Wright, J. (2013) A systematic review of the relationships between social capital and socioeconomic inequalities in health: a contribution to understanding the psychosocial pathway of health inequalities. Int J Equity Health; 12: 54.

van, der Heide., I, van Rijn, R. M., Robroek, S. J., Burdorf, A., & Proper, K. I. (2013) Is retirement good for your health? A systematic review of longitudinal studies. BMC Public Health; 13: 1180.

Vohra, J., Marmot, M. G., Bauld, L., & Hiatt, R. A. (2016) Socioeconomic position in childhood and cancer in adulthood: a rapid-review. J Epidemiol Community Health; 70: 629-634.

von dem, Knesebeck, O., Vonneilich, N., & Kim, T. J. (2017) Public awareness of poverty as a determinant of health: survey results from 23 countries. Int J Public Health; https://doi.org/10.1007/s00038-017-1035-9

Walsh, D., Bendel, N., Jones, R., & Hanlon, P. (2010). *Investigating a "Glasgow Effect": why do equally deprived UK cities experience different health outcomes?* Glasgow Centre for Population Health, Glasgow, Briefing Paper 25 Findings.

Wane, K., Berry, K., Kidner, C., & Georghiou, N. (2016). *SPICe Briefing: New Social Security Powers* 16/45, Scottish Parliament, Edinburgh.

Wang, Z., McLoone, P., & Morrison, D. S. (2015) Diet, exercise, obesity, smoking and alcohol consumption in cancer survivors and the general population: a comparative study of 16 282 individuals. Br J Cancer; 112: 572-575.

Ward, E., Jemal, A., Cokkinides, V., Singh GK, Cardinez C, Ghafoor A, & Thun M (2004) Cancer Disparities by Race/Ethnicity and Socioeconomic Status. CA: A Cancer Journal for Clinicians; 54: 78-93.

Watt, R. G. (2002) Emerging theories into the social determinants of health: implications for oral health promotion. Community Dent Oral Epidemiol; 30: 241-247.

Watt, R. G. (2007) From victim blaming to upstream action: tackling the social determinants of oral health inequalities. Community Dent Oral Epidemiol; 35: 1-11.

Watt, R. G. (2011) GPs at the Deep End. British Journal of General Practice; 61: 66-67.

Watt, R. G. & Sheiham, A. (2012) Integrating the common risk factor approach into a social determinants framework. Community Dent Oral Epidemiol.; 40: 289-296.

WCRF/AICR 2007, *Food, Nutrition, Physical Activity, and the Prevention of Cancer: a Global Perspective* American Institute for Cancer Research, Washington, DC.

Weiderpass, E. & Pukkala, E. (2006) Time trends in socioeconomic differences in incidence rates of cancers of gastro-intestinal tract in Finland. BMC Gastroenterology; 6: 41.

Wentzensen, I. M., Mirabello, L., Pfeiffer, R. M., & Savage, S. A. (2011) The Association of Telomere Length and Cancer: a Meta-analysis. Cancer Epidemiology Biomarkers & Prevention; 20: 1238-1250.

WHO (2011). World Conference on Social Determinants of Health: Rio Political Declaration on Social Determinants of Health, WHO, ed., WHO, Rio de Janeiro, 1-7.

WHO (2003). *Diet, nutrition and the prevention of chronic diseases 916*, World Health Organisation, Geneva.

Wild, C. P. (2012) The exposome: from concept to utility. Int J Epidemiol; 41: 24-32.

Wild, S. H., Fischbacher, C. M., Brock, A., Griffiths, C., & Bhopal, R. (2006) Mortality from all cancers and lung, colorectal, breast and prostate cancer by country of birth in England and Wales, 2001-2003. British Journal of Cancer; 94: 1079-1085.

Wilford, J., McMahon, A. D., Peters, J., Pickvance, S., Jackson, A., Blank, L., Craig, D., O'Rourke, A., & Macdonald, E. B. (2008) Predicting job loss in those off sick. Occup.Med (Lond); 58: 99-106.

Wilkinson, R and Pickett, K (2009) The Spirit Level The Equality Trust. [website] https://www.equalitytrust.org.uk/spirit-level [accessed 25-10-2017].

Willett, W. C. (2009) Diet and Nutrition. In *Cancer Epidemiology and Prevention*, D. Schottenfeld & J. Fraumeni, eds., Oxford University Press, New York.

Williams, C. (1995) Healthy eating: clarifying advice about fruit and vegetables. BMJ; 310: 1453-1455.

Winn, D. M., Lee, Y. C., Hashibe, M., & Boffetta, P. (2015) The INHANCE consortium: toward a better understanding of the causes and mechanisms of head and neck cancer. Oral Diseases; 21: 685-693.

Women's Own. Margaret Thatcher: a life in quotes. In The Guardian Online, 8-4-2013. [website] https://www.theguardian.com/politics/2013/apr/08/margaret-thatcher-quotes [accessed September 2017).

Woodward, A. & Kawachi, I. (2000) Why reduce health inequalities? Journal of Epidemiology and Community Health; 54: 923-929.

Working Group 21 (2016). *IARC identifies eight additional cancer sites linked to overweight and obesity*, IARC, Lyon, France, Press Release 247.

Working Party of the Royal College of Physicians (2001). *UK: Alcohol-can the NHS afford it? Recommendations for a coherent alcohol strategy for hospitals*, Royal College of Physicians, London.

World Bank Group. Country Income Groups (World Bank Classification). 2016.

World Health Organisation (2003). *Social Determinants of Health.The Solid Factors*, World Health Organisation, Copenhagen, 2nd Edition.

World Health Organisation (2006). *Air Quality Guidelines Global Update 2005*, World Health Organisation, Copenhagen.

Wu, X. F., Amos, C. I., Zhu, Y., Zhao, H., Grossman, B. H., Shay, J. W., Luo, S., Hong, W. K., & Spitz, M. R. (2003) Telomere dysfunction: A potential cancer predisposition factor. Journal of the National Cancer Institute; 95: 1211-1218.

Yen, Y. C. & Lung, F. W. (2013) Older adults with higher income or marriage have longer telomeres. Age and Ageing; 42: 234-239.

## Appendices

Appendix 4.1 Scottish Longitudinal Study (SLS) Project Clearance Form


**SLS PROJECT CLEARANCE FORM**

SLS Unit, Room 1G1,
Ladywell House,
Ladywell Road,
Edinburgh,
EH12 7TF.

21 October 2009

Dear Katharine

Your project proposal "*Areal versus Individual Measures of SEP – How do they compare in predicting cancer incidence?*" has now been assessed by the SLS Research Board. The board decided that your proposal has been:

☒ Cleared as it stands

☐ Cleared as it stands, with some suggestions for change/improvement

☐ Cleared under condition there are minor changes

☐ Cleared under condition there are major changes

☐ Rejected

With kind regards,

Claire Boag, SLS project manager

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

Appendix 5.1 NHS NSS PAC Approval letter

**Information Services Division**

Area 151A
Gyle Square
1 South Gyle Crescent
Edinburgh, EH12 9EB
Telephone 0131 275 6000
Fax 0131 275 7606
**www.isdscotland.org**

NHS National Services Scotland

Mrs Katharine Sharpe
Information Consultant
NHS NSS
1 Gyle Square
EDINBURGH
EH12 9EB

Date 25.04.2013
Your Ref
Our Ref 100/12

Enquiries to Janet Murray
Direct Line 0131 275 6954
Email janet.murray1@nhs.net

Dear Mrs Sharpe

**Scottish Health Survey and Scottish Cancer Registry Linkage to explore risk behaviour association, socioeconomic inequalities and lung, upper-aero digestive tract and all cancer risk**

The Privacy Advisory Committee has considered and approved your application for a data linkage in support of the above study.

*Conditions applied:* None

*Time period:* As specified

*Points highlighted:* None

The approval of the Committee is for a period of 5 years from the date of this letter. Any change to the terms of your application, including changes in data user(s), additional data fields or extension of the time period approved must be requested through Susan Kerr, PAC Administrator on 0131 275 6445 or nss.pac@nhs.net

In order to progress your request please contact Janey Read at Janey.Read@nhs.net or telephone 0131 275 6703.

Please note that the following details about your application will be published under the following headings on the PAC website at http://www.nhsnss.org/pages/corporate/pac_meetings_and_decision_making.php later this year:

| No | Title | Type | Summary | Date sent to PAC | PAC Responses | NSS Decision | Date Completed |
|---|---|---|---|---|---|---|---|
| | | | | | | | |

If you have any queries about this please contact Patricia Ruddy patricia.ruddy@nhs.net.

Kind regards.

Yours sincerely

JS Murray

Dr Janet Murray
Consultant in Public Health Medicine

cc Janey Read

**Appendix 5.2 University of Glasgow MVLS Ethics Committee Approval**

University of Glasgow | College of Medical, Veterinary & Life Sciences

26th June 2013

Dear Mrs. Katharine Sharpe

**MVLS College Ethics Committee**

***Project Title:*** Scottish Health Survey and Scottish Cancer Registry Linkage to explore risk behaviour association, socioeconomic inequalities and lung, upper-aero digestive tract and all cancer risk
***Project No:*** 200120069

The College Ethics Committee has reviewed your application and has agreed that there is no objection on ethical grounds to the proposed study. They are happy therefore to approve the project, subject to the following conditions

- The research should be carried out only on the sites, and/or with the groups defined in the application.
- Any proposed changes in the protocol should be submitted for reassessment, except when it is necessary to change the protocol to eliminate hazard to the subjects or where the change involves only the administrative aspects of the project. The Ethics Committee should be informed of any such changes.
- If the study does not start within three years of the date of this letter, the project should be resubmitted.
- You should submit a short end of study report to the Ethics Committee within 3 months of completion.

Yours sincerely

Prof. Andrew C. Rankin
Deputy Chair, College Ethics Committee

Andrew C. Rankin
Professor of Medical Cardiology
BHF Glasgow Cardiovascular Research Centre
College of Medical, Veterinary & Life Sciences
University of Glasgow, G12 8TA
Tel: 0141 211 4833
Email: Andrew.rankin@glasgow.ac.uk

## Appendix 5.3 Scottish Health Survey (SHeS) socioeconomic and behaviour variable list and description by SHeS year

| | SCOTTISH HEALTH SURVEY YEAR: ECONOMIC ACTIVITY | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 1995 | | 1998 | | 2003 | | 2008, 2009, 2010, 2011 | |
| Variable | Variable / Description | Category code, description | Variable / Description | Category code, description | Variable / Description | Category code, description | Variable / Description | Category code, description |
| Economic activity | ECSTA Economic activity status | -9.00 M Not answered<br>-8.00 M Don't know<br>-6.00 M Schedule not obtained<br>-2.00 M Schedule not applicable<br>-1.00 M Item not applicable<br>1.00 FT work<br>2.00 PT work<br>3.00 Work - unspecified hrs<br>4.00 Unemployed<br>5.00 Permanently sick<br>6.00 Retired<br>7.00 Keeping house<br>8.00 FT student (no job)<br>9.00 Other inactive | ECONACT Economic activity | -9.00 M Not answered/refused<br>-8.00 M Don't know<br>-6.00 M schedule not obtained<br>-1.00 M not applicable<br>1.00 In employment<br>2.00 Unemployed<br>3.00 Retired<br>4.00 Other economically inactive | ECONACT Economic activity | -9 No answer/refused<br>-8 Don't know<br>-2 Schedule not applicable<br>-1 Item not applicable<br>1 In employment<br>2 ILO unemployed<br>3 Retired<br>4 Other economically inactive | ECONAC08 Economic status | 1 In education<br>2 In paid employment, self-employed or on government<br>3 Perm unable to work<br>4 Looking for/intending to look for paid work<br>5 Retired<br>6 Looking after home/family<br>7 Doing something else.<br>Lo through -1 Missing or Inactive |

| | SCOTTISH HEALTH SURVEY YEAR: SIMD | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 1995 | | 1998 | | 2003 | | 2008, 2009, 2010, 2011 | |
| Variable | Variable / Description | Category code, description | Variable / Description | Category code, description | Variable / Description | Category code, description | Variable / Description | Category code, description |
| SIMD | SCSIMD2012 | 1 Most deprived<br>5 Least deprived | SCSIMD2012 | 1 Most deprived<br>5 Least deprived | SIMD5 | 1 Least deprived<br>5 Most deprived | SIMD5_RP | 1 Least deprived<br>5 Most deprived |

## Appendix 5.3 Scottish Health Survey (SHeS) socioeconomic and behaviour variable list and description by SHeS year

| | SCOTTISH HEALTH SURVEY YEAR: HIGHEST QUALIFICATION | | | |
|---|---|---|---|---|
| | 1995 | | 1998 | |
| Variable | Variable / Description | Category code, description | Variable / Description | Category code, description |
| Highest educational qualifications | TOPQUAL5 Highest education level | -9.00 M Not answered<br>-8.00 M Don't know<br>-6.00 M Schedule not obtained<br>-2.00 M Schedule not applicable<br>-1.00 M Item not applicable<br>1.00 Degree or equivalent<br>2.00 Higher, below degree<br>3.00 A level or equivalent<br>4.00 GCSE A-C or equivalent<br>5.00 GCSE D-G or equivalent | TOPQUAL Highest education level | -9 M Not answered/refused<br>-8 M Don't know<br>-6 M schedule not obtained<br>-1 M not applicable<br>1 Degree or degree level qualification<br>2 SCE Higher/CSYS (Certificate of Sixth Year)<br>3 SCE Ordinary ( O Grades) Bands A - C<br>4 Standard Grade (Level 1 - 3)<br>5 SLC Lower<br>6 SUPE Lower or Ordinary<br>7`O level passes (Grade A - C if after 1)<br>8 GCSE (grade A - C)<br>9 CSE Grade 1<br>10 School Certificate or Metric<br>11 SCE Ordinary ( O Grades) Bands D & E<br>12 Standard Grade (Level 4, 5)<br>13 CSE Grades 2 - 5<br>14 GCE `O Grades D & E (if after 1975)<br>15 GCSE (Grades D, E, F, G)<br>16 CSE ungraded<br>17 Foreign qualifications<br>18 Other academic qualifications<br>19 No academic qualifications |

| | SCOTTISH HEALTH SURVEY YEAR: HIGHEST QUALIFICATION | | | |
|---|---|---|---|---|
| | 2003 | | 2008, 2009, 2010, 2011 | |
| Variable | Variable / Description | Category code, description | Variable / Description | Category code, description |
| Highest educational qualifications continued | HEDQUAL Highest educational qualification | 1 School leaving certificate/ NNQ Access Unit<br>2 O grade / Standard grade / GCSE / CSE<br>3 GSVQ found / SVQ level 1 or 2 / Scotvec module<br>4 Higher grade / A level / CSYS<br>5 GSVQ advanced / SQV level 3 / ONC, OND<br>6 City and Guilds<br>7 HNC / HND / SQV level 4 or 5<br>8 First degree / Higher degree<br>9 Professional qualifications<br>10 None of these qualifications | hedqul08 (D) Highest educational qualification revised 2008 | 1 Degree or higher<br>2 HNC/D or equivalent<br>3 Higher grade or equivalent<br>4 Standard or equivalent<br>5 Other school level<br>6 No qualifications<br>-8 Don't know<br>-9 Not answered<br>-2 Schedule not applicable<br>-1 Item not applicable. |

## Appendix 5.3 Scottish Health Survey (SHeS) socioeconomic and behaviour variable list and description by SHeS year

| | SCOTTISH HEALTH SURVEY YEAR: OCCUPATIONAL SOCIAL CLASS | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 1995 | | 1998 | | 2003 | | 2008, 2009, 2010, 2011 | |
| Variable | Variable / Description | Category code, description | Variable / Description | Category code, description | Variable / Description | Category code, description | Variable / Description | Category code, description |
| Occupational Social Class | SOCCLS Respondent's social class | -1.00 M Item not applicable<br>1.00 I Professional<br>2.00 II Intermediate<br>3.00 III Skilled non manual<br>4.00 III Skilled manual<br>5.00 IV Partly Skilled<br>6.00 V Unskilled<br>7.00 Armed Forces<br>8.00 Not fully described | SCCIEG6 (D) social class of chief income earner | -9.00 M Not answered/refused<br>-8.00 M Don't know<br>-6.00 M schedule not obtained<br>-1.00 M not applicable<br>1.00 I Professional<br>2.00 II Managerial technical<br>3.10 IIN Skilled non-manual<br>3.20 IIIM Skilled manual<br>4.00 IV Semi-skilled manual<br>5.00 V Unskilled manual | SCCIEG7: (D) Social Class of Chief Income Earner - I, II, III N, III M, IV, V, Others | 1 I Professional<br>2 II Managerial technical<br>3 IIIN Skilled non manual<br>4 IIIM Skilled manual<br>5 IV Semiskilled manual<br>6 V Unskilled manual<br>7 Others | schrpg7 (D) Social Class of HRP - I, II, IIIN, IIIM, IV, V, Others | 1 I Professional<br>2 II Managerial<br>3 IIIN Skilled non<br>4 IIIM Skilled manual<br>5 IV Semiskilled<br>6 V Unskilled manual<br>7 Others. |

| | SCOTTISH HEALTH SURVEY YEAR: CAR OWNERSHIP | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 1995 | | 1998 | | 2003 | | 2008, 2009, 2010, 2011 | |
| Variable | Variable / Description | Category code, description | Variable / Description | Category code, description | Variable / Description | Category code, description | Variable / Description | Category code, description |
| Car ownership | NUMCAR | 1 None<br>2 One<br>3 Two<br>4 Three or more<br>-9 Not answered | numcars | * M Not answered/refused<br>* M Don't know<br>* M schedule not obtained<br>* M not applicable<br>1 One<br>2 Two<br>3 Three or more | NumCars Number of cars available | 1 One<br>2 Two<br>3 Three or more | NumCars Number of cars available | 1 One<br>2 Two<br>3 Three or more |
| | | | | | CAR: Car or van available | 1 Yes<br>2 No | Car: Car or van available | 1 Yes<br>2 No |

## Appendix 5.3 Scottish Health Survey (SHeS) socioeconomic and behaviour variable list and description by SHeS year

| | SCOTTISH HEALTH SURVEY YEAR: HOUSING TENURE | | | | |
|---|---|---|---|---|---|
| | 1995 | | | 1998 | |
| Variable | Variable / Description | Category code, description | | Variable / Description | Category code, description |
| Housing tenure | TENURE Housing tenure (broad categories) | -9.00 M | Not answered (9) | ownorent | -9 M Not answered/refused |
| | | -6.00 M | Schedule not obtained | | -8 M Don't know |
| | | -2.00 M | Schedule not applicable | | -6 M schedule not obtained |
| | | -1.00 M | Item not applicable | | -1 M not applicable |
| | | 1.00 | Owner-occupier | | 1 Owns with mortgage/loan |
| | | 2.00 | Rents LA | | 2 Owns outright |
| | | 3.00 | Rents privately | | 3 Rents from local authority/new town |
| | | | | | 4 Rents from housing association |
| | | | | | 5 Rents - privately, unfurnished |
| | | | | | 6 Rents - privately, furnished |
| | | | | | 7 Rents from employer |
| | | | | | 8 Rents - other with payment |
| | | | | | 9 Rent free |

| | SCOTTISH HEALTH SURVEY YEAR: HOUSING TENURE | | | |
|---|---|---|---|---|
| | 2003 | | 2008, 2009, 2010, 2011 | |
| Variable | Variable / Description | Category code, description | Variable / Description | Category code, description |
| Housing tenure continued | OWNORENT Household tenure | 1 Owns with mortgage/loan | OwnRnt08 Household tenure | 1 Buying it with the help of a mortgage or loan |
| | | 2 Owns outright | | 2 Own it outright |
| | | 3 Rents from local authority/new town | | 3 Pay part rent and part mortgage (shared ownership) |
| | | 4 Rents from housing association | | 4 Rent it |
| | | 5 Rents - privately, unfurnished | | 5 Tied accommodation (e.g. where the accommodation goes with your job) |
| | | 6 Rents - privately, furnished | | 6 Live here rent free (including rent-free in relative's/friend's property) |
| | | 7 Rents from employer | | |
| | | 8 Rents - other with payment | | |
| | | 9 Rent free | | |

## Appendix 5.3 Scottish Health Survey (SHeS) socioeconomic and behaviour variable list and description by SHeS year

| | SCOTTISH HEALTH SURVEY YEAR: SMOKING VARIABLES | | | |
|---|---|---|---|---|
| | 1995 | | 1998 | |
| Variable | Variable / Description | Category code, description | Variable / Description | Category code, description |
| Smoking status | SMOKEVER Ever smoked a cigarette, cigar or pipe | -9.00 M Not answered<br>-8.00 M Don't know<br>-1.00 M Item not applicable<br>1.00 Yes<br>2.00 No | cigst1 (D) Cigarette Smoking Status - Never/Ex-regular/Ex-occasional/ Current | 1 Never smoked cigarettes at all<br>2 Used to smoke cigarettes occasionally<br>3 Used to smoke cigarettes regularly<br>4 Current cigarette smoker |
| | SMOKENOW Do you smoke cigarettes at all nowadays | -9.00 M Not answered<br>-8.00 M Don't know<br>-1.00 M Item not applicable<br>1.00 Yes<br>2.00 No | | |
| | SMOKEREG Smoked cigarettes regularly/occasionally | -1.00 M Item not applicable<br>1.00 Smoked cigarettes regularly, at least 1 per day<br>2.00 Smoked them only occasionally<br>3.00 Never really smoked cigarettes, just tried them once or twice | | |
| Number of cigarettes smoked | NUMSMOK how many cigarettes did you smoke in a day | -9.00 M Not answered<br>-8.00 M Don't know<br>-7.00 M Smokes roll-ups<br>-1.00 M Item not applicable | CIGST2: (D) Cigarette Smoking Status - Banded current smokers | 1 Light smokers, under 10 a day<br>2 Moderate smokers, 10 to under 20 a day<br>3 Heavy smokers, 20 or more a day<br>4 Don't know number smoked a day<br>5 Non-smoker |

| | SCOTTISH HEALTH SURVEY YEAR: SMOKING VARIABLES | | | |
|---|---|---|---|---|
| | 2003 | | 2008, 2009, 2010, 2011 | |
| Variable | Variable / Description | Category code, description | Variable / Description | Category code, description |
| Smoking status | CIGST1: (D) Cigarette Smoking Status - Never/Ex-regular/Ex- | 1 Never smoked cigarettes at all<br>2 Used to smoke cigarettes occasionally<br>3 Used to smoke cigarettes regularly<br>4 Current cigarette smoker | cigst1 (D) Cigarette Smoking Status - Never/Ex-regular/Ex-occasional/ Current | 1 Never smoked cigarettes at all<br>2 Used to smoke cigarettes occasionally<br>3 Used to smoke cigarettes regularly<br>4 Current cigarette smoker. |
| Number of cigarettes smoked | CIGST2: (D) Cigarette Smoking Status - Banded current smokers | 1 Light smokers, under 10 a day<br>2 Moderate smokers, 10 to under 20 a day<br>3 Heavy smokers, 20 or more a day<br>4 Don't know number smoked a day<br>5 Non-smoker | cigst2 (D) Cigarette Smoking Status - Banded current smokers | 1 Light smokers, under 10 a day<br>2 Moderate smokers, 10 to under 20 a day<br>3 Heavy smokers, 20 or more a day<br>4 Don’t know number smoked a day<br>5 Non-smoker |

## Appendix 5.3 Scottish Health Survey (SHeS) socioeconomic and behaviour variable list and description by SHeS year

| | SCOTTISH HEALTH SURVEY YEAR: SMOKING VARIABLES CONTINUED | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 1995 | | 1998 | | 2003 | | 2008, 2009, 2010, 2011 | |
| Variable | Variable / Description | Category code, description | Variable / Description | Category code, description | Variable / Description | Category code, description | Variable / Description | Category code, description |
| Age started smoking | SMOKYRS How many years did you smoke | -9.00 M Not answered<br>-8.00 M Don't know<br>-1.00 M Item not applicable | SMOKYRS How many years did you smoke | -9.00 M Not answered<br>-8.00 M Don't know<br>-1.00 M Item not | SMOKYRS How many years did you smoke | -9.00 M Not answered<br>-8.00 M Don't know<br>-1.00 M Item not | SMOKYRS How many years did you smoke | -9.00 M Not answered<br>-8.00 M Don't know<br>-1.00 M Item not |
| Duration smoked | STARTSMK How old when started to smoke | -9.00 M Not answered<br>-8.00 M Don't know<br>-7.00 M Never smoked<br>-1.00 M Item not applicable | STARTSMK How old when started to smoke cigarettes | -9.00 M Not answered<br>-8.00 M Don't know<br>-7.00 M Never smoked<br>-1.00 M Item not | STARTSMK How old when started to smoke | -9.00 M Not answered<br>-8.00 M Don't know<br>-7.00 M Never smoked<br>-1.00 M Item not | STARTSMK How old when started to smoke | -9.00 M Not answered<br>-8.00 M Don't know<br>-7.00 M Never<br>-1.00 M Item not |

## Appendix 5.3 Scottish Health Survey (SHeS) socioeconomic and behaviour variable list and description by SHeS year

| SCOTTISH HEALTH SURVEY YEAR: ALCOHOL CONSUMPTION VARIABLES | | | | | | | |
|---|---|---|---|---|---|---|---|
| 1995 | | 1998 | | 2003 | | 2008, 2009, 2010, 2011 | |
| Variable / Description | Category code, description | Variable / Description | Category code, description | Variable / Description | Category code, description | Variable / Description | Category code, description |
| ALCOHOLM Alcohol consumption per week men | -9 M not answered<br>-1 M item not applicable<br>1 Ex-drinker<br>2 Non/occasional drinker<br>3 1-10 units per week<br>4 11-21 units per week<br>5 Over 21 units per week | ALCBASMT (D) Alcohol consumption: men | -9.00 M Not answered/refused<br>-8.00 M Don't know<br>-6.00 M schedule not obtained<br>-1.00 M not applicable<br>1.00Never drunk alcohol<br>2.00Ex-drinker<br>3.00Under 1 per week<br>4.00Over 1-10<br>5.00Over 10-21<br>6.00Over 21-35<br>7.00Over 35-50<br>8.00Over 50 units per week | ALCBASMT (D) Alcohol consumption: men | 1 Never drunk alcohol<br>2 Ex-drinker<br>3 Under 1 per week<br>4 Over 1-10<br>5 Over 10-21<br>6 Over 21-35<br>7 "Over 35-50"<br>8 'Over 50 units per week'. | alcbsmt (D) Alcohol consumption: men | 1 Never drunk alcohol<br>2 Ex-drinker<br>3 Under 1 per week<br>4 Over 1-10<br>5 Over 10-21<br>6 Over 21-35<br>7 "Over 35-50"<br>8 Over 50 units per week. |
| ALCOHOLW Alcohol consumption per week women | -9 M not answered<br>-1 M item not applicable<br>1 Ex-drinker<br>2 Non/occasional drinker<br>3 1-10 units per week<br>4 11-21 units per week<br>5 Over 21 units per week | ALCBASWT (D) Alcohol consumption: women | -9.00 M Not answered/refused<br>-8.00 M Don't know<br>-6.00 M schedule not obtained<br>-1.00 M not applicable<br>1.00Never drunk alcohol<br>2.00Ex-drinker<br>3.00Under 1 per week<br>4.00Over 1-7<br>5.00Over 7-14<br>6.00Over 14-21<br>7.00Over 21-35<br>8.00Over 35 | ALCBASWT (D) Alcohol consumption: women | 1 Never drunk alcohol<br>2 Ex-drinker<br>3 Under 1 per week<br>4 Over 1-7<br>5 Over 7-14<br>6 Over 14-21<br>7 Over 21-35<br>8 Over 35 | alcbswt (D) Alcohol consumption: women. | 1 Never drunk alcohol<br>2 Ex-drinker<br>3 Under 1 per week<br>4 Over 1-7<br>5 Over 7-14<br>6 Over 14-21<br>7 "Over 21-35"<br>8 Over 35. |
| PERDAY: Number of units drunk per day | -9 M not answered<br>-1 M item not applicable | D7GROUP Units drunk on heaviest day in past week | -9.00 M Not answered/refused<br>-8.00 M Don't know<br>-6.00 M Not obtained<br>-1.00 M not applicable<br>1.00 Under 2 units<br>2.00 2, under 3 units<br>3.00 3, under 4 units<br>4.00 4, under 5 units<br>5.00 5, under 6 units<br>6.00 6, under 8 units<br>7.00 8 or more units | D7GROUP Units drunk on heaviest day in past week | -9.00 M Not answered/refused<br>-8.00 M Don't know<br>-6.00 M Not obtained<br>-1.00 M not applicable<br>1.00 Under 2 units<br>2.00 2, under 3 units<br>3.00 3, under 4 units<br>4.00 4, under 5 units<br>5.00 5, under 6 units<br>6.00 6, under 8 units<br>7.00 8 or more units | D7UT08_2 Units drunk on heaviest day in past week revised for alcopops and wine | -9.00 M Not<br>-8.00 M Don't know<br>-6.00 M Not obtained<br>-1.00 M not applicable |

## Appendix 5.3 Scottish Health Survey (SHeS) socioeconomic and behaviour variable list and description by SHeS year

| SCOTTISH HEALTH SURVEY YEAR: FRUIT & VEGETABLES VARIABLES | | | | | | | |
|---|---|---|---|---|---|---|---|
| 1995 | | 1998 | | 2003 | | 2008, 2009, 2010, 2011 | |
| Variable / Description | Category code, description | Variable / Description | Category code, description | Variable / Description | Category code, description | Variable / Description | Category code, description |
| FRUIT How often do you eat fresh fruit | -9.00 M Not answered | FRUIT How often do you eat fresh fruit | -9 M Not answered/refused | PORFV (D) Total portion of fruit and vegetables yesterday | 0 None | porfv (D) Total portion of fruit and veg | 0 None |
| | -8.00 M Don't know | | -8 M Don't know | | 1 Less than 1 portion | | 1 Less than 1 portion |
| | -1.00 M Not applicable | | -6 M schedule not obtained | | 2 '>1, < 2 | | 2 '>1, < 2 |
| | 1.00 6 or more times a day | | -1 M not applicable | | 3 =>2, <3 | | 3 =>2, <3 |
| | 2.00 4 or 5 times a day | | 1 6 or more times a day | | 4 =>3, <4 | | 4 =>3, <4 |
| | 3.00 2 to 3 times a day | | 2 4 or 5 times a day | | 5 =>4, <5 | | 5 =>4, <5 |
| | 4.00 Once a day | | 3 2 or 3 times a day | | 6 =>5, <6 | | 6 =>5, <6 |
| | 5.00 5 or 6 times a week | | 4 Once a day | | 7 =>6, <7 | | 7 =>6, <7 |
| | 6.00 2 to 4 times a week | | 5 5 or 6 times a week | | 8 =>7, <8 | | 8 =>7, <8 |
| | 7.00 Once a week | | 6 2 to 4 times a week | | 9 =>8 | | 9 =>8 |
| | 8.00 1 to 3 times per month | | 7 Once a week | | | | |
| | 9.00 Less often or never | | 8 1 to 3 times per month | | | | |
| | | | 9 Less often or never | | | | |
| GREENVEG How often cooked green vegetables eaten | -9.00 M Not answered | GREENVEG How often cooked green vegetables eaten | -9 M Not answered/refused | | | | |
| | -8.00 M Don't know | | -8 M Don't know | | | | |
| | -6.00 M schedule not | | -6 M schedule not obtained | | | | |
| | -1.00 M Item not applicable | | -1 M not applicable | | | | |
| | 1.00 6 or more times a day | | 1 6 or more times a day | | | | |
| | 2.00 4 or 5 times a day | | 2 4 or 5 times a day | | | | |
| | 3.00 2 to 3 times a day | | 3 2 or 3 times a day | | | | |
| | 4.00 Once a day | | 4 Once a day | | | | |
| | 5.00 5 or 6 times a week | | 5 5 or 6 times a week | | | | |
| | 6.00 2 to 4 times a week | | 6 2 to 4 times a week | | | | |
| | 7.00 Once a week | | 7 Once a week@ | | | | |
| | 8.00 1 to 3 times per month | | 8 1 to 3 times per month | | | | |
| | 9.00 Less often or never | | 9 Less often or never | | | | |

## Appendix 5.3 Scottish Health Survey (SHeS) socioeconomic and behaviour variable list and description by SHeS year

| SCOTTISH HEALTH SURVEY YEAR: PHYSICAL ACTIVITY VARIABLE | | | | | | | |
|---|---|---|---|---|---|---|---|
| 1995 | | 1998 | | 2003 | | 2008, 2009, 2010, 2011 | |
| Variable / Description | Category code, description | Variable / Description | Category code, description | Variable / Description | Category code, description | Variable / Description | Category code, description |
| EXNUM How many times per week do you exercise | -9.00 M Not answered<br>-8.00 M Don't know<br>-1.00 M Item not applicable<br>1 Less than once a week<br>2 Once a week<br>3 2-3 times a week<br>4 4-5 times a week<br>5 6+ times a week | ADTOT30C Number of days per week any activities 30 minutes + | -8.00 M Don't know<br>-6.00 M Not obtained<br>-1.00 M not applicable<br>0 None<br>1 Less than 1<br>2 1 or 2 a week<br>3 3 or 4 a week<br>4 5 or more a week | ADTOT30C: (D) Adults: Total days per week active 30 minutes + moderate | 0 None<br>1 Less than 1<br>2 1 or 2 a week<br>3 3 or 4 a week<br>4 5 or more a | adtt15cN (D) Number of days per week any activities 30 minutes +, 15-29 min sessions included (sports = moderate effort). | 0 None<br>1 Less than 1<br>2 1 or 2 a week<br>3 3 or 4 a week<br>4 5 or more a week.<br>Missing (lo thru -1) |