THE EFFECTS OF ALCOHOL INTAKE, ADIPOSITY AND SMOKING ON THE RISK OF COLORECTAL CANCER IN UK BIOBANK

A thesis submitted for the degree of
DOCTOR OF PHILOSOPHY

Alex Whitmarsh

Division of Population Medicine, Cardiff University

2017

DECLARATION

This work has not been submitted in substance for any other degree or award at this or any other university or place of learning, nor is being submitted concurrently in candidature for any degree or other award.

Signed …… (candidate) Date ……

STATEMENT 1

This thesis is being submitted in partial fulfilment of the requirements for the degree of PhD.

Signed …… (candidate) Date ……

STATEMENT 2

This thesis is the result of my own independent work/investigation, except where otherwise stated, and the thesis has not been edited by a third party beyond what is permitted by Cardiff University's Policy on the Use of Third Party Editors by Research Degree Students. Other sources are acknowledged by explicit references. The views expressed are my own.

Signed …… (candidate) Date ……

STATEMENT 3

I hereby give consent for my thesis, if accepted, to be available online in the University's Open Access repository and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed …… (candidate) Date ……

Acknowledgements

I would like to thank my supervisors, first of all for providing me with the opportunity to undertake this PhD and also for the support and constant encouragement required to complete it. I would also like to thank my family and friends for all their support.

Summary

The burden of non-communicable diseases, including cancer, is growing globally. Epidemiological studies have shown that lifestyle factors can increase the risk of these diseases. Colorectal cancer represents the third most common cancer for men and women in the UK. This thesis investigated the relationships between three classic lifestyle risk factors - alcohol intake, adiposity and smoking - and colorectal cancer in the UK Biobank cohort.

UK Biobank is a cohort of 500,000 men and women aged 40-69 recruited between 2006-2010. Participants were followed-up for cancer and death registrations until 31st March 2014 through linkage with national datasets. Cox proportional hazards models were used to analyse these data.

This thesis found that alcohol intake was associated with colorectal cancer for men but not for women. For men, there was a dose-response relationship between alcohol intake and colorectal cancer. Results were similar using non-drinkers or never drinkers as the reference group.

Body mass index (BMI), waist circumference (WC) and waist to hip ratio (WHR) were each strongly associated with colon cancer for men but there was only slight evidence for an association with WC and WHR for women. Modelling BMI and WC/WHR together, the associations for WC/WHR remained while the association for BMI was attenuated, indicating that WC/WHR may be more directly associated with colorectal cancer risk than BMI.

While former cigarette smokers had an increased risk of colorectal cancer compared to never smokers, there was no clear evidence of an increased risk for current cigarette smokers. Furthermore, former cigarette smokers with ≥40 years smoking duration had a higher risk than current cigarette smokers with ≥40 years duration.

In conclusion, this thesis found that alcohol intake, adiposity and smoking are each associated with colorectal cancer risk. The prevalence of these risk factors should be minimised in order to prevent disease.

Contents

Chapter 1 Introduction

There has been a dramatic shift worldwide in disease patterns as non-communicable diseases (NCDs) (which mainly affect adults in later life) have replaced communicable diseases (which mainly affect infants and children) as the predominant causes of morbidity and mortality.[1, 2] NCDs now represent the leading cause of death globally, responsible for 68% of all deaths in 2012.[3] The global population is predicted to reach 8.5 billion by 2030 with the population aged 60 or over growing fastest.[4] Thus, the burden of NCDs will continue to grow worldwide. Whilst the burden of NCDs has increased as a result of more people surviving to later adulthood, it is clear that a significant proportion of NCDs are preventable.[3]

Cancer represents one of the leading NCDs and a major cause of morbidity and mortality globally. In 2012, there were an estimated 14.1 million new cancer cases and 8.2 million cancer deaths.[5] By 2030, the annual incidence of new cancer cases is expected to increase by more than 50%, to 22 million.[6] In the UK, it is predicted that 1 in 2 people born after 1960 will be diagnosed with some form of cancer in their lifetime.[7]

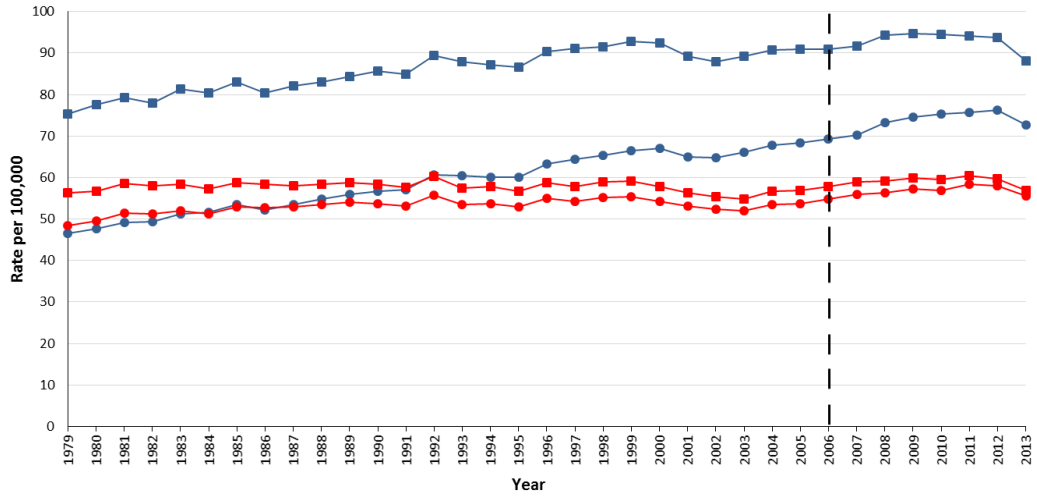
1.1 Colorectal Cancer Epidemiology

Worldwide, there were an estimated 1.4 million cases of colorectal (or bowel) cancer in 2012.[5] In the UK, 22,957 men and 18,155 women were diagnosed with colorectal cancer in 2013, making it the third most common type of cancer for both men and women.[8] The predicted lifetime risk of colorectal cancer is 7.3% for men and 5.5% for women.[8]

Trends of colorectal cancer in Great Britain show that the number of people diagnosed with colorectal cancer has increased in recent decades. From 1979-1981 to 2011-2013, the crude incidence rate of colorectal cancer (i.e. the proportion of people diagnosed with colorectal cancer) per year in Great Britain increased by 56.7% for men and 14.8% for women (Figure 1.1.1). This increase was largely due to the increase in the age of the population during this time and age standardised incidence rates showed more moderate

trends; nevertheless, the age standardised incidence rate increased by 17.0% for men and 3.2% for women from 1979-1981 to 2011-2013.

Figure 1.1.1 Colorectal Cancer Incidence Rates in Great Britain between 1979 and 2013[8]



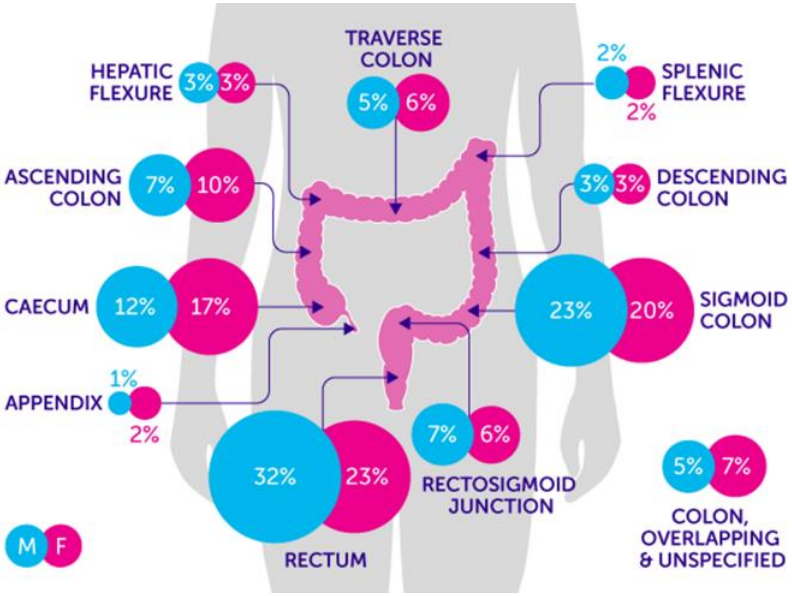
Blue circles represent crude incidence rates of colorectal cancer for men. Red circles represent crude incidence rates of colorectal cancer for women. Blue squares represent age standardised incidence rates for men. Red squares represent age standardised incidence rates for women. The dotted line represents the introduction of the bowel screening programme in England.

Cancer sites are classified according to the International Classification of Diseases for Oncology (published by the World Health Organisation (WHO)).[9] Colon cancer is defined as cancer diagnosed in the appendix through to the sigmoid colon and rectal cancer is defined as cancer diagnosed in the rectosigmoid junction and rectum (see Figure 1.1.2 for the anatomy of the large bowel). 61% of colorectal cancers diagnosed in men and 70% of colorectal cancers diagnosed in women are located in the colon. Proximal colon is usually defined as the appendix to the splenic flexure and distal colon is usually defined as the descending colon and sigmoid colon.

Colorectal cancer also represents the third most common cause of cancer death for both men and women in the UK with 16,187 deaths attributable to colorectal cancer in 2012.[8] However, the age-standardised mortality rate has decreased dramatically since 1971 (Figure 1.1.3) and the 10-year survival rate for men and women diagnosed with colorectal cancer has increased from 22% in 1971-1972 to 56% in 2010-2011.[8] The increase in survival is a result of improved treatment as well as earlier detection.

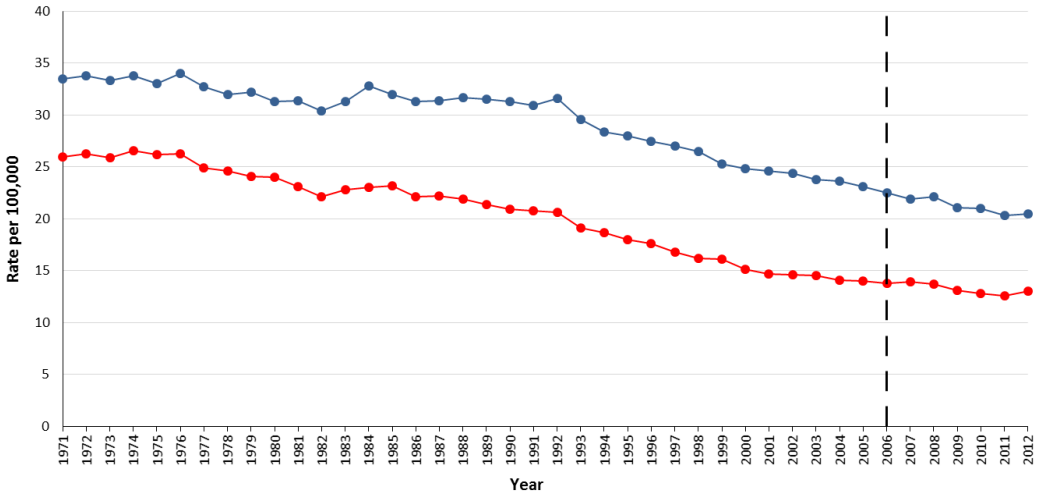
Survival for colorectal cancer is strongly related to the stage of disease at diagnosis; there is a much higher chance of survival associated with earlier stage of disease diagnosis.[8, 10, 11]

Figure 1.1.2 Anatomy of the Large Bowel and Distribution of Colorectal Cancers[8]



Anatomy of the large bowel. Blue and pink circles show the percentage of colorectal cancer cases diagnosed by anatomical subsite for men and women, respectively.

Figure 1.1.3 Colorectal Cancer Mortality Rates in UK between 1971 and 2012[8]



Blue circles represent age standardised mortality rates of colorectal cancer for men. Red circles represent age standardised mortality rates of colorectal cancer for women. The dotted line represents the introduction of the bowel screening programme in England.

The importance of early detection for survival led to the introduction of colorectal screening programmes which have been shown to reduce colorectal cancer incidence

and mortality.[12-14] In the UK, colorectal screening began in 2006, 2007 and 2008 in England, Scotland and Wales respectively.[15-17] All men and women aged 60-74 (50-74 in Scotland) are invited to complete a guaiac faecal occult blood test at home every two years. People with a positive test are then offered colonoscopy to identify any polyps or cancers. However, the uptake of colorectal screening remains disappointingly low with only 55% uptake.[15] Also, colorectal screening is not recommended in less developed countries where resources are lower and the incidence of colorectal cancer is not yet sufficiently high.[18]

Most cases of colorectal cancer develop from a small growth called a polyp or adenoma. It is predicted that the progression from adenoma to cancer takes approximately 10-20 years though not all adenomas will develop into cancers.[19, 20] Hence, colorectal incidence and mortality can be reduced by detecting and treating colorectal polyps at an early stage.[21]

1.1.1 Risk Factors for Colorectal Cancer

The incidence rate of colorectal cancer varies greatly between countries; the highest age-standardised incidence rates are in Australia/New Zealand, Europe and North America and the lowest rates are in Africa and South-Central Asia.[5, 22] The incidence of colorectal cancer is also increasing in many countries, particularly economically transitioning countries.[22, 23] Migrant studies also show that the incidence of colorectal cancer increases rapidly among migrants after moving from areas with low rates of colorectal cancer to areas with high rates.[19, 24]

These data suggest an important role of lifestyle for the risk of colorectal cancer. Over the last few decades numerous epidemiological studies have investigated the relationship between a number of lifestyle factors and colorectal cancer risk and there is currently convincing evidence that alcohol intake, adiposity, smoking, physical inactivity, red meat consumption and processed meat consumption can increase the risk of colorectal cancer.[25-29] Therefore, many cases of colorectal cancer may be prevented by people adopting healthier lifestyles. It was estimated that 54% of colorectal cancers in the UK in 2010 were preventable.[30]

Genetics and hereditary factors are also known to influence an individual's risk of colorectal cancer and it is estimated that inherited factors are important in 15-35% of cases.[31] People with inherited syndromes, such as Lynch syndrome (LS) and familial adenomatous polyposis (FAP), have a particularly high risk of colorectal cancer. The lifetime risk of colorectal cancer is 50-80% for individuals with LS and virtually 100% for those with FAP.[32, 33] However, these syndromes are only estimated to be responsible for approximately 5% of colorectal cancer.[31, 33] Other genetic factors increasing the risk of colorectal cancer are less well understood.

1.2 Alcohol Intake, Adiposity and Smoking

This thesis focused on alcohol intake, adiposity and smoking since they represent three of the most important modifiable lifestyle factors for the reduction of disease burden worldwide.[3] Alcohol consumption is causally related to more than 200 health conditions and in 2012, approximately 3.3 million (6%) of all global deaths were attributable to alcohol consumption.[34] Wide variation in alcohol consumption exists between countries with the highest levels found in Europe, Australia/New Zealand and North America.[34] In 2013, 34% of men and 26% of women in Great Britain reported drinking above the recommended guidelines (four units for men and three units for women) on at least one occasion in the previous week and 19% of men and 12% of women reported binge drinking (more than eight units for men and six units for women) on at least one occasion in the previous week.[35]

The rising prevalence of obesity worldwide has been described as a global obesity pandemic.[36] In the past 30 years obesity has risen in practically all countries.[37] The prevalence of obesity has more than doubled since 1980 and, in 2014, more than 1.9 billion (39%) adults aged 18 or older were overweight or obese and 600 million (13%) were obese.[38] Overweight and obesity were estimated to cause 3.4 million deaths in 2010.[39] In England, 67.1% of men and 57.2% of women aged 16 and over are overweight or obese. Between 1993 and 2013, the prevalence of obesity in England increased from 13.2% in 1993 to 26.0% in men and from 16.4% to 23.8% in women.[40]

Six million people die each year worldwide as a result of tobacco smoking and this is increasing.[41] At least half of lifelong smokers are killed by smoking and, on average,

lifelong smokers lose at least ten years of life.[42] Though the prevalence of smoking has decreased in many high-income countries, it is increasing in many low- and middle-income countries. It is estimated that one billion people will die as a result of smoking in the current century (compared to 100 million in the previous century) with the large majority occurring in low- and middle-income countries unless there is widespread cessation.[43] In 2013, 19% of people in Great Britain were cigarette smokers.[44] This has decreased from 46% in 1974 due to greater awareness of the health risks as well as a wide range of public health interventions. However, smoking remains the leading cause of preventable death in Great Britain.[44]

Thus, alcohol intake, adiposity and smoking represent three of the most important risk factors for reducing the burden of disease worldwide. Furthermore, epidemiological studies have shown that alcohol intake, adiposity and smoking can increase an individual's risk of colorectal cancer. Based on the existing evidence for the associations between these risk factors and colorectal cancer risk, it was predicted that 11.6, 13.0 and 8.1% of colorectal cancer cases diagnosed in the UK in 2010 were attributable to alcohol intake, adiposity and smoking, respectively.[30] However, there are inconsistent results between studies and there remain many important questions about these relationships.

1.3 Alcohol Intake and Colorectal Cancer

Both the World Cancer Research Fund and the International Agency for Research on Cancer concluded that there is convincing evidence from numerous cohort studies that alcohol intake is causally related to colorectal cancer.[25, 26] Both organisations also highlighted a potential threshold with an increased risk only for intake above 30 grams of alcohol (approximately equal to three small glasses of wine or half pints of beer) per day. This conclusion was largely based on a study which pooled results from eight separate cohort studies.[45] However, few other studies have so far been able to analyse the association in similar detail. Thus, more studies are required to investigate the shape of the relationship between alcohol and colorectal cancer risk and whether lower levels of intake are associated with an increased risk.

It is also unclear whether alcohol increases the risk of colorectal cancer equally for both men and women. The World Cancer Research Fund concluded that alcohol was a convincing cause of colorectal cancer only for men.[25] Overall, the evidence that alcohol intake increases the risks of colorectal cancer is clearer for men than for women, but this may simply be a result of studies of men including a much higher proportion of heavy drinkers than studies of women.

Furthermore, many studies of alcohol intake and colorectal cancer use non-drinkers as the reference group i.e. both former drinkers and never drinkers. However, there is much debate about the reference group used for analyses of alcohol intake and the inclusion of former drinkers in the reference group may lead to underestimated risks in current drinkers.[46]

1.4 Adiposity and Colorectal Cancer

Body mass index (BMI), defined as the weight in kilograms divided by the square of height in metres, is the most common measure used to assess adiposity. There is consistent evidence that BMI is more strongly related to colon cancer than rectal cancer.[25] Studies have also found that the association between BMI and colorectal cancer is stronger for men than for women though the reason for this difference is not known.[25]

Waist circumference (WC) and waist to hip ratio (WHR) are also associated with colorectal cancer risk.[25] WC and WHR provide alternative measures of adiposity that focus on abdominal adiposity. There is evidence from one cohort study that suggests WC and WHR are more important measures of adiposity for colorectal cancer risk and thus the different results observed for men and women for the association between BMI and colorectal cancer may be due to the different body fat distributions of men and women though this needs to be confirmed in further analyses.[47]

Many studies have relied on self-reported data to measure adiposity. However, self-reported measures of adiposity are likely to underestimate actual adiposity. For example, people tend to underestimate their weight and overestimate their height,

resulting in underestimated measures of BMI.[48] This may lead to overestimation of the associations with colorectal cancer risk.

1.5 Smoking and Colorectal Cancer

Both former smokers and current smokers have an increased risk of colorectal cancer in comparison to never smokers. A recent meta-analysis found that, compared to never smokers, the pooled relative risk and 95% confidence interval was 1.20 (1.10–1.30) for current smokers and 1.18 (1.12–1.25) for former smokers.[49] It remains unclear exactly why current smokers and former smokers appear to experience a similar level of risk though it may indicate that smoking mainly affects early stages of colorectal cancer and so the effects of smoking persist for many years, even after cessation.

Studies have also analysed more detailed measures of smoking exposure such as smoking intensity and smoking duration.[26, 29, 50] Most studies found that these measures are positively associated with colorectal cancer. However, most studies combine former and current smokers when analysing intensity or duration which may lead to misleading results and few studies have investigated the effects of intensity and duration together.

It is also important to investigate how the risk of colorectal cancer differs according to smoking cessation at different times. Although former smokers have a similar level of risk to current smokers, studies have generally found that the risk of colorectal cancer decreases with earlier cessation.[26, 29] Furthermore, it remains unclear whether the reduced risk is due to earlier cessation or shorter duration of smoking.

1.6 Limitations of Existing Studies and the Need for Further Research

In summary, while it is known that alcohol intake, adiposity and smoking are related to colorectal cancer risk, there are a number of important questions that remain about these associations and many existing studies have a number of limitations. For example, alcohol intake, adiposity and smoking represent three complex risk factors, yet many studies include only very basic information about these risk factors. Studies also often include limited information on other important risk factors. Consideration of confounders is essential in observational studies since many aspects of lifestyle are

correlated in the general population. The quality of data in existing studies is also very variable and there is a need for more high quality studies with detailed information on a wide range of risk factors.
It may also be important to consider how associations are modified by exposure to other risk factors. Different levels of exposure to effect modifiers could contribute to differences in results between different studies. For example, the association between alcohol intake and colorectal cancer risk may be modified by levels of BMI.[45] However, in general, there is a lack of evidence for effect modifiers from existing studies.

Furthermore, though generally analysed as a single disease, colorectal cancer actually represents a heterogeneous group of diseases[51, 52] and studies have found different associations according to different colorectal subsites. Therefore, more research is needed based on large cohort studies that are able to investigate the subsites separately.

Thus, there is a clear need to investigate the associations between alcohol intake, adiposity and smoking and colorectal cancer in greater detail. A greater understanding of these associations may provide insight into why some studies find contrasting results and lead to a greater understanding of the mechanisms linking these risk factors to colorectal cancer and thus to greater opportunities for prevention. To achieve this, it is necessary to employ cohort studies with detailed information on a wide range of risk factors on a large number of people.

1.7 UK Biobank

Data from UK Biobank were used in this thesis to investigate the associations between alcohol intake, adiposity and smoking and colorectal cancer. UK Biobank is a prospective cohort study of half a million men and women mostly aged 40-69 recruited from the general population in England, Scotland and Wales.[53] UK Biobank was established by the Medical Research Council and the Wellcome Trust with the aim to provide insights into the complex interplay of lifestyle, environment and genetics in promoting a wide range of diseases. Men and women were invited to an assessment centre where they gave very detailed information on health and lifestyle factors, completed physical measures and gave biological samples. Recruitment of participants finished in 2010. Follow-up data on participants for outcomes was ascertained through

linkage with national datasets.[53] Data from UK Biobank is available to all bona fide researchers undertaking health-related research that is in the public good. The large size of UK Biobank, combined with the detailed information on a wide range of health and lifestyle factors, makes it ideal to contribute to a better understanding of how alcohol intake, adiposity and smoking are associated with colorectal cancer risk. Further information on the UK Biobank cohort is provided in Chapter 3.

1.8 Summary

In summary, the burden of cancer is rising dramatically worldwide, representing a major challenge for health services. The prevention of disease is vital in efforts to alleviate the burden of cancer. Colorectal cancer is the third most commonly diagnosed cancer in males and the second in females worldwide.[5] The risk of colorectal cancer is strongly associated with lifestyle factors, including alcohol intake, adiposity and smoking, three classic causes of disease. However, many important questions remain about the relationships between these risk factors and colorectal cancer.

The aim of this thesis is to investigate the relationships between alcohol intake, adiposity and smoking and colorectal cancer risk in detail using data from the UK Biobank cohort in order to contribute to a greater understanding of how these risk factors contribute to the risk of colorectal cancer.

The first objective of this thesis is to conduct detailed literature reviews in order to identify and summarise the existing evidence for the relationships between alcohol intake, adiposity and smoking and colorectal cancer. Following these literature reviews, more specific objectives will be described for this thesis.

These literature reviews and thesis objectives are included in Chapter 2. Chapter 3 provides further information on the UK Biobank cohort. The methods used in the analyses of this thesis are described in Chapter 4. The main results for the associations between alcohol intake, adiposity and smoking and colorectal cancer are presented in Chapters 5, 6 and 7. Chapter 8 details the final discussion of this thesis.

Chapter 2 Literature Reviews

The aim of this chapter is to review the existing literature and summarise the evidence for the effects of alcohol intake, adiposity and smoking in relation to colorectal cancer risk. Three separate literature reviews were conducted, one for each risk factor. The general methodology used to complete the three literature reviews was similar across the reviews and is described below.

2.1 Overall Search Strategy

Alcohol intake, adiposity and smoking represent three of the most important modifiable lifestyle factors for long-term health and disease prevention[3, 39] and have been extensively investigated in relation to a wide range of diseases including colorectal cancer. This means that there exist a large number of studies investigating these relationships and numerous reviews and meta-analyses have previously attempted to identify the relevant literature for each risk factor. These reviews and meta-analyses were used as a starting point to identify the relevant literature for these reviews.

Literature searches were also conducted for each risk factor using PubMed in order to identify the most recent literature not included in previous reviews or meta-analyses, using similar search criteria to the previous reviews and meta-analyses. The literature searches overlapped in time with the existing reviews/meta-analyses to ensure that no articles were missed (and to verify that the literature searches identified similar articles). References of identified articles were carefully searched for further relevant articles. The specific search strategy used for each risk factor is described within each literature review.

These literature reviews focused on evidence from prospective cohort studies. Prospective cohort studies are generally considered to be of greater quality than retrospective case-control studies since case-control studies may be more prone to selection bias and are unable to preclude potential recall bias.[54, 55]

2.2 Alcohol Intake and Colorectal Cancer Literature Review

In this section, the existing evidence for an association between alcohol intake and colorectal cancer risk is reviewed. The relationship between alcohol intake and colorectal cancer has been investigated in numerous prospective cohort studies and the World Cancer Research Fund (WCRF) and the International Agency for Research on Cancer (IARC) have both concluded that alcohol intake is causally related to colorectal cancer risk.[26, 56] Questionnaires are usually employed to assess individuals' average intake of different alcoholic beverages. The most common measurement of alcohol intake in the literature is grams of alcohol. In the UK, the alcoholic content of beverages is usually measured in units with one unit being approximately equal to eight grams of alcohol.[57, 58]

2.2.1 Search Strategy

The WCRF published reports in 2007 and 2011 and the IARC published reports in 2010 and 2012 summarising the evidence from prospective cohort studies for an association between alcohol intake and colorectal cancer.[25, 26, 56, 59] The most recent report from the WCRF searched the Medline database for articles published up to May 2010.[56] Details on the literature search were not provided in the IARC reports though the most recent report included articles published in 2009.[26] Recent meta-analyses[60-63] and review articles[64] were also searched for relevant articles.

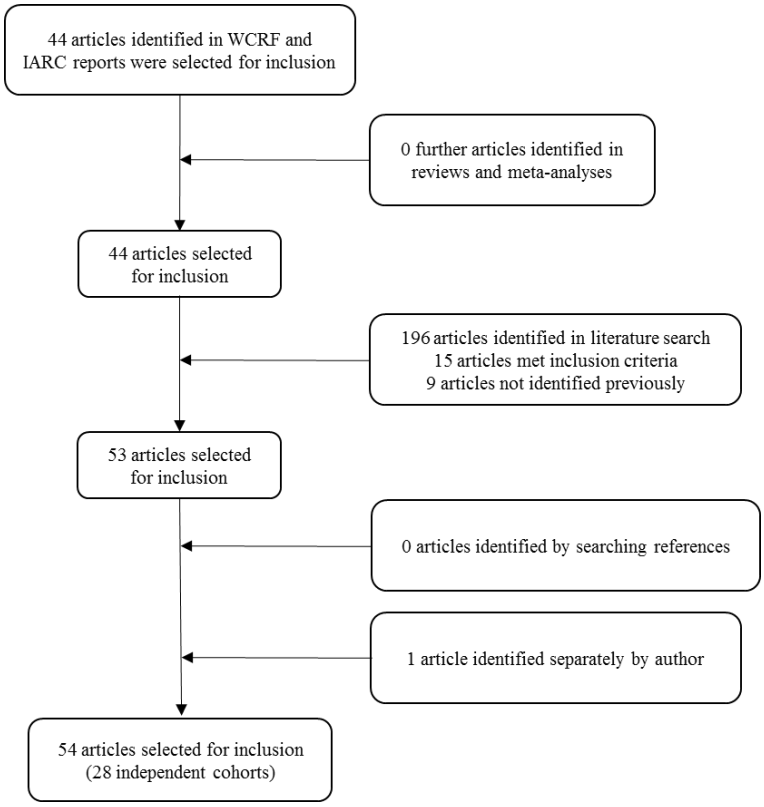
To complement the articles identified by these reports and meta-analyses, PubMed was searched for articles reporting human studies, written in English, and published between 1st January 2008 and 31st March 2015, using the MeSH terms "Colorectal Neoplasms" and at least one of "Ethanol", "Alcohol Drinking" or "Alcoholic Beverages". 196 articles were retrieved.

Articles were selected for inclusion in this review if they (i) were based on prospective cohort studies (including nested case-control studies), (ii) analysed total alcohol intake in relation to colorectal cancer incidence (or colorectal subsites) in the general population and (iii) reported risk estimates (hazard ratios (HR), relative risks (RR), odds ratios (OR)) and confidence intervals (CI). Analyses that reported results only by other

factors were not included. Analyses based on fewer than 30 cases were excluded. References of included articles were carefully examined for additional articles.

Figure 2.2.1 shows a flow diagram of the literature search process, illustrating the numbers of articles identified from the different sources. Overall, 54 articles, based on 28 independent cohorts, met the inclusion criteria for this literature review.

Figure 2.2.1 Flowchart of Literature Search Process for Alcohol Intake and Colorectal Cancer



WCRF = World Cancer Research Fund, IARC = International Agency for Research and Cancer

Table 2.2.1, presented at the end of this review, provides more details on the articles (identified in the literature search) mentioned in this review. Where necessary, more than one article from the same cohort is included in Table 2.2.1. At least one article from each cohort identified in the literature search is included (normally the article with the largest number of cases). Results are included for the highest vs lowest exposure comparison and separate results are included for colorectal cancer, colon cancer and rectal cancer where possible. Articles are ordered by publication year.

2.2.2 Alcohol Intake

The evidence is fairly consistent that alcohol intake is associated with an elevated risk of colorectal cancer, particularly alcohol intake above 30 g/day. A potential threshold effect was first suggested by a pooled analysis of data from eight cohort studies in five countries in North America and Europe, including 490,000 men and women and 4,700 cases of colorectal cancer. This analysis by Cho et al. found that an increased risk of colorectal cancer was restricted to people drinking at least 30 g/d.[45] Compared with non-drinkers, the RRs and 95% CIs for people drinking >0-<5, 5-<15, 15-<30, 30-<45 and ≥45 g/d were 0.94 (0.86-1.03), 0.97 (0.88-1.06), 1.01 (0.86-1.18), 1.16 (0.99-1.36) and 1.41 (1.16-1.72).

Analyses of the individual cohorts (with longer follow-up) also provided support for a threshold at 30 g/d. The Netherlands Cohort Study and the Health Professionals Follow-up Study both found evidence of an increased risk of colorectal cancer only for people drinking above 30 g/d.[65, 66] Two studies of women (the Iowa Women's Health Study and the Canadian National Breast Screening Study) did not find evidence for an increased risk of colorectal cancer for women drinking ≥30 g/d[67, 68] but there was evidence of an increased risk in the Nurses' Health Study of women.[66]

Data from two large cohorts also suggested that there is a threshold for the association between alcohol intake and colorectal cancer. The European Prospective Investigation into Cancer and Nutrition (EPIC), which included almost 500,000 men and women and 1,833 cases of colorectal cancer, found an increased risk only above 30 g/d; the HRs and 95% CIs comparing 4.9-14.9, 15-29.9, 30-59.9 and ≥60 g/d with 0.1-4.9 g/d were 1.05 (0.92-1.19), 1.03 (0.88-1.20), 1.26 (1.06-1.49) and 1.64 (1.29-2.08).[69] The Million Women Study included over 4,000 cases of colon cancer and over 2,000 cases of rectal cancer. Women mainly reported low to moderate alcohol intake. No association was observed between alcohol intake and colon cancer but drinking ≥15 drinks per week (approximately ≥20 g/d) was associated with an increased risk of rectal cancer (HR, 1.25; 95% CI, 1.06-1.49).[70]

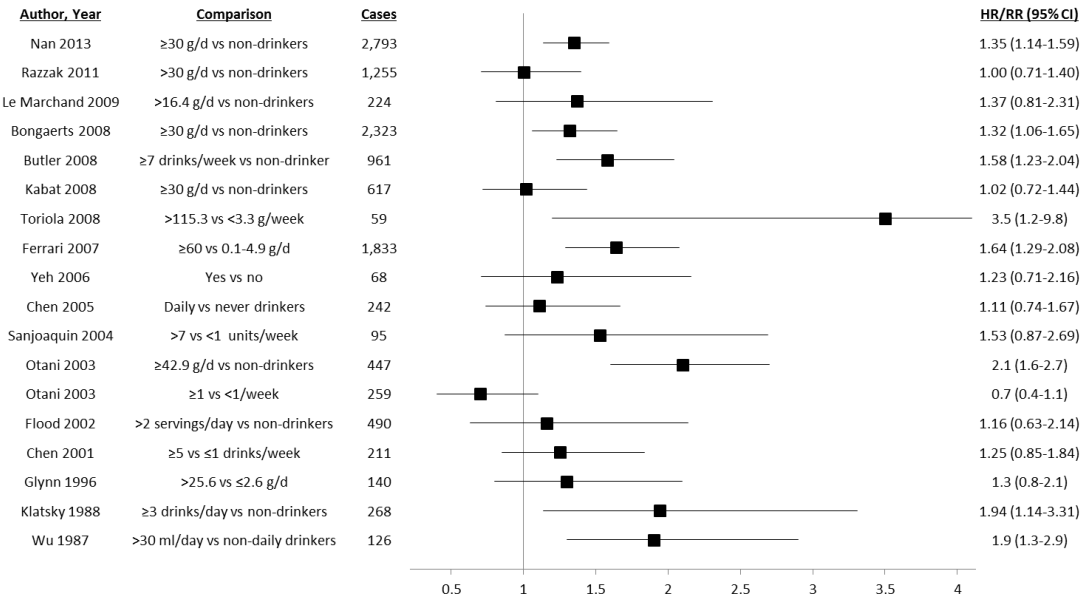
Mizoue et al. also conducted a pooled analysis, this time of five cohort studies in Japan, including 1,724 cases for men.[71] Results from this analysis did not seem to suggest a

threshold for the effects of alcohol intake. Compared with non-drinkers, the RRs and 95% CIs for men drinking 0.1-<23, 23-<46, 46-<69, 69-<92 and ≥92 g/d were 1.22 (0.92-1.61), 1.42 (1.21-1.66), 1.95 (1.53-2.49), 2.15 (1.74-2.64) and 2.96 (2.27-3.86). There seemed to be a particularly strong association between alcohol intake and colorectal cancer for Japanese men compared to other studies in Western populations although this study analysed a much wider range of alcohol intake than other studies. However, re-analysing the data using the same categories as the Western pooled analysis by Cho et al.,[45] alcohol intake still appeared to be more strongly associated with colorectal cancer among Japanese men compared to Western men.[71]

In general, other studies have not been able to investigate the relationship between alcohol intake and colorectal cancer in such a precise manner. A number of studies still found evidence for an increased risk of colorectal cancer with increased alcohol intake.[72-76] The Singapore Chinese Health Study found that drinking at least seven drinks/week increased the risk of colorectal cancer compared with non-drinkers (HR, 1.58; 95% CI, 1.23-2.04).[72] Results from a study of Danish men and women by Pedersen et al. indicated a dose-response association between alcohol and rectal cancer but no association with colon cancer.[73] One study of Japanese American men in Hawaii found that drinking ≥22.35 g/d increased the risk of colon cancer (HR, 1.39; 95% CI, 1.05-1.83) and rectal cancer (HR, 2.30; 95% CI, 1.43-4.69) compared with non-drinkers.[74]

Overall, the evidence seems consistent that a high intake of alcohol (particularly above 30 g/d) increases a person's risk of colorectal cancer. It is less clear whether more moderate alcohol intake increases the risk. Figure 2.2.2 shows the pattern of results for the association between alcohol intake and colorectal cancer from different prospective cohort studies.

Figure 2.2.2 Results from Prospective Cohort Studies for the Association between Alcohol Intake and Colorectal Cancer



Results show the comparison between the highest intake category and the reference group from prospective cohort studies. Squares represent HRs/RRs. Bars represent 95% CIs.

Sex

There is some uncertainty about whether the effects of alcohol intake on colorectal cancer risk are consistent for men and women. Results from meta-analyses[60, 62] have shown stronger associations between alcohol intake and colorectal cancer for men than for women and the WCRF stated that alcohol is a cause of colorectal cancer for men but only a probable cause for women.[56] However, rather than a true difference in the association for men and women, this could simply reflect the fact that on average men are heavier drinkers than women and so the stronger associations may be due to higher levels of intake being analysed among men. For example, the highest category of intake in the Japanese pooled analysis of men was ≥92 g/d. Mizoue et al. also performed a pooled analysis for Japanese women and the highest category of intake was ≥23 g/d.[71] However, these differences are not reflected in meta-analyses which simply pool highest vs lowest results. A further possibility is that there are different patterns of confounding for men and women or perhaps even a difference in the accuracy of reporting.

In contrast to the stronger effect for men suggested by meta-analyses, no individual studies have found evidence for a different effect for men and women.[45, 65, 66, 69, 73, 77]

These studies have the advantage (compared with meta-analyses) of being able to take into account alcohol consumption more precisely when comparing results for men and women. For example, the RRs (95% CIs) from the Western pooled analysis comparing ≥45 g/d with non-drinkers were 1.41 (1.11-1.79) for men and 1.41 (0.98-2.02) for women.[45] Comparing ≥30 g/d with non-drinkers in the Netherlands Cohort Study, the HRs (95% CIs) for colorectal cancer risk were 1.61 (1.17-2.20) for men and 1.82 (1.06-3.11) for women.[65] The risk of colorectal cancer was also similar for men (HR, 1.38; 95% CI, 1.00-1.69 for ≥30 g/d vs non-drinkers) and women (HR, 1.38; 95% CI, 1.11-1.72 for ≥30 g/d vs non-drinkers) in the Health Professionals Follow-up Study and the Nurses' Health Study.[66] The risk of colon and rectal cancer associated with a 15 g/day increase was similar for men and women in the EPIC study.[69] Thus, there is no strong evidence that the effects of alcohol intake on colorectal cancer risk are different for men and women.

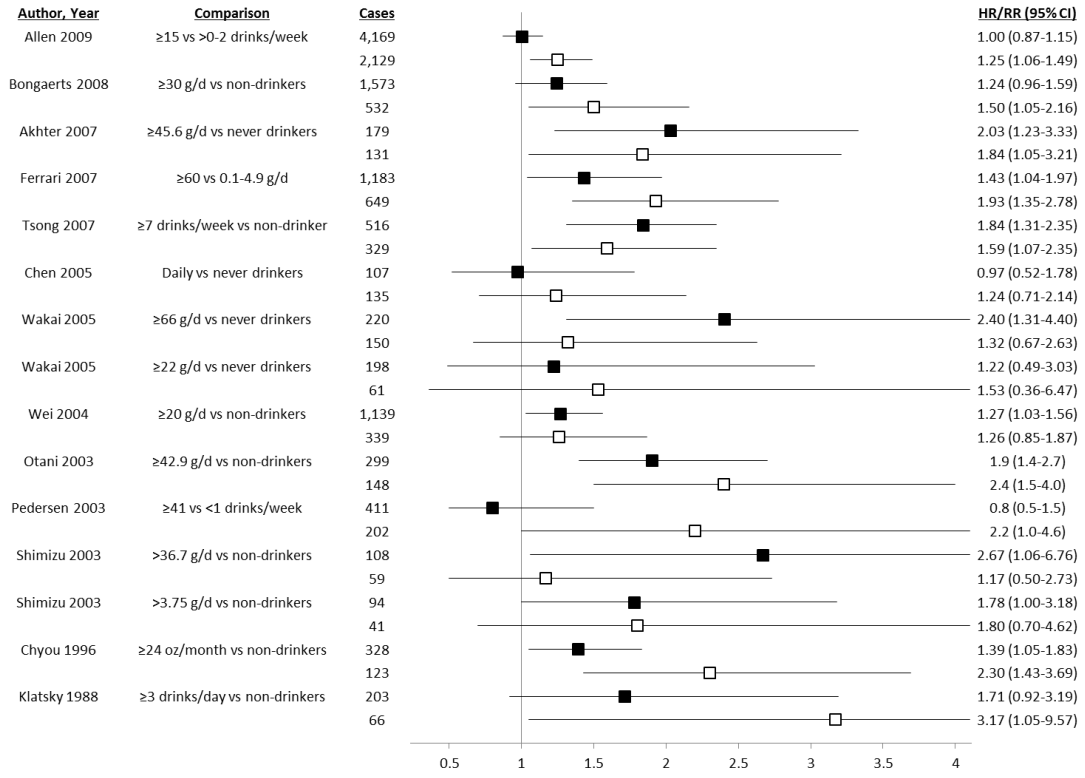
Colorectal Subsites

It seems clear that alcohol intake increases the risk of both colon and rectal cancer though the association may be slightly stronger for rectal cancer. One meta-analysis which pooled results for the highest vs. lowest categories of intake from cohort studies found RRs (95% CIs) of 1.50 (1.25-1.79) for colon cancer and 1.63 (1.35-1.97) for rectal cancer.[60] Another meta-analysis pooled results from case-control and cohort studies of both incidence and mortality using a "dose-response" approach.[62] The RRs for colon and rectal cancer were 1.15 (1.06-1.24) and 1.23 (1.13-1.35) for moderate drinking (12.6-49.9 g/day) and 1.43 (1.23-1.67) and 1.59 (1.18-2.15) for heavy drinking (≥50 g/day).

The majority of studies (including the EPIC study and the Netherlands Cohort Study) showed stronger associations for rectal cancer than colon cancer.[65, 69, 74, 75, 78, 79] Also, the Million Women Study and the Danish study by Pedersen et al. found that alcohol intake was associated only with rectal cancer and not with colon cancer.[70, 73] However, the results were similar for colon and rectal cancer in the Western pooled analysis[45] and among the men of the Japanese pooled analysis (results were stronger for rectal cancer for Japanese women).[71] The Singapore Chinese Health Study found a stronger association for colon cancer[77] and the Health Professionals Follow-up Study found an

association only for colon cancer.[80] Figure 2.2.3 shows the association between alcohol and colon and rectal cancer from different cohorts. Only cohorts which presented results for both colon and rectal cancer are included.

Figure 2.2.3 Results from Prospective Cohort Studies for the Association between Alcohol Intake and Colon Cancer and Rectal Cancer



Results show the comparison between the highest intake category and the reference group from prospective cohort studies which presented results for colon cancer and rectal cancer. Squares represent HRs/RRs. Bars represent 95% CIs. Black squares represent results for colon cancer and white squares represent results for rectal cancer.

The proximal and distal colon differ in terms of anatomy, incidence rates and embryonic origins[51, 81] and thus the effects of alcohol intake on cancer risk may differ for the proximal and distal colon. All studies seem to agree that there is a stronger association between alcohol intake and distal colon cancer than between alcohol intake and proximal colon cancer[45, 65, 69, 80, 82] except for Pedersen et al. which found null results for colon cancer overall as well as for proximal and distal colon cancer.[73] In the Western pooled analysis the results for proximal and distal colon cancer were 1.35 (0.97-1.89) and 1.66 (1.17-2.36) comparing ≥45 g/d with non-drinkers.[45] The results were 0.92 (0.51-1.66) and 1.68 (1.08-2.62) for >60 vs 0.1-4.9 g/d in the EPIC study.[69] Meta-

analysis results for heavy drinking (≥50 g/day) were 1.38 (0.96-1.98) for proximal colon cancer and 2.46 (1.38-4.40) for distal colon cancer.[62]

Beverage Types

The increased risk of colorectal cancer due to alcohol intake does not appear to differ by beverage type. Different studies have employed slightly different methods when investigating whether certain beverages are more strongly related to colorectal cancer though there seems to be general agreement amongst studies that the increased risk is due to overall alcohol intake rather than specific beverages.[45, 65, 69, 70, 73, 77, 80]

Summary

Alcohol intake is positively associated with colorectal cancer risk. It is possible that there exists a threshold for increased risk at 30 g/d. Alcohol intake seems to have a similar effect on cancer risk for both men and women. Alcohol intake increases the risk of both colon and rectal cancer though alcohol intake seems to have a slightly stronger effect on rectal cancer. The effect seems to be stronger for distal colon cancer than proximal colon cancer.

2.2.3 Effect Modifiers

It is important to consider if there are any important effect modifiers such that certain groups of people are at a particularly elevated risk due to alcohol intake. Investigating effect modifiers may explain why studies find different results and may contribute to the understanding of underlying mechanisms. A number of studies have investigated different factors as possible modifiers of the relationship between alcohol intake and colorectal cancer.

Alcohol Intake and BMI

Analysing the effect of alcohol intake on colorectal cancer risk by BMI, the Western pooled analysis found evidence for a stronger effect of alcohol amongst leaner people; the HRs (95% CIs) comparing ≥30 g/d with non-drinkers were 1.84 (1.27-2.67), 1.23

(0.91-1.65) and 1.08 (0.88-2.33) for people with BMI <22, 22-<25 and ≥25 kg/m^2.[45] The same pattern was found for men in the Japanese pooled study using the same BMI categories; the results for ≥69 g/d vs non-drinkers were 3.25 (2.12-4.99), 2.12 (1.57-2.87) and 1.83 (1.26-2.67) for men with BMI <22, 22-<25 and ≥25 kg/m^2.[71]

These two large pooled analyses of subjects in different populations provide fairly strong evidence for an effect modification of BMI on the relationship between alcohol intake and colorectal cancer yet there is a dearth of evidence from other studies. The Norfolk arm of the EPIC cohort found no association between alcohol intake and colorectal cancer for people with BMI <25, 25-30 or >30 kg/m^2 however the analysis included few cases and only analysed drinking ≥8 g/d compared with non-drinkers.[83] One case-control study actually found the opposite effect i.e. alcohol intake was associated with colorectal cancer among people with BMI ≥30 kg/m^2 but not among people with BMI ≤30 kg/m^2.[84]

Folate

Folate is a B vitamin found in foods such as leafy green vegetables, asparagus, broccoli and liver. Folate intake is thought to be protective against colorectal cancer risk[85-87] and alcohol is known to inhibit the bioavailability and metabolism of folate.[88, 89] Thus, it is hypothesised that alcohol intake is more strongly related to colorectal cancer among people with low folate intake.

However, the evidence for an association between folate intake and colorectal cancer risk from cohort studies is not clear. While some studies indicate an inverse association between folate intake and colorectal cancer risk,[72, 90-94] other studies do not find clear evidence for an association.[68, 95-101] One reason contributing to the conflicting results is that different study populations will have different distributions of folate intake and different studies have considered very different comparisons of folate intake. Another reason for the conflicting results could be measurement error. Accurately measuring folate intake using questionnaires presents a real challenge and error in measuring folate intake will bias associations towards the null.[102] Given the difficulties in accurately measuring folate intake using questionnaires, measures of circulating folate may

provide a more accurate reflection of average folate intake however studies of circulating folate also provide conflicting results.[103-106]

While evidence from cohort studies may indicate an inverse association between folate intake and colorectal cancer, randomised controlled trials (RCTs) of folic acid have generally failed to show an effect on colorectal adenoma recurrence[107-111] or incidence.[112] RCTs are often considered to be the gold standard of epidemiological studies for evaluating the effectiveness of a treatment since randomisation should minimise differences between the treatment groups.[113] Associations between other nutrients or dietary factors and cancer observed in observational studies have also not been confirmed in RCTs.[114-116] This difference may be because a collection of dietary (and possibly lifestyle) factors might work together to prevent cancer but it is very difficult to disentangle the separate effects in cohort studies. However, there are other differences that could possibly explain the difference. For example, the RCTs considered colorectal adenoma as the outcome compared to colorectal cancer in the cohort studies and most of the RCTs were conducted among people with a history of adenoma. Also, the RCTs only considered relatively short periods of treatment and follow-up; cancer has a long latency period and folate intake in cohort studies may more accurately reflect long-term intake.

Alcohol Intake and Folate

A number of cohort studies of alcohol intake and colorectal cancer have investigated a possible interaction by folate intake though results are conflicting. Neither of the two large pooled analyses found important differences in the association between alcohol intake and colorectal cancer according to tertiles of folate intake.[45, 71] The Netherlands Cohort Study also found no evidence for an interaction between alcohol intake and folate intake.[65] In contrast, the EPIC study found that the effect of alcohol intake on colorectal cancer risk was modified by folate intake; the HRs and 95% CIs for an increase of 15 g/day of alcohol for low, middle and high tertiles of folate intake were 1.13 (1.06-1.20), 1.09 (1.03-1.15) and 1.03 (0.98-1.09).[69]

Analyses of folate intake and colorectal cancer have also investigated whether the association varies by alcohol intake and also find conflicting results. Though some

studies have found a stronger inverse association between folate intake and colorectal cancer among people with higher alcohol intake,[91, 93] other studies do not support such an interaction.[68, 90, 96-98, 117-119]

Alcohol Intake and Smoking

Few cohort studies have investigated the possible interaction between alcohol intake and smoking status though results seem to suggest that the association between alcohol intake and colorectal cancer may be stronger for people with greater smoking exposure. Comparing an intake of ≥30 g/d with non-drinkers for never, former and current smokers, the results from the Western pooled analysis were 1.17 (0.84-1.63), 1.26 (1.00-1.58) and 1.42 (1.11-1.83).[45] Similarly, in the EPIC study, an increase of 15 g/d was associated with a greater risk among current smokers (HR, 1.23; 95% CI, 1.12-1.36) than among never (HR, 1.15; 95% CI, 1.03-1.28) or former smokers (HR, 1.11; 95% CI, 0.97-1.28).[69] The Singapore Chinese Health Study suggested a stronger effect of alcohol intake for ever smokers than for never smokers for colon cancer but not for rectal cancer.[77] Two smaller Japanese studies did not find strong evidence for an interaction.[82, 120]

Summary

Two large pooled analyses in different populations support a stronger association between alcohol intake and colorectal cancer risk for people with lower BMI. However, there is an absence of evidence from other studies and one case-control study found evidence for an opposite interaction between alcohol intake and BMI. An interaction between alcohol intake and folate intake has been investigated in a number of studies but the results are conflicting and equivocal. Few studies have investigated an interaction between alcohol intake and smoking for the risk of colorectal cancer though the evidence so far seems to suggest a synergistic effect of alcohol intake and smoking.

2.2.4 Reference Groups

The choice of reference group is an important consideration in analyses of alcohol intake. There are a number of advantages and disadvantages for different choices of

reference group (see section 4.1.2 for a more detailed discussion on the choice of reference groups) and different choices could lead to very different results. Many studies of alcohol intake and colorectal cancer, including the two pooled analyses,[45, 71] have used non-drinkers as the reference group (i.e. never drinkers and former drinkers combined). Using non-drinkers as the reference group could possibly underestimate the association between alcohol intake and colorectal cancer (compared to results using never drinkers) since former drinkers may have an increased risk of colorectal cancer compared to never drinkers. Few studies have analysed alcohol intake and colorectal cancer using never drinkers as the reference group, perhaps because it was not possible to separately identify former and never drinkers or because never drinkers represented a small group. Other studies, including EPIC and the Million Women Study have used some definition of light drinkers as the reference group.[69, 70]

Very few studies have compared results using different reference groups. The two pooled analyses did include sensitivity analyses to assess the effect of including former drinkers in the reference group (though data from the Western pooled analysis was very limited).[45, 71] Four of the cohorts in the Western pooled analysis had data on alcohol intake five to ten years before baseline. Restricting the analysis to these cohorts, the results were similar when past drinkers (in the last five or ten years) were included (RR, 1.56; 95% CI, 1.20-2.04 for ≥45 g/d vs non-drinkers) and excluded (RR, 1.65; 95% CI, 1.24-2.21 for ≥45 g/d vs never drinkers) from the reference group.[45] The Japanese pooled analysis also found similar results when former drinkers were excluded from the reference group.[71]

2.2.5 Mechanisms

Colorectal cancer is a complex disease and it remains unclear exactly how alcohol intake increases the risk of colorectal cancer. It seems likely that acetaldehyde, the first metabolite of ethanol and a known carcinogen in animals, plays a key role in colorectal carcinogenesis.[26, 121] Alcohol is absorbed into the blood from the stomach and small intestine and is circulated around the body. Alcohol is primarily metabolised in a two-step process. First alcohol dehydrogenase (ADH) enzymes oxidise ethanol to acetaldehyde which is then oxidised to acetate by aldehyde dehydrogenase (ALDH) enzymes.

These dehydrogenase enzymes exist in different variants with different levels of activity.[122-124] For example, the *ALDH2*2* allele, relatively common in East Asian populations but essentially absent in Europeans, is virtually inactive. This causes acetaldehyde to accumulate in the body which results in facial flushing and other adverse effects. These effects are particularly severe for homozygous carriers who generally become abstainers whereas heterozygous carriers may tolerate alcohol consumption.

Hence, polymorphisms of these enzymes should be associated with different levels of colorectal cancer risk since these polymorphisms will lead to different levels of acetaldehyde as a result of alcohol intake. However, since the different polymorphisms may be predictive of alcohol use, it is important that the actual alcohol intake is also considered at the same time. Studies of alcohol intake and colorectal cancer investigating dehydrogenase enzymes do not provide strong evidence that the association between alcohol intake and colorectal cancer is modified by any of these polymorphisms.[125-137] Most of these studies included a relatively small number of cases and the confounding adjustment was quite poor in a number of studies.

2.2.6 Summary

Numerous studies have shown that high alcohol intake is associated with an increased risk of colorectal cancer and both the WCRF and the IARC declared in their most recent reports that alcohol intake is a cause of colorectal cancer.[25, 26, 56, 59] Alcohol intake is associated with both colon and rectal cancer. The association may be slightly stronger for rectal cancer and seems to be stronger for distal colon cancer than proximal colon cancer (though few studies have investigated colon subsites). The largest studies indicate that an increased risk of colorectal cancer may be limited to people drinking above 30 g/d. It is less clear whether lower levels of alcohol intake increase colorectal cancer risk. The association appears to be similar for men and women. The risk does not seem to differ by beverage type and whether the risk differs according to other factors is unclear.

Table 2.2.1 Prospective Cohort Studies of Alcohol Intake and Colorectal Cancer

Ref.	Author, year	Study name, location	Population	Follow-up period (average follow-up)	Age	Sex	Number of subjects	Exposure	Site	Number of cases	RR (95% CI)
66	Nan 2013	Nurses' Health Study Health Professionals Follow-up Study, USA	Nurses Health professionals	1980 - 2008 (28 yrs) 1986 - 2008 (22 yrs)	34-59 40-75	B	134,730	≥30 g/d vs non-drinkers	CRC	2,793	1.35 (1.14-1.59)
67	Razzak 2011	Iowa Women's Health Study, USA	General population	1986 - 2002 (17 yrs)	55-69	F	38,001	>30 g/d vs non-drinkers	CRC	1,255	1.00 (0.71-1.40)
138	Shin 2011	Korea	General population	1996 - 2003	30-80	M	869,725	>1 bottle of Korean distilled spirits vs non-drinkers	PCC	536	1.2 (0.9-1.6)
									DCC	751	1.5 (1.2-2.0)
									RC	1,535	1.2 (1.0-1.4)
						F	395,501	>1 bottle of Korean distilled spirits vs non-drinkers	PCC	236	1.6 (0.7-3.5)
									DCC	225	0.5 (0.1-1.9)
									RC	551	1.7 (1.0-2.8)

Ref.	Author, year	Study name, location	Population	Follow-up period (average follow-up)	Age	Sex	Number of subjects	Exposure	Site	Number of cases	RR (95% CI)
139	Park 2010	EPIC-Norfolk EPIC-Oxford Guernsey Study Oxford Vegetarian Study MRC National Survey of Health and Development UK Women's Cohort Study Whitehall II, UK	General population		40-77 32-84 39-78 26-79 43 44-78 41-62	B	1,734*	≥30 vs >0-<5 g/d	CRC CC RC	458 308 150	1.30 (0.86-1.95) 1.47 (0.89-2.43) 1.01 (0.48-2.11)
70	Allen 2009	Million Women Study, UK	General population	1996 - 2006 (7.2 yrs)	50-64	F	1,280,296	≥15 vs >0-2 drinks/week	CC RC	4,169 2,129	1.00 (0.87-1.15) 1.25 (1.06-1.49)
140	Le Marchand 2009	Multiethnic Cohort Study, USA	General population	2001 - 2006	53-88	B	411*	>16.4 g/d vs non-drinkers	CRC	224	1.37 (0.81-2.31)
83	Park 2009	European Prospective Investigation into Cancer and Nutrition-Norfolk, UK	General population	1993 - 2006 (11 yrs)	40-79	B	24,244	≥21 units/week vs non-drinkers	CRC CC RC	386 256 122	0.70 (0.44-1.13) 0.59 (0.32-1.09) 0.94 (0.43-2.09)

Ref.	Author, year	Study name, location	Population	Follow-up period (average follow-up)	Age	Sex	Number of subjects	Exposure	Site	Number of cases	RR (95% CI)
65	Bongaerts 2008	Netherlands Cohort Study, Netherlands	General population	1986 - 1999 (13.3 yrs)	55-69	B	4,118†	≥30 g/d vs non-drinkers	CRC	2,323	1.32 (1.06-1.65)
									CC	1,573	1.24 (0.96-1.59)
									RC	532	1.50 (1.05-2.16)
72	Butler 2008	Singapore Chinese Health Study, Singapore	General population	1993 - 2005 (9.8 yrs)	45-74	B	61,321	≥7 drinks/week vs non-drinker	CRC	961	1.58 (1.23-2.04)
68	Kabat 2008	Canadian National Breast Screening Study, Canada	General population	1980 - 2000 (16.4 yrs)	40-59	F	49,654	≥30 g/d vs non-drinkers	CRC	617	1.02 (0.72-1.44)
71	Mizoue 2008	Japan Public Health Center-based Prospective Study I	General population	1990 - 2004	40-59	M	98,265	≥92 g/d vs non-drinkers	CRC	1,724	2.96 (2.27-3.86)
				1993 - 2004	40-69				CC	1,093	3.44 (2.50-4.72)
				1988 - 2001	40-79				RC	629	2.10 (1.16-3.83)
		Japan Public Health Center-based Prospective Study II		1990 - 2001	40-64	F	111,498	≥23 g/d vs non-drinkers	CRC	1,078	1.57 (1.11-2.21)
				1992 - 1999	≥35				CC	736	1.66 (1.12-2.46)
									RC	338	2.39 (1.18-4.88)
		Japan Collaborative Cohort Study									
		Miyagi Cohort Study									
		Takayama Study, Japan									

Ref.	Author, year	Study name, location	Population	Follow-up period (average follow-up)	Age	Sex	Number of subjects	Exposure	Site	Number of cases	RR (95% CI)
80	Thygesen 2008	Health Professionals Follow-up Study, USA	Health professionals	1986 - 2002 (16 yrs)	40-75	M	47,432	>45 g/d vs non-drinkers	CRC	868	1.75 (1.21-2.52)
141	Toriola 2008	Findrink Study, Finland	General population	1984 - 2005 (16.7 yrs)	42-60	M	2,627	>115.3 vs <3.3 g/week	CRC	59	3.5 (1.2-9.8)
82	Akhter 2007	Miyagi Cohort Study, Japan	General population	1990 - 2001 (11 yrs)	40-64	B	21,199	≥45.6 g/d vs never drinkers	CRC	307	1.91 (1.32-2.78)
									CC	179	2.03 (1.23-3.33)
									RC	131	1.84 (1.05-3.21)
69	Ferrari 2007	European Prospective Investigation into Cancer and Nutrition, Europe	General population	1992 - 2002 (6.2 yrs)	35-70	B	478,732	≥60 vs 0.1-4.9 g/d	CRC	1,833	1.64 (1.29-2.08)
									CC	1,183	1.43 (1.04-1.97)
									RC	649	1.93 (1.35-2.78)
77	Tsong 2007	Singapore Chinese Health Study, Singapore	General population	1993 - 2004 (8.9 yrs)	45-74	B	61,321	≥7 drinks/week vs non-drinker	CRC	845	1.84 (1.31-2.58)
									CC	516	1.84 (1.31-2.35)
									RC	329	1.59 (1.07-2.35)
142	Yeh 2006	Taiwan	General population	1990 - 2001 (10 yrs)	30-65	M	10,923	Drinkers vs non-drinkers	CRC	68	1.23 (0.71-2.16)

Ref.	Author, year	Study name, location	Population	Follow-up period (average follow-up)	Age	Sex	Number of subjects	Exposure	Site	Number of cases	RR (95% CI)
[79]	Chen 2005	China	General population	1990 - 2001 (10.6 yrs)	≥30	B	64,100	Daily vs never drinkers	CRC	242	1.11 (0.74-1.67)
									CC	107	0.97 (0.52-1.78)
									RC	135	1.24 (0.71-2.14)
[143]	Wakai 2005	Japan Collaborative Cohort Study, Japan	General population	1988 - 1997 (7.6 yrs)	40-79	M	23,708	≥66 g/d vs never drinkers	CC	220	2.40 (1.31-4.40)
									RC	150	1.32 (0.67-2.63)
						F	34,028	≥22 g/d vs never drinkers	CC	198	1.22 (0.49-3.03)
									RC	61	1.53 (0.36-6.47)

Ref.	Author, year	Study name, location	Population	Follow-up period (average follow-up)	Age	Sex	Number of subjects	Exposure	Site	Number of cases	RR (95% CI)
[45]	Cho 2004	Alpha-Tocopherol Beta-Carotene Cancer Prevention Study, Finland Canadian National Breast Screening Study, Canada Health Professionals Follow-up Study, USA Iowa Women's Health Study, USA Netherlands Cohort Study, Netherlands New York State Cohort, USA Nurses' Health Study, USA Sweden Mammography Cohort, Sweden	General population	1985 - 1995 1980 - 1993 1986 - 1996 1986 - 1998 1986 - 1993 1980 - 1987 1980 - 1996 1987 - 1998	50-69 40-59 40-75 55-69 55-69 15- 107 34-59 40-76	B	489,979	≥45 g/d vs non-drinkers	CRC CC RC	4,687 3,291 1,370	1.41 (1.16-1.72) 1.45 (1.14-1.83) 1.49 (1.04-2.12)
[144]	Sanjoaquin 2004	Oxford Vegetarian Study, UK	Vegetarians and non-vegetarians	1980 - 1999 (17 yrs)	16-89	B	10,998	>7 vs <1 units/week	CRC	95	1.53 (0.87-2.69)

Ref.	Author, year	Study name, location	Population	Follow-up period (average follow-up)	Age	Sex	Number of subjects	Exposure	Site	Number of cases	RR (95% CI)
145	Su 2004	National Health and Nutrition Examination Survey I Epidemiologic Follow-Up Study, USA	General population	1982 - 1993	36-87	B	10,418	≥1 drink/day vs non-drinkers	CC	111	1.69 (1.03-2.79)
78	Wei 2004	Nurses' Health Study Health Professionals Follow-up Study, USA	Nurses Health professionals	1980 - 2000 (20 yrs) 1986 - 2000 (14 yrs)	34-59 40-75	B	134,365	≥20 g/d vs non-drinkers	CC	1,139	1.27 (1.03-1.56)
									RC	339	1.26 (0.85-1.87)
120	Otani 2003	Japan Public Health Center-based Prospective Study, Japan	General population	1990 - 1999	40-69	M	42,540	≥42.9 g/d vs non-drinkers	CRC	447	2.1 (1.6-2.7)
									CC	299	1.9 (1.4-2.7)
									RC	148	2.4 (1.5-4.0)
						F	47,464	≥1 vs <1/week	CRC	259	0.7 (0.4-1.1)
73	Pedersen 2003	Copenhagen Centre for Prospective Population Studies, Denmark	General population	1964/70/76-1999 (14.7 yrs)	23-95	B	29,132	≥41 vs <1 drinks/week	CC	411	0.8 (0.5-1.5)
									RC	202	2.2 (1.0-4.6)
146	Shimizu 2003	Takayama Study, Japan	General population	1993 - 2000	≥35	M	13,392	>36.7 g/d vs non-drinkers	CC	108	2.67 (1.06-6.76)
									RC	59	1.17 (0.50-2.73)
						F	15,659	>3.75 g/d vs non-drinkers	CC	94	1.78 (1.00-3.18)
									RC	41	1.80 (0.70-4.62)

Ref.	Author, year	Study name, location	Population	Follow-up period (average follow-up)	Age	Sex	Number of subjects	Exposure	Site	Number of cases	RR (95% CI)
98	Flood 2002	Breast Cancer Detection Demonstration Project Follow-up Study, USA	General population	1987 - 1998 (8.5 yrs)	40-93	F	45,264	>2 servings/day vs non-drinkers	CRC	490	1.16 (0.63-2.14)
135	Chen 2001	Physicians' Health Study, USA	Physicians	1982 - 1995	40-84	M	1,113*	≥5 vs ≤1 drinks/week	CRC	211	1.25 (0.85-1.84)
147	Singh 1998	Adventist Health Study, USA	Seventh-day Adventists	1977 - 1982 (6 yrs)	≥25	B	32,051	≥1 vs <1/week	CC	157	2.05 (1.00-4.23)
74	Chyou 1996	Honolulu Heart Program, Hawaii	General population	1965 - 1995 (16.2 yrs)	45-65	M	7,945	≥24 oz/month vs non-drinkers	CC	328	1.39 (1.05-1.83)
									RC	123	2.30 (1.43-3.69)
148	Glynn 1996	Alpha-Tocopherol Beta-Carotene Cancer Prevention Study, Finland	Smokers	1985 - 1993	50-69	M	27,109	>25.6 vs ≤2.6 g/d	CRC	140	1.3 (0.8-2.1)
149	Kreger 1992	Framingham Study, USA	General population	1948 - 1988	30-62	M	2,336	Per 1oz/week	CC	56	1.00 (0.97-1.04)
						F	2,873	Per 1oz/week	CC	66	0.95 (0.87-1.04)

Ref.	Author, year	Study name, location	Population	Follow-up period (average follow-up)	Age	Sex	Number of subjects	Exposure	Site	Number of cases	RR (95% CI)
[75]	Klatsky 1988	Kaiser Permanente Study, USA	General population	1978 - 1984		B	10,303*	≥3 drinks/day vs non-drinkers	CRC	268	1.94 (1.14-3.31)
									CC	203	1.71 (0.92-3.19)
									RC	66	3.17 (1.05-9.57)
[76]	Wu 1987	USA	Retirement community	1981 - 1985		B	11,644	>30 ml/day vs non-daily drinkers	CRC	126	1.9 (1.3-2.9)

M = males, F = females, CRC = colorectal cancer, CC = colon cancer, PCC = proximal colon cancer, DCC = distal colon cancer, RC = rectal cancer

*Number of controls in a nested case-control study

†Number of non-cases in case-cohort analysis

2.3 Adiposity and Colorectal Cancer Literature Review

This section reviews the existing evidence for an association between adiposity and colorectal cancer, focusing on body mass index, waist circumference and waist to hip ratio. BMI is widely used in epidemiological studies to assess adiposity. A large number of studies have presented results for BMI and colorectal cancer and the WCRF concluded that there is convincing evidence that BMI is causally related to colorectal cancer.[25] Fewer studies have investigated WC and WHR though the WCRF also concluded that there was convincing evidence that WC and WHR are causally related to colorectal cancer.[25] This review first discusses the evidence for the association between BMI and colorectal cancer. The evidence for WC and WHR is then discussed. Other measures such as adiposity in early adulthood and weight change are also considered in this review. Possible effect modifiers and potential mechanisms are then described.

2.3.1 Search Strategy

The WCRF published reports in 2007 and 2011 summarising the evidence from cohort studies for an association between adiposity (specifically BMI, WC and WHR) and colorectal cancer.[25, 56] The WCRF searched the Medline database for articles published up to December 2009. There were a number of recent meta-analyses of adiposity and colorectal cancer that were also searched for relevant studies.[150-156] These meta-analyses are summarised in Table 2.3.1. Review articles were also searched for relevant studies.[157, 158]

To complement the articles identified by these reports and meta-analyses, PubMed was searched for articles reporting human studies, written in English, and published between 1st January 2008 and 31st March 2015, using the MeSH terms “Colorectal Neoplasms” and at least one of “Body Mass Index”, “Body Weight”, “Waist Circumference” or “Waist-Hip Ratio”. 748 articles were retrieved.

Table 2.3.1 Meta-analyses of Adiposity and Colorectal Cancer

Ref.	Author, Year	Search Date	Adiposity Measure	Number of Studies	Main Results	
					Comparison	RR (95% CI)
[150]	Ma 2013	01/2012	BMI	41* P	Highest vs lowest category	M CC 1.55 (1.47-1.63) M RC 1.24 (1.11-1.38) F CC 1.23 (1.10-1.37) F RC 1.07 (1.01-1.14)
			WC	13 P	Highest vs lowest category	M CC 1.81 (1.46-2.42) M RC 1.28 (0.99-1.66) F CC 1.50 (1.25-1.79) F RC 1.50 (1.03-2.18)
[151]	Robsahm 2013	12/2010	BMI	17 P	Highest vs lowest category	B PCC 1.24 (1.08-1.42) B DCC 1.59 (1.34-1.89) B RC 1.23 (1.02-1.48)
[152]	Ning 2010	02/2008	BMI	44† P 14 R	≥30 vs <23.0 kg/m^2	M CC 1.60 (1.53-1.69) M RC 1.30 (1.17-1.43) F CC 1.25 (1.12-1.39) F RC 1.14 (1.02-1.27)
[153]	Harriss 2009	12/2007	BMI	28 P	5 kg/m^2 increase	M CC 1.24 (1.20-1.28) M RC 1.09 (1.06-1.12) F CC 1.09 (1.04-1.14) F RC 1.02 (0.99-1.04)
[154]	Larsson 2007	04/2007	BMI	31† P	5 kg/m^2 increase	M CC 1.30 (1.25-1.35) M RC 1.12 (1.09-1.16) F CC 1.12 (1.07-1.18) F RC 1.03 (0.99-1.08)
			WC	6 P	10 cm increase	M CC 1.33 (1.19-1.49) M RC 1.12 (1.03-1.22) F CC 1.16 (1.09-1.23) F RC 1.09 (0.99-1.20)
			WHR	7 P	0.1 increase	M CC 1.43 (1.19-1.71) M RC 1.22 (0.81-1.83) F CC 1.20 (1.08-1.33) F RC 1.15 (0.95-1.39)
[155]	Moghaddam 2007	04/2007	BMI	23 P 8 R	≥30 vs <25.0 kg/m^2	M CC 1.51 (1.42-1.61) M RC 1.29 (1.19-1.40) F CC 1.16 (1.01-1.34) F RC 1.08 (0.92-1.26)
			WC	10 P	Highest vs lowest category	B CRC 1.50 (1.35-1.67)

Ref.	Author, Year	Search Date	Adiposity Measure	Number of Studies	Main Results	
					Comparison	RR (95% CI)
[156]	Dai 2007	01/2007	BMI	15 P	≥30 vs 18.5-24.9 kg/m^2	M CC 1.71 (1.33-2.19) M RC 1.75 (1.17-2.62) F CC 1.10 (0.92-1.32) F RC 1.12 (0.84-1.49)
			WC	6 P	Highest vs lowest category	M CC 1.68 (1.36-2.08) M RC 1.26 (0.90-1.77) F CC 1.48 (1.19-1.84) F RC 1.23 (0.81-1.86)
			WHR	6 P	Highest vs lowest category	M CC 1.91 (1.46-2.49) M RC 1.93 (1.19-3.13) F CC 1.49 (1.23-1.81) F RC 1.20 (0.81-1.78)

*Some studies included more than once.

†Included 5 studies of cancer mortality.

BMI, body mass index; WC, waist circumference; WHR, waist to hip ratio; P, prospective studies; R, retrospective studies; M, males; F females; B, both males and females; CRC, colorectal cancer; CC, colon cancer; PCC, proximal colon cancer; DCC, distal colon cancer; RC, rectal cancer.

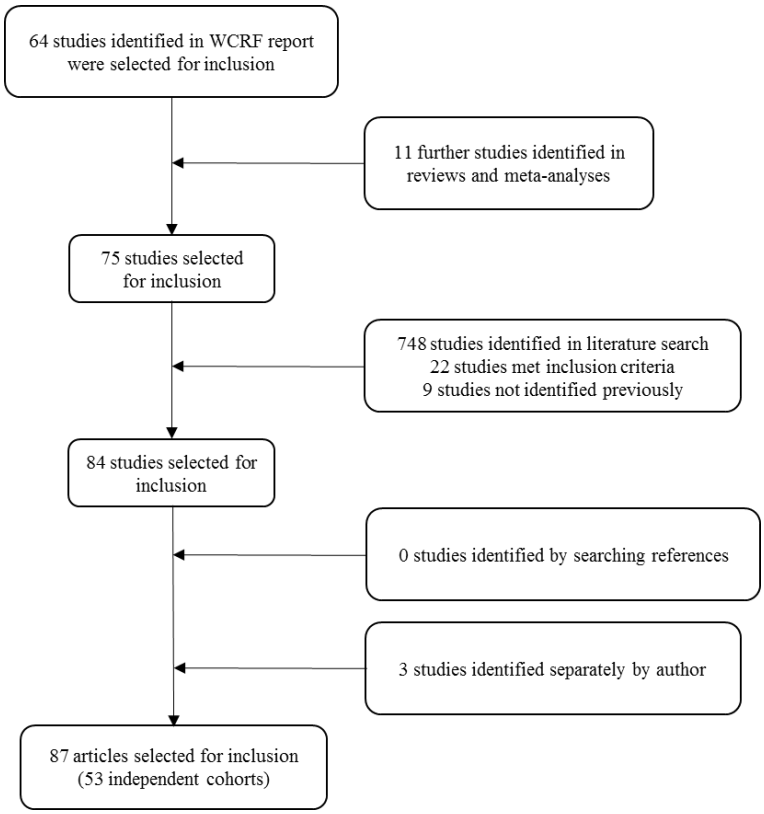
Articles were selected for inclusion in this review if they (i) were prospective cohort studies (including nested case-control studies and historical cohort studies), (ii) analysed measures of current adiposity (BMI, WC or WHR) assessed at baseline in relation to colorectal cancer incidence (or colorectal subsites) in the general population at least 18 years old and (iii) reported risk estimates (HRs, RRs, ORs) and CIs. Analyses that reported results by other factors were not included. Analyses based on fewer than 30 cases were ignored. References of included articles were carefully examined for additional articles.

Figure 2.3.1 shows a flow diagram of the literature search process, illustrating the number of articles identified from the different sources. Overall, 87 articles met the inclusion criteria for this literature review. These 87 articles were based on 53 independent cohorts.

Table 2.3.6, presented at the end of this review, shows the key characteristics of the articles (identified in the literature search) mentioned in this review. Results are shown for BMI, WC and WHR and results are presented separately for colon and rectal cancer and men and women where possible. At least one article from each cohort identified in

the literature search is included. More than one article from the same cohort is included if they include analyses of different measures of adiposity.

Figure 2.3.1 Flowchart of Literature Search Process for Adiposity and Colorectal Cancer



WCRF = World Cancer Research Fund

2.3.2 Body Mass Index

The most common measure of adiposity is BMI, defined as the weight in kilograms divided by the square of the height in metres (kg/m^2). The WHO has defined categories of BMI which are widely used in epidemiological studies of adiposity. BMI below 18.5 kg/m^2 is defined as underweight, BMI between 18.5-<25.0 kg/m^2 is defined as normal weight, BMI ≥25.0 kg/m^2 is defined as overweight and BMI ≥30.0 kg/m^2 is defined as obese.[159]

Prospective cohort studies have consistently found an increased risk of colon cancer associated with excess body weight as measured by BMI. BMI is also associated with rectal cancer risk though the evidence is less clear. The associations between BMI and colon and rectal cancer are stronger for men than for women.[150, 152, 153]

The NIH-AARP Diet and Health Study recruited 307,708 men and 209,436 women aged 50-71 years old in 1995-1996 who were followed-up until 2000.[160] The results for the association between BMI and colon and rectal cancer for men and women are shown in Table 2.3.2. There was strong evidence for an association between BMI and colon cancer risk for men with increasing levels of BMI associated with greater risk. BMI was also associated with colon cancer risk for women but the evidence was weaker. There was no clear evidence for an increased risk of rectal cancer for men or women.

Table 2.3.2 Body Mass Index and the Risk of Colon Cancer and Rectal Cancer in the NIH-AARP Diet and Health Study[160]

	Colon cancer			Rectal cancer	
BMI, kg/m^2	**Cases**	**HR (95% CI)**	**BMI, kg/m^2**	**Cases**	**HR (95% CI)**
Men			Men		
18.5-<23	136	1.00	18.5-<23	74	1.00
23-<25	260	1.11 (0.90-1.37)	23-<25	101	0.78 (0.58-1.06)
25-<27.5	479	1.22 (1.01-1.48)	25-<27.5	218	1.01 (0.77-1.31)
27.5-<30	367	1.44 (1.18-1.76)	27.5-<30	135	0.96 (0.72-1.28)
30-<32.5	219	1.53 (1.23-1.90)	30-<32.5	74	0.94 (0.68-1.30)
32.5-<35	110	1.57 (1.22-2.03)	32.5-<35	42	1.10 (0.75-1.61)
35-<40	76	1.71 (1.29-2.27)	≥35	33	1.00 (0.68-1.58)
≥40	29	2.39 (1.59-3.58)			
Women			Women		
18.5-<23	151	1.00	18.5-<23	60	1.00
23-<25	141	1.20 (0.95-1.51)	23-<25	49	1.05 (0.72-1.53)
25-<27.5	172	1.29 (1.03-1.60)	25-<27.5	60	1.13 (0.79-1.63)
27.5-<30	106	1.31 (1.01-1.68)	27.5-<30	37	1.16 (0.76-1.76)
30-<32.5	77	1.28 (0.97-1.69)	30-<32.5	26	1.09 (0.68-1.75)
32.5-<35	42	1.13 (0.80-1.60)	32.5-<35	14	0.95 (0.52-1.71)
35-<40	52	1.46 (1.06-2.02)	≥35	32	1.44 (0.92-2.25)
≥40	28	1.49 (0.98-2.25)			

In the Nurses' Health Study and Health Professionals Follow-up Study, 87,733 female nurses and 46,632 male health professionals were followed-up from 1980 and 1986, respectively, until 2000 with repeat questionnaires on weight (and other variables) every two years.[78] Similar results to the NIH-AARP Diet and Health Study were seen for colon cancer with a strong, dose-response association for men and a weaker association for women (Table 2.3.3). There was no evidence for an association with rectal cancer for men but BMI appeared to be more strongly related to rectal cancer than colon cancer for women (though the CIs were wide).

Table 2.3.3 Body Mass Index and the Risk of Colon Cancer and Rectal Cancer in the Health Professionals Follow-up Study and the Nurses' Health Study[78]

	Colon cancer		Rectal cancer	
BMI, kg/m^2	**Cases**	**RR (95% CI)**	**Cases**	**RR (95% CI)**
Men				
<23	57	1.00	24	1.00
23-<25	119	1.33 (0.97-1.83)	42	1.16 (0.70-1.94)
25-<30	225	1.54 (1.15-2.07)	55	0.93 (0.57-1.53)
≥30	51	1.85 (1.26-2.72)	11	1.03 (0.49-2.14)
Women				
<23	210	1.00	56	1.00
23-<25	141	1.10 (0.88-1.36)	46	1.37 (0.92-2.02)
25-<30	207	1.11 (0.91-1.35)	68	1.40 (0.98-2.01)
≥30	113	1.28 (1.10-1.62)	34	1.56 (1.01-2.42)

The EPIC study analysed the risk of colon cancer and rectal cancer among 129,731 men and 238,546 women recruited from ten European countries.[47] Results were presented according to sex-specific quintiles of BMI (Table 2.3.4). There was only evidence for an increased risk of colon cancer for men in the highest quintile of BMI; there was no evidence for an association for colon cancer for women or for rectal cancer for men or women.

Table 2.3.4 Body Mass Index and the Risk of Colon Cancer and Rectal Cancer in the European Investigation into Cancer and Nutrition[47]

	Colon cancer		Rectal cancer	
BMI, kg/m^2	**Cases**	**HR (95% CI)**	**Cases**	**HR (95% CI)**
Men				
<23.6	64	1.00	52	1.00
23.6-25.3	85	1.18 (0.85-1.63)	52	0.88 (0.60-1.30)
25.4-27.0	74	1.00 (0.71-1.41)	58	0.96 (0.66-1.40)
27.1-29.3	88	1.19 (0.85-1.66)	69	1.11 (0.77-1.62)
≥29.4	110	1.55 (1.12-2.15)	64	1.05 (0.72-1.55)
Women				
<21.7	87	1.00	47	1.00
21.7-23.5	96	0.92 (0.68-1.23)	44	0.78 (0.51-1.18)
23.6-25.7	120	1.02 (0.77-1.35)	72	1.14 (0.78-1.66)
25.8-28.8	137	1.09 (0.83-1.45)	63	0.95 (0.65-1.41)
≥28.9	123	1.04 (0.79-1.42)	65	1.06 (0.78-1.51)

Bhaskaran et al. utilised primary care data on over five million UK adults to analyse the relationship between BMI and colorectal cancer.[161] During the defined follow-up period, over 13,000 cases of colon cancer and 6,000 cases of rectal cancer were identified. Besides measured height and weight, data on smoking status, alcohol use, diabetes and deprivation were also available. For men there was a piecewise linear

relationship between BMI and colon cancer such that there was an association only for men with BMI 22-34 kg/m^2 (which included the large majority of men). The HR and 99% CI for a 5 kg/m^2 increase for BMI <22, 22-34 and >34 kg/m^2 was 0.92 (0.69-1.23), 1.23 (1.17-1.30) and 0.97 (0.81-1.15). For women, the increased risk was more modest with no evidence for non-linearity (HR, 1.05; 99% CI, 1.01-1.08 for 5 kg/m^2 increase). For rectal cancer there was no evidence for effect modification by gender or a non-linear relationship (HR, 1.04; 99% CI, 1.00-1.08 for 5 kg/m^2 increase for men and women).

The large majority of studies investigating adiposity and colorectal cancer risk have been carried out in Western countries. For the same level of BMI, Asian people generally have a higher percentage of body fat than white people and a greater risk of type 2 diabetes and cardiovascular disease.[162] Thus, it is possible that BMI represents a stronger risk factor for colorectal cancer in Asian populations.

One study pooled data from eight Japanese cohorts including almost 5,000 cases of colorectal cancer.[163] Compared to men and women with BMI 23-<25 kg/m^2, men and women with BMI ≥30 kg/m^2 had an increased risk of colon cancer (Table 2.3.5). There was also slight evidence of a decreased risk of colon cancer for men and women with BMI <19 kg/m^2. For rectal cancer, there was evidence of an increased risk only for men with BMI ≥30 kg/m^2.

In a study of Chinese men and women, men (but not women) in the highest quintile of BMI had an increased risk of colon cancer but there was no clear evidence of an association with rectal cancer for men or women.[164] Jee et al. analysed data from over one million Korean men and women who completed a medical examination as part of the national health system, identifying 8,703 cases of colorectal cancer for men and 3,640 cases for women.[165] Compared with men with BMI 23.0-24.9 kg/m^2, men with BMI ≥30 kg/m^2 had an increased risk of colon cancer but not rectal cancer. Men with BMI <20.0 and 20.0-22.9 kg/m^2 had a decreased risk of colon cancer and rectal cancer. For women, there was no evidence for an increased risk of colon or rectal cancer for women with BMI ≥30 kg/m^2 but there was evidence for a decreased risk of colon cancer for women with BMI <20.0 and 20.0-22.9 kg/m^2. There was also slight evidence of a

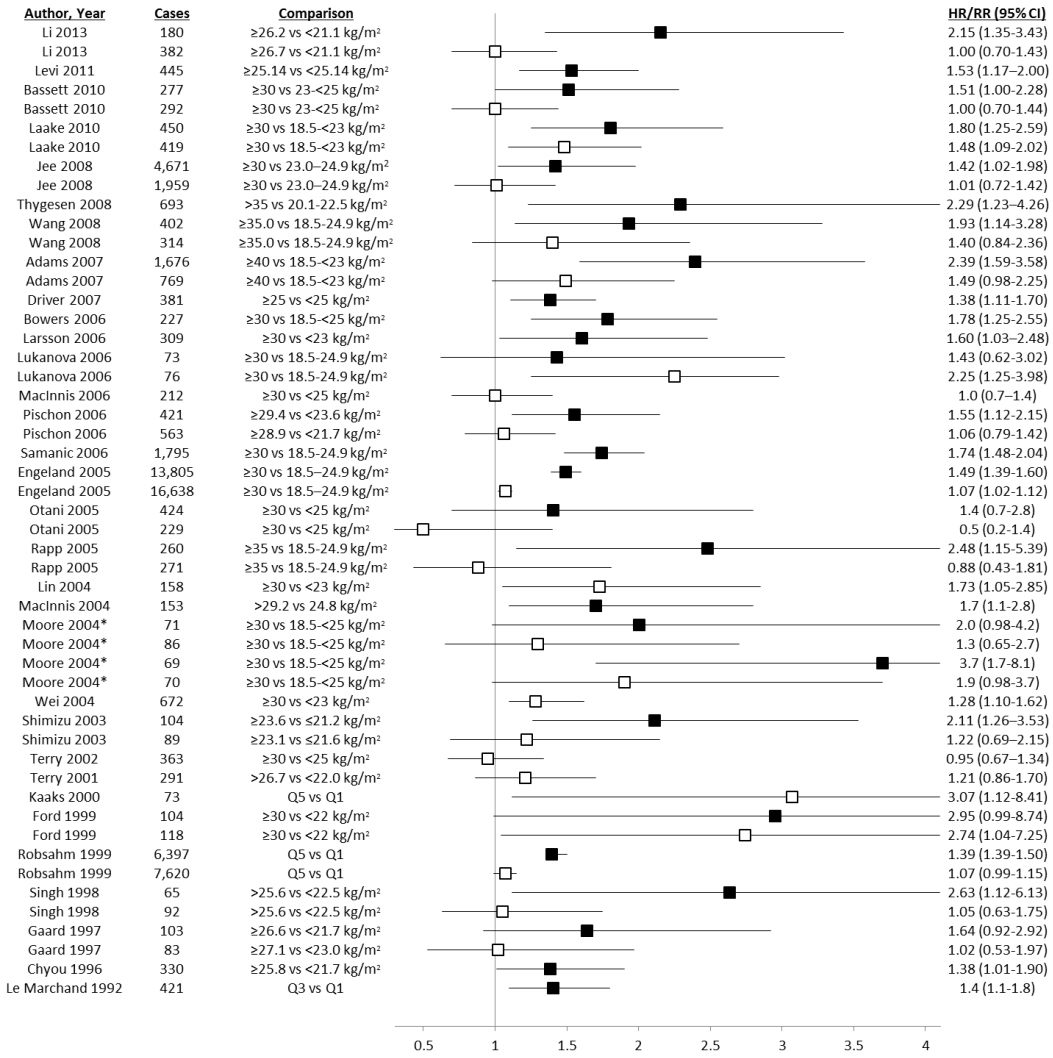
decreased risk of rectal cancer for women with BMI <20.0 kg/m^2. (Results were adjusted only for age and smoking status).

Table 2.3.5 Body Mass Index and the Risk of Colon Cancer and Rectal Cancer in the Pooled Analysis of Japanese Studies[163]

	Colon cancer		Rectal cancer	
BMI, kg/m^2	**Cases**	**HR (95% CI)**	**Cases**	**HR (95% CI)**
Men				
<19	98	0.84 (0.67-1.04)	59	0.92 (0.69-1.23)
19-<21	317	0.94 (0.81-1.08)	179	0.95 (0.79-1.15)
21-<23	473	0.86 (0.76-0.97)	325	1.09 (0.93-1.28)
23-<25	512	1.00	284	1.00
25-<27	319	1.16 (1.01-1.34)	158	1.04 (0.86-1.27)
27-<30	168	1.27 (1.07-1.52)	80	1.17 (0.91-1.52)
≥30	32	1.37 (0.96-1.98)	26	1.85 (1.23-2.78)
Women				
<19	76	0.80 (0.61-1.04)	53	1.31 (0.95-1.81)
19-<21	215	1.00 (0.83-1.20)	97	0.98 (0.76-1.27)
21-<23	330	1.03 (0.88-1.21)	147	0.94 (0.74-1.18)
23-<25	512	1.00	284	1.00
25-<27	217	1.18 (0.99-1.41)	80	0.88 (0.66-1.17)
27-<30	136	1.22 (0.99-1.51)	54	0.92 (0.67-1.27)
≥30	48	1.39 (1.02-1.90)	20	1.33 (0.82-2.15)

Figure 2.3.2 and Figure 2.3.3 show the pattern of results by sex across the individual cohorts included in this review for the association between BMI and colon cancer and rectal cancer. Only one set of results is included for each cohort.

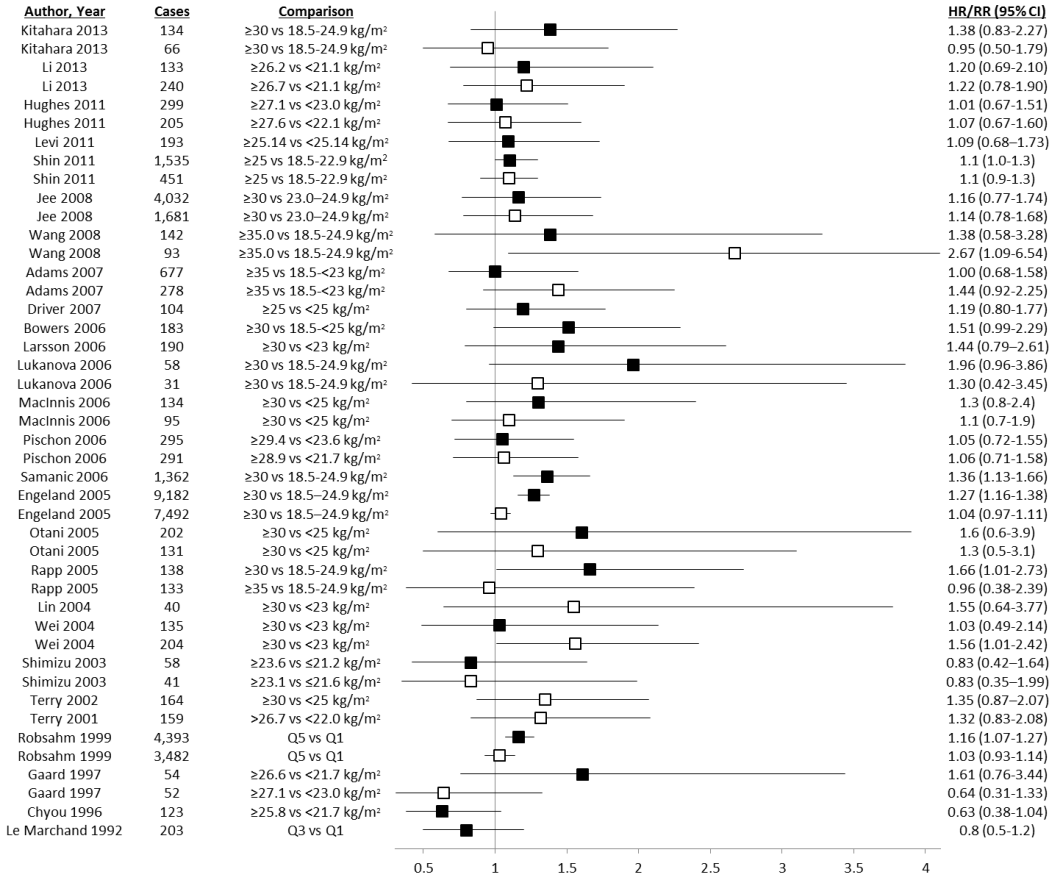
Figure 2.3.2 Results from Prospective Cohort Studies for the Association between Body Mass Index and Colon Cancer Risk



Results show the comparison between the highest BMI category and the reference group from prospective cohort studies. Squares represent HRs/RRs. Bars represent 95% CIs. Black squares represent results for men and white squares represent results for women.

*Separate results presented for different age groups.

Figure 2.3.3 Results from Prospective Cohort Studies for the Association between Body Mass Index and Rectal Cancer Risk



Results show the comparison between the highest BMI category and the reference group from prospective cohort studies. Squares represent HRs/RRs. Bars represent 95% CIs. Black squares represent results for men and white squares represent results for women.

Proximal Colon Cancer and Distal Colon Cancer

Studies have also investigated the associations with proximal colon and distal colon cancer.[138, 163, 166-180] In the NIH-AARP Diet and Health Study, results were similar for proximal and distal colon cancer for men (HR, 1.57; 95% CI, 1.09-2.25 and HR, 1.68; 95% CI, 1.05-2.68, comparing ≥35 vs 18.5-21.9 kg/m^2) and women (HR, 1.15; 95% CI, 0.82-1.63 and HR, 1.35; 95% CI, 0.81-2.25).[167] Comparing highest and lowest quintiles of BMI in the Netherlands Cohort Study, results were also similar for proximal and distal colon cancer for men (HR, 1.35; 95% CI, 0.93-1.98 and HR, 1.38; 95% CI, 0.95-1.98) and women (HR, 0.91; 95% CI, 0.65-1.28 and HR, 1.04; 95% CI, 0.72-1.50).[168] The pooled Japanese study also found similar results for proximal and distal colon cancer, for men and women, though the CIs were very wide.[163]

In contrast, however, some studies have found evidence for a stronger association for distal colon cancer. For example, the HRs and 95% CIs from a Norwegian cohort study comparing BMI ≥30 vs 18.5-<23.0 kg/m^2 were 1.17 (0.68-2.00) for proximal colon cancer and 3.26 (1.79-5.95) for distal colon cancer for men.[169] Results for women were 1.48 (1.09-2.02) and 1.65 (1.01-2.70). The Nurses' Health Study also found a stronger association for distal colon cancer (RR, 1.96; 95% CI, 1.18-3.25 for ≥29 vs <21 kg/m^2) than proximal colon cancer (RR, 1.26; 95% CI, 0.71-2.23).[180] Two other studies found slight evidence for a stronger association for proximal colon cancer.[174, 175]

Meta-analysis results support a stronger association for distal colon cancer.[150, 151, 153] Pooling highest vs lowest results from prospective cohort studies, the RRs and 95% CIs were 1.24 (1.08-1.42) for proximal colon cancer and 1.59 (1.34-1.89) for distal colon cancer.[151]

Differences between Studies

It remains unclear if there are reasons why different studies have found different results. Meta-analyses found that associations were stronger from studies with self-reported height and weight than from studies that measured height and weight.[152-154] For example, one meta-analysis found that the RR (95% CI) for a 5 kg/m^2 increase in BMI in relation to colon cancer for men was 1.32 (1.21-1.44) for studies using self-reported data and 1.22 (1.19-1.26) for studies using measured data.[153] There is expected to be measurement error when asking participants to report height and weight data. However, random measurement error would attenuate associations towards the null.[102] Thus, since analysing self-reported height and weight results in stronger associations, it seems that BMI tends to be underestimated in these studies.

In fact, it has been shown in numerous studies that both men and women tend to overestimate height and underestimate weight, resulting in underestimated BMI.[48] Also, although studies generally find high correlations between self-reported and measured BMI, there is large variation at the individual level with the extent of underestimation increasing with increasing BMI and increasing age for both men and women.[181-183] It was estimated in one study that 41% of obese men and 27% of obese women were

misclassified into a lower BMI category using self-reported measures of height and weight.[181]

It is possible that other aspects of study design or analysis may also contribute to the differences in results observed between studies. Studies have adjusted for different sets of confounders and the quality of the measurement of these confounders will vary between studies. Another possibility is that the different results are due to the differences in study populations. The distributions of different lifestyle factors will differ between studies. If any of these factors influence the association between BMI and colorectal cancer, this would result in different studies finding different results.

For example, the Million Women Study (including 4,008 cases of colorectal cancer) found a null association between BMI and colorectal cancer risk overall.[184] However, stratifying the analysis by menopausal status, BMI was associated with an increased risk of colorectal cancer among premenopausal women (HR, 1.61; 95% CI, 1.05-2.48 per 10 kg/m^2 increase) but not postmenopausal women who were never users of hormone replacement therapy (HR, 0.99; 95% CI, 0.88-1.12 per 10 kg/m^2 increase). Thus, this could mean that the association between BMI and colorectal cancer for women depends on the proportions of pre- and post-menopausal women. The potential effects of menopause and other risk factors on the association between BMI and colorectal cancer are described in section 2.3.6.

Summary

There is convincing evidence for an association between BMI and colon cancer. The evidence for the relationship between BMI and rectal cancer is less consistent. Associations are stronger for men than for women. The association appears to be stronger for the distal colon than the proximal colon. Studies relying on self-reported height and weight may overestimate results.

2.3.3 Waist Circumference and Waist to Hip Ratio

BMI is the most commonly studied measure of adiposity since it is very straightforward to assess in large populations; almost all cohorts include data on height and weight

(whether measured or self-reported). This has resulted in a multitude of studies relating BMI to colorectal cancer, albeit with varying quality.[150, 152] As described above, BMI is strongly associated with colon cancer risk for men with a weaker association for women. The exact reasons for the discrepancy in results between men and women remain unclear.

One potential explanation is that men and women have different body compositions. BMI is a rather crude measure of weight adjusted for height; it does not discriminate fat mass from lean mass and ignores fat distribution. Men tend to store excess weight centrally (abdominal adiposity) whereas women are more likely to store excess weight in the thighs and buttocks (gluteofemoral adiposity).[185] Also, visceral adipose tissue (stored within the abdominal cavity) is more metabolically active and secretes greater levels of cytokines and hormones compared to subcutaneous adipose tissue (stored just below the skin).[185] Hence, visceral adiposity is thought to be of particular importance for colorectal cancer[157, 158] and it has been hypothesised that measures of abdominal adiposity such as WC and WHR may be more directly associated with colorectal cancer risk for both men and women.[47]

Results from the EPIC study provided strong support for the hypothesis that measures of central adiposity are more important than BMI for assessing colorectal cancer risk. As described in the previous section, BMI was associated with an increased risk of colon cancer for men but not women.[47] In contrast, WC and WHR were associated with an increased risk of colon cancer for both men (HR, 1.39; 95% CI, 1.01-1.93 for highest vs lowest quintile of WC and HR, 1.51; 95% CI, 1.06-2.15 for highest vs lowest quintile of WHR) and women (HR, 1.48; 95% CI, 1.08-2.03 for highest vs lowest quintile of WC and HR, 1.52; 95% CI, 1.12-2.05 for highest vs lowest quintile of WHR). Furthermore, WC and WHR were still associated with colon cancer risk for women after further adjustment for weight (as well as height) (HR, 1.44; 95% CI, 0.92-2.26 and HR, 1.46; 95% CI, 1.06-2.00).[47] However, there was no longer evidence for an association with WC or WHR for men (HR, 1.01; 95% CI, 0.62-1.65 and HR, 1.18; 95% CI, 0.79-1.76). Hence, these results suggest that WC and WHR convey important information about colon cancer risk beyond BMI for women.

Results from the Melbourne Collaborative Cohort Study followed a similar pattern. BMI (HR, 1.7; 95% CI, 1.1-2.8 for highest vs lowest quartile), WC (HR, 2.1; 95% CI, 1.3-3.5) and WHR (HR, 2.1; 95% CI, 1.3-3.4) were all associated with colon cancer risk for men.[173] Women in the highest tertile of BMI were not at an increased risk of colon cancer (HR, 1.0; 95% CI, 0.7-1.4) but there was evidence of an increased risk for women in the highest tertiles of WC (HR, 1.4; 95% CI, 1.0-1.9) and WHR (HR, 1.7; 95% CI, 1.1-2.4).[172]

Other studies, however, do not provide such clear support. Results from the NIH-AARP Diet and Health Study were in complete contrast to those from the EPIC study. BMI, WC and WHR were each associated with an increased risk of colon cancer for men but there was no association for women for any of these measures.[186] The HRs and 95% CIs comparing the highest and lowest quintiles for BMI, WC and WHR for men were 1.42 (1.19-1.68), 1.45 (1.16-1.82) and 1.29 (1.10-1.52). The corresponding results for women were 0.96 (0.74-1.23), 0.90 (0.63-1.27) and 0.90 (0.70-1.15). After adjusting for BMI, results were attenuated but WC (HR, 1.32; 95% CI, 1.03-1.70) and WHR (HR, 1.17; 95% CI, 0.99-1.38) were both still associated with colon cancer for men, suggesting WC/WHR may provide important information beyond BMI for colon cancer risk for men.

One study found that both BMI (HR, 1.93; 95% CI, 1.14-3.28 for ≥35.0 vs 18.5-24.9 kg⁄m^2) and WC (HR, 2.05; 95% CI, 1.29-2.35 for ≥120 vs <95 cm) were strongly associated with colon cancer risk for men.[187] The associations were weaker for women (HR, 1.40; 95% CI, 0.84-2.36 for ≥35.0 vs 18.5-24.9 kg⁄m^2 and HR, 1.54; 95% CI, 1.00-2.37 for ≥110 vs <85 cm). Mutually adjusting for BMI and WC, the associations with WC seemed to remain for men and women while the associations with BMI were attenuated (though the CIs were very wide).

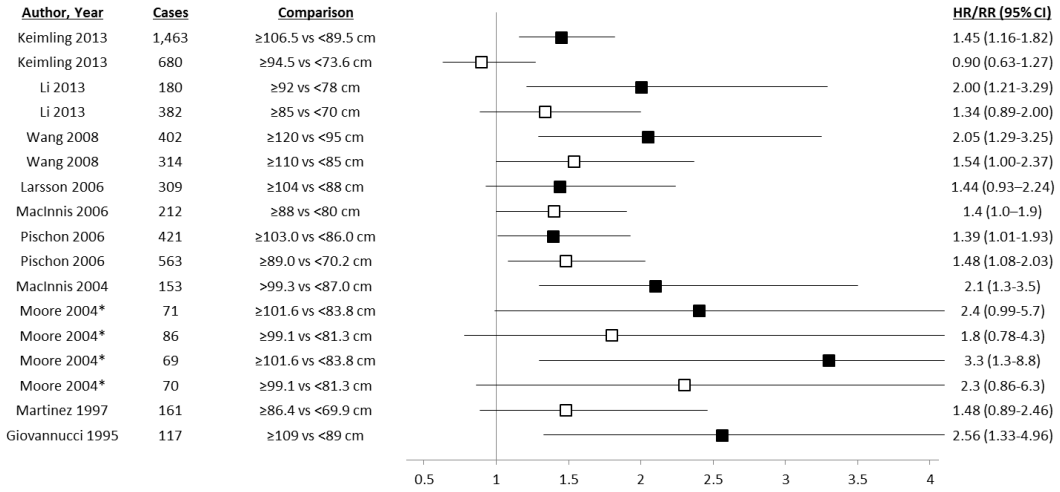
The Iowa Women's Health Study found a similar increased risk of colorectal cancer comparing high and low quartiles of BMI (HR, 1.29; 95% CI, 1.10-1.51), WC (HR, 1.32; 95% CI, 1.11-1.56) and WHR (HR, 1.28; 95% CI, 1.08-1.50).[188] A cohort of Chinese men and women found a strong association between BMI and colon cancer for men (HR, 2.15; 95% CI, 1.35-3.43 for highest vs lowest quintile) but no association for women (HR, 1.00; 95% CI, 0.70-1.43).[164] WC and WHR were also strongly associated

with colon cancer risk for men (HR, 2.00; 95% CI, 1.21-3.29 for WC and HR, 1.97; 95% CI, 1.19-3.24 for WHR) but there was no association for women (HR, 1.34; 95% CI, 0.89-2.00 for WC and HR, 0.96; 95% CI, 0.69-1.34 for WHR).

Hence, it is unclear whether WC or WHR represent more accurate predictors of colon cancer risk and whether the stronger association between BMI and colon cancer for men than for women is due to different body compositions. Though the EPIC study found similar associations between WC/WHR and colon cancer for men and women, most studies find a stronger association for men.[150, 154]

Figure 2.3.4 and Figure 2.3.5 show the results by sex from different cohorts for the association between WC and colon cancer and rectal cancer. Only one set of results is included for each cohort.

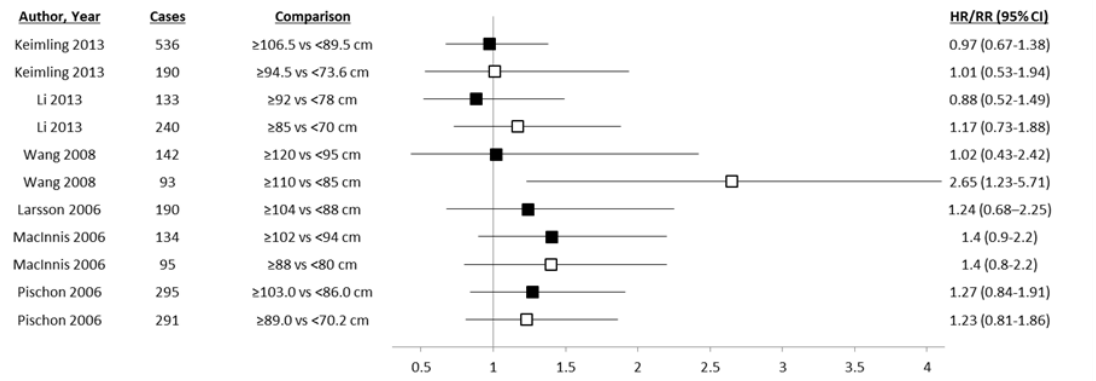
Figure 2.3.4 Results from Prospective Cohort Studies for the Association between Waist Circumference and Colon Cancer Risk



Results show the comparison between the highest WC category and the reference group from prospective cohort studies. Squares represent HRs/RRs. Bars represent 95% CIs. Black squares represent results for men and white squares represent results for women.

*Separate results presented for different age groups.

Figure 2.3.5 Results from Prospective Cohort Studies for the Association between Waist Circumference and Rectal Cancer Risk



Results show the comparison between the highest WC category and the reference group from prospective cohort studies. Squares represent HRs/RRs. Bars represent 95% CIs. Black squares represent results for men and white squares represent results for women.

Summary

BMI is predominantly used to measure adiposity due to its simplicity. However, BMI may be limited as a measure of adiposity since it ignores fat distribution. Measures of central adiposity such as WC and WHR are also associated with colon cancer risk, particularly for men. Results are conflicting for women. It remains unclear whether WC or WHR may be more directly associated with colorectal cancer risk than BMI, for men or women.

2.3.4 Early Adulthood Adiposity

Another question that remains about the relationship between adiposity and colorectal cancer is how adiposity throughout life influences lifetime risk. Adiposity during early adulthood and weight change throughout life may be just as important as attained weight for determining colorectal cancer risk. People with similar levels of excess weight at baseline (i.e. cohort initiation) may have had very different trajectories of weight gain and this may result in different levels of risk.

Few studies have included contemporaneous measures of early adulthood adiposity.[189-192] All of these studies were historical cohort studies, relying on existing data on exposure and follow-up. The main disadvantage of historical cohort studies is that, since data were collected for other purposes, the quality of data can be poor. For example, the

data on confounders were very limited in these studies. Results from these studies do suggest that excess weight in early adulthood is associated with colon cancer risk for men (none of these studies included results for women).[189, 191, 192] The largest study included measured height and weight data (recorded as part of an examination for military service) for over a million Jewish Israeli men aged 16-19 from 1967 to 2005 who were followed up for cancer incidence until the end of 2006.[189] The HR (95% CI) for colon cancer comparing men above with men below the 85th percentile of BMI (25.14 kg⁄m^2) was 1.53 (1.17-2.00).[189]

Other studies have relied on participant recall in order to analyse the effects of early adiposity on later colorectal cancer risk. Since self-reported current weight involves significant measurement error, there should be even greater concern about the accuracy of recalled weight. Studies comparing measured weight during early adulthood and recalled weight many years later show that people tend to underestimate their previous weight and also that people's recall is influenced by current BMI.[193-195]

Overall, studies that used participants' recall to analyse early adulthood BMI do not provide strong evidence that early adulthood BMI increases the risk of colorectal cancer.[164, 167, 168, 188, 196-198] However, Zhang et al. did find that BMI at age 18 was associated with an increased risk of colorectal cancer for women in the Nurses' Health Study (HR, 1.71; 95% CI, 1.16-2.51 for ≥30 vs <18.5 kg⁄m^2) though there was no association for BMI at age 21 for men in the Health Professionals Follow-up Study (HR, 1.11; 95% CI, 0.70-1.76 for ≥30 vs <18.5 kg⁄m^2).[198] Two other studies also found slight evidence for an increased risk of colon cancer with BMI at age 18/20 for men but no evidence for women.[167, 168]

There are two main issues concerning the analysis of early adiposity with colorectal cancer risk. The first issue is that there is a much narrower range of BMI during early adulthood, meaning that many studies will lack sufficient statistical power to evaluate the risk due to excess BMI during early adulthood. The second issue is that people with higher BMI during early adulthood will be more likely to have higher BMI later in life. Thus, it is very difficult to identify whether early adulthood adiposity is associated with an increased risk of colorectal cancer independent of later adiposity. Zhang et al. was the only study that considered adult BMI when investigating the effect of early

adulthood BMI on colorectal cancer risk and actually found evidence to support an independent effect of early adulthood BMI.[198] Adjusting for adult BMI, early adulthood BMI was still associated with colon cancer risk among women (HR, 1.44; 95% CI, 1.06-1.95).

Summary

Adiposity during early adulthood may increase the risk of colorectal cancer later in life. Crucially, however, it is unknown whether early adiposity is related to colorectal cancer risk independent of adiposity in later life.

2.3.5 Weight Change

A number of studies have analysed the relationship between weight change and colorectal cancer risk. Weight change may represent a useful measure of adiposity since adult weight gain generally occurs through accumulating fat mass and also because adult weight gain tends to accumulate centrally and central adiposity may be of particular importance for colorectal cancer risk, as described above.[199] Also, weight gain is simple for the general public to understand.

Most studies have analysed weight change from early adulthood to baseline, relying on self-reported weight during early adulthood.[164, 167, 168, 171, 188, 196, 197, 200, 201] In general, these studies have found that weight change is associated with colorectal cancer risk. For example, weight gain ≥20 kg from age 18/21 to baseline was associated with an increased risk of colorectal cancer compared to weight change <2 kg for men in the Health Professionals Study and women in the Nurse's Health Study.[200] Weight gain ≥20 kg compared to weight change <2 kg from age 20 to baseline was also associated with an increased risk of colorectal cancer in the EPIC study.[201]

The main issue with weight change from early adulthood is that it is unclear what the effect of weight change may be, independent of later attained weight. On average, individuals with the largest weight gain since early adulthood will have the highest weight at baseline. Also, many individuals who maintained a stable weight since early adulthood will have a normal BMI at baseline since there is a much narrower range of

BMI during early adulthood (compared to later adulthood). Only two studies presented results attempting to separate the effects of weight change since early adulthood from the effects of attained adiposity later in life.[188, 201]

Other studies have analysed weight change during later adulthood using data from prospective repeat assessments.[169, 200, 202, 203] None of these studies have found strong evidence for a relationship between weight change and colorectal cancer for men or women though one study found an increased risk for men (but not women) who gained ≥10 kg when restricting the analysis to men with BMI ≥25 kg/m^2 at baseline.[169]

Most studies investigating weight change have also included separate categories for weight loss. However, it is unclear whether weight loss is associated with a decreased risk of colorectal cancer. This is partly because the reference group in these analyses was people who maintain a stable weight which will include many people with a low BMI. Studies should focus on people with excess weight at baseline to investigate whether intentional weight loss reduces the excess risk of cancer due to excess weight. There is currently a lack of evidence for the effects of intentional weight loss on the risk of cancer.[204] Hence, given the current prevalence of obesity in many countries,[36] this represents an important research question.

Summary

Weight gain from early adulthood is associated with an increased risk of colorectal cancer. Similar to the relationship with BMI, weight gain from early adulthood is primarily associated with an increased risk of colon cancer and the association appears to be stronger for men than women.[199, 205] However, the effect of weight gain independent of attained weight in later adulthood is unknown.

2.3.6 Effect Modifiers

As mentioned above, it is unclear why different studies find contrasting results. One possible explanation is that the relationship between adiposity and colorectal cancer is modified by other risk factors. Study populations may have very different distributions

of other risk factors and if these risk factors modify the relationship between adiposity and colorectal cancer risk, this will lead to different results.

Adiposity and Menopause and Hormone Replacement Therapy Use

Early studies of BMI and colorectal cancer found that the association was stronger among younger women,[178, 206] suggesting that the relationship between BMI and colorectal cancer could be modified by changes due to menopause. Another study was able to stratify by menopausal status at baseline and found evidence for an association for premenopausal women (HR, 1.88; 95% CI, 1.24-2.86 for ≥30 vs <25 kg⁄m^2) but no association for postmenopausal women (HR, 0.73; 95% CI, 0.48-1.10).[177] This finding was supported by the Million Women Study which also found an association only for pre-menopausal women (described above).[184] Meta-analysis results have also indicated a greater effect of BMI on colorectal cancer risk for pre-menopausal women than post-menopausal women.[152]

Oestrogen may be protective against colorectal cancer. A number of prospective cohort studies found evidence for an inverse association between hormone replacement therapy (HRT) use and colorectal cancer risk.[64] Furthermore, the Women's Health Initiative (WHI), an RCT of HRT use, found a 44% decreased risk (HR, 0.56; 95% CI, 0.38-0.81) of colorectal cancer among post-menopausal women given oestrogen plus progestin versus placebo.[207] However, there is some doubt about this result since cancers in the treatment group were diagnosed at a more advanced stage[207] and there was no evidence for a lower colorectal cancer mortality in the treatment group.[208] Also, a separate RCT within the WHI found no decreased risk of colorectal cancer for women given oestrogen only (though this was based on few cases).[209]

Recent prospective studies[210-214] (though not all)[215] have generally shown that HRT use is associated with a decreased colorectal cancer risk and a recent meta-analysis of RCTs, cohort studies and case-control studies found that both combined and oestrogen only HRT use were associated with lower colorectal cancer risk.[216]

Some studies have investigated the relationship between endogenous oestrogen levels and colorectal cancer risk for post-menopausal women though the results are

inconsistent.[217-220] One study did find inverse associations between oestrogen levels and colorectal cancer risk[220] though another study indicated a positive association.[217]

After menopause, oestrogen levels decrease and adipose tissue becomes the main source of endogenous oestrogens. Consequently, BMI is positively associated with circulating oestrogen levels in post-menopausal women.[221, 222] Thus, this may explain why there is a stronger association between BMI and colorectal cancer for pre-menopausal women. It is hypothesised that, for post-menopausal women, the higher risk of colorectal cancer for women with greater adiposity is counterbalanced by the beneficial effect of excess adiposity on oestrogen levels.

However, the evidence for the effect modification by menopause status is not entirely consistent. Two cohort studies of post-menopausal women found a strong association between BMI and colorectal cancer risk[188, 223] and another study found a positive association for post-menopausal women (HR, 1.76; 95% CI, 1.13-2.74 for ≥30 vs 18.5-22.9 kg⁄m^2) but no association for pre-menopausal women (HR, 0.79; 95% CI, 0.30-2.10).[169] In the NIH-AARP Diet and Health Study, BMI was associated with colon cancer risk for women aged 50-62 and 63-66 at baseline (though there was no association for women aged 67-71).[160]

Studies have also investigated how hormone replacement therapy (HRT) use may affect the relationship between adiposity and colorectal cancer risk. The hypothesis described above would predict a stronger relationship for current users of hormone replacement therapy.[206] Evidence by HRT use is fairly scarce and not all studies support this prediction. The NIH-AARP Diet and Health Study did find that BMI was positively associated with colorectal cancer risk for current HRT users (HR, 1.13; 95% CI, 1.01-1.26 for 5 kg⁄m^2 increase) but was unrelated to cancer risk among former HRT users (HR, 1.02; 95% CI, 0.83-1.26) and never HRT users (HR, 0.97; 95% CI, 0.89-1.06).[167] Another study also found slight evidence for an interaction; there seemed to be a positive association between BMI and colorectal cancer for ever users but an inverse association for never users.[166] However, BMI was not associated with colon cancer risk for HRT users or non-users in the EPIC study and WC and WHR were both associated with colon cancer risk among HRT non-users only.[47] Two further studies did not find

clear evidence that the association between BMI and colorectal cancer risk differed by HRT use.[174, 224]

Adiposity and Physical Activity

Physical activity is inversely associated with colon cancer.[225] Given the strong relation between physical activity and adiposity, it is surprising that more studies have not investigated a possible interaction between these two factors.[168, 170, 175, 192] In a study of Swedish men, the risk of colorectal cancer was analysed according to categories of BMI and physical activity. Compared to men with the highest BMI and the least amount of physical activity, men with the lowest BMI and the most physical activity had the lowest cancer risk though there was no clear evidence for an interaction.[170] A similar analysis was performed in the Netherlands Cohort Study with WC and physical activity but there was no clear pattern, for men or women.[168]

A recent meta-analysis actually found that there was a slightly stronger association between BMI and colorectal cancer risk among studies that adjusted for physical activity.[152] Since BMI is associated with physical activity, adjusting for physical activity should result in a weaker association. It is unclear what may explain this result. The authors hypothesised that including data on physical activity could be indicative of greater study quality but is not clear how this could relate to results. For example, greater study quality could mean that BMI was measured with greater accuracy. However, since people tend to underestimate BMI,[48] results using self-reported data would overestimate the association between BMI and colorectal cancer, meaning that greater study quality would produce lower results. Greater study quality could also mean that studies included more confounders and/or that confounders were more accurately measured but this would also likely lead to lower results.

Adiposity and Smoking

Cigarette smoking is another risk factor that is associated with both colorectal cancer and BMI.[226, 227] Again, few studies have explored a potential interaction between these two risk factors. In a study of Singapore Chinese men and women, BMI was associated with colon cancer risk among never smokers but not ever smokers.[228] Another study

similarly found stronger results when restricting to never smokers[171] but Bhaskaran et al. found similar associations for BMI and colon and rectal cancer, both overall and restricting to never smokers.[161]

Summary

BMI seems to have a stronger effect on colorectal cancer risk among pre-menopausal women than among post-menopausal women. This may be explained by a protective effect of oestrogen on colorectal cancer since adipose tissue becomes an important source of endogenous oestrogen after menopause. The results according to hormone therapy use are equivocal. Evidence for interactions with physical activity and smoking are insufficient.

2.3.7 Mechanisms

The exact mechanisms relating adiposity and excess weight with increased colorectal cancer risk remain unclear. Insulin resistance, a condition where the body uses insulin less effectively and which results in elevated levels of insulin (and also insulin-like growth factor (IGF)-1), may be the most important factor relating adiposity to colorectal cancer development. Adiposity is strongly associated with insulin resistance and insulin and IGF-1 have been shown to promote cell proliferation and inhibit apoptosis.[157, 158, 185, 229] The role of insulin resistance is also supported by the evidence of an increased risk of colorectal cancer for men and women with type 2 diabetes.[230] Adipose tissue also produces various hormones and cytokines, known collectively as adipokines or adipocytokines, which have been related to cancer development.[158, 185] Another possible mechanism relates to inflammation, supported by evidence that individuals with chronic inflammatory bowel disease have a higher risk of colorectal cancer and that aspirin and other anti-inflammatory drugs are associated with a lower risk.[157]

2.3.8 Summary

Substantial evidence indicates an association between BMI and colorectal cancer risk. The evidence is strongest for colon cancer and for men. The mechanisms underlying the stronger association for men remain unclear. BMI is a simple measure of weight

adjusted for height but does not take into account fat distribution. WC and WHR are also associated with increased colorectal cancer risk though it remains unclear what the most appropriate measure of adiposity is for predicting colorectal cancer risk. Early adulthood adiposity and weight gain may both be associated with increased colorectal cancer risk but their effects independent of attained adiposity are unknown. The effect of BMI on colorectal cancer risk may be stronger among pre-menopausal women than post-menopausal women.

Table 2.3.6 Prospective Cohort Studies of Adiposity and Colorectal Cancer

Ref.	Author, year	Study name, location	Population	Assessment method[a]	Follow-up period (average follow-up)	Age	Sex[b]	Number of subjects	Comparison[c]	Site[d]	Number of cases	HR/RR (95% CI)
161	Bhaskaran 2014	UK	General population	M	1987 - 2012	≥16	B	5,243,978	BMI, ≥35.0 vs 18.5-24.9 kg/m^2	CC	13,465	1.36 (1.23-1.51)*
										RC	6,123	1.18 (1.01-1.38)*
197	Han 2014	Atherosclerosis Risk in Communities, USA	General population	M	1987 - 2006	45-64	M	6,332	BMI, ≥30 vs 18.5-<25.0 kg/m^2	CRC	151	1.14 (0.91-1.44)
							F	7,569	BMI, ≥30 vs 18.5-<25.0 kg/m^2	CRC	147	1.01 (0.80-1.26)
202	Steins Bisschop 2014	European Prospective Investigation into Cancer–Physical Activity, Nutrition, Alcohol, Cessation of Smoking, Eating Study, Europe	General population	S	1992 - 2008 (6.8 yrs)	25-70	M	91,231	BMI, >30 vs <25 kg/m^2	CC	480	1.49 (1.13-1.97)
										RC	354	1.13 (0.81-1.58)
							F	237,550	BMI, >30 vs <25 kg/m^2	CC	781	1.24 (0.99-1.56)
										RC	393	1.09 (0.78-1.51)

Ref.	Author, year	Study name, location	Population	Assessment method[a]	Follow-up period (average follow-up)	Age	Sex[b]	Number of subjects	Comparison[c]	Site[d]	Number of cases	HR/RR (95% CI)
186	Keimling 2013	NIH-AARP Diet & Health Study, USA	Membership organisation of over 50s	S	1996 - 2006 (10 yrs)	50-72	M	124,208	BMI, Q5 vs Q1	CC	1,463	1.42 (1.19-1.68)
										RC	536	0.90 (0.68-1.19)
									WC, ≥106.5 vs <89.5 cm	CC	1,463	1.45 (1.16-1.82)
										RC	536	0.97 (0.67-1.38)
									WHR, Q5 vs Q1	CC	1,463	1.29 (1.10-1.52)
										RC	536	1.08 (0.82-1.43)
							F	78,969	BMI, Q5 vs Q1	CC	680	0.96 (0.74-1.23)
										RC	190	1.26 (0.79-2.01)
									WC, ≥94.5 vs <73.6 cm	CC	680	0.90 (0.63-1.27)
										RC	190	1.01 (0.53-1.94)
									WHR, Q5 vs Q1	CC	680	0.90 (0.70-1.15)
										RC	190	1.13 (0.69-1.86)
166	Kitahara 2013	Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial, USA	General population	S	1993 - 2009 (11.9 yrs)	55-74	M	36,912	BMI, ≥30 vs 18.5-24.9 kg/m^2	PCC	275	1.48 (1.05-2.09)
										DCC	131	1.48 (0.90-2.42)
										RC	134	1.38 (0.83-2.27)
							F	37,562	BMI, ≥30 vs 18.5-24.9 kg/m^2	PCC	254	1.23 (0.89-1.69)
										DCC	88	0.66 (0.36-1.21)
										RC	66	0.95 (0.50-1.79)

Ref.	Author, year	Study name, location	Population	Assessment method[a]	Follow-up period (average follow-up)	Age	Sex[b]	Number of subjects	Comparison[c]	Site[d]	Number of cases	HR/RR (95% CI)
164	Li 2013	Shanghai Men's Health Study, China	General population	M	2002 - 2009 (5.5 yrs)	40-74	M	61,283	BMI, ≥26.2 vs <21.1 kg/m^2	CC	180	2.15 (1.35-3.43)
										RC	133	1.20 (0.69-2.10)
									WC, ≥92 vs <78 cm	CC	180	2.00 (1.21-3.29)
										RC	133	0.88 (0.52-1.49)
									WHR, ≥0.95 vs <0.85	CC	180	1.97 (1.19-3.24)
										RC	133	1.24 (0.69-2.26)
		Shanghai Women's Health Study, China	General population	M	1997 - 2009 (11.0 yrs)	40-70	F	72,972	BMI, ≥26.7 vs <21.1 kg/m^2	CC	382	1.00 (0.70-1.43)
										RC	240	1.22 (0.78-1.90)
									WC, ≥85 vs <70 cm	CC	382	1.34 (0.89-2.00)
										RC	240	1.17 (0.73-1.88)
									WHR, ≥0.85 vs <0.77	CC	382	0.96 (0.69-1.34)
										RC	240	1.11 (0.74-1.66)
167	Renehan 2012	NIH-AARP Diet & Health Study, USA	Membership organisation of over 50s	S	1996 - 2006 (10 yrs)	50-72	M	168,294	BMI, ≥35 vs 18.5-21.9 kg/m^2	CC	2,070	1.53 (1.16-2.03)
										RC	762	1.43 (0.90-2.28)
							F	105,385	BMI, ≥35 vs 18.5-21.9 kg/m^2	CC	962	1.23 (0.93-1.64)
										RC	282	1.28 (0.76-2.16)

Ref.	Author, year	Study name, location	Population	Assessment method[a]	Follow-up period (average follow-up)	Age	Sex[b]	Number of subjects	Comparison[c]	Site[d]	Number of cases	HR/RR (95% CI)
168	Hughes 2011	Netherlands Cohort Study, Netherlands	General population	S	1986 - 2002 (16.3 yrs)	55-69	M	1,365†	BMI, ≥27.1 vs <23.0 kg/m^2	PCC	327	1.35 (0.90-1.98)
										DCC	427	1.38 (0.95-1.98)
										RC	299	1.01 (0.67-1.51)
									WC, Q5 vs Q1	PCC	327	1.32 (0.81-2.15)
										DCC	427	2.56 (1.55-4.24)
										RC	299	1.33 (0.77-1.29)
							F	1,832†	BMI, ≥27.6 vs <22.1 kg/m^2	PCC	459	0.91 (0.65-1.28)
										DCC	327	1.04 (0.72-1.50)
										RC	205	1.07 (0.67-1.60)
									WC, Q5 vs Q1	PCC	459	1.46 (0.98-2.18)
										DCC	327	1.15 (0.74-1.80)
										RC	205	1.07 (0.59-1.93)

Ref.	Author, year	Study name, location	Population	Assessment method[a]	Follow-up period (average follow-up)	Age	Sex[b]	Number of subjects	Comparison[c]	Site[d]	Number of cases	HR/RR (95% CI)
163	Matsuo 2011	Japan Public Health Center-based Prospective Study I	General population	S	1990 - 2006	40-59	M	157,927	BMI, ≥30 vs 23-<25 kg/m^2	CC	1,919	1.37 (0.96–1.98)
					1993 - 2006	40-69				RC	1,111	1.85 (1.23–2.78)
		Japan Public Health Center-based Prospective Study II			1988 - 2001	40-79	F	183,457	BMI, ≥30 vs 23-<25 kg/m^2	CC	1,534	1.39 (1.02–1.90)
					1990 - 2003	40-64				RC	735	1.33 (0.82–2.15)
		Japan Collaborative Cohort Study			1984 - 1992	≥40						
		Miyagi Cohort Study-I			1985 - 2000	40-103						
		Miyagi Cohort Study-II			1992 - 1999	≥35						
		Aichi Cohort Study			1994 - 2003	40-79						
		Ohsaki Cohort Study										
		Takayama Study, Japan, Japan										
228	Odegaard 2011	Singapore Chinese Health Study, Singapore	General population	S	1993 - 2007 (11.5 yrs)	45-74	B	51,251	BMI, ≥27.5 vs 21.5-24.4 kg/m^2	CC	596	1.48 (1.13-1.92)
										RC	384	0.93 (0.64-1.36)

Ref.	Author, year	Study name, location	Population	Assessment method[a]	Follow-up period (average follow-up)	Age	Sex[b]	Number of subjects	Comparison[c]	Site[d]	Number of cases	HR/RR (95% CI)
138	Shin 2011	Korea	General population	M	1996 - 2003	30-80	M	869,725	BMI, ≥25 vs 18.5-22.9 kg/m^2	PCC	536	1.1 (0.9-1.4)
										DCC	751	1.6 (1.4-2.0)
										RC	1,535	1.1 (1.0-1.3)
							F	395,501	BMI, ≥25 vs 18.5-22.9 kg/m^2	PCC	236	1.5 (1.1-2.0)
										DCC	225	1.3 (0.9-1.8)
										RC	451	1.1 (0.9-1.3)
196	Bassett 2010	Melbourne Collaborative Cohort Study, Australia	General population	M	1990 - 2007 (14.0 yrs)	27-80	M	16,188	BMI, ≥30 vs 23-<25 kg/m^2	CC	277	1.51 (1.00-2.28)
							F	23,438	BMI, ≥30 vs 23-<25 kg/m^2	CC	292	1.00 (0.70-1.44)
190	Burton 2010	Glasgow Alumni Cohort, Scotland	University students	M	1948 - 2008 (47.5 yrs)	<30	B	12,206	BMI, >25 vs 19-23 kg/m^2	CC	71	1.12 (0.44–2.82)
										RC	41	1.46 (0.51–4.61)
169	Laake 2010	Norwegian Counties Study, Norway	General population	M	1974 - 2005 (23.2 yrs)	20-49	M	38,822	BMI, ≥30 vs 18.5-<23 kg/m^2	CC	450	1.80 (1.25-2.59)
							F	37,357	BMI, ≥30 vs 18.5-<23 kg/m^2	CC	419	1.48 (1.09-2.02)

Ref.	Author, year	Study name, location	Population	Assessment method[a]	Follow-up period (average follow-up)	Age	Sex[b]	Number of subjects	Comparison[c]	Site[d]	Number of cases	HR/RR (95% CI)
188	Oxentenko 2010	Iowa Women's Health Study, USA	General population	S	1986 - 2005	55-69	F	36,941	BMI, ≥40 vs 18.5–24.9 kg/m^2	CRC	1,464	1.56 (1.10-2.22)
									WC, ≥96.53 vs ≤77.15 cm	CRC	1,464	1.32 (1.11-1.56)
									WHR, ≥0.90 vs ≤0.78	CRC	1,464	1.28 (1.08-1.50)
231	Cnattingius 2009	Sweden	Twins	S	1973 - 2006	15-47	B	23,337	BMI, ≥25.0 vs 18.5-24.9 kg/m^2	CRC	210	1.60 (1.15-2.23)
217	Gunter 2008	Women's Health Initiative Observational Study, USA	General population	M	1993 - 2004 (6.4 yrs)	50-79	F	809†	BMI, ≥30 vs 18.5-<25.0 kg/m^2	CRC	438	1.55 (1.13–2.13)
									WC, ≥93.0 vs 75.0 cm	CRC	438	1.47 (1.04–2.09)
									WHR, ≥0.85 vs <0.75	CRC	438	1.82 (1.22–2.70)
165	Jee 2008	Korea	General population	M	1992 - 2006 (10.8 yrs)	30-95	M	770,556	BMI, ≥30 vs 23.0–24.9 kg/m^2	CC	4,671	1.42 (1.02-1.98)
										RC	4,032	1.16 (0.77-1.74)
							F	443,273	BMI, ≥30 vs 23.0–24.9 kg/m^2	CC	1,959	1.01 (0.72-1.42)
										RC	1,681	1.14 (0.78-1.68)
223	Song 2008	Korea	General population	M	1994 - 2003 (8.8 yrs)	40-64	F	152,772	BMI, ≥30 vs 21.0-22.9 kg/m^2	CC	453	2.18 (1.43-3.33)
										RC	482	0.91 (0.55-1.52)

Ref.	Author, year	Study name, location	Population	Assessment method[a]	Follow-up period (average follow-up)	Age	Sex[b]	Number of subjects	Comparison[c]	Site[d]	Number of cases	HR/RR (95% CI)
187	Wang 2008	Cancer Prevention Study-II Nutrition Cohort, USA	General population	S	1997 - 2005 (7.7 yrs)		M	44,068	BMI, ≥35.0 vs 18.5-24.9 kg/m^2	CC	402	1.93 (1.14-3.28)
										RC	142	1.38 (0.58-3.28)
									WC, ≥120 vs <95 cm	CC	402	2.05 (1.29-3.25)
										RC	142	1.02 (0.43-2.42)
							F	51,083	BMI, ≥35.0 vs 18.5-24.9 kg/m^2	CC	314	1.40 (0.84-2.36)
										RC	93	2.67 (1.09-6.54)
									WC, ≥110 vs <85 cm	CC	314	1.54 (1.00-2.37)
										RC	93	2.65 (1.23-5.71)
82	Akhter 2007	Miyagi Cohort Study, Japan	General population	S	1990 - 2001 (11 yrs)	40-64	B	21,199	≥25.0 vs <18.5 kg/m^2	CRC	307	1.61 (0.59-4.40)
160	Adams 2007	NIH-AARP Diet & Health Study, USA	Membership organisation of over 50s	S	1995 - 2000 (5 yrs)	50-71	M	307,708	BMI, ≥40 vs 18.5-<23 kg/m^2	CC	1,676	2.39 (1.59-3.58)
									BMI, ≥35 vs 18.5-<23 kg/m^2	RC	677	1.00 (0.68-1.58)
							F	209,436	BMI, ≥40 vs 18.5-<23 kg/m^2	CC	769	1.49 (0.98-2.25)
									BMI, ≥35 vs 18.5-<23 kg/m^2	RC	278	1.44 (0.92-2.25)
232	Driver 2007	Physicians' Health Study, USA	Physicians	S	1982 - 2004	40-84	M	21,581	BMI, ≥25 vs <25 kg/m^2	CC	381	1.38 (1.11-1.70)
										RC	104	1.19 (0.80-1.77)

Ref.	Author, year	Study name, location	Population	Assessment method[a]	Follow-up period (average follow-up)	Age	Sex[b]	Number of subjects	Comparison[c]	Site[d]	Number of cases	HR/RR (95% CI)
233	Lundqvist 2007	2 cohorts from Sweden and 1 cohort from Finland	Twins	S	1969 - 2004 (22.0 yrs)	43-96	B	24,821	BMI, ≥30 vs 18.5-<25.0 kg/m^2	CC	513	1.3 (0.9-1.8)
										RC	324	0.7 (0.4-1.2)
					1973 - 2004 (28.4 yrs)	18-47	B	43,328	BMI, ≥30 vs 18.5-<25.0 kg/m^2	CC	204	1.1 (0.5-2.5)
										RC	154	0.9 (0.3-2.5)
184	Reeves 2007	Million Women Study, UK	General population	S	1996 - 2004 (5.4 yrs)	50-64	F	1,222,630	BMI, ≥30 vs 22.5-24.9 kg/m^2	CRC	4,008	1.01 (0.94-1.09)
224	Wang 2007	Cancer Prevention Study-II Nutrition Cohort, USA	General population	S	1992 - 2003		F	73,842	BMI, ≥30 vs 18.5–24.9 kg/m^2	CRC	814	1.19 (0.97-1.45)
234	Ahmed 2006	Atherosclerosis Risk in Communities Study, USA	General population	M	1987 - 2000 (11.5 yrs)	45-64	B	14,109	BMI, ≥35 vs <25 kg/m^2	CRC	194	1.54 (0.9-2.8)
									WC, ≥102/88 vs <102/88 cm	CRC	194	1.40 (1.0-1.9)
									WHR, ≥0.98 vs. <0.88	CRC	194	1.67 (1.1-2.5)

Ref.	Author, year	Study name, location	Population	Assessment method[a]	Follow-up period (average follow-up)	Age	Sex[b]	Number of subjects	Comparison[c]	Site[d]	Number of cases	HR/RR (95% CI)
235	Bowers 2006	Alpha-Tocopherol, Beta-Carotene Cancer Prevention Study, Finland	Smokers	M	1985 - 2002 (14.1 yrs)		M	28,983	BMI, ≥30 vs 18.5-<25 kg/m^2	CC	227	1.78 (1.25-2.55)
										RC	183	1.51 (0.99-2.29)
170	Larsson 2006	Cohort of Swedish Men, Sweden	General population	S	1997 - 2005 (7.1 yrs)	45-79	M	45,906	BMI, ≥30 vs <23 kg/m^2	CC	309	1.60 (1.03–2.48)
										RC	190	1.44 (0.79–2.61)
									WC, ≥104 vs <88 cm	CC	309	1.44 (0.93–2.24)
										RC	190	1.24 (0.68–2.25)
236	Lukanova 2006	Northern Sweden Health and Disease Cohort, Sweden	General population	M	1985 - 2003 (8.2 yrs)	30-60	M	33,424	BMI, ≥30 vs 18.5-24.9 kg/m^2	CC	73	1.43 (0.62-3.02)
										RC	58	1.96 (0.96-3.86)
							F	35,362	BMI, ≥30 vs 18.5-24.9 kg/m^2	CC	76	2.25 (1.25-3.98)
										RC	31	1.30 (0.42-3.45)

Ref.	Author, year	Study name, location	Population	Assessment method[a]	Follow-up period (average follow-up)	Age	Sex[b]	Number of subjects	Comparison[c]	Site[d]	Number of cases	HR/RR (95% CI)
237	MacInnis 2006	Melbourne Collaborative Cohort Study, Australia	General population	M	1990 - 2003 (10.3 yrs)	27-75	M	16,867	BMI, ≥30 vs <25 kg/m^2	RC	134	1.3 (0.8-2.4)
									WC, ≥102 vs <94 cm	RC	134	1.4 (0.9-2.2)
									WHR, ≥0.95 vs <0.90	RC	134	1.2 (0.8-1.8)
							F	24,247	BMI, ≥30 vs <25 kg/m^2	RC	95	1.1 (0.7-1.9)
									WC, ≥88 vs <80 cm	RC	95	1.4 (0.8-2.2)
									WHR, ≥0.80 vs <0.75	RC	95	1.4 (0.8-2.4)
172	MacInnis 2006	Melbourne Collaborative Cohort Study, Australia	General population	M	1990 - 2003 (10.4 yrs)	27-75	F	24,072	BMI, ≥30 vs <25 kg/m^2	CC	212	1.0 (0.7–1.4)
									WC, ≥88 vs <80 cm	CC	212	1.4 (1.0–1.9)
									WHR, ≥0.80 vs <0.75	CC	212	1.7 (1.1–2.4)

Ref.	Author, year	Study name, location	Population	Assessment method[a]	Follow-up period (average follow-up)	Age	Sex[b]	Number of subjects	Comparison[c]	Site[d]	Number of cases	HR/RR (95% CI)
47	Pischon 2006	European Prospective Investigation into Cancer and Nutrition (EPIC), Europe	General population	M	1992 - 2004 (6.1 yrs)	25-70	M	129,731	BMI, ≥29.4 vs <23.6 kg/m^2	CC	421	1.55 (1.12-2.15)
										RC	295	1.05 (0.72-1.55)
									WC, ≥103.0 vs <86.0 cm	CC	421	1.39 (1.01-1.93)
										RC	295	1.27 (0.84-1.91)
									WHR, ≥0.990 vs <0.887	CC	421	1.51 (1.06-2.15)
										RC	295	1.93 (1.19-3.13)
							F	238,546	BMI, ≥28.9 vs <21.7 kg/m^2	CC	563	1.06 (0.79-1.42)
										RC	291	1.06 (0.71-1.58)
									WC, ≥89.0 vs <70.2 cm	CC	563	1.48 (1.08-2.03)
										RC	291	1.23 (0.81-1.86)
									WHR, ≥0.846 vs <0.734	CC	563	1.52 (1.12-2.05)
										RC	291	1.20 (0.81-1.79)
171	Samanic 2006	Swedish Foundation for Occupational Safety and Health of the Construction Industry, Sweden	Constructio n workers	M	1971 - 1999 (19 yrs)	18-67	M	362,552	BMI, ≥30 vs 18.5-24.9 kg/m^2	CC	1,795	1.74 (1.48-2.04)
										RC	1,362	1.36 (1.13-1.66)
142	Yeh 2006	Taiwan	General population	M	1990 - 2001	30-65	M	10,923	BMI, >28.6 vs <24.2 kg/m^2	CRC	68	1.98 (0.91-4.30)

Ref.	Author, year	Study name, location	Population	Assessment method[a]	Follow-up period (average follow-up)	Age	Sex[b]	Number of subjects	Comparison[c]	Site[d]	Number of cases	HR/RR (95% CI)
238	Engeland 2005	Norway	General population	M	1963 - 2001 (23 yrs)	20-74	M	962,901	BMI, ≥30 vs 18.5–24.9 kg/m^2	CC	13,805	1.49 (1.39-1.60)
										RC	9,182	1.27 (1.16-1.38)
							F	1,037,077	BMI, ≥30 vs 18.5–24.9 kg/m^2	CC	16,638	1.07 (1.02-1.12)
										RC	7,492	1.04 (0.97-1.11)
239	Otani 2005	Japan Public Health Center-based Prospective Study, Japan	General population	S	1990 - 2001	40-69	M	49,158	BMI, ≥30 vs <25 kg/m^2	CC	424	1.4 (0.7-2.8)
										RC	202	1.6 (0.6-3.9)
							F	53,791	BMI, ≥30 vs <25 kg/m^2	CC	229	0.5 (0.2-1.4)
										RC	131	1.3 (0.5-3.1)
240	Rapp 2005	Vorarlberg Health Monitoring and Promotion Program Study Cohort, Austria	General population	M	1985 - 2002 (9.9 yrs)	≥19	M	67,447	BMI, ≥35 vs 18.5-24.9 kg/m^2	CC	260	2.48 (1.15-5.39)
									BMI, ≥30 vs 18.5-24.9 kg/m^2	RC	138	1.66 (1.01-2.73)
							F	78,484	BMI, ≥35 vs 18.5-24.9 kg/m^2	CC	271	0.88 (0.43-1.81)
										RC	133	0.96 (0.38-2.39)
174	Lin 2004	Women's Health Study, USA	Health professionals	S	1993 - 2003 (8.7 yrs)	≥45	F	37,671	BMI, ≥30 vs <23 kg/m^2	CC	158	1.73 (1.05-2.85)
										RC	40	1.55 (0.64-3.77)

Ref.	Author, year	Study name, location	Population	Assessment method[a]	Follow-up period (average follow-up)	Age	Sex[b]	Number of subjects	Comparison[c]	Site[d]	Number of cases	HR/RR (95% CI)
173	MacInnis 2004	Melbourne Collaborative Cohort Study, Australia	General population	M	1990 - 2002	27-75	M	16,556	BMI, >29.2 vs 24.8 kg/m^2	CC	153	1.7 (1.1-2.8)
									WC, >99.3 vs <87.0 cm	CC	153	2.1 (1.3-3.5)
									WHR, >0.96 vs <0.88	CC	153	2.1 (1.3-3.4)
175	Moore 2004	Framingham Study, USA	General population	M	1948 - 1999	30-54	M	1,684	BMI, ≥30 vs 18.5-<25 kg/m^2	CC	71	2.0 (0.98-4.2)
									WC, ≥101.6 vs <83.8 cm	CC	71	2.4 (0.99-5.7)
							F	2,080	BMI, ≥30 vs 18.5-<25 kg/m^2	CC	86	1.3 (0.65-2.7)
									WC, ≥99.1 vs <81.3 cm	CC	86	1.8 (0.78-4.3)
						55-79	M	1,661	BMI, ≥30 vs 18.5-<25 kg/m^2	CC	69	3.7 (1.7-8.1)
									WC, ≥101.6 vs <83.8 cm	CC	69	3.3 (1.3-8.8)
							F	2,141	BMI, ≥30 vs 18.5-<25 kg/m^2	CC	70	1.9 (0.98-3.7)
									WC, ≥99.1 vs <81.3 cm	CC	70	2.3 (0.86-6.3)

Ref.	Author, year	Study name, location	Population	Assessment method[a]	Follow-up period (average follow-up)	Age	Sex[b]	Number of subjects	Comparison[c]	Site[d]	Number of cases	HR/RR (95% CI)
144	Sanjoaquin 2004	Oxford Vegetarian Study, UK	General population	S	1980 - 1999 (17 yrs)	16-89	B	10,998	BMI, ≥25 vs <20 kg/m^2	CRC	92	0.74 (0.36-1.53)
78	Wei 2004	Health Professionals Follow-up Study, USA	Health professionals	S	1986 - 2000 (13 yrs)	40-75	M	46,632	BMI, ≥30 vs <23 kg/m^2	CC	467	1.85 (1.26-2.72)
										RC	135	1.03 (0.49-2.14)
		Nurses' Health Study, USA	Nurses	S	1980 - 2000 (20 yrs)	34-59	F	87,733	BMI, ≥30 vs <23 kg/m^2	CC	672	1.28 (1.10-1.62)
										RC	204	1.56 (1.01-2.42)
176	Saydah 2003	CLUE II Cohort, USA	General population	S	1989 - 2000	≥18	B	346‡	BMI, ≥30 vs <25 kg/m^2	CC	132	1.79 (1.02-3.13)
										RC	41	1.64 (0.68-3.94)
146	Shimizu 2003	Takayama Study, Japan	General population	S	1993 - 2000	≥35	M	13,392	BMI, ≥23.6 vs ≤21.2 kg/m^2	CC	104	2.11 (1.26–3.53)
										RC	58	0.83 (0.42–1.64)
							F	15,659	BMI, ≥23.1 vs ≤21.6 kg/m^2	CC	89	1.22 (0.69–2.15)
										RC	41	0.83 (0.35–1.99)
177	Terry 2002	National Breast Screening Study, Canada	General population	S	1980 - 1993 (10.6 yrs)	40-59	F	89, 835	BMI, ≥30 vs <25 kg/m^2	CC	363	0.95 (0.67–1.34)
										RC	164	1.35 (0.87–2.07)

Ref.	Author, year	Study name, location	Population	Assessment method[a]	Follow-up period (average follow-up)	Age	Sex[b]	Number of subjects	Comparison[c]	Site[d]	Number of cases	HR/RR (95% CI)
241	Nilsen 2001	Norway	General population	M	1984 - 1996 (10.8 yrs)	≥20	M	36,975	≥27.2 vs ≤23.0 kg/m^2	CRC	354	1.07 (0.80–1.42)
							F	38,244	≥27.5 vs ≤21.8 kg/m^2	CRC	358	0.98 (0.71–1.34)
178	Terry 2001	Sweden	General population	S	1987 - 1998 (9.6 yrs)	40-76	F	61,463	BMI, >26.7 vs <22.0 kg/m^2	CC	291	1.21 (0.86-1.70)
										RC	159	1.32 (0.83-2.08)
242	Kaaks 2000	New York University Women's Health Study, USA	Women attending mammography screening	S	1985 - 1998	35-65	F	134‡	BMI, Q5 vs Q1	CC	73	3.07 (1.12-8.41)
243	Ford 1999	National Health and Nutrition Examination Survey, USA	General population	M	1971 - 1992	25-74	M	5,506	BMI, ≥30 vs <22 kg/m^2	CC	104	2.95 (0.99-8.74)
							F	7,914	BMI, ≥30 vs <22 kg/m^2	CC	118	2.74 (1.04-7.25)
179	Robsahm 1999	Norway	Tuberculosis screening programme		1963 - 1989	30-69	M	532,300	BMI, Q5 vs Q1	CC	6,397	1.39 (1.39-1.50)
										RC	4,393	1.16 (1.07-1.27)
							F	590,552	BMI, Q5 vs Q1	CC	7,620	1.07 (0.99-1.15)
										RC	3,482	1.03 (0.93-1.14)

Ref.	Author, year	Study name, location	Population	Assessment method[a]	Follow-up period (average follow-up)	Age	Sex[b]	Number of subjects	Comparison[c]	Site[d]	Number of cases	HR/RR (95% CI)
244	Schoen 1999	Cardiovascular Health Study, USA	General population	M	1989 - 1996 (6.4 yrs)	≥65	B	5,849	BMI, ≥28.5/29.6 vs <23.9/23.2 kg/m^2	CRC	102	1.4 (0.8-2.5)
									WC, ≥104.1/101.2 vs <91.0/82.0 cm	CRC	102	2.2 (1.2-4.1)
									WHR, ≥1.01/0.96 vs <0.83/0.83	CRC	102	2.6 (1.4-4.8)
147	Singh 1998	Adventist Health Study, USA	Seventh-day Adventists	S	1976 - 1982 (6 yrs)	≥25	M		BMI, >25.6 vs <22.5 kg/m^2	CC	65	2.63 (1.12-6.13)
							F		BMI, >25.6 vs <22.5 kg/m^2	CC	92	1.05 (0.63-1.75)
245	Gaard 1997	Norway	General population	M	1977 - 1991 (11.3 yrs)	20-54	M	31,507	BMI, ≥26.6 vs <21.7 kg/m^2	CC	103	1.64 (0.92-2.92)
										RC	54	1.61 (0.76-3.44)
							F	30,666	BMI, ≥27.1 vs <23.0 kg/m^2	CC	83	1.02 (0.53-1.97)
										RC	52	0.64 (0.31-1.33)

Ref.	Author, year	Study name, location	Population	Assessment method[a]	Follow-up period (average follow-up)	Age	Sex[b]	Number of subjects	Comparison[c]	Site[d]	Number of cases	HR/RR (95% CI)
180	Martinez 1997	Nurses' Health Study, USA	Nurses	S	1980 - 1992 (12 yrs)	34-59	F	89,488	BMI, ≥29 vs <21 kg/m^2	CC	393	1.45 (1.02-2.07)
									WC, ≥86.4 vs <69.9 cm	CC	161	1.48 (0.89-2.46)
									WHR, ≥0.83 vs <0.73	CC	161	1.48 (0.88-2.49)
246	Tulinius 1997	Cardiovascular Risk Factor Study, Iceland	General population	M	1968-1995	33-60	M	11,366	BMI, per 1 kg/m^2	CRC	193	1.04 (1.00-1.08)
74	Chyou 1996	Honolulu Heart Program, USA	General population	M	1965 - 1995	46-68	M	7,945	BMI, ≥25.8 vs <21.7 kg/m^2	CC	330	1.38 (1.01-1.90)
										RC	123	0.63 (0.38-1.04)
247	Thune 1996	Finland	General population	M	1972-1991	20-49	M	53,242	BMI, per 10 kg/m^2	CC	236	1.25 (1.01-1.53)
										RC	170	0.99 (0.60-1.63)
							F	28,274	BMI, per 10 kg/m^2	CC	99	0.93 (0.57-1.52)
										RC	58	0.96 (0.51-1.82)

Ref.	Author, year	Study name, location	Population	Assessment method[a]	Follow-up period (average follow-up)	Age	Sex[b]	Number of subjects	Comparison[c]	Site[d]	Number of cases	HR/RR (95% CI)
248	Giovannucci 1995	Health Professionals Follow-up Study, USA	Health professionals	S	1986 - 1992 (5 yrs)	40-75	M	47,723	BMI, ≥29 vs <22 kg/m^2	CC	203	1.48 (0.89-2.46)
									WC, ≥109 vs <89 cm	CC	117	2.56 (1.33-4.96)
									WHR, ≥0.99 vs <0.90	CC	117	3.41 (1.52-7.66)
191	Le Marchand 1992	Hawaii	General population	S	1972-1986	45-69	M	52,539	BMI, Q3 vs Q1	CC	421	1.4 (1.1-1.8)
										RC	203	0.8 (0.5-1.2)
192	Lee 1992	Harvard Alumni Health Study, USA	University alumni	S	1962 -1988		M	17,595	BMI, ≥26.0 vs <22.5 kg/m^2	CC	290	1.52 (1.06-2.17)
75	Klatsky 1988	USA	Members of health plan	M	1978 - 1984		B	10,303	BMI, per 0.1 kg/m^2	CC	203	1.04 (1.02-1.06)
										RC	66	1.00 (0.96-1.04)
76	Wu 1987	USA	Retirement community	S	1981 - 1985		M	4,141	BMI, ≥35 vs ≤31 kg/m^2	CRC	58	2.40 (1.1-5.4)
							F	7,421	BMI, ≥34 vs ≤29 kg/m^2	CRC	68	1.19 (0.7-2.2)

[a] S = self-reported, M = measured

[b] M = male, F = female, B = both

c BMI = body mass index, WC = waist circumference, WHR = waist to hip ratio

[d] CRC = colorectal cancer, CC = colon cancer, PCC = proximal colon cancer, DCC = distal colon cancer, RC = rectal cancer

* 99% CIs

† Case-cohort analysis

‡ Nested case-control study

2.4 Smoking and Colorectal Cancer Literature Review

In this section, the evidence for an association between tobacco smoking and colorectal cancer risk is reviewed. This association has been controversial. Many early studies did not find that smoking was associated with colorectal cancer risk.[249] Recent studies have more consistently shown an association though the increased risk of colorectal cancer is fairly modest. In their 2004 monograph on tobacco smoke and the risk of cancer (based on the evidence available through 2002), the IARC stated that it was not possible to conclude that tobacco smoking is causally associated with colorectal cancer risk, mainly due to concerns about confounding by other risk factors.[250] However, since that report, a number of prospective cohort studies have continued to support a relationship between smoking and colorectal cancer and, consequently, in their 2012 monograph, it was concluded that smoking is causally related to colorectal cancer risk.[26]

Exposure to tobacco smoke is usually assessed in epidemiological studies through the use of questionnaires. Exposure can be defined in a number of ways. The simplest measure of exposure, referred to as smoking status, classifies people as never, former or current smokers. Other measures attempt to classify people more precisely according to different aspects of exposure such as smoking intensity (number of cigarettes smoked per day), smoking duration, pack-years of smoking or age at initiation. Time since cessation and age at cessation are used to try to measure how the risk differs for former smokers who quit smoking at different times.

2.4.1 Search Strategy

The IARC published reports in 2004 and 2012[26, 250] and the U.S. Department of Health and Human Services (USDHHS) published reports in 2004 and 2014[29, 251] summarising the evidence from prospective cohort studies for an association between smoking and colorectal cancer. The USDHHS searched the Medline database as well as Web of Science and Embase for articles published up to December 2009.[29] Details on the literature search were not provided in the IARC reports though they included articles published in 2009.[26] Recent meta-analyses[49, 50, 61, 226, 252] and review articles[64] of smoking and colorectal cancer risk were also searched for relevant articles.

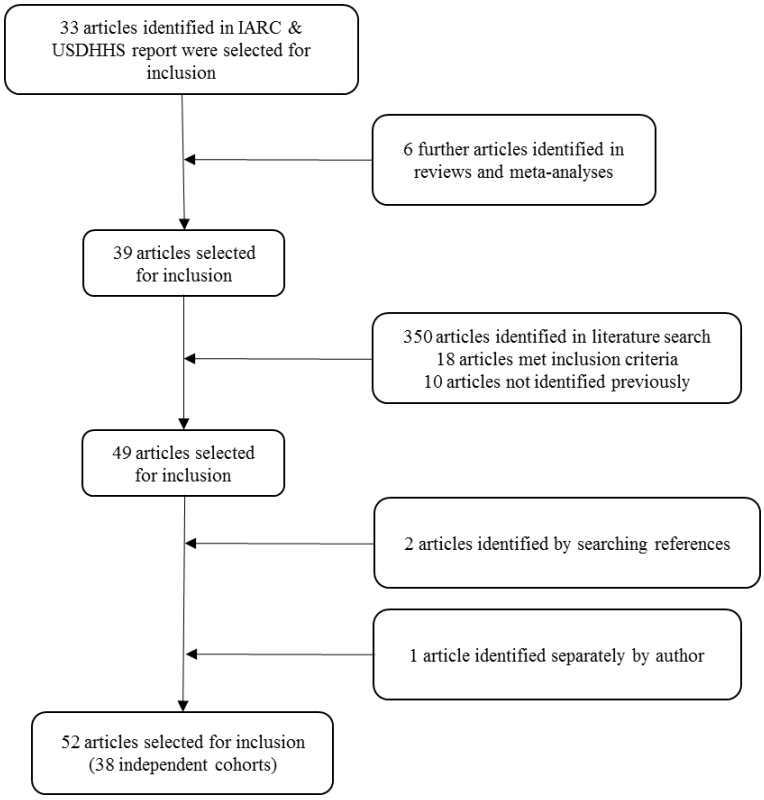
To complement the articles identified by these reports and meta-analyses, PubMed was searched for articles reporting human studies, written in English, and published between 1st January 2008 and 31st March 2015, using the MeSH terms "Colorectal Neoplasms" and at least one of "Tobacco Use", "Tobacco" or "Smoking Cessation". 350 articles were retrieved.

Articles were selected for inclusion in this review if they (i) were based on prospective cohort studies (including nested case-control studies), (ii) analysed smoking status (separating current and former smokers), smoking duration, smoking intensity, pack-years, age at initiation, time since cessation or age at cessation in relation to colorectal cancer incidence (including colorectal subsites) in the general population and (iii) reported risk estimates (HRs, RRs, ORs) and CIs. Analyses that reported results by other factors were not included. Analyses based on fewer than 30 cases were ignored. References of included articles were carefully examined for additional articles.

This literature review excluded analyses solely focused on forms of tobacco use other than cigarettes (e.g. cigar or pipe smokers). However, articles generally did not provide detailed information on the questionnaires used and so it is not always clear how smokers of other forms of tobacco were defined in analyses. For example, a cigar smoker may be excluded from the analysis, they may be defined as a current smoker or they may be defined as a former smoker or never smoker based on previous cigarette use.

Figure 2.4.1 shows a flow diagram of the literature search process, illustrating the number of articles identified from the different sources. Overall, 52 articles met the inclusion criteria for this literature review. These 52 articles were based on 38 independent cohorts.

Figure 2.4.1 Flowchart of Literature Search Process for Smoking and Colorectal Cancer



IARC = International Agency for Research and Cancer, USDHHS = U.S. Department of Health and Human Services

Table 2.4.2, presented at the end of this review shows the key characteristics of the articles (identified in the literature search) mentioned in this review. Results are presented for each measure of exposure (smoking status, smoking duration, smoking intensity, pack-years, age at initiation, time since cessation, age at cessation) included in the article. Where possible, these results are presented for colorectal cancer overall and for men and women combined. At least one article from each cohort is included. More than one article from the same cohort is included if they include analyses of different measures of exposure.

2.4.2 Smoking Status

Smoking status represents the simplest form of evaluating an individuals' smoking exposure; people are categorised as never, former or current smokers. Numerous prospective cohort studies have evaluated the relationship between smoking status and colorectal cancer and these studies have been summarised in meta-analyses.[49, 50, 61, 226, 252] Four of these meta-analyses included studies published up to 2008 and one included

studies published up to 2013. Though there were differences in how each meta-analysis was conducted, there was large agreement in results.

Both former and current smokers have an increased risk of colorectal cancer. In one meta-analysis of 28 prospective cohort studies, the RR (95% CI) was 1.18 (1.12-1.25) for former smokers and 1.20 (1.10-1.30) for current smokers, compared with never smokers.[49] For both former and current smokers, there was a stronger association for rectal cancer than for colon cancer. For former smokers, the RR (95% CI) was 1.24 (1.11-1.37) for rectal cancer and 1.13 (1.05-1.21) for colon cancer. For current smokers, the RR (95% CI) was 1.36 (1.15-1.61) for rectal cancer and 1.11 (1.02-1.21) for colon cancer.

There was also evidence for a difference in results between men and women. Male current smokers had an increased risk of colorectal cancer (RR, 1.38; 95% CI, 1.22-1.56) but there was no evidence of an increased risk for female current smokers (RR, 1.06; 95% CI, 0.95-1.18).[49] The risk of colorectal cancer was similar for male (RR, 1.23; 95% CI, 1.09-1.40) and female (RR, 1.18; 95% CI, 1.08-1.28) former smokers.

Table 2.4.1 Results from Meta-analysis of Smoking and Colon and Rectal Cancer by Cheng et al.[50]

	Colon cancer				**Rectal cancer**			
	Former smokers		**Current smokers**		**Former smokers**		**Current smokers**	
	n	**RR (95% CI)**	**n**	**RR (95% CI)**	**n**	**RR (95% CI)**	**n**	**RR (95% CI)**
Overall	21	1.16 (1.11-1.22)	23	1.09 (1.01-1.18)	19	1.20 (1.11-1.30)	20	1.24 (1.16-1.39)
Men	7	1.18 (1.05-1.33)	8	1.09 (0.92-1.30)	7	1.23 (1.01-1.48)	8	1.36 (1.08-1.71)
Women	7	1.19 (1.09-1.30)	8	1.08 (0.97-1.21)	7	1.27 (1.05-1.52)	8	1.16 (0.97-1.40)

n = number of studies.

The most recent meta-analysis, by Cheng et al., included the most detailed results, presenting estimates separately according to colorectal subsite and sex (see Table 2.4.1).[50] While there was a similar risk of colon and rectal cancer for former smokers, current smokers had a lower risk of colon cancer than rectal cancer. Male current smokers had a greater risk of rectal cancer than female current smokers but the risk was similar for men and women across the other analyses.

Although it seems clear that smoking is a risk factor for colorectal cancer, individual studies find conflicting results for former and current smokers, men and women and colorectal subsites. For example, the EPIC study followed over 450,000 men and women from ten European countries over an average of nine years, during which 2,741 subjects were diagnosed with colorectal cancer.[253] In the EPIC study, the evidence for an association between cigarette smoking was stronger for colon cancer than rectal cancer and stronger for former smokers than current smokers. For colon cancer, the HR (95% CI) was 1.21 (1.08-1.36) for former smokers and 1.13 (0.98-1.31) for current smokers. For rectal cancer, the HR (95% CI) was 1.10 (0.94-1.30) for former smokers and 0.98 (0.80-1.19) for current smokers.

The Cancer Prevention Study II Nutrition Cohort included 51,365 men and 73,386 women aged 50-74 recruited in 1992.[254] 1,962 cases of colorectal cancer were identified during follow-up to 2005. The risk of colorectal cancer was similar for former and current smokers and for men and women. For men, the HR (95% CI) was 1.26 (1.09-1.45) for former smokers and 1.24 (0.96-1.59) for current smokers. For women, the HR (95% CI) was 1.19 (1.04-1.37) for former smokers and 1.30 (1.01-1.68) for current smokers. Analysing colon and rectal cancer separately (for men and women overall), the association for former smokers was similar for colon cancer (HR, 1.19; 95% CI, 1.06-1.34) and rectal cancer (HR, 1.26; 95% CI, 1.02-1.55) whereas the association for current smokers was stronger for colon cancer (HR, 1.28; 95% CI, 1.04-1.57) than for rectal cancer (HR, 0.97; 95% CI, 0.63-1.47).

In contrast, in the Singapore Chinese Health Study, former (HR, 1.45; 95% CI, 1.04-2.01) and current smokers (HR, 1.63; 95% CI, 1.23-2.17) both had an increased risk of rectal cancer but there was no increased risk of colon cancer for former smokers (HR, 0.96; 95% CI, 0.73-1.27) or current smokers (HR, 0.83; 95% CI, 0.64-1.06).[77]

Parajuli et al. analysed data for 600,000 Norwegian men and women recruited from four different health surveys.[255, 256] Range of follow-up was 4-33 years across the four surveys. 3,998 cases of colon cancer and 2,176 cases of rectal cancer were identified during follow-up. There was an increased risk of colon cancer for both female former smokers (HR, 1.16; 95% CI, 1.02-1.31) and current smokers (HR, 1.22; 95% CI, 1.10-1.36).[255, 256] In contrast, there was an increased risk of colon cancer for male former

smokers (HR, 1.14; 95% CI, 1.02-1.27) but not current smokers (HR, 1.03; 95% CI, 0.92-1.15). For rectal cancer, results were very similar for female former (HR, 1.26; 95% CI, 1.05-1.52) and current smokers (HR, 1.29; 95% CI, 1.10-1.51) and male former (HR, 1.28; 95% CI, 1.11-1.50) and current smokers (HR, 1.26; 95% CI, 1.09-1.45).

The Women's Health Initiative identified 1,242 cases of colorectal cancer during an average of eight years follow-up.[257] Current smokers had an increased risk of rectal cancer (HR, 1.95; 95% CI, 1.10-3.47) but not colon cancer (HR, 1.03; 95% CI, 0.77-1.38). For former smokers, there was no clear evidence of an increased risk for colon (HR, 1.12; 95% CI, 0.97-1.29) or rectal cancer (HR, 1.15; 95% CI, 0.80-1.67). In a study of female teachers (California Teachers Study), 1,205 cases of colorectal cancer were diagnosed during follow-up between 1995 and 2009.[258] There was some evidence that current smokers had an increased risk of colon (HR, 1.25; 95% CI, 0.94-1.66) and rectal cancer (HR, 1.36; 95% CI, 0.85-2.20) though the CIs were wide due to a small number of cases among current smokers. Former smokers seemed to have a lower risk of colon (HR, 1.10; 95% CI, 0.95-1.26) and rectal cancer (HR, 1.10; 95% CI, 0.86-1.42) than current smokers.

Proximal Colon Cancer and Distal Colon Cancer

Most studies that have investigated colon subsites separately support a stronger association for proximal colon cancer than for distal colon cancer.[253, 255, 259, 260] For example, in the EPIC study, the HRs (95% CIs) for proximal colon cancer were 1.25 (1.04-1.50) for former smokers and 1.31 (1.06-1.64) for current smokers and the HRs (95% CIs) for distal colon cancer were 1.13 (0.95-1.36) for former smokers and 0.91 (0.73-1.14) for current smokers.[253] Parajuli et al. also found that female former and current smokers had a greater risk of proximal colon cancer (HR, 1.22; 95% CI, 1.02-1.45 for former smokers and HR, 1.37; 95% CI, 1.18-1.59 for current smokers) than distal colon cancer (HR, 1.15; 95% CI, 0.94-1.41 for former smokers and HR, 1.12; 95% CI, 0.93-1.34 for current smokers) though results were less clear for men.[255] However, in the cohort of female teachers, there seemed to be an increased risk of distal colon cancer for former smokers (HR, 1.19; 95% CI, 0.92-1.55) and current smokers (HR, 1.56; 95% CI, 0.95-2.54) but evidence was weaker for proximal colon cancer (HR,

1.06; 95% CI, 0.90-1.25 for former smokers and HR, 1.14; 95% CI, 0.80-1.61 for current smokers).[258]

In the most recent meta-analysis, the RRs and 95% CIs for former smokers were 1.30 (1.15-1.48) for proximal colon cancer and 1.14 (0.97-1.33) for distal colon cancer. For current smokers, the results were 1.31 (1.13-1.52) for proximal colon cancer and 0.98 (0.84-1.14) for distal colon cancer.[50] These estimates were based on results from four studies.

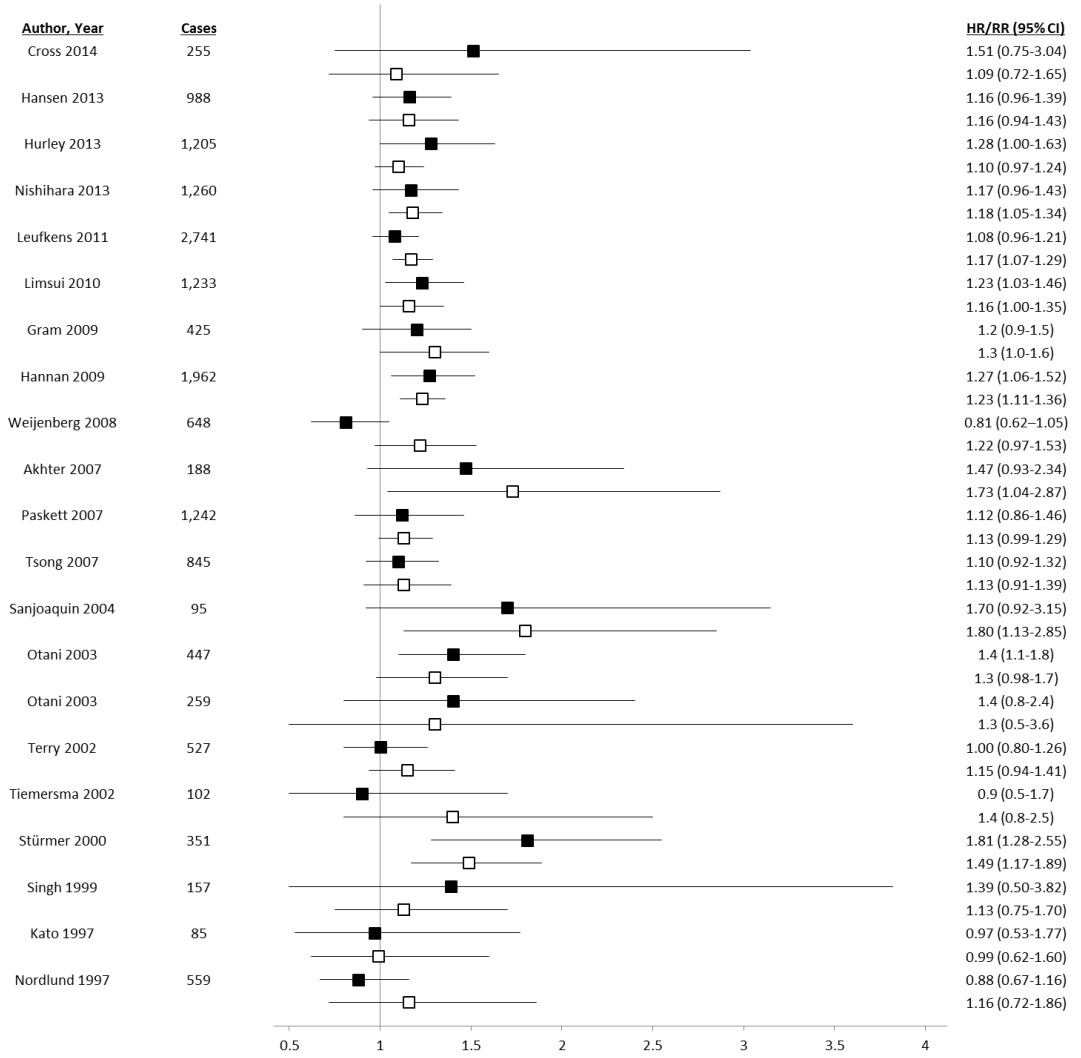
Risk for Former Smokers and Current Smokers

As mentioned above, meta-analysis results find a similar risk of colorectal cancer for former smokers and current smokers.[49, 61, 226, 252] Figure 2.4.2 illustrates the pattern of results for colorectal cancer risk for former and current smokers from different cohorts. Only cohorts which included results for both former and current smokers were included.

The reason why former smokers have an elevated risk of colorectal cancer similar to that of current smokers remains unclear. It could be that former smokers have a greater lifetime exposure on average than current smokers at baseline. However, on average, former smokers will have a shorter duration of smoking than current smokers. Also, prospective studies have shown that people who smoke fewer cigarettes are more likely to quit smoking.[261-266] Thus, it seems clear that current smokers will generally have had a greater exposure to tobacco smoke than former smokers.

Confounding must always be considered in observational studies. Based on baseline data, studies find that, on average, compared to current smokers, former smokers are more likely to have a higher socioeconomic status, to complete more physical activity, to eat less red and processed meat and to attend bowel screening. Former smokers, however, tend to have a higher BMI than current smokers. Since BMI is associated with colorectal cancer risk,[56] it is possible that the association between former smokers and colorectal cancer is confounded by BMI. However, many prospective cohort studies have adjusted for BMI and meta-analyses found no clear difference in results between studies that did and did not adjust for BMI.[49, 50] Thus, it seems unlikely that the risk of colorectal cancer among former smokers is due to confounding by other risk factors.

Figure 2.4.2 Results from Prospective Cohort Studies for the Association between Smoking Status and Colorectal Cancer Risk



Results show the comparison between current smokers or former smokers and never smokers from prospective cohort studies. Squares represent HRs/RRs. Bars represent 95% CIs. Black squares represent results for current smokers and white squares represent results for former smokers.

Alternatively, the similar risks observed for former and current smokers may reflect that tobacco smoke promotes early colorectal carcinogenesis but has less influence on later stages of tumour progression, meaning that the effect of smoking persists for many years, even after cessation. However, current smokers would still be expected to have a higher risk on average since many former smokers will have a short duration of exposure and would not be expected to have an increased risk.

Other Factors Affecting the Association between Smoking and Colorectal Cancer

It is also unclear why studies find rather conflicting results, for example why some studies find an increased risk for both former smokers and current smokers whereas others find an increased risk only for former smokers or a higher risk for former smokers. The meta-analysis by Cheng et al. did find some evidence that results may be affected by certain study characteristics.[50] For example, smoking was more strongly associated with colon and rectal cancer among studies with fewer than 500 cases compared with studies with more than 500 cases. This may indicate publication bias though the authors reported no evidence of publication bias based on Begg's test. There was also a difference in results for current smokers according to study quality score. The RRs (95% CIs) for the highest and lowest quality studies were 1.15 (1.05-1.27) and 1.00 (0.89-1.12) for colon cancer and 1.34 (1.15-1.55) and 1.10 (0.97-1.25) for rectal cancer. However, this quality score was not described in detail and another meta-analysis did not find different results according to study quality.[49] There were no important differences in results between studies that did or did not adjust for alcohol intake, BMI, physical activity or family history of colorectal cancer.[49, 50]

Interestingly, meta-analyses found some evidence that results for current smokers may depend on length of follow-up. One meta-analysis found RRs and 95% CIs of 1.08 (0.95-1.22), 1.17 (1.08-1.29) and 1.23 (1.07-1.41) for studies with follow-up <10, 10-<25 and ≥25 years.[61] Similarly, Cheng et al. found that current smokers had a greater risk of colon and rectal cancer in studies with ≥10 years of follow-up than in studies with <10 years follow-up[50] (although another meta-analysis did not find a difference in results for current smokers between studies with <12 and ≥12 years of follow-up).[49]

That the risk of colorectal cancer among current smokers would increase with longer follow-up seems rather puzzling. It could be because smoking duration among current smokers will increase with increasing follow-up as participants age. However, this should not affect results since all studies adjust for age. It must also be mentioned that if the HR does increase with increasing follow-up, this would violate the assumptions required for Cox proportional hazards models.[267]

As follow-up increases, more current smokers will quit smoking and become former smokers. This would seem to predict a lower risk for current smokers with increasing follow-up. However, the higher risk could be related to weight gain among former smokers. Smokers tend to have a lower BMI than non-smokers and smoking cessation is associated with weight gain.[227, 268] Since BMI is a major risk factor for colorectal cancer,[56, 157] weight gain among former smokers may explain why the risk of colorectal cancer among baseline current smokers increases as more baseline current smokers quit smoking. However, this is not entirely clear since excess adiposity is thought to increase colorectal cancer risk as a result of increasing insulin resistance.[157, 185] Although smoking is associated with lower BMI, smoking is also associated with insulin resistance. Consequently, although former smokers tend to gain weight, smoking cessation may actually improve insulin sensitivity.[227, 268]

Another possibility is that duration of follow-up serves as a proxy for induction period. Giovannucci et al. hypothesised that smoking acts as an initiator of colorectal carcinogenesis and that an induction period of 35-40 years is required to observe an increased risk of colorectal cancer.[269-271] This hypothesis was based on the fact that early studies of smoking and colorectal cancer did not find evidence of an association yet studies consistently found a strong association between cigarette smoking and colorectal adenomas (recognised as precursor lesions for colorectal cancer). Given that colorectal cancer develops over a number of decades, studies may have failed to find evidence for an association with colorectal cancer because they did not take into account this long induction period.

Support for this hypothesis was provided by analyses of the Health Professionals Follow-up Study and the Nurses' Health Study.[269, 270] For example, men in the Health Professionals Follow-up Study reported the average number of cigarettes smoked at different ages (≤14, 15-19, 20-29, etc.).[269] Total pack-years of smoking up to baseline was associated with an increased risk of colorectal adenoma (RR, 1.67; 95% CI, 1.25-2.22 for ≥40 vs 0 pack-years) as well as colorectal cancer (HR, 1.45; 95% CI, 1.01-2.09 for ≥40 vs 0 pack-years). Pack-years of smoking before age 30 was associated with the risk of large adenomas (≥1 cm) (RR, 2.47; 95% CI, 1.48-4.11 for ≥16 vs 0 pack-years) and colorectal cancer (HR, 1.66; 95% CI, 1.15-2.40 for ≥16 vs 0 pack-years) but not small adenomas (<1 cm) (RR, 1.19; 95% CI, 0.75-1.88 for ≥16 vs 0 pack-years). After

adjusting for pack-years of smoking after age 30, the association with large adenomas became slightly weaker (RR, 2.25; 95% CI, 1.20-4.21) whereas the association with colorectal cancer became stronger (HR, 2.02; 95% CI, 1.27-3.23). In contrast however, after adjusting for pack-years before age 30, there was only evidence of an association for pack-years of smoking after age 30 with small adenomas.

Summary

Tobacco smoking is associated with colorectal cancer. The association is similar for men and women and for colon and rectal cancer. The association may be stronger for proximal colon cancer than distal colon cancer. Former smokers have a similar risk of colorectal cancer to current smokers, possibly indicating an effect of smoking on early colorectal carcinogenesis.

2.4.3 Smoking Duration

Smoking status offers a very narrow picture of a person's smoking history. If tobacco smoke increases the risk of colorectal cancer, it would be expected that a greater exposure to tobacco smoke is associated with a greater risk. One measure of tobacco smoke exposure is duration of smoking. Many studies have investigated the relationship between smoking duration and colorectal cancer risk and most find evidence that a longer duration is associated with a greater risk.

In the Iowa Women's Health Study, the HRs and 95% CIs for the risk of colorectal cancer for smoking duration 1-19, 20-39 and ≥40 years, compared to never smokers, were 1.17 (0.94-1.46), 1.05 (0.88-1.24) and 1.40 (1.17-1.68).[272] In the California Teachers Study, there was only evidence of an increased risk of colorectal cancer for smokers with duration ≥40 years (HR, 1.27; 95% CI, 1.03-1.57).[258] There was evidence for an increased risk of colorectal cancer associated with smoking more than 20 years in the Women's Health Initiative though people who smoked for 20-29 years (HR, 1.36; 95% CI, 1.12-1.66) had a higher risk than people who smoked for 30-39 or ≥40 years.[257]

Parajuli et al. found that smoking duration was associated with colon and rectal cancer risk for both men and women.[255, 256] Dose-response relationships were evident for each

analysis. There seemed to be an increased risk of cancer across all categories of smoking duration i.e. including for 1-19 years, except for the risk of colon cancer among men where there was an increased risk only for ≥40 years duration.

Cheng et al. found that the risk of colon cancer increased by 2, 5 and 10% with an increase of 10, 20 and 40 years of smoking. The corresponding increases in risk for rectal cancer were 6, 13 and 24%.[50] In another meta-analysis including prospective and retrospective studies, smoking duration was modelled as a continuous variable. The risk of colorectal cancer did not begin to increase until after approximately 10 years of smoking and reached statistical significance only after 30 years of smoking.[226]

One issue with the analyses of smoking duration is that most studies consider ever smokers overall (i.e. former and current smokers combined) and do not analyse the effect of smoking duration on colorectal cancer risk separately for former smokers and current smokers. This may lead to rather misleading results; although some former and current smokers may have a similar duration of smoking up to baseline, it is plausible that they experience different risks of cancer since current smokers continue to add to their exposure during follow-up.

Smoking duration was analysed separately for former and current smokers in the EPIC study.[253] For former smokers, smoking less than 20 years was not associated with colorectal cancer risk whereas smoking 20-29 years (HR, 1.27; 95% CI, 1.10-1.47) and ≥30 years (HR, 1.24; 95% CI, 1.07-1.44) was associated with colorectal cancer risk. The association between smoking duration and colorectal cancer risk was less clear for current smokers. There was evidence for an increased risk for 30-39 years (HR, 1.19; 95% CI, 1.00-1.41) but not for ≥40 years (HR, 0.94; 95% CI, 0.78-1.14).

In the Singapore Chinese Health Study, smoking duration was analysed separately for former and current smokers for rectal cancer only (there was no evidence for an association with colon cancer for smoking status or smoking duration).[77] For former smokers, there was no increased risk of rectal cancer for duration <40 years (HR, 1.15; 95% CI, 0.78-1.69) but there was an increased risk for former smokers with duration ≥40 years (HR, 1.93; 95% CI, 1.20-3.09). In contrast, there was an increased risk for current smokers with <40 years (HR, 1.58; 95% CI, 1.09-2.29) and ≥40 years duration

(HR, 1.78; 95% CI, 1.27-2.50). In the Cancer Prevention Study II Nutrition Cohort, current smokers with 40-49 (HR, 1.32; 95% CI, 1.02-1.72) and ≥50 years duration (HR, 1.38; 95% CI, 1.04-1.84) had an increased risk of colorectal cancer whereas there was no evidence of an increased risk for current smokers with <40 years duration (HR, 1.02; 95% CI, 0.69-1.49).[254]

Another potential issue when analysing smoking duration is that smokers with longer duration may tend to have been heavier smokers. Earlier age at onset of smoking is associated with nicotine dependence and heavier smoking.[273-275] Therefore, it is possible that part of the increased risk for longer duration may be due to heavier smoking intensity.

Results for smoking duration were adjusted for smoking intensity in a pooled analysis by Gong et al.[260] Data were pooled from five cohort studies (including the Women's Health Initiative) and three case-control studies. The analysis included a total of 6,796 cases of colorectal cancer and 7,770 controls. Adjusting for smoking intensity (never smoker, <20, 20, >20 cigarettes/day), there was no evidence of an increased risk of colorectal cancer for ever smokers with a smoking duration <20 years whereas there was an increased risk for ever smokers with at least 20 years duration. The ORs (95% CIs) for <10, 10-20, 20-30, 30-40 and ≥40 years were 0.94 (0.78-1.13), 1.07 (0.93-1.24), 1.29 (1.11-1.50), 1.29 (1.12-1.48) and 1.28 (1.10-1.49).

Summary

Smoking duration is associated with colorectal cancer risk. Studies should analyse smoking duration separately for former smokers and current smokers.

2.4.4 Smoking Intensity

Similar to smoking duration, most prospective cohort studies support a dose-response relationship between the number of cigarettes smoked per day and colorectal cancer risk.

The risk of colorectal cancer was associated with the number of cigarettes per day in the Iowa Women's Health Study. The HRs and 95% CIs for 1-19, 20 and >20 cigarettes per day were 1.10 (0.94-1.29), 1.28 (1.06-1.55) and 1.32 (1.04-1.69).[272] Parajuli et al. found that smoking intensity was associated with colon and rectal cancer risk for both men and women. The HRs and 95% CIs associated with smoking at least 20 cigarettes per day were 1.16 (0.97-1.28) and 1.31 (1.11-1.55) for colon and rectal cancer for men and 1.28 (1.06-1.55) and 1.38 (1.05-1.81) for colon and rectal cancer for women.[255, 256] In the Women's Health Initiative, smoking at least 25 cigarettes per day was associated with a HR of 1.41 (95% CI, 1.14-1.76).[257]

Similar to smoking duration, most studies analyse the association between smoking intensity and colorectal cancer for ever smokers. In the EPIC study, average number of cigarettes/day was associated with colorectal cancer risk for former smokers.[253] The HRs and 95% CIs for ≤9, 10-14 and ≥15 cigarettes/day were 1.08 (0.92-1.26), 1.18 (0.98-1.42) and 1.26 (1.08-1.47). Current smokers who smoked ≥20 cigarettes/day at baseline also had an increased risk of colorectal cancer (HR, 1.15; 95% CI, 0.98-1.36).

Adjusting for smoking duration, Gong et al. found very similar increased risks of colorectal cancer for ever smokers who smoked <20, 20 and >20 cigarettes/day.[260] The ORs and 95% CIs were 1.28 (1.11-1.48), 1.30 (1.09-1.55) and 1.31 (1.01-1.70).

Studies have not investigated the risk of colorectal cancer according to duration and intensity together. For example, the risk of smoking 20 cigarettes a day will likely be different for someone who smoked for 10 years and someone who smoked for 40 years. Similarly, the risk of smoking 40 years may be different for someone who smokes 5 cigarettes a day and someone who smokes 30 cigarettes a day. Further studies should categorise people by both duration and intensity.

Summary

Greater number of cigarettes per day is associated with a greater risk of colorectal cancer. More studies need to investigate the effects of duration and intensity together.

2.4.5 Pack-years

Both smoking duration and smoking intensity are associated with colorectal cancer risk. Pack-years are defined as the duration of smoking multiplied by the number of cigarettes smoked per day divided by 20. Thus, pack-years combine smoking duration and intensity to provide a single measure of cumulative exposure.

Pack-years were associated with colorectal cancer risk in the Iowa Women's Health Study though there was only clear evidence for an increased risk for at least 40 pack-years (HR, 1.39; 95% CI, 1.14-1.70).[272] In the California Teachers Study, there was slight evidence for an increased risk of colorectal cancer for 21-30 pack-years (HR, 1.19; 95% CI, 0.91-1.56) and stronger evidence for ≥31 pack-years (HR, 1.37; 95% CI, 1.12-1.67).[258]

Parajuli et al. found that pack-years of smoking was associated with colon cancer risk for men and women though the association was slightly weaker for men (HR, 1.14; 95% CI, 0.99-1.31 for ≥20 pack-years) than for women (HR, 1.33; 95% CI, 1.11-1.57 for ≥20 pack-years).[255] Pack-years of smoking was also associated with rectal cancer risk for both men (HR, 1.35; 95% CI, 1.14-1.58) and women (HR, 1.47; 95% CI, 1.13-1.91).[256]

Thus, increasing pack-years appear to be associated with increasing risk of colorectal cancer, which is also supported by meta-analysis results.[49, 50, 226, 252] However, the use of pack-years to represent smoking exposure has received much criticism.[276-278] The main issue with pack-years is that it equates duration and intensity; for example, smoking 10 cigarettes per day for 20 years and smoking 20 cigarettes for 10 years are both equal to 10 pack-years. Thus, the use of pack-years assumes that both components are related to risk in the same way.

Another issue with pack-years is that most studies only ask for very limited information about the number of cigarettes smoked. Most studies only ask current smokers to report the number of cigarettes they smoke at baseline. Calculating pack-years based on baseline smoking assumes that current smokers have smoked the same amount of cigarettes throughout the entire duration of their smoking. Former smokers are generally asked to report their average number of cigarettes which in theory should be sufficient

to provide an accurate measure of pack-years but it is not clear how accurately people are able to provide such information.

Summary

Pack-years are associated with colorectal cancer risk. Pack-years of smoking may not be a suitable measure of smoking exposure since it assumes that the effects of duration and intensity are the same.

2.4.6 Age at initiation

Another aspect of smoking exposure is age at initiation. People who start smoking at an earlier age, on average, will have a greater level of exposure than people who start smoking later in life. Studies do find an association between age at initiation and colorectal cancer risk though the evidence is not as consistent as for duration. The main difficulty when analysing age at initiation is that the large majority of smokers begin smoking during a very narrow age range.

People who began smoking before age 15 had an increased risk of colorectal cancer (HR, 1.32; 95% CI, 1.02-1.71) compared to never smokers in the Singapore Chinese Health Study but there was no evidence of an increased risk after age 15 (HR, 1.07; 95% CI, 0.90-1.27).[77] Parajuli et al. found evidence that women who began smoking ≤16, 17-19, 20-24 and ≥25 all had an increased risk of colon cancer though the risk was greatest for ≤16 (HR, 1.48; 95% CI, 1.21-1.81). The evidence was weaker for men; the HR (95% CI) for the risk of colon cancer for men who began smoking ≤16 was 1.15 (0.99-1.34).[255] For rectal cancer, both men and women who began smoking ≤19, 20-24 and ≥25 had an increased risk. However, the risk was greatest for 20-24 for both men (HR, 1.35; 95% CI, 1.13-1.61) and women (HR, 1.45; 95% CI, 1.18-1.78).[256]

In the EPIC study, there seemed to be an increased risk for each category of age at initiation for former smokers, though the greatest risk was for former smokers who started between ages 17-19; the HRs (95% CIs) for ≤16, 17-19 and ≥20 compared to never smokers were 1.12 (0.96-1.31), 1.24 (1.09-1.41) and 1.13 (0.99-1.30).[253] The risk of colorectal cancer was similar for women who started before 20 (HR, 1.14; 95% CI,

0.97-1.33) and women who started after 20 (HR, 1.12; 95% CI, 0.95-1.31) in the Women's Health Initiative.[257]

Since the large majority of smokers begin smoking during a very narrow age range, age at initiation and smoking duration are very highly correlated. Hence, it is not clear whether age at initiation is associated with colorectal cancer risk independent of smoking duration since studies are unable to separate the effects of age at initiation on colorectal cancer risk from the effects of duration. Furthermore, it is unknown in general how the timing of smoking exposure may affect the risk of cancer. For example, does a 50 year old who starts smoking have the same risk as a 50 year old lifelong smoker? Hence, the effects of timing versus duration remains an important question for studies of smoking.

One possible issue with age at initiation, besides problems with recall, is differentiating between starting smoking and starting smoking regularly. For example, a person may try their first cigarette many years before taking up smoking as a regular habit. So, although someone begins smoking at a very early age, they might not experience a greater risk because they only smoked very occasionally for a number of years. This will affect analyses of age at initiation as well as duration.

As described in section 2.4.3, people who take up smoking at an earlier age are more likely to be heavy smokers and to be nicotine dependent.[273-275] Therefore, people who start smoking later may have a lower risk of colorectal cancer because they are less dependent and tend to smoke fewer cigarettes than people who started earlier.

Summary

Earlier age at initiation of smoking is associated with a greater risk of colorectal cancer. People who start smoking earlier will generally have a longer duration of smoking and so the effect of age at initiation on colorectal cancer independent of smoking duration is not known.

2.4.7 Smoking Cessation

The risk of colorectal cancer is similar for former and current smokers, suggesting that the adverse effects of smoking may persist for many years after smoking cessation. A number of studies have investigated how the risk of colorectal cancer differs with cessation at different times by analysing time since cessation and age at cessation.

Time since cessation was associated with colorectal cancer risk in the Cancer Prevention Study II Nutrition Cohort though there was no increased risk of colorectal cancer for former smokers who quit at least 31 years ago.[254] Compared to never smokers, the HRs and 95% CIs for former smokers who quit ≥31, 21-30, 11-20 and 1-10 years before baseline were 1.03 (0.89-1.19), 1.28 (1.10-1.49), 1.33 (1.14-1.55) and 1.48 (1.27-1.73). In a cohort of female teachers, smoking cessation less than 5 years ago was associated with an increased risk of colorectal cancer (HR, 1.44; 95% CI, 1.03-2.01) whereas former smokers who quit at least 20 years ago did not have an increased risk (HR, 0.98; 95% CI, 0.83-1.14).[258] Former smokers who quit smoking less than 20 years ago had an increased risk of colorectal cancer in the EPIC study though the highest risk was actually observed for former smokers who quit 15-19 years ago (HR, 1.34; 95% CI, 1.12-1.60).[253] There was no clear evidence of an increased risk for former smokers who quit 20-24 (HR, 1.11; 95% CI, 0.91-1.35) and ≥25 years ago (HR, 1.08; 95% CI, 0.92-1.26).

Gong et al. found evidence for an increased risk of colorectal cancer for the most recent quitters, adjusting for pack-years of smoking.[260] Former smokers who quit within the last 15 years had the highest risk of cancer (OR, 1.47; 95% CI, 1.21-1.78) compared with never smokers. The risk was similar to never smokers after 25-35 (OR, 1.15; 95% CI, 0.85-1.55) and ≥35 years since cessation (OR, 0.74; 95% CI, 0.47-1.18).

Given that former smokers overall have a similar risk of colorectal cancer to current smokers and that time since cessation is associated with colorectal cancer risk, it is to be expected that the most recent quitters have a higher risk than current smokers. However, this result still seems surprising since recent quitters and current smokers will have a similar level of exposure up to baseline. One possible explanation for the higher risk among the most recent quitters could relate to overall health status. Former smokers

may quit smoking as a result of ill health whereas the fact that current smokers continue to smoke may be a signal of satisfactory health.

Only a few studies have investigated age at cessation though they seem to agree that former smokers who quit smoking before age 40 are not at an increased risk of colorectal cancer compared to never smokers. People who quit smoking after 40 have an increased risk though it is not clear whether the risk increases with increasing age at cessation. For example, in the Cancer Prevention Study II Nutrition Cohort, HRs and 95% CIs for quitting smoking <40, 40-49, 50-59 and ≥60 years old were 1.05 (0.91-1.22), 1.31 (1.13-1.52), 1.44 (1.24-1.66) and 1.29 (1.08-1.54).[254] Similarly, in the Women's Health Initiative, women who quit smoking <30 (HR, 0.94; 95% CI, 0.74-1.23) and 30-39 years old (HR, 0.86; 95% CI, 0.67-1.11) did not have an increased risk of colorectal whereas women who quit smoking 40-49 (HR, 1.26; 95% CI, 1.02-1.56) and ≥50 years old (HR, 1.28; 95% CI, 1.07-1.55) did have an increased risk.[257] In the pooled analysis by Gong et al. (which included the Women's Health Initiative), the ORs and 95% CIs for smokers who quit smoking <40, 40-50 and ≥50 years old were 1.03 (0.79-1.34), 1.28 (0.96-1.70) and 1.31 (1.01-1.70).[260]

The main difficulty when investigating smoking cessation and colorectal cancer risk is that the large majority of smokers begin smoking during a narrow age range. This means that the former smokers with the longest time since cessation will generally be the same smokers with the shortest smoking duration. Consequently, it is very difficult to disentangle the effects of duration and cessation and it is not possible to say for certain whether the lower risk for someone who quit smoking 30 years ago is due to the long time since cessation or because of the short smoking duration.

It seems plausible that the association between smoking cessation and risk of colorectal cancer differs according to smoking intensity; former light smokers who quit recently may have a lower risk of colorectal cancer than former heavy smokers who quit recently. However, no studies have investigated the effects of smoking cessation and smoking intensity together and it is not known how the risk due to cessation at different times may differ according to the amount smoked.

Summary

Time since cessation is associated with colorectal cancer risk. The risk of cancer for former smokers is comparable to the risk for never smokers after 20-30 years of cessation. The most recent quitters have a higher risk of colorectal cancer than current smokers. Age at cessation is also associated with colorectal cancer risk. Former smokers who quit before 40 do not appear to have an increased risk compared to never smokers.

2.4.8 Effect Modifiers

It is possible that the association between smoking and colorectal cancer is modified by the level of other risk factors. Few studies have investigated effect modifiers for smoking and colorectal cancer.

Smoking and Alcohol Intake

Alcohol intake and smoking are known to act synergistically in cancers of the upper aerodigestive tract[26] though there is insufficient evidence for colorectal cancer. Both smoking and alcohol drinking were associated with colorectal cancer risk in the Singapore Chinese Health Study though there was no evidence for an interaction between the two risk factors.[77] There was also no evidence for an interaction in a Japanese cohort study.[120] The EPIC study reported a non-significant p-value for the effect modification by alcohol intake.[253]

Smoking and BMI

BMI represents a major risk factor for colorectal cancer.[25] On average, smokers have a lower BMI than non-smokers. However, BMI is positively associated with the number of cigarettes smoked per day among smokers.[227, 279] Hence, it may be important to consider these risk factors together when investigating colorectal cancer risk. In the pooled analysis by Gong et al., the OR (95% CI) comparing ever vs never smokers was 1.14 (1.10–1.27) for people with BMI <25 kg/m^2 and 1.24 (1.13–1.35) for people with BMI ≥25 kg/m^2.[260] Parajuli et al. investigated the risk for ever smokers compared to never smokers for different categories of BMI for rectal cancer only. The risk increased

with increasing BMI for men; the HR (95% CI) comparing ever smokers to never smokers were 1.17 (1.01-1.36) for BMI <25 kg/m^2, 1.33 (1.15-1.54) for BMI 25-29 kg/m^2 and 1.53 (1.20-1.95) for BMI ≥30 kg/m^2. The corresponding results for women were 1.18 (1.01-1.39), 1.39 (1.15-1.70) and 1.39 (1.15-1.70).[256] There was no clear difference in results comparing ≥31 pack-years versus never smokers by BMI category in a cohort of female teachers[258] and the EPIC study reported a non-significant p-value for the effect modification by BMI.[253]

Summary

There is insufficient evidence that the association between smoking and colorectal cancer is modified by alcohol intake. There is slight evidence that the effect of smoking on colorectal cancer risk increases with increasing BMI.

2.4.9 Colorectal Adenomas

Studies have also investigated the relationship between smoking and colorectal adenomas. Colorectal adenomas are recognised precursor lesions for most colorectal cancers.[19, 280] Generally, these studies have found a much stronger association for colorectal adenomas than for colorectal cancer. In a meta-analysis of prospective and retrospective studies of smoking and colorectal adenomas through 2008, Botteri et al. found a pooled RR and 95% CI of 2.14 (1.86-2.46) for current smokers and 1.47 (1.29-1.67) for former smokers.[281] In a similar meta-analysis of smoking and colorectal cancer by the same authors, the corresponding RR and 95% CI was 1.07 (0.99-1.16) for current smokers and 1.17 (1.11-1.22) for former smokers.[226]

Since the adenoma-carcinoma sequence accounts for the vast majority of colorectal cancers, it is unclear why there should be such a discrepancy in results for colorectal adenomas and colorectal cancer. Different explanations have been proposed.[282]

First, the discrepancy in results may be explained by the idea that tobacco smoke acts as an initiator of colorectal carcinogenesis. However, even if smoking had no effect on later tumour progression, the effects of smoking on colorectal adenoma and cancer should still be similar. The only possible explanation for a lower risk of cancer than

adenoma would be if smoking promoted the initiation of adenomas that were less likely to develop into colorectal cancer. However, Botteri et al. actually found that the association between smoking and colorectal adenoma was stronger for "high-risk" adenomas than "low-risk" adenomas.[281]

Thus, on its own, that smoking only affects early colorectal carcinogenesis is insufficient to explain the difference in results for adenomas and cancer. However, an effect of smoking mainly on early carcinogenesis implies a long induction period for colorectal cancer risk and perhaps the difference in results is due to studies failing to account for this long induction period, as hypothesised by Giovannucci et al.[269-271] However, as described above, most studies still find a fairly modest risk of colorectal cancer even for people with 30 or 40 years of smoking duration.

Most studies of smoking and colorectal adenomas are retrospective case-control studies. Thus the higher risk of colorectal adenomas could be due to recall bias (i.e. people diagnosed with adenoma being more likely to recall smoking) or due to other differences in methodology between the two study types. However, this seems unlikely since the meta-analysis by Botteri et al. found a similar association between smoking and colorectal cancer for cohort and case-control studies.[226]

Another idea put forward to explain the difference in results relates to the selection of controls in case-control studies.[282, 283] Studies of colorectal adenoma are generally based on direct evaluation of the large bowel and so it is known that the controls are free of adenomas. In contrast, population-based controls are normally used in studies of colorectal cancer. Since these population-based controls have not undergone any kind of screening, a high proportion of controls will have prevalent colorectal adenomas. Thus, it has been argued that the inclusion of these controls with adenomas is responsible for the lower association observed for colorectal cancer risk.[283] Indeed, the authors included results showing that the effect of smoking on colorectal cancer became stronger after excluding controls with adenomas.

However, since smoking is associated with colorectal adenoma risk, restricting the control group to people free of adenomas reduces the prevalence of smoking in the reference group. Thus, rather than eliminating a bias of results towards the null,

excluding people with adenoma from the control group could be considered as introducing a bias of results away from the null. This idea was described in an article by Poole.[284] Poole also highlighted that the effect of smoking on colorectal cancer would be equal to that of adenoma when there is a single pathway for colorectal cancer i.e. all colorectal cancers develop from adenoma. However, the effect of smoking would be weaker for cancer when there are multiple pathways.

Summary

The association between smoking and colorectal adenoma is much stronger than the association between smoking and colorectal cancer. Different explanations have been proposed though none seem to provide a sufficient explanation. The fact that the discrepancy in results between colorectal adenoma and colorectal cancer is not observed for other risk factors such as alcohol intake and adiposity[62, 63, 150, 153, 285-288] implies that the discrepancy relates to how smoking influences colorectal carcinogenesis rather than a methodological difference between studies of adenoma and cancer.

2.4.10 Molecular Subtypes

Accumulating evidence indicates that smoking is strongly associated with a subtype of colorectal cancers characterised by high microsatellite instability (MSI), CpG island methylator phenotype (CIMP) positive status and positive BRAF mutation status. In the Iowa Women's Health Study, the risk of these colorectal cancer subtypes for current smokers was approximately twice the risk for never smokers (HR, 1.99; 95% CI, 1.26-3.14 for MSI-high cancers, HR, 1.88; 95% CI, 1.22-2.90 for CIMP-positive cancers and HR, 1.92; 95% CI, 1.22-3.02 for BRAF mutation-positive cancers).[272] In contrast, there was no evidence of an association between smoking and microsatellite stable/MSI-low cancers, CIMP-negative cancers or BRAF mutation-negative cancers.

The risk of these colorectal cancer subtypes was also investigated in the Health Professionals Follow-up Study and the Nurses' Health Study.[289] Compared to never smokers, current smokers had a two-fold risk of CIMP-high cancers (HR, 2.08; 95% CI, 1.35-3.20) but no increased risk of CIMP-low cancers (HR, 1.12; HR, 0.89-1.41) and a two-fold risk of MSI-high cancers (HR, 2.05; 95% CI, 1.29-3.26) but no increased risk

of microsatellite stable cancers (HR, 1.14; 95% CI, 0.91-1.42). The results were more similar for BRAF mutation-positive (HR, 1.38; 95% CI, 0.84-2.25) and BRAF mutation-negative cancers (HR, 1.22; 95% CI, 0.98-1.52).

Case-control studies also support a stronger association between smoking and these molecularly defined subtypes of colorectal cancer.[290-294] Furthermore, CIMP-high, MSI-high and BRAF mutation positive cancers occur more frequently in the proximal colon[52, 272, 289, 295] which is in agreement with the stronger association observed between smoking and proximal colon cancer.[50]

These molecular subtypes of cancer are thought to arise along a distinct pathway. Until recently, colorectal cancer was often thought of as a single disease, with conventional adenomas being the sole precursors to colorectal cancer. Indeed, the majority of colorectal cancers develop from adenomatous polyps via the adenoma-carcinoma sequence.[280] However, it is now accepted that colorectal cancer is a heterogeneous disease and develops through multiple pathways.[296-299] Approximately 30% of cancers are now recognised to develop along a serrated neoplasia pathway. Cancers arising from serrated adenomas are characterised by high MSI, CIMP positive status and positive BRAF mutation status.

The fact that there are multiple pathways for colorectal cancer may help to explain the discrepancy in results for adenomas and cancer though this is not immediately clear. Smoking appears to be primarily associated with cancers developing along the serrated pathway yet serrated adenomas are actually more difficult to detect during screening.[300, 301] If serrated adenomas are more likely to be missed in studies of adenomas, this would seem to predict that smoking would be more strongly related to colorectal cancer than adenomas. In contrast, a stronger association for adenomas would be expected if smoking were primarily associated with the risk of conventional adenomas and not with the risk of serrated adenomas since the serrated adenomas (that are more likely to be missed in studies of adenoma) would "dilute" results for analyses of cancer.

Some studies have investigated the risk of smoking on conventional adenomas and serrated adenomas separately. One study analysed data from three RCTs of antioxidants, calcium and aspirin.[302] All subjects had at least one adenoma removed at

baseline and were followed up for the occurrence of new adenomas for up to four years. Out of 2,667 subjects with follow-up data, 973 had at least one conventional adenoma and 633 had at least one serrated adenoma. While there was an increased risk of one or more conventional adenomas for current smokers (RR, 1.29; 95% CI, 1.11-1.49) and former smokers (RR, 1.18; 95% CI, 1.05-1.32), there was a higher risk of one or more serrated adenomas for current smokers (RR, 2.01; 95% CI, 1.66-2.44) and former smokers (RR, 1.42; 95% CI, 1.20-1.68).

Three case-control studies also found that smoking was more strongly associated with the risk of serrated adenomas than conventional adenomas.[303-305] Comparing current to never smokers, the ORs and 95% CIs in the three studies for conventional adenomas only were 1.56 (0.99-2.44), 1.96 (1.61-2.38) and 1.8 (1.5-2.1). The ORs and 95% CIs for serrated adenomas only were 3.00 (1.93-4.66), 4.44 (3.47-5.67) and 4.4 (3.7-5.2).

Summary

The association between smoking and colorectal cancer seems to be restricted to cancers with high MSI, CIMP positive status and positive BRAF mutation status. These cancers develop from serrated adenomas which are precursors to approximately a third of colorectal cancers. Smoking is more strongly associated with serrated adenomas than conventional adenomas. These results need to be considered when trying to explain the discrepancy in results between colorectal adenoma and cancer.

2.4.11 Mechanisms

Cigarette smoke contains numerous carcinogens including polycyclic aromatic hydrocarbons, aromatic amines, and N-nitrosamines.[306, 307] These carcinogens can reach the colorectal mucosa through the digestive system or the circulatory system.[249] Carcinogens from cigarette smoke form DNA adducts and subsequent mutations can result in the loss of normal growth control mechanisms and cancer.[306, 308] However, not much is known in detail about how tobacco smoke increases colorectal cancer risk. In particular, smoking seems to greatly increase the risk of a subtype of colorectal cancers that develop along a serrated pathway yet the mechanisms for these neoplasms are not currently well explained in the literature.

2.4.12 Summary

Cigarette smoking is associated with an increased risk of colorectal cancer. The risk is similar for men and women and smoking increases the risk of both colon and rectal cancer though studies find conflicting results. Smoking is more strongly related to proximal colon cancer than distal colon cancer. Former smokers suffer a similar level of risk as current smokers which may indicate that smoking mainly affects the early stages of colorectal carcinogenesis. Both smoking duration and smoking intensity are positively associated with colorectal cancer risk though it is unclear how these factors may interact to increase risk. It is also unclear how the risk of colorectal cancer differs with smoking cessation at different times. Compared to the risk of colorectal cancer, smoking is strongly associated to the risk of colorectal adenomas. The explanation for this discrepancy remains unknown but is possibly related to the stronger effects of smoking on a subset of cancers arising along a serrated pathway.

Table 2.4.2 Prospective Cohort Studies of Smoking and Colorectal Cancer Risk

Ref.	Author, year	Study name, location	Population	Follow-up period (average follow-up)	Age	Sex[a]	Number of subjects	Site[b]	Number of cases	Exposure[c]	HR/RR (95% CI)
309	Cross 2014	Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial, USA	General population	1993 - 2001	55-74	B	254*	CRC	255	C v N	1.51 (0.75-3.04)
										F v N	1.09 (0.72-1.65)
										>40 D v N	1.09 (0.47-2.55)
										≥30 I v N	1.11 (0.34-3.62)
										>60 PY v N	1.30 (0.38-4.42)
										>20 YSC v N	0.94 (0.59-1.48)
256	Parajuli 2014	The Oslo Study I The Norwegian Counties Study The 40 Years Cohort The CONOR Study, Norway	General population	1972 - 2007 (14 yrs)	19-67	M	299,376	RC	1,336	C v N	1.26 (1.09-1.45)
										F v N	1.28 (1.11-1.50)
										≥30 D v N	1.31 (1.09-1.59)
										≥20 I v N	1.31 (1.11-1.55)
										≥20 PY v N	1.35 (1.14-1.58)
										≤19 AI v N	1.28 (1.08-1.50)
						F	302,866	RC	840	C v N	1.29 (1.10-1.51)
										F v N	1.26 (1.05-1.52)
										≥30 D v N	1.54 (1.11-2.12)
										≥20 I v N	1.38 (1.05-1.81)
										≥20 PY v N	1.47 (1.13-1.91)
										≤19 AI v N	1.35 (1.10-1.67)

Ref.	Author, year	Study name, location	Population	Follow-up period (average follow-up)	Age	Sex[a]	Number of subjects	Site[b]	Number of cases	Exposure[c]	HR/RR (95% CI)
310	Hansen 2013	Diet, Cancer and Health Study, Denmark	General population	1995 - 2009 (13 yrs)	51-64	B	54,208	CRC	988	C v N	1.16 (0.96-1.39)
										F v N	1.16 (0.94-1.43)
258	Hurley 2013	California Teachers Study, USA	Teachers	1995 - 2009	22-104	F	122,264	CRC	1,205	C v N	1.28 (1.00-1.63)
										F v N	1.10 (0.97-1.24)
289	Nishihara 2013	Health Professionals Follow-up Study and Nurses' Health Study, USA	Health professionals and nurses	1980 - 2008	35-75	B	134,204	CRC	1,260	C v N	1.17 (0.96-1.43)
										F v N	1.18 (1.05-1.34)
										≥40 PY v N	1.28 (1.08-1.51)
										≥40 YSC v N	1.23 (0.99-1.54)

Ref.	Author, year	Study name, location	Population	Follow-up period (average follow-up)	Age	Sex[a]	Number of subjects	Site[b]	Number of cases	Exposure[c]	HR/RR (95% CI)
255	Parajuli 2013	The Oslo Study I The Norwegian Counties Study The 40 Years Cohort The CONOR Study, Norway	General population	1972 - 2007 (14 yrs)	19-67	M	299,376	CC	2,152	C v N	1.03 (0.92-1.15)
										F v N	1.14 (1.02-1.17)
										≥40 D v N	1.29 (1.05-1.59)
										≥20 I v N	1.16 (0.97-1.28)
										≥20 PY v N	1.14 (0.99-1.31)
										≤16 AI v N	1.15 (0.99-1.34)
						F	302,866	CC	1,846	C v N	1.22 (1.10-1.36)
										F v N	1.16 (1.02-1.31)
										≥40 D v N	1.47 (1.11-1.95)
										≥20 I v N	1.28 (1.06-1.55)
										≥20 PY v N	1.33 (1.11-1.57)
										≤16 AI v N	1.48 (1.21-1.81)
311	Doubeni 2012	NIH-AARP Diet and Health Study, USA	General population	1995 - 2006	50-71	B	506,488	CRC	7,676	>20 I v N (CO)	1.37 (1.22-1.53)
										>20 I v N (FO)	1.41 (1.33-1.50)

Ref.	Author, year	Study name, location	Population	Follow-up period (average follow-up)	Age	Sex[a]	Number of subjects	Site[b]	Number of cases	Exposure[c]	HR/RR (95% CI)
[260]	Gong 2012	Genetics and Epidemiology of Colorectal Cancer Consortium	5 cohort and 3 case-control studies		20-79	B	7,770*	CRC	6,796	C v N	1.26 (1.11-1.43)
										F v N	1.18 (1.09-1.27)
										≥40 D v N	1.28 (1.10-1.49)
										>20 I v N	1.28 (1.10-1.49)
										>60 PY v N	1.37 (1.16-1.62)
										≥35 YSC v N	0.74 (0.47-1.18)
										<40 AC v N	1.03 (0.79-1.34)
[253]	Leufkens 2011	European Prospective Investigation into Cancer and Nutrition, Europe	General population	1991 - 2006 (8.7 yrs)	35-70	B	465,879	CRC	2,741	C v N	1.08 (0.96-1.21)
										F v N	1.17 (1.07-1.29)
										≥40 D v N (CO)	0.94 (0.78-1.14)
										≥30 D v N (FO)	1.24 (1.07-1.44)
										≥15 I v N (CO)	1.07 (0.88-1.30)
										≥15 I v N (FO)	1.26 (1.08-1.47)
										≤16 AI v N (CO)	0.95 (0.77-1.17)
										≤16 AI v N (FO)	1.12 (0.96-1.31)
										≥25 YSC v N	1.08 (0.92-1.26)
[312]	Nordenval l 2011	Sweden	Construction workers	1971 - 2007 (24 yrs)	15-82	M	200,142	CC	1,367	≥25 D v N	1.12 (0.98-1.28)
										≥25 I v N	1.22 (0.94-1.60)
								RC	1,006	≥25 D v N	1.18 (1.01-1.39)
										≥25 I v N	1.21 (0.88–1.66)

Ref.	Author, year	Study name, location	Population	Follow-up period (average follow-up)	Age	Sex[a]	Number of subjects	Site[b]	Number of cases	Exposure[c]	HR/RR (95% CI)
138	Shin 2011	National Health Insurance Corporation Study, Korea	General population	1996 - 2003	30-80	M	869,725	PCC	536	C v N	1.0 (0.8-1.2)
										F v N	1.0 (0.8-1.3)
								DCC	751	C v N	0.9 (0.8-1.1)
										F v N	1.4 (1.2-1.7)
								RC	1,535	C v N	1.0 (0.9-1.1)
										F v N	1.1 (1.0-1.3)
						F	395,501	PCC	236	C v N	0.7 (0.4-1.2)
										F v N	1.1 (0.5-2.6)
								DCC	225	C v N	1.1 (0.7-1.9)
										F v N	0.7 (0.2-2.4)
								RC	551	C v N	1.0 (0.7-1.4)
										F v N	0.9 (0.5-1.7)
272	Limsui 2010	Iowa Women's Health Study, USA	General population	1986 - 2002	55-69	F	37,399	CRC	1,233	C v N	1.23 (1.03-1.46)
										F v N	1.16 (1.00-1.35)
										≥40 D v N	1.40 (1.17-1.68)
										>20 I v N	1.32 (1.04-1.69)
										≥40 PY v N	1.39 (1.14-1.70)
										≤30 AI v N	1.19 (1.05-1.36)

Ref.	Author, year	Study name, location	Population	Follow-up period (average follow-up)	Age	Sex[a]	Number of subjects	Site[b]	Number of cases	Exposure[c]	HR/RR (95% CI)
259	Gram 2009	Norwegian Women and Cancer Study, Norway	General population	1996 - 2005	30-69	F	68,160	CRC	425	C v N	1.2 (0.9-1.5)
										F v N	1.3 (1.0-1.6)
										≥30 D v N	1.3 (1.0-1.6)
										≥15 I v N	1.1 (0.7-1.7)
										≥20 PY v N	1.3 (1.0-1.8)
										<20 AI v N	1.3 (1.0-1.6)
										≥20 YSC v N	1.1 (0.8-1.5)
254	Hannan 2009	Cancer Prevention Study II Nutrition Cohort, USA	General population	1992 - 2005	50-74	B	124,751	CRC	1,962	C v N	1.27 (1.06-1.52)
										F v N	1.23 (1.11-1.36)
										≥50 D v N (CO)	1.38 (1.04-1.84)
										≥31 YSC v N	1.03 (0.89-1.19)
										<40 AC v N	1.05 (0.91-1.22)
313	Nöthlings 2009	Multiethnic Cohort Study, USA	General population	1993 -	45-75	B	1,522*	CRC	1,009	>30 PY v N	1.51 (1.17-1.95)

Ref.	Author, year	Study name, location	Population	Follow-up period (average follow-up)	Age	Sex[a]	Number of subjects	Site[b]	Number of cases	Exposure[c]	HR/RR (95% CI)
314	Hooker 2008	USA	General population	1963 - 1978 (15 yrs)	≥25	M	20,926	RC	66	C v N	3.05 (1.19-7.82)
										F v N	2.63 (0.98-7.05)
										≥40 D v N (CO)	4.40 (1.58-12.29)
										≥20 I v N (CO)	2.88 (0.98-8.48)
										≥40 PY v N (CO)	2.68 (0.97-7.44)
										<17 AI v N	3.33 (1.19-9.33)
						F	24,823	RC	54	C v N	0.93 (0.48-1.80)
										F v N	0.62 (0.19-2.08)
										≥40 D v N (CO)	1.00 (0.30-3.31)
										≥20 I v N (CO)	1.13 (0.33-3.91)
										≥40 PY v N (CO)	1.08 (0.41-2.85)
										<17 AI v N	0.83 (0.18-3.75)
				1975 - 1994 (19 yrs)	≥25	M	21,780	RC	77	C v N	1.80 (0.88-3.67)
										F v N	1.92 (0.98-3.78)
										≥20 I v N (CO)	2.27 (0.98-5.28)
						F	26,372	RC	77	C v N	1.57 (0.89-3.76)
										F v N	1.87 (1.02-3.45)
										≥20 I v N (CO)	2.84 (1.22-6.62)
315	Shankar 2008	Singapore Chinese Health Study, Singapore	General population	1993 - 2005	45-74	B	61,320	CRC	931	F v N	1.1 (0.92-1.4)
										>22 I v N (CO)	0.83 (0.56-1.2)

Ref.	Author, year	Study name, location	Population	Follow-up period (average follow-up)	Age	Sex[a]	Number of subjects	Site[b]	Number of cases	Exposure[c]	HR/RR (95% CI)
316	Weijenberg 2008	Netherlands Cohort Study on Diet and Cancer, Netherlands	General population	1989 - 1994 (5 yrs)	57-72	B	4,083†	CRC	648	C v N	0.81 (0.62–1.05)
										F v N	1.22 (0.97-1.53)
										>40 D v N	0.86 (0.65-1.13)
										>20 I v N	1.16 (0.89-1.52)
										>30 PY v N	1.12 (0.84-1.49)
										<17 AI v N	1.18 (0.89-1.56)
										>30 YSC v N	0.78 (0.45-1.33)
317	Akhter 2007	Miyagi Cohort Study, Japan	General population	1990 - 1997 (7 yrs)	40-64	M	21,695	CRC	188	C v N	1.47 (0.93-2.34)
										F v N	1.73 (1.04-2.87)
										≥40 D v N (CO)	1.59 (0.89-2.86)
										≥20 I v N (CO)	1.60 (0.99-2.58)
										≤18 AI v N (CO)	1.86 (0.97-3.58)
232	Driver 2007	Physicians' Health Study, USA	Physicians	1982 - 2004 (22 yrs)	40-84	M	21,581	CC	351	F v N	1.50 (1.19-1.89)
										≥40 I v N (CO)	1.53 (1.02-2.29)
								RC	100	F v N	1.13 (0.73-1.75)
										≥40 I v N (CO)	1.92 (1.01-3.66)

Ref.	Author, year	Study name, location	Population	Follow-up period (average follow-up)	Age	Sex[a]	Number of subjects	Site[b]	Number of cases	Exposure[c]	HR/RR (95% CI)
257	Paskett 2007	Women's Health Initiative, USA	General population	1993 - 2005 (7.8 yrs)	50-79	F	146,877	CRC	1,242	C v N	1.12 (0.86-1.46)
										F v N	1.13 (0.99-1.29)
										≥40 D v N	1.23 (0.97-1.55)
										≥25 I v N	1.41 (1.14-1.76)
										<20 AI v N	1.12 (0.95-1.31)
										≥40 YSC v N	0.93 (0.71-1.22)
										<30 AC v N	0.94 (0.72-1.23)
77	Tsong 2007	Singapore Chinese Health Study, Singapore	General population	1993 - 2004 (8.9 yrs)	45-74	B	61,321	CRC	845	C v N	1.10 (0.92-1.32)
										F v N	1.13 (0.91-1.39)
										≥40 D v N	1.19 (0.97-1.45)
										≥13 I v N	1.18 (0.98-1.43)
										<15 AI v N	1.32 (1.02-1.71)

Ref.	Author, year	Study name, location	Population	Follow-up period (average follow-up)	Age	Sex[a]	Number of subjects	Site[b]	Number of cases	Exposure[c]	HR/RR (95% CI)
318	Yun 2005	National Health Insurance Corporation Study, Korea	General population	1996 - 2000 (4 yrs)	≥30	M	733,134	CC	417	C v N	0.81 (0.63-1.05)
										F v N	1.37 (1.06-1.77)
										≥30 D v N (CO)	0.96 (0.69-1.33)
										≥30 D v N (FO)	2.08 (1.29-3.37)
										≥20 I v N (CO)	0.76 (0.51-1.15)
								RC	453	C v N	0.97 (0.76-1.24)
										F v N	1.17 (0.91-1.52)
										≥30 D v N (CO)	1.12 (0.82-1.52)
										≥30 D v N (FO)	0.61 (0.27-1.41)
										≥20 I v N (CO)	1.05 (0.74-1.50)
319	Jee 2004	Korean Cancer Prevention Study, Korea	General population	1993 - 2001	30-95	M	830,139	CC	1,633	C v N	0.8 (0.7-1.0)
										F v N	1.1 (1.0-1.3)
144	Sanjoaquin 2004	Oxford Vegetarian Study, UK	General population	1980 - 1999 (17 yrs)	16-89	B	10,998	CRC	95	C v N	1.70 (0.92-3.15)
										F v N	1.80 (1.13-2.85)
120	Otani 2003	Japan Public Health Center-based Prospective Study, Japan	General population	1990 - 1999	40-70	M	42,540	CRC	447	C v N	1.4 (1.1-1.8)
										F v N	1.3 (0.98-1.7)
										≥40 PY v N (CO)	1.4 (0.99-1.8)
						F	47,464	CRC	259	C v N	1.4 (0.8-2.4)
										F v N	1.3 (0.5-3.6)

Ref.	Author, year	Study name, location	Population	Follow-up period (average follow-up)	Age	Sex[a]	Number of subjects	Site[b]	Number of cases	Exposure[c]	HR/RR (95% CI)
146	Shimizu 2003	Takayama Study, Japan	General population	1993 - 2000 (8 yrs)	≥36	M	13,392	CC	104	>20 PY v N	1.37 (0.81-2.32)
								RC	57	>20 PY v N	2.44 (1.12-5.30)
						F	15,659	CC	77	>10 PY v N	0.77 (0.30-1.96)
								RC	38	>10 PY v N	0.94 (0.21-4.16)
320	Wakai 2003	Japan Collaborative Cohort Study, Japan	General population	1988 - 1997 (7.6 yrs)	40-79	M	25,260	CC	219	C v N	1.23 (0.85-1.78)
										F v N	1.07 (0.72-1.59)
										≥40 D v N	1.07 (0.71-1.61)
										≥40 I v N	0.69 (0.33-1.43)
										≥60 PY v N	0.68 (0.34-1.37)
										<20 AI v N	1.04 (0.62-1.74)
										≥20 YSC v N	0.79 (0.41-1.52)
								RC	147	C v N	0.83 (0.55-1.26)
										F v N	0.88 (0.56-1.39)
										≥40 D v N	0.72 (0.45-1.16)
										≥40 I v N	0.80 (0.38-1.69)
										≥60 PY v N	0.78 (0.38-1.59)
										<20 AI v N	1.18 (0.69-1.99)
										≥20 YSC v N	0.53 (0.22-1.28)
						F	34,619	CC	189	C v N	1.06 (0.55-2.02)
										F v N	1.07 (0.39-2.92)
								RC	57	C v N	0.36 (0.05-2.65)
										F v N	1.05 (0.14-7.69)

Ref.	Author, year	Study name, location	Population	Follow-up period (average follow-up)	Age	Sex[a]	Number of subjects	Site[b]	Number of cases	Exposure[c]	HR/RR (95% CI)
321	Terry 2002	Canadian National Breast Screening Study, Canada	Trial of mammography screening	1980 - 1993 (10.6 yrs)	40-59	F	89,835	CRC	527	C v N	1.00 (0.80-1.26)
										F v N	1.15 (0.94-1.41)
										≥40 D v N	1.30 (0.68-2.47)
										≥40 I v N	0.71 (0.35-1.43)
322	Tiemersma 2002	Monitoring Project on Cardiovascular Disease Risk Factors, Netherlands	General population	1987 - 1998	20-59	B	535*	CRC	102	C v N	0.9 (0.5-1.7)
										F v N	1.4 (0.8-2.5)
										>25 D v N	1.2 (0.7-2.1)
										>14 I v N	1.5 (0.9-2.6)
										>15 YSC v N	1.1 (0.5-2.3)
323	Terry 2001	Sweden	General population	1967 - 1997 (22.5 yrs)	42-81	B	17,118	CRC	498	F v N	1.0 (0.8-1.4)
										≥21 I v N (CO)	3.1 (1.4-7.1)
324	Stürmer 2000	Physicians' Health Study, USA	Physicians	1982 - 1995 (13 yrs)	40-84	M	22,011	CRC	351	C v N	1.81 (1.28-2.55)
										F v N	1.49 (1.17-1.89)
										≥20 I v N (CO)	2.14 (1.45-3.14)
										≥20 I v N (FO)	1.31 (1.00-1.73)
										>40 PY v N	1.68 (1.20-2.35)
325	Knekt 1998	Finland	General population	1966 - 1994	≥15	B	56,973	CRC	457	F v N	1.02 (0.74-1.39)
										≥15 I v N (CO)	1.04 (0.73-1.48)

Ref.	Author, year	Study name, location	Population	Follow-up period (average follow-up)	Age	Sex[a]	Number of subjects	Site[b]	Number of cases	Exposure[c]	HR/RR (95% CI)
147	Singh 1998	Adventist Health Study, USA	Seventh-day Adventists	1976 - 1983	≥25	B	32,051	CRC	157	C v N	1.39 (0.50-3.82)
										F v N	1.13 (0.75-1.70)
326	Kato 1997	New York University Women's Health Study, USA	General population	1985 - 1994 (7.1 yrs)	34-65	F	14,727	CRC	85	C v N	0.97 (0.53-1.77)
										F v N	0.99 (0.62-1.60)
327	Nordlund 1997	Sweden	General population	1963 - 1989 (26 yrs)	18-69	F	26,032	CRC	559	C v N	0.88 (0.67-1.16)
										F v N	1.16 (0.72-1.86)
										≥16 I v N (CO)	1.42 (0.77-2.60)
										≤19 AI v N (CO)	0.97 (0.49-1.92)
246	Tulinius 1997	Cardiovascular Risk Factor Study, Iceland	General population	1968 - 1995	33-60	F	11,580	CRC	145	F v N	1.12 (0.68-1.86)
										≥25 I v N (CO)	2.48 (0.99-6.19)
74	Chyou 1996	USA	General population	1965 - 1995	46-65	M	7,944	CC	330	C v N	1.42 (1.09-1.85)
										F v N	1.27 (0.95-1.70)
										≥31 PY v N	1.48 (1.13-1.94)
								RC	123	C v N	1.95 (1.25-3.04)
										F v N	1.31 (0.78-2.20)
										≥31 PY v N	1.92 (1.23 -2.99)

Ref.	Author, year	Study name, location	Population	Follow-up period (average follow-up)	Age	Sex[a]	Number of subjects	Site[b]	Number of cases	Exposure[c]	HR/RR (95% CI)
328	Engeland 1996	Norway	General population	1966 - 1993	33-72	M	11,863	CC	230	C v N	1.2 (0.8-1.6)
										F v N	1.0 (0.6-1.5)
								RC	139	C v N	1.6 (1.0-2.6)
										F v N	0.8 (0.4-1.6)
						F	14,269	CC	300	C v N	1.1 (0.8-1.4)
										F v N	1.3 (0.9-2.0)
								RC	141	C v N	0.8 (0.5-1.3)
										F v N	1.3 (0.8-2.4)
329	Nyren 1996	Sweden	Construction workers	1971 - 1991 (17.6 yrs)		M	134,985	CC	713	C v N	0.98 (0.82-1.17)
										F v N	1.02 (0.84-1.24)
										≥41 D v N (CO)	0.99 (0.72-1.35)
										≥21 D v N (FO)	1.08 (0.84-1.40)
										≥25 I v N	1.07 (0.63-1.82)
								RC	505	C v N	1.16 (0.94-1.44)
										F v N	1.22 (0.97-1.54)
										≥41 D v N (CO)	1.08 (0.73-1.60)
										≥21 D v N (FO)	1.06 (0.76-1.48)
										≥25 I v N	1.08 (0.58-2.03)

Ref.	Author, year	Study name, location	Population	Follow-up period (average follow-up)	Age	Sex[a]	Number of subjects	Site[b]	Number of cases	Exposure[c]	HR/RR (95% CI)
269	Giovannucci 1994	Health Professionals Follow-up Study, USA	Health professionals	1986 - 1990 (4 yrs)	40-75	M	47,935	CRC	238	C v N	1.36 (0.84-2.20)
								CRC	238	≥40 PY vs N	1.45 (1.01-2.09)
270	Giovannucci 1994	Nurses' Health Study, USA	Nurses	1976 - 1990 (14 yrs)	30-55	F	118,334	CRC	586	C v N	0.91 (0.74-1.12)
75	Klatsky 1988	USA	General population	1978 - 1984		B	10,572†	CC	203	F v N	1.03 (0.74-1.43)
										≥20 I v N (CO)	1.35 (0.78-2.35)
								RC	66	F v N	1.28 (0.71-2.28)
										≥20 I v N (CO)	1.01 (0.37-2.79)
76	Wu 1987	USA	Retirement community	1981 - 1985		M	4,163	CRC	58	C v N	1.80 (0.6-5.2)
										>20 YSC v N	1.71 (0.8-3.6)
						F	7,456	CRC	68	C v N	1.35 (0.7-1.0)
										>20 YSC v N	1.61 (0.8-3.0)

[a] M = male, F = female, B = both

[b] CRC = colorectal cancer, CC = colon cancer, RC = rectal cancer, PCC = proximal colon cancer, DCC = distal colon cancer

[c] C = current smokers, F = former smokers, N = never smokers, D = duration, I = intensity, PY = pack-years, AI = age at initiation, YSC = years since cessation, AC = age at cessation, CO = current smokers only, FO = former smokers only

* Case-control analysis

† Case-cohort analysis

2.5 Thesis Objectives

The main aim of this thesis was to investigate the relationships between alcohol intake, adiposity and smoking and the risk of colorectal cancer (and colorectal cancer subsites) using data from UK Biobank. These literature reviews identified a number of important gaps in the existing literature for these relationships that this thesis will aim to answer. Thus, the objectives of this thesis are:

i. To examine the shape of the relationship between alcohol intake and colorectal cancer and the possibility of a threshold at 30 g/d.

ii. To investigate whether the relationship between alcohol intake and colorectal cancer differs for men and women.

iii. To explore how the choice of different reference groups can impact on the results for the association between alcohol intake and colorectal cancer.

iv. To investigate whether the level of BMI affects the relationship between alcohol intake and colorectal cancer.

v. To explore reasons why BMI is more strongly related to colorectal cancer risk for men than for women.

vi. To investigate how different measures of adiposity are related to colorectal cancer risk.

vii. To explore reasons why former and current smokers experience a similar risk of colorectal cancer.

viii. To investigate how the effects of smoking duration and smoking intensity increase the risk of colorectal cancer.

ix. To investigate how the risk of colorectal cancer differs according to smoking cessation at different times.

Chapter 3 UK Biobank

In this chapter, the UK Biobank cohort is described in detail including its creation, the recruitment of participants and the data available. At the end, characteristics of the cohort are presented.

Since the middle of the previous century, epidemiological studies have provided vital evidence for the prevention of disease through environmental and lifestyle factors that has had a tremendous impact on public health. The role of genetics in disease has also been the focus of much research, particularly following the success of the Human Genome Project at the turn of the century.[330, 331] However, it has been recognised for many years that an individual's risk of disease depends on the complex interplay between lifestyle, environment and genetic factors.[332] Therefore, a number of large, prospective studies including detailed information on lifestyle, environment and genetics have been established in recent years with the aim to improve the understanding of the major determinants of complex diseases and why some people develop certain diseases and others do not.

Prospective cohort studies have a number of important advantages in comparison to retrospective case-control studies for the investigation of lifestyle, environment and genetic factors in relation to a variety of diseases.[54, 55, 333] First of all, since the exposures are assessed before the onset of disease, prospective studies avoid recall bias where the presence of disease may affect participants' responses. Moreover, prospective studies are able to investigate diseases that cannot be investigated using retrospective studies such as dementia or rapidly fatal conditions. Prospective studies also allow the investigation of a wide range of diseases whereas case-control studies by definition are only able to investigate a single disease. This means that prospective cohorts can be used to investigate the overall effects of a single exposure on multiple diseases. However, prospective studies require significant investment since large numbers of people need to be followed-up for many years in order to accumulate sufficient numbers of disease cases.

The UK Biobank cohort was established by the Medical Research Council and the Wellcome Trust with the aim to investigate how the complex interplay between

lifestyle, environment and genetics increases the risk of a wide range of diseases.[53, 334] UK Biobank is a very large and detailed prospective study including over 500,000 men and women aged mostly 40-69 recruited between 2006 and 2010. At baseline participants provided detailed information on their lifestyle, health and environmental factors, completed physical measures and provided biological samples. Follow-up on health related outcomes including deaths, cancers and hospital admissions is achieved through linkage with national health datasets.

Several other large-scale prospective cohorts have been established in recent years with information on lifestyle, environment and genetics including the 500,000 person China Kadoorie Biobank, the 500,000 person European Prospective Investigation into Cancer and Nutrition and the 150,000 person Mexico City Prospective Study.[335-337] In comparison to these cohorts, UK Biobank aims to support a much wider range of health related research and includes detailed information on a wider range of exposures and health related outcomes.

UK Biobank represents the most detailed large-scale cohort ever established and it is hoped that this unique combination of large size and detailed information means that this ambitious cohort remains a rich resource for epidemiological research for many years to come and is able to make a real difference to the health of future generations.

3.1 Initiation

The initiation of UK Biobank goes back to May 1999 when the Medical Research Council and the Wellcome Trust hosted a workshop to discuss the potential value of establishing a large prospective cohort that would be able to investigate lifestyle, environment and genetic factors in relation to a wide range of complex diseases.[338, 339] Soon after, both the Medical Research Council and the Wellcome Trust agreed in principle to fund the project, then known as the UK Population Biomedical Collection. A protocol for the cohort was developed based on wide consultation with specialist experts as well as public consultation. Following positive international peer review from 12 experts, a full protocol for the project was published in February 2002. Initial funding of £45 million (£20 million each from the Medical Research Council and the

Wellcome Trust and £5 million from the Department of Health) was announced in April 2002.[339]

3.2 Pilot Studies

Two pilot studies were conducted before main recruitment began. A small pilot study was carried out in early 2005 across six centres recruiting about 300 people. The principal aim of this initial pilot study was to evaluate the feasibility and suitability of the assessment methods and to identify areas for improvement.

A more comprehensive pilot study was conducted between March and June 2006.[338, 340] This pilot study assessed about 4,000 people at an assessment centre in Stockport. This time the main aim was to assess the entire recruitment process from invitation to assessment to sample collection and storage at the throughput required for the main recruitment. These participants were included in the final UK Biobank cohort.

Names and addresses of people aged 40-69 living within the vicinity of the assessment centre were sought from four local National Health Service primary care trusts. These people were sent an invitation letter with a provisional appointment plus an information leaflet and a pre-paid reply form. A freephone service was also available for participants to change their appointment or ask any questions. In total 59,383 primary invitation letters were sent.[340]

Whilst the requested data was provided rapidly by one trust, data from another trust was delayed by several weeks and data from the other two trusts could not be obtained before the end of the pilot study. This presented a number of issues for recruitment. First, it meant that invitation letters were sent only three to four weeks before the provisional appointments instead of the proposed six to eight weeks. The delays in accessing contact details also meant that there was an uneven pattern of invitations which caused large spikes in demand for the call centre staff and increased waiting times which may have negatively affected participation. Furthermore, the two trusts that did not provide data actually covered the area immediately adjacent to the assessment centre which resulted in few people living within two miles being invited who may have

been more likely to participate. These problems highlighted the importance of identifying participants from national data sources to ensure smooth operation.

Overall, about 4,000 participants were recruited in the pilot study, giving a response rate of approximately 6.7%. A higher response rate was expected for the main study since the response rate in the pilot study was affected by the difficulties with the primary care trusts. It was also found that the response rate could be improved by sending pre-visit reminders; the non-attendance rate was halved from 20% to 10% using pre-visit reminders.[340] Also, a number of people were unable to arrange a convenient appointment at the end of recruitment in the pilot study which negatively affected the response rate; this should have impacted the response rate in the main study less since assessment centres were open longer. Greater promotion of UK Biobank during the main study should also have encouraged greater participation.

A 10% random sample of people who attended the assessment centre were asked to complete an anonymous postal survey about the assessment visit. 65% of people responded.[340] Responses showed that participants had a good understanding of what the study would involve and that participants found the amount of information asked and the length of the assessment visit acceptable (though there was a tendency to say it was too long and virtually no participants said it was too short). There was also a good understanding of what they were agreeing to by consenting to take part in the study. One undesirable finding was that 29% of participants mentioned "to have a health check" as one of the main reasons for participating in UK Biobank. Thus, following the pilot study, the information included in the invitation materials was revised for the main study.

Approximately 10,000 people declined to participate in the pilot study (most people did not respond) using either the pre-paid response form or by phone and approximately 70% of these people were willing to provide reasons. The most common reasons given were too busy (mentioned by 2,166 people), too unwell (935 people) and too far to travel/too inconvenient (531 people).[340]

Following the pilot study, a revised protocol was assessed by the Wellcome Trust's Study Design Expert Group, the independent Ethics and Governance Council, and a

specially convened International Review Panel.[338] The panel concluded that "UK Biobank has the potential, in ways that are not currently available elsewhere, to support a wide range of research, particularly investigations into complex interactions of various exposures, including genetic and lifestyle factors in the pathways to disease and health"[341] and unanimously recommended that full scale recruitment should begin without delay.[338] In August 2006 the study was approved with £61 million funding from the Medical Research Council, the Wellcome Trust, the Department of Health, the Scottish Government and the North West Regional Development Agency.[341] Main recruitment began in April 2007 and lasted over three years with the final participant recruited in October 2010.

3.3 Main Recruitment

3.3.1 Identification and Invitation

In the UK virtually all members of the general population are registered with a general practitioner (GP) through the National Health Service (NHS). NHS records were used to identify people aged 40-69 years old to invite to participate in UK Biobank.[53] The data made available to UK Biobank about potential participants were restricted to name, address, sex, date of birth, NHS number and name and address of GP.[338]

People aged 40-69 years old living nearby to one of the 22 assessment centres were sent an invitation letter with a provisional appointment to participate in UK Biobank. Invitations were sent at least six to eight weeks before the provisional appointment.

With the invitation letter, people also received an information leaflet which gave more detailed information on the purpose of UK Biobank, how people had been identified for invitation, what consenting to take part would mean (including being able to withdraw at anytime) and what the assessment visit would involve. People were advised that the assessment visit would last about two to three hours. People who wanted to find out more information were encouraged to telephone the Participant Resource Centre (PRC) free of charge or visit the study website.

To accept the provisional appointment or decline to take part people could phone the PRC, return the reply form using the pre-paid envelope or go to UK Biobank's website. People could change the provisional appointment to a more convenient time by calling the PRC; most assessment centres were open Monday to Friday from 8am to 7pm and Saturday from 8am to 5pm. If people confirmed an appointment they were asked to provide a telephone number or e-mail address so that they could receive a reminder closer to the date. Also people who confirmed an appointment were sent written confirmation including a map with directions to the assessment centre and a short pre-visit questionnaire for questions that might be more difficult to answer precisely on the day.

3.3.2 Assessment Centres

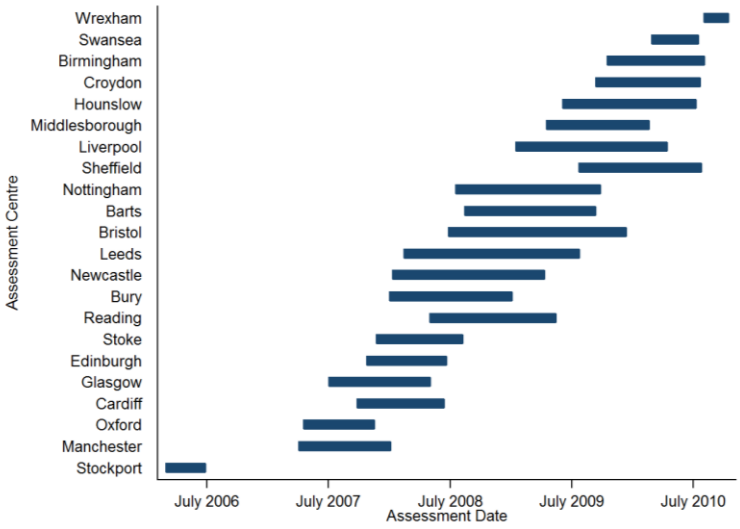
The first requirement for establishing an assessment centre was a large number of people aged 40-69 years old living within close proximity in order to maintain a high throughput of participants and reduce recruitment costs. Therefore, all assessment centres were generally located in large cities. It was also important that assessment centres were conveniently located for public transport links, nearby parking and easy disabled access. Figure 3.3.1 shows a map of the 22 assessment centres used for recruiting participants.

Figure 3.3.1 Map of UK Biobank Assessment Centres

Taken from www.biobank.ctsu.ox.ac.uk/crystal/

The recruitment process was coordinated centrally with up to six assessment centres open at any time.[53] Each assessment centre recruited around 100 participants per day and most assessment centres were open approximately 8-14 months. Figure 3.3.2 shows the timeline of recruitment for each assessment centre.

Figure 3.3.2 Timelines for Recruitment by UK Biobank Assessment Centre



Period of recruitment at each UK Biobank assessment centre.

3.3.3 Assessment

The baseline assessment consisted of a touchscreen questionnaire, a verbal interview with a trained member of staff and a series of physical measures. The data available from the baseline assessment are summarised in Table 3.3.1. A number of questions or measures were added during recruitment and so were available only for a subset of the entire cohort.

Participants first provided consent to be part of UK Biobank using the touchscreen and a digital signature pad. Participants were given a physical copy of this consent form. Then participants completed the touchscreen questionnaire which included detailed questions on socio-demographics, lifestyle, environment and health. The touchscreen questionnaire was important in enabling the recruitment of a large number of participants at a relatively low cost since a number of participants could complete the questionnaire at the same time with minimal supervision from staff. It also meant participants could be guided through the questionnaire in order to only answer the

questions pertinent to them and it allowed validity checks for questions e.g. not accepting extreme values or asking participants if they are sure. The full touchscreen questionnaire can be found on the UK Biobank website (www.ukbiobank.ac.uk) and information on each variable can be viewed in the interactive data showcase available on the study website.

Table 3.3.1 Data Collected at the UK Biobank Baseline Assessment[53]

Questionnaire and interview	
Sociodemographic	Social class; ethnicity; employment status; marital status; education; income; car ownership
Family history and early life exposures	Family history of major diseases; birth weight; breast feeding; maternal smoking; childhood body size; residence at birth
Psychosocial factors	Neurosis; depression (including bi-polar spectrum disorder); social support
Environmental factors	Current address; current (or last) occupation; domestic heating and cooking fuel; housing; means of travel; shift work; mobile phone use; sun exposure
Lifestyle	Smoking; alcohol consumption; physical activity; diet; sleep
Health status	Medical history; medications; disability; hearing; sight; sexual and reproductive history
Hearing threshold	Speech reception threshold*
Cognitive function	Pairs matching; reaction time; prospective memory*; fluid intelligence*; numeric memory†
Physical measures	
Blood pressure and heart rate	two automated measures, one minute apart
Grip strength	Left- and right-hand grip strength
Anthropometrics	Standing and sitting height; weight and bio-impedance; hip and waist circumference
Spirometry	Up to three measures
Bone density‡	Calcaneal ultrasound
Arterial stiffness¶	Pulse wave velocity
Eye examination§	Refractive index, intraocular pressure; acuity; retinal photograph; optical coherence tomography
Fitness test§	Cycle ergometry with electrocardiogram (ECG) heart rate monitoring

* assessed in 170,000 participants;
† assessed in 50,000 participants;
‡measured in one heel for 170,000 participants and in both heels for 320,000 participants;
¶ measured in 170,000 participants;
§ measured in 100,000 participants

doi:10.1371/journal.pmed.1001779.t002

Next was the verbal interview with a member of staff designed to gain more information on early life factors, employment, medications and medical history. Two measurements of blood pressure were taken at this stage. Participants then went to the physical measurements station. Physical measures included height, hip circumference, WC, weight, bioelectrical impedance, hand grip strength, ultrasound bone densitometry and spirometry. The final stage of the assessment centre was to provide biological samples.

Upon leaving the assessment centre participants were provided with some basic feedback on their physical measures. While it was important that feedback be kept to a minimum so that participants did not sign up for UK Biobank thinking it was a health check, it was seen as impractical and inappropriate to conceal information on physical measures from participants. So participants received a hard copy of some of the physical measurements such as blood pressure, weight, BMI and heel bone ultrasound along with

a simple interpretation (e.g. good/borderline/high) and information on seeking further advice.

3.4 Response

The overall response rate for UK Biobank was 5.5%.[334] Achieving a high response rate was never an objective of UK Biobank; planning for recruitment was based on a response rate of 10%.[338]

It is widely accepted that people are generally more reluctant to take part in epidemiological studies nowadays and response rates have been declining for a number of years.[342] Furthermore, the UK Biobank assessment was particularly demanding; participants were asked to take three hours out of their day to take part in the study.

However, although a low response rate was expected for UK Biobank, it is not clear why the study failed to achieve the projected 10% response or why the response rate was lower than that of the pilot study.

Unfortunately, there is no detailed information available about the recruitment of participants that could possibly explain why the target response rate was not achieved. The pilot report did include some basic information on factors affecting confirmation rates (based on invitation prior to 1st May 2006 to avoid the impact of the assessment centre closure on confirmation rates).[340] For example, younger people and people from more deprived areas were less likely to confirm an appointment. Hence, one possibility for the low response rate would be if these people were oversampled in the main recruitment.

There are also other factors that may have contributed to the low response rate. For example, it seems that there was not much local promotion of UK Biobank during recruitment. Hence, the mailed invitation would likely be the first occasion when most people learned of the study and many people may have simply ignored it. The lack of promotion may have been due to constraints on resources i.e. it was more economical to have a low response rate and invite greater numbers of people than to increase the response rate by investing in promotion.

The original protocol mentioned that UK Biobank would recruit participants from 35 assessment centres, each open for about six months.[338] In reality, there were fewer assessment centres that were generally open much longer. Again, it is possible that it was decided to keep fewer centres open for longer as it was more economical than establishing new centres though of course this will have negatively affected response rates as more people living further away from the assessment centres would have to have been invited.

Early reports suggested that UK Biobank was achieving a response rate of approximately 10%.[343, 344] It is not known whether this does mean that there was a drop-off in response rates in later assessment centres or perhaps response rates were calculated slightly differently. Unfortunately, there is no detailed information about response rates by assessment centre, although it was reported that the response rate achieved in Bristol was much higher than in other assessment centres.[344, 345] Though it cannot be known for certain, this was thought to be related to the greater awareness of the benefits of cohort studies as a result of the locally-based Avon Longitudinal Study of Parents and Children.[346] This is interesting because it indicates that although response rates for epidemiological research are declining, there may be factors that can increase response rates, in particular through increasing awareness of the benefits of such research.

3.5 Representativeness

The aim of all epidemiological research is to acquire knowledge that can be applied to improve the health of the population. To do this, some form of investigation is undertaken in a sample of the population and it is hoped that the evidence found can be generalised to the population of interest. Investigators must always think carefully about any potential reasons why the results found in the sample may not be generalisable to the larger population. Representativeness of the sample has often been thought of as a crucial component of generalisability but this may not be the case.[347]

First of all, representativeness is vitally important in certain contexts. The validity of the conclusions from censuses and surveys are completely dependent on the representativeness of the sample used.[348]

However, cohort studies are used to investigate associations between risk factors and disease and are not used to describe the characteristics of a larger population. A distinction should first be made between representativeness and response rates. A high response rate obviously signifies that the recruited participants will closely resemble the population of interest, provided that the sample of people approached was adequately drawn. However, a low response rate does not necessarily signify a non-representative sample or cohort. A cohort with a low response rate could still be representative if each member of the population of interest had an equal probability of taking part. In reality though a low response rate generally signifies high levels of self-selection among participants which will result in non-representativeness.

The British Doctors Study is often mentioned when talking about representativeness.[42] Male doctors are clearly not representative of the general population. However, the findings of this study on smoking and health have been considered broadly applicable to the general population. In fact, the non-representativeness is often considered as a strength of the study since it meant that smokers and non-smokers were very similar in most regards except for the exposure of interest.

The only reason why non-representativeness may result in results that are not generalisable to a larger population is if the association between exposure and outcome was modified by a third factor which differed in prevalence between the cohort and the population of interest. In this case, it would be necessary to investigate the association according to exposure to this third factor and representativeness could actually be detrimental for this. For example, if an association between exposure and outcome were modified by ethnicity, a representative sample would likely be unable to provide accurate estimates of the association within different ethnic groups since there will be only small numbers of ethnic minorities. The most efficient thing to do from a statistical point of view when an association is modified by another factor, rather than recruiting a representative sample, would be to recruit equal numbers of people from each subgroup of interest.

In other words, results from a representative cohort would represent the overall average association in the general population. However, this does not mean that this association is equally generalisable to everybody in the general population. Generalisation includes

an understanding of the full set of circumstances in which results apply and when they do not.[347] For example, some hypothetical exposure could be equally harmful to men as it is beneficial to women but this would not mean that the exposure is safe in the general population. Furthermore, representativeness is not a static idea. Surveys must be continuously updated. If the proportion of men in the general population were to increase, then the hypothetical exposure would appear to be harmful. However, the association for men and the association for women would not change.

Also, the self-selection of participants at baseline may be advantageous if it leads to a more motivated group of participants who are more willing to undergo further measures. Thus, the low response rates at baseline should be balanced against the response rates for follow-up measures.

UK Biobank is not a representative cohort but representativeness was not a priority for UK Biobank and heterogeneity was emphasised over representativeness. UK Biobank is of a sufficient size that it will be possible to obtain reliable estimates of risk factor-disease associations according to a number of subgroups (even though the proportions of people represented in these subgroups may not match those in the general population).

3.6 Further Measurements

A number of further assessments have been carried out (and more will be carried out in the future) in order to further enhance the UK Biobank cohort. There are two reasons for conducting further assessments; one is to gain information on changes in participants' exposure over time and the other is to expand the amount of data available in order to enable an even broader scope for research.

The first of these enhancements was the 24 hour recall dietary questionnaire. This web questionnaire asked participants about their food and drink intake during the previous 24 hours. It included questions on about 200 commonly consumed foods and drinks and took about 15 minutes to complete. This questionnaire was added to the assessment centre towards the end of recruitment. It was later sent out to all participants (approximately 320,000) who provided UK Biobank with an e-mail address on four

separate occasions over one year between 2011-2012 in order to account for seasonal variation and gain a more reliable estimate of a participant's average diet. Approximately 210,000 participants completed at least one questionnaire and 80,000 participants completed at least three questionnaires.

Between August 2012 and June 2013 (on average approximately five years after the baseline assessment), 20,000 participants were completely re-assessed at the UK Biobank co-ordinating centre in Stockport. The response rate for this re-assessment was 21%.[349] Similar re-assessments of subsets of participants are planned to be carried out roughly every three to five years.

Starting in June 2013 participants were invited by e-mail to wear a physical activity monitor for a week. Participants who agreed to take part were sent a device to wear on their wrist continuously for seven days before returning the device to UK Biobank in a pre-paid envelope. The data obtained from these devices will be used to provide an objective measure of physical activity. Approximately 100,000 participants agreed to wear the physical activity monitors.

The UK Biobank Imaging Study began in May 2014 and plans to conduct detailed imaging scans of the vital organs of 100,000 participants. The scans will include the brain, heart, tissue and bones. These data available on such a large number of people is unprecedented and will further support research into a wide range of diseases.

3.7 Follow-up

Follow-up data on health outcomes is achieved mainly through linkage with routinely available national health datasets.[53] Data are currently available on cancer diagnoses and deaths for all participants as well as data on hospital in-patient episodes. Data on cancer diagnoses are provided to UK Biobank by the Health and Social Care Information Centre for participants in England and Wales and by the Information Services Division of NHS National Services Scotland for participants in Scotland. Data on deaths are provided to UK Biobank by the Health and Social Care Information Centre for participants in England and Wales and by the National Health Service Central Register, Scotland for participants in Scotland. Follow-up data are updated periodically and

information from further datasets will be added in the future. The follow-up data available on cancers and deaths used in the analyses for this thesis are described in further detail in section 4.5.

3.8 Characteristics of UK Biobank Cohort

Complete data on baseline and follow-up for cancers and deaths were available for 502,642 participants (229,175 men and 273,467 women). Table 3.8.1 shows the dates each assessment centre was open and also the number of participants recruited from each assessment centre.

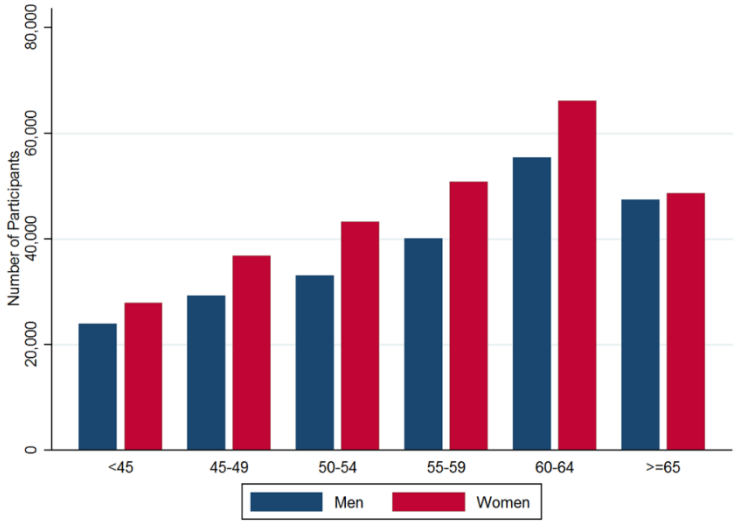
Table 3.8.1 Assessment Dates and Participants by Assessment Centre

Assessment centre	Recruitment dates	Participants
Stockport	13/03/2006 - 14/06/2006	3,798
Manchester	16/04/2007 - 22/12/2007	13,940
Oxford	30/04/2007 - 03/11/2007	14,062
Cardiff	08/10/2007 - 31/05/2008	17,883
Glasgow	16/07/2007 - 19/04/2007	18,653
Edinburgh	07/11/2007 - 07/06/2007	17,201
Stoke	05/12/2007 - 26/07/2008	19,440
Reading	14/05/2008 - 02/05/2009	29,421
Bury	14/01/2008 - 20/12/2008	28,335
Newcastle	23/01/2008 - 28/03/2009	37,009
Leeds	27/02/2008 - 11/07/2009	44,211
Bristol	09/07/2008 - 28/11/2009	43,017
Barts	27/08/2008 - 29/08/2008	12,584
Nottingham	30/07/2008- 12/09/2009	33,880
Sheffield	05/08/2009 - 13/07/2009	30,397
Liverpool	28/01/2009 - 01/04/2010	32,822
Middlesborough	29/04/2009 - 06/02/2010	21,289
Hounslow	17/06/2009 - 26/06/2010	28,880
Croydon	24/09/2009 - 09/07/2010	27,386
Birmingham	29/10/2009 - 21/07/2010	25,503
Swansea	11/03/2010 - 03/07/2010	2,282
Wrexham	16/08/2010 - 01/10/2010	649

Participants were mostly aged 40-69 at baseline; the age range of the cohort was 37-73; 2,419 participants were aged 70, seven participants were aged 71-73 and seven

participants were aged 37-39. The age distribution of the cohort is shown in Figure 3.8.1. 61% of participants were aged 55 or over.

Figure 3.8.1 Age Distribution of UK Biobank Participants by Sex



94% of participants reported a white ethnic background (Table 3.8.2). The 2011 national census found that the white ethnic group accounted for 86.0% of the population, decreasing from 91.3% in 2001.[350]

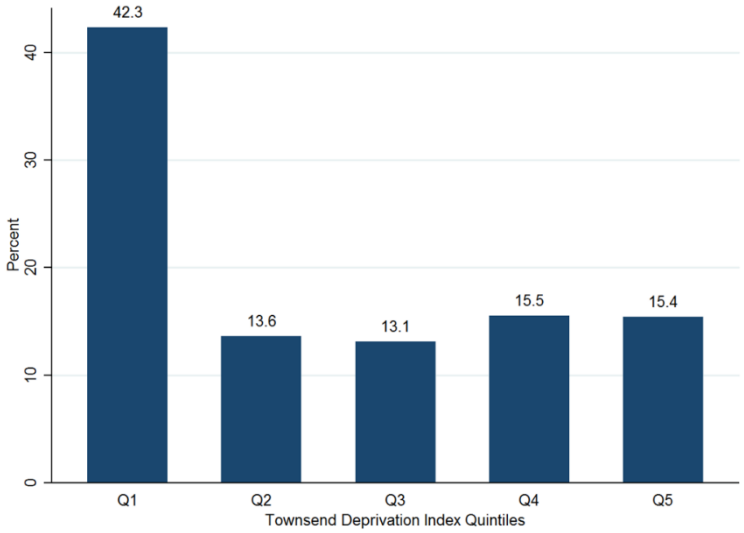
Table 3.8.2 Ethnic Background of UK Biobank Participants

Ethnic background	Participants	%
White	472,837	94.2
Asian	9,882	2.0
Black	8,066	1.6
Chinese	1,574	0.3
Mixed	2,958	0.6
Other ethnic group	4,560	0.9
Do not know	217	0.0
Prefer not to answer	1,662	0.3

Participants were assigned a Townsend deprivation index score based on data from the 2001 national census. Participants were assigned the score corresponding to the output area (lowest geographical area for which census data are calculated) in which their postcode at baseline was located. Townsend deprivation index scores from the 2001 census for England and Wales by ward are available at census.ukdataservice.ac.uk/get-data/related/deprivation. These data were used to calculate quintiles of Townsend

deprivation in the general population. The percentage of participants included in each quintile is shown in Figure 3.8.2. A higher Townsend score represents a higher level of deprivation. Thus, 42% of participants lived in areas in the lowest quintile of deprivation.

Figure 3.8.2 Distribution of Townsend Deprivation Index Score by National Quintiles of Deprivation



Proportion of UK Biobank participants within quintiles of Townsend deprivation index based on national distribution of Townsend scores. Quintile 1 represents the lowest levels of deprivation and quintile 5 represents the highest levels of deprivation.

For alcohol intake, smoking, BMI and physical activity, participants' responses were compared with data from the 2008 Health Survey for England (HSE).[351] Data from the 2008 HSE were acquired from www.ukdataservice.ac.uk and were restricted to men and women aged 40-69 in order to facilitate comparisons with UK Biobank participants.

Overall, participants in UK Biobank consumed more alcohol on average than people in the general population (Table 3.8.3 and Table 3.8.4). For example, 51.3% of male participants and 36.6% of female participants reported drinking at least three times a week in UK Biobank compared to 44.8% of men and 32.0% of women in the 2008 HSE. Also, 2.8% of male participants and 5.9% of female participants in UK Biobank reported never drinking alcohol compared to 3.8% of men and 6.5% of in the 2008 HSE.

Table 3.8.3 Average Alcohol Intake Frequency in UK Biobank

	Men		Women	
Alcohol intake frequency	**n**	**%**	**n**	**%**
Daily or almost daily	57,914	25.3	43,878	16.1
Three or four times a week	59,552	26.0	55,911	20.5
Once or twice a week	59,137	25.9	70,191	25.7
One to three times a month	20,359	8.9	35,515	13.0
Special occasions only	16,860	7.4	41,170	15.1
Former	8,128	3.6	9,988	3.7
Never	6,409	2.8	15,988	5.9
Prefer not to answer	379	0.2	378	0.1

Table 3.8.4 Average Alcohol Intake Frequency in HSE 2008

	Men		Women	
Alcohol intake frequency	**n**	**%**	**n**	**%**
Almost every day	641	18.5	483	11.9
Five or six days a week	254	7.3	193	4.8
Three or four days a week	653	18.9	622	15.3
Once or twice a week	936	27.1	1,011	24.9
Once or twice a month	341	9.9	487	12.0
Once every couple of months	146	4.2	352	8.7
Once or twice a year	179	5.2	402	9.9
Not at all in last 12 months	24	0.7	50	1.2
Former	152	4.4	198	4.9
Never	132	3.8	262	6.5
Do not know/ No answer	2	0.1	1	0.0

There was a much lower prevalence of cigarette smoking in UK Biobank than in HSE 2008 (Table 3.8.5 and Table 3.8.6). There was a slight difference in definitions with respect to current occasional smokers though this did not account for the large difference in cigarette smoking prevalence.

Table 3.8.5 Cigarette Smoking in UK Biobank

	Men		Women	
Cigarette smoking	**n**	**%**	**n**	**%**
Current regular cigarette smoker	17,944	7.9	18,551	6.8
Former regular cigarette smoker	62,203	27.2	55,043	20.2
Never regular cigarette smoker	147,488	64.5	198,287	72.6
Prefer not to answer/ Do not know	1,075	0.5	1,138	0.4

Table 3.8.6 Cigarette Smoking in HSE 2008

	Men		Women	
Cigarette smoking	**n**	**%**	**n**	**%**
Current cigarette smoker	715	20.6	777	19.1
Former regular cigarette smoker	1,183	34.1	1,049	25.8
Never regular cigarette smoker	1,563	45.0	2,235	55.0
Do not know/ No answer	10	0.3	5	0.1

There was a lower proportion of men and women with BMI ≥30 kg/m^2 in UK Biobank than in the HSE 2008 (Table 3.8.7 and Table 3.8.8).

Table 3.8.7 Body Mass Index in UK Biobank

	Men		Women	
BMI (kg/m^2)	**n**	**%**	**n**	**%**
<18.5	547	0.2	2,079	0.8
18.5 - <25	56,763	25.0	105,695	38.9
25 - <30	112,269	49.3	99,906	36.7
30 - <40	54,900	24.1	57,683	21.2
≥40	3,058	1.3	6,649	2.4

Table 3.8.8 Body Mass Index in HSE 2008

	Men		Women	
BMI (kg/m^2)	**n**	**%**	**n**	**%**
<18.5	10	0.3	28	0.8
18.5 - <25	710	23.3	1,183	33.8
25 - <30	1,384	45.5	1,243	35.5
30 - <40	893	29.3	933	26.6
≥40	47	1.5	118	3.4

In UK Biobank, 38.2% of male participants and 34.4% of female participants reported doing ≥30 minutes of at least moderate activity at least five times per week. This was slightly higher than the proportion of men (35.5%) and women (29.9%) who reported a similar level of activity in the 2008 HSE.

A lower proportion of men and women reported having less than fair health in UK Biobank than in the 2008 HSE (Table 3.8.9 and Table 3.8.10). Fewer men (34.9%) and women (29.4%) in UK Biobank reported having a long-standing illness than men (49.0%) and women (49.7%) in the 2008 HSE.

Table 3.8.9 Overall Health Status in UK Biobank

	Men		Women	
Health status	**n**	**%**	**n**	**%**
Excellent	35,544	15.5	46,346	17.0
Good	127,487	55.7	161,597	59.2
Fair	52,461	22.9	52,935	19.4
Poor	12,010	5.3	10,771	4.0
Prefer not to answer/ Do not know	1,217	0.5	1,355	0.5

Table 3.8.10 Overall Health Status in HSE 2008

	Men		Women	
Health status	**n**	**%**	**n**	**%**
Very good	1,060	30.5	1,295	31.9
Good	1,452	41.8	1,686	41.5
Fair	689	19.9	740	18.2
Bad	201	5.8	253	6.2
Very bad	66	1.9	91	2.2
Do not know	3	0.1	1	0.0

Chapter 4 Methods

This chapter provides more detailed information on the UK Biobank dataset and the statistical methods used to analyse the data. The data available for each of the three risk factors of interest are described, including the derivation of the variables used in the analysis; sections on alcohol intake and smoking also include a brief discussion of methods for measuring and analysing these exposures in epidemiological studies. Other variables relevant to the analyses are also described. Then, the outcome data are described in detail. Lastly, the statistical methods utilised in the analysis are described.

4.1 Alcohol Data

4.1.1 Measuring Alcohol Intake

Accurately measuring alcohol intake is a serious challenge for epidemiological studies. People's patterns of consumption can be very variable and alcoholic drinks come in a wide range of sizes and alcoholic strengths. Studies rely on self-reported data which mean that people may misreport their alcohol intake. In fact, it is well known that estimates of total alcohol consumption from surveys greatly underestimate total alcohol consumption based on national sales data.[34] In the UK, only about 60% of total alcohol consumption is accounted for in surveys.[352]

Questionnaires

There exist a number of different types of questionnaire for acquiring data on average alcohol intake.[353, 354] Semi-quantitative food frequency questionnaires are commonly used in epidemiological studies. People are asked to report their typical intake using a single question which usually includes frequency categories and also quantity categories for more frequent drinkers, for example "never in last 12 months, less than once per month, 1-3 times per month, 1-2 times per week, 3-4 times per week, 5-6 times per week, 1-2 drinks per day, 3-4 drinks per day, 4-6 drinks per day, >6 drinks per day." This kind of question is commonly used because of its simplicity but has the disadvantage that it confounds frequency and quantity i.e. there is no distinction

between someone who drinks one drink once a week and someone who drinks five drinks once a week.

Another common method for assessing alcohol intake in epidemiological studies and surveys is the quantity frequency (QF) method. People are asked to report how often they usually drink alcohol (frequency) and then how much alcohol they usually drink on one of these days (quantity). The QF method is a simple measure that should be able to capture average alcohol intake for both regular and infrequent or occasional drinkers. However, the main issue with the QF method is the inability to distinguish different drinking patterns.

Some more recent questionnaires have attempted to obtain data on patterns of drinking and binge drinking. For example, some questionnaires may ask people, in addition to the quantity frequency questions, how often they drank above a specified amount or the maximum amount they drank on a single day in a specified reference period. The graduated frequencies (GF) method attempts to obtain information on average alcohol intake as well as drinking patterns.[353, 355] People are first asked how often they drink a large number of "standard" drinks (e.g. 12 or more). People are then asked to report how often they drink decreasing quantities of "standard" drinks (e.g. 8-11, 5-7, etc.). The advantage of graduated frequencies is that they are able to measure the variability of intake and in particular heavy drinking patterns. A disadvantage of the GF method is that it is complex and obviously more intensive for people to complete. There are also concerns about the administration of the method and that some people may over-report their intake.[356]

Pattern of Drinking

In recent years, there has been an increased focus on the adverse health effects of binge drinking and on methods to reduce binge drinking. As described above, one major limitation of many alcohol questionnaires is the inability to distinguish different patterns of alcohol intake by simply asking people to report their average intake. For many people alcohol intake is a variable behaviour and is characterised by a regular pattern of drinking interspersed with heavier drinking occasions. When reporting average alcohol intake, people may tend to ignore infrequent or atypical episodes of drinking and simply

report their "usual" intake. This could be an important reason why alcohol intake is underestimated by self-reported data.

One study asked people to report their average alcohol intake using a simple QF method and also included further questions about atypical drinking and drinking during special occasions.[357] Including the responses for atypical/special occasion drinking, the estimate for total alcohol consumption increased from 63% to 79% of total alcohol intake based on national sales data. Hence, it seems likely that questionnaires can better characterise people's average alcohol intake by including questions on pattern of intake. However, there still remained a significant difference between total alcohol intake estimated by the questionnaire and by national sales data.

It must also be mentioned that sales data are not a `gold standard' for measuring the accuracy or reliability of alcohol questionnaires and it is incorrect to assume that one measure of alcohol intake is more accurate simply because people report a higher alcohol intake on average. For example, the 2004 Australian National Drug Strategy Household Survey compared the amount of alcohol reported using the QF method, the GF method and also questions on alcohol intake yesterday.[358] Whilst the QF and GF methods resulted in a similar estimate for total alcohol intake (50 and 52% of total alcohol intake based on sales data), alcohol intake yesterday produced the highest estimate for total alcohol intake (57%). This makes sense since people will have less difficulty accurately recalling their intake from the day before. However, alcohol intake yesterday is clearly an inadequate measure for average alcohol intake. For example, alcohol intake yesterday will overestimate the number of people who are non-drinkers because it will not only include actual non-drinkers but also drinkers who did not happen to drink yesterday.

Estimating Drinks

Another difficulty when assessing alcohol intake is that alcoholic beverages are available in a wide variety of sizes and strengths. This means that a typical alcoholic drink can contain very different amounts of alcohol for two different individuals and also that the typical alcohol intake for an individual can be quite variable. Questionnaires do not generally ask people to provide information on size and strength

of alcoholic beverages. Many questionnaires define a "standard" drink and ask participants to report their alcohol intake in terms of "standard" drinks. However, people have difficulties understanding these definitions and tend to overestimate how much alcohol is in a "standard" drink.[359] The estimates of total alcohol consumption from the 2004 Australian National Drug Strategy Household Survey mentioned above were based on "standard" drinks. People were also asked to report their alcohol intake yesterday using more detailed questions including options for different drink sizes and strengths and this method accounted for 81% of total national alcohol consumption based on sales data.[358]

Diaries

Asking participants to record their alcohol intake using diaries can avoid many of the issues for questionnaires. The use of diaries precludes any problems with recall, they automatically capture information on the pattern of intake and participants should be able to provide more accurate information on the actual drink size and strength. The main issue with diaries is that they are demanding for participants and impractical for long periods of time. Hence, people are normally asked to record their alcohol intake over a relatively short period e.g. one week. However, this week may not represent an individual's typical alcohol intake over a longer period of time.

4.1.2 Alcohol Reference Group

The choice of reference group for analyses of alcohol intake may have important consequences for the results though it is not clear what the ideal choice is. There has been much debate about the choice of reference group for analyses of alcohol intake. This has generally centred around studies of coronary heart disease and overall mortality where studies have found evidence for a J-shaped relationship such that low or moderate alcohol intake is associated with the lowest level of risk.[360, 361]

Non-drinkers

Many early studies of alcohol and mortality that found a lower risk for light/moderate drinkers used non-drinkers or current abstainers (i.e. both never drinkers and former

drinkers combined) as the reference group. Shaper and colleagues consequently argued that the lower risk of mortality observed among this group compared to non-drinkers may be a result of including former drinkers in the reference group.[46] Many former drinkers will quit drinking as a result of illness, often referred to as "sick-quitters", and it is likely that these people will have an elevated risk of mortality. In fact, meta-analyses have shown a lower protective effect of alcohol on coronary heart disease and mortality when restricting to results using never drinkers as the reference group.[360, 361]

Furthermore, former drinkers tend to exhibit a number of characteristics associated with greater levels of disease and poor health.[362-364] For example, in the British Regional Heart Study, former drinkers had the highest prevalence of disease, the highest prevalence of medication use and were most likely to report poor health status.[362] Former drinkers are also likely to be heavy smokers and to have a low socioeconomic status. Also, one prospective study found that people who developed a longstanding illness were much more likely to quit alcohol drinking.[365]

Another issue with the use of non-drinkers as the reference group is that the proportion of former drinkers will vary between studies. This means that the risk of disease in the reference group will vary between studies because it will depend on the relative proportions of former drinkers and never drinkers. Despite these reasons against the use of non-drinkers, studies continue to use non-drinkers as the reference group, presumably because the questionnaire only asked about current drinking and so could not separate current abstainers from lifetime abstainers.

Never drinkers

Thus, non-drinkers should be avoided as the reference group where possible. An alternative is to use never drinkers or lifetime abstainers as the reference group. However, it has been argued that never drinkers also do not provide an ideal reference group.[46, 366, 367] In many countries alcohol consumption is a widespread behaviour whereas lifelong abstainers represent a small proportion of the population. Hence, never drinkers may not represent a reliable reference group for many studies due to the small number of people, particularly if the outcome of interest is fairly rare. It is important

that different studies use the same or similar reference groups in order to facilitate the comparison of results between studies.

There is also a concern about confounding when using never drinkers as the reference group.[366, 368] Never drinkers represent a small selected group and may abstain from drinking for a variety of reasons. Thus, never drinkers may differ from drinkers in terms of a number of other health related characteristics.[363, 364, 368, 369] In general, light or moderate drinking is associated with a number of favourable health related behaviours.

Another potential reason against using never drinkers as the reference group is because of misclassification of lifelong abstainers. The U.S. National Alcohol Survey, conducted in 1984 with follow-ups in 1990 and 1992, found that approximately half of the respondents reporting never drinking in 1992 had reported some level of drinking at an earlier survey.[366] The 1958 British Birth Cohort Study found a similar pattern with 67% of people reporting never drinking at age 45 also having reported some level of drinking at an earlier age (16, 23, 33 or 42).[370] In both studies however these participants mainly reported occasional or very light drinking at earlier timepoints. Also since there will be similar misclassification amongst occasional or very light drinkers,[370] this is not necessarily an argument against the use of never drinkers as the reference group but rather for improved alcohol measures.

In the Health Survey for England, participants were first asked if they ever drink alcohol nowadays. Participants who said no were then asked whether they meant never or very occasionally. Remarkably, almost 40% of participants who reported not drinking alcohol based on the first question said they drink very occasionally.[351, 371] This indicates that, rather than simply poor recall from participants, some of the misclassification is due to poor questionnaire design and that, in order to identify never drinkers, the difference between never and hardly ever should be made more explicit.[372]

Light Drinkers

Another option for the reference group is light or occasional drinkers. Compared with never drinkers, light drinkers may provide a larger, more reliable reference group and also a less distinct group. However, the exposure needs to be well defined when using

light drinkers as the reference group. For example, some light drinkers will be former heavy drinkers. These former heavy drinkers will likely have a higher risk of disease meaning that the overall risk for light drinkers depends on the proportion who used to be heavier drinkers (this is similar to the issue with using non-drinkers as the reference group). Furthermore, as described above, it is difficult to obtain accurate data on patterns of drinking. Hence, some light drinkers could be infrequent binge drinkers who also may be at an elevated risk of disease.

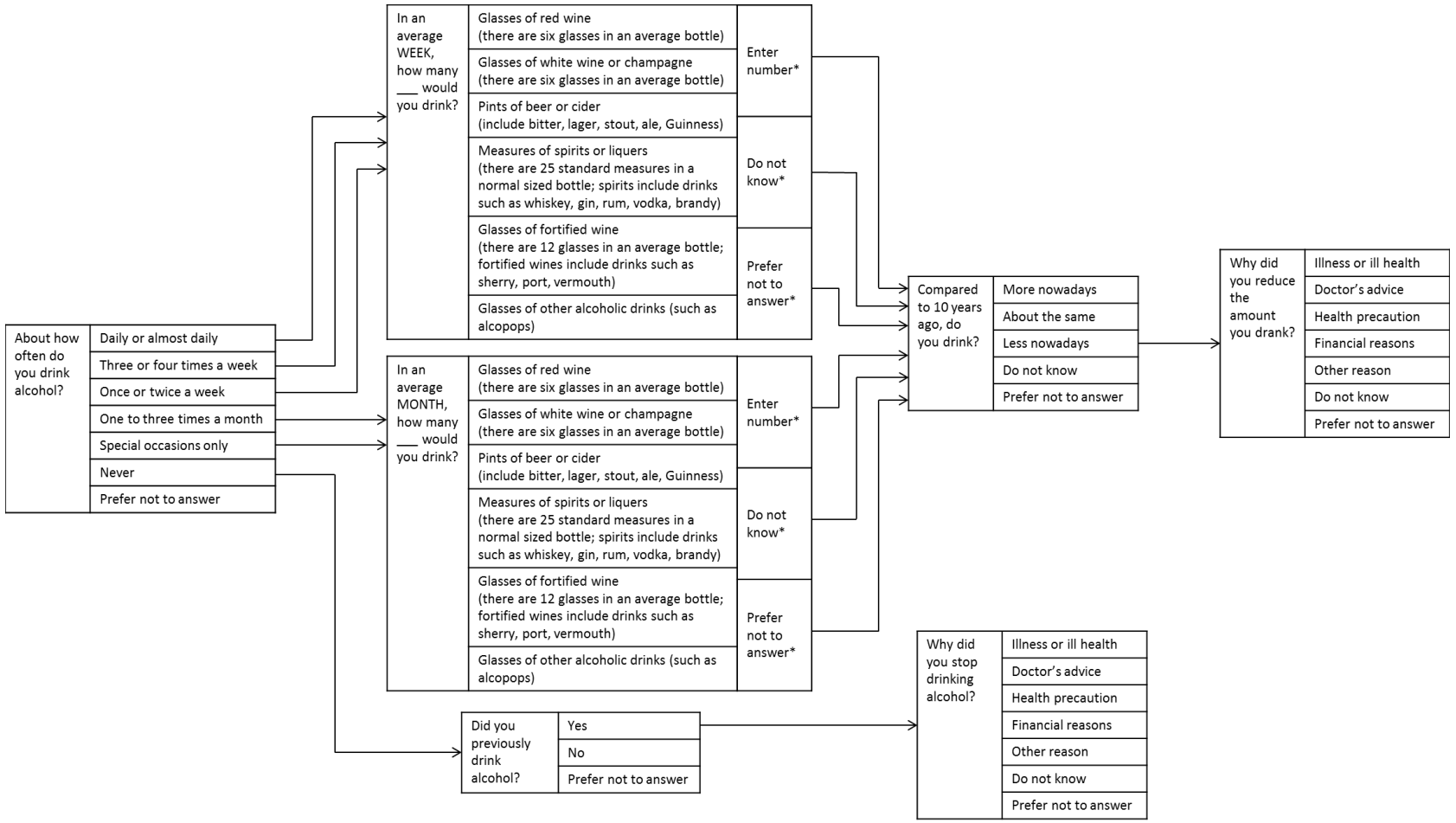
Thus, the risk of disease for light drinkers in each study will depend on the proportion of former heavy drinkers and binge drinkers. However, it is important that the reference group be well defined so that comparisons can be made between different studies. Therefore, light drinkers may not be a suitable choice for the reference group in the absence of detailed information on lifetime exposure and patterns of drinking.

4.1.3 UK Biobank Questionnaire

Questions

Participants answered questions on alcohol intake as part of the touchscreen questionnaire. All questions included in the touchscreen questionnaire can be viewed on the study website (www.ukbiobank.ac.uk). The structure and the wording of the alcohol questions can be seen in Figure 4.1.1. Participants were first asked how often they drink alcohol. Participants who responded “never” were asked if they previously drank alcohol and participants who previously drank alcohol were asked why they stopped drinking.

Figure 4.1.1 UK Biobank Questions on Alcohol Intake



The main questions on alcohol intake answered by UK Biobank participants as part of the touchscreen questionnaire.

*For each question on specific beverages, participants could either enter a number, say "do not know" or "prefer not to answer". Regardless of their answers to these questions all participants were then asked about their drinking ten years ago.

Participants who reported drinking "daily or almost daily", "three or four times a week" or "once or twice a week" were asked to report their average weekly intake of different alcoholic beverages (red wine, white wine/champagne, beer/cider, spirits, fortified wine, other). Participants who reported drinking "one to three times a month" or "special occasions only" were asked to report their average monthly consumption of the same alcoholic beverages. Unfortunately the monthly questions were only added during recruitment which resulted in missing data (further explained below).

Participants who reported drinking were also asked how their current drinking at baseline compared to ten years ago. If participants said they drank less nowadays they were asked why they reduced their intake. Finally, participants were also asked "When you drink alcohol is it usually with meals?"

Missing Data

Altogether, 113,903 participants reported drinking "one to three times a month" or "special occasions only". According to the structure of the questionnaire these participants should have then reported their average monthly intake of different alcoholic beverages. However, 73,061 (64%) of these participants had missing data for each of these monthly questions. These data were missing because the monthly questions were only introduced during recruitment and were not included in the original touchscreen questionnaire.

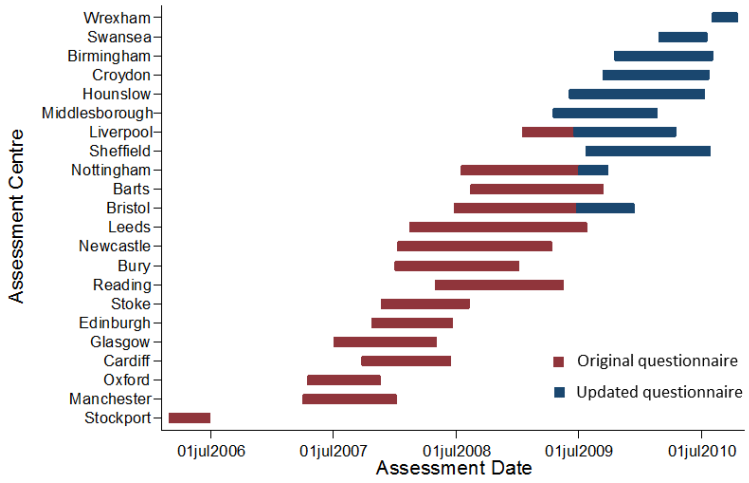
In the original questionnaire all participants who reported drinking (from "daily or almost daily" to "special occasions only") were asked about their average weekly consumption of specific alcoholic beverages. Hence, participants who completed the original questionnaire and reported drinking "one to three times a month" or "special occasions only" answered the weekly beverage questions instead of the monthly beverage questions.

Participants who reported drinking "one to three times a month" or "special occasions only" would have found it difficult to accurately report their average weekly intake and it is likely that many of these participants simply answered zero for each question. This

is presumably why the monthly questions were added (although they were a rather late addition) and also why the weekly data were not made available for these participants.

The updated questionnaire was introduced around May-July 2009 (Figure 4.1.2). Only three assessment centres changed from the original to the updated questionnaire during recruitment. All other assessment centres either used the original questionnaire or the updated questionnaire throughout recruitment. The questions on other alcoholic drinks (both per week and per month) were also only added as part of the updated questionnaire so participants who answered the original questionnaire had missing data for other alcoholic drinks.

Figure 4.1.2 Introduction of the Updated Questionnaire by Assessment Centre in UK Biobank



Recruitment period for each assessment centre. Red indicates that the original questionnaire was used. Participants who answered the original questionnaire and reported drinking “one to three times a month” or “special occasions only” had missing data for the monthly beverage questions since these questions were added during recruitment. Blue indicates the use of the updated questionnaire.

Table 4.1.1 Frequency of Alcohol Intake in UK Biobank by Age and Sex

	Men								Women							
	<50 yrs		50-59 yrs		≥60 yrs		Total		<50 yrs		50-59 yrs		≥60 yrs		Total	
Alcohol Intake Frequency	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
Daily or almost daily	9,538	17.9	18,311	25.0	30,064	29.2	57,913	25.3	8,193	12.7	15,125	16.1	20,560	17.9	43,878	16.1
Three or four times a week	13,442	25.3	19,999	27.3	26,111	25.4	59,552	26.0	13,849	21.4	20,389	21.7	21,672	18.9	55,910	20.4
Once or twice a week	15,767	29.6	18,926	25.9	24,441	23.8	59,134	25.8	18,211	28.2	24,609	26.2	27,368	23.9	70,188	25.7
One to three times a month	6,169	11.6	6,300	8.6	7,890	7.7	20,359	8.9	9,808	15.2	12,127	12.9	13,580	11.8	35,515	13.0
Special occasions only	4,197	7.9	5,002	6.8	7,660	7.5	16,859	7.4	9,021	13.9	13,363	14.2	18,786	16.4	41,170	15.1
Never	2,056	3.9	1,705	2.3	2,646	2.6	6,407	2.8	3,463	5.4	4,710	5.0	7,815	6.8	15,988	5.9
Former	1,791	3.4	2,659	3.6	3,677	3.6	8,127	3.6	1,918	3.0	3,446	3.7	4,623	4.0	9,987	3.7
Missing/Prefer not to answer*	255	0.5	250	0.3	319	0.3	824	0.4	230	0.4	280	0.3	321	0.3	831	0.3
Total	53,215	100.0	73,152	100.0	102,808	100.0	229,175	100.0	64,693	100.0	94,049	100.0	114,725	100.0	273,467	100.0

*898 people had missing data for alcohol questions, 605 said prefer not to answer to first question, 152 people said prefer not to answer for question on past drinking.

Responses

Table 4.1.1 shows how often participants reported drinking alcohol. Participants who reported never drinking alcohol were split into never drinkers and former drinkers according to whether they reported drinking in the past. 51.3% of men reported drinking at least three or four times a week and 77.1% of men reported drinking at least once or twice a week. The corresponding figures for women were 36.5% and 62.2%. Older men and women were more likely to drink daily or almost daily. The proportion of former drinkers was similar for men (3.6%) and women (3.7%) but the proportion of never drinkers among women (5.9%) was more than double that among men (2.8%).

4.1.4 Alcohol Analysis Variable

To obtain a measure of overall consumption, grams of alcohol per day were calculated for each participant based on their responses to the alcoholic beverage questions. The grams of alcohol contained in an alcoholic drink depends on the size of the drink and the strength of the alcohol. The sizes and strengths used to calculate the grams of alcohol for each alcoholic beverage are shown in Table 4.1.2. The size of each alcoholic beverage was calculated using the information in the questionnaire (Figure 4.1.1). For example the red and white wine questions stated that an average bottle contains six glasses of wine. Therefore, assuming an average bottle of wine contains 750ml, the drink size for red and white wine was defined as 125ml.

Table 4.1.2 Defintion of Grams of Alcohol for Different Alcoholic Beverages used in Analyses

Alcoholic beverage	Beverage size (ml)	g/100ml	g
Red wine	125	9.6	12.0
White wine	125	9.0	11.3
Beer	568.3	3.5	19.9
Spirits	30	30.7	9.2
Fortified wine	62.5	15.4	9.6

The alcohol strengths (g/100ml) for each alcoholic beverage were taken from the Oxford WebQ,[373] an online dietary questionnaire used by UK Biobank to assess participants' dietary intake. The alcohol strengths used by Oxford WebQ were based on

values from McCance and Widdowson's The Compositions of Foods.[58, 374] These g/100ml values were multiplied by the drink sizes (before dividing by 100) to obtain the number of grams of alcohol for each alcoholic beverage.

For each participant, the grams of alcohol for each alcoholic beverage were multiplied by the reported number of drinks of that beverage consumed in an average week or month. These values were summed to give the total grams of alcohol per week or per month before converting to grams of alcohol per day.

The questions on other alcoholic drinks were ignored when calculating grams of alcohol per day for two reasons. First, the questions were rather imprecise in comparison with the other questions on alcoholic beverages; other drinks could include a wide range of alcoholic beverages and the indication of the drink size was vague. Secondly, the large majority of participants reported zero intake for other alcoholic drinks; 99% of participants who answered the weekly question on other alcoholic drinks reported zero intake and 95% of participants who answered the monthly question reported zero intake.

Alcohol intake was not defined for a number of participants. 898 participants were excluded because they had missing data for the first question on alcohol intake frequency and all following alcohol questions. 757 participants were excluded because they said "prefer not to answer" for the alcohol intake frequency question or for past alcohol intake question. 7,445 participants who said "do not know" or "prefer not to answer" for at least one of the alcoholic beverage questions were also excluded.

2,350 participants were excluded for reporting a number of alcoholic drinks inconsistent with their reported alcohol intake frequency. This included participants who reported drinking "daily or almost daily" and less than four drinks per week, "three or four times a week" and less than two drinks per week, "once or twice a week" and less than one drink per week, and "one to three times a month" and less than one drink per month. Participants who reported drinking "special occasions only" were not excluded for reporting a low number of drinks.

Table 4.1.3 Mean Grams of Alcohol per Day Consumed by Men and Women in UK Biobank, by Alcohol Frequency

	Men						Women					
	Red wine	**White wine**	**Beer**	**Spirits**	**Fortified wine**	**Total**	**Red wine**	**White wine**	**Beer**	**Spirits**	**Fortified wine**	**Total**
Daily or almost daily	13.1	5.5	21.2	5.1	0.4	45.3	10.3	9.7	2.3	3.1	0.6	26.0
Three or four times a week	7.1	3.0	15.0	2.4	0.2	27.6	6.3	5.6	1.8	1.8	0.4	15.8
Once or twice a week	2.8	1.3	9.2	1.3	0.1	14.9	2.7	2.6	1.4	1.3	0.2	8.3
One to three times a month	0.6	0.4	2.1	0.3	0.1	3.6	0.6	0.7	0.5	0.4	0.1	2.1
Special occasions only	0.3	0.2	0.7	0.1	0.0	1.3	0.2	0.2	0.1	0.1	0.0	0.6

Alcohol grams per day was not calculated for 73,061 participants who reported drinking "one to three times a month" or "special occasions only" due to the missing data described above. (These missing data were imputed using multiple imputation as a form of sensitivity analysis for the analysis of alcohol intake and colorectal cancer (described in detail in section 4.6.5)).

Table 4.1.3 shows the mean grams of alcohol per day reported by male and female drinkers for the different alcoholic beverages and overall. Results are presented by alcohol frequency. Within each frequency category, men reported drinking substantially more alcohol than women. For men the majority of alcohol was consumed in the form of beer and red wine and for women the majority of alcohol was consumed in the form of red and white wine.

Categorical Analysis Variable

The following categories of alcohol intake were defined: never drinkers, former drinkers, <5 g/d, 5-<15 g/d, 15-<30 g/d 30-<45 g/d and ≥45 g/d. These categories were based on those of a large pooled analysis of alcohol intake and colorectal cancer.[45] Participants who reported not drinking alcohol at baseline and not previously drinking alcohol were defined as never drinkers. Participants who reported not drinking alcohol at baseline and reported previously drinking alcohol were defined as former drinkers. Table 4.1.4 shows the number of participants in each category of the alcohol intake variable. 418,131 participants (217,928 men and 200,203 women) were included in this variable.

Table 4.1.4 Alcohol Intake Variable

	Men		**Women**		**Overall**	
Alcohol intake	**n**	**%**	**n**	**%**	**n**	**%**
Never	6,407	3.2	15,988	7.3	22,395	5.4
Former	8,127	4.1	9,987	4.6	18,114	4.3
<5 g/d	17,154	8.6	46,582	21.4	63,736	15.2
5-<15 g/d	48,041	24.0	82,547	37.9	130,588	31.2
15-<30 g/d	57,929	28.9	46,348	21.3	104,277	24.9
30-<45 g/d	31,748	15.9	11,758	5.4	43,506	10.4
≥45 g/d	30,797	15.4	4,718	2.2	35,515	8.5
Total	200,203	100.0	217,928	100.0	418,131	100.0

Reference Group

Given the potential issues with using non-drinkers or light drinkers as the reference group described in section 4.1.2, never drinkers were used as the reference group in analyses of the categorical alcohol intake variable. The disadvantage of never drinkers is that they represent a small group which can result in wide CIs, particularly for subgroup analyses. However, in the absence of detailed information on past drinking or patterns of drinking, it was decided that never drinkers represented the most consistent and well-defined reference group.

As a form of sensitivity analysis (and to provide a comparison), results were also calculated using a slightly different reference group which combined never drinkers and consistent light drinkers. Consistent light drinkers were defined as participants who reported drinking <5 g/d and reported drinking "about the same" or "more nowadays" as ten years before baseline. 27,252 participants met the definition of consistent light drinkers.

Continuous Analysis Variable

A categorical variable offers a very simple way to compare the risk of disease associated with different levels of exposure and therefore the interpretation of the results is straightforward. However, results will depend on the arbitrary choice of cut-points and results would suggest that the risk suddenly increases as a cut-point is crossed. In contrast, a continuous variable avoids the arbitrary cut-points. It also avoids the issues with the choice of the reference group. However, a number of different models can be fit to a continuous variable and it can be difficult to decide the most appropriate model.

For the continuous alcohol grams per day variable, never drinkers were defined as zero g/d while former drinkers were excluded. Participants who reported drinking "special occasions only" and reported zero intake for each alcoholic beverage were also excluded. One g/d was added to the continuous variable to accommodate different analysis models (e.g. logarithmic and negative power transformations). Thus, the continuous variable was defined for 391,645 participants (190,006 men and 201,639 women).

Alternative Variable

As described above, a number of participants who reported drinking "one to three times a month" or "special occasions only" had missing data and so could not be included in the analysis variables. In the analyses of adiposity and smoking, an alternative alcohol variable was used in order to avoid excluding these participants.

This alternative alcohol variable was similar to the original categorical variable. Never and former drinkers were defined as before. Also, participants who reported drinking "once or twice a week", "three to four times a week" or "daily or almost daily" were defined by their average grams of alcohol per day as before. However, in contrast to the original variable, participants who reported drinking "one to three times a month" or "special occasions only" were not categorised by their average grams of alcohol per day but were included as distinct categories (Table 4.1.5).

Table 4.1.5 Alternative Alcohol Intake Variable

	Men		**Women**		**Overall**	
Alcohol intake	**n**	**%**	**N**	**%**	**n**	**%**
Never	6,407	2.9	15,988	6.0	22,395	4.5
Former	8,127	3.6	9,987	3.7	18,114	3.7
Special occasions only	16,859	7.5	41,170	15.3	58,029	11.8
One to three times a month	20,359	9.1	35,515	13.2	55,874	11.3
<5 g/d	5,895	2.6	20,937	7.8	26,832	5.4
5-<15 g/d	46,392	20.7	81,904	30.5	128,296	26.0
15-<30 g/d	57,874	25.8	46,343	17.3	104,217	21.2
30-<45 g/d	31,746	14.1	11,757	4.4	43,503	8.8
≥45 g/d	30,794	13.7	4,718	1.8	35,503	7.2
Total	224,453	100.0	268,319	100.0	492,772	100.0

This variable was not ideal since it combined measures of average alcohol intake and average frequency of intake. However, only participants with a low frequency of alcohol intake were classified by their frequency. These participants were expected to have a very similar low level of average alcohol intake. For example, among participants with observed data, 95% of participants who reported drinking "one to three times a month" reported drinking less than 6.5 g/d and 95% of participants who reported drinking "special occasions only" reported drinking less than 3.1 g/d. Therefore,

including these categories should not have resulted in significant misclassification. In contrast, there was much more variation in average alcohol intake among participants who reported drinking alcohol more frequently and so it was important that these participants were still classified by their average alcohol intake rather than their frequency of intake. However, it should not have made an important difference whether participants who reported drinking "one to three times a month" or "special occasions only" were classified by their average alcohol intake or their average frequency of intake. In fact, among participants with complete data for alcohol intake, results for BMI and overall smoking status were practically identical whether adjusting for the original categorical variable or the alternative variable (see Table A-1 and Table A-2 in the appendix).

An alternative approach to avoid excluding those participants with missing data from the analysis would be to impute the data using multiple imputation. Multiple imputation was used to impute these missing data in a sensitivity analysis for the relationship between alcohol intake and colorectal cancer (see section 4.6.5). However, it was decided not to use this method when considering alcohol intake as a confounder in the adiposity and smoking analyses. First of all, as mentioned above, the alternative variable was shown to give practically identical results as the original categorical variable among participants with complete data, suggesting that the use of this alternative variable throughout the adiposity and smoking analyses should produce valid results.

Furthermore, multiple imputation has a number of limitations and is highly dependent on the correct specification of the imputation model.[375-377] This can be a complex problem when using multiple imputation for multiple analyses with many variables and possible interactions between variables and a poorly specified imputation model may lead to biased results.

Multiple imputation is also dependent on the assumption that the missing data are missing at random i.e. that, conditional on the observed data available in the dataset, all participants have an equal probability of having missing data.[375, 376] However, if the probability that a participant has missing data depends on the unobserved values, then multiple imputation can lead to biased results. When using multiple imputation for a

number of analyses, a large number of analysis variables must be included in the imputation model. However, each of these variables will likely include some missing data and this reduces the amount of missing data that can be imputed (since fewer participants will have complete data for all variables in the imputation model). To solve this issue, data can be imputed for multiple analysis variables simultaneously using chained equations.[375] However, if the data for any of these variables are not missing at random, multiple imputation can lead to biased results.

4.2 Adiposity Data

This section provides information on the measures of adiposity used in this thesis to investigate the association between adiposity and colorectal cancer. All measures of adiposity were completed by a trained member of staff at the baseline assessment centre. The following measures of adiposity were considered in this thesis: body mass index, waist circumference, waist to hip ratio, waist to height ratio and % body fat.

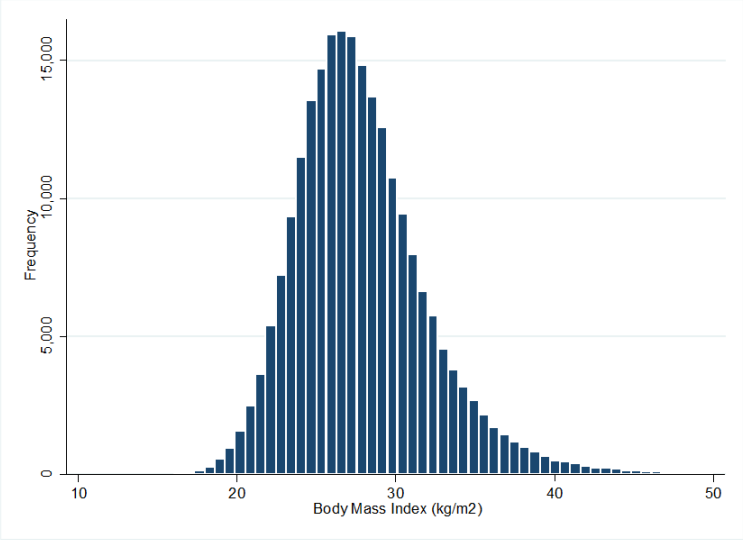
Body Mass Index

BMI is the most commonly used measure of adiposity in the general population. It is defined as the weight in kilograms divided by the square of the height in metres (kg/m^2) and thus provides a very simple measure of weight adjusted for height.

In UK Biobank height and weight were both measured at baseline by the assessment centre staff. Participants were required to stand barefoot for measurements of height and weight. Height was measured to the nearest centimetre and weight was measured to the nearest 100g.

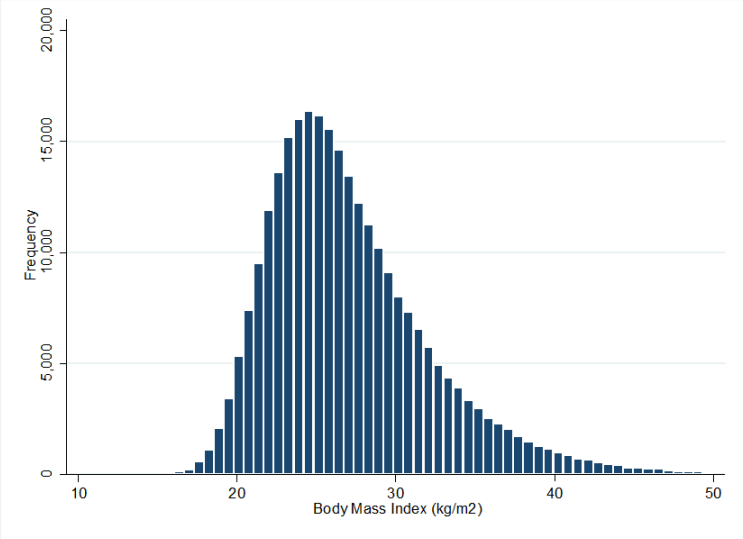
227,529 men and 272,008 women provided data on BMI. 30 men and 96 women with BMI <16 kg/m^2 were excluded because these values were considered as outliers (no participants were excluded for a high value of BMI). BMI was defined for 227,499 men and 271,912 women. Figure 4.2.1 and Figure 4.2.2 show the distribution of BMI for men and women.

Figure 4.2.1 Body Mass Index for Men



Distribution of body mass index for men.

Figure 4.2.2 Body Mass Index for Women



Distribution of body mass index for women.

Participants were classified into approximate quintiles of BMI, separately for men and women (Table 4.2.1). Participants were also classified into three categories of BMI following the WHO definitions;[159] <25.0, 25.0-<30.0 and ≥30.0 kg/m^2 (Table 4.2.2). (Men and women with BMI <18.5 kg/m^2 were not defined separately as there were only 517 men and 1,983 women with BMI <18.5 kg/m^2).

Table 4.2.1 Body Mass Index Quintiles for Men and Women

Men		Women	
BMI (kg/m^2)	**Participants**	**BMI (kg/m^2)**	**Participants**
<24.45	45,428	<22.89	54,479
24.45-<26.40	45,519	22.89-<25.04	54,314
26.40-<28.28	43,633	25.04-<27.34	54,034
28.28-<30.84	47,440	27.34-<30.80	54,646
≥30.84	45,479	≥30.80	54,439
Total	227,499	Total	271,912

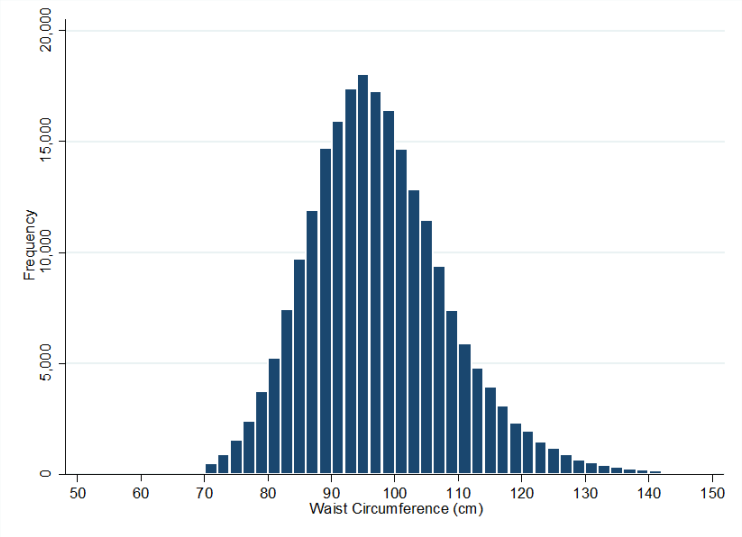
Table 4.2.2 Body Mass Index Categories

BMI (kg/m^2)	**Men**	**Women**
<25.0	57,277	107,676
25.0-<30.0	112,265	99,905
≥30.0	57,957	64,331
Total	227,499	271,912

Waist Circumference

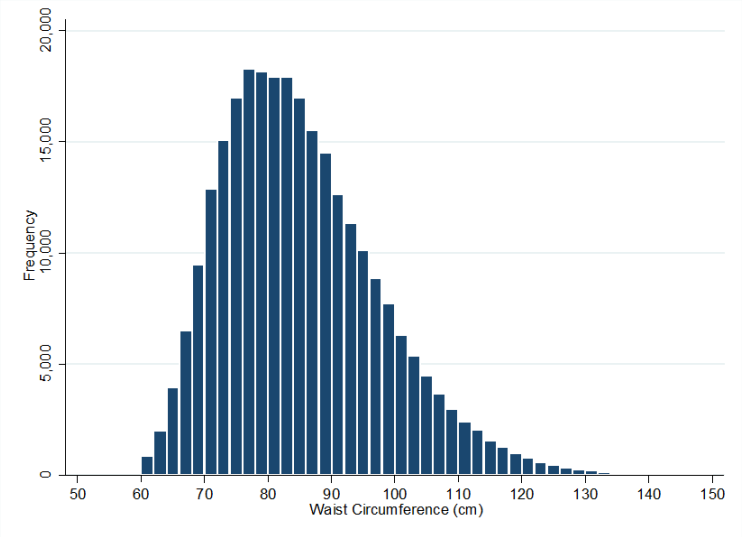
WC was measured to the nearest centimetre by the assessment centre staff using a tape measure. 228,099 men and 272,383 women provided data on WC. 408 men with a WC less than 70 cm and 396 women with a WC less than 60 cm were excluded because these values were considered as outliers (no participants were excluded for a high value of WC), leaving 227,691 men and 271,987 women. Figure 4.2.3 and Figure 4.2.4 show the distribution of WC for men and women.

Figure 4.2.3 Waist Circumference for Men



Distribution of waist circumference for men.

Figure 4.2.4 Waist Circumference for Women



Distribution of waist circumference for women.

Approximate sex-specific quintiles were created for the analysis of WC (Table 4.2.3).

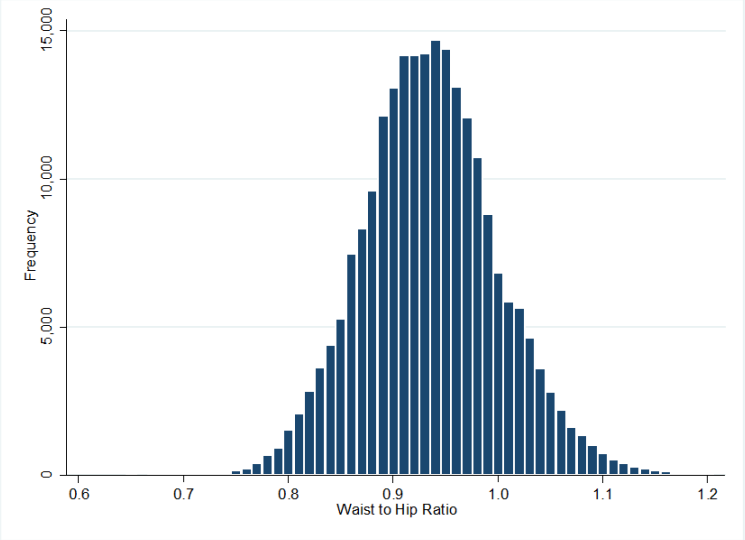
Table 4.2.3 Waist Circumference Quintiles for Men and Women

Men		Women	
WC (cm)	**Participants**	**WC (cm)**	**Participants**
≤87	43,468	≤74	59,183
88-93	48,046	75-80	54,106
94-98	44,165	81-86	51,753
99-105	46,532	87-95	56,178
≥106	45,480	≥96	50,767
Total	227,691	Total	271,987

Waist to Hip Ratio

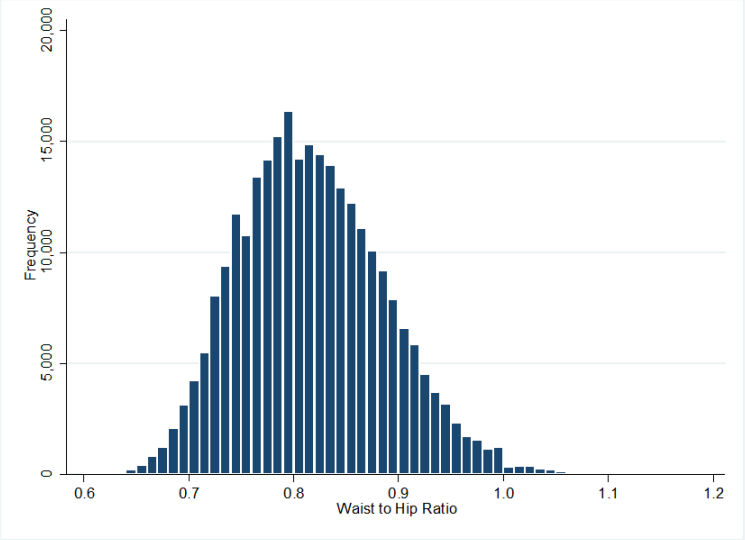
Hip circumference was also measured to the nearest centimetre by assessment centre staff using a tape measure. WHR was calculated as the ratio between WC and hip circumference. 228,037 men and 272,340 women provided data on both WC and hip circumference. 427 men with WC less than 70 cm or hip circumference less than 80 cm and 405 women with WC less than 60 cm or hip circumference less than 70 cm were excluded (because these values were considered as outliers), leaving 227,610 men and 271,935 women. Figure 4.2.5 and Figure 4.2.6 show the distribution of WHR for men and women.

Figure 4.2.5 Waist to Hip Ratio for Men



Distribution of waist to hip ratio for men.

Figure 4.2.6 Waist to Hip Ratio for Women



Distribution of waist to hip ratio for women.

Approximate sex-specific quintiles were created for the analysis of WHR (Table 4.2.4).

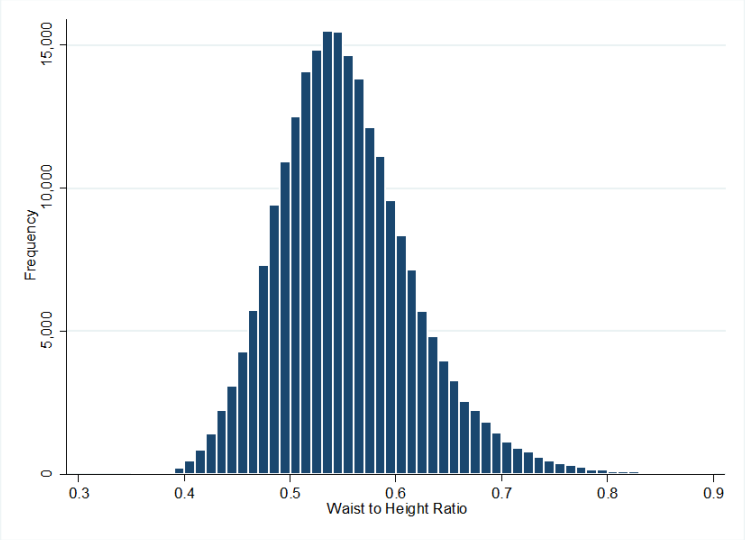
Table 4.2.4 Waist to Hip Ratio Quintiles for Men and Women

Men		**Women**	
WHR	**Participants**	**WHR**	**Participants**
<0.883	45,618	<0.758	55,428
0.883-<0.918	44,701	0.758-<0.796	54,195
0.918-<0.950	46,286	0.796-<0.832	54,229
0.950-<0.990	45,822	0.832-<0.876	52,680
≥0.990	45,183	≥0.876	55,403
Total	227,610	Total	271,935

Waist to Height Ratio

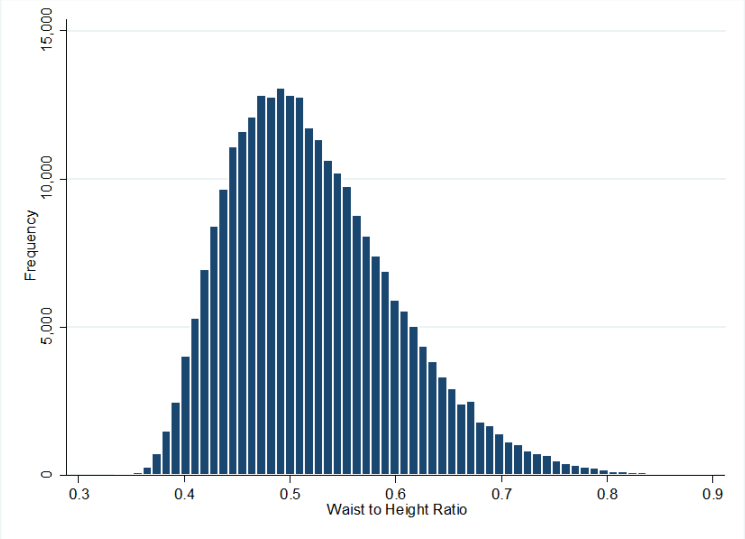
Waist to height ratio (WHtR) was calculated as the ratio between waist circumference and height (see section 4.4.1 for information on height). 227,457 men and 271,848 women had data on both WC and height. 386 men with WC less than 70 cm and 363 women with WC less than 60 cm were excluded, leaving 227,071 men and 271,485 women. Figure 4.2.7and Figure 4.2.8 show the distribution of WHtR for men and women.

Figure 4.2.7 Waist to Height Ratio for Men



Distribution of waist to height ratio for men.

Figure 4.2.8 Waist to Height Ratio for Women



Distribution of waist to height ratio for women.

Approximate sex-specific quintiles were created for the analysis of WHtR (Table 4.2.5).

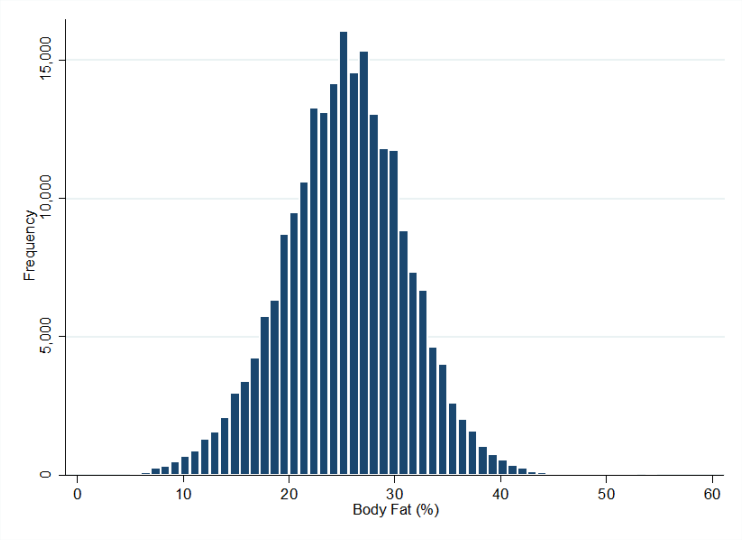
Table 4.2.5 Waist to Height Ratio Quintiles for Men and Women

Men		Women	
WHtR	**Participants**	**WHtR**	**Participants**
<0.500	44,448	<0.453	54,039
0.500-<0.532	46,366	0.453-<0.492	54,582
0.532-<0.562	45,120	0.492-<0.532	53,654
0.562-<0.602	45,481	0.532-<0.586	55,040
≥0.602	45,656	≥0.586	54,170
Total	227,071	Total	271,485

% Body Fat

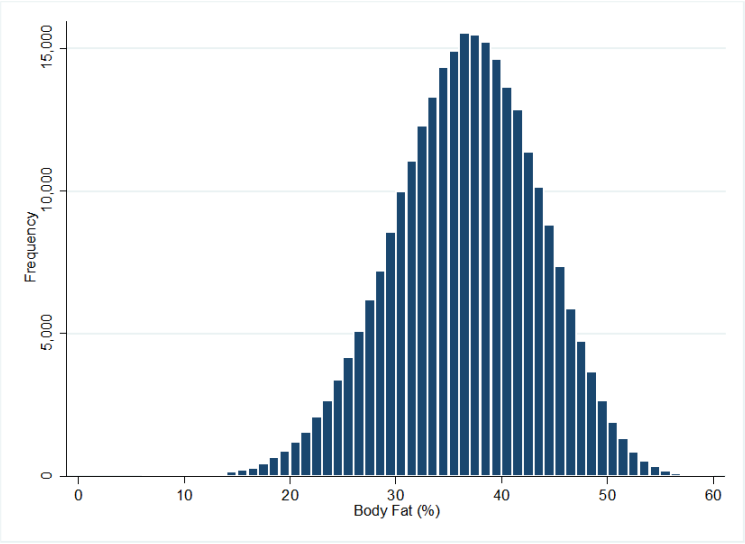
% body fat of participants was measured using bioelectrical impedance. Bioelectrical impedance involves sending a low electrical current through the body and then measuring the resistance to this current. Since the current flows more easily through muscle which contains a high proportion of water and less easily through fat tissue, it is possible to estimate body fat. UK Biobank used a Tanita BC418MA body composition analyser to measure % body fat. Data on % body fat were available for 223,889 men and 268,343 women. Figure 4.2.9 and Figure 4.2.10 show the distribution of % body fat for men and women.

Figure 4.2.9 % Body Fat for Men



Distribution of % body fat for men.

Figure 4.2.10 % Body Fat for Women



Distribution of % body fat for women.

Approximate sex-specific quintiles were created for the analysis of % body fat (Table 4.2.6).

Table 4.2.6 % Body Fat Quintiles for Men and Women

Men		Women	
% body fat	**Participants**	**% body fat**	**Participants**
≤20.6	45,490	≤30.8	53,980
20.7-24.0	44,978	30.9-35.0	53,546
24.1-26.8	44,845	35.1-38.5	53,586
26.9-30.0	43,903	38.6-42.5	54,017
≥30.1	44,673	≥42.6	53,214
Total	223,889	Total	268,343

4.3 Smoking Data

4.3.1 Measuring Smoking Exposure

Information on tobacco use is generally obtained in epidemiological studies through the use of self-reported questionnaires. These questionnaires are broadly similar across studies. However, these questionnaires have clear limitations and there are a number of issues that should be kept in mind when evaluating these studies.

The simplest measure of tobacco exposure, included in practically all studies, is smoking status which classifies people as never, former or current smokers. This is usually accomplished with the use of one or two questions. For example, a questionnaire may ask people whether they currently smoke tobacco and then ask people who do not currently smoke tobacco whether they smoked tobacco in the past.

When measuring any behaviour or exposure that can be considered socially undesirable, there are concerns about the accuracy of self-reported data. A systematic review of studies comparing measures of self-reported smoking status and cotinine (a biomarker of exposure to tobacco smoke) found that the self-reported data underestimated the prevalence of smoking in most studies.[378] However, different studies use a wide range of cut-off points to define smokers and it should be recognised that biochemical measures of smoking can be affected by many factors including metabolism, inhalation and brand of cigarettes.[378]

Defining participants as never, former or current smokers is clearly a rather simple approach and, as well as the concerns about the misclassification of participants using self-reported data, the smoking status variable has a number of other limitations. For example, current smokers will likely include both regular smokers and occasional smokers although these two groups will clearly have a very different level of exposure. Also, smoking status often does not differentiate between different forms of tobacco use and so will combine cigarette smokers and other smokers e.g. cigar or pipe smokers who may differ in terms of pattern of use as well as tobacco type.

It can also be difficult to clearly differentiate between never smokers and former smokers when defining smoking status. For example, many people may have tried smoking once or twice but never smoked regularly. Using the strictest sense of the terms, these people are former smokers. However, given their low exposure, it seems inappropriate to combine these people with former regular smokers. Also, many of these people likely identify themselves as never smokers. Thus, a common approach in questionnaires is to define an upper threshold for never smokers which is often defined as smoking 100 times in their lifetime.[379]

Questionnaires also attempt to measure other aspects of exposure to tobacco smoke such as smoking intensity and smoking duration. These questions generally focus on cigarette use. To assess smoking intensity, current smokers are normally asked to report how many cigarettes they smoke per day on average and former smokers are asked to report how many cigarettes they used to smoke per day on average. However, self-reports of cigarettes per day may not be very accurate. For example, there is a large amount of digit preference for multiples of ten when reporting the number of cigarettes per day. While some degree of digit preference may be expected since cigarettes are commonly sold in packets of ten and 20 in many countries, this does not seem to provide a sufficient explanation. For example, one study found that while the distribution of cigarettes per day reported by smokers showed large spikes for multiples of ten (and particularly 20), the distribution of a biochemical measure of smoking showed no such spikes, suggesting that the distribution of self-reported cigarettes per day does not reflect actual cigarette consumption.[380]

Another study of smokers who were part of a cessation programme similarly showed large digit preference for multiples of ten when asked to report the average number of cigarettes they smoked per day.[381] These people were also asked to concurrently record (using an electronic device) each cigarette they smoked throughout the day for a period of two weeks before the initiation of the programme (subjects were instructed not to change their smoking habits). The distribution of cigarettes per day based on the concurrent daily reports showed a more even spread, again suggesting that more simple self-report measures of cigarettes per day are limited in providing accurate data.

To assess smoking duration, questionnaires ask people to report their age when they started smoking. In general, there is a fairly narrow range for age at initiation since the vast majority of smokers begin smoking before adulthood. In a U.S. survey, 88.2/99.0% of adults who had ever been daily cigarette smokers reported trying their first cigarette by the time they were 18/26 years old and 65.1/96.2% percent reported beginning daily smoking by 18/26 years of age.[382] However, age at smoking initiation may not be easily reported. Generally, people are asked to report their age at initiation decades later meaning that there may be error due to difficulties with recall. Furthermore, there may be a number of stages from first cigarette to regular use and different people will follow different trajectories.[382, 383] Former smokers are also asked to report when they quit

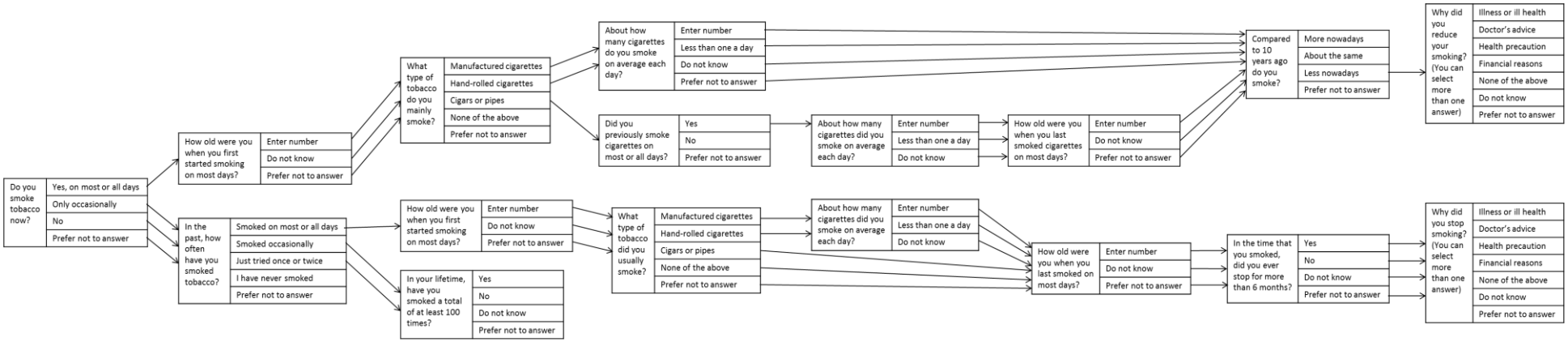
smoking. This may include some error, possibly depending on how long ago people quit smoking and the manner in which they converted from regular smoking to no smoking.

4.3.2 UK Biobank Questionnaire

Questions

Participants answered questions on tobacco use as part of the touchscreen questionnaire. The section of the questionnaire on tobacco use was very detailed with a wide range of questions. The structure and wording for most of the questions on tobacco use can be seen in Figure 4.3.1.

Figure 4.3.1 UK Biobank Questions on Tobacco Use



The main questions on tobacco use answered by UK Biobank participants as part of the touchscreen questionnaire.

Participants were first asked if they currently smoke tobacco. Participants who reported smoking tobacco on most or all days were asked what type of tobacco they mainly smoke. If they reported smoking cigarettes (manufactured or hand-rolled), they were then asked how many cigarettes they smoke on an average day. If participants reported smoking cigars or pipes, they were asked if they previously smoked cigarettes on most or all days and, if they had previously smoked cigarettes on most or all days, they were asked how many cigarettes they smoked on an average day.

If participants reported that they do not smoke tobacco or smoke tobacco only occasionally, they were then asked about their past smoking. Participants who reported smoking on most or all days in the past were asked what type of tobacco they usually smoked and, if they reported smoking cigarettes (manufactured or hand-rolled), they were asked how many cigarettes they used to smoke on an average day. Participants who reported smoking occasionally in the past or having just tried once or twice were asked if they had ever smoked at least 100 times in their lifetime.

Responses

Table 4.3.1 provides a detailed description of participants' responses to the main questions on tobacco use. Participants who responded "smoked occasionally" or "just tried once or twice" to the question on past tobacco use were defined as former occasional smokers. Not all subcategories sum to the higher categories due to other responses provided by participants, for example responding "prefer not to answer".

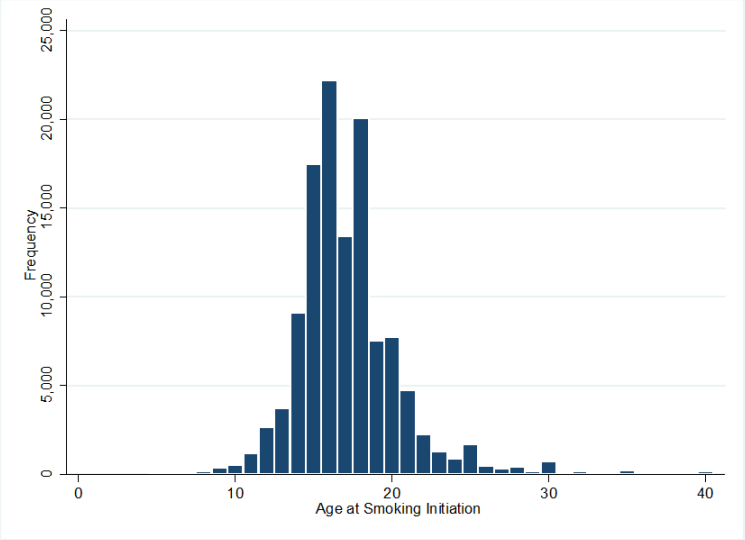
Table 4.3.1 Tobacco Use in UK Biobank by Sex

	Men		Women		Overall	
Smoking status	**n**	**%**	**n**	**%**	**n**	**%**
Daily smokers	20,565	9.0	18,689	6.8	39,254	7.8
Cigarette smoker	17,943		18,551		36,494	
Cigar/pipe smoker	2,563		105		2,668	
Occasional smoker	8,052	3.5	5,683	2.1	13,735	2.7
Former daily smoker	2,900		2,134		5,034	
Cigarette smoker	2,576		2,127		4,703	
Cigar/pipe smoker	317		5		322	
Former occasional smoker	5,114		3,520		8,634	
Smoked ≥100 times	4,580		3,118		7,698	
Smoked <100 times	381		250		631	
Non-smoker	199,909	87.2	248,423	90.8	448,332	89.2
Former daily smoker	62,892		53,025		115,917	
Cigarette smoker	57,917		52,838		110,755	
Cigar/pipe smoker	4,886		114		5,000	
Former occasional smoker	57,161		72,674		129,835	
Smoked ≥100 times	26,115		27,071		53,186	
Smoked <100 times	29,223		41,883		71,106	
Never smoker	79,050		121,851		200,901	
Missing/Prefer not to answer*	649	0.3	672	0.2	1,321	0.3
Total	229,175	100.0	273,467	100.0	502,642	100.0

*892 participants had missing data for questions on tobacco use. 429 participants reported prefer not to answer to question on current tobacco use.

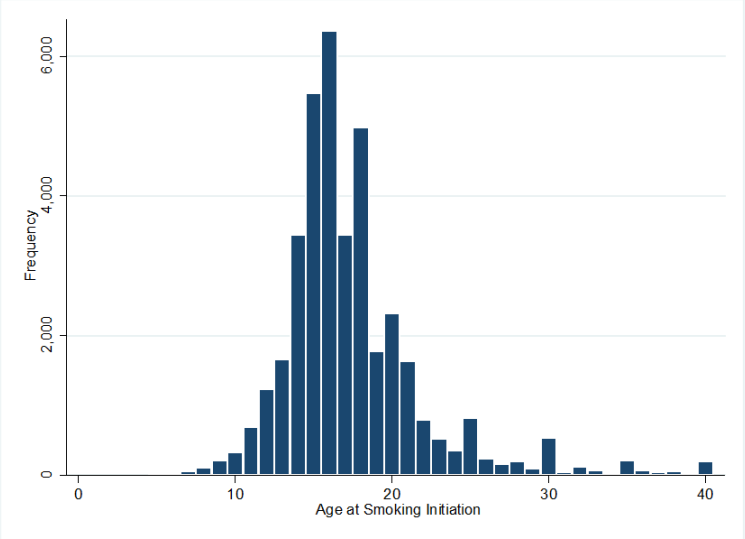
Figure 4.3.2 and Figure 4.3.3 show the age at smoking initiation reported by former and current smokers. 60.5% of former smokers and 51.6% of current smokers reported that they started smoking on most days between 15 and 18 years old.

Figure 4.3.2 Age at Smoking initiation by Former Smokers



Age reported starting smoking on most days by former smokers.

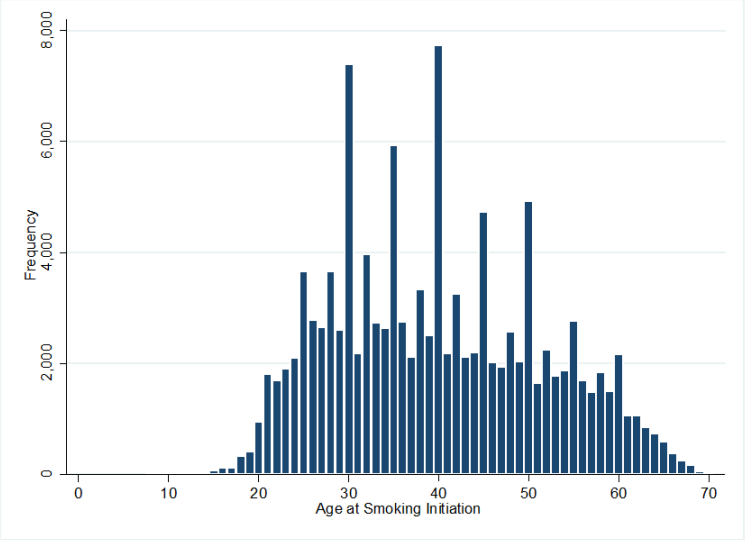
Figure 4.3.3 Age at Smoking Initiation by Current Smokers



Age reported starting smoking on most days by current smokers.

Figure 4.3.4 shows the age at smoking cessation reported by former smokers. The distribution shows a clear digit preference for multiples of five/ten.

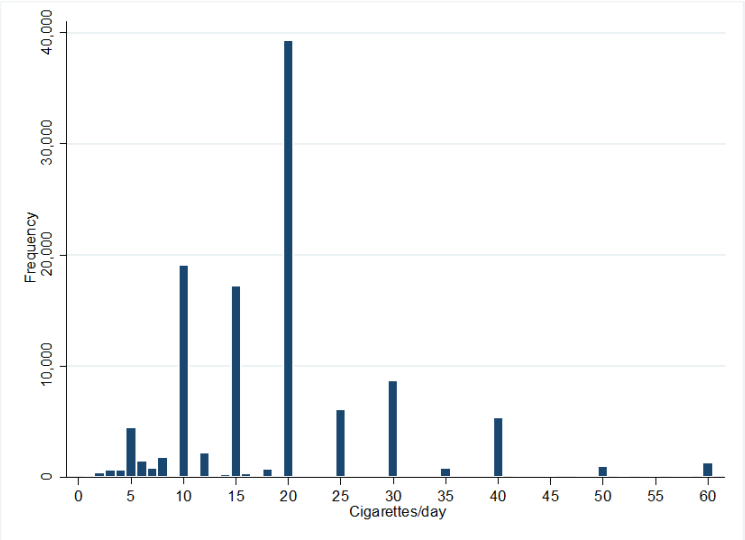
Figure 4.3.4 Age at Smoking Cessation



Age reported stopping smoking on most days by former smokers.

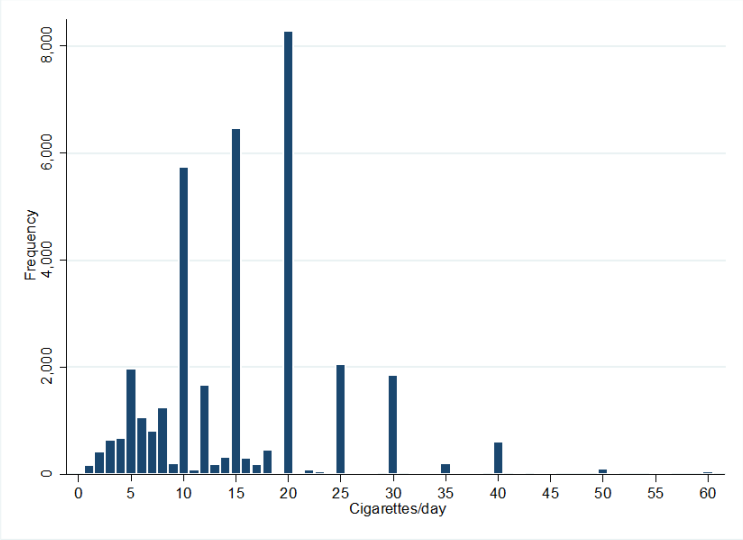
Of the 120,951 participants who reported smoking on most or all days in the past, 115,458 reported smoking cigarettes. Figure 4.3.5 shows a graph of the number of cigarettes smoked per day reported by these participants. 39,254 participants reported currently smoking tobacco on most or all days. Of these participants, 36,494 participants reported smoking cigarettes. Figure 4.3.6 shows the distribution of the number of cigarettes per day reported by these participants. Both graphs showed a clear digit preference for multiples of five/ten.

Figure 4.3.5 Number of Cigarettes Smoked per Day by Former Cigarette Smokers



Number of cigarettes smoked per day reported by participants who reported smoking cigarettes most or all days in the past.

Figure 4.3.6 Number of Cigarettes Smoked per Day by Current Cigarette Smokers



Number of cigarettes smoked per day reported by participants who reported smoking cigarettes most or all days.

4.3.3 Smoking Analysis Variables

Different variables were created in order to investigate the association between smoking and colorectal cancer. The first variable focused on overall smoking status in order to include as many participants as possible whereas further analysis variables focused on cigarette smoking since this was by far the most common form of tobacco use.

Overall Smoking Status

Overall smoking status was defined in order to try and analyse the range of exposures reported by participants. Participants were defined as follows:

Never ever

- Participants who reported currently not smoking tobacco and never smoking tobacco in the past.

Former occasional <100

- Participants who reported currently not smoking tobacco, reported smoking occasionally in the past or just tried once or twice and reported smoking less than 100 times.

Former occasional ≥100

- Participants who reported currently not smoking tobacco, reported smoking occasionally in the past or just tried once or twice and reported smoking at least 100 times.

Former daily

- Participants who reported currently not smoking tobacco and reported smoking on most or all days in the past.

Always occasional

- Participants who reported currently smoking tobacco occasionally and reported smoking occasionally in the past or just tried once or twice and reported smoking at least 100 times.

Occasional, former daily

- Participants who reported currently smoking tobacco occasionally and reported smoking on most or all days in the past.

Daily smokers

- Participants who reported currently smoking tobacco on most or all days.

Overall smoking status was defined for 493,096 participants (Table 4.3.2).

Table 4.3.2 Overall Smoking Status

Overall smoking status	Men	Women	Overall
Never ever	79,050	121,851	200,901
Former occasional <100	29,223	41,883	71,106
Former occasional ≥100	26,115	27,071	53,186
Former daily	62,892	53,025	115,917
Always occasional	4,580	3,118	7,698
Occasional, former daily	2,900	2,134	5,034
Daily	20,565	18,689	39,254
Total	225,325	267,771	493,096

Since this variable included categories with relatively few participants, a further variable was created which combined categories in order to define participants as never

smokers, former smokers or current smokers. Table 4.3.3 shows how these categories were combined to create the new variable.

Table 4.3.3 Overall Smoking Status

Overall smoking status	Never smoker	Former smoker	Current smoker
Never ever	200,901		
Former occasional <100	71,106		
Former occasional ≥100		53,186	
Former daily		115,917	
Always occasional			7,698
Occasional, former daily			5,034
Daily			39,254
Total	272,007	169,103	51,986

Cigarette Smoking Status

A further variable was created which restricted former smokers to former daily cigarette smokers and current smokers to current daily cigarette smokers. Participants were classified as never smokers, former cigarette smokers or current cigarette smokers according to the following criteria:

Never smoker

- Participants who reported currently not smoking tobacco and never smoking tobacco in the past

OR

- Participants who reported currently not smoking tobacco, reported smoking occasionally in the past or just tried once or twice and reported smoking less than 100 times.

Former cigarette smoker

- Participants who reported currently not smoking tobacco, reported smoking on most or all days in the past and reported smoking cigarettes.

Current cigarette smoker

- Participants who reported currently smoking tobacco on most or all days and reported smoking cigarettes.

Cigarette smoking status was defined for 419,256 participants (Table 4.3.4).

Table 4.3.4 Cigarette Smoking Status

Cigarette smoking status	Men	Women	Overall
Never smoker	108,273	163,734	272,007
Former cigarette smoker	57,917	52,838	110,755
Current cigarette smoker	17,943	18,551	36,494
Total	184,133	235,123	419,256

Cigarette Smoking Intensity

Next, former and current cigarette smokers were categorised according to the number of cigarettes they reported smoking per day (Table 4.3.5). Former and current cigarette smokers were classified slightly differently; the greater number of former cigarette smokers allowed a more precise classification using more categories. 685 former cigarette smokers and 321 current cigarette smokers said “do not know” or “prefer not to answer” for the questions on number of cigarettes smoked per day and were excluded.

Table 4.3.5 Cigarette Smoking Intensity

Cigarette smoking intensity	Participants
Never	272,007
Former, ≤10	27,844
Former, 11-15	19,137
Former, 16-20	39,331
Former, ≥21	23,758
Current, ≤10	12,953
Current, 11-20	17,989
Current, ≥21	5,231
Total	418,250

Age at Cigarette Smoking Initiation

Participants were also classified according to their age at initiation of smoking (Table 4.3.6). Participants who reported smoking cigarettes currently or in the past were asked to report how old they were when they started smoking on most days. 552 former cigarette smokers and 420 current cigarette smokers responded “do not know” or “prefer not to answer”.

Table 4.3.6 Age at Cigarette Smoking Initiation

Cigarette smoking initiation	Participants
Never	272,007
Former, ≤14	16,613
Former, 15-16	36,886
Former, 17-18	30,714
Former, ≥19	25,990
Current, ≤15	12,559
Current, 16-17	9,132
Current, ≥18	14,383
Total	418,284

Age at Cigarette Smoking Cessation

Former cigarette smokers were asked to report their age when they stopped smoking on most days. 483 participants responded “do not know” or “prefer not to answer”. Table 4.3.7 shows the analysis variable created.

Table 4.3.7 Age at Cigarette Smoking Cessation

Cigarette smoking cessation	Participants
Never	272,007
≤29	24,198
30-39	33,765
40-49	27,843
≥50	24,466
Total	382,279

Years since Cigarette Smoking Cessation

Former cigarette smokers were also classified according to the time since cessation by calculating the difference between age at baseline and the age they stopped smoking on most days (Table 4.3.8).

Table 4.3.8 Years since Cigarette Smoking Cessation

Years since cigarette smoking cessation	Participants
Never	272,007
≤9	29,709
10-19	25,417
20-29	29,685
≥30	25,461
Total	382,279

Cigarette Smoking Duration

For former cigarette smokers, smoking duration was calculated as the difference between the reported age when participants started smoking on most days and the reported age when participants last smoked on most days. For current cigarette smokers, smoking duration was calculated as the difference between the reported age when participants started smoking on most days and the age at baseline. 971 former cigarette smokers responded “do not know” or “prefer not to answer” for at least one of age started smoking and age stopped smoking and 420 current cigarette smokers responded “do not know” or “prefer not to answer” for age started smoking. Table 4.3.9 shows the analysis variable for cigarette smoking duration.

Table 4.3.9 Cigarette Smoking Duration

Cigarette smoking duration	Participants
Never	272,007
Former, ≤9	16,171
Former, 10-19	34,722
Former, 20-29	28,978
Former, 30-39	19,951
Former, ≥40	9,962
Current, ≤29	8,204
Current, 30-39	12,750
Current, ≥40	15,120
Total	417,865

4.4 Other Analysis Variables

This section provides further information for other variables relevant to the analyses.

4.4.1 Confounder Variables

Analyses of alcohol intake were adjusted for BMI and overall smoking status, analyses of adiposity were adjusted for alcohol intake and overall smoking status and analyses of smoking were adjusted for alcohol intake and BMI. All analyses of men and women were adjusted for sex. In addition, all analysis models were adjusted for the following confounder variables: Townsend deprivation index, red meat intake, processed meat intake, family history of bowel cancer and bowel cancer screening. Analyses of alcohol intake and smoking were also adjusted for height. These variables were selected as confounders since there was strong evidence that these variables influence colorectal cancer risk.[19, 25, 26, 28, 384-386] These variables are described below.

Townsend Deprivation Index

The Townsend deprivation index is a measure of socioeconomic status based on measures of unemployment, household overcrowding, non-home ownership and non-car ownership.[387] A higher Townsend deprivation index indicates a higher level of deprivation. Townsend deprivation index scores were calculated based on census output areas. Participants were assigned the index score for the output area in which their

postcode was located. 627 participants had missing data. Participants were categorised into approximate quintiles of deprivation; the categories were <−3.93, −3.93-<−2.76, −2.76-<−1.30, −1.30-<1.35 and ≥1.35.

Red Meat Intake

In the touchscreen questionnaire, participants were asked separate questions about their average intake of beef, lamb/mutton and pork. The left column of Table 4.4.1 shows the options available for each question. Participants could also respond "do not know" or "prefer not to answer". Data were missing for 898 participants and 6,107 participants who responded "do not know" or "prefer not to answer" to at least one of the questions were excluded. The right column of Table 4.4.1 shows the values assigned to each response. Values from the three questions were added together and participants were categorised as eating red meat ≤1, >1-<3 and ≥3 times a week (Table 4.4.2).

Table 4.4.1 Coding Values for Average Intake of Beef, Lamb/Mutton and Pork

Response category	Assigned number
Never	0
Less than once a week	0.25
Once a week	1
2-4 times a week	3
5-6 times a week	5.5
Once or more daily	7

Table 4.4.2 Red Meat Intake

	Men		Women		Overall	
Red meat intake	**n**	**%**	**n**	**%**	**n**	**%**
≤once/week	94,330	41.8	133,607	49.5	227,937	46.0
>1-<3 times/week	73,250	32.5	83,572	31.0	156,822	31.6
≥3 times/week	58,001	25.7	52,878	19.6	110,879	22.4
Total	225,581	100.0	270,057	100.0	495,638	100.0

Processed Meat Intake

In the touchscreen questionnaire, participants were asked how often they eat processed meats. Options were the same as for the questions on beef, lamb/mutton and pork. 896

participants had missing data and 1,337 participants responded “do not know” or “prefer not to answer”. Participants were categorised as eating processed meat <once, once and >once per week (Table 4.4.3).

Table 4.4.3 Processed Meat Intake

Processed meat intake	Men		Women		Overall	
	n	%	n	%	n	%
<once/week	61,186	26.8	138,020	50.7	199,206	39.8
once/week	67,929	29.8	78,142	28.7	146,071	29.2
>once/week	98,939	43.4	56,193	20.6	155,132	31.0
Total	228,054	100.0	272,355	100.0	500,409	100.0

Height

Height was measured at baseline. Unfortunately, the height data were not included in the dataset. However, data on weight and BMI were available. Thus, height was calculated as the square root of weight (kg) divided by BMI (kg/m^2):

$$\sqrt{\frac{weight}{body\ mass\ index}}$$

3,231 participants did not have data on BMI or weight and so height was defined for 499,411 participants.

Family History of Bowel Cancer

Questions relating to illnesses of family members were asked in the touchscreen questionnaire. 7,359 participants reported being adopted and thus did not answer these questions. 907 participants had missing data for these questions. Participants were first asked if their father was still alive. If they responded other than “do not know” or “prefer not to answer”, they were asked to select from a list of illnesses (including bowel cancer) which illnesses their father had ever suffered. Participants were then asked the same questions with regards to their mother. Participants were also asked how

many brothers they have and also how many sisters they have. If they had at least one sibling they were asked to select the illnesses suffered by any of their siblings.

Three variables were first created to identify whether there was a history of bowel cancer for participants' father, mother and siblings (Table 4.4.4). 25,107 participants reported that their father suffered from bowel cancer. Participants who responded "do not know" to which illnesses their father had suffered (41,979 participants) or responded "do not know" to whether their father was alive (6,156 participants) were included as no history of bowel cancer. 319 participants responded "prefer not to answer" to whether their father was alive and 1,103 participants responded "prefer not to answer" to which illnesses their father had suffered and were excluded.

Table 4.4.4 History of Bowel Cancer in Family Members

History of bowel cancer	Father	Mother	Siblings
No	467,847	469,952	483,482
Yes	25,107	23,242	9,486
Total	492,954	493,194	492,968

23,242 participants reported a history of bowel cancer for their mother. 1,095 participants responded "do not know" to whether their mother was alive and 25,908 participants responded "do not know" to which illnesses their mother had suffered. These participants were included as no history of bowel cancer. 241 participants responded "prefer not to answer" to whether their mother was alive and 941 participants responded "prefer not to answer" to which illnesses their mother had suffered and were excluded.

9,486 participants reported that a sibling had suffered from bowel cancer. 59,382 participants reported having no brothers and no sisters and were included as no history of bowel cancer. 30,402 participants responded "do not know" to which illnesses their siblings had suffered and were included as no history of bowel cancer. 434 participants responded "do not know" to both of the questions on the number of brothers and the number of sisters or "do not know" to one and zero to the other and were included as no history of bowel cancer. 869 participants responded "prefer not to answer" to which illnesses their siblings had suffered and were excluded. 539 participants responded

"prefer not to answer" to at least one of the questions on the number of brothers and number of sisters and did not report siblings in the other question and were excluded.

Finally, participants were defined as having a family history of bowel cancer if there was a history of bowel cancer for at least one family member (father, mother or sibling) whereas participants were defined as having no family history if there was no history for any family member. 491,623 participants were included in this variable (Table 4.4.5).

Table 4.4.5 Family History of Bowel Cancer

	Men		Women		Overall	
Family history of bowel cancer	**n**	**%**	**n**	**%**	**n**	**%**
No	198,436	88.6	238,531	89.1	436,967	88.9
Yes	25,485	11.4	29,158	10.9	54,643	11.1
Total	223,921	100.0	267,689	100.0	491,610	100.0

Bowel screening

Participants were asked in the touchscreen questionnaire "Have you ever had a screening test for bowel (colorectal) cancer? (Please include tests for blood in the stool/faeces or a colonoscopy or a sigmoidoscopy)". Data were missing for 925 participants. 8,800 participants responded "do not know" or "prefer not to answer". Thus, 492,917 participants were included in the variable for bowel screening (Table 4.4.6).

Table 4.4.6 Bowel Screening

	Men		Women		Overall	
Bowel screening	**n**	**%**	**n**	**%**	**n**	**%**
No	149,541	67.0	187,821	69.6	337,362	68.4
Yes	73,628	33.0	81,927	30.4	155,555	31.6
Total	223,169	100.0	269,748	100.0	492,917	100.0

The higher rate of screening among men was surprising since women have slightly higher rates of bowel screening uptake than men in the general population.[15, 388] Bowel screening was only offered to men and women aged 60-74 though UK Biobank includes men and women aged 40-69 so it was possible that the greater proportion of men than

women reporting screening was related to the difference in age groups. However, comparing the proportion of men and women reporting screening by age group, men were more likely to report screening than women within each age group (except <45 where the proportion of men and women reporting screening was similar) (Table 4.4.7).

Table 4.4.7 All Bowel Screening Responses by Age and Sex

	Age group						Total
Bowel screening (%)	**<45**	**45-49**	**50-54**	**55-59**	**60-64**	**≥65**	
Men							
Prefer not to answer	0.2	0.2	0.1	0.1	0.1	0.1	0.1
Do not know	2.3	2.6	2.6	2.8	1.9	2.0	2.3
No	87.8	84.0	79.9	75.6	48.2	43.9	65.4
Yes	9.7	13.2	17.3	21.5	49.8	54.0	32.2
Women							
Prefer not to answer	0.1	0.1	0.1	0.1	0.1	0.1	0.1
Do not know	1.1	1.1	1.1	1.1	1.0	1.2	1.1
No	89.1	85.9	81.8	78.8	50.8	46.7	68.8
Yes	9.7	12.9	17.0	20.0	48.2	52.0	30.0

4.4.2 Other Variables

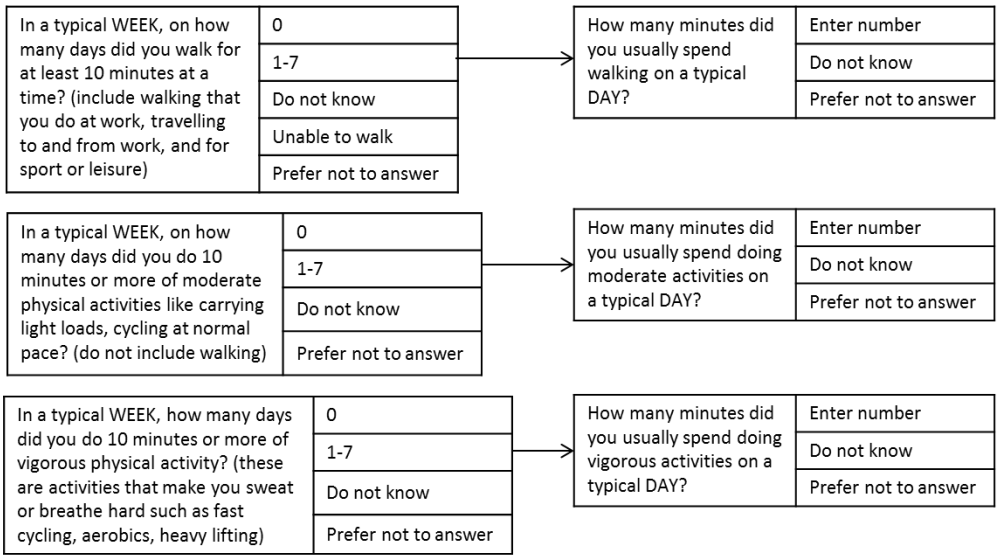
This section describes other variables that were not included as confounders but were used in the analyses e.g. for the investigation of potential effect modifiers.

Physical activity

Physical activity is associated with colorectal cancer risk.[25] However, physical activity was not included as a confounder in the analyses presented in this thesis. Analysis results were very similar with and without adjustment for physical activity (results for alcohol intake and overall smoking status and colorectal cancer with and without adjustment for physical activity are shown in Table A-3 and Table A-4 in the appendix) (results for BMI and colon and rectal cancer with and without adjustment for physical activity are shown in section 6.2.6) and unfortunately physical activity could not be defined for a large number of participants.

Participants gave information on physical activity in the touchscreen questionnaire during the baseline assessment. Participants were asked to report the number of days in a typical week they walk at least ten minutes. If participants entered a number between one and seven, they were then asked to report the average duration of walking on a typical day. Similar questions were asked regarding moderate activity and vigorous activity. These questions are shown in Figure 4.4.1. In the pilot study, participants were given five categories to choose from for each question on average duration. These were re-coded as follows: "less than 30 mins" = 15 minutes, "30 minutes to 1 hour" = 45 minutes, "1 to 2 hours" = 90 minutes, "2 to 4 hours" = 180 minutes and "more than 4 hours" = 300 minutes.

Figure 4.4.1 Questions on Average Physical Activity



These questions were based on the International Physical Activity Questionnaire (IPAQ) short form.[389] There exist guidelines for the scoring of these data available online at www.sites.google.com/site/theipaq/scoring-protocol.

For each activity, the frequency was multiplied by the average duration to obtain the average number of minutes per week. To calculate MET-minutes per week (MET = metabolic equivalent), the number of minutes per week of walking, moderate activity and vigorous activity were multiplied by 3.3, 4.0 and 8.0 METs, respectively. Participants were classified into three categories (low, moderate, high) of physical activity according to the following definitions:

Low

- No activity or some activity is reported but less than the definition for moderate or high physical activity.

Moderate

- Three or more days of vigorous activity of at least 20 minutes per day

OR

- Five or more days of moderate activity and/or walking of at least 30 minutes per day

OR

- Five or more days of any combination of walking, moderate activity or vigorous activity accumulating at least 600 MET-minutes per week.

High

- Vigorous activity on at least three days and accumulating at least 1,500 MET-minutes per week

OR

- Seven or more days of any combination of walking, moderate activity or vigorous activity accumulating at least 3,000 MET-minutes per week.

Participants were excluded if they responded "do not know" or "prefer not to answer" to any of the questions. Participants were excluded if the sum of the walking, moderate activity and vigorous activity duration variables was greater than 960 minutes (16 hours). Responses of less than ten minutes were replaced with zero and responses greater than 180 minutes were truncated. Thus, physical activity was defined for 388,771 participants (Table 4.4.8).

Table 4.4.8 IPAQ Physical Activity Variable

	Men		Women		Overall	
Physical activity	**n**	**%**	**n**	**%**	**n**	**%**
Low	35,205	18.9	36,681	18.1	71,886	18.5
Moderate	71,372	38.3	86,571	42.7	157,943	40.6
High	79,560	42.7	79,382	39.2	158,942	40.9
Total	186,137	100.0	202,634	100.0	388,771	100.0

Unfortunately, 111,487 participants responded "do not know" or "prefer not to answer" to at least one of these questions and so could not be included in this variable.

Therefore, an alternative physical activity variable was created which excluded fewer participants (by focusing on responses to the number of days only and avoiding excluding participants for responding "do not know" or "prefer not to answer" when possible). This variable classified participants into three categories of physical activity using the following definitions:

Low

- Fewer than three days of moderate activity and fewer than two days of vigorous activity.

Medium

- Three or more days of moderate activity

OR

- Two or more days of vigorous activity.

High

- Three or more days of vigorous activity.

This variable included 464,246 participants (Table 4.4.9).

Table 4.4.9 Alternative Physical Activity Variable

	Men		Women		Overall	
Physical activity	**n**	**%**	**n**	**%**	**n**	**%**
Low	59,750	27.8	70,160	28.1	129,910	28.0
Medium	75,170	35.0	104,298	41.8	179,468	38.7
High	79,838	37.2	75,030	30.1	154,868	33.4
Total	214,758	100.0	249,488	100.0	464,246	100.0

Among the 388,771 participants with data for both physical activity variables, there was reasonable agreement between the two variables (Table 4.4.10). For example, only 1.9% of participants included in the low category in the IPAQ variable were included in the high category in the alternative variable and only 2.5% of participants included in the

high category in the IPAQ variable were included in the low category in the alternative variable.

Table 4.4.10 Comparison of Physical Activity Variables

	Alternative variable			
IPAQ variable (%)	**Low**	**Medium**	**High**	**Total**
Low	80.4	17.7	1.9	100.0
Medium	32.0	58.6	9.4	100.0
High	2.5	26.7	70.8	100.0

Unfortunately, this alternative physical activity variable still resulted in the exclusion of 37,517 participants who responded "do not know" or "prefer not to answer". Multiple imputation could have been considered in order to "fill in" the physical activity values for these participants. However, valid multiple imputation relies on the assumption that the probability that a participant has missing data is independent of the true value[375] and this assumption seemed questionable for physical activity data. For example, a participant may have responded "do not know" or "prefer not to answer" because they did a lot of physical activity and found it too taxing to try and answer all the questions accurately. Alternatively, a participant may have responded "do not know" or "prefer not to answer" because they did not do much physical activity but did not want to admit it. Hence, multiple imputation was not used for physical activity data.

Another approach for dealing with missing data is to add an extra category for people with missing data. However, this approach would combine people with different levels of exposure and can produce biased results.[113, 390] Despite this fact, this method is fairly commonly used in epidemiological studies.

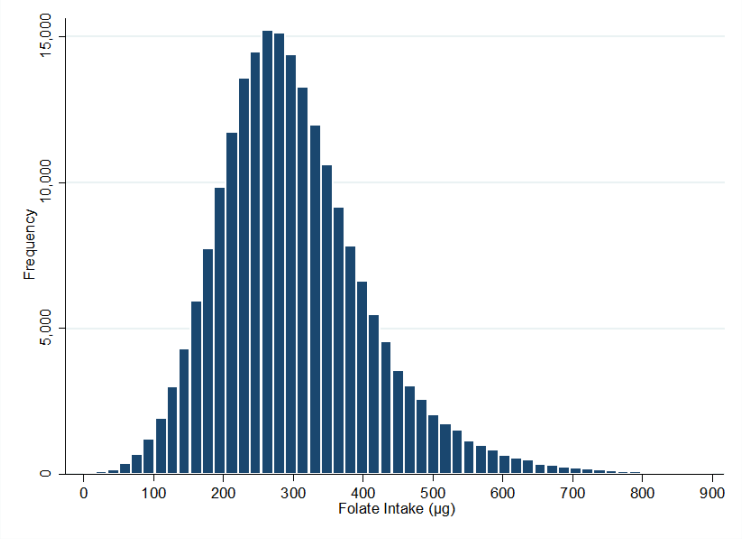
Folate

To obtain further information on dietary factors, an online questionnaire was created, asking participants to report the food and drink consumed during the previous 24 hours. This was included as part of the assessment centre during the later stages of recruitment. Later, an invitation was sent to participants by e-mail to complete the questionnaire on four separate occasions. Hence, the questionnaire may have been completed by

participants up to five times. 211,063 participants completed at least one questionnaire and 5,772 participants completed all five questionnaires.

For each questionnaire, folate intake was calculated based on the reported consumption of food and drinks. The use of supplements was not included in the calculation of folate intake. The estimated folate intake values across the questionnaires were added and then divided by the number of questionnaires completed in order to obtain a single value of average folate intake for each participant. Figure 4.4.2 shows the distribution of average folate intake. Median folate intake was 288.7 µg.

Figure 4.4.2 Average Folate Intake



Average folate intake reported in online dietary questionnaire.

Menopause

Female participants were asked if they had had their menopause in the touchscreen questionnaire. 476 women had missing data for this question. 535 women responded "prefer not to answer". 31,185 participants responded "not sure - had a hysterectomy" and 11,735 participants responded "not sure - other reason". 64,092 women had not experienced menopause at baseline and 165,444 women had experienced menopause. Of these 165,444 women, 87,418 women reported never using hormone replacement therapy (HRT). Only pre-menopausal women and post-menopausal women who had never used HRT were included in this analysis variable (Table 4.4.11).

Table 4.4.11 Menopause Status for Women

Menopause	n	%
Pre-menopause	64,092	42.3
Post-menopause, never HRT users	87,418	57.7
Total	151,510	100.0

4.5 Outcome Data on Cancers and Deaths

4.5.1 Data Sources

Outcome data on cancers and deaths are obtained via linkage with national datasets. UK Biobank participants are flagged in these datasets and the relevant data are sent to UK Biobank on a regular basis.

Data on cancer diagnoses in the UK are collected from a variety of sources including hospital admissions, pathology laboratories, cancer screening programmes and death records. Cancer diagnoses in England are recorded by eight regional cancer registries which together form the National Cancer Registration Service (part of Public Health England) and cancer diagnoses in Wales are recorded by the Welsh Cancer Intelligence Surveillance Unit (part of Public Health Wales). These data are then submitted to the Office for National Statistics (ONS) and are made available to UK Biobank through the Health and Social Care Information Centre (with permission from the ONS). Cancer diagnoses in Scotland are recorded by the Scottish Cancer Registry and are provided to UK Biobank by the Information Services Division of NHS National Services Scotland.

Maintaining national cancer registries is a complex activity and data can never be 100% accurate and complete. There exist a number of measures for assessing the quality and validity of cancer registry data though their interpretation is not straightforward.[391, 392] One indication for the accuracy of cancer registry data is the proportion of cases for which the only information available is from a death certificate, known as death certificate only (DCO) registrations. For these registrations, the incidence date is unknown and thus the date of death is used instead. Thus, a DCO registration indicates that the cancer could not be identified while the person was alive and a high percentage of DCO registrations signifies incompleteness (although a low percentage does not guarantee the opposite). UK cancer registries aim to have a DCO rate of less than 2%.[393]

In 2012 the English regional cancer registries migrated to a single cancer registration system resulting in a considerable extra workload. As a result there was a noticeable increase in the rates of DCO registrations (2.5% for males and 2.8% for females) though these figures were still extremely good.[393] The average rates across the cancer registries of the UK and Ireland for 2012 were 1.1% for males and 1.3% for females. Further data on performance indicators of UK and Ireland cancer registries are available at www.ukiacr.org/kpis.

Data on deaths in England and Wales are compiled by the ONS and are provided to UK Biobank through the Health and Social Care Information Centre. Data on deaths for Scotland are collected and provided to UK Biobank by the National Health Service Central Register, Scotland.

4.5.2 Death Data

For each participant who died, the following data were available:

- Date of death
- Age at death*
- Primary cause of death (ICD-10)
- Secondary cause(s) of death (ICD-10)

*Age at death was calculated by UK Biobank as the difference between date of birth and date of death.

Causes of death were coded according to the 10th revision of the International Classification of Diseases (ICD-10), published by the WHO.[394]

4.5.3 Cancer Data

For each participant diagnosed with cancer, the following data were available (for each diagnosis):

- Date of cancer diagnosis
- Age at cancer diagnosis*
- Type of cancer (ICD-10)

- Type of cancer (ICD-9)
- Histology of neoplasm
- Behaviour of neoplasm

*Age at cancer diagnosis was calculated by UK Biobank as the difference between date of birth and date of cancer diagnosis.

Data were available on both prevalent and incident cancer diagnoses. The type of cancer was coded according to the International Classification of Diseases, 9th Revision (ICD-9) or 10th Revision (ICD-10). The use of the 10th revision replaced the use of the 9th revision between 1995 and 1996. Thus, all incident cancers are coded according to the 10th revision. Histology and behaviour were coded according to the International Classification of Diseases for Oncology 3rd Edition (ICD10-O-3).[9] Neoplasms were coded as one of the following behaviour types:

- 0, Benign
- 1, Uncertain whether benign or malignant
- 2, Carcinoma in situ
- 3, Malignant, primary site
- 6, Malignant, metastatic site
- 9, Malignant, uncertain whether primary or metastatic site

4.5.4 Outcome Definition

Baseline assessment and follow-up data on cancers and deaths were available on 502,642 participants. Follow-up data were complete up to 31st March 2014.

12 participants were excluded for erroneous follow-up data: three participants had a death date earlier than their assessment date, two participants had a primary cause of death but no death date, six participants had a date of death but no primary cause, one participant had a cancer diagnosis date later than the date of death.

26,857 participants who had a cancer diagnosis (other than non-melanoma skin cancer) (ICD-10 C00 - C97 except C44, ICD-9 140-209 except 173) before recruitment were excluded. 3,991 participants who reported suffering from inflammatory bowel disease,

Crohn's disease or ulcerative colitis were excluded because these participants may have a substantially elevated risk of colorectal cancer.[19]

Participants were censored at the date of their first cancer diagnosis (other than non-melanoma skin cancer) (ICD-10 C00 - C97 except C44), date of death or end of follow-up (31st March 2014), whichever came first. Participants were defined as colorectal cancer cases if their first cancer diagnosis was colorectal cancer (ICD-10 C18.0 - C20). If participants were diagnosed with colorectal cancer and another type of cancer on the same date they were defined as colorectal cancer cases. One participant whose first incident cancer diagnosis was colorectal cancer was excluded because the behaviour was coded as "malignant, uncertain whether primary or metastatic site".

Finally, the outcome was defined for 471,781 participants. 20,070 participants were censored at the date of a cancer diagnosis (including colorectal cancer) and 3,337 participants were censored at the date of death. All other participants were censored at 31st March 2014 (end of follow-up).

2,302 participants were defined as colorectal cancer cases. The definitions and the number of cases for colorectal, colon, proximal colon, distal colon and rectal cancer are shown in Table 4.5.1. Participants with diagnoses at different subsites on the same date were defined as cases for each analysis e.g. 21 participants had a colon cancer diagnosis and a rectal cancer diagnosis on the same date and were defined as cases for both the colon and rectal cancer analyses.

Proximal colon cancer was defined as cancer located in the caecum through to the splenic flexure (C180.0 - C18.5) and distal colon cancer was defined as cancer located in the descending colon and sigmoid colon (C18.6 - C18.7). Nine participants were defined as cases for both proximal and distal colon cancer. 96 participants who had cancer classified as "overlapping lesion of colon" (C18.8) or "colon, unspecified" (C18.9) (and did not have a separate proximal colon cancer diagnosis) were excluded from the analysis of proximal colon cancer. 103 participants who had cancer classified as "overlapping lesion of colon" (C18.8) or "colon, unspecified" (C18.9) (and did not have a separate distal colon cancer diagnosis) were excluded from the analysis of distal colon cancer.

Table 4.5.1 Definitions of Colorectal Cancer and Colorectal Cancer Subsites for All Analyses

ICD-10 code	Cancer type	Number of cases*	Colorectal	Colon	Proximal colon	Distal colon	Rectum
C18.0	Caecum	276	✓	✓	✓		
C18.1	Appendix	40	✓	✓	✓		
C18.2	Ascending colon	207	✓	✓	✓		
C18.3	Hepatic flexure	78	✓	✓	✓		
C18.4	Transverse colon	112	✓	✓	✓		
C18.5	Splenic flexure	54	✓	✓	✓		
C18.6	Descending colon	97	✓	✓		✓	
C18.7	Sigmoid colon	611	✓	✓		✓	
C18.8	Overlapping lesion of colon	9	✓	✓			
C18.9	Colon, unspecified	100	✓	✓			
C19	Rectosigmoid junction	157	✓				✓
C20	Rectum	638	✓				✓
	Number of participants		2,302	1,532	751	700	791

*Number of cases for each cancer type do not sum to total number of participants with colorectal cancer due to multiple diagnoses; 71 participants had two colorectal cancer diagnoses on the same date and three participants had three colorectal cancer diagnoses on the same date.

4.6 Statistical Methods

This section describes the statistical methods used in this thesis to analyse the UK Biobank data. All statistical analyses were performed using Stata 13.

4.6.1 Cox Proportional Hazards Models

Survival analysis methods are used to analyse the time until a certain event. One important difference between survival data and other types of data is that the variable of interest (the time until the event) is not necessarily observed for all subjects i.e. not all subjects will experience the event during the follow-up time of the study. People who do not experience the event during follow-up are censored.

In survival analysis, the survival function $S(t)$ is defined as the probability that an individual survives (i.e. does not experience the event) until at least time t. The hazard function $h(t)$ represents the probability of an event at time t (conditional on survival until t).

The Cox proportional hazards model is widely used for the analysis of survival data.[267] The mathematical form of the Cox proportional hazards model is:

$$\log\big(h(t)\big) = \log(h_0(t)) + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k$$

where $h_0(t)$ represents the baseline hazard function (corresponding to an individual for whom all exposure variables are equal to zero) and x_1 to x_k are the k exposure variables.

One of the main advantages of the Cox proportional hazards model over other methods of survival analysis is that it allows for the investigation of the effects of multiple variables on survival simultaneously. Also, it requires no assumptions about the form of the baseline hazard function. However, the key assumption of the Cox model is that the ratio of the hazard functions comparing different exposure groups is constant over time i.e. the hazards are proportional to each other over time. This is known as the proportional hazards assumption (investigated for the analyses in this thesis in section 4.6.2).

Age was used as the underlying time variable in all analyses since the risk of colorectal cancer is expected to change primarily according to age whereas there should be no important relation between time in the study and colorectal cancer risk. Only the month and year of birth were available for participants so age was calculated using the 15th day of the month as the day of birth.

The proportional hazards model assumes that there are no tied survival times (since the time to the event is continuous). This would mean that each event would occur at a unique time and thus the participants at risk of the event at each event time would be clear. However, tied events do occur in survival data since it is only possible to measure time with limited precision. For this analysis the Breslow method was used for handling ties.[395] When considering each individual tied event, this method assumes that all other participants who suffered the event at the tied time were at risk of the event at that time.

Hence, Cox proportional hazard models were used in this thesis to analyse the associations between alcohol intake, adiposity and smoking and colorectal cancer risk, adjusting for other risk factors. Age was used as the underlying time variable and time in the study was calculated up to the date of diagnosis of colorectal cancer, the date of diagnosis of any other cancer (excluding non-melanoma skin cancer), the date of death or the end of follow-up (31st March 2014).

4.6.2 Proportional Hazards Assumption

The use of the Cox proportional hazards model relies on the assumption that the HR is constant over time i.e. that the hazard functions are proportional to each other over time. This assumption can be assessed in a number of ways. One method is to plot -ln[-ln($S(t)$)] against ln(t) for each exposure group. These plots are usually referred to as "log-log" plots. If the hazard functions are proportional, then the lines representing the different exposure groups should be approximately parallel.[396]

Figure 4.6.1 shows the log-log plot for the exposure groups defined for alcohol intake. Overall, the lines for the different alcohol intake exposure groups were approximately parallel, suggesting that there was no reason to reject the proportional hazards assumption for the analysis of alcohol intake. The greater variation observed at the start

was due to the relatively few cases in this cohort at an earlier age. Log-log plots were similar for BMI (Figure 4.6.2 and Figure 4.6.3) and overall smoking status (Figure 4.6.4), and so the proportional hazards assumption was not rejected for these analyses either. (Separate plots are presented for men and women for BMI due to the evidence from the existing literature that the association between BMI and colorectal cancer differs for men and women (see section 2.3)).

Figure 4.6.1 Log-log Plot of Alcohol Intake

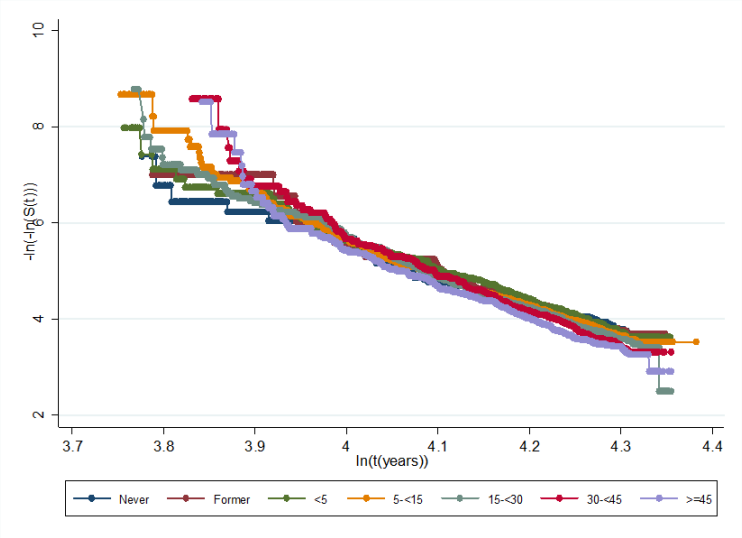


Figure 4.6.2 Log-log Plot of Body Mass Index Quintiles for Men

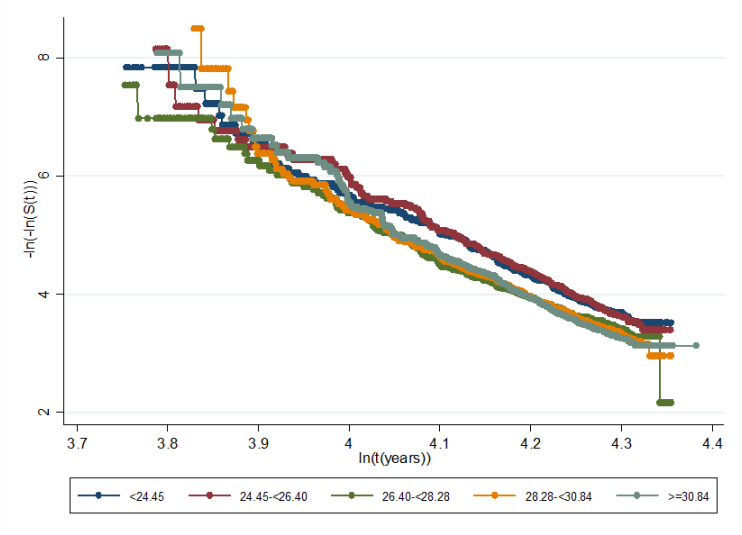


Figure 4.6.3 Log-log Plot of Body Mass Index Quintiles for Women

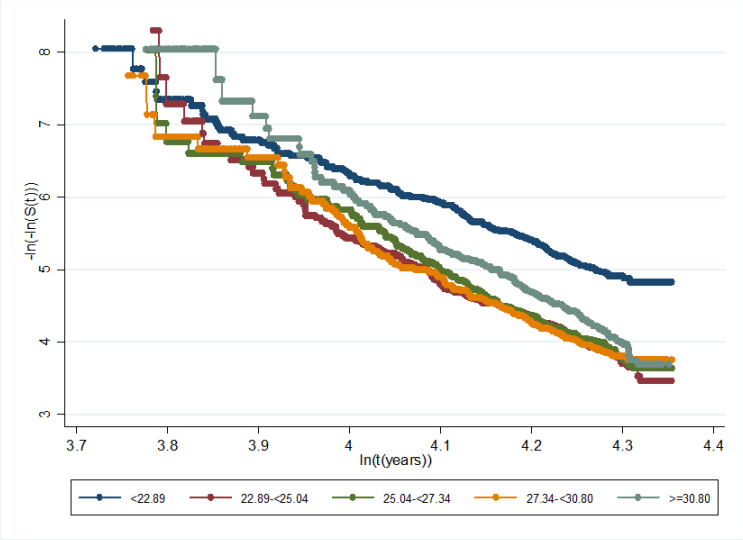
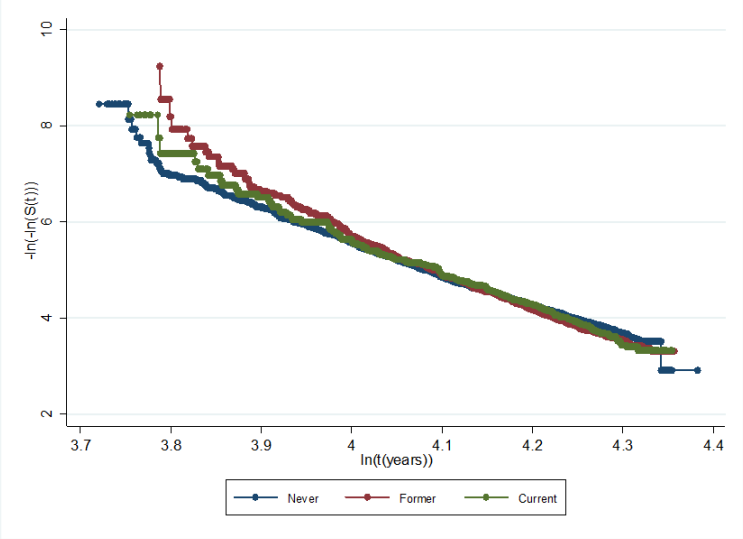
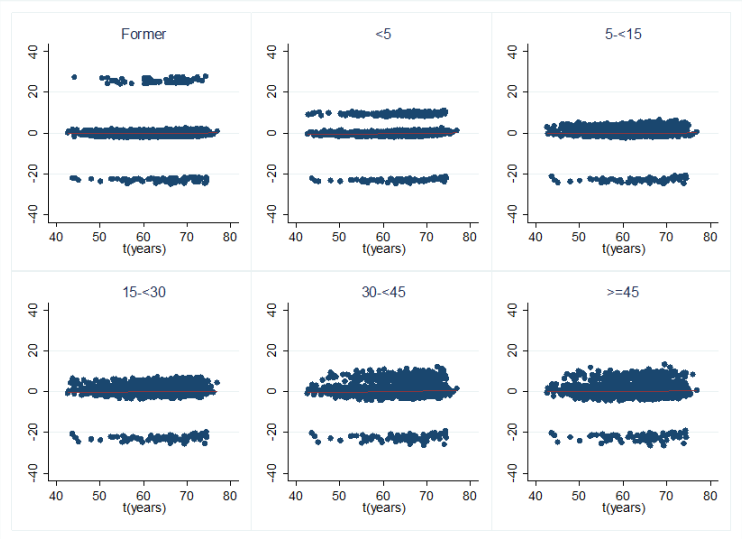


Figure 4.6.4 Log-log Plot of Overall Smoking Status



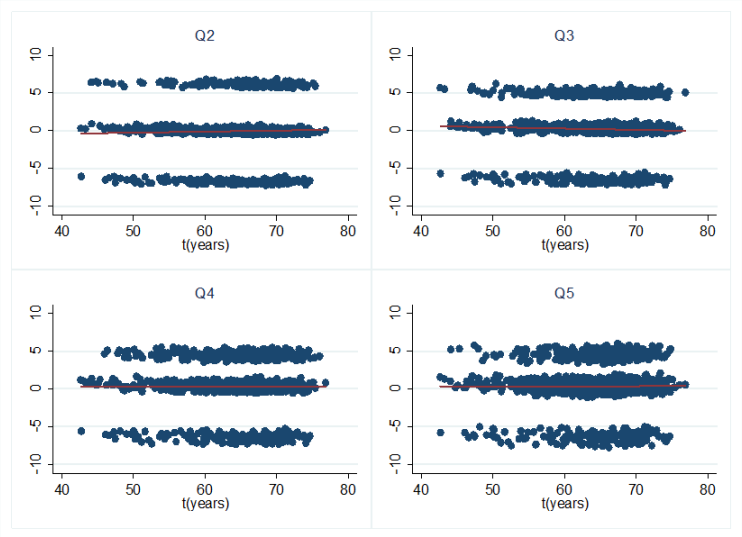
Another method for assessing the proportionality hazards assumption is using scaled Schoenfeld residuals. Scaled Schoenfeld residuals represent the difference between the observed and predicted values of the covariates at each event (calculated separately for each covariate).[396] Hence, the proportional hazards assumption can be assessed by plotting the scaled Schoenfeld residuals against follow-up time. If the proportional hazards assumption is valid, there should be no systematic trend in the residuals in relation to time. Figure 4.6.5 - Figure 4.6.8 show the scaled Schoenfeld residuals plotted against follow-up time. Since there were no clear trends in the residuals against follow-up time, the proportional hazards assumption was deemed appropriate.

Figure 4.6.5 Scaled Schoenfeld Residuals against Follow-up Time (Age) for Alcohol Intake



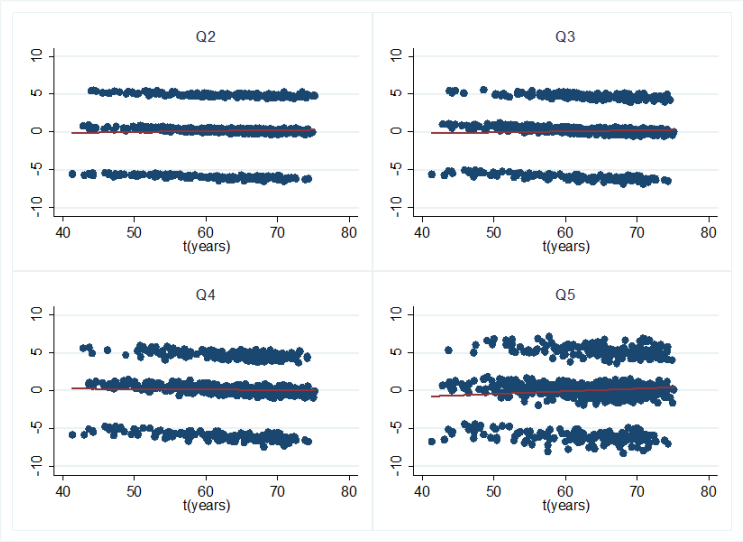
Scaled Schoenfeld residuals over follow-up time by category of alcohol intake. Red line represents the fitted regression line.

Figure 4.6.6 Scaled Schoenfeld Residuals against Follow-up Time (Age) for BMI Quintiles for Men



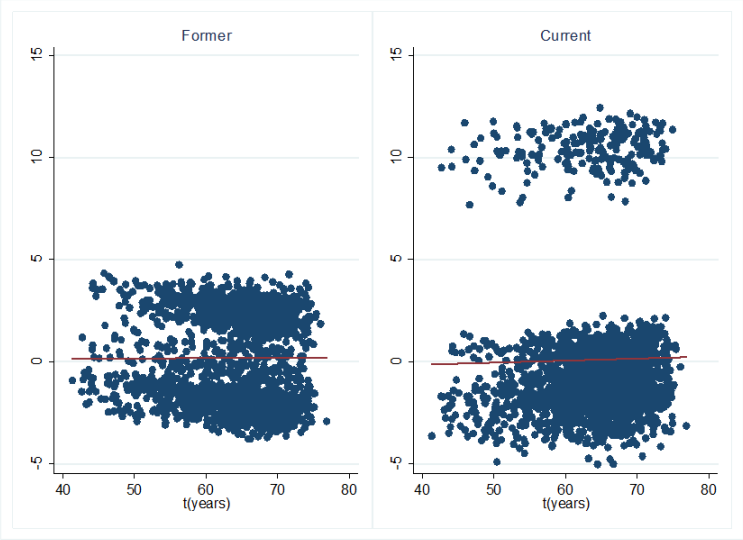
Scaled Schoenfeld residuals over follow-up time by quintiles of BMI for men. Red line represents the fitted regression line.

Figure 4.6.7 Scaled Schoenfeld Residuals against Follow-up Time (Age) for BMI Quintiles for Women



Scaled Schoenfeld residuals over follow-up time by quintiles of BMI for men. Red line represents the fitted regression line.

Figure 4.6.8 Scaled Schoenfeld Residuals against Follow-up Time (Age) for Overall Smoking Status



Scaled Schoenfeld residuals over follow-up time by category of overall smoking status. Red line represents the fitted regression line.

4.6.3 Fractional Polynomials

As mentioned in section 4.1.4, alcohol intake was considered as a continuous analysis variable as well as a categorical analysis variable. A continuous analysis variable may

be considered to be a more efficient use of the available data since it avoids the arbitrary cut-points used for categorical analysis variables. However, the use of continuous analysis variables requires an assumption to be made about the shape of the relationship between the variable of interest and the outcome. Continuous analysis variables are often modelled using linear models but the linearity assumption must always be evaluated. When a linear model is not suitable, the relationship can often be modelled using more complex polynomial models e.g. x^2 or $\ln(x)$.

Fractional polynomial analysis was used in this thesis to evaluate the linearity assumption for the relationship between alcohol intake and colorectal cancer and to identify the "best-fitting" model.[397, 398] The advantage of fractional polynomial analysis is that it is a very flexible method since it allows the investigation of different polynomial models based on a wide range of powers.

In this thesis, all possible polynomial models of one (e.g. cx^a) or two terms (e.g. $cx^a + dx^b$) were considered with powers from the set {-2, -1, -0.5, 0, 0.5, 1, 2, 3} (where 0 represents the natural logarithm). (When the same power was selected for a two term polynomial model, the second term was multiplied by the natural logarithm). From all these possible models, the one and two term polynomial models with lowest deviance (-2 log likelihood) were identified and a likelihood ratio test was used to compare these models.

4.6.4 Population Attributable Fraction

The population attributable fraction (PAF) is an estimate of the proportion of disease in a population due to exposure to a certain risk factor.[399] In other words, it estimates the proportion of cases that would be prevented if no participants were exposed to the risk factor of interest, while the distribution of other risk factors in the population stays the same. Therefore, it assumes a causal association between the exposure and disease. The PAF is calculated as

$$PAF = p_d \frac{RR - 1}{RR}$$

Where p_d is the proportion of cases exposed to the risk factor of interest and RR is the adjusted relative risk.[399]

PAFs were estimated for a range of factors considered in this thesis: alcohol intake, BMI, overall smoking status, Townsend deprivation index, red meat intake, processed meat intake, family history of colorectal cancer and physical activity. These variables were all modelled together using Cox Proportional Hazards models (results were also adjusted for sex and colorectal screening). Following the Cox models, PAFs and 95% CIs were calculated using the Stata punafcc command, based on the adjusted results from the Cox models.[400]

For alcohol intake, exposure was defined as >14 units/week, following the latest guidelines (grams of alcohol were divided by 7.9 to convert to units).[58, 401] Former drinkers were included in the non-exposed group (≤14 units/week). Excluding the participants who reported drinking "one to three times a month" or "special occasions only" and had missing data will overestimate the exposure prevalence in the population (since those with missing data will be more likely to be in the non-exposed group) and overestimate the PAF. Among participants who reported drinking "one to three times a month" or "special occasions only" and had complete data, more than 99.5% of participants reported drinking ≤14 units per week. Therefore, to calculate the PAF for alcohol intake in this cohort, it was assumed that those participants with missing data drank ≤14 units per week and were included in the non-exposure group.

Exposure for adiposity was defined as BMI ≥25 kg/m^2 and exposure for smoking was defined as ever smoking (current smokers and former smokers). For Townsend deprivation index, exposure was defined as a Townsend index above the lowest quintile of deprivation, based on 2001 national census data (see section 3.8). Exposure for red meat intake and processed meat intake was >once/week. Participants were included in the exposure group if they did have a family history of colorectal cancer. Exposure for physical activity was not completing at least 150 minutes of moderate activity or 75 minutes of vigorous activity (or an equivalent combination) per week.[402]

PAF results for are shown in section 7.3. The PAF depends on the prevalence of exposure in the cohort. UK Biobank participants were highly self-selected which means

that they will likely be healthier and more likely to follow a healthy lifestyle on average compared to the general population. For example, there was a very low prevalence of current smoking among participants (although alcohol intake appeared to be slightly higher on average than in the general population) (see Chapter 3). Therefore, the PAF values presented in this thesis are likely to be biased and do not necessarily represent the PAF estimates for the general population.

4.6.5 Multiple Imputation

Missing Data and Multiple Imputation

Missing data have always been a pervasive problem in epidemiological research and for a long time the best approach for dealing with missing data (besides not having any) was to ignore subjects with missing data and simply perform a complete case analysis. However, excluding participants with missing data results in a loss in precision and can lead to biased results in certain contexts.[376] Recently, the use of statistical techniques to "fill in" these missing values has become more widespread as they are now easily available in many standard statistical software packages.

The first step when dealing with missing data is to try to understand why the data are missing. Missing data are commonly classified into three types.[403] Data are missing completely at random (MCAR) when the probability of being missing does not depend on any observed or missing data i.e. the probability of being missing is the same for all people/observations. Data are missing at random (MAR) when the probability of being missing depends on observed data but not on missing data (for example if younger people are less likely to have their blood pressure measured). Data are missing not at random (MNAR) when the probability depends on the values of the missing data (for example if people with low blood pressure are less likely to have their blood pressure measured).

If data are MCAR, the results of a complete case analysis (ignoring people with missing data) will not be biased, though there may be a loss of precision. A complete case analysis may lead to biased results when data are MAR. For example, if younger people are less likely to have their blood pressure measured than older people, then the estimate

of mean blood pressure using only people with complete data will overestimate the true value of mean blood pressure in the population (given that older people tend to have a higher blood pressure than younger people).

Multiple imputation is a common method for the handling of missing data. The underlying motivation for multiple imputation is fairly intuitive. A first approach for "filling in" missing values might be to simply replace the missing data with the mean value of the observed data but this would distort associations towards the null. Instead of filling in missing values using the mean value of the observed data, a regression model based on the observed data could be used to predict the missing values. Hence, the association between the variables in the observed data would be preserved in the missing data. However, the imputed values would be too precise since this method ignores the variation in the regression model i.e. all the imputations would lie perfectly on the regression line. A better approach would be to introduce random error into the imputations based on the residual variance from the regression model in order to preserve the variability in the data. However, this approach is still flawed because the imputed values would be treated as if they were known i.e. the uncertainty of the predictions would not be reflected in the final results. Multiple imputation addresses this problem by repeating this stochastic approach a number of times to create a number of different imputed datasets and then estimating the quantities of interest in each imputed dataset before combining these estimates to take into account the variation between imputations.

Since multiple imputation uses the relationships among the observed data to fill in the missing values, valid multiple imputation relies on the assumption that the missing data are MAR i.e. the probability that the data are missing cannot be related to the unobserved values (MNAR). Otherwise, there would be no way to reliably estimate the missing values using only the observed data. In reality though, it is not possible to prove that the missing data are not MNAR using only the observed data. Hence, careful consideration of why data are missing is necessary when performing multiple imputation.

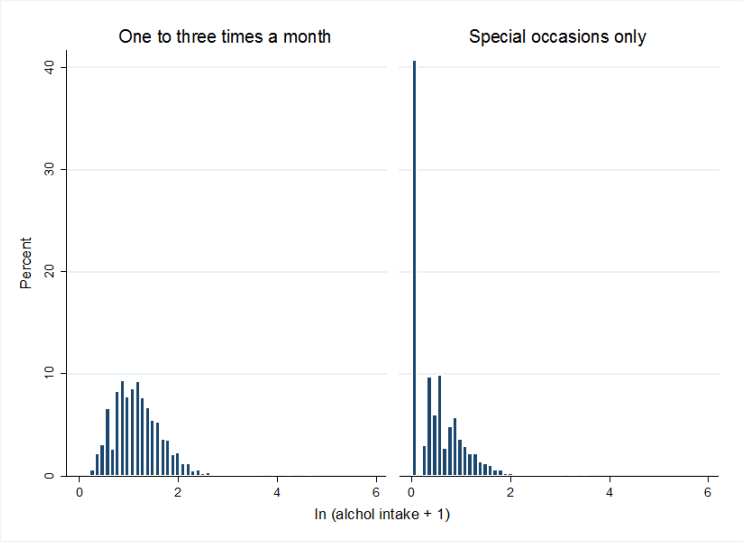
Missing Alcohol Intake Data in UK Biobank

The missing data for alcohol intake were explained in section 4.1.3. Briefly, 73,061 participants who reported drinking "one to three times a month" or "special occasions only" were not included in the alcohol grams per day variable since they had missing data for the questions on the average monthly intake of different alcoholic beverages. These data were missing because these participants completed an earlier version of the questionnaire that did not include these questions. Instead, these participants answered questions on their average weekly intake of different alcoholic beverages but these data were not made available.

The probability that the alcohol data were missing depended on the assessment date as well as assessment centre since not all centres used the same questionnaire at the same times (Figure 4.1.2 shows when the new questionnaire was introduced at each assessment centre). The probability that the data were missing also depended on alcohol frequency since there were missing data only for participants reporting either "special occasions only" or "one to three times a month". Crucially, since the data were missing due to a change in the questionnaire, the probability that the alcohol data were missing was unrelated to the values of the missing data (given alcohol frequency). Hence, alcohol grams per day data were MAR, conditional on assessment date, assessment centre and alcohol frequency.

The distribution of alcohol grams per day among participants with observed alcohol data was very positively skewed. Figure 4.6.9 shows a histogram of the natural logarithm of alcohol grams per day (plus one g/d to avoid undefined values) for participants who reported drinking "special occasions only" and "one to three times a month". The distribution of the natural logarithm of alcohol grams per day was approximately Normal among participants who reported "one to three times a month" but the distribution for "special occasions only" was not Normal, due to the large number of participants who reported drinking zero grams per day.

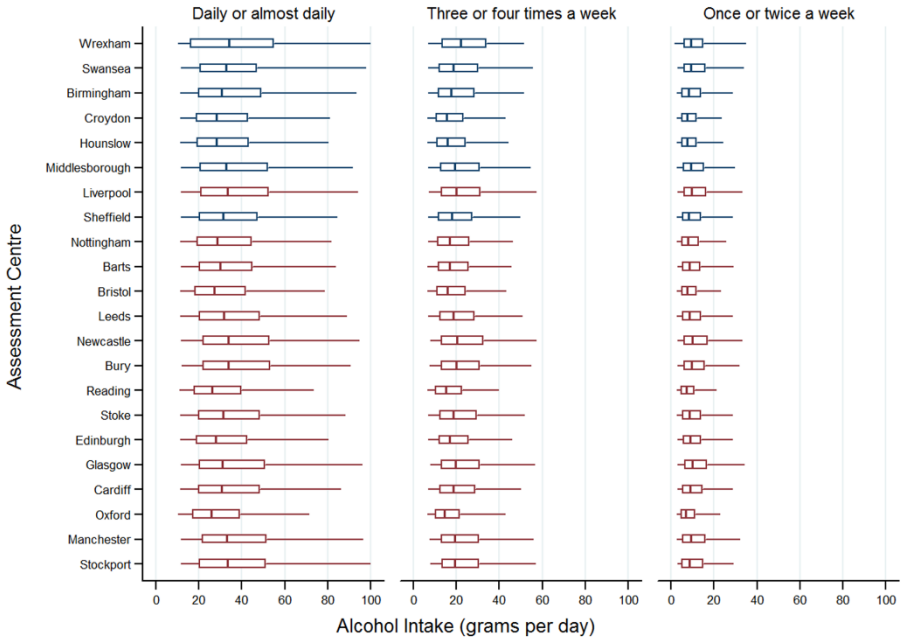
Figure 4.6.9 Distribution of Alcohol Intake by Alcohol Frequency for Participants Drinking "One to Three Times a Month" or "Special Occasions Only"



Histogram of ln(alcohol grams per day + 1) by alcohol frequency for participants who reported drinking "one to three times a month" or "special occasions only".

Figure 4.6.10 shows the distribution of alcohol grams per day by assessment centre for participants who reported drinking "daily or almost daily", "three or four times a week" or "once or twice a week". The distribution of alcohol intake was similar across assessment centres, within alcohol frequency. The range of median alcohol intake across assessment centres for the three categories of alcohol frequency was 26.0-34.3, 14.8-22.4 and 7.4-10.3 g/d.

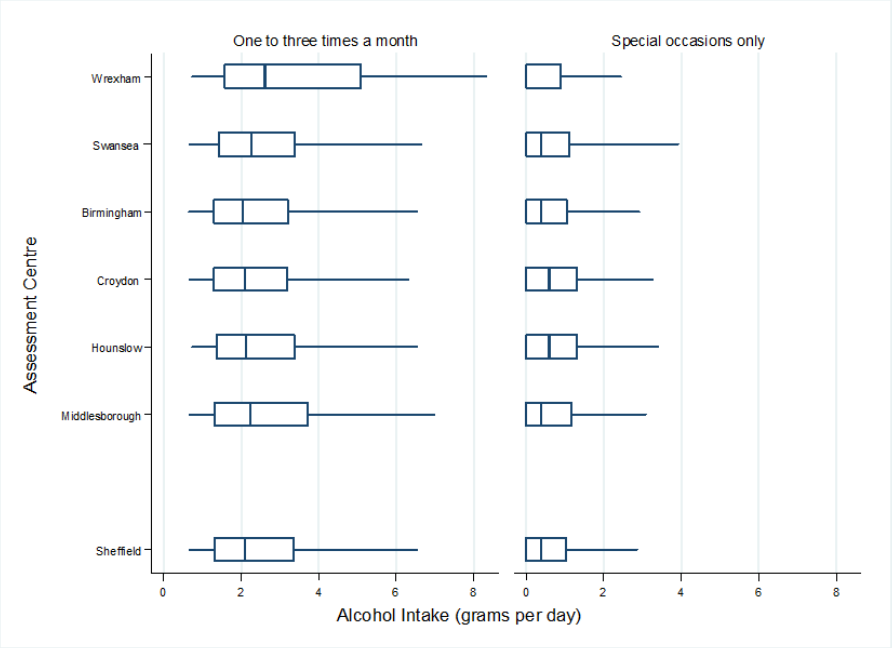
Figure 4.6.10 Alcohol Intake Distribution by Alcohol Frequency and Assessment Centre for Participants Drinking at Least "Once or Twice a Week"



Boxes show 25th, 50th and 75th percentiles of alcohol intake. Lines show the 5th and 95th percentiles of alcohol intake. Blue indicates that the assessment centre recruited participants after the change in questionnaire. Red indicates assessment centres which recruited participants before the change in questionnaire. (Bristol, Nottingham and Liverpool changed questionnaire during recruitment).

There were slight differences by assessment centre. For example, within each category of alcohol frequency, participants from Oxford reported the lowest alcohol intake. The assessment centres are sorted approximately by assessment dates; hence, there seemed to be no trend in the amount of alcohol intake reported by assessment date, within alcohol frequency. Also, there seemed to be less variability between assessment centres with lower alcohol frequency. Figure 4.6.11 shows the distribution of alcohol grams per day by assessment centre for participants who reported drinking "one to three times a month" or "special occasions only".

Figure 4.6.11 Alcohol Intake Distribution by Alcohol Frequency and Assessment Centre Less Than "One to Three Times a Month"



Boxes show 25th, 50th and 75th percentiles of alcohol intake. Lines show the 5th and 95th percentiles of alcohol intake.

Multiple Imputation Method

Multiple imputation was used in this thesis to impute average alcohol intake (grams/day) for participants with missing data. A number of participants were excluded from other analysis variables as well as alcohol intake because participants responded "do not know" or "prefer not to answer". Data for these variables could also have been imputed using multiple imputation by chained equations (MICE).[375] However, since participants chose not to answer, it is plausible that these data would not be MAR and so multiple imputation of these data could lead to biased results. Similarly, alcohol intake was imputed only for participants with missing data and not for participants who did answer the questionnaire but were excluded e.g. for responding "do not know".

Multiple imputation can be considered in three stages.[375] The first stage is to create m imputed datasets based on an imputation model which predicts missing values of a variable z using data from a set of variables $x_1, \ldots, x_n$, which are complete.

This analysis used predictive mean matching (PMM) to impute missing values of alcohol grams per day. PMM is commonly used for the imputation of continuous variables when the Normality assumption is not met.[375, 404] PMM begins by fitting a linear regression model of z on $x_1, \dots, x_n$ using the observed data in order to obtain predictions for each missing value of z. Then, for each individual with missing z, PMM identifies the k individuals with the closest observed value of z and one of these k values is chosen at random to replace the missing value. Hence, the distribution of imputed values should closely match the distribution of the observed values, conditional on the values of the other variables in the imputation model.

The default value for k in Stata was one. However, this would result in the same individual being selected for each imputation. This means there would be no uncertainty in the imputed values which would be equivalent to single imputation. Hence, k was set to ten in this analysis.[405]

In the second stage, each of these m datasets is analysed, individually but using identical methods, to obtain m estimates for the quantities of interest. The final stage is to combine these m estimates using Rubin's rules.[406] If Q_i represents the estimate for the quantity of interest from the i^{th} imputed dataset and W_i is the estimated variance of Q_i, then the combined estimate Q is

$$Q = \frac{1}{m}\sum_{i=1}^{m} Q_i$$

and the total variance of Q is

$$Var(Q) = W + \left(1 + \frac{1}{m}\right)B$$

where W is the within-imputation variance

$$W = \frac{1}{m}\sum_{i=1}^{m} W_i$$

and B is the between-imputation variance

$$B = \frac{1}{m-1}\sum_{i=1}^{m}(Q_i - Q)^2 .$$

Hence, the variation within and between imputations is included in the analysis.

Multiple Imputation Model

It is vital for valid imputation that the outcome is included in the imputation model.[407] Otherwise, there will be no association between alcohol intake and colorectal cancer among individuals with imputed data and so results using multiple imputation would be biased towards the null. Similarly, it is crucial that all confounder variables considered in analyses using imputed data are included in the imputation model.

Hence, the imputation model used in this analysis included all the confounders included in the analysis model for alcohol intake (sex, overall smoking status, BMI, Townsend deprivation index, red meat intake, processed meat intake, height, family history of colorectal cancer and bowel screening) as well as age at baseline. The question on alcohol intake ten years ago was also included. For the outcome, it is recommended to include the event indicator variable and the Nelson-Aalen estimator of the cumulative hazard when using Cox proportional hazards models.[408] Since separate analyses were performed for colon and rectal cancer, outcome variables for colon and rectal cancer were included in the imputation model.

Since alcohol grams per day was strongly associated with alcohol frequency, missing values were imputed by alcohol frequency i.e. separate imputation models were used for participants who reported drinking "special occasions only" and for participants who reported drinking "one to three times a month".

Alcohol data for participants who reported "special occasions only" and "one to three times a month" were practically either completely observed or completely missing according to assessment date/assessment centre (see Figure 4.1.2). Thus, it was not possible to estimate the effect of assessment date/assessment centre on alcohol grams per day for the earlier centres with missing data. Hence, these variables were not included in the imputation model despite them being highly predictive of the missing alcohol data. Therefore, this analysis assumed no relationship between assessment date or assessment centre and alcohol intake (within alcohol frequency). Although alcohol grams per day did differ by assessment centre (see Figure 4.6.10), this variation is more

likely explained by differences in other lifestyle factors rather than differences in geographic location. Also, the variation was smaller among people with lower alcohol frequency.

Table 4.6.1 Comparison of Variables for Participants with Non-missing and Missing Alcohol Data

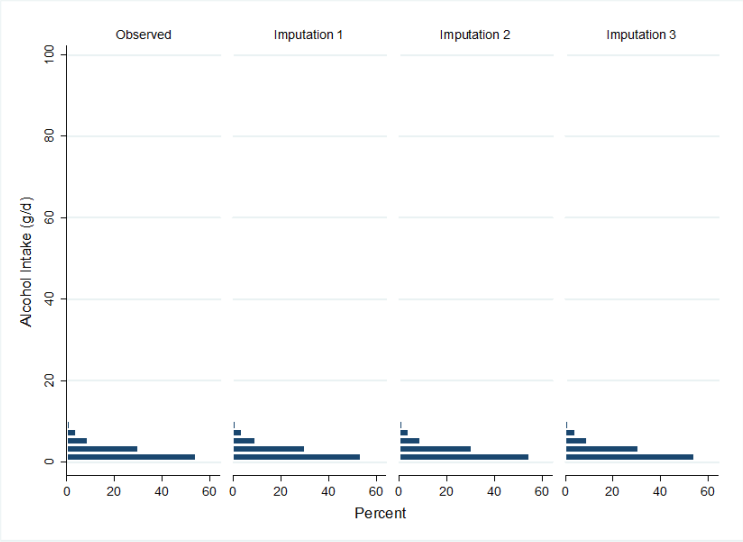
	One to three times a month		Special occasions only	
	Non-missing (n = 16,401)	**Missing (n = 31,653)**	**Non-missing (n = 17,170)**	**Missing (n = 30,647)**
Age (mean), yrs	55.3	55.1	56.7	56.7
Sex, % male	37.1	36.3	29.4	28.5
Current smoker, %	9.9	10.5	10.5	12.4
BMI, mean, kg/m^2	28.0	28.0	28.5	28.5
Height (mean), cm	167.6	167.2	165.6	165.3
Townsend index, % in most deprived quintile	21.7	20.0	28.3	26.3
Red meat intake, % ≥3 times/week	17.6	19.4	18.4	20.2
Processed meat intake % >once/week	28.8	27.6	26.4	25.9
Family history, %	10.4	10.8	10.7	11.0
Bowel screening, %	36.6	22.7	40.5	25.4
Alcohol intake ten years ago, % less nowadays	63.6	59.5	61.8	57.5
Colorectal cancer cases/10,000 person-years	7.5	8.4	7.6	8.8

Altogether, 33,571 participants had complete data for alcohol grams per day as well as all variables included in the imputation model. 62,300 participants had complete data for all variables included in the imputation model but had missing data for alcohol intake. Table 4.6.1 compares the characteristics of participants with non-missing and missing alcohol data. There were slight differences for some variables. For example, participants with missing alcohol data were slightly more likely to be current smokers, less likely to live in areas associated with highest deprivation and more likely to eat red meat at least 3 times a week. There was a much higher proportion of participants with a history of bowel screening among people with non-missing data than among people with missing data since national screening programmes were introduced during UK Biobank recruitment.[15-17] Thus, many participants who attended earlier assessment centres would not have been invited to attend screening until after recruitment. Participants with missing alcohol data had a slightly higher number of cases per 10,000

person-years because these participants were recruited earlier and so followed-up until an older age.

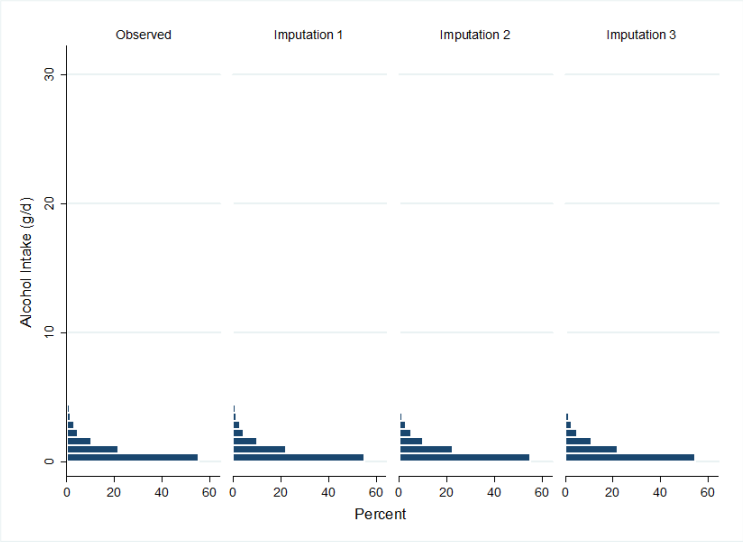
The number of imputations had to be decided for the analysis. There are no agreed guidelines for the number of imputations required. One rule of thumb is that the number of imputations should be at least equal to the percentage of observations with missing data.[375] 64.0% of participants who reported drinking "one to three times a month" and 63.3% of participants who reported drinking "special occasions only" had missing alcohol intake data. Hence, this analysis created 70 imputations (the effect of the choice for the number of imputations on the final results was investigated in the analysis (see section **Error! Reference source not found.**)). Figure 4.6.12 and Figure 4.6.13 compare the distribution of alcohol grams per day for participants with complete alcohol data with the distribution of the alcohol grams per day data in the first three imputations for participants with missing data. The distribution of alcohol grams per day in the imputations closely matched the distribution of alcohol grams per day in the observed data (for both "one to three times a month" and "special occasions only").

Figure 4.6.12 Distribution of Observed and Imputed Data for Alcohol Intake for "One to Three Times a Month"



Distribution of alcohol grams per day among participants who reported drinking "one to three times a month" with observed data and distribution of alcohol grams per day as calculated in the first three imputations for participants with missing data.

Figure 4.6.13 Distribution of Observed and Imputed Data for Alcohol Intake for "Special Occasions Only"



Distribution of alcohol grams per day among participants who reported drinking "special occasions only" with observed data and distribution of alcohol grams per day as calculated in the first three imputations for participants with missing data.

Chapters 5, 6 and 7 present the results for the analyses of alcohol intake, adiposity and smoking in relation to colorectal cancer. Separate discussion sections for the three risk factors are included in each chapter. An overall discussion of the thesis is included in Chapter 8.

Analyses in this thesis were complete-case analyses (except for the sensitivity analyses in Chapter 5 using multiple imputation), excluding participants with missing data or who responded "do not know" or "prefer not to answer". A disadvantage of this approach is that it leads to the exclusion of many participants. An alternative approach is to include an additional "dummy" category for each variable for those participants who could not be included in the other categories. The main analyses for alcohol intake, BMI and overall smoking status were also performed using this alternative approach and results are shown in Table A-5 - Table A-7 in the appendix (compare results with Table 5.2.2, Table 6.1.2, Table 6.1.3 and Table 7.1.2). Results for alcohol intake and smoking were very similar. For BMI, there was a slight difference in results, mainly for colon cancer for men; the HR for the highest quintile of BMI was 1.61 in the main analysis and 1.53 when including "dummy" categories.

Chapter 5 Alcohol Intake and Colorectal Cancer

5.1 Data Summary

Before presenting the results for the analysis of alcohol intake and colorectal cancer, summary tables are presented to show the relationships between the main analysis variables and colorectal cancer by sex (Table 5.1.1 - Table 5.1.11). These tables show the number of cases per 10,000 person-years (C/10KPY) and therefore provide an absolute measure of risk for these variables. Note that these results are not adjusted for other variables.

The absolute risk increased with each category of intake, from 7.2 C/10KPY for never drinkers to 16.8 C/10KPY for ≥45 g/d. However, there was no such trend for women where never drinkers had a higher absolute risk (8.9 C/10KPY) compared to all categories of current drinking.

The absolute risk increased monotonically with increasing BMI for men, from 9.4 C/10KPY for the lowest quintile to 14.5 C/10KPY for the highest quintile. Women in the lowest quintile of BMI did have the lowest risk (6.1 C/10KPY) but there was no clear trend in risk since women in the highest quintile had the next lowest level of risk (7.2 C/10KPY). The absolute risk for men increased from 7.2 in the lowest quintile to 16.9 C/10KPY in the highest quintile for WC and from 7.0 to 17.2 C/10KPY for WHR for men. For women, the increase was from 5.6 to 8.7 C/10KPY for WC and from 5.8 to 9.5 C/10KPY for WHR.

Former smokers had the highest absolute risk for men (15.9 C/10KPY) followed by current smokers (11.7 C/10KPY) and then never smokers (9.3 C/10KPY). For women, the absolute risks for never smokers, former smokers and current smokers were 7.1, 8.9 and 6.4 C/10KPY respectively.

Men who consumed more red meat and processed had a higher absolute risk of colorectal cancer. Men and women with a family history of bowel cancer and a history of bowel screening had a higher absolute risk.

Also, Table 5.1.12 shows results for the association between the confounder variables used in the analyses and colorectal cancer risk. There was no evidence for an association between Townsend deprivation index and colorectal cancer. Men with a high intake of red meat and processed meat had an increased risk of colorectal cancer but evidence was less clear for women. Height was positively associated with colorectal cancer risk. Participants with a family history of colorectal cancer had an increased risk of colorectal cancer and participants with a history of bowel screening had a lower risk.

Table 5.1.1 Colorectal Cancer Cases per 10,000 Person-years by Alcohol Intake and Sex in UK Biobank

	Alcohol intake (grams/day)						
	Former	**Never**	**<5**	**5-<15**	**15-<30**	**30-<45**	**≥45**
Men							
Person-years	37,757	30,641	71,949	227,317	278,444	153,671	148,995
Cases	36	22	63	243	341	219	250
Cases/10,000 person-years	9.5	7.2	8.8	10.7	12.2	14.3	16.8
Women							
Person-years	45,911	74,383	196,993	390,551	221,231	56,102	22,393
Cases	44	66	135	281	160	48	10
Cases/10,000 person-years	9.6	8.9	6.9	7.2	7.2	8.6	4.5

Table 5.1.2 Colorectal Cancer Cases per 10,000 Person-years by BMI and Sex in UK Biobank

	BMI (kg/m^2)				
	<24.45	**24.45-<26.40**	**26.40-<28.28**	**28.28-<30.84**	**≥30.84**
Men					
Person-years	217,112	218,711	209,047	226,974	217,554
Cases	205	213	268	321	316
Cases/10,000 person-years	9.4	9.7	12.8	14.1	14.5
	<22.89	**22.89-<25.04**	**25.04-<27.34**	**27.34-<30.80**	**≥30.80**
Women					
Person-years	258,203	256,816	254,415	257,517	255,599
Cases	157	197	216	211	183
Cases/10,000 person-years	6.1	7.7	8.5	8.2	7.2

Table 5.1.3 Colorectal Cancer Cases per 10,000 Person-years by WC and Sex in UK Biobank

	WC (cm)				
	≤87	**88-93**	**94-98**	**99-105**	**≥106**
Men					
Person-years	209,426	231,055	211,952	222,144	215,783
Cases	151	239	270	299	364
Cases/10,000 person-years	7.2	10.3	12.7	13.5	16.9
	≤74	**75-80**	**81-86**	**87-95**	**≥96**
Women					
Person-years	282,172	256,826	243,626	263,752	236,394
Cases	158	196	198	206	206
Cases/10,000 person-years	5.6	7.6	8.1	7.8	8.7

Table 5.1.4 Colorectal Cancer Cases per 10,000 Person-years by WHR and Sex in UK Biobank

	WHR				
	<0.883	**0.883-<0.918**	**0.918-<0.950**	**0.950-<0.990**	**≥0.990**
Men					
Person-years	222,937	216,822	221,726	217,534	210,999
Cases	156	193	283	329	362
Cases/10,000 person-years	7.0	8.9	12.8	15.1	17.2
	<0.758	**0.758-<0.796**	**0.796-<0.832**	**0.832-<0.876**	**≥0.876**
Women					
Person-years	265,802	257,807	256,930	247,422	254,583
Cases	154	166	218	183	243
Cases/10,000 person-years	5.8	6.4	8.5	7.4	9.5

Table 5.1.5 Colorectal Cancer Cases per 10,000 Person-years by Overall Smoking Status and Sex in UK Biobank

	Overall smoking status		
	Never smokers	**Former smokers**	**Current smokers**
Men			
Person-years	524,453	418,660	135,880
Cases	489	667	159
Cases/10,000 person-years	9.3	15.9	11.7
Women			
Person-years	777,234	372,227	113,734
Cases	548	332	73
Cases/10,000 person-years	7.1	8.9	6.4

Table 5.1.6 Colorectal Cancer Cases per 10,000 Person-years by Townsend Deprivation Index and Sex in UK Biobank

	Townsend deprivation index				
	<−3.93	**−3.93-<2.76**	**−2.76-<−1.30**	**−1.30-<1.35**	**≥1.35**
Men					
Person-years	224,067	217,964	215,265	214,239	224,176
Cases	278	291	255	246	262
Cases/10,000 person-years	12.4	13.4	11.8	11.5	11.7
Women					
Person-years	260,540	257,497	261,358	258,637	249,935
Cases	207	209	180	196	175
Cases/10,000 person-years	7.9	8.1	6.9	7.6	7.0

Table 5.1.7 Colorectal Cancer Cases per 10,000 Person-years by Red Meat Intake and Sex in UK Biobank

	Red meat intake		
	≤once/week	**>1-<3 times/week**	**≥3 times/week**
Men			
Person-years	450,877	350,956	278,090
Cases	458	451	406
Cases/10,000 person-years	10.2	12.9	14.6
Women			
Person-years	628,005	394,641	250,789
Cases	431	337	187
Cases/10,000 person-years	6.9	8.5	7.5

Table 5.1.8 Colorectal Cancer Cases per 10,000 Person-years by Processed Meat Intake and Sex in UK Biobank

	Processed meat intake		
	<once/week	**once/week**	**>once/week**
Men			
Person-years	293,115	326,031	472,804
Cases	298	414	615
Cases/10,000 person-years	10.2	12.7	13.0
Women			
Person-years	652,199	368,607	263,558
Cases	488	291	185
Cases/10,000 person-years	7.5	7.9	7.0

Table 5.1.9 Colorectal Cancer Cases per 10,000 Person-years by Height and Sex in UK Biobank

	Height (cm)				
	≤170	**171-174**	**175-177**	**178-181**	**≥182**
Men					
Person-years	244,989	232,454	189,344	211,185	211,423
Cases	347	256	241	246	233
Cases/10,000 person-years	14.2	11.0	12.7	11.6	11.0
	≤157	**158-161**	**162-164**	**165-168**	**≥169**
Women					
Person-years	278,345	290,895	242,209	258,815	212,285
Cases	212	223	175	201	153
Cases/10,000 person-years	7.6	7.7	7.2	7.8	7.2

Table 5.1.10 Colorectal Cancer Cases per 10,000 Person-years by Family History of Bowel Cancer and Sex in UK Biobank

	Family history of bowel cancer	
	No	**Yes**
Men		
Person-years	951,540	120,555
Cases	1,108	198
Cases/10,000 person-years	11.6	16.4
Women		
Person-years	1,126,331	136,034
Cases	814	127
Cases/10,000 person-years	7.2	9.3

Table 5.1.11 Colorectal Cancer Cases per 10,000 Person-years by Bowel Screening and Sex in UK Biobank

	Bowel screening	
	No	**Yes**
Men		
Person-years	748,656	319,605
Cases	831	479
Cases/10,000 person-years	11.1	15.0
Women		
Person-years	918,481	353,633
Cases	647	305
Cases/10,000 person-years	7.0	8.6

Table 5.1.12 Confounder Variables and the Risk of Colorectal Cancer in UK Biobank

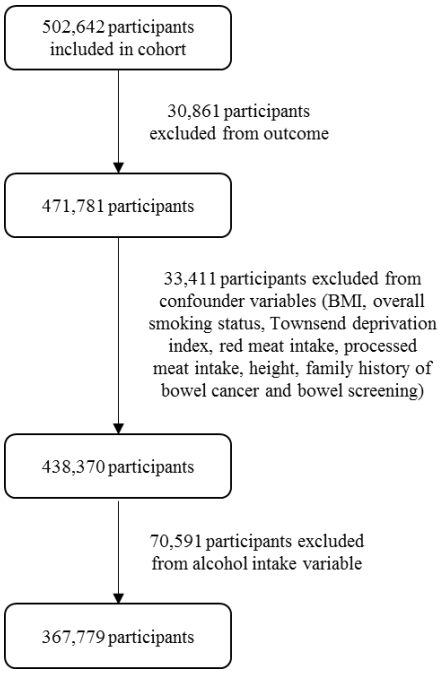
	Overall		Men		Women	
	Cases	**HR (95% CI)***	**Cases**	**HR (95% CI)***	**Cases**	**HR (95% CI)***
Sex						
Women	886	1.00				
Men	1,223	1.19 (1.04-1.36)				
Townsend index						
Q1	455	1.00	260	1.00	195	1.00
Q2	468	1.06 (0.93-1.20)	272	1.09 (0.92-1.29)	196	1.02 (0.84-1.24)
Q3	408	0.95 (0.83-1.09)	243	1.01 (0.85-1.21)	165	0.87 (0.70-1.07)
Q4	405	1.01 (0.88-1.16)	224	1.01 (0.85-1.21)	181	1.01 (0.82-1.23)
Q5	373	1.01 (0.88-1.16)	224	1.08 (0.90-1.30)	149	0.92 (0.74-1.15)
Red meat intake						
≤once/week	821	1.00	423	1.00	398	1.00
>1-<3 times/week	736	1.15 (1.04-1.27)	421	1.12 (0.97-1.28)	315	1.20 (1.03-1.39)
≥3 times/week	552	1.14 (1.02-1.27)	379	1.20 (1.04-1.39)	173	1.02 (0.85-1.22)
Processed meat intake						
<once/week	725	1.00	271	1.00	454	1.00
Once/week	646	1.08 (0.97-1.20)	381	1.17 (1.00-1.37)	265	1.03 (0.88-1.20)
>once/week	738	1.10 (0.99-1.23)	571	1.23 (1.06-1.43)	167	0.94 (0.78-1.12)
Height (per 10 cm)	2,109	1.11 (1.04-1.19)	1,223	1.08 (0.99-1.17)	886	1.16 (1.04-1.29)
Family history of colorectal cancer						
No	1,799	1.00	1,031	1.00	768	1.00
Yes	310	1.24 (1.10-1.40)	192	1.30 (1.11-1.52)	118	1.16 (0.95-1.41)
Bowel screening						
No	1,369	1.00	772	1.00	597	1.00
Yes	740	0.87 (0.79-0.95)	451	0.87 (0.77-0.98)	289	0.87 (0.75-1.01)

* Results adjusted for alcohol intake (never, former, special occasions only, one to three times a month, <5 g/d, 5-<15 g/d, 15-<30 g/d, 30-<45 g/d, ≥45 g/d), BMI (<25, 25-<30, ≥30 kg/m^2), overall smoking status (never, former, current).

5.2 Results

Cox proportional hazards models using age as the primary time variable were used to estimate the association between alcohol intake and colorectal cancer risk. Participants were followed-up from the date of the baseline assessment until the date of any cancer diagnosis, date of death or 31st March 2014, whichever came first. All analyses presented in this section were adjusted for sex, BMI, overall smoking status, Townsend deprivation index, red meat intake, processed meat intake, height, family history of colorectal cancer and bowel screening. This analysis included a total of 367,779 participants (177,075 men and 190,704 women) (Figure 5.2.1). 1,806 participants (1,108 men and 698 women) were diagnosed with colorectal cancer during a median follow-up of 5.07 years (range 3.50-8.05 years).

Figure 5.2.1 Flowchart of Exclusions for Analysis of Alcohol Intake and Colorectal Cancer



Flowchart of reasons for excluding participants from the analysis of alcohol intake and colorectal cancer.

5.2.1 Main Results

Table 5.2.1 compares characteristics of the participants included in the analysis by alcohol intake category and sex. Female never and former drinkers were slightly older than current drinkers and heavier drinkers were slightly younger than lighter drinkers.

Heavier drinkers as well as former drinkers were more likely to be current smokers. Never and former drinkers were more likely to live in areas with higher deprivation than current drinkers. Heavy drinkers reported the highest intakes of red meat and processed meat. Light drinkers had the highest rates of bowel screening for men and women. Male never drinkers had the lowest rates of screening whereas female heavy drinkers had the lowest rates.

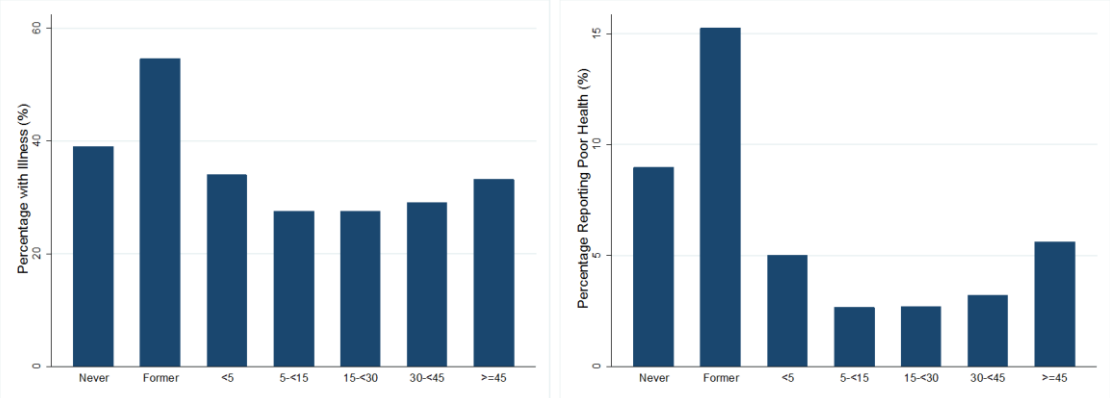
The main results for the association between alcohol intake and colorectal cancer are shown in Table 5.2.2. There was a noticeable difference between results adjusted for sex only and results adjusted for all confounder variables. Table A-8 in the appendix shows results adjusting for different confounder variables individually. Adjusting for smoking status resulted in the biggest attenuation in results.

Compared to never drinkers, there was only clear evidence of an increased risk of colorectal cancer for participants drinking ≥45 g/d (HR, 1.33; 95% CI, 1.02-1.74) (Table 5.2.2). The association between alcohol intake and colorectal cancer risk was stronger for men than women. For men, there appeared to be evidence of an increased risk associated with intake ≥15 g/d; the HRs (95% CIs) were 1.48 (0.93-2.37) for 15-<30 g/d, 1.61 (1.00-2.58) for 30-45 g/d and 1.86 (1.16-2.99) for ≥45 g/d. (CIs were wide due to a small number of cases in the reference group for men). In contrast, there was no clear evidence of an association for women.

P-values for test for trend were calculated by assigning to participants in each category of alcohol intake the median value within their category (former drinkers were excluded and never drinkers were included as 0) and modelling this variable as a continuous variable. The p-values for the test for trend were <0.0001 overall, <0.0001 for men and 0.7962 for women. The p-values for the test for trend were also calculated using data from the re-assessment. Instead of using the median value of baseline alcohol intake for each category, participants were assigned the median value of alcohol intake at re-assessment (among those in the baseline category who were re-assessed). This analysis excluded baseline never drinkers and former drinkers. The p-values for the test for trend were <0.0001 overall, <0.0001 for men and 0.7608 for women. Hence, results for test for trend were similar when taking repeated measures into account.

There was no clear evidence of an increased risk of colorectal cancer for former drinkers compared to never drinkers (HR, 0.97; 95% CI, 0.70-1.34). Some former drinkers will have quit drinking as a result of ill health and it is plausible that these former drinkers will have an increased risk of colorectal cancer. In UK Biobank, former drinkers were most likely to report a long-standing illness or poor health status at baseline (Figure 5.2.2).

Figure 5.2.2 Percentage of Participants Reporting Long-standing Illness and Poor Health, by Alcohol Intake



Percentage of participants reporting a long-standing illness and poor health by categories of alcohol intake.

In the touchscreen questionnaire, former drinkers were asked why they stopped drinking alcohol with the choice of five responses. “Illness or ill health”, “doctor’s advice” and “health precaution” were classified as illness related reasons and “financial reasons” and “other reasons” were classified as other reasons. Compared to never drinkers, the HR (95% CI) for colorectal cancer risk was 1.00 (0.67-1.48) for former drinkers who reported illness related reasons and 0.86 (0.56-1.32) for former drinkers who reported other reasons.

Table 5.2.1 Characteristics of UK Biobank Cohort by Alcohol Intake and Sex

	Alcohol intake (grams/day)						
	Former	**Never**	**<5**	**5-<15**	**15-<30**	**30-<45**	**≥45**
Men							
Number of participants	6,713	5,180	14,942	42,628	51,640	28,414	27,558
Age, mean (SD), years	56.8 (8.1)	55.8 (9.0)	56.9 (8.4)	56.8 (8.2)	56.7 (8.1)	56.7 (8.0)	56.5 (7.8)
Overall smoking status, % current smoker	17.4	7.9	9.4	7.2	9.6	13.5	21.9
BMI, mean (SD), kg/m^2	28.3 (5.1)	27.8 (4.8)	27.7 (4.6)	27.4 (4.1)	27.6 (3.9)	27.9 (3.9)	28.3 (4.1)
Townsend index, % in most deprived quintile	37.1	35.2	21.3	14.9	15.8	17.3	22.9
Red meat intake, % ≥3 times/week	24.9	28.0	19.9	21.5	24.7	28.2	34.8
Processed meat intake, % >once/week	38.2	35.3	39.0	39.2	42.2	47.0	54.5
Family history, %	11.5	9.1	10.7	11.1	11.3	12.1	12.3
Colorectal screening, %	32.5	26.5	38.0	32.4	32.5	32.7	31.0
Women							
Number of participants	8,381	13,322	40,556	72,672	41,167	10,456	4,150
Age, mean (SD), years	57.1 (7.8)	57.4 (8.3)	56.5 (8.1)	56.1 (7.9)	55.5 (7.8)	55.2 (7.6)	54.1 (7.6)
Overall smoking status, % current smoker	13.8	5.4	6.6	6.3	10.0	16.5	25.3
BMI, mean (SD), kg/m^2	28.2 (6.1)	28.2 (5.9)	27.4 (5.4)	26.3 (4.6)	26.3 (4.5)	26.7 (4.6)	27.2 (4.9)
Townsend index, % in most deprived quintile	31.1	29.1	19.5	14.5	16.2	18.6	23.3
Red meat intake, % ≥3 times/week	17.2	19.6	17.1	19.3	21.9	24.4	28.1
Processed meat intake, % >once/week	18.7	18.7	20.7	19.6	21.4	24.3	26.5
Family history, %	11.7	10.1	10.5	10.9	11.2	10.9	11.7
Colorectal screening, %	33.7	30.1	34.7	28.9	28.0	27.7	26.0

Table 5.2.2 Alcohol Intake and the Risk of Colorectal Cancer in UK Biobank

	Alcohol intake (grams/day)							
	Former	**Never**	**<5**	**5-<15**	**15-<30**	**30-<45**	**≥45**	**P-trend§**
Overall								
Person-years	75,602	92,950	251,412	582,471	471,550	197,596	160,697	
Cases	70	79	187	499	479	248	244	
Sex adjusted HR (95% CI)	1.04 (0.76-1.44)	1.00	0.92 (0.71-1.20)	1.02 (0.81-1.30)	1.15 (0.90-1.46)	1.33 (1.03-1.72)	1.56 (1.20-2.02)	<0.0001
Multivariate HR (95% CI)*	0.97 (0.70-1.34)	1.00	0.89 (0.68-1.16)	0.96 (0.75-1.22)	1.04 (0.81-1.33)	1.16 (0.89-1.51)	1.33 (1.02-1.74)	<0.0001
Men								
Person-years	33,393	26,099	66,413	212,948	261,178	144,277	139,594	
Cases	30	19	59	234	328	204	234	
Multivariate HR (95% CI)*	1.07 (0.60-1.90)	1.00	1.13 (0.67-1.90)	1.35 (0.84-2.15)	1.48 (0.93-2.37)	1.61 (1.00-2.58)	1.86 (1.16-2.99)	<0.0001
Women								
Person-years	42,210	66,851	184,999	369,523	210,373	53,319	21,103	
Cases	40	60	128	265	151	44	10	
Multivariate HR (95% CI)*	1.04 (0.69-1.55)	1.00	0.80 (0.59-1.09)	0.81 (0.61-1.08)	0.82 (0.60-1.12)	0.97 (0.65-1.44)	0.59 (0.30-1.17)	0.7962

*Adjusted for sex, BMI (<25, 25-<30, ≥30 kg/m^2), overall smoking status (never, former, current), Townsend deprivation index (quintiles), red meat intake (≤1, >1-<3, ≥3 times/week), processed meat (<1, 1, >1 time/week), height (continuous), family history of colorectal cancer (yes/no) and history of colorectal screening (yes/no).

§ P-values for test for trend were calculated by assigning participants the median value of their category of alcohol intake and this variable was modelled as a continuous variable (former drinkers were excluded for this analysis).

Participants who reported drinking alcohol at baseline were asked how their alcohol intake had changed in comparison to ten years ago. Table 5.2.3 shows the responses of participants by category of alcohol intake. For each category, a similar proportion of participants reported drinking "about the same" though the proportion reporting "less nowadays" decreased with increasing alcohol intake. Compared to the overall analysis, restricting the analysis to only those participants who reported drinking "about the same" as ten years ago, the HRs for each category of alcohol intake were similar except for ≥45 g/d which increased; the HRs (95% CIs) associated with drinking <5, 5-<15, 15-<30, 30-<45 and ≥45 g/d were 0.90 (0.66-1.24), 0.92 (0.70-1.20), 0.96 (0.72-1.27), 1.06 (0.76-1.46) and 1.51 (1.09-2.07).

Table 5.2.3 Alcohol Intake Compared to Ten Years before Baseline

Alcohol intake vs ten years ago	Alcohol intake (grams/day)									
	<5		5-<15		15-<30		30-<45		≥45	
	n	%	n	%	n	%	n	%	n	%
Less nowadays	30,868	55.6	51,643	44.8	36,107	38.9	13,514	34.8	9,536	30.1
About the same	21,203	38.2	44,073	38.2	34,065	36.7	14,460	37.2	12,526	39.5
More nowadays	2,882	5.2	19,111	16.6	22,220	23.9	10,774	27.7	9,572	30.2
Do not know	458	0.8	461	0.4	408	0.4	118	0.3	72	0.2
Prefer not to answer	87	0.2	12	0.0	7	0.0	4	0.0	2	0.0

Participants who reported drinking alcohol at baseline were also asked "when you drink alcohol is it usually with meals?" Participants responded "yes", "it varies" or "no". For the ease of writing, participants who responded "yes" will be referred to as "meal drinkers", participants who responded "it varies" as "variable drinkers" and participants who responded "no" as "non-meal drinkers".

The risk of colorectal cancer was compared between the three categories. Since average alcohol intake differed between the categories (mean alcohol intake was 16.0, 23.3 and 26.4 g/d for "meal drinkers", "variable drinkers" and "non-meal drinkers"), this analysis was adjusted for continuous alcohol intake. Compared with "meal drinkers", the HR (95% CI) was 0.95 (0.85-1.06) for "variable drinkers" and 0.96 (0.84-1.10) for "non-meal drinkers".

Results for men also showed no evidence for differences by drinking pattern (HR, 1.00; 95% CI, 0.86-1.15 for "variable drinkers" and HR, 1.00; 95% CI, 0.85-1.18 for "non-

meal drinkers") but results for women suggested that "variable drinkers" and "non-meal drinkers" may have a lower risk of colorectal cancer than "meal drinkers" (HR, 0.89; 95% CI, 0.74-1.07 for "variable drinkers" and HR, 0.87; 95% CI, 0.67-1.12 for "non-meal drinkers").

Colorectal Subsites

Investigating separately the risk of colon and rectal cancer, there was still no evidence that alcohol intake was associated with an increased risk for women (Table 5.2.4). In contrast, alcohol intake seemed to be strongly associated with colon cancer risk (HR, 1.85; 95% CI, 0.99-3.45 for ≥45 g/d) and rectal cancer risk (HR, 1.91; 95% CI, 0.92-3.96 for ≥45 g/d) for men (though CIs were very wide).

Associations with proximal and distal colon cancer were also investigated (Table 5.2.5). Participants drinking 30-≤45 g/d (HR, 1.66; 95% CI, 1.04-2.66) and ≥45 g/d (HR, 1.52; 95% CI, 0.94-2.48) had an increased risk of distal colon cancer but there was no clear evidence of an increased risk for proximal colon cancer.

Table 5.2.4 Alcohol Intake and the Risk of Colon Cancer and Rectal Cancer in UK Biobank

	Alcohol intake (grams/day)							
	Former	**Never**	**<5**	**5-<15**	**15-<30**	**30-<45**	**≥45**	**P-trend§**
Colon cancer								
Overall								
Cases	51	52	130	339	305	164	141	
HR (95% CI)*	1.12 (0.76-1.66)	1.00	0.96 (0.69-1.33)	1.04 (0.77-1.40)	1.09 (0.81-1.48)	1.32 (0.95-1.82)	1.34 (0.96-1.88)	0.0017
Men								
Cases	20	11	39	143	192	132	136	
HR (95% CI)*	1.22 (0.58-2.55)	1.00	1.29 (0.66-2.52)	1.41 (0.76-2.61)	1.48 (0.80-2.74)	1.77 (0.95-3.29)	1.85 (0.99-3.45)	0.0020
Women								
Cases	31	41	91	196	113	32	5	
HR (95% CI)*	1.22 (0.76-1.96)	1.00	0.87 (0.60-1.26)	0.93 (0.66-1.32)	0.98 (0.68-1.42)	1.14 (0.71-1.83)	0.48 (0.19-1.22)	0.9457
Rectal cancer								
Overall								
Cases	20	27	60	166	176	88	104	
HR (95% CI)*	0.74 (0.41-1.32)	1.00	0.81 (0.51-1.28)	0.86 (0.57-1.29)	0.95 (0.63-1.44)	0.98 (0.63-1.54)	1.30 (0.83-2.03)	0.0024
Men								
Cases	11	8	20	92	136	75	99	
HR (95% CI)*	0.95 (0.38-2.37)	1.00	0.91 (0.40-2.07)	1.27 (0.62-2.63)	1.49 (0.72-3.05)	1.43 (0.68-2.99)	1.91 (0.92-3.96)	0.0008
Women								
Cases	9	19	40	74	44	13	5	
HR (95% CI)*	0.66 (0.30-1.47)	1.00	0.72 (0.41-1.24)	0.60 (0.36-1.01)	0.55 (0.31-0.97)	0.71 (0.34-1.46)	0.74 (0.27-2.02)	0.5643

*Adjusted for sex, BMI (<25, 25-<30, ≥30 kg/m^2), overall smoking status (never, former, current), Townsend deprivation index (quintiles), red meat intake (≤1, >1-<3, ≥3 times/week), processed meat (<1, 1, >1 time/week), height (continuous), family history of colorectal cancer (yes/no) and history of colorectal screening (yes/no).

§ P-values for test for trend were calculated by assigning participants the median value of their category of alcohol intake and this variable was modelled as a continuous variable (former drinkers were excluded for this analysis).

Table 5.2.5 Alcohol Intake and the Risk of Proximal Colon Cancer and Distal Colon Cancer in UK Biobank

	Former	Never	<5	5-<15	15-<30	30-<45	≥45	P-trend§
Proximal colon cancer								
Cases	25	22	75	180	152	61	62	
HR (95% CI)*	1.29 (0.73-2.31)	1.00	1.28 (0.79-2.06)	1.30 (0.83-2.03)	1.28 (0.81-2.02)	1.13 (0.68-1.87)	1.36 (0.82-2.27)	0.7638
Distal colon cancer								
Cases	23	24	49	144	131	96	74	
HR (95% CI)*	1.09 (0.61-1.94)	1.00	0.80 (0.49-1.30)	0.95 (0.61-1.47)	1.01 (0.65-1.58)	1.66 (1.04-2.66)	1.52 (0.94-2.48)	<0.0001

*Adjusted for sex, BMI (<25, 25-<30, ≥30 kg/m^2), overall smoking status (never, former, current), Townsend deprivation index (quintiles), red meat intake (≤1, >1-<3, ≥3 times/week), processed meat (<1, 1, >1 time/week), height (continuous), family history of colorectal cancer (yes/no) and history of colorectal screening (yes/no).

§ P-values for test for trend were calculated by assigning participants the median value of their category of alcohol intake and this variable was modelled as a continuous variable (former drinkers were excluded for this analysis).

5.2.2 Continuous Alcohol Intake Variable

Alcohol intake was further analysed as a continuous variable. Fractional polynomials were used to investigate the "best-fitting" model for the association. A linear model was used to model alcohol intake since there was no evidence of a non-linear relationship (p-value for the likelihood ratio test comparing the "best-fitting" 2-degree polynomial with the linear model was 0.42). The HR and 95% CI associated with an increase of 10 g/d was 1.05 (1.03-1.07) (Table 5.2.6). (Results were presented per 10 g/d since 10 g of alcohol is approximately equal to a small glass of wine, half pint of beer or measure of spirits). Results were very similar for men but there was no evidence of an association for women.

To investigate the effect of including never drinkers in the continuous analysis, models were refitted excluding never drinkers from the continuous alcohol intake variable. There was no evidence for a non-linear relationship and the results for an increase of 10 g/d were very similar to the previous results (Table 5.2.6).

Table 5.2.6 Continuous Alcohol Intake and the Risk of Colorectal Cancer in UK Biobank

Alcohol intake (grams/day)	Overall	Men	Women
Overall			
Cases	1,713	1,065	648
HR (95% CI) per 10g/d*	1.05 (1.03-1.07)	1.05 (1.03-1.07)	0.98 (0.91-1.05)
Excluding never drinkers			
Cases	1,634	1,046	588
HR (95% CI) per 10g/d*	1.05 (1.03-1.07)	1.05 (1.03-1.07)	0.99 (0.92-1.06)
Excluding "extreme" values			
Cases	1,681	1,033	648
HR (95% CI) per 10g/d*	1.06 (1.03-1.09)	1.06 (1.03-1.10)	0.99 (0.92-1.06)

*Adjusted for sex, BMI (<25, 25-<30, ≥30 kg/m^2), overall smoking status (never, former, current), Townsend deprivation index (quintiles), red meat intake (≤1, >1-<3, ≥3 times/week), processed meat (<1, 1, >1 time/week), height (continuous), family history of colorectal cancer (yes/no) and history of colorectal screening (yes/no).

Since very large values of alcohol intake may impact on results (UK Biobank participants reported drinking up to 900 g/d), models were refitted again excluding participants with alcohol intake >96 g/d (the 99th percentile of alcohol intake among drinkers). Again, there was no evidence for a non-linear relationship and results were very similar (Table 5.2.6).

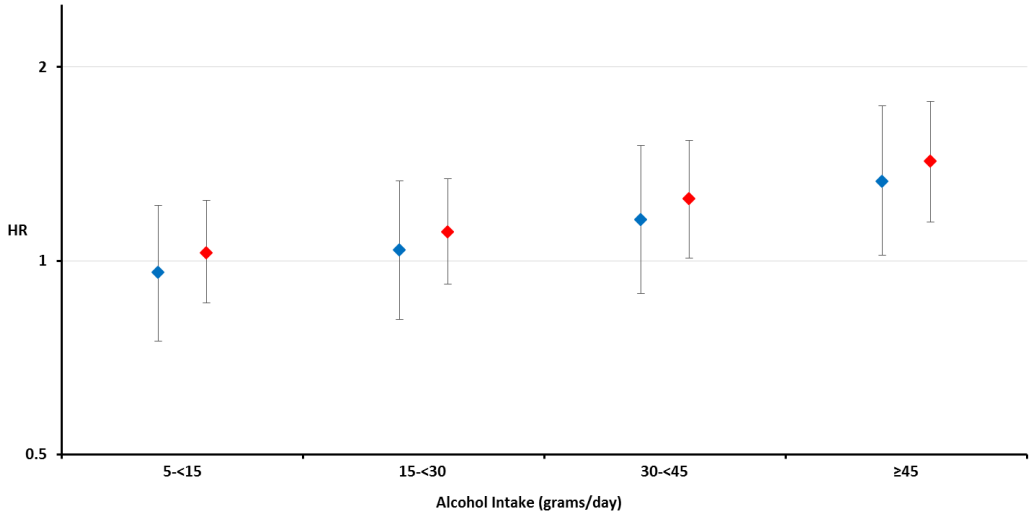
5.2.3 Alternative Reference Group

24,085 participants who reported drinking <5 g/d at baseline reported drinking "about the same" or "more nowadays" compared to ten years before baseline (Table 5.2.3). These participants were defined as consistent light drinkers and were combined with never drinkers to form an alternative reference group. Table 5.2.7 presents the HRs and 95% CIs for the association between alcohol intake and colorectal cancer using this alternative reference group. Participants who reported drinking <5 g/d at baseline and reported drinking "less nowadays" compared to ten years ago were classified separately (<5 (reduced)).

The HRs using never drinkers and consistent light drinkers as the reference group were slightly higher than those using only never drinkers (Figure 5.2.3). CIs were narrower due to the larger number of cases included in the new reference group. Consequently, in contrast to overall results using never drinkers as the reference group, overall results using the alternative reference group did show evidence of an increased risk for participants drinking 30-<45 g/d and slight evidence of an increased risk for participants drinking 15-<30 g/d.

HRs for men decreased slightly when including consistent light drinkers in the reference group, for example the HR associated with drinking 30-<45 g/d decreased from 1.61 to 1.54 and the HR associated with drinking ≥45 g/d decreased from 1.86 to 1.79. For women, there remained no evidence for an association.

Figure 5.2.3 Comparison of Association between Alcohol Intake and Colorectal Cancer Using Different Reference Groups



Blue diamonds represent HRs using never drinkers as the reference group. Red diamonds represent HRs using never drinkers and consistent light drinkers combined as the reference group. Error bars represent 95% CIs.

Table 5.2.8 presents the results for the association between alcohol intake and colon and rectal cancer using the alternative reference group. HRs for the association with colon cancer were very similar using the alternative reference group (Figure 5.2.4). However, there was a decrease in the HRs for men using the alternative reference group.

In contrast to the overall results using never drinkers as the reference group, there was evidence for an association between alcohol intake and rectal cancer risk when including consistent light drinkers in the reference group (Table 5.2.8 and Figure 5.2.5). The HRs for the association between alcohol intake and rectal cancer increased for men and thus there was evidence that drinking above 15 g/d is associated with an increased risk of rectal cancer compared to the alternative reference group.

Table 5.2.7 Alcohol Intake and the Risk of Colorectal Cancer in UK Biobank (Alternative Reference Group)

	Alcohol intake (grams/day)							
	Former	**Never / <5 (consistent)**	**<5 (reduced)**	**5-<15**	**15-<30**	**30-<45**	**≥45**	**P-trend§**
Overall								
Cases	70	160	103	499	479	248	244	
HR (95% CI)*	1.04 (0.78-1.38)	1.00	0.96 (0.75-1.23)	1.03 (0.86-1.24)	1.11 (0.92-1.34)	1.25 (1.01-1.54)	1.43 (1.15-1.77)	<0.0001
Men								
Cases	30	41	36	234	328	204	234	
HR (95% CI)*	1.03 (0.64-1.65)	1.00	1.10 (0.70-1.72)	1.29 (0.93-1.81)	1.43 (1.03-1.98)	1.54 (1.10-2.17)	1.79 (1.28-2.51)	<0.0001
Women								
Cases	40	119	67	265	151	44	10	
HR (95% CI)*	1.18 (0.82-1.69)	1.00	0.92 (0.68-1.25)	0.92 (0.74-1.15)	0.94 (0.73-1.20)	1.10 (0.77-1.57)	0.68 (0.35-1.30)	0.8463

*Adjusted for sex, BMI (<25, 25-<30, ≥30 kg/m^2), overall smoking status (never, former, current), Townsend deprivation index (quintiles), red meat intake (≤1, >1-<3, ≥3 times/week), processed meat (<1, 1, >1 time/week), height (continuous), family history of colorectal cancer (yes/no) and history of colorectal screening (yes/no).

§ P-values for test for trend were calculated by assigning participants the median value of their category of alcohol intake and this variable was modelled as a continuous variable (former drinkers were excluded for this analysis).

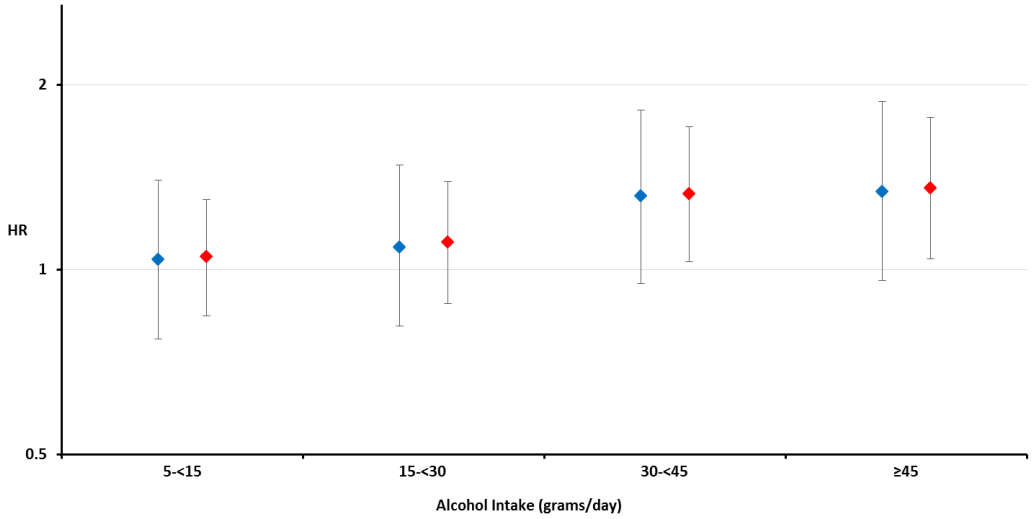
Table 5.2.8 Alcohol Intake and the Risk of Colon Cancer and Rectal Cancer in UK Biobank (Alternative Reference Group)

	Former	**Never / <5 (consistent)**	**<5 (reduced)**	**5-<15**	**15-<30**	**30-<45**	**≥45**	**P-trend§**
Colon cancer								
Overall								
Cases	51	111	68	339	305	164	141	
HR (95% CI)*	1.13 (0.81-1.59)	1.00	0.93 (0.69-1.26)	1.05 (0.84-1.30)	1.11 (0.88-1.39)	1.33 (1.03-1.71)	1.36 (1.04-1.77)	0.0015
Men								
Cases	20	27	22	143	192	132	136	
HR (95% CI)*	1.03 (0.57-1.84)	1.00	1.01 (0.57-1.78)	1.18 (0.78-1.79)	1.25 (0.83-1.87)	1.48 (0.98-2.26)	1.55 (1.02-2.37)	0.0019
Women								
Cases	31	84	46	196	113	32	5	
HR (95% CI)*	1.32 (0.87-1.99)	1.00	0.91 (0.64-1.31)	1.01 (0.78-1.31)	1.06 (0.79-1.42)	1.23 (0.81-1.86)	0.52 (0.21-1.28)	0.8881
Rectal cancer								
Overall								
Cases	20	50	37	166	176	88	104	
HR (95% CI)*	0.88 (0.52-1.48)	1.00	1.07 (0.70-1.65)	1.02 (0.74-1.41)	1.14 (0.82-1.57)	1.17 (0.81-1.69)	1.55 (1.08-2.23)	0.0024
Men								
Cases	11	14	14	92	136	75	99	
HR (95% CI)*	1.14 (0.51-2.52)	1.00	1.28 (0.61-2.68)	1.52 (0.86-2.67)	1.78 (1.02-3.09)	1.71 (0.96-3.05)	2.28 (1.29-4.04)	0.0008
Women								
Cases	9	36	23	74	40	13	5	
HR (95% CI)*	0.83 (0.40-1.74)	1.00	0.99 (0.59-1.68)	0.76 (0.50-1.14)	0.70 (0.44-1.11)	0.90 (0.47-1.72)	0.94 (0.36-2.43)	0.5660

*Adjusted for sex, BMI (<25, 25-<30, ≥30 kg/m^2), overall smoking status (never, former, current), Townsend deprivation index (quintiles), red meat intake (≤1, >1-<3, ≥3 times/week), processed meat (<1, 1, >1 time/week), height (continuous), family history of colorectal cancer (yes/no) and history of colorectal screening (yes/no).

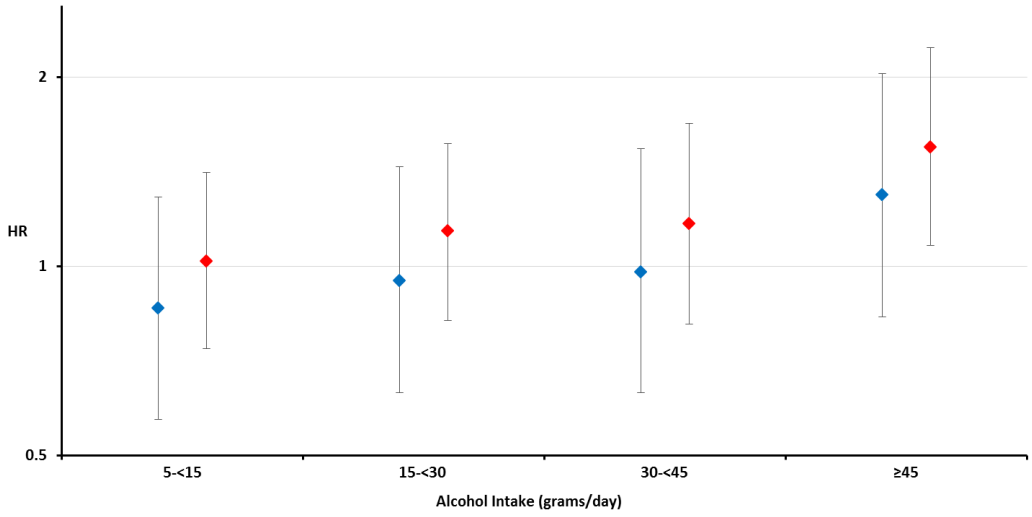
§ P-values for test for trend were calculated by assigning participants the median value of their category of alcohol intake and this variable was modelled as a continuous variable (former drinkers were excluded for this analysis).

Figure 5.2.4 Comparison of Association between Alcohol Intake and Colon Cancer Using Different Reference Groups



Blue diamonds represent HRs using never drinkers as the reference group. Red diamonds represent HRs using never drinkers and consistent light drinkers combined as the reference group. Error bars represent 95% CIs.

Figure 5.2.5 Comparison of Association between Alcohol Intake and Rectal Cancer Using Different Reference Groups



Blue diamonds represent HRs using never drinkers as the reference group. Red diamonds represent HRs using never drinkers and consistent light drinkers combined as the reference group. Error bars represent 95% CIs.

5.2.4 Effect Modifiers

Alcohol intake was modelled as a continuous variable when investigating the potential effect of other variables on the association between alcohol intake and colorectal cancer risk in order to maximise statistical power (former drinkers were not included).

Table 5.2.9 Alcohol Intake and the Risk of Colorectal Cancer in UK Biobank by Other Variables

	Alcohol intake (per 10 g/d)	**P-interaction**
BMI (kg/m^2)		
<25		
Cases	459	
HR (95% CI)*†	1.03 (0.98-1.08)	
25-<30		
Cases	815	
HR (95% CI)*†	1.05 (1.02-1.09)	
≥30		
Cases	439	
HR (95% CI)*†	1.05 (1.02-1.08)	0.5153
Smoking status		
Never		
Cases	763	
HR (95% CI)*‡	1.05 (1.02-1.08)	
Ever		
Cases	950	
HR (95% CI)*‡	1.04 (1.02-1.07)	0.9875
Folate intake (µg)		
<289.43		
Cases	285	
HR (95% CI)*†‡	1.03 (0.97-1.09)	
≥289.43		
Cases	318	
HR (95% CI)*†‡	1.02 (0.97-1.08)	0.6647

*Adjusted for sex, Townsend deprivation index (quintiles), red meat intake (≤1, >1-<3, ≥3 times/week), processed meat (<1, 1, >1 time/week), height (continuous), family history of colorectal cancer (yes/no) and history of colorectal screening (yes/no).

† Adjusted for overall smoking status (never, former, current).

‡ Adjusted for BMI (<25, 25-<30, ≥30 kg/m^2).

Results did not show clear evidence that the association between alcohol intake and colorectal cancer differed by BMI category (Table 5.2.9). The HR (95% CI) associated with a 10 g/d increase was 1.03 (0.98-1.08) for BMI <25 kg/m^2, 1.05 (1.02-1.09) for BMI 25-<30 kg/m^2 and 1.05 (1.02-1.08) for BMI ≥30 kg/m^2. There was also no evidence that the association differed for never smokers (HR, 1.05; 95% CI, 1.02-1.08) and for ever smokers (HR, 1.04; 95% CI, 1.02-1.07).

The association between alcohol intake and colorectal cancer was investigated according to average folate intake as measured by the online 24 hour recall dietary questionnaire. 167,941 participants included in this analysis also had data on average folate intake. Median folate intake among these participants was 289.4 µg. There was no clear evidence that the association between alcohol intake and colorectal cancer differed for participants with above or below median folate intake.

5.2.5 Multiple Imputation

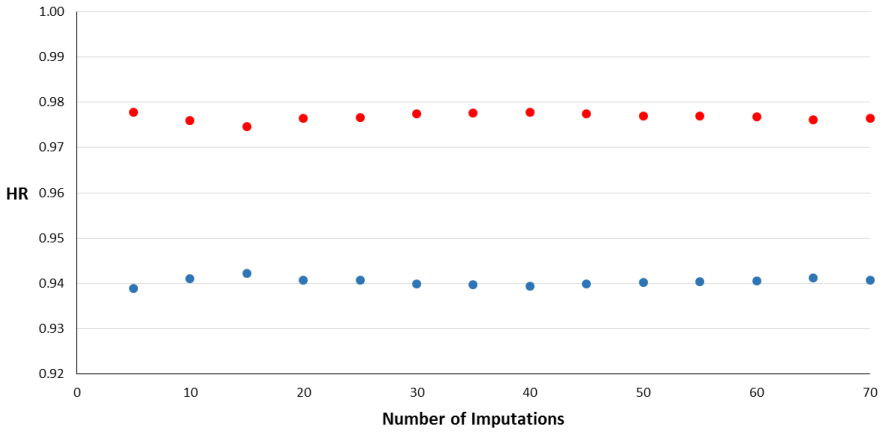
Multiple imputation was used to impute missing alcohol grams per day data for participants who reported drinking "one to three times a month" and "special occasions only". 16,472 participants who reported drinking "one to three times a month" had complete data for alcohol grams per day as well as all confounder variables. In contrast, 31,653 participants who reported drinking "one to three times a month" had complete data for all confounder variables but had missing data for alcohol grams per day. Of the participants who reported drinking "special occasions only" and who had complete data for all confounder variables, 17,557 had complete alcohol grams per day data and 30,647 participants had missing data.

Thus, alcohol grams per day data were imputed for those participants with missing data using multiple imputation. Missing data were imputed separately for participants who reported drinking "one to three times a month" and "special occasions only". Following multiple imputation, this analysis included 430,079 participants (197,289 men and 232,790 women) and 2,103 cases of colorectal cancer (1,221 cases for men and 882 cases for women).

Results for the association between alcohol intake and colorectal cancer, colon cancer and rectal cancer are shown in Table 5.2.10, for the entire cohort and men and women separately. Results were very similar to the results of the complete case analysis (Table 5.2.2 and Table 5.2.4). The most notable difference was the increase for the HR for colorectal cancer associated with <5 g/d for men from 1.13 to 1.20.

This analysis was based on 70 imputations. Figure 5.2.6 shows that there was very little change in the HRs for <5 and 5-<15 g/d when restricting the analysis to different numbers of imputations. Thus, results were unlikely to change with more imputations.

Figure 5.2.6 Association between Alcohol Intake and Colorectal Cancer by Number of Imputations



HR calculated using different numbers of imputations. Blue dots show the HR for the risk of colorectal cancer associated with <5 g/d and red dots show the HR associated with 5-<15 g/d as calculated based on different numbers of imputations.

It was also decided to repeat the analysis with the alternative reference group using the imputed data. All participants (including those with missing alcohol grams per day data) were asked how their alcohol drinking compared to ten years ago. As before, participants who drank <5 g/d and reported drinking "about the same" or "more nowadays" compared to ten years before baseline were combined with never drinkers to form the reference group. Results for colorectal cancer, colon cancer and rectal cancer are shown in Table 5.2.11.

HRs for colorectal cancer (overall and for men and women) were similar to HRs in the complete case analysis using the alternative reference group (Table 5.2.7). CIs were

narrower in this analysis due to more cases being included in the reference group. As a result, there was now slight evidence of an increased risk including for men drinking 5-<15 g/d (HR, 1.30; 95% CI, 0.98-1.71).

HRs were also similar to the complete case analysis results for colon and rectal cancer (Table 5.2.8). While alcohol intake was associated with colon cancer (HR, 1.64; 95% CI, 1.15-2.33 for ≥45 g/d) and rectal cancer risk (HR, 2.09; 95% CI, 1.33-3.29 for ≥45 g/d) for men, there remained no clear evidence of an association for women.

Table 5.2.10 Alcohol Intake and the Risk of Colorectal Cancer in UK Biobank using Multiple Imputation

	Alcohol intake (grams/day)							
HR (95% CI)*	**Former**	**Never**	**<5**	**5-<15**	**15-<30**	**30-<45**	**≥45**	**P-trend§**
Colorectal cancer								
Overall	0.98 (0.71-1.35)	1.00	0.94 (0.74-1.20)	0.98 (0.77-1.24)	1.05 (0.82-1.34)	1.19 (0.91-1.54)	1.36 (1.05-1.78)	<0.0001
Men	1.08 (0.61-1.92)	1.00	1.20 (0.74-1.95)	1.37 (0.86-2.20)	1.51 (0.95-2.40)	1.63 (1.02-2.63)	1.90 (1.18-3.05)	<0.0001
Women	1.04 (0.69-1.55)	1.00	0.85 (0.64-1.12)	0.81 (0.61-1.08)	0.82 (0.61-1.12)	0.97 (0.65-1.45)	0.60 (0.30-1.18)	0.6540
Colon cancer								
Overall	1.13 (0.76-1.66)	1.00	1.02 (0.76-1.37)	1.04 (0.77-1.40)	1.09 (0.81-1.48)	1.32 (0.96-1.82)	1.35 (0.97-1.89)	0.0015
Men	1.23 (0.59-2.57)	1.00	1.36 (0.73-2.54)	1.41 (0.76-2.62)	1.49 (0.81-2.75)	1.78 (0.96-3.31)	1.87 (1.00-3.47)	0.0012
Women	1.21 (0.76-1.94)	1.00	0.93 (0.66-1.30)	0.93 (0.66-1.31)	0.97 (0.67-1.40)	1.13 (0.70-1.81)	0.48 (0.19-1.21)	0.8661
Rectal cancer								
Overall	0.75 (0.42-1.34)	1.00	0.82 (0.54-1.25)	0.89 (0.59-1.35)	0.99 (0.65-1.50)	1.03 (0.66-1.60)	1.37 (0.88-2.13)	0.0004
Men	0.96 (0.39-2.40)	1.00	1.02 (0.48-2.17)	1.33 (0.64-2.75)	1.54 (0.75-3.15)	1.48 (0.71-3.09)	1.97 (0.95-4.09)	0.0001
Women	0.68 (0.31-1.51)	1.00	0.70 (0.42-1.16)	0.62 (0.37-1.04)	0.57 (0.32-1.00)	0.73 (0.36-1.52)	0.77 (0.28-2.10)	0.6292

*Adjusted for sex, BMI (<25, 25-<30, ≥30 kg/m^2), overall smoking status (never, former, current), Townsend deprivation index (quintiles), red meat intake (≤1, >1-<3, ≥3 times/week), processed meat (<1, 1, >1 time/week), height (continuous), family history of colorectal cancer (yes/no) and history of colorectal screening (yes/no).

§ P-values for test for trend were calculated by assigning participants the median value of their category of alcohol intake and this variable was modelled as a continuous variable (former drinkers were excluded for this analysis).

Table 5.2.11 Alcohol Intake and the Risk of Colorectal Cancer in UK Biobank (Alternative Reference Group) using Multiple Imputation

	Alcohol intake (grams/day)							
HR (95% CI)*	**Former**	**Never / <5 (consistent)**	**<5 (reduced)**	**5-<15**	**15-<30**	**30-<45**	**≥45**	**P-trend§**
Colorectal cancer								
Overall	0.99 (0.76-1.30)	1.00	0.93 (0.78-1.11)	1.00 (0.86-1.16)	1.07 (0.92-1.25)	1.21 (1.01-1.45)	1.39 (1.15-1.67)	<0.0001
Men	1.02 (0.66-1.58)	1.00	1.20 (0.88-1.65)	1.30 (0.98-1.71)	1.42 (1.09-1.86)	1.55 (1.17-2.04)	1.80 (1.36-2.37)	<0.0001
Women	1.11 (0.79-1.56)	1.00	0.84 (0.68-1.04)	0.87 (0.72-1.05)	0.88 (0.71-1.09)	1.04 (0.75-1.45)	0.64 (0.34-1.21)	0.6540
Colon cancer								
Overall	1.08 (0.79-1.47)	1.00	0.94 (0.77-1.16)	1.00 (0.83-1.20)	1.05 (0.87-1.27)	1.26 (1.01-1.58)	1.30 (1.02-1.64)	0.0014
Men	1.08 (0.63-1.84)	1.00	1.27 (0.86-1.88)	1.24 (0.88-1.76)	1.31 (0.93-1.83)	1.56 (1.10-2.22)	1.64 (1.15-2.33)	0.0012
Women	1.20 (0.81-1.77)	1.00	0.84 (0.66-1.08)	0.92 (0.74-1.15)	0.96 (0.75-1.24)	1.11 (0.75-1.64)	0.47 (0.19-1.15)	0.8778
Rectal cancer								
Overall	0.85 (0.52-1.39)	1.00	0.93 (0.68-1.27)	1.01 (0.77-1.33)	1.12 (0.85-1.47)	1.17 (0.85-1.60)	1.55 (1.14-2.12)	0.0004
Men	1.02 (0.50-2.09)	1.00	1.15 (0.67-1.96)	1.41 (0.89-2.23)	1.63 (1.05-2.52)	1.57 (0.99-2.50)	2.09 (1.33-3.29)	0.0001
Women	0.85 (0.42-1.71)	1.00	0.84 (0.57-1.24)	0.78 (0.55-1.10)	0.71 (0.47-1.08)	0.92 (0.50-1.70)	0.97 (0.38-2.43)	0.6300

*Adjusted for sex, BMI (<25, 25-<30, ≥30 kg/m^2), overall smoking status (never, former, current), Townsend deprivation index (quintiles), red meat intake (≤1, >1-<3, ≥3 times/week), processed meat (<1, 1, >1 time/week), height (continuous), family history of colorectal cancer (yes/no) and history of colorectal screening (yes/no).

§ P-values for test for trend were calculated by assigning participants the median value of their category of alcohol intake and this variable was modelled as a continuous variable (former drinkers were excluded for this analysis).

5.3 Discussion

5.3.1 Alcohol Intake and Colorectal Cancer

Though the relationship between alcohol intake and colorectal cancer has been investigated in a number of prospective cohort studies,[26, 56, 63] few studies have analysed alcohol intake with a similar level of precision as this analysis of UK Biobank data. Among the more detailed studies, there seems to be a suggestion that only alcohol intake above a certain threshold increases the risk of colorectal cancer.[45, 69, 70] A recent review of studies (published between 2008 and 2014) investigating alcohol intake and colorectal cancer concluded that alcohol intake below 30 g/d was not associated with an increased risk of colorectal cancer.[409] Overall results from this analysis of UK Biobank only found clear evidence of an increased risk for participants drinking ≥45 g/d. However, this analysis also found important differences in results between men and women.

5.3.2 Sex

Alcohol intake was strongly associated with colorectal cancer risk for men in this study. There did not appear to be a threshold for the effects of alcohol intake on colorectal cancer risk; compared to never drinkers, there was slight evidence of an increased risk for men drinking 5-<15 g/d when using the alternative reference group. Furthermore, the association for men appeared to be much stronger than in other studies of men in Western populations. For example, the HR (95% CI) comparing ≥45 g/d vs non-drinkers in the Western pooled analysis was 1.41 (1.11-1.79) and the HR (95% CI) comparing ≥45 g/d vs never drinkers in this analysis was 1.86 (1.16-2.99).[45] CIs were much wider in this analysis but results were similar (with narrower CIs) when using the alternative reference group and the multiple imputed data. In contrast to results for men, there was no evidence of an association between alcohol intake and colorectal cancer for women.

The WCRF classified alcohol intake as a cause of colorectal cancer for men but only as a probable cause for women.[56] Furthermore, meta-analysis results supported a stronger association for men than women.[63] However, it is likely that this is because women on

average have a lower alcohol intake and so there are fewer studies with results for heavy alcohol intake among women. For example, the most recent meta-analysis identified 20 studies (both case-control and cohort studies) reporting results for heavy intake (≥50 g/d) for men but only four studies for women.[63]

Compared with meta-analyses, individual studies provide more reliable evidence for investigating a differential effect of alcohol intake on colorectal cancer risk according to sex since they are able to use the same categories of alcohol intake for both men and women. Overall, individual studies seem to support a similar association between alcohol intake and colorectal cancer risk for both men and women.[45, 65, 66, 69] For example, in the pooled analysis by Cho et al., the HR (95% CI) comparing ≥45 g/d with non-drinkers was 1.41 (1.11-1.79) for men and 1.41 (0.98-2.02) for women.[45] However, the Canadian National Breast Screening Study (HR, 1.02; 95% CI, 0.72-1.44 for ≥30 g/d vs non-drinkers) and the Iowa Women's Health Study (HR, 1.00; 95% CI, 0.71-1.40 for ≥30 g/d vs non-drinkers) (both included in Cho et al.) both failed to find evidence of an association for women.[67, 68]

It is not clear why this analysis found evidence of an association only for men when other studies found evidence for both men and women. There were a relatively small number of colorectal cancer cases among women drinking ≥30 g/d in this study. However, there seemed to be a clear difference in results for men and women, even for 15-<30 g/d.

One possible explanation could be due to different patterns of consumption among men and women. Men are more likely to binge drink than women.[35] Thus, if heavy episodes of alcohol consumption were particularly associated with colorectal cancer risk, the association between alcohol intake and colorectal cancer would be stronger for men. That men are more likely to binge drink than women also means that men would be more likely to underestimate their average alcohol intake[357] which would also lead to a stronger association for men than women. However, while this explanation may be plausible, it does not explain why other studies do not generally find a stronger association for men.

Men and women also differ in terms of alcoholic beverages consumed which could lead to different results if different beverages are not similarly related to colorectal cancer risk. Hence, a further analysis was performed analysing alcohol intake from wine, beer and spirits separately (Table 5.3.1). Results for each beverage type were adjusted for alcohol intake from the other two beverage types. Former drinkers were excluded from the analysis. While beer intake was associated with colorectal cancer risk there appeared to be no association between wine intake and colorectal cancer risk (results were similar when restricted to men (data not shown)). Hence, it is possible that the different results for men and women for the association between total alcohol intake and colorectal cancer are due to the fact that men tend to drink more beer whereas women tend to drink more wine.

Table 5.3.1 Intake of Different Alcoholic Beverages and the Risk of Colorectal Cancer in UK Biobank

Alcohol intake (g/d)	Wine intake		Beer intake		Spirits intake	
	Cases	**HR (95% CI)***	**Cases**	**HR (95% CI)***	**Cases**	**HR (95% CI)***
0	399	1.00	775	1.00	1,072	1.00
>0-<5	306	0.98 (0.84-1.15)	200	0.96 (0.81-1.12)	357	0.90 (0.80-1.02)
5-<15	604	1.04 (0.91-1.18)	368	1.02 (0.88-1.17)	234	1.12 (0.97-1.29)
15-<30	290	1.04 (0.89-1.21)	204	1.21 (1.02-1.45)	49	1.09 (0.81-1.45)
30-<45	98	1.14 (0.91-1.43)	87	1.38 (1.08-1.75)	17	0.94 (0.58-1.52)
≥45	39	1.04 (0.74-1.45)	102	1.61 (1.28-2.03)	7	1.03 (0.49-2.17)
P-trend§		0.3173		<0.0001		0.6334

*Adjusted for sex, BMI (<25, 25-<30, ≥30 kg/m^2), overall smoking status (never, former, current), Townsend deprivation index (quintiles), red meat intake (≤1, >1-<3, ≥3 times/week), processed meat (<1, 1, >1 time/week), height (continuous), family history of colorectal cancer (yes/no) and history of colorectal screening (yes/no). Results for wine adjusted for intake of beer and spirits (0, >0-<5, 5-<15, 15-<30, 30-<45, ≥45 g/d), results for beer adjusted for intake of wine and spirits (0, >0-<5, 5-<15, 15-<30, 30-<45, ≥45 g/d) and results for spirits adjusted for intake of beer and wine (0, >0-<5, 5-<15, 15-<30, 30-<45, ≥45 g/d). Former drinkers were excluded from this analysis.

§ P-values for test for trend were calculated by assigning participants the median value of their category of wine/beer/spirits intake and this variable was modelled as a continuous variable.

An analysis of the Norfolk participants of the EPIC study found that people who drank ≥7 units/week of wine had a decreased risk of colorectal cancer compared to people who drank no wine (HR, 0.61; 95% CI, 0.40–0.94).[83] A Danish study also found slight evidence that wine intake decreased the risk of colon cancer (HR, 0.5; 95% CI, 0.2-1.0

for ≥14 vs 0 glasses of wine/week) and also that the risk of rectal cancer for people drinking ≥14 drinks per week was higher if they did not drink any wine.[73] Both of these studies cited evidence that resveratrol, found in red wine, may be protective against cancer. However, this is controversial since it is doubtful that resveratrol is present in wine at sufficient levels to reduce cancer risk.[410] Furthermore, most other studies have not found evidence for different associations by alcoholic beverage.[45, 65, 69, 70, 80] For example, the association between alcohol intake and rectal cancer was similar for women who drank wine exclusively and for women who drank other alcoholic drinks in the Million Women Study.[70]

5.3.3 Continuous Alcohol Intake Variable

Alcohol intake was also analysed as a continuous variable. A continuous variable does not require the creation of arbitrary cut-points such as for a categorical variable and so can be considered a more efficient use of the available data. It also avoids the need to explicitly define a reference group which can be a controversial decision for analyses of alcohol intake. However, results using a continuous variable are dependent on the choice of the analysis model.

Alcohol intake was modelled as a linear variable in this analysis since there was no evidence for a non-linear relationship using fractional polynomial analysis. Overall, each additional 10 g/d of alcohol was associated with a 5% increased risk of colorectal cancer. Results were similar when excluding never drinkers and excluding "extreme" levels of intake. Analysing alcohol intake as a continuous variable separately for men and women, there was no evidence for an association for women.

5.3.4 Colorectal Cancer Subsites

This analysis found that alcohol intake was associated with an increased risk of both colon cancer and rectal cancer for men though results were slightly different when using different reference groups. Using never drinkers as the reference group, there seemed to be a similar association for colon cancer and rectal cancer. However, using never drinkers and consistent light drinkers as the reference group, alcohol intake was more strongly associated with rectal cancer than colon cancer.

Other studies have found differing results for the associations with colon cancer and rectal cancer. Alcohol intake was associated with both colon and rectal cancer in the Western and Japanese pooled analyses with no clear differences in results.[45, 71] The EPIC cohort found that alcohol intake was more strongly associated with rectal cancer risk than colon cancer risk and the Million Women Study found that alcohol intake was only associated with rectal cancer risk.[69, 70] The Singapore Chinese Health Study found a stronger association between alcohol intake and colon cancer than rectal cancer and the Health Professionals Follow-up Study found evidence only for an association with colon cancer and not rectal cancer.[77, 78]

Overall, there is convincing evidence that alcohol intake increases the risk of both colon and rectal cancer and the evidence seems to indicate a slightly stronger association for rectal cancer.[60-62] For example, one meta-analysis found that the pooled RRs (95% CIs) associated with moderate drinking (12.6-49.9 g/d) were 1.15 (1.06-1.24) for colon cancer and 1.23 (1.13-1.35) for rectal cancer and that the pooled RRs (95% CIs) associated with heavy drinking (≥50 g/d) were 1.43 (1.23-1.67) for colon cancer and 1.59 (1.18-2.15) for rectal cancer.[62]

The current analysis found that drinking ≥30 g/d was associated with an increased risk of distal colon cancer compared to never drinkers while alcohol intake was not associated with proximal colon cancer. Few studies have investigated the association between alcohol intake and proximal colon cancer and distal colon cancer though these studies seem to agree that alcohol intake is more strongly associated with the risk of distal colon cancer.[45, 65, 69, 80, 82]

5.3.5 Reference Group

Non-drinkers

Many studies of alcohol intake and colorectal cancer have used non-drinkers as the reference group. The use of non-drinkers as the reference group has been argued against, particularly for studies of mortality and heart disease, since a number of former drinkers will have quit drinking due to illness.[46, 367] In UK Biobank, former drinkers were most likely to report a long-standing illness or poor health status. Furthermore,

prospective studies have shown that people reporting long standing illness or poor health are more likely to quit drinking alcohol.[365, 411]

However, the choice of reference group has not been given much discussion in studies of alcohol intake and colorectal cancer. Most studies have used non-drinkers as the reference group. This is presumably because many questionnaires only asked about baseline alcohol intake and did not include questions on past drinking. The Western pooled analysis did find that results were similar whether including or excluding former drinkers from the reference group (though it was only possible to identify people who had quit drinking in the previous five or ten years).[45]

Only a few studies have actually reported results comparing the risk of colorectal cancer between former drinkers and never drinkers; one Japanese study[82] did not find an increased risk of colorectal cancer for former drinkers (HR, 1.08; 95% CI, 0.64-1.85) though another Japanese study[143] found an increased risk of colon cancer (HR, 2.01; 95% CI, 1.09-3.68) but not rectal cancer (HR, 1.25; 95% CI, 0.66-2.38) for former drinkers. The EPIC cohort did not include results but reported that former drinkers had a similar risk to never drinkers.[69]

In UK Biobank, there was no evidence that former drinkers had an increased risk of colorectal cancer compared with never drinkers (overall or for men and women separately). Furthermore, results did not suggest that the risk of colorectal cancer differed for former drinkers who quit for "illness related reasons" or for "other reasons".

Consequently, results using non-drinkers instead of never drinkers as the reference group were similar to the main results for the association between alcohol intake and colorectal cancer. The HRs (95% CIs) comparing <5, 5-<15, 15-<30, 30-<45 and ≥45 g/d with non-drinkers for men were 1.08 (0.74-1.58), 1.29 (0.95-1.76), 1.43 (1.05-1.93), 1.54 (1.12-2.11) and 1.79 (1.31-2.44). The corresponding results for women were 0.79 (0.61-1.03), 0.80 (0.63-1.01), 0.81 (0.62-1.05), 0.95 (0.66-1.37) and 0.59 (0.30-1.13). Therefore, there was no evidence in this study that the association between alcohol intake and colorectal cancer may be underestimated in studies using non-drinkers as the reference group.

Never Drinkers

Never drinkers were used as the reference group in this analysis since they represent a consistent, well-defined group; while the risk of colorectal cancer for non-drinkers in each study will depend on the relative proportions of never drinkers and former drinkers, the risk for never drinkers should be reasonably consistent across studies. However, the main issue with using never drinkers as the reference group is that they represent a relatively small group which results in wide CIs.

The choice of never drinkers as the reference group has also been criticised because of the high levels of misclassification of never drinkers.[366] For example, one U.S. survey of alcohol intake found that approximately half of the people who reported being never drinkers in a 1992 survey had previously reported drinking in the 1984 or 1990 survey.[366] A British study also found that 67% of people who reported being never drinkers at age 45 previously reported some level of drinking at age 16, 23, 33 or 42.[370] However, in both these studies, these participants had mainly reported very light or occasional drinking in the past.

20,346 UK Biobank participants were completely re-assessed between 2012 and 2013, a mean of 4.4 years (range 2.1-7.0 years) after the baseline assessment. 710 participants reported being never drinkers at re-assessment. 532 of these participants reported being never drinkers at baseline, meaning that 25.1% of participants reporting never drinking at re-assessment had previously reported drinking. However, similar to the studies mentioned above, most of these participants reported a low level of drinking; 7.8% reported being former drinkers, 15.6% reported drinking "special occasions only" and 1.4% reported drinking "one to three times a month". Although this does not exactly represent the proportion of never drinkers at baseline who may be misclassified, it does suggest that there should not have been a large amount of misclassification and so this should not have had a large impact on results.

Light Drinkers

A further choice for the reference group when analysing alcohol intake is light drinkers. For example, the EPIC cohort used alcohol drinkers drinking <5 g/d as the reference

group and the Million Women Study used alcohol drinkers drinking ≤2 drinks/week.[69, 70] However, similar to non-drinkers, light drinkers represent a heterogeneous group. For example, some light drinkers will be former heavy drinkers and some light drinkers may be binge drinkers. Hence, the risk of colorectal cancer among light drinkers will depend on the relative proportions of people with different past exposures or different patterns of drinking. Furthermore, while some people quit drinking entirely as a result of illness, other people who suffer from illness will simply reduce their intake and become light or occasional drinkers.[365]

Alternative Reference Group

To further evaluate the impact of the choice of the reference group on the association between alcohol intake and colorectal cancer, it was also decided to investigate results using an alternative reference group. Ideally, it would have been possible to identify participants who had forever been at most light or occasional drinkers and had never been binge drinkers. These participants would represent a larger group than never drinkers on their own but would still represent a well-defined group who should experience the lowest risk of colorectal cancer. Furthermore, it would reduce concerns about the misclassification of never drinkers and hardly ever drinkers. Unfortunately, UK Biobank did not include sufficiently detailed information on the pattern of drinking or the lifetime exposure to alcohol intake.

Participants in UK Biobank were asked about how their drinking compared to ten years before baseline and thus consistent light drinkers were defined as participants who reported drinking <5 g/d at baseline and reported drinking "about the same" or "more nowadays" compared to ten years before baseline. Using never drinkers and consistent light drinkers combined as the reference group, the HRs for men for the association between alcohol intake and colorectal cancer were slightly lower to those using never drinkers as the reference group. However, since the CIs were narrower, the evidence for an increased risk was slightly stronger and there was even slight evidence of an increased risk for men drinking 5-<15 g/d. The HRs decreased for the association with colon cancer using the alternative reference group though there was still evidence for an association between alcohol intake and colon cancer risk. In contrast, the HRs increased for rectal cancer and there was clearer evidence of an increased risk for men drinking

≥15 g/d. There remained no evidence of an association for women between alcohol intake and colorectal, colon or rectal cancer.

5.3.6 Pattern of Drinking

The pattern of drinking may be important when considering the association between alcohol intake and colorectal cancer risk. For example, it was shown that the protective effects of moderate alcohol drinking for coronary heart disease were absent among people with irregular heavy drinking occasions (>60 g per occasion at least monthly).[412] However, no studies have presented detailed analyses on the pattern of drinking and the risk of colorectal cancer.

UK Biobank did not include detailed questions about patterns of drinking or binge drinking. However, participants were asked "when you drink alcohol is it usually with meals?" In an analysis of a small Italian study which asked a similar question, people who reported drinking outside of meals had a much higher risk of death compared to people who reported drinking with meals, adjusting for total alcohol intake.[413]

In the current analysis, participants were classified as meal drinkers, variable drinkers and non-meal drinkers. It was expected that variable drinkers and non-meal drinkers would be most likely to have heavy episodes of drinking. Therefore, if heavy episodes of drinking carried a greater risk of colorectal cancer compared to more consistent intake, these participants would be expected to have the highest risk. Furthermore, variable drinkers and non-meal drinkers may be expected to have a greater risk of colorectal cancer since they will be more likely to underestimate their average alcohol intake as a result of their irregular drinking pattern.

Results comparing variable drinkers and non-meal drinkers with meal drinkers (adjusting for total alcohol intake) did not show an increased risk of colorectal cancer. It is not clear why non-meal drinkers did not have a higher risk but it may be related to the crudeness of the question as a measure of binge drinking. Interestingly, results were similar when not adjusting for total alcohol intake despite the fact that variable drinkers and non-meal drinkers had a higher alcohol intake on average than meal drinkers (data not shown).

5.3.7 Effect Modifiers

Body Mass Index

The Western pooled analysis by Cho et al. was the first study to provide evidence that the association between alcohol intake and colorectal cancer risk may be modified by BMI.[45] Cho et al. found that alcohol intake was strongly associated with colorectal cancer risk for people with low BMI but there was no association for people with high BMI. The HR (95% CI) associated with ≥30 g/d vs non-drinkers was 1.84 (1.27-2.67) for BMI <22 kg/m^2, 1.23 (0.91-1.65) for BMI 22-<25kg/m^2 and 1.08 (0.88-1.33) for BMI ≥25 kg/m^2.

The effect modification by BMI was also investigated by Mizoue et al. in the Japanese pooled analysis.[71] Similar to the Western pooled analysis, alcohol intake had a stronger effect on colorectal cancer risk among men with low BMI though there was evidence that alcohol intake was associated with colorectal cancer within each category of BMI; the HR (95% CI) associated with drinking ≥69 g/d compared with non-drinkers was 3.25 (21.2-4.99) for BMI <22 kg/m^2, 2.12 (1.57-2.87) for BMI 22-<25kg/m^2 and 1.83 (1.26-2.67) for BMI ≥25 kg/m^2. A stronger effect of alcohol intake on colorectal cancer risk for people with lower BMI may also contribute to the stronger association observed for Japanese men compared to Western men.

However, results from this analysis of UK Biobank did not find a stronger association between alcohol intake and colorectal cancer risk among people with lower BMI. There is scarce evidence from other epidemiological studies though one case-control study actually found that alcohol intake increased the risk of colorectal cancer for obese subjects (BMI ≥30 kg/m^2) but there was no association for non-obese subjects.[84] Hence, more studies should investigate the possible interaction between alcohol intake and BMI.

Smoking

Alcohol intake and tobacco smoking are both known risk factors for colorectal cancer and it is plausible that alcohol intake and tobacco smoking have a synergistic effect on

cancer risk, such that the combined effect of both exposures is superior to the sum of the individual exposures. For example, there is evidence that alcohol intake and tobacco smoking have a synergistic effect on the risk of cancers of the upper aerodigestive tract.[26] There is not much existing evidence for a synergistic effect in relation to colorectal cancer.

In UK Biobank, although the evidence for an interaction was not clear, results did suggest a stronger association for ever smokers than for never smokers. Results from other studies also seem to show a stronger association between alcohol intake and colorectal cancer for current and former smokers.[45, 69, 77] For example, the HRs (95% CIs) comparing 30 g/d with non-drinking in the Western pooled analysis were 1.17 (0.84-1.63) for never smokers, 1.26 (1.00-1.58) for former smokers and 1.42 (1.11-1.83) for current smokers.[45] In the EPIC study, the HR (95% CI) associated with an increase of 15 g/d was higher for current smokers (HR, 1.23; 95% CI, 1.12-1.36) than for never smokers (HR, 1.15; 95% CI, 1.03-1.28) or former smokers (HR, 1.11; 95% CI, 0.97-1.28).[69]

Folate

The final effect modifier considered in this analysis was folate intake. Alcohol is known to inhibit folate metabolism and folate intake may be associated with a lower colorectal cancer risk.[85, 88, 89] Thus, the adverse effects of alcohol on folate are thought to be a potential mechanism for how alcohol increases an individual's risk of colorectal cancer.[414] This would predict a stronger association between alcohol intake and colorectal cancer among people with a lower folate intake. Unfortunately, data on average folate intake was available only for a subset of participants in UK Biobank and there was no clear evidence of a difference in the association for participants with below or above median folate intake.

Overall, the existing evidence for an interaction between alcohol intake and folate intake for colorectal cancer is unclear. Although results from the EPIC study did support a stronger association between alcohol intake and colorectal cancer for people with lower folate intake,[69] results from the Western and Japanese pooled analyses did not find evidence that the association differed by folate intake.[45, 71] Results from studies

of folate intake and colorectal cancer also find inconsistent results for an interaction with alcohol intake.[90, 91, 93, 117]

5.3.8 Multiple Imputation

Results for the association between alcohol intake and colorectal cancer using multiple imputation were generally very similar to the results among participants with complete data. This was expected since only participants who reported drinking "one to three times a month" or "special occasions only" had missing alcohol grams per day data in this dataset.

There was a notable increase for the HR for the risk of colorectal cancer associated with <5 g/d for men. Hence, this seems to suggest that, among male participants who reported drinking "one to three times a month" or "special occasions only", men with missing data had a greater risk of colorectal cancer than men with complete data. This could possibly be explained if men with missing data tended to drink more alcohol than men with complete data, so that the imputed values for alcohol intake underestimated their "actual" values of intake. Since men with missing data were recruited earlier than participants with complete data, this could happen if average alcohol intake decreased in the general population during recruitment. In fact, there does seem to have been a slight decrease in alcohol intake in the general population between 2006 and 2010 (beginning and end of recruitment).[35] However, these differences were small and it seems likely that changes in alcohol intake during recruitment will have been partly captured by alcohol frequency, meaning that differences within participants reporting "one to three times a month" or "special occasions only" should have been very minor.

Participants with missing data were recruited earlier and so had a longer period of follow-up on average than participants with complete data. Therefore, there was a much higher proportion of cases of colorectal cancer among participants with missing data. However, this should not have impacted on results since the Cox proportional hazards model accounts for different follow-up time between participants.

5.3.9 Summary

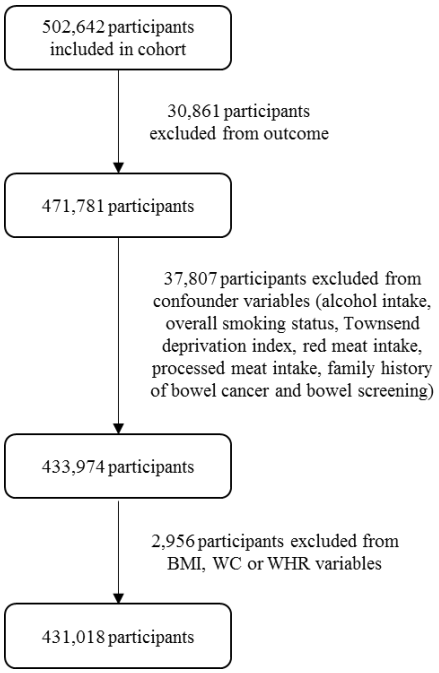
This analysis found that alcohol intake was strongly associated with colorectal cancer risk for men. However, this analysis did not find evidence of an association for women. Alcohol intake was associated with both colon cancer and rectal cancer risk but was more strongly associated to distal colon cancer than proximal colon cancer risk. Never drinkers were used as the reference group in this analysis. Results using non-drinkers as the reference group were similar to results using never drinkers as the reference group. This analysis was not able to identify any effect modifiers for the association between alcohol intake and colorectal cancer. Results using multiple imputation for participants with missing alcohol data were similar to results of the complete case analysis.

Chapter 6 Adiposity and Colorectal Cancer

6.1 Results

To assess the relationship between adiposity and colorectal cancer risk, data on BMI, WC and WHR were used. Analyses were restricted to participants with available data on BMI, WC and WHR. Thus, the analysis of adiposity and colorectal cancer included 197,443 men and 233,575 women (Figure 6.1.1). Data were analysed using Cox proportional hazards models using age as the underlying time variable. Participants were followed-up from the date of the baseline assessment until the date of any cancer diagnosis, the date of death or the end of follow-up (31st March 2014), whichever came first. Participants were followed up for a median 5.18 years (range 3.50-8.05 years). 1,219 men and 884 women were diagnosed with colorectal cancer. All analyses were adjusted for alcohol intake, overall smoking status, Townsend deprivation index, red meat intake, processed meat intake, family history of colorectal cancer and bowel screening. Analyses of WC and WHR were additionally adjusted for height. Results are not presented for colorectal cancer or for men and women overall as there is consistent evidence that the association differs by colorectal subsite and sex.[150, 152, 153]

Figure 6.1.1 Flowchart of Exclusions for Analysis of Adiposity and Colorectal Cancer



Flowchart of reasons for excluding participants from the analysis of adiposity and colorectal cancer.

6.1.1 Body Mass Index

Table 6.1.1 shows some basic characteristics of the analysis cohort by sex-specific quintiles of BMI. Men with higher BMI were more likely to be heavy drinkers. Men and women with the lowest BMI were most likely to be current smokers. Men and women with the highest BMI were more likely to live in areas with higher deprivation. Men and women with higher BMI reported higher intakes of red meat and processed meat.

Quintiles of BMI for men showed a strong association with the risk of colon cancer (Table 6.1.2). Compared to men in the lowest quintile of BMI, the HRs (95% CIs) were 1.52 (1.19-1.95) for men in the fourth quintile and 1.63 (1.27-2.09) for men in the highest quintile. However, there was no clear evidence for an association between quintiles of BMI and the risk of rectal cancer (HR, 1.13; 95% CI, 0.84-1.51 for fifth vs first quintile). Table A-9 - Table A-12 in the appendix show the effects of adjusting for different confounders on results.

Compared to women with the lowest BMI, there was no evidence that women with the highest BMI had an increased risk of colon cancer (HR, 1.07; 95% CI, 0.82-1.39) or rectal cancer (HR, 0.88; 95% CI, 0.56-1.40) (Table 6.1.3). However, there was slight evidence that women in the second quintile of BMI had an increased risk of rectal cancer compared to women in the first quintile (HR, 1.41; 95% CI, 0.94-2.12).

BMI categories based on the WHO definitions of overweight (≥25 kg/m^2) and obesity (≥30 kg/m^2) were also considered. Compared to men with BM <25 kg/m^2, men with BMI 25-<30 and ≥30 kg/m^2 had an increased risk of colon cancer but evidence was weaker for an increased risk of rectal cancer. While there was no evidence of an increased risk of colon cancer for women with BMI ≥30 kg/m^2, there was slight evidence that women with BMI ≥30 kg/m^2 had a decreased risk of rectal cancer compared to women with BMI <25 kg/m^2. (Restricting the reference group to 18.5-<25 kg/m^2, results for colon cancer and rectal cancer for men and women were very similar (results not shown)).

Table 6.1.1 Characteristics of UK Biobank Cohort by Sex Specific BMI Quintiles and Sex

	BMI (kg/m^2)				
Men	**<24.45**	**24.45-<26.40**	**26.40-<28.28**	**28.28-<30.84**	**≥30.84**
Number of participants	39,155	39,890	38,170	41,200	39,028
Age, mean (SD), years	56.0 (8.4)	56.6 (8.2)	56.7 (8.2)	56.9 (8.1)	56.7 (7.9)
Alcohol intake, % ≥45 g/d	10.9	12.5	14.1	15.9	16.2
Overall smoking status, % current smoker	15.2	11.8	11.2	11.1	11.1
Townsend index, % in most deprived quintile	20.3	17.0	17.1	18.9	23.8
Red meat intake, % ≥3 times/week	21.7	23.0	24.8	27.1	30.7
Processed meat intake, % >once/week	39.9	40.4	42.1	44.3	49.5
Family history, %	10.5	11.2	11.3	11.8	12.0
Colorectal screening, %	30.6	31.5	32.3	32.4	31.8
Women	**<22.89**	**22.89-<25.04**	**25.04-<27.34**	**27.34-<30.80**	**≥30.80**
Number of participants	47,150	47,203	46,458	46,751	46,013
Age, mean (SD), years	54.5 (8.2)	55.8 (8.0)	56.6 (7.9)	57.1 (7.9)	56.5 (7.8)
Alcohol intake, % ≥45 g/d	1.6	1.7	1.9	2.0	1.8
Overall smoking status, % current smoker	9.8	8.5	8.5	8.5	8.1
Townsend index, % in most deprived quintile	17.1	15.5	16.8	19.4	26.2
Red meat intake, % ≥3 times/week	15.5	18.0	19.6	21.5	23.0
Processed meat intake, % >once/week	16.7	18.4	19.8	21.8	26.0
Family history, %	10.2	10.7	11.2	11.0	11.2
Colorectal screening, %	26.6	28.3	29.9	30.6	29.9

Table 6.1.2 BMI and the Risk of Colon Cancer and Rectal Cancer for Men in UK Biobank

		Colon cancer			Rectal cancer		
BMI (kg/m^2)	**Person-years**	**Cases**	**HR (95% CI)***	**HR (95% CI)†**	**Cases**	**HR (95% CI)***	**HR (95% CI)†**
<24.45	197,756	99	1.00	1.00	83	1.00	1.00
24.45-<26.40	201,583	118	1.11 (0.85-1.45)	1.06 (0.81-1.39)	81	0.91 (0.67-1.24)	0.89 (0.65-1.21)
26.40-<28.28	192,524	159	1.56 (1.21-2.00)	1.46 (1.13-1.87)	98	1.15 (0.86-1.54)	1.09 (0.81-1.47)
28.28-<30.84	207,592	184	1.67 (1.30-2.13)	1.52 (1.19-1.95)	113	1.23 (0.92-1.63)	1.14 (0.86-1.52)
≥30.84	196,336	186	1.81 (1.42-2.32)	1.63 (1.27-2.09)	106	1.23 (0.93-1.65)	1.13 (0.84-1.51)
P-trend§			<0.0001	<0.0001		0.0377	0.1628
<25	250,280	137	1.00	1.00	99	1.00	1.00
25-<30	494,608	372	1.30 (1.07-1.58)	1.22 (1.00-1.49)	249	1.22 (0.96-1.53)	1.16 (0.92-1.46)
≥30	250,903	237	1.66 (1.34-2.05)	1.50 (1.21-1.85)	133	1.29 (0.99-1.67)	1.19 (0.91-1.55)
P-trend§			<0.0001	0.0002		0.0706	0.2461

* Results not adjusted for confounders.

† Results adjusted for alcohol intake (never, former, special occasions only, one to three times a month, <5 g/d, 5-<15 g/d, 15-<30 g/d, 30-<45 g/d, ≥45 g/d), overall smoking status (never, former, current), Townsend deprivation index (quintiles), red meat intake (1, >1-<3, ≥3 times/week), processed meat (<1, 1, >1 time/week), family history of colorectal cancer (yes/no).

§ P-values for test for trend were calculated by assigning participants the median value of their category of BMI and this variable was modelled as a continuous variable.

Table 6.1.3 BMI and the Risk of Colon Cancer and Rectal Cancer for Women in UK Biobank

		Colon cancer			Rectal cancer		
BMI (kg/m^2)	**Person-years**	**Cases**	**HR (95% CI)***	**HR (95% CI)†**	**Cases**	**HR (95% CI)***	**HR (95% CI)†**
<22.89	239,626	106	1.00	1.00	38	1.00	1.00
22.89-<25.04	239,747	134	1.15 (0.89-1.49)	1.15 (0.89-1.48)	59	1.44 (0.96-2.16)	1.41 (0.94-2.12)
25.04-<27.34	235,738	137	1.13 (0.88-1.46)	1.13 (0.87-1.45)	60	1.42 (0.95-2.14)	1.38 (0.92-2.08)
27.34-<30.80	237,556	143	1.14 (0.88-1.46)	1.13 (0.88-1.45)	53	1.22 (0.80-1.85)	1.17 (0.77-1.78)
≥30.80	232,894	127	1.08 (0.84-1.40)	1.07 (0.82-1.39)	39	0.94 (0.60-1.47)	0.88 (0.56-1.40)
P-trend			0.7827	0.8433		0.2995	0.1921
<25	474,532	237	1.00	1.00	97	1.00	1.00
25-<30	435,007	257	1.05 (0.88-1.26)	1.05 (0.88-1.25)	105	1.07 (0.81-1.42)	1.05 (0.79-1.38)
≥30	276,021	153	1.01 (0.83-1.24)	1.01 (0.82-1.24)	47	0.77 (0.54-1.09)	0.73 (0.51-1.04)
P-trend			0.8827	0.9410		0.1617	0.0981

* Results not adjusted for confounders.

† Results adjusted for alcohol intake (never, former, special occasions only, one to three times a month, <5 g/d, 5-<15 g/d, 15-<30 g/d, 30-<45 g/d, ≥45 g/d), overall smoking status (never, former, current), Townsend deprivation index (quintiles), red meat intake (1, >1-<3, ≥3 times/week), processed meat (<1, 1, >1 time/week), family history of colorectal cancer (yes/no).

It is important to consider the potential for reverse causation for the association between adiposity and colorectal cancer i.e. participants' adiposity at baseline could be influenced by undiagnosed colorectal cancer. This would mean that the association between adiposity and colorectal cancer would change over follow-up time. To investigate this possibility, BMI was modelled as a time varying coefficient i.e. a product term for BMI and follow-up time was included in the Cox proportional hazards model (BMI was modelled as a continuous variable). If the association did vary with time, the coefficient for this term would be non-zero. The p-values, testing the null hypothesis that this term was zero versus the two-sided alternative, were 0.460 for men and 0.032 for women for the association with colon cancer and 0.714 for men and 0.917 for women for the association with rectal cancer. For colon cancer for women, the HR associated with the product term (between BMI and time) was extremely close to 1 (1.000007). Therefore, although the p-value was statistically significant (at the 5% significance level), it is unlikely that there was important reverse causation for these results.

In the touchscreen questionnaire, participants were asked whether their weight had changed compared to one year ago. 61.5% of men and 50.6% of women reported that they weighed about the same as one year ago. The analysis for the association between BMI and colorectal cancer risk was repeated among these participants (Table 6.1.4). Results were fairly similar for men; there was clear evidence for an association between BMI and colon cancer risk but the evidence for an association with rectal cancer risk was less clear. For women, there was still no clear evidence for an association with colon cancer and the increased risk of rectal cancer for women in the second BMI quintile appeared to be even greater in this analysis.

Table 6.1.4 BMI and the Risk of Colon Cancer and Rectal Cancer in UK Biobank for Participants with No Weight Change from One Year Ago

		Colon cancer		Rectal cancer	
BMI (kg/m^2)	**Person-years**	**Cases**	**HR (95% CI)***	**Cases**	**HR (95% CI)***
Men					
<24.45	153,944	79	1.00	62	1.00
24.45-<26.40	139,530	79	0.98 (0.72-1.34)	57	0.93 (0.65-1.33)
26.40-<28.28	119,317	101	1.41 (1.05-1.89)	57	1.06 (0.73-1.52)
28.28-<30.84	113,053	99	1.40 (1.04-1.88)	65	1.23 (0.87-1.75)
≥30.84	86,849	82	1.52 (1.11-2.08)	51	1.27 (0.87-1.86)
P-trend			0.0011		0.0809
Women					
<22.89	172,664	79	1.00	30	1.00
22.89-<25.04	141,846	86	1.19 (0.88-1.62)	46	1.67 (1.05-2.65)
25.04-<27.34	111,759	68	1.12 (0.81-1.56)	28	1.21 (0.72-2.04)
27.34-<30.80	93,745	61	1.15 (0.82-1.61)	27	1.33 (0.78-2.25)
≥30.80	79,141	53	1.24 (0.87-1.77)	14	0.82 (0.43-1.57)
P-trend			0.3146		0.3973

* Results adjusted for alcohol intake (never, former, special occasions only, one to three times a month, <5 g/d, 5-<15 g/d, 15-<30 g/d, 30-<45 g/d, ≥45 g/d), overall smoking status (never, former, current), Townsend deprivation index (quintiles), red meat intake (1, >1-<3, ≥3 times/week), processed meat (<1, 1, >1 time/week), family history of colorectal cancer (yes/no) and history of colorectal screening (yes/no).

BMI was strongly associated with the risk of proximal colon cancer for men (HR, 1.75; 95% CI, 1.24-2.49 for highest vs lowest quintiles) (Table 6.1.5). There was also slight evidence for an increased risk of distal colon cancer though the association appeared to be weaker (HR, 1.41; 95% CI, 0.97-2.05 for highest vs lowest quintile). For women, there was no evidence for an association between BMI and proximal colon cancer or distal colon cancer (though women in the second BMI quintile appeared to have an increased risk of distal colon cancer).

Table 6.1.5 BMI and the Risk of Proximal Colon Cancer and Distal Colon Cancer in UK Biobank

	Proximal colon cancer		Distal colon cancer	
BMI (kg/m^2)	**Cases**	**HR (95% CI)***	**Cases**	**HR (95% CI)***
Men				
<24.45	51	1.00	44	1.00
24.45-<26.40	52	0.96 (0.65-1.41)	62	1.19 (0.81-1.75)
26.40-<28.28	73	1.40 (0.98-2.01)	76	1.44 (0.99-2.09)
28.28-<30.84	98	1.74 (1.23-2.44)	82	1.37 (0.95-1.98)
≥30.84	91	1.75 (1.24-2.49)	82	1.41 (0.97-2.05)
P-trend		<0.0001		0.0797
Women				
<22.89	55	1.00	40	1.00
22.89-<25.04	60	0.98 (0.68-1.41)	63	1.43 (0.96-2.13)
25.04-<27.34	72	1.13 (0.79-1.61)	62	1.35 (0.91-2.01)
27.34-<30.80	71	1.08 (0.76-1.54)	67	1.38 (0.93-2.06)
≥30.80	59	0.99 (0.68-1.44)	58	1.25 (0.83-1.89)
P-trend		0.9938		0.5925

* Results adjusted for alcohol intake (never, former, special occasions only, one to three times a month, <5 g/d, 5-<15 g/d, 15-<30 g/d, 30-<45 g/d, ≥45 g/d), overall smoking status (never, former, current), Townsend deprivation index (quintiles), red meat intake (1, >1-<3, ≥3 times/week), processed meat (<1, 1, >1 time/week), family history of colorectal cancer (yes/no) and history of colorectal screening (yes/no) and height (continuous).

6.1.2 Waist Circumference and Waist to Hip Ratio

While BMI is the most common measure of adiposity in the general population, WC and WHR are also used to assess adiposity. Table 6.1.6 shows the correlation coefficients between the three adiposity measures for men and women. WC was highly correlated with BMI (0.88 for men and women) and WHR (0.79 for men and 0.74 for women) for both men and women. The correlations between BMI and WHR were weaker (0.59 for men and 0.46 for women).

Table 6.1.6 Correlation between BMI, WC and WHR

Correlation coefficients	BMI	WC	WHR
Men			
BMI	1.00		
WC	0.88	1.00	
WHR	0.59	0.79	1.00
Women			
BMI	1.00		
WC	0.88	1.00	
WHR	0.46	0.74	1.00

Table 6.1.7 WC and the Risk of Colon Cancer and Rectal Cancer in UK Biobank

		Colon cancer		Rectal cancer	
WC (cm)	**Person-years**	**Cases**	**HR (95% CI)***	**Cases**	**HR (95% CI)***
Men					
≤87	191,786	76	1.00	60	1.00
88-93	212,882	127	1.31 (0.98-1.74)	97	1.29 (0.94-1.78)
94-98	195,031	157	1.63 (1.23-2.14)	99	1.35 (0.97-1.86)
99-105	202,313	171	1.60 (1.21-2.10)	114	1.41 (1.03-1.93)
≥106	193,779	215	2.02 (1.54-2.64)	111	1.39 (1.01-1.92)
P-trend			<0.0001		0.0687
Women					
≤74	263,697	108	1.00	44	1.00
75-80	239,409	128	1.17 (0.91-1.52)	56	1.26 (0.85-1.87)
81-86	225,830	136	1.23 (0.95-1.59)	47	1.05 (0.69-1.59)
87-95	242,507	141	1.14 (0.89-1.47)	52	1.04 (0.69-1.56)
≥96	214,117	134	1.25 (0.96-1.62)	50	1.12 (0.74-1.70)
P-trend			0.1540		0.8357

* Results adjusted for alcohol intake (never, former, special occasions only, one to three times a month, <5 g/d, 5-<15 g/d, 15-<30 g/d, 30-<45 g/d, ≥45 g/d), overall smoking status (never, former, current), Townsend deprivation index (quintiles), red meat intake (1, >1-<3, ≥3 times/week), processed meat (<1, 1, >1 time/week), family history of colorectal cancer (yes/no), history of colorectal screening (yes/no) and height (continuous).

Table 6.1.8 WHR and the Risk of Colon Cancer and Rectal Cancer in UK Biobank

		Colon cancer		Rectal cancer	
WHR	**Person-years**	**Cases**	**HR (95% CI)***	**Cases**	**HR (95% CI)***
Men					
<0.883	206,085	84	1.00	61	1.00
0.883-<0.918	200,001	104	1.10 (0.82-1.46)	77	1.12 (0.80-1.57)
0.918-<0.950	203,347	162	1.52 (1.17-1.98)	108	1.41 (1.03-1.93)
0.950-<0.990	198,071	181	1.61 (1.24-2.09)	116	1.44 (1.05-1.97)
≥0.990	188,287	215	1.87 (1.44-2.42)	119	1.45 (1.06-1.99)
P-trend			<0.0001		0.0091
Women					
<0.758	249,292	109	1.00	38	1.00
0.758-<0.796	240,077	107	0.93 (0.71-1.22)	49	1.25 (0.82-1.92)
0.796-<0.832	238,261	144	1.19 (0.92-1.53)	60	1.47 (0.98-2.22)
0.832-<0.876	227,351	130	1.08 (0.83-1.39)	42	1.04 (0.67-1.62)
≥0.876	230,578	157	1.22 (0.95-1.57)	60	1.41 (0.93-2.14)
P-trend			0.0750		0.3243

* Results adjusted for alcohol intake (never, former, special occasions only, one to three times a month, <5 g/d, 5-<15 g/d, 15-<30 g/d, 30-<45 g/d, ≥45 g/d), overall smoking status (never, former, current), Townsend deprivation index (quintiles), red meat intake (1, >1-<3, ≥3 times/week), processed meat (<1, 1, >1 time/week), family history of colorectal cancer (yes/no), history of colorectal screening (yes/no) and height (continuous).

There was a strong association between WC and colon cancer risk for men (Table 6.1.7). There was also evidence of an association between WC and rectal cancer risk for men. The HRs (95% CIs) comparing the highest and lowest quintiles of WC were 2.02 (1.54-2.64) for colon cancer and 1.39 (1.01-1.92) for rectal cancer. The association between WC and colon cancer risk appeared to be weaker for women than for men though there was still slight evidence of an increased risk of colon cancer for women in the highest WC quintile (HR, 1.25; 95% CI, 0.96-1.62). There was no clear evidence for an association between WC and rectal cancer risk for women.

Results for WHR showed a similar pattern to results for WC (Table 6.1.8). WHR was associated with the risk of colon cancer (HR, 1.87; 95% CI, 1.44-2.42 for highest vs lowest quintile) and rectal cancer (HR, 1.45; 95% CI, 1.06-1.99 for highest vs lowest quintile) for men. There was slight evidence of an increased risk of colon cancer for women in the highest quintile of WHR (HR, 1.22; 95% CI, 0.95-1.57). In contrast to

results for WC, women in the highest quintile of WHR also appeared to have an increased risk of rectal cancer (HR, 1.41; 95% CI, 0.93-2.14).

Table 6.1.9 and Table 6.1.10 show the associations between WC and WHR and proximal and distal colon cancer for men and women. WC and WHR were both strongly associated with proximal colon cancer risk for men; the HRs (95% CIs) associated with the highest quintile compared with the lowest quintile were 2.53 (1.70-3.78) for WC and 2.19 (1.49-3.24) for WHR. Men in the highest quintiles of WC and WHR also had an increased risk of distal colon cancer though the association was weaker than for proximal colon cancer. Women in the highest quintiles of WC and WHR had an increased risk of proximal colon cancer but not distal colon cancer.

Table 6.1.9 WC and the Risk of Proximal Colon Cancer and Distal Colon Cancer in UK Biobank

	Proximal colon cancer		**Distal colon cancer**	
WC (cm)	**Cases**	**HR (95% CI)***	**Cases**	**HR (95% CI)***
Men				
≤87	33	1.00	39	1.00
88-≤93	60	1.49 (0.98-2.29)	61	1.16 (0.77-1.73)
94-≤98	81	2.07 (1.38-3.11)	71	1.33 (0.90-1.97)
99-≤105	86	2.02 (1.35-3.04)	80	1.32 (0.89-1.94)
≥106	105	2.53 (1.70-3.78)	95	1.53 (1.04-2.24)
P-trend		<0.0001		0.0190
Women				
≤74	47	1.00	47	1.00
75-≤80	63	1.32 (0.90-1.92)	56	1.18 (0.80-1.74)
81-≤86	69	1.42 (0.98-2.06)	63	1.31 (0.90-1.92)
87-≤95	72	1.33 (0.92-1.93)	65	1.20 (0.82-1.76)
≥96	66	1.45 (0.99-2.13)	59	1.23 (0.83-1.82)
P-trend		0.0923		0.3989

* Results adjusted for alcohol intake (never, former, special occasions only, one to three times a month, <5 g/d, 5-<15 g/d, 15-<30 g/d, 30-<45 g/d, ≥45 g/d), overall smoking status (never, former, current), Townsend deprivation index (quintiles), red meat intake (1, >1-<3, ≥3 times/week), processed meat (<1, 1, >1 time/week), family history of colorectal cancer (yes/no) and history of colorectal screening (yes/no) and height (continuous).

Table 6.1.10 WHR and the Risk of Proximal Colon Cancer and Distal Colon Cancer in UK Biobank

	Proximal colon cancer		Distal colon cancer	
WHR	**Cases**	**HR (95% CI)***	**Cases**	**HR (95% CI)***
Men				
<0.883	36	1.00	44	1.00
0.883-<0.918	54	1.39 (0.91-2.11)	46	0.88 (0.58-1.33)
0.918-<0.950	87	2.03 (1.38-3.01)	67	1.11 (0.76-1.62)
0.950-<0.990	90	2.03 (1.37-3.00)	82	1.25 (0.87-1.82)
≥0.990	98	2.19 (1.49-3.24)	107	1.56 (1.09-2.23)
P-trend		<0.0001		0.0009
Women				
<0.758	46	1.00	50	1.00
0.758-<0.796	55	1.13 (0.77-1.68)	42	0.80 (0.53-1.20)
0.796-<0.832	65	1.27 (0.87-1.85)	74	1.32 (0.92-1.90)
0.832-<0.876	67	1.31 (0.90-1.92)	59	1.05 (0.72-1.55)
≥0.876	84	1.56 (1.08-2.26)	65	1.08 (0.74-1.58)
P-trend		0.0132		0.4442

* Results adjusted for alcohol intake (never, former, special occasions only, one to three times a month, <5 g/d, 5-<15 g/d, 15-<30 g/d, 30-<45 g/d, ≥45 g/d), overall smoking status (never, former, current), Townsend deprivation index (quintiles), red meat intake (1, >1-<3, ≥3 times/week), processed meat (<1, 1, >1 time/week), family history of colorectal cancer (yes/no) and history of colorectal screening (yes/no) and height (continuous).

Next, analyses were performed which modelled quintiles of BMI and WC simultaneously (Table 6.1.11) and quintiles of BMI and WHR simultaneously (Table 6.1.12). For men, there was no longer an association between BMI and colon cancer risk or rectal cancer risk when including WC or WHR in the analysis model. In contrast, results for WC and WHR were similar for colon and rectal cancer risk with and without BMI though results were slightly attenuated for colon cancer.

For women, there was still no evidence that BMI was associated with colon cancer or rectal cancer risk. Results were also similar for WC and WHR after including BMI in the analysis model though there was now slight evidence of an increased risk of rectal cancer for women in the highest quintile of WC as well as women in the highest quintile of WHR.

Table 6.1.11 BMI and WC and the Risk of Colon Cancer and Rectal Cancer in UK Biobank

	Colon cancer		Rectal cancer	
	Cases	**HR (95% CI)***	**Cases**	**HR (95% CI)***
Men				
BMI				
<24.45	99	1.00	83	1.00
24.45-<26.40	118	0.92 (0.69-1.24)	81	0.76 (0.54-1.07)
26.40-<28.28	159	1.17 (0.85-1.60)	98	0.88 (0.61-1.28)
28.28-<30.84	184	1.14 (0.80-1.61)	113	0.90 (0.60-1.36)
≥30.84	186	1.07 (0.72-1.60)	106	0.90 (0.55-1.46)
P-trend		0.7720		0.8007
WC				
≤87	76	1.00	60	1.00
88-≤93	127	1.27 (0.93-1.73)	97	1.40 (0.98-1.98)
94-≤98	157	1.50 (1.07-2.11)	99	1.46 (0.98-2.18)
99-≤105	171	1.42 (0.98-2.06)	114	1.49 (0.96-2.32)
≥106	215	1.83 (1.20-2.78)	111	1.45 (0.86-2.42)
P-trend		0.0040		0.2086
Women				
BMI				
<22.89	106	1.00	38	1.00
22.89-<25.04	134	1.06 (0.80-1.40)	59	1.39 (0.89-2.17)
25.04-<27.34	137	0.99 (0.72-1.35)	60	1.35 (0.81-2.23)
27.34-<30.80	143	0.94 (0.66-1.35)	53	1.02 (0.57-1.83)
≥30.80	127	0.81 (0.53-1.24)	39	0.59 (0.29-1.21)
P-trend		0.2111		0.0420
WC				
≤74	108	1.00	44	1.00
75-≤80	128	1.17 (0.89-1.56)	56	1.13 (0.73-1.75)
81-≤86	136	1.27 (0.92-1.73)	47	0.97 (0.59-1.60)
87-≤95	141	1.24 (0.87-1.77)	52	1.14 (0.65-1.98)
≥96	134	1.50 (0.98-2.29)	50	1.79 (0.92-3.49)
P-trend		0.0731		0.0824

* Results adjusted for alcohol intake (never, former, special occasions only, one to three times a month, <5 g/d, 5-<15 g/d, 15-<30 g/d, 30-<45 g/d, ≥45 g/d), overall smoking status (never, former, current), Townsend deprivation index (quintiles), red meat intake (1, >1-<3, ≥3 times/week), processed meat (<1, 1, >1 time/week), family history of colorectal cancer (yes/no), history of colorectal screening (yes/no) and height (continuous). BMI results were adjusted for WC (quintiles) and WC results were adjusted for BMI (quintiles).

Table 6.1.12 BMI and WHR and the Risk of Colon Cancer and Rectal Cancer in UK Biobank

	Colon cancer		Rectal cancer	
	Cases	**HR (95% CI)***	**Cases**	**HR (95% CI)***
Men				
BMI				
<24.45	99	1.00	83	1.00
24.45-<26.40	118	0.98 (0.74-1.29)	81	0.81 (0.59-1.11)
26.40-<28.28	159	1.26 (0.96-1.65)	98	0.95 (0.69-1.30)
28.28-<30.84	184	1.25 (0.95-1.64)	113	0.95 (0.69-1.31)
≥30.84	186	1.27 (0.95-1.69)	106	0.91 (0.66-1.29)
P-trend		0.0652		0.9321
WHR				
<0.883	84	1.00	61	1.00
0.883-<0.918	104	1.03 (0.77-1.39)	77	1.15 (0.81-1.62)
0.918-<0.950	162	1.39 (1.05-1.84)	108	1.44 (1.03-2.02)
0.950-<0.990	181	1.41 (1.06-1.89)	116	1.46 (1.03-2.07)
≥0.990	215	1.58 (1.17-2.14)	119	1.47 (1.02-2.13)
P-trend		0.0002		0.0270
Women				
BMI				
<22.89	106	1.00	38	1.00
22.89-<25.04	134	1.12 (0.87-1.45)	59	1.36 (0.90-2.06)
25.04-<27.34	137	1.08 (0.85-1.40)	60	1.31 (0.86-1.99)
27.34-<30.80	143	1.06 (0.81-1.38)	53	1.08 (0.69-1.68)
≥30.80	127	0.98 (0.74-1.31)	39	0.80 (0.49-1.31)
P-trend		0.6321		0.0815
WHR				
<0.758	109	1.00	38	1.00
0.758-<0.796	107	0.93 (0.71-1.22)	49	1.25 (0.81-1.92)
0.796-<0.832	144	1.18 (0.91-1.53)	60	1.49 (0.98-2.26)
0.832-<0.876	130	1.08 (0.82-1.41)	42	1.08 (0.68-1.72)
≥0.876	157	1.24 (0.94-1.63)	60	1.57 (1.00-2.46)
P-trend		0.0574		0.0763

* Results adjusted for alcohol intake (never, former, special occasions only, one to three times a month, <5 g/d, 5-<15 g/d, 15-<30 g/d, 30-<45 g/d, ≥45 g/d), overall smoking status (never, former, current), Townsend deprivation index (quintiles), red meat intake (1, >1-<3, ≥3 times/week), processed meat (<1, 1, >1 time/week), family history of colorectal cancer (yes/no), history of colorectal screening (yes/no) and height (continuous). BMI results were adjusted for WHR (quintiles) and WHR results were adjusted for BMI (quintiles).

6.1.3 Waist to Height Ratio

Waist to height ratio was also analysed for men and women (Table 6.1.13). WHtR was strongly associated with colon cancer risk for men (HR, 1.97; 95% CI, 1.50-2.59 for highest vs lowest quintile) while the association with rectal cancer was less clear. There was slight evidence that women with greater WHtR had a higher risk of colon cancer; however, for rectal cancer, women in the second (HR, 1.82; 95% CI, 1.19-2.78) and third quintiles (HR, 1.58; 95% CI, 1.02-2.48) of WHtR had a particularly increased risk.

Table 6.1.13 Waist to Height Ratio and the Risk of Colon Cancer and Rectal Cancer in UK Biobank

	Colon cancer		Rectal cancer	
WHtR	**Cases**	**HR (95% CI)***	**Cases**	**HR (95% CI)***
Men				
<0.500	75	1.00	32	1.00
0.500-<0.532	118	1.31 (0.98-1.75)	64	1.11 (0.89-1.54)
0.532-<0.562	167	1.74 (1.32-2.29)	57	1.36 (0.99-1.86)
0.562-<0.602	182	1.78 (1.36-2.35)	47	1.25 (0.91-1.72)
≥0.602	204	1.97 (1.50-2.59)	49	1.31 (0.95-1.80)
P-trend		<0.0001		0.0996
Women				
<0.453	93	1.00	63	1.00
0.453-<0.492	129	1.22 (0.93-1.60)	84	1.82 (1.19-2.78)
0.492-<0.532	136	1.24 (0.95-1.62)	110	1.58 (1.02-2.45)
0.532-<0.586	148	1.27 (0.97-1.66)	108	1.24 (0.79-1.97)
≥0.586	141	1.28 (0.97-1.68)	116	1.35 (0.85-2.14)
P-trend		0.2465		0.6647

* Results adjusted for alcohol intake (never, former, special occasions only, one to three times a month, <5 g/d, 5-<15 g/d, 15-<30 g/d, 30-<45 g/d, ≥45 g/d), overall smoking status (never, former, current), Townsend deprivation index (quintiles), red meat intake (1, >1-<3, ≥3 times/week), processed meat (<1, 1, >1 time/week), family history of colorectal cancer (yes/no) and history of colorectal screening (yes/no).

6.1.4 Percent Body Fat

194,277 men and 230,531 women additionally had data on percent body fat, as measured by bioelectrical impedance. Percent body fat was strongly correlated with BMI; correlation coefficients were 0.80 for men and 0.85 for women. Table 6.1.14

shows the results for the association between quintiles of percent body fat and colorectal cancer for men and women. Percent body fat was associated with colon cancer risk for men (HR, 1.59; 95% CI, 1.23-2.05 for highest vs lowest quintile) but not women (HR, 0.98; 95% CI, 0.76-1.27 for highest vs lowest quintile). There was no clear evidence for an association between percent body fat and rectal cancer for men or women.

Table 6.1.14 Percent Body Fat and the Risk of Colon Cancer and Rectal Cancer in UK Biobank

	Colon cancer		Rectal cancer	
Body fat (%)	**Cases**	**HR (95% CI)***	**Cases**	**HR (95% CI)***
Men				
≤20.6	90	1.00	73	1.00
20.7-24.0	115	1.11 (0.84-1.46)	81	0.98 (0.71-1.34)
24.1-26.8	149	1.31 (1.00-1.70)	97	1.08 (0.80-1.47)
26.9-30.0	169	1.41 (1.09-1.83)	110	1.17 (0.87-1.59)
≥30.1	209	1.59 (1.23-2.05)	109	1.07 (0.79-1.45)
P-trend		<0.0001		0.4120
Women				
≤30.8	115	1.00	39	1.00
30.9-35.0	115	0.87 (0.67-1.13)	58	1.31 (0.87-1.97)
35.1-38.5	133	0.95 (0.74-1.22)	57	1.22 (0.81-1.84)
38.6-42.5	142	0.97 (0.76-1.25)	45	0.93 (0.60-1.43)
≥42.6	138	0.98 (0.76-1.27)	45	0.94 (0.61-1.46)
P-trend		0.8184		0.3394

* Results adjusted for alcohol intake (never, former, special occasions only, one to three times a month, <5 g/d, 5-<15 g/d, 15-<30 g/d, 30-<45 g/d, ≥45 g/d), overall smoking status (never, former, current), Townsend deprivation index (quintiles), red meat intake (1, >1-<3, ≥3 times/week), processed meat (<1, 1, >1 time/week), family history of colorectal cancer (yes/no) and history of colorectal screening (yes/no).

6.1.5 Effect Modifiers

Physical Activity

187,797 men and 215,976 women also had available data for physical activity (using the alternative physical activity variable). The association between BMI, WC and WHR and colon cancer was investigated across different levels of physical activity, modelling

BMI, WC and WHR as continuous variables in order to maximize statistical power (Table 6.1.15).

Table 6.1.15 BMI, WC, WHR and the Risk of Colon Cancer by Physical Activity

	Low physical activity		Medium physical activity		High physical activity	
	Cases	**HR (95% CI)***	**Cases**	**HR (95% CI)***	**Cases**	**HR (95% CI)***
BMI (per 5 kg/m^2)						
Men	202	1.06 (0.91-1.24)	292	1.31 (1.15-1.49)	205	1.29 (1.09-1.52)
			P-interaction = 0.0684			
Women	175	0.95 (0.82-1.10)	256	1.05 (0.93-1.19)	160	0.97 (0.81-1.16)
			P-interaction = 0.3600			
WC (per 10 cm)						
Men	202	1.11 (0.99-1.25)	292	1.24 (1.12-1.37)	205	1.27 (1.12-1.44)
			P-interaction = 0.2139			
Women	175	1.11 (0.99-1.25)	256	1.06 (0.95-1.17)	160	1.00 (0.87-1.15)
			P-interaction = 0.5899			
WHR (per 0.1)						
Men	202	1.26 (1.01-1.57)	292	1.30 (1.08-1.57)	205	1.54 (1.24-1.91)
			P-interaction = 0.4857			
Women	175	1.46 (1.19-1.79)	256	1.04 (0.87-1.25)	160	1.02 (0.81-1.30)
			P-interaction = 0.0567			

* Results adjusted for alcohol intake (never, former, special occasions only, one to three times a month, <5 g/d, 5-<15 g/d, 15-<30 g/d, 30-<45 g/d, ≥45 g/d), overall smoking status (never, former, current), Townsend deprivation index (quintiles), red meat intake (1, >1-<3, ≥3 times/week), processed meat (<1, 1, >1 time/week), family history of colorectal cancer (yes/no) and history of colorectal screening (yes/no).

BMI was associated with the risk of colon cancer among men who reported medium physical activity (HR, 1.31; 95% CI, 1.15-1.49 per 5 kg/m^2) and high physical activity (HR, 1.29; 95% CI, 1.09-1.52 per 5 kg/m^2) but there was no association for men reporting low physical activity (HR, 1.06; 95% CI, 0.91-1.24 per 5 kg/m^2). Results for WC for men were similar though there was slight evidence for an association for men reporting low physical activity (HR, 1.11; 95% CI, 0.99-1.25 per 10 cm). WHR was associated with the risk of colon cancer within each category of physical activity though the association was strongest for men who reported high physical activity (HR, 1.54; 95% CI, 1.24-1.91 per 0.1).

BMI was not associated with colon cancer risk for women reporting low, medium or high physical activity. For WC and WHR, there was an association only for women reporting low physical activity; the HR (95% CI) was 1.11 (0.99-1.25) for WC and 1.46 (1.19-1.79) for WHR.

Menopause

57,957 women reported being pre-menopausal at baseline and 73,719 women reported being post-menopausal at baseline and never having used HRT. Analysing the associations between BMI, WC and WHR and colon cancer risk separately for pre-menopausal women and post-menopausal women who had never taken HRT, there was no clear evidence for an association for either group of women (Table 6.1.16).

Table 6.1.16 BMI, WC, WHR and the Risk of Colon Cancer by Menopause Status

	Pre-menopause		**Post-menopause**	
	Cases	**HR (95% CI)***	**Cases**	**HR (95% CI)***
BMI (per 5 kg/m^2)	75	0.92 (0.72-1.16)	224	1.02 (0.89-1.16)
		P-interaction = 0.3938		
WC (per 10 cm)	75	0.94 (0.78-1.15)	224	1.08 (0.97-1.20)
		P-interaction = 0.2242		
WHR (per 0.1)	75	1.00 (0.70-1.41)	224	1.13 (0.94-1.37)
		P-interaction = 0.4711		

* Results adjusted for alcohol intake (never, former, special occasions only, one to three times a month, <5 g/d, 5-<15 g/d, 15-<30 g/d, 30-<45 g/d, ≥45 g/d), overall smoking status (never, former, current), Townsend deprivation index (quintiles), red meat intake (1, >1-<3, ≥3 times/week), processed meat (<1, 1, >1 time/week), family history of colorectal cancer (yes/no) and history of colorectal screening (yes/no). Results for WC and WHR additionally adjusted for height (continuous).

6.2 Discussion

6.2.1 Body Mass Index

BMI was clearly associated with colon cancer risk for men but there was no strong evidence for an association with rectal cancer risk. For women, there was no evidence of an association between BMI and colon cancer risk. Surprisingly, compared to women

with BMI <25 kg/m^2, women with BMI ≥30 kg/m^2 seemed to have a decreased risk of rectal cancer. Analysing BMI quintiles, it seemed that the decreased risk observed was explained by the fact that women with BMI 22.89-<25.04 kg/m^2 had an increased risk of rectal cancer compared to women with BMI <22.89 kg/m^2. However, it remains unclear why there would be an increased risk of rectal cancer for women with BMI 22.89-<25.04 kg/m^2 but not for women with BMI ≥30.80 kg/m^2.

Numerous studies have previously investigated the association between BMI and colon and rectal cancer. Overall, studies have shown that the association is stronger for men than women and is stronger for colon cancer than rectal cancer.[150, 152, 153] For example, a recent meta-analysis, pooling results from the highest vs lowest comparisons, found that the pooled RRs (95% CIs) were 1.55 (1.47-1.63) for colon cancer and 1.24 (1.11-1.38) for rectal cancer for men. For women, the RRs (95% CIs) were 1.23 (1.10-1.37) for colon cancer and 1.07 (1.01-1.14) for rectal cancer.[150]

Hence, results from UK Biobank were in broad agreement with results from meta-analyses, showing a strong association between BMI and colon cancer risk for men. It is not clear why this analysis of UK Biobank did not find evidence for an association between BMI and colon cancer for women or for an association between BMI and rectal cancer for men or women. However, individual cohorts do find slightly different results and other cohorts have also found similar results to this analysis.

For example, the EPIC cohort also found evidence of an association between BMI and colon cancer only for men and no evidence of an association between BMI and rectal cancer for men or women. The HRs (95% CIs) comparing the highest and lowest quintiles of BMI were 1.55 (1.12-2.15) for men and 1.06 (0.79-1.42) for women for colon cancer and 1.05 (0.72-1.55) for men and 1.06 (0.71-1.58) for women for rectal cancer.[47] The NIH-AARP Diet and Health Study also found evidence for a strong association between BMI and colon cancer for men but no evidence for an association with rectal cancer. However, there was evidence that BMI was associated with colon and rectal cancer for women.[160]

The associations between BMI and proximal colon cancer and distal colon cancer were also investigated. For men, BMI was strongly associated with proximal colon cancer

risk and there was slight evidence of an increased risk of distal colon cancer for men in the highest BMI quintile. For women, there was no clear evidence for an association between BMI and proximal colon cancer or distal colon cancer.

In contrast to these results, evidence from other cohort studies indicates a stronger association between BMI and distal colon cancer.[150, 151, 153] One meta-analysis which pooled results comparing the highest and lowest categories of BMI from 16 cohort studies found a pooled RR (95% CI) of 1.24 (1.08-1.42) for proximal colon cancer and 1.59 (1.34-1.89) for distal colon cancer.[151]

A strength of the current analysis, compared to many other studies of BMI and colorectal cancer, was that height and weight were measured at the baseline assessment. Many other studies have relied on self-reported height and weight. However, since people tend to underestimate their weight as well as overestimate their height,[48] self-reported data results in BMI being underestimated for many participants. Consequently, studies using self-reported height and weight data will tend to overestimate the association between BMI and colorectal cancer, as found in meta-analyses.[152-154]

The precise mechanisms linking adiposity to colorectal cancer are unclear. However, insulin resistance likely plays an important role in colorectal carcinogenesis.[157, 158, 185] This is supported by studies that found that people with type 2 diabetes have an increased risk of colorectal cancer.[415, 416] Furthermore, adipose tissue is not simply an inert tissue that stores fat; it also produces various hormones, growth factors and cytokines known as adipokines which may increase the risk of colorectal cancer.[157, 158, 185]

While this analysis found evidence for an association between BMI and colon cancer for men, there was no evidence that BMI was associated with colon cancer risk for women. Although other studies have found an association between BMI and colorectal cancer for women, it is clear that BMI is more strongly associated with colorectal cancer risk for men than for women. The reason for this difference remains unknown.

One hypothesis for the difference for men and women is related to the different body compositions of men and women.[47] The proportion of body fat and the way body fat is

distributed differs between men and women.[417] Overall, women tend to store greater levels of body fat than men. However, there are two types of body fat: subcutaneous body fat and visceral body fat.[418] Subcutaneous fat is stored just beneath the skin whereas visceral fat is stored within the abdominal cavity. Women tend to store more fat in the thighs and buttocks whereas men store more abdominal fat. Hence, women tend to have greater levels of subcutaneous fat and men tend to have greater levels of visceral fat.[185, 417] Visceral fat is more metabolically active than subcutaneous fat and secretes greater numbers of cytokines and hormones.[185, 229] Also, visceral fat is associated with a greater risk of insulin resistance.[418, 419]

BMI is a crude measure of body weight adjusted for height and is unable to account for how body fat is distributed throughout the body. Thus, the stronger association between BMI and colorectal cancer observed for men may be because BMI is more closely related with visceral adiposity for men than for women. Hence, measures of abdominal adiposity such as WC and WHR may represent more precise predictors of colorectal cancer risk for both men and women.

Results from the EPIC cohort provided strong evidence for this hypothesis.[47] While BMI was associated with colon cancer risk only for men and not for women, WC and WHR were associated with colon cancer risk for both sexes. The HRs (95% CIs) associated with the highest vs lowest quintile of WC were 1.39 (1.01-1.93) for men and 1.48 (1.08-2.03) for women and the HRs (95% CIs) associated with the highest vs lowest quintile of WHR were 1.51 (1.06-2.15) for men and 1.52 (1.12-2.05) for women. Furthermore, after adjusting for body weight as well as height, the associations for women remained whereas the associations for men were attenuated, indicating that WC and WHR provided information about colon cancer risk beyond BMI for women.

6.2.2 Waist Circumference and Waist to Hip Ratio

For men, this analysis found that, similar to BMI, WC and WHR were strongly associated with colon cancer risk. For women, while there was no association between BMI and colon cancer, there was slight evidence for an association between WC and WHR and colon cancer risk. Thus, this seemed to suggest that WC and WHR were better predictors of colon cancer risk than BMI for women, in agreement with results

from the EPIC study. Also, modelling BMI and WC/WHR together in this analysis, the associations with WC and WHR for men remained while there was no longer an association between BMI and colon cancer, suggesting that WC and WHR provided better measures of colon cancer risk than BMI for men.

However, in contrast to the EPIC study, the association between WC/WHR and colon cancer risk still appeared to be much weaker for women than for men. Hence, results in this analysis did not seem to support the hypothesis described above to explain the differences in the association between BMI and colon cancer risk for men and women.

Other studies have also investigated the associations between BMI, WC and WHR and colon cancer risk for men and women. For example, in the NIH-AARP Diet and Health Study, BMI, WC and WHR were each associated with colon cancer risk for men; HRs (95% CIs) comparing highest and lowest quintiles were 1.42 (1.1-1.68) for BMI, 1.45 (1.16-1.82) for WC and 1.29 (1.10-1.52) for WHR.[186] In contrast, there was no evidence for an association for women with any of these adiposity measures; HRs (95% CIs) were 0.96 (0.74-1.23) for BMI, 0.90 (0.63-1.27) for WC and 0.90 (0.70-1.15) for WHR. Further adjusting for BMI, results for WC and WHR were slightly attenuated for men while there was still no evidence for an association for women. Analysis of the Shanghai Men's Health Study and the Shanghai Women's Health Study also found that BMI, WC and WHR were each strongly associated with colon cancer risk for men but not for women.[164]

One cohort of U.S. men and women found that both BMI and WC were strongly associated with colon cancer risk for men (HR, 1.93; 95% CI, 1.14-3.28 for ≥35.0 vs 18.5-24.9 kg⁄m^2 and HR, 2.05; 95% CI, 1.29-2.35 for ≥120 vs <95 cm).[187] For women, there also seemed to be a similar increased risk of colon cancer for both BMI and WC (HR, 1.40; 95% CI, 0.84-2.36 for ≥35.0 vs 18.5-24.9 kg⁄m^2 and HR, 1.54; 95% CI, 1.00-2.37 for ≥110 vs <85 cm). Results were also presented for BMI and WC mutually adjusted for one another for men and women and results were clearly attenuated for BMI but not for WC.

While BMI was associated with colon cancer risk only for men and not for women in the Melbourne Collaborative Cohort Study, WC and WHR were associated with colon

cancer risk for both men and women (though the associations were slightly stronger for men).[172, 173] In the Iowa Women's Health Study, there were similar associations with colon cancer risk for BMI (HR, 1.29; 95% CI, 1.10-1.51), WC (HR, 1.32; 95% CI, 1.11-1.56) and WHR (HR, 1.28; 95% CI, 1.08-1.50), comparing highest and lowest quartiles.[188]

Thus, all studies found strong associations between BMI, WC and WHR and colon cancer risk for men. Furthermore, studies that found positive associations between WC/WHR and colon cancer risk and further adjusted for BMI generally found that WC and WHR were still associated with colon cancer risk, indicating that WC and WHR may be more directly associated with colon cancer risk.

Results were less clear for women. Some studies did show an association between WC and WHR but these studies gave conflicting results about whether WC/WHR were more strongly associated with colon cancer risk than BMI. Furthermore, results from these studies still showed that the association between WC/WHR and colon cancer was weaker for women than for men. Hence, it seems unlikely that the difference in the association between BMI and colon cancer risk between men and women is due to the difference in body composition between men and women.

Visceral Abdominal Adiposity

However, it should be pointed out that while WC and WHR measure total abdominal fat, they do not differentiate between subcutaneous abdominal fat and visceral abdominal fat. WC is strongly correlated with total abdominal fat for both men and women. In contrast, WC is only moderately correlated with visceral abdominal fat though this correlation is stronger for men than for women.[420] In other words, the difference in results between men and women for measures of abdominal adiposity may be due to the fact that WC and WHR are not strong predictors of visceral fat.

Some case-control studies have actually measured visceral fat using computed tomography and found that visceral fat is associated with colorectal adenoma risk[421-425] and colorectal cancer risk.[426] Furthermore, visceral fat was associated with colorectal adenoma risk for both men and women.[421-423] Hence, it is possible that the stronger

associations between BMI/WC/WHR and colorectal cancer risk observed for men than women are due to how closely these measures represent visceral adiposity for men and women. Larger prospective studies will be required to investigate these relationships in greater detail and confirm the role of visceral fat for colorectal cancer risk.

6.2.3 Waist to Height Ratio

WHtR is another measure of adiposity that has received attention recently.[427] Recent research indicates that WHtR is a better predictor of cardiovascular risk factors including hypertension and type 2 diabetes than BMI.[428, 429] However, so far no studies have investigated WHtR and colorectal cancer risk.

WHtR is a measure of abdominal adiposity, similar to WC and WHR. WHtR was strongly correlated with WC in UK Biobank (correlation coefficient was 0.94 for men and 0.97 for women) and results were very similar for WC and for WHtR. WHtR was strongly associated with colon cancer risk for men whereas the association appeared to be weaker for women. There was slight evidence for an association with rectal cancer for men. However, one difference in the results for WHtR was that there was a large increased risk of rectal cancer for women in the second and third quintiles of WHtR, compared to women in the lowest quintile, but not for women with greater WHtR.

6.2.4 Percent Body Fat

Percent body fat, measured by bioelectrical impedance represents another method for assessing adiposity. Compared to other measures of adiposity that rely on data that can be recorded easily and cheaply, bioelectrical impedance analysis is more costly and so it is less widely used to assess adiposity.

One other cohort, the Melbourne Collaborative Cohort Study, previously investigated the relationship between body fat as measured by bioelectrical impedance and colorectal cancer risk.[172, 173, 237] For colon cancer, the HR (95% CI) associated with highest vs lowest tertile of percent body fat was 1.2 (0.9-1.6) for women. Results for percent body fat for men were not presented for colon cancer (though fat mass was associated with colon cancer risk for men). The HRs (95% CIs) comparing the highest and lowest

tertiles of percent body fat for rectal cancer risk were 1.5 (1.0-2.3) for men and 1.3 (0.8-2.2) for women.

In UK Biobank, percent body fat was strongly associated with colon cancer risk for men but not women and was not associated with rectal cancer risk for men or women. Hence, results for percent body fat were similar to results for BMI which was not surprising since percent body fat and BMI were highly correlated in this cohort.

6.2.5 Menopause

Another potential reason for the difference in results between men and women is due to the changes brought about by menopause among women. Studies have found a stronger association between BMI and colorectal cancer among pre-menopausal women than among post-menopausal women.[177, 184] For example, the Million Women Study found that the HR (95% CI) associated with a 10 kg/m^2 increase for colorectal cancer risk was 1.61 (1.05-2.48) for pre-menopausal women and 0.99 (0.88-1.12) for post-menopausal women who had never used HRT.[184] Meta-analysis results also support a stronger association between BMI and colorectal cancer for pre-menopausal women (RR, 1.20; 95% CI, 1.03-1.32 per 5 kg/m^2 increase) than post-menopausal women (RR, 1.09; 95% CI, 1.00-1.19).[152]

HRT use is associated with a lower risk of colorectal cancer.[64, 207, 216] Hence, oestrogen may have a protective effect for colorectal cancer. Menopause is associated with a decline in oestrogen levels. After menopause, adipose tissue becomes a major source of oestrogen and BMI is correlated with oestrogen levels in postmenopausal women.[222] Thus, it is hypothesised that the association between BMI and colorectal cancer is weaker among post-menopausal women since the adverse effects of greater adiposity are off-set by the beneficial effects of greater adiposity on oestrogen levels.[177, 206]

This analysis did not find any clear evidence that the association between BMI, WC or WHR and colon cancer risk differed among pre-menopausal women and post-menopausal women who had never used HRT. However, this cohort included a relatively small number of pre-menopausal women and there were relatively few cases of colon cancer among these women. Thus, the fact that this analysis did not find an

association among pre-menopausal women does not provide strong evidence that adiposity is not associated with colon cancer risk among these women.

However, other studies also cast doubt on a stronger association for pre-menopausal women. One study found that BMI was associated with colon cancer risk for post-menopausal women (HR, 1.76; 95% CI, 1.13-2.74 for ≥30 vs 18.5-22.9 kg⁄m^2) but not pre-menopausal women (HR, 0.79; 95% CI, 0.30-2.10).[169] A large prospective analysis of primary care data including 2.9 million women did not find different associations between BMI and colon or rectal cancer by menopause status.[161] Two other studies of post-menopausal women found a strong association between BMI and colorectal and colon cancer risk[188, 223] and the NIH-AARP Diet and Health Study found that BMI was associated with colon cancer risk for women aged 50-62 and 63-66 though there was no association for women aged 67-71 at baseline.[160]

Hence, there is no consistent evidence that adiposity is more strongly associated with colon cancer risk for post-menopausal women. Interestingly, after menopause, as oestrogen levels fall, women tend to gain weight and in particular to gain visceral fat.[417, 430] Since visceral fat is thought to be of particular importance for colorectal cancer risk,[157] this may actually suggest that there should be a stronger relation between adiposity and colorectal cancer among post-menopausal women.

6.2.6 Physical Activity

Confounder

Physical activity is strongly related to adiposity and physical activity is related to the risk of colon cancer.[25, 151, 225] Hence, physical activity represents an important confounder for analyses of adiposity and colorectal cancer. However, the analyses for the association between adiposity and colorectal cancer risk presented in this thesis were not adjusted for physical activity.

Physical activity was not included as a confounder in the analysis models as there was no evidence that results for the association between adiposity and colorectal cancer were strongly affected by the inclusion of physical activity and also because the inclusion of

physical activity would have led to the exclusion of a large number of participants (who responded “do not know” or “prefer not to answer” to at least one of the questions on physical activity) (see section 4.4.2). Table 6.2.1 shows results for the association between BMI and colorectal cancer with and without adjustment for physical activity (using the IPAQ physical activity variable), restricted to 164,406 men and 177,187 women with complete data for all variables.

One possible solution would have been to have imputed the physical activity data for those participants who were excluded using multiple imputation. However, since the data were missing as a result of participants choosing not to respond, the assumption that the data were missing at random was questionable and thus multiple imputation may have led to biased results.[375]

Table 6.2.1 BMI and the Risk of Colon Cancer and Rectal Cancer in UK Biobank with and without Adjustment for Physical Activity

		Colon cancer			Rectal cancer		
BMI (kg/m^2)	**Person-years**	**Cases**	**HR (95% CI)***	**HR (95% CI)†**	**Cases**	**HR (95% CI)***	**HR (95% CI)†**
Men							
<24.45	167,759	83	1.00	1.00	69	1.00	1.00
24.45-<26.40	170,793	93	1.02 (0.76-1.37)	1.01 (0.75-1.36)	70	0.92 (0.66-1.28)	0.92 (0.66-1.29)
26.40-<28.28	161,390	140	1.58 (1.20-2.07)	1.56 (1.19-2.06)	81	1.10 (0.79-1.52)	1.11 (0.80-1.53)
28.28-<30.84	171,793	148	1.54 (1.17-2.01)	1.51 (1.15-1.99)	96	1.20 (0.88-1.64)	1.22 (0.89-1.67)
≥30.84	157,964	142	1.62 (1.23-2.13)	1.57 (1.19-2.07)	70	0.96 (0.68-1.34)	1.00 (0.71-1.40)
P-trend			<0.0001	0.0001		0.7830	0.5920
Women							
<22.89	190,701	81	1.00	1.00	24	1.00	1.00
22.89-<25.04	188,452	97	1.11 (0.82-1.49)	1.10 (0.82-1.48)	41	1.60 (0.97-2.66)	1.60 (0.97-2.65)
25.04-<27.34	180,201	104	1.17 (0.88-1.57)	1.17 (0.87-1.56)	45	1.77 (1.07-2.91)	1.77 (1.07-2.91)
27.34-<30.80	176,003	106	1.19 (0.89-1.60)	1.18 (0.88-1.58)	44	1.73 (1.04-2.86)	1.72 (1.04-2.85)
≥30.80	164,121	87	1.10 (0.81-1.50)	1.08 (0.79-1.48)	27	1.16 (0.66-2.04)	1.15 (0.65-2.03)
P-trend			0.5661	0.6600		0.8999	0.9329

* Adjusted for alcohol intake (never, former, special occasions only, one to three times a month, <5 g/d, 5-<15 g/d, 15-<30 g/d, 30-<45 g/d, ≥45 g/d), overall smoking status (never, former, current), Townsend deprivation index (quintiles), red meat intake (1, >1-<3, ≥3 times/week), processed meat (<1, 1, >1 time/week), family history of colorectal cancer (yes/no), history of colorectal screening (yes/no) and height (continuous).

† Also adjusted for IPAQ physical activity (low, moderate, high).

However, since there exists convincing evidence for a relationship between physical activity and colon cancer risk,[25, 151, 225] it is unclear why there was no difference in results for BMI in this analysis. It could have been because self-reported physical activity is a very limited measure of physical activity and includes significant measurement error.[431, 432] For example, the adherence to physical activity recommendations in a representative sample of the U.S. population was substantially lower using accelerometer measured activity compared to self-reported activity.[433] Therefore, it is possible that adjustment for physical activity did not affect the association between BMI and colorectal cancer as a result of measurement error for physical activity. Yet, physical activity did appear to be inversely associated with colon cancer risk in the analysis models, particularly for men (data not shown).

Also, results from a meta-analysis of BMI and colorectal cancer indicated that the association between BMI and colorectal cancer was actually stronger among studies that adjusted for physical activity.[152] Hence, the relationship between BMI and physical activity for colorectal cancer risk may not be as straightforward as it seems.

Effect Modifier

Physical activity was also considered as a potential effect modifier for the association between adiposity and colon cancer risk in this thesis. For this analysis, an alternative physical activity variable was used since it excluded fewer participants (see section 4.4.2). A stronger association between adiposity and colorectal cancer among people with a low level of physical activity may be hypothesised since physical activity may counteract some of the harmful effects of adiposity e.g. by increasing insulin sensitivity.[434, 435]

This analysis found that WC and WHR were associated with colon cancer risk among women who reported low physical activity but not among women reporting medium or high physical activity. In contrast, BMI, WC and WHR were more strongly associated with colon cancer for men reporting medium and high physical activity than for men reporting low physical activity.

Few other studies have presented results for an interaction between adiposity and physical activity and these studies do not provide convincing evidence that the association between adiposity and colorectal cancer is modified by physical activity. One study analysed men according to categories of BMI and physical activity.[170] Compared with men with high BMI and low physical activity, men with low BMI and high physical activity had a lower risk of colorectal cancer (HR, 0.42; 95% CI, 0.22-0.78), however, there was no evidence for an interaction. A similar analysis was performed in the Netherlands Cohort Study using categories of trouser size and physical activity and there was also no clear evidence for an interaction.[168]

6.2.7 Summary

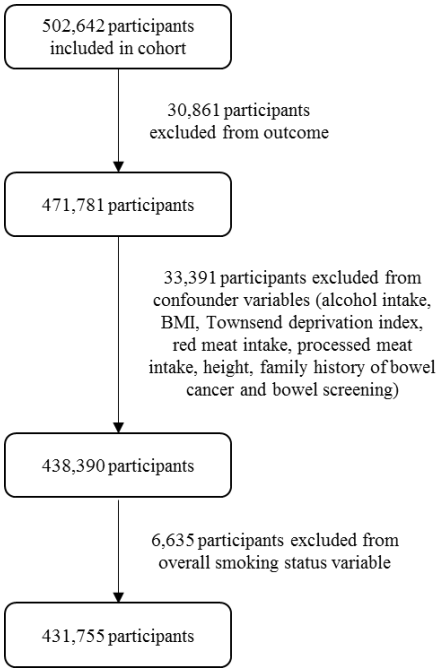
BMI was strongly associated with colon cancer risk for men but not for women. There was no clear evidence of an association between BMI and rectal cancer risk for men or women. One hypothesis for the stronger relationship between BMI and colorectal cancer for men than for women is that it is due to the different body compositions for men and women i.e. BMI is more closely related to abdominal or visceral adiposity for men than for women. However, the results from this analysis for WC and WHR did not support this hypothesis since, although there was evidence that WC and WHR were associated with colon cancer risk for women, the associations between WC/WHR and colon cancer were still much stronger for men than for women. Modelling BMI and WC/WHR together, the associations for WC/WHR remained while there was no longer an association between BMI and colon cancer risk, suggesting that measures of abdominal adiposity may be more directly associated with colon cancer risk. Adiposity was more strongly related to the risk of proximal colon cancer than distal colon cancer. There was some evidence that the associations between adiposity and colon cancer risk differed according to the level of physical activity though results differed for men and women. This analysis was not able to find evidence for different associations according to menopause status among women.

Chapter 7 Smoking and Colorectal Cancer

7.1 Results

The analysis of the association between smoking and colorectal cancer included 431,755 participants (197,800 men and 233,955 women) (Figure 7.1.1). Data were analysed using Cox proportional hazards models using age as the underlying time variable. Participants were followed-up from the date of the baseline assessment until the date of any cancer diagnosis, the date of death or the end of follow-up (31st March 2014), whichever came first. Participants were followed up for a median 5.18 years (range 3.50-8.05 years). During follow-up, 2,109 participants (1,223 men and 886 women) were diagnosed with colorectal cancer. All analyses were adjusted for sex, alcohol intake, BMI, Townsend deprivation index, red meat intake, processed meat intake, height, family history of colorectal cancer and bowel screening.

Figure 7.1.1 Flowchart of Exclusions for Analysis of Smoking and Colorectal Cancer



Flowchart of reasons for excluding participants from the analysis of smoking and colorectal cancer.

7.1.1 Main Results

Table 7.1.1 shows the characteristics of the analysis cohort by category of overall smoking status and by sex. Participants with a greater smoking exposure were more

likely to drink ≥45 grams of alcohol per day and to live in the most deprived areas. Also, participants with a greater smoking exposure ate more red meat and processed meat. Former daily smokers had the highest rates of bowel cancer screening for men and women. Daily smokers had the lowest rates of screening for men and always occasional smokers had the lowest rates for women.

Compared to never ever smokers, former daily smokers had an increased risk of colorectal cancer (HR, 1.20; 95% CI, 1.08-1.34) (Table 7.1.2). In contrast, there was no clear evidence of an increased risk for other categories of smoking, including daily smokers (HR, 1.09; 95% CI, 0.92-1.30). Results for former daily and daily smokers were similar for men and women. However, there was slight evidence that women but not men who smoked occasionally in the past (more than 100 times) and no longer smoked at baseline had an increased risk of colorectal cancer (HR, 1.19; 95% CI, 0.95-1.48). Also, men who smoked fewer than 100 times in the past had a lower risk of colorectal cancer than never ever smokers (HR, 0.77; 95% CI, 0.62-0.97).

Table 7.1.1 Characteristics of UK Biobank Cohort by Overall Smoking Status and Sex

	Never	Former occ <100	Former occ ≥100	Former daily	Always occ	Occ, former daily	Daily
Men							
Number of participants	70,071	26,238	23,158	54,439	4,030	2,494	17,370
Age, mean (SD), years	55.5 (8.2)	54.9 (8.3)	57.4 (8.0)	59.2 (7.4)	54.2 (8.3)	55.0 (8.5)	55.0 (8.2)
Alcohol intake, % ≥45 g/d	8.5	8.3	12.7	19.2	21.0	20.7	27.0
BMI, mean (SD), kg/m^2	27.6 (4.2)	27.1 (4.0)	27.7 (4.0)	28.7 (4.2)	28.0 (4.1)	27.8 (4.2)	27.1 (4.4)
Townsend index, % in most deprived quintile	16.9	14.8	16.5	19.5	25.3	28.8	37.9
Red meat intake, % ≥3 times/week	23.8	22.4	24.1	27.1	26.5	29.1	33.0
Processed meat intake, % >once/week	40.8	43.3	39.3	44.2	41.7	46.7	55.1
Family history, %	10.8	10.3	11.9	12.5	11.7	10.5	11.3
Colorectal screening, %	28.9	29.0	34.3	38.2	27.1	28.3	24.7
Women							
Number of participants	107,086	37,073	23,885	45,582	2,731	1,789	15,809
Age, mean (SD), years	55.9 (8.1)	55.8 (8.1)	56.6 (7.9)	57.5 (7.6)	52.8 (8.0)	54.3 (8.3)	54.3 (7.9)
Alcohol intake, % ≥45 g/d	0.7	1.2	1.7	3.2	3.6	4.6	5.5
BMI, mean (SD), kg/m^2	27.0 (5.2)	26.5 (5.0)	26.8 (5.0)	27.8 (5.3)	26.5 (4.9)	27.0 (5.2)	26.8 (5.3)
Townsend index, % in most deprived quintile	16.7	14.2	16.7	22.0	25.2	31.2	37.3
Red meat intake, % ≥3 times/week	19.5	18.7	18.4	19.7	17.1	19.3	22.8
Processed meat intake, % >once/week	19.8	21.7	17.9	20.8	17.3	21.0	26.2
Family history, %	10.5	10.8	11.4	11.6	9.6	10.8	10.6
Colorectal screening, %	27.8	29.2	30.8	33.4	21.2	25.4	23.8

Table 7.1.2 Overall Smoking Status and Cigarette Smoking Status and the Risk of Colorectal Cancer in UK Biobank

	Overall				Men			Women		
	Person-years	**Cases**	**HR (95% CI)***	**HR (95% CI)†**	**Person-years**	**Cases**	**HR (95% CI)†**	**Person-years**	**Cases**	**HR (95% CI)†**
Overall smoking status										
Never ever	900,144	742	1.00	1.00	355,705	360	1.00	544,438	382	1.00
Former occasional <100	321,017	231	0.89 (0.76-1.03)	0.88 (0.76-1.02)	132,953	100	0.77 (0.62-0.97)	188,064	131	1.00 (0.82-1.22)
Former occasional ≥100	236,639	247	1.10 (0.95-1.27)	1.07 (0.93-1.24)	116,252	143	0.99 (0.82-1.21)	120,388	104	1.19 (0.95-1.48)
Former daily	502,821	682	1.28 (1.15-1.42)	1.20 (1.08-1.34)	272,029	478	1.19 (1.04-1.38)	230,792	204	1.16 (0.97-1.39)
Always occasional	34,336	27	1.00 (0.68-1.47)	0.93 (0.63-1.37)	20,518	21	0.98 (0.63-1.52)	13,818	6	0.76 (0.34-1.71)
Occasional, former daily	21,754	22	1.20 (0.79-1.84)	1.13 (0.74-1.72)	12,634	15	1.08 (0.64-1.81)	9,120	7	1.22 (0.58-2.59)
Daily	168,328	158	1.17 (0.98-1.39)	1.09 (0.92-1.30)	87,464	106	1.12 (0.90-1.40)	80,863	52	1.04 (0.77-1.40)
Overall smoking status										
Never smokers	1,221,160	973	1.00	1.00	488,658	460	1.00	732,503	513	1.00
Former smokers	739,460	929	1.26 (1.15-1.38)	1.20 (1.09-1.32)	388,280	621	1.21 (1.07-1.37)	351,180	308	1.17 (1.01-1.36)
Current smokers	224,418	207	1.18 (1.02-1.37)	1.10 (0.95-1.29)	120,617	142	1.16 (0.95-1.41)	103,801	65	1.02 (0.78-1.33)
Cigarette smoking status										
Never smokers	1,221,160	973	1.00	1.00	488,658	460	1.00	732,503	513	1.00
Former cigarette smokers	481,055	648	1.34 (1.21-1.48)	1.27 (1.14-1.41)	250,996	445	1.32 (1.15-1.51)	230,059	203	1.16 (0.98-1.38)
Current cigarette smokers	156,268	134	1.15 (0.96-1.37)	1.09 (0.91-1.32)	75,995	82	1.14 (0.89-1.45)	80,273	52	1.05 (0.79-1.41)

* Adjusted for sex only.

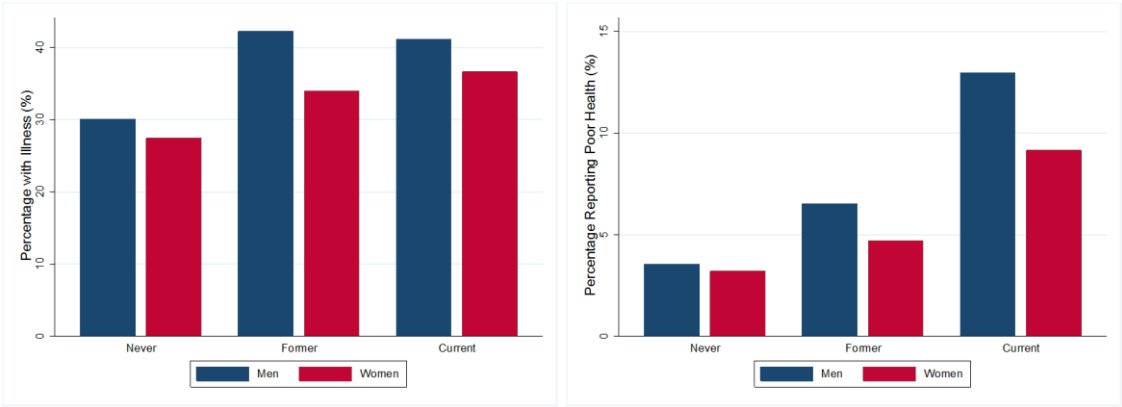
† Adjusted for sex, alcohol intake (never, former, special occasions only, one to three times a month, <5 g/d, 5-<15 g/d, 15-<30 g/d, 30-<45 g/d, ≥45 g/d), BMI (<25, 25-<30, ≥30 kg/m^2), Townsend deprivation index (quintiles), red meat intake (1, >1-<3, ≥3 times/week), processed meat (<1, 1, >1 time/week), height (continuous), family history of colorectal cancer (yes/no) and history of colorectal screening (yes/no).

The relationship between smoking and colorectal cancer was next evaluated using a simpler variable that combined the groups of the first variable. Both male (HR, 1.21; 95% CI, 1.07-1.37) and female former smokers (HR, 1.17; 95% CI, 1.01-1.36) had an increased risk of colorectal cancer compared to never smokers. There was slight evidence of an increased risk for male current smokers (HR, 1.16; 95% CI, 0.95-1.41) but not for female current smokers (HR, 1.02; 95% CI, 0.78-1.33). Table A-13 in the appendix shows the effects of adjusting for different confounders on results.

Next, the relationship between cigarette smoking status and colorectal cancer was evaluated. Results were similar to the previous variable for men and women except male former cigarette smokers had a slightly higher risk of colorectal cancer than male former smokers (HR, 1.32; 95% CI, 1.15-1.51).

Thus, former cigarette smokers had a greater risk of colorectal cancer than current cigarette smokers. This could be related to differences in health status between the two groups. For example, former cigarette smokers may have had greater levels of illness whereas the fact that current cigarette smokers continued to smoke could indicate that they were generally in good health and relatively free of major illnesses. However, the proportion of participants reporting a long-standing illness was similar among former and current cigarette smokers while approximately twice as many current cigarette smokers reported poor health as former cigarette smokers (Figure 7.1.2).

Figure 7.1.2 Participants Reporting Long-standing Illness and Poor Health, by Cigarette Smoking Status and Sex



Former cigarette smokers were also asked to provide reasons why they quit smoking. Participants could select all options that applied from “illness or ill health”, “doctor’s

advice", health precaution", "financial reasons" or could select "none of the above". Compared with never smokers, the HR (95% CI) was 1.22 (1.09-1.38) for former cigarette smokers who selected at least one of the illness related reasons (first three options) and 1.37 (1.16-1.60) for former cigarette smokers who selected other reasons.

Table 7.1.3 Cigarette Smoking Status and the Risk of Colon Cancer and Rectal Cancer in UK Biobank

	Cigarette smoking status		
	Never	**Former**	**Current**
Colon cancer			
Overall			
Cases	658	415	80
HR (95% CI)*	1.00	1.22 (1.08-1.39)	1.00 (0.79-1.27)
Men			
Cases	274	276	40
HR (95% CI)*	1.00	1.36 (1.14-1.62)	0.96 (0.68-1.35)
Women			
Cases	384	139	40
HR (95% CI)*	1.00	1.03 (0.85-1.26)	1.08 (0.77-1.51)
Rectal cancer			
Overall			
Cases	323	242	54
HR (95% CI)*	1.00	1.38 (1.16-1.65)	1.23 (0.91-1.66)
Men			
Cases	188	175	42
HR (95% CI)*	1.00	1.29 (1.04-1.60)	1.36 (0.96-1.93)
Women			
Cases	135	67	12
HR (95% CI)*	1.00	1.57 (1.16-2.14)	0.93 (0.51-1.70)

*Adjusted for sex, alcohol intake (never, former, special occasions only, one to three times a month, <5 g/d, 5-<15 g/d, 15-<30 g/d, 30-<45 g/d, ≥45 g/d), BMI (<25, 25-<30, ≥30 kg/m^2), Townsend deprivation index (quintiles), red meat intake (1, >1-<3, ≥3 times/week), processed meat (<1, 1, >1 time/week), height (continuous), family history of colorectal cancer (yes/no) and history of colorectal screening (yes/no).

Next, the associations between cigarette smoking and colon and rectal cancer were investigated (Table 7.1.3). Male former cigarette smokers had an increased risk of colon cancer (HR, 1.36; 95% CI, 1.14-1.62) but there was no evidence of an increased risk for

male current cigarette smokers (HR, 0.96; 95% CI, 0.68-1.35). There was no evidence of an increased risk of colon cancer for former or current cigarette smokers for women.

For rectal cancer, both former (HR, 1.29; 95% CI, 1.04-1.60) and current cigarette smokers (HR, 1.36; 95% CI, 0.96-1.93) had an increased risk for men. Female former cigarette smokers had a particularly high risk of rectal cancer (HR, 1.57; 95% CI, 1.16-2.14) but there was no evidence of an increased risk for female current cigarette smokers (HR, 0.93; 95% CI, 0.51-1.70).

Analysing proximal and distal colon cancer, there was only clear evidence of an increased risk of distal colon cancer for former cigarette smokers (HR, 1.44; 95% CI, 1.20-1.74) (Table 7.1.4).

Table 7.1.4 Cigarette Smoking Status and the Risk of Proximal Colon Cancer and Distal Colon Cancer in UK Biobank

	Cigarette smoking status		
	Never	**Former**	**Current**
Proximal colon cancer			
Cases	333	178	45
HR (95% CI)*	1.00	1.06 (0.87-1.28)	1.23 (0.89-1.69)
Distal colon cancer			
Cases	281	216	31
HR (95% CI)*	1.00	1.44 (1.20-1.74)	0.83 (0.57-1.21)

*Adjusted for sex, alcohol intake (never, former, special occasions only, one to three times a month, <5 g/d, 5-<15 g/d, 15-<30 g/d, 30-<45 g/d, ≥45 g/d), BMI (<25, 25-<30, ≥30 kg/m^2), Townsend deprivation index (quintiles), red meat intake (1, >1-<3, ≥3 times/week), processed meat (<1, 1, >1 time/week), height (continuous), family history of colorectal cancer (yes/no) and history of colorectal screening (yes/no).

7.1.2 Smoking Duration and Intensity

The effects of smoking duration and smoking intensity on colorectal cancer risk were also investigated for former and current cigarette smokers. Compared to never smokers, former cigarette smokers with ≥40 years of duration had the highest risk of colorectal cancer (HR, 1.65; 95% CI, 1.33-2.05) (Table 7.1.5). Former cigarette smokers with ≥40 years of duration also had an increased risk of both colon and rectal cancer. Results

seemed to indicate a dose-response relationship between smoking duration and rectal cancer but results were less clear for colon cancer since there was an increased risk for participants who smoked ≤9 years (HR, 1.32; 95% CI, 1.01-1.74).

Table 7.1.5 Smoking Duration and the Risk of Colorectal Cancer in UK Biobank

Smoking duration (years)	Colorectal cancer		Colon cancer		Rectal cancer	
	Cases	HR (95% CI)*	Cases	HR (95% CI)*	Cases	HR (95% CI)*
Never smoker	973	1.00	658	1.00	323	1.00
Former cigarette smoker						
≤9	81	1.26 (1.00-1.58)	56	1.32 (1.01-1.74)	26	1.17 (0.78-1.74)
10-19	184	1.20 (1.02-1.41)	116	1.15 (0.94-1.40)	70	1.32 (1.01-1.72)
20-29	150	1.14 (0.96-1.36)	91	1.05 (0.84-1.32)	61	1.35 (1.02-1.79)
30-39	128	1.30 (1.08-1.57)	80	1.23 (0.97-1.56)	50	1.48 (1.09-2.02)
≥40	97	1.65 (1.33-2.05)	67	1.70 (1.31-2.20)	31	1.58 (1.08-2.32)
P-trend		0.0510		0.1286		0.2063
Current cigarette smoker						
≤29	10	0.77 (0.41-1.45)	4	0.47 (0.17-1.27)	6	1.31 (0.57-3.01)
30-39	33	1.01 (0.71-1.44)	17	0.84 (0.51-1.37)	16	1.26 (0.75-2.11)
≥40	90	1.18 (0.95-1.48)	58	1.15 (0.87-1.52)	32	1.22 (0.84-1.77)
P-trend		0.1235		0.3446		0.1880

*Adjusted for sex, alcohol intake (never, former, special occasions only, one to three times a month, <5 g/d, 5-<15 g/d, 15-<30 g/d, 30-<45 g/d, ≥45 g/d), BMI (<25, 25-<30, ≥30 kg/m^2), Townsend deprivation index (quintiles), red meat intake (1, >1-<3, ≥3 times/week), processed meat (<1, 1, >1 time/week), height (continuous), family history of colorectal cancer (yes/no) and history of colorectal screening (yes/no).

There was also slight evidence that current cigarette smokers with duration ≥40 years had an increased risk of colorectal cancer (HR, 1.18; 95% CI, 0.95-1.48) though the risk was lower than for former cigarette smokers with duration ≥40 years.

Former cigarette smokers who smoked 11-15, 16-20 and ≥21 cigarettes per day had an increased risk of colorectal cancer (Table 7.1.6). These participants also had an increased risk of colon and rectal cancer risk. However, while participants who smoked ≥21 cigarettes per day had the highest risk of colon cancer (HR, 1.44; 95% CI, 1.17-1.78), participants who smoked 11-15 cigarettes per day had the highest risk of rectal cancer (HR, 1.68; 95% CI, 1.23-2.29).

Current cigarette smokers who smoked ≥21 cigarettes per day had a high risk of rectal cancer though the CI was wide (HR, 1.81; 95% CI, 1.05-3.13). There was no evidence of an association for colon cancer.

Table 7.1.6 Smoking Intensity and the Risk of Colorectal Cancer in UK Biobank

	Colorectal cancer		**Colon cancer**		**Rectal cancer**	
	Cases	**HR (95% CI)***	**Cases**	**HR (95% CI)***	**Cases**	**HR (95% CI)***
Never smoker	973	1.00	658	1.00	323	1.00
Former cigarette smoker						
≤10	122	1.05 (0.87-1.27)	76	0.96 (0.76-1.22)	47	1.24 (0.91-1.70)
11-15	118	1.43 (1.18-1.73)	74	1.34 (1.05-1.71)	47	1.68 (1.23-2.29)
16-20	229	1.26 (1.08-1.46)	146	1.21 (1.01-1.46)	84	1.33 (1.04-1.71)
≥21	174	1.39 (1.17-1.65)	116	1.44 (1.17-1.78)	60	1.33 (1.00-1.78)
P-trend		0.1735		0.0592		0.7969
Current cigarette smoker						
≤10	42	1.04 (0.76-1.42)	29	1.08 (0.74-1.57)	13	0.94 (0.54-1.65)
11-20	63	1.03 (0.79-1.33)	37	0.93 (0.67-1.30)	26	1.18 (0.79-1.78)
≥21	28	1.42 (0.97-2.09)	14	1.16 (0.68-1.99)	14	1.81 (1.05-3.13)
P-trend		0.0658		0.4615		0.0455

*Adjusted for sex, alcohol intake (never, former, special occasions only, one to three times a month, <5 g/d, 5-<15 g/d, 15-<30 g/d, 30-<45 g/d, ≥45 g/d), BMI (<25, 25-<30, ≥30 kg/m^2), Townsend deprivation index (quintiles), red meat intake (1, >1-<3, ≥3 times/week), processed meat (<1, 1, >1 time/week), height (continuous), family history of colorectal cancer (yes/no) and history of colorectal screening (yes/no).

Among former cigarette smokers, the effects of smoking duration and smoking intensity were analysed together (Table 7.1.7). Participants with the greatest level of exposure (participants who smoked ≥16 cigarettes per day for ≥40 years) had the highest risk of colorectal cancer compared to never smokers (HR, 1.72; 95% CI, 1.33-2.21).

There were some slight differences for colon and rectal cancer. Participants who smoked ≤15 or ≥16 cigarettes per day for ≥40 years both had a similar increased risk of colon cancer. However, for rectal cancer, there was only evidence of an increased risk for participants who smoked ≥16 cigarettes per day. There seemed to be a greater risk of rectal cancer for participants who smoked ≤15 cigarettes per day than for participants

who smoked ≥16 cigarettes per day among participants who smoked for ≤29 years. This pattern was reversed for participants who smoked for ≥30 years.

There was also slight evidence of an increased risk of colon cancer (HR, 1.41; 95% CI, 0.98-2.03) and rectal cancer (HR, 1.53; 95% CI, 0.94-2.49) for former cigarette smokers with the lowest level of exposure (participants who smoked ≤15 cigarettes per day for ≤9 years) compared to never smokers.

Table 7.1.7 Smoking Duration and Smoking Intensity for Former Cigarette Smokers and the Risk of Colorectal Cancer in UK Biobank

Smoking duration and smoking intensity	Colorectal cancer		Colon cancer		Rectal cancer	
	Cases	HR (95% CI)*	Cases	HR (95% CI)*	Cases	HR (95% CI)*
Never Smoker	973	1.00	658	1.00	323	1.00
≤9 yrs, ≤15 cigs/d	48	1.46 (1.09-1.95)	31	1.41 (0.98-2.03)	17	1.53 (0.94-2.49)
≤9 yrs, ≥16 cigs/d	31	1.00 (0.70-1.44)	24	1.20 (0.80-1.80)	7	0.64 (0.30-1.35)
10-19 yrs, ≤15 cigs/d	70	1.11 (0.87-1.42)	40	0.94 (0.68-1.30)	32	1.53 (1.06-2.20)
10-19 yrs, ≥16 cigs/d	114	1.27 (1.04-1.55)	76	1.30 (1.02-1.66)	38	1.19 (0.85-1.68)
20-29 yrs, ≤15 cigs/d	51	1.08 (0.82-1.44)	28	0.88 (0.60-1.28)	24	1.54 (1.02-2.34)
20-29 yrs, ≥16 cigs/d	99	1.18 (0.96-1.46)	63	1.15 (0.88-1.50)	37	1.26 (0.89-1.79)
30-39 yrs, ≤15 cigs/d	37	1.07 (0.77-1.49)	25	1.07 (0.72-1.60)	12	1.05 (0.59-1.88)
30-39 yrs, ≥16 cigs/d	91	1.42 (1.14-1.77)	55	1.31 (0.99-1.74)	38	1.69 (1.20-2.40)
≥40 yrs, ≤15 cigs/d	29	1.53 (1.05-2.22)	23	1.75 (1.15-2.66)	7	1.17 (0.55-2.48)
≥40 yrs, ≥16 cigs/d	68	1.72 (1.33-2.21)	44	1.70 (1.23-2.30)	24	1.78 (1.16-2.73)

*Adjusted for sex, alcohol intake (never, former, special occasions only, one to three times a month, <5 g/d, 5-<15 g/d, 15-<30 g/d, 30-<45 g/d, ≥45 g/d), BMI (<25, 25-<30, ≥30 kg/m^2), Townsend deprivation index (quintiles), red meat intake (1, >1-<3, ≥3 times/week), processed meat (<1, 1, >1 time/week), height (continuous), family history of colorectal cancer (yes/no) and history of colorectal screening (yes/no).

7.1.3 Smoking Initiation

Age at smoking initiation was also investigated for former and current cigarette smokers (Table 7.1.8). For former cigarette smokers, there were no clear trends in results, for colon cancer or rectal cancer risk. Former cigarette smokers with the earliest age at initiation (≤14 years) did have the highest risk of colon cancer (HR, 1.41; 95% CI, 1.08-1.82) however there was an increased risk for participants who started smoking at age

≥19 years (HR, 1.27; 95% CI, 1.03-1.57) while there was no increased risk for participants who started at age 17-18 years (HR, 1.01; 95% CI, 0.81-1.25). Results for rectal cancer showed an approximately similar level of risk for each category of age at initiation.

For current cigarette smokers, there was no evidence of an association between age at smoking and colon cancer risk but participants who started smoking ≤15 and 16-17 years seemed to have an increased risk of rectal cancer.

Table 7.1.8 Smoking Initiation and the Risk of Colorectal Cancer in UK Biobank

Smoking initiation (years)	Colorectal cancer		Colon cancer		Rectal cancer	
	Cases	**HR (95% CI)***	**Cases**	**HR (95% CI)***	**Cases**	**HR (95% CI)***
Never smoker	973	1.00	658	1.00	323	1.00
Former cigarette smoker						
≤14	104	1.41 (1.14-1.73)	66	1.41 (1.08-1.82)	38	1.40 (0.99-1.98)
15-16	230	1.32 (1.14-1.53)	148	1.30 (1.08-1.56)	86	1.41 (1.10-1.81)
17-18	158	1.10 (0.93-1.31)	96	1.01 (0.81-1.25)	64	1.32 (1.00-1.74)
≥19	152	1.29 (1.09-1.54)	102	1.27 (1.03-1.57)	52	1.37 (1.02-1.84)
P-trend		0.4940		0.5284		0.8182
Current cigarette smoker						
≤15	46	1.09 (0.81-1.48)	24	0.90 (0.60-1.36)	22	1.40 (0.90-2.18)
16-17	33	1.07 (0.76-1.52)	15	0.76 (0.45-1.27)	18	1.62 (1.00-2.62)
≥18	54	1.11 (0.84-1.46)	40	1.23 (0.89-1.70)	14	0.84 (0.49-1.44)
P-trend		0.8537		0.3238		0.1202

*Adjusted for sex, alcohol intake (never, former, special occasions only, one to three times a month, <5 g/d, 5-<15 g/d, 15-<30 g/d, 30-<45 g/d, ≥45 g/d), BMI (<25, 25-<30, ≥30 kg/m^2), Townsend deprivation index (quintiles), red meat intake (1, >1-<3, ≥3 times/week), processed meat (<1, 1, >1 time/week), height (continuous), family history of colorectal cancer (yes/no) and history of colorectal screening (yes/no).

7.1.4 Smoking Cessation

Participants who quit smoking after age 50 had an increased risk of colon cancer (HR, 1.43; 95% CI, 1.18-1.68) and rectal cancer (HR, 1.39; 95% CI, 1.05-1.85) (Table 7.1.9). However, for rectal cancer, participants who quit smoking 40-49 had the highest risk compared to never smokers (HR, 1.67; 95% CI, 1.28-2.17). Participants who quit

smoking before 30 also appeared to have an increased risk of colon and rectal cancer compared to never smokers.

Results for years since cessation did seem to show clearer trends for the risk of colon cancer and rectal cancer. The HRs (95% CIs) for participants who quit smoking ≤9 years before baseline were 1.45 (1.17-1.78) for colon cancer and 1.53 (1.15-2.04). Participants who quit at least 30 years before baseline still appeared to have an increased risk of rectal cancer compared to never smokers (HR, 1.24; 95% CI, 0.94-1.65).

Table 7.1.9 Smoking Cessation and the Risk of Colorectal Cancer in UK Biobank

	Colorectal cancer		**Colon cancer**		**Rectal cancer**	
	Cases	**HR (95% CI)***	**Cases**	**HR (95% CI)***	**Cases**	**HR (95% CI)***
Age at cessation (yrs)						
Never smoker	973	1.00	658	1.00	323	1.00
≤29	126	1.27 (1.05-1.53)	80	1.23 (0.97-1.49)	47	1.35 (0.99-1.84)
30-39	169	1.12 (0.95-1.33)	113	1.14 (0.93-1.36)	58	1.12 (0.84-1.49)
40-49	158	1.25 (1.05-1.49)	89	1.06 (0.85-1.36)	72	1.67 (1.28-2.17)
≥50	190	1.41 (1.20-1.66)	130	1.43 (1.18-1.68)	62	1.39 (1.05-1.85)
P-trend		0.1144		0.1738		0.3819
Years since cessation						
Never smoker	973	1.00	658	1.00	323	1.00
≤9	162	1.46 (1.23-1.73)	107	1.45 (1.17-1.78)	58	1.53 (1.15-2.04)
10-19	136	1.26 (1.05-1.51)	82	1.14 (0.90-1.44)	54	1.47 (1.09-1.97)
20-29	174	1.21 (1.03-1.43)	112	1.18 (0.96-1.45)	64	1.30 (0.98-1.71)
≥30	171	1.15 (0.97-1.35)	111	1.12 (0.91-1.38)	63	1.24 (0.94-1.65)
P-trend		0.0174		0.0452		0.1829

*Adjusted for sex, alcohol intake (never, former, special occasions only, one to three times a month, <5 g/d, 5-<15 g/d, 15-<30 g/d, 30-<45 g/d, ≥45 g/d), BMI (<25, 25-<30, ≥30 kg/m^2), Townsend deprivation index (quintiles), red meat intake (1, >1-<3, ≥3 times/week), processed meat (<1, 1, >1 time/week), height (continuous), family history of colorectal cancer (yes/no) and history of colorectal screening (yes/no).

Analysing years since cessation and smoking intensity together, participants who quit smoking ≤9 years before baseline and smoked ≥16 cigarettes per day had an increased risk of colon cancer (HR, 1.51; 95% CI, 1.17-1.95) (Table 7.1.10). Participants who smoked ≥16 cigarettes per day and quit 20-29 years before baseline also had an

increased risk but evidence for an increased risk was less clear for those who quit 10-19 years before baseline. Among participants who smoked ≤15 cigarettes per day, there was an increased risk of colon cancer only for those who quit ≤9 years before baseline.

For rectal cancer, participants who smoked ≥16 cigarettes per day had an increased risk if they quit smoking ≤9 (HR, 1.80; 95% CI, 1.30-2.49) or 10-19 years before baseline (HR, 1.58; 95% CI, 1.12-2.23). However, among participants who smoked ≤15 cigarettes per day, there was an increased risk only for participants who quit 20-29 or ≥30 years before baseline.

Table 7.1.10 Smoking Cessation and Smoking Intensity and the Risk of Colorectal Cancer in UK Biobank

Years since cessation and smoking intensity	Colorectal cancer		Colon cancer		Rectal cancer	
	Cases	HR (95% CI)*	Cases	HR (95% CI)*	Cases	HR (95% CI)*
Never Smoker	973	1.00	658	1.00	323	1.00
≤9 yrs, ≥16 cigs/d	111	1.59 (1.30-1.95)	69	1.51 (1.17-1.95)	44	1.80 (1.30-2.49)
≤9 yrs, ≤15 cigs/d	51	1.25 (0.94-1.65)	38	1.36 (0.98-1.89)	14	1.04 (0.61-1.78)
10-19 yrs, ≥16 cigs/d	92	1.34 (1.08-1.67)	54	1.20 (0.90-1.59)	38	1.58 (1.12-2.23)
10-19 yrs, ≤15 cigs/d	44	1.14 (0.84-1.54)	28	1.06 (0.73-1.56)	16	1.25 (0.76-2.08)
20-29 yrs, ≥16 cigs/d	113	1.25 (1.03-1.53)	79	1.35 (1.06-1.71)	35	1.10 (0.77-1.57)
20-29 yrs, ≤15 cigs/d	61	1.17 (0.90-1.51)	33	0.94 (0.66-1.33)	29	1.66 (1.13-2.43)
≥30 yrs, ≥16 cigs/d	87	1.06 (0.85-1.33)	60	1.12 (0.85-1.47)	27	0.95 (0.63-1.42)
≥30 yrs, ≤15 cigs/d	82	1.24 (0.99-1.56)	50	1.12 (0.84-1.50)	34	1.55 (1.08-2.22)

*Adjusted for sex, alcohol intake (never, former, special occasions only, one to three times a month, <5 g/d, 5-<15 g/d, 15-<30 g/d, 30-<45 g/d, ≥45 g/d), BMI (<25, 25-<30, ≥30 kg/m^2), Townsend deprivation index (quintiles), red meat intake (1, >1-<3, ≥3 times/week), processed meat (<1, 1, >1 time/week), height (continuous), family history of colorectal cancer (yes/no) and history of colorectal screening (yes/no).

7.1.5 Effect Modifiers

Finally, the relationship between cigarette smoking status and colorectal cancer risk was investigated according to different levels of BMI (Table 7.1.11). Within each category of BMI, former cigarette smokers had an increased risk of colorectal cancer compared to never smokers. Current cigarette smokers had an increased risk compared to never smokers only among participants with BMI <25 kg/m^2 (HR, 1.38; 95% CI, 1.02-1.87).

Table 7.1.11 Cigarette Smoking Status and the Risk of Colorectal Cancer in UK Biobank by Body Mass Index

	Cigarette smoking status			
BMI	**Never**	**Former**	**Current**	**P-interaction**
<25 kg/m^2				
Cases	305	121	56	
HR (95% CI)*	1.00	1.22 (0.98-1.52)	1.38 (1.02-1.87)	
25-<30 kg/m^2				
Cases	432	314	52	
HR (95% CI)*	1.00	1.31 (1.12-1.53)	1.02 (0.76-1.37)	
≥30 kg/m^2				
Cases	236	213	26	
HR (95% CI)*	1.00	1.20 (0.99-1.46)	0.87 (0.58-1.32)	0.3538

*Adjusted for sex, alcohol intake (never, former, special occasions only, one to three times a month, <5 g/d, 5-<15 g/d, 15-<30 g/d, 30-<45 g/d, ≥45 g/d), Townsend deprivation index (quintiles), red meat intake (1, >1-<3, ≥3 times/week), processed meat (<1, 1, >1 time/week), height (continuous), family history of colorectal cancer (yes/no) and history of colorectal screening (yes/no).

7.2 Discussion

7.2.1 Smoking Status and Colorectal Cancer

Tobacco smoking and cigarette smoking were associated with colorectal cancer risk in this analysis. The detail of the baseline questionnaire meant that different categories of tobacco exposure could be assessed in this analysis. Most analyses of tobacco smoking simply classify individuals as never, former or current smokers and thus combine slightly different levels of exposure without much consideration due to the limitations of the questionnaire. However, the current analysis was able to analyse exposure to tobacco smoke more precisely.

For example, compared to never ever smokers, participants who had smoked occasionally in the past or just tried smoking once or twice and had smoked less than 100 times did not have an increased risk of colorectal cancer, as expected. The question on smoking 100 times is often used to define a threshold of exposure for never smoking, given that many people will have tried smoking but will have a very low level of exposure.[379] Thus, in all further analyses in this section, these participants were combined with never ever smokers to form the reference group never smokers.

Both male and female former daily smokers had an increased risk of colorectal cancer compared to never ever smokers. However, evidence was weaker for an increased risk for daily smokers. Results were unclear for participants who were occasional smokers at baseline due to the relatively small number of cases among these participants. Also, there was very slight evidence that women who had smoked occasionally in the past and had smoked more than 100 times had an increased risk of colorectal cancer, despite the fact that these participants were never regular smokers.

Next, some of these categories were combined to create a simpler smoking status variable (never smokers, former smokers, current smokers), comparable to other analyses. Also, to provide further comparison, a further variable was created which focused on cigarette smoking status and restricted former smokers to former daily cigarette smokers and current smokers to current daily cigarette smokers.

Results for overall smoking status and cigarette smoking status were fairly similar. For women, there was a similar increased risk of colorectal cancer for former smokers and former cigarette smokers. There was no evidence of an increased risk for current smokers or current cigarette smokers. Results for current smokers and current cigarette smokers were also similar for men. However, there was a slight difference between former smokers and former cigarette smokers with a slightly greater risk of colorectal cancer for former cigarette smokers. Other studies should obtain more detailed information on smoking exposure and also provide more information for how analysis variables were derived.

Thus, this analysis found clear evidence that former cigarette smokers had an increased risk of colorectal cancer but not current cigarette smokers. Overall, it seems clear from existing studies that both former and current smokers have an increased risk of colorectal cancer. For example, a recent meta-analysis found that the pooled RR (95% CI) was 1.20 (1.10-1.30) for current smokers and 1.18 (1.12-1.25) for former smokers.[49]

It is not clear why current cigarette smokers did not appear to have an increased risk of colorectal cancer in this study. Individual studies have found contrasting results for former and current smokers. For example, former smokers (HR, 1.17; 95% CI, 1.07-1.29) also had a greater risk of colorectal cancer than current smokers (HR, 1.08; 95%

CI, 0.96-1.21) in the EPIC study.[253] However, the Cancer Prevention Study II Nutrition Cohort found a similar risk of colorectal cancer for former smokers (HR, 1.23; 95% CI, 1.11-1.36) and current smokers (HR, 1.27; 95% CI, 1.06-1.52)[254] and current smokers (HR, 1.28; 95% CI, 1.00-1.63) had a greater risk of colorectal cancer than former smokers (HR, 1.10; 95% CI, 0.97-1.24) in a cohort of female teachers.[258]

One possible explanation for the conflicting results is because of the varying definitions of smoking status between studies. Many epidemiological studies include only very basic questions on smoking exposure and smoking status is a fairly crude measure. For example, many studies will likely define both occasional and regular smokers as current smokers.

However, although varying definitions of smoking status may contribute to conflicting results between studies, it remains unclear why this study found a greater risk of colorectal cancer for former cigarette smokers than for current cigarette smokers despite the more precise definition of this variable.

7.2.2 Colorectal Cancer Subsites

Colon and Rectal Cancer

The relationships between cigarette smoking status and colon and rectal cancer risk were also investigated in this study. Results were slightly different for men and women. Male former cigarette smokers had an increased risk of colon cancer but there was no evidence of an increased risk for male current cigarette smokers. Male former cigarette smokers and male current cigarette smokers appeared to have a similar increased risk of rectal cancer. There was no evidence that female former cigarette smokers or current cigarette smokers had an increased risk of colon cancer while female former cigarette smokers had a particularly increased risk of rectal cancer.

Overall, evidence from prospective cohort studies supports a relationship between smoking and both colon and rectal cancer risk though smoking seems to be slightly more strongly related to rectal cancer risk.[49, 50] For example, in the most recent meta-analysis, the pooled RRs (95% CIs) for former smokers were 1.20 (1.11-1.30) for rectal

cancer and 1.16 (1.11-1.22) for colon cancer and the pooled RRs (95% CIs) for current smokers were 1.24 (1.16-1.39) for rectal cancer and 1.09 (1.01-1.18) for colon cancer.[50] Thus, similar to results for men in this analysis, former and current smokers have a similar risk of rectal cancer while former smokers have a greater risk of colon cancer than current smokers.

Parajuli et al. also presented results for colon and rectal cancer for men and women.[255, 256] Similar to the results for UK Biobank, male former smokers (HR, 1.28; 95% CI, 1.11-1.50) and current smokers (HR, 1.26; 95% CI, 1.09-1.45) had a similar increased risk of rectal cancer. Also, male former smokers had an increased risk of colon cancer (HR, 1.14; 95% CI, 1.02-1.27) but male current smokers did not have an increased risk (HR, 1.03; 95% CI, 0.92-1.15). However, results differed to UK Biobank for women. Parajuli et al. found an increased risk of colon cancer for both female former smokers (HR, 1.16; 95% CI, 1.02-1.31) and current smokers (HR, 1.22; 95% CI, 1.10-1.36) and an increased risk of rectal cancer for both female former smokers (HR, 1.26; 95% CI, 1.05-1.52) and current smokers (HR, 1.29; 95% CI, 1.10-1.51).

Other studies have also found slightly different results. For example, in the EPIC study, both former and current smokers appeared to have a greater risk of colon cancer than rectal cancer and former smokers appeared to have a greater risk than current smokers for both colon cancer and rectal cancer.[253] The HRs (95% CIs) were 1.21 (1.08-1.36) for former smokers and 1.13 (0.98-1.31) for current smokers for colon cancer and 1.10 (0.94-1.30) for former smokers and 0.98 (0.80-1.19) for current smokers for rectal cancer. While both former smokers and current smokers in the Cancer Prevention Study II Nutrition Cohort had an increased risk of colon cancer (HR, 1.19; 95% CI, 1.06-1.34 for former smokers and HR, 1.28; 95% CI, 1.04-1.57 for current smokers), only former smokers had an increased risk of rectal cancer (HR, 1.26; 95% CI, 1.02-1.55 for former smokers and HR, 0.97; 95% CI, 0.63-1.47 for current smokers).[254]

Proximal and Distal Colon Cancer

This analysis also investigated the relationships between cigarette smoking status and proximal colon cancer and distal colon cancer risk. Former cigarette smokers had an increased risk of distal colon cancer but not proximal colon cancer. It was unclear

whether current cigarette smokers had an increased risk of proximal colon cancer or distal colon cancer. Few studies have investigated the risk of proximal and distal colon cancer. The meta-analysis by Cheng et al. suggested that smoking is more strongly related to proximal colon cancer risk though results were based on only four studies.[50]

7.2.3 Smoking Duration and Intensity

Smoking Duration

Overall, studies generally find a positive association between smoking duration and colorectal cancer risk.[26, 29, 49, 226] However, most studies of smoking and colorectal cancer have combined former and current smokers when analysing smoking duration. This may lead to misleading results since current smokers will continue to add to their exposure during follow-up whereas the number of years since cessation will increase for former smokers. Also, the distributions of smoking duration are very different for former and current smokers e.g. very few current smokers will have a duration less than ten years.

Smoking duration was associated with colorectal cancer risk for former cigarette smokers in UK Biobank. Former cigarette smokers with ≥40 years duration had the highest risk of colorectal cancer. Compared to never smokers, there was evidence of an increased risk for all former cigarette smokers, including those with ≤9 years duration. Current cigarette smokers with ≥40 years duration also had an increased risk of colorectal cancer but the risk was much more moderate compared to the risk for former cigarette smokers with ≥40 years duration (in fact, results for current cigarette smokers with ≥40 years duration were similar to results for former cigarette smokers with ≤9 years duration).

The EPIC study also investigated smoking duration separately for former and current smokers.[253] However, there were no clear trends for former or current smokers. Compared to never smokers, former smokers who smoked 20-29 and ≥30 years both had an increased risk of colon cancer but so did former smokers who smoked ≤9 years. For rectal cancer, only former smokers who smoked 20-29 years had an increased risk and there was no evidence of an increased risk for former smokers who smoked ≥30

years. There was no evidence that current smokers with ≥40 years duration had an increased risk of colon cancer or rectal cancer.

In contrast to results from the EPIC study, the Cancer Prevention Study II Nutrition Cohort found an increased risk of colorectal cancer for current smokers who smoked 40-49 (HR, 1.32; 95% CI, 1.02-1.72) and ≥50 years (HR, 1.38; 95% CI, 1.04-1.84). There was no increased risk for current smokers with <40 years duration (HR, 1.02; 95% CI, 0.69-1.49).[254]

Smoking Intensity

Again, existing studies do show a positive association between smoking intensity and colorectal cancer risk.[26, 29, 49, 226] Similar to analyses of smoking duration, it may be important to analyse smoking intensity separately for former and current smokers since current smokers continue to accumulate exposure during follow-up. Another reason is that smoking intensity normally represents two slightly different measures for former and current smokers in most studies; while current smokers are asked to report how many cigarettes they currently smoke per day, former smokers are asked to report the average number of cigarettes they smoked per day in the past.

Former cigarette smokers who smoked 11-15, 16-20 and ≥21 cigarettes per day each had an increased risk of colon cancer and rectal cancer compared to never smokers in UK Biobank. However, while smoking ≥21 cigarettes per day carried the highest risk of colon cancer, smoking 11-15 cigarettes per day was associated with a higher risk of rectal cancer than smoking 16-20 or ≥21 cigarettes per day. Current cigarette smokers who smoked ≥21 cigarettes per day appeared to have a particularly high risk of rectal cancer (though the CI was wide) but there was no evidence of an increased risk of colon cancer.

The EPIC study also found that smoking intensity was associated with colon and rectal cancer risk for former smokers.[253] Former smokers who smoked 10-14 (HR, 1.30; 95% CI, 1.04-1.63) and ≥15 cigarettes per day (HR, 1.25; 95% CI, 1.03-1.53) both had an increased risk of colon cancer while only former smokers who smoked ≥15 cigarettes per day had an increased risk of rectal cancer (HR, 1.26; 95% CI, 0.98-1.61). In contrast

to UK Biobank results for current smokers, there was only evidence of an increased risk of colon cancer for current smokers smoking ≥20 cigarettes per day (HR, 1.20; 95% CI, 0.97-1.50).

The Physicians' Health Study also analysed smoking intensity separately for former and current smokers and found a much higher risk of colorectal cancer for current smokers who smoked ≥20 cigarettes per day (HR, 2.14; 95% CI, 1.45-3.14) than for former smokers who smoked cigarettes ≥20 per day (HR, 1.31; 95% CI, 1.00-1.73) (though CIs were quite wide).[324] In a cohort of Japanese men, the HR (95% CI) for current smokers smoking ≥20 cigarettes per day was 1.60 (0.99-2.58).[317]

Smoking Duration and Intensity

Studies of smoking and colorectal cancer have not investigated the effects of smoking duration and smoking intensity together. It is important to consider smoking duration and intensity together because the effects of smoking duration on colorectal cancer risk may be affected by smoking intensity and vice versa. For example, individuals who start smoking earlier will tend to have a longer duration than individuals who start later but they may also have a greater nicotine dependence and thus be heavier smokers.[273-275] Also, individuals who smoke fewer cigarettes are more likely to quit smoking and therefore will have a shorter smoking duration on average.[261-266] Furthermore, it is plausible that people with a long duration of heavy smoking will have a particularly elevated risk of colorectal cancer (yet this is ignored in most studies).

A number of studies have analysed the relationship between pack-years of smoking, which combines smoking duration and smoking intensity, and colorectal cancer. However, the pack-years variable is flawed and has received much criticism.[276, 277] The main issue with pack-years is that it confounds smoking duration and smoking intensity when it seems clear that these two aspects do not affect cancer risk in the same way.

Pack-years is also flawed due to the limitations of measuring smoking duration and smoking intensity. Pack-years is intended to be a measure of lifetime exposure. Yet current smokers are generally only asked to report their current smoking intensity. Thus,

it is necessary to assume that current smokers have always smoked this number of cigarettes, including when people report starting smoking at very young ages.

This analysis investigated the joint effects of smoking duration and smoking intensity among former cigarette smokers (current smokers were not included because of the relatively small number of cases). Former cigarette smokers who smoked for ≥40 years and smoked ≤15 or ≥16 cigarettes per day both had an increased risk of colorectal cancer compared to never smokers though there was a greater risk for participants who smoked ≥16 cigarettes per day. There was also evidence that participants who smoked ≥16 cigarettes per day for 10-39 years had an increased risk of colorectal cancer but there was no evidence of an increased risk for participants who smoked ≤15 cigarettes per day.

Thus, results from this analysis found that the association between smoking duration and colorectal cancer differed according to smoking intensity (and vice versa). Hence, more studies should try to classify people by both smoking duration and smoking intensity in order to better classify the additional risks due to smoking exposure. Surprisingly, this analysis also found an increased risk of colorectal cancer for former cigarette smokers with the lowest exposure. In fact, these participants appeared to have an approximately similar risk as participants who smoked ≥16 cigarettes per day for 30-39 years or ≤15 cigarettes per day for ≥40 years. Possible reasons for such an increased risk among these participants are unclear.

7.2.4 Smoking Initiation

There were no clear trends in the results for age at smoking initiation and colon cancer and rectal cancer risk in this analysis. For example, although former cigarette smokers with the earliest age at initiation (≤14) had the highest risk of colon cancer, there was also an increased risk for former cigarette smokers who started smoking ≥19 but not for former cigarette smokers who started smoking 17-18. For rectal cancer, there was a similar increased risk for each category of age at initiation. There was slight evidence that current cigarette smokers who began smoking ≤15 or 16-17 had an increased risk of rectal cancer but there was no evidence for an association between age at initiation and colon cancer risk.

Other studies also do not find clear trends for the relationship between smoking initiation and colorectal cancer. For example, in the EPIC study, there was an increased risk of colon cancer for former smokers who began smoking 17-19 or ≥20 but not for former smokers who began ≤16.[253] However, for rectal cancer, only former smokers who began smoking ≤16 appeared to have an increased risk. There was no clear association for current smokers.

Parajuli et al. investigated age at smoking initiation for men and women and found a clear trend for the associations with colon cancer but evidence for a trend was less clear for rectal cancer.[255, 256] Also, the Women's Health Initiative found a similar risk of colorectal cancer for women who started smoking before and after age 20.[257]

7.2.5 Smoking Cessation

Age at Cessation

Age at cessation and years since cessation were both considered in this analysis to investigate the effects of smoking cessation at different times on colorectal cancer risk. Participants who quit ≥50 years old had an increased risk of both colon and rectal cancer. However, there also seemed to be an increased risk of colon and rectal cancer for participants who quit ≤29 years old.

Few other studies have investigated age at cessation but results seem to indicate that there is an increased risk of colorectal cancer compared to never smokers only for former smokers who quit smoking ≥40 years old.[254, 257, 260] For example, the HRs (95% CIs) in the Cancer Prevention Study II Nutrition Cohort were 1.05 (0.91-1.22), 1.31 (1.13-1.52), 1.44 (1.24-1.66) and 1.29 (1.08-1.54) for quitting smoking <40, 40-49, 50-59 and ≥60 years old.[254] Thus, it was slightly surprising that former smokers who quit ≤29 years old in UK Biobank appeared to have an increased risk of colorectal cancer compared with never smokers.

Years since Cessation

In contrast to the results for age at cessation, there were clearer trends for the associations between years since cessation and colon and rectal cancer. Interestingly, participants who quit smoking ≥30 years ago appeared to still have an increased risk of rectal cancer (but not colon cancer) compared to never smokers.

There was a clear trend between years since cessation and colorectal cancer risk in the Cancer Prevention Study II Nutrition Cohort with no evidence of an increased risk for former smokers who quit more than 30 years before baseline; the HRs (95% CIs) for ≥31, 21-30, 11-20 and 1-10 years since cessation were 1.03 (0.89-1.19), 1.28 (1.10-1.49), 1.33 (1.14-1.55) and 1.48 (1.27-1.73).[254] In the pooled analysis by Gong et al., there was evidence of an increased risk of colorectal cancer for former smokers who quit 0-15 and 15-25 years ago but not for former smokers who quit 25-35 or ≥35 years ago. [260]

The EPIC study did not find a clear relationship between years since smoking cessation and the risk of colon cancer or rectal cancer.[253] For example, although there was no evidence of an increased risk of colon or rectal cancer for former smokers who quit 20-24 or ≥25 years ago, former smokers who quit 15-19 years ago had the highest risk of colon cancer (HR, 1.36; 95% CI, 1.09-1.70) and rectal cancer (HR, 1.30; 95% CI, 0.96-1.76).

Years since Cessation and Smoking Intensity

This analysis also investigated the combined effects of years since cessation and smoking intensity. Participants with the highest exposure (quit smoking ≤9 years before baseline and smoked ≥16 cigarettes per day) had the greatest risk of colon and rectal cancer. However, there were slightly different patterns of results for colon and rectal cancer. Participants who smoked ≥16 cigarettes per day seemed to have an increased risk of colon cancer if they quit ≤29 years before baseline and participants who smoked ≤15 cigarettes per day seemed to have an increased risk of colon cancer only if they quit ≤9 years before baseline. For rectal cancer, while there was an increased risk for participants who smoked ≥16 cigarettes per day and quit ≤19 years before baseline,

among participants who smoked ≤15 cigarettes per day, there was only an increased risk for participants who quit ≥20 years before baseline. Similar to the analysis of smoking duration and intensity, it is unclear why former cigarette smokers with supposedly the lowest exposure should have an increased risk of rectal cancer compared to never smokers.

7.2.6 Smoking and Body Mass Index

On average, smokers have a lower BMI than non-smokers, likely as a result of the effects of smoking on metabolism as well as suppressing appetite.[227] However, heavy smokers also tend to have a higher BMI than light smokers and non-smokers, most likely due to a clustering of negative behaviours (e.g. low physical activity, poor diet and high alcohol intake).[227] Since BMI represents a major risk factor for colorectal cancer, it may be important to consider the effects of smoking and BMI together.

This analysis found a fairly similar increased risk for former cigarette smokers within each category of BMI, compared to never smokers. However, there was a clear difference for current cigarette smokers as there was an increased risk only for current cigarette smokers with BMI <25 kg/m^2. Furthermore, within participants with BMI <25 kg/m^2, current cigarette smokers seemed to have a higher risk of colorectal cancer than former cigarette smokers.

One possible explanation for a higher risk for current cigarette smokers with lower BMI than for current cigarette smokers with higher BMI could be related to insulin resistance. Higher insulin resistance is thought to be an important factor linking BMI and colorectal cancer risk.[157, 185] Although smoking is associated with a lower BMI, smoking may actually increase insulin resistance.[227] Thus, compared to never smokers with a lower BMI, current cigarette smokers with a lower BMI may be at an increased risk of colorectal cancer due to the direct negative effects of smoking as well as the indirect effects through insulin resistance. In contrast, never smokers with a higher BMI will have an increased insulin resistance and so the effects of current smoking on insulin resistance may be less important.

Only a few studies have previously presented results for smoking and BMI. A cohort of female teachers found no clear difference in results comparing ≥31 pack-years with never smokers for women with BMI <25 kg/m^2, 25-29 kg/m^2 and BMI ≥30 kg/m^2.[258] Results from the pooled analysis by Gong et al. supported a stronger effect of smoking on colorectal cancer risk for people with higher BMI and Parajuli et al. also supported a stronger effect of smoking on rectal cancer risk for people with higher BMI.[256, 260] However, both of these analyses compared ever vs never smokers which makes results more difficult to interpret since former smokers tend to gain weight after cessation.[227]

7.2.7 Former and Current Smokers

Perhaps the biggest remaining question about the relationship between smoking and colorectal cancer is why former smokers appear to have a similar (or even greater) risk than current smokers. Yet, surprisingly, there is very little discussion about these results in most studies of smoking and colorectal cancer.

One possible explanation why current smokers do not have a greater risk than former smokers could be because of confounding. However, this seems unlikely since although former cigarette smokers had a slightly higher BMI than current cigarette smokers on average in UK Biobank, current cigarette smokers had a higher intake of alcohol and red and processed meat, were more likely to live in areas of high deprivation and were less likely to have undergone bowel cancer screening. Another possible explanation could be that a number of former smokers have a particularly elevated risk of colorectal cancer because they quit smoking as a result of ill health. In contrast, the fact that current smokers continue to smoke could be considered a marker of reasonable health. However, current cigarette smokers reported the highest levels of poor health in this cohort. Furthermore, results for former cigarette smokers did not appear to differ according to reasons for quitting smoking (though this question was very crude).

The similar risks for former and current smokers may possibly be related to the long induction period for the effects of smoking on colorectal cancer hypothesised by Giovannucci et al., meaning that the effects of smoking persist for many years, even after cessation.[50, 249, 252, 269]

Though smoking is only associated with a moderately increased risk of colorectal cancer, smoking has been shown to be strongly related to the risk of colorectal adenomas. Meta-analyses by Botteri et al. found that the pooled RR (95% CI) for current vs never smokers was 2.14 (1.86-2.46) for colorectal adenomas and 1.07 (0.99-1.16) for colorectal cancer.[226, 281] Since colorectal adenomas are precursors to the majority of colorectal cancers,[19, 280] the risk of cancer would be expected to be at least as large as the risk for adenomas.[282] Thus, to explain this discrepancy, Giovannucci et al. hypothesised that smoking mainly affects the early stages of colorectal carcinogenesis and that an induction period of 30-40 years is required to observe an increased risk of cancer.

First of all, it is not clear that this hypothesis does explain the greater risk of colorectal adenomas than colorectal cancer. In UK Biobank, there was still only a modest increased risk for current smokers with more than 40 years duration. More pertinently, a long induction period does not seem to explain why current smokers do not have a greater risk of colorectal cancer than former smokers.

Even if smoking mainly affects earlier stages of colorectal carcinogenesis, current smokers should still be expected to have a greater risk of colorectal cancer. Both former and current smokers will have started smoking at a similar age (median age at initiation in UK Biobank was 17 for former cigarette smokers and 16 for current cigarette smokers). However, while current smokers will have continued to accumulate a number of decades of smoking, a number of former smokers will have quit relatively early and so would be expected to have a low risk of colorectal cancer. In other words, current smokers should always be expected to have a greater level of exposure than former smokers. The only way this could be offset is if former smokers tended to be much heavier smokers than current smokers though this seems unlikely since heavier smokers are less likely to quit smoking.[261-263]

Furthermore, a long induction period does not explain why the most recent quitters had a much higher risk of colorectal cancer than current smokers in UK Biobank despite having a similar level of exposure up to baseline, or why former smokers with ≥40 years duration had a much higher risk than current smokers with ≥40 years duration.

Protective Effect?

So why do former and current smokers have a similar risk of colorectal cancer? And why do current smokers not have a similar risk to recent quitters? A similar risk for former and current smokers could be explained if, although smoking promoted early colorectal carcinogenesis, smoking also had some protective effect during later stages of carcinogenesis. Individuals with ulcerative colitis are at a substantially increased risk of colorectal cancer[19, 436, 437] and it has been found that smoking reduces the risk of ulcerative colitis.[438, 439] Hence, it may be plausible that current smoking does have some protective effect on colorectal cancer risk.

This hypothesis would explain why current smokers do not have a greater risk than former smokers and also why the most recent former smokers have the highest risk of colorectal cancer. While recent quitters and current smokers have a similar exposure up to baseline, the harmful effects of smoking on early carcinogenesis could be offset for current smokers if active smoking offered some form of protection against late stages of carcinogenesis. Furthermore, this hypothesis would also explain why smoking is more strongly related to the risk of colorectal adenomas than colorectal cancer.

This hypothesis could also explain why results from meta-analyses suggest that the risk of colorectal cancer for current smokers increases with longer follow-up.[50, 61] This result is surprising since more current smokers will become former smokers with longer follow-up. However, if current smoking offered some form of protection against colorectal cancer, it could make sense that the risk increased as more current smokers stopped smoking.

7.2.8 Summary

In conclusion, this analysis found that smoking increases the risk of colorectal cancer risk. While there was an increased risk for former cigarette smokers, it was not clear whether current cigarette smokers in UK Biobank had an increased risk of colorectal cancer. Only male former cigarette smokers had an increased risk of colon cancer. For rectal cancer, there was an increased risk for male former and current cigarette smokers and female former cigarette smokers. There seemed to be dose-response relationships

for smoking duration and years since smoking cessation though results were less clear for smoking intensity, initiation and age at cessation. Former cigarette smokers with the shortest duration or longest time since cessation still seemed to have a higher risk of colorectal cancer than never smokers. Results for analyses combining smoking intensity and smoking duration/years since cessation were unclear since there seemed to be an increased risk for participants with the lowest level of exposure. Despite no evidence for an increased risk overall, current cigarette smokers with BMI <25 kg/m^2 did seem to have an increased risk of colorectal cancer compared to never smokers with BMI <25 kg/m^2. It is unclear why former cigarette smokers had a higher risk of colorectal cancer than current cigarette smokers or why former cigarette smokers with ≥40 years duration had a much higher risk than current cigarette smokers with ≥40 years duration.

7.3 Population Attributable Fractions

Table 7.3.1 shows the population attributable fraction estimates (presented as percentages) and 95% CIs for different risk factors, overall and for men and women separately. 378,263 participants (178,468 men and 199,795 women) were included in this analysis and there were 1,833 cases (1,096 men and 737 women).

Table 7.3.1 Population Attributable Fractions for Different Risk Factors in UK Biobank

	PAF (%) (95% CI)		
Exposure	**Overall**	**Men**	**Women**
Alcohol intake, >14 units/week	6.62 (2.44, 10.61)	12.69 (5.89, 19.01)	-0.39 (-4.46, 3.52)
BMI, ≥25 kg/m^2	9.82 (2.65, 16.46)	16.43 (5.74, 25.91)	3.54 (-6.03, 12.25)
Smoking, ever smoker	7.95 (3.22, 12.46)	9.79 (2.71, 16.36)	5.09 (-1.05, 10.86)
Townsend index, >lowest quintile	-2.86 (-8.01, 2.04)	-0.67 (-7.29, 5.54)	-6.12 (-14.51, 1.67)
Red meat intake, >once/week	8.68 (3.18, 13.87)	11.10 (3.35, 18.22)	5.69 (-2.17, 12.95)
Processed meat intake, >once/week	2.48 (-0.94, 5.79)	6.52 (1.10, 11.64)	-2.12 (-5.64, 1.27)
Family history, yes	3.52 (1.61, 5.39)	4.29 (1.74, 6.78)	2.34 (-0.54, 5.15)
Physical activity, <75/150 mins moderate/vigorous activity/week	0.69 (-2.87, 4.12)	-0.81 (-5.41, 3.60)	2.68 (-3.03, 8.08)

For alcohol intake, it was estimated that 6.62% of cases were due to alcohol intake >14 units/week i.e. 6.62% of cases in this cohort could have been avoided if all participants drank ≤14 units/week. Estimates were very different for men and women since there was no evidence that alcohol intake was associated with colorectal cancer for women in

this cohort; the PAF estimate was 12.69 for men and -0.39 for women. It was estimated that 9.82% of cases were attributable to participants having BMI ≥25 kg/m^2 (16.43% for men and 3.54% for women) and that 7.95% of cases were attributable to participants being ever smokers (9.79% for men and 5.09% for women).

8.68% of cases could have been avoided if participants reduced their intake of red meat to ≤once/week and reducing processed meat intake to ≤once/week was estimated to reduce the number of cases for men (6.52%) but not for women (-2.12%). 3.52% of cases were attributable to family history of colorectal cancer.

Chapter 8 Discussion

This thesis investigated the relationship between three classic lifestyle risk factors and colorectal cancer. Alcohol intake, adiposity and smoking represent three of the most important modifiable risk factors for a wide range of NCDs, that are currently the leading cause of death worldwide.[3] Yet these diseases are often the result of unhealthy lifestyles and are largely preventable. For example, it is predicted that one in two people in the UK born after 1960 will be diagnosed with cancer in their lifetime.[7] However, based on existing evidence, it is also predicted that approximately 40% of cancers are attributable to lifestyle factors.[30] Therefore, research to improve our understanding of how these risk factors contribute to disease is highly important.

The main objectives for this thesis were described in section 2.5. The findings for these objectives are summarised below, separately for alcohol intake, adiposity and smoking.

8.1 Alcohol Intake and Colorectal Cancer

Results from different studies have found evidence for a threshold for the association between alcohol intake and colorectal cancer with an increased risk only for people drinking at least approximately 30 g/d.[45, 65, 66, 69, 70, 409] In contrast, this thesis found a dose-response relationship between alcohol intake and colorectal cancer risk for men with evidence of an increased risk including for men drinking 5-<15 g/d (when using the alternative reference group and multiple imputed data). Furthermore, fractional polynomial analysis did not find evidence for a non-linear relationship between alcohol intake and colorectal cancer. Hence, results from this thesis did not support a threshold for alcohol intake and suggest that any alcohol intake should be minimised to prevent colorectal cancer, at least for men.

In contrast to the results for men, this thesis did not find evidence for an association between alcohol intake and colorectal cancer for women (whether analysing alcohol intake as a categorical or continuous variable). It was unclear why this thesis found an association for men but not for women, particularly when other studies have not found evidence for a stronger association for men.[45, 65, 66, 69] One possible explanation could be because men and women tend to prefer different alcoholic beverages and there was

evidence that beer intake but not wine intake was associated with colorectal cancer risk in UK Biobank. However, it is not clear if this provides a valid explanation since most other studies do not find differences according to alcoholic beverages.[45, 65, 69, 70] Also, it may not be surprising that these studies do not find differences by alcoholic beverages since they do not find evidence for different associations for men and women despite their different preferences for alcoholic beverages.

Few studies have investigated the association between alcohol intake and colorectal cancer using different reference groups despite the fact that the choice of reference group is a controversial issue in analyses of alcohol intake.[46] Given the detail available in this cohort, it was possible to investigate different choices of reference group and it was hypothesised that the association between alcohol intake and colorectal cancer could be underestimated in many studies that used non-drinkers as the reference group since former drinkers would likely have an increased risk of colorectal cancer compared to never drinkers. This could potentially explain differences in results between studies e.g. why this thesis seemed to find a stronger association between alcohol intake and colorectal cancer for men than other studies of men conducted in Western populations or why there was no evidence of a threshold for men. However, former drinkers did not have an increased risk of colorectal cancer in this study compared to never drinkers and results were similar using never drinkers or non-drinkers as the reference group.

The main disadvantage of using never drinkers as the reference group is that they represent a small group and so include a relatively small number of cases of colorectal cancer which leads to wide CIs. Therefore, the association between alcohol intake and colorectal cancer was also investigated using an alternative reference group which combined never drinkers and select light drinkers who reported not reducing their drinking since ten years before baseline. (Ideally, it would have been possible to identify participants who had never been heavy drinkers and who had never been binge drinkers but the data available on past drinking and pattern of drinking were limited.) Results for men using this alternative reference group were fairly similar to results using never drinkers as the reference group; the HRs were slightly attenuated though there was still a clear trend between alcohol intake and colorectal cancer risk. For women, there was still no evidence for an association.

This thesis did not find a clear difference in the association between alcohol intake and colorectal cancer risk by BMI. Hence, these results were not in agreement with results from the Western pooled analysis and the Japanese pooled analysis which showed a stronger association between alcohol intake and colorectal cancer for people with lower BMI.[45, 71] Future studies should investigate the relationship between alcohol intake and colorectal cancer by BMI.

8.2 Adiposity and Colorectal Cancer

This thesis found a strong association between BMI and colon cancer risk for men but no association for women. There was no clear evidence for an association with rectal cancer risk for men or women. Numerous other studies have also found that BMI is more strongly associated with both colon cancer risk and rectal cancer risk for men than for women.[150] However, the reasons for this difference remain unclear. One hypothesis for the difference is due to the different body compositions of men and women. Visceral fat, which is stored abdominally, may be particularly important for colorectal cancer risk and the different associations for men and women may be because BMI is more strongly correlated with abdominal fat for men than for women.[47, 185]

However, results from this thesis did not appear to support this hypothesis. Although there was slight evidence that WC and WHR were associated with colon cancer risk for women, WC and WHR were still much more strongly related to colon cancer risk for men than for women. Other studies investigating BMI, WC and/or WHR also did not support the hypothesis that the difference was due to different body compositions.[164, 172, 173, 186-188] Hence, it seems that the different body compositions for men and women do not explain the difference in associations, at least based on the evidence for WC and WHR. However, these abdominal measures of adiposity still do not represent precise measures of visceral fat[420] and hence the different body compositions of men and women may still contribute to the different associations for men and women.

The other possible reason investigated in this thesis for the different associations in men and women was because of the effects of menopause on oestrogen levels which may be protective against colorectal cancer. Meta-analysis results do indicate that the association between BMI and colorectal cancer is stronger for pre-menopausal women

than post-menopausal women.[152] However, this thesis found no evidence of an association for pre-menopausal women or for post-menopausal women who were never HRT users. Not many studies have included results by menopause status but there are other studies that also do not support a stronger association for pre-menopausal women than for post-menopausal women.[161, 169] Thus, further research is required to understand how menopause and oestrogen may influence the association between BMI and colorectal cancer.

Although results for BMI, WC and WHR in this thesis did not appear to explain the discrepancy between men and women for the association between BMI and colon cancer risk, results did seem to suggest that WC and WHR may be more directly associated with colon cancer risk than BMI. First of all, WC and WHR seemed to be associated with colon cancer risk for women whereas there was no association for BMI. Furthermore, when modelling BMI and WC or WHR simultaneously, the associations for WC and WHR remained while the associations for BMI were severely attenuated. Also, other studies found similar results when adjusting analyses of WC/WHR for BMI.[47, 186, 187]

8.3 Smoking and Colorectal Cancer

Compared to never smokers in UK Biobank, former cigarette smokers had an increased risk of colorectal cancer but there was no increased risk for current cigarette smokers. Other studies have also shown a slightly greater risk for former smokers than current smokers.[226, 252] However, the similar risks for former and current smokers are rarely discussed in studies of smoking and colorectal cancer and the reasons why current smokers do not have a greater risk are unclear. This thesis tried to consider different possible explanations. For example, it seems unlikely that results were due to confounding since results in this thesis were adjusted for a number of confounders associated with colorectal cancer.

Another reason could be related to the overall health status of former and current smokers i.e. quitting smoking may be related to higher levels of disease whereas current smokers, since they continue to smoke, may find themselves in reasonable health. However, there was no evidence in this thesis to support this idea since former cigarette

smokers did not report higher levels of long-standing illness or poor health and also former cigarette smokers who quit for illness related reasons did not appear to have a greater risk than former cigarette smokers who quit for other reasons.

Finally, another reason considered was because the effects of smoking may persist for many years.[50, 252] However, this does not explain why current cigarette smokers with 40 years duration appeared to have a much lower risk of colorectal cancer than former cigarette smokers with a similar duration in this analysis. Thus, the reason why current smokers do not have a greater risk of colorectal cancer than former smokers remains an important unanswered question for future studies.

This thesis also investigated the effects of smoking duration and smoking intensity on the risk of colorectal cancer and, in particular, tried to investigate how these measures interact for colorectal cancer risk. Results did suggest that former cigarette smokers with the highest exposure i.e. the longest duration and heaviest intensity had the greatest risk of colorectal cancer. However, the overall pattern of risk was not very clear and former cigarette smokers with the lowest exposure also appeared to have a particularly high risk of colorectal cancer.

Results for age at cessation were not clear in this analysis. While there seemed to be no increased risk for former cigarette smokers who quit smoking 30-39 years old, former cigarette smokers who quit ≤29 years old did appear to have an increased risk of colorectal cancer. This was in contrast to results from other studies which did not find an increased risk for former smokers who quit before 40 years old.[254, 257, 260] The pattern of results was slightly clearer when investigating years since cessation. However, there was still evidence that former cigarette smokers who quit at least 30 years ago had a slightly increased risk of colorectal cancer compared to never smokers though other studies seem to indicate that former smokers with the longest time since cessation no longer had an increased risk.[253, 254, 260]

8.4 Strengths and Limitations

The analyses carried out in this thesis had a number of strengths and limitations. First of all, UK Biobank was a prospective cohort study, meaning that participants were asked

to provide data on risk factors before being followed-up using health records to discover which participants suffered particular diseases. This is very important when attempting to draw causal inferences from epidemiological studies since the outcome may influence participants' recall of exposure (known as recall bias) in retrospective studies, leading to biased results.[54, 55]

Very few other studies combine such a large number of participants with detailed information on a wide range of health and lifestyle factors. This meant that this thesis was able to investigate the associations between these risk factors and colorectal cancer in greater detail than many other studies. For example, in contrast to many existing studies of alcohol intake and colorectal cancer which have used non-drinkers as the reference group, this thesis was able to separately identify never and former drinkers and therefore compare results using different reference groups. Also, the questions on smoking in UK Biobank were very detailed in comparison to many other studies which often include only very basic questions in order to identify current smokers and separate never smokers from former smokers.

Since UK Biobank included detailed information on health and lifestyle, it was possible to adjust for a number of confounders known to be important for colorectal cancer risk, thus reducing the possibility that the associations observed between the risk factors of interest and the outcome were due to confounding by other factors. Analyses in this thesis, however, did not adjust for physical activity, despite the fact that physical activity is associated with colon cancer and so represents an important confounder, [25, 225] due to a large number of participants who chose not to answer the questions on physical activity. However, results with and without adjustment for physical activity were found to be very similar (for participants with complete data on physical activity), suggesting that the exclusion of physical activity from analysis models did not have a large effect on results.

Establishing a cohort study with a wide range of information on a large number of participants can be very expensive and time-consuming. It can also take many years before sufficient numbers of cases accumulate. Routinely collected data (e.g. from general practitioners or from hospital admissions) are increasingly being used in epidemiological research as they allow easy access to data on a very large number of

people who have often been followed-up for many years. However, since these data are generally collected for purposes other than epidemiological research, the use of routine data for epidemiological research has important limitations.[440, 441] For example, the data available from routine sources can be very limited and important information may not be available (e.g. it may not include data on a number of important confounders). Furthermore, it is not always clear exactly how data are generated and there may be important variation in coding between different individuals/organisations.

While UK Biobank did include detailed information on alcohol intake, adiposity and smoking, these data also had certain limitations. For example, there was no detailed information about alcohol intake, adiposity or smoking throughout life. Furthermore, the use of questionnaires may not provide an accurate picture of people's average exposure since questionnaires generally include fairly basic questions which are often poor measures of the exposure of interest. For example, questionnaires on alcohol intake generally ask people to report their average alcohol intake. However, when answering such questions, people tend to ignore atypical or occasional drinking episodes, which is one reason why alcohol intake is underestimated using self-reported data.[357] Another example is the large amount of digit preference for multiples of ten when participants are asked to report the average number of cigarettes they smoke per day.[380, 381]

A strength of these analyses was that the adiposity variables were measured at baseline. This was important because relying on self-reported height and weight is known to underestimate BMI which can lead to biased results.[48] However, analyses of alcohol intake and smoking relied on self-reported data which meant that alcohol intake and smoking prevalence may have been underestimated in UK Biobank.[352, 378]

UK Biobank included only a small proportion of current cigarette smokers which limited the analyses that could be carried out among these participants. For example, it was difficult to evaluate the effects of smoking duration and smoking intensity among current cigarette smokers. In 2008, 21% of men and 20% of women in England reported smoking.[442] In UK Biobank, only 12.5% of men and 8.9% of women reported smoking (either daily or occasionally). Assuming that a representative sample of the population was invited to participate in UK Biobank, this suggests that smokers were much less

likely to participate than non-smokers. Previous studies have also demonstrated that smokers are less likely to take part in surveys and epidemiologic studies.[443-445]

The low response rate achieved by UK Biobank meant that participants were highly self-selected and therefore unlikely to be representative of the greater population. High levels of self-selection may lead to the inclusion of fewer people with unhealthy lifestyle behaviours or ill health and this may be a concern for the validity of the results in this thesis. On average, compared to the general population, UK Biobank participants lived in areas with less deprivation, were more likely to have a healthy lifestyle (though they were more likely to be heavy alcohol drinkers) and were less likely to have poor health (see Chapter 3). Representativeness is not necessarily an important issue for the investigation of associations in cohort studies i.e. the proportion of smokers is unimportant when investigating how smoking relates to the outcome of interest. However, the self-selection of participants may still be a concern for results if, for example, there were major differences between participating and non-participating smokers.

Another limitation of this analysis was that a large number of participants had missing data on alcohol intake due to changes to the alcohol questions during recruitment. In a sensitivity analysis, multiple imputation was used to impute alcohol intake values for these participants. Overall, results using multiple imputation were very similar to the complete case results. This was not surprising since only participants who reported drinking "one to three times a month" or "special occasions only" had missing data. Thus, the only major difference was in the estimate of the HR associated with alcohol intake <5 g/d. These results were presented as a sensitivity analysis since among participants who reported drinking "one to three times a month" or "special occasions only", the majority actually had missing data.

Instead of using multiple imputation in the analyses of adiposity and smoking, it was decided to use an alternative analysis variable to adjust for alcohol intake in the analysis models. Using multiple imputation for multiple analyses requires a much more complex imputation model and can potentially lead to biased results if this imputation model is not correctly specified.[375] Furthermore, this imputation model would need to include a large number of variables which would reduce the number of participants with complete

data on all independent variables in the imputation model. This issue is normally resolved by imputing data for more variables however this can lead to biased results if missing data are not missing at random, which may be likely when participants choose not to answer.[376] Thus, given these limitations of multiple imputation and given that the alternative analysis variable was shown to give practically identical results as the original grams per day variable (among participants with complete data), it was decided to use the alternative variable for analyses of adiposity and smoking. Furthermore, based on the imputation model used for the sensitivity analysis, it was possible to calculate basic results for BMI and overall smoking status using the imputed alcohol intake values and these results showed no important differences to the results presented in this thesis (see Table A-14 and Table A-15 in the appendix).

Follow-up data on cancers and deaths are periodically updated as UK Biobank receives data from the cancer and death registries. This presented a challenge for this thesis since results (and consequently conclusions) could change significantly as further data were made available. This was made more difficult as there were significant delays in these follow-up data becoming available which impeded the progress of analyses. Ultimately, the analyses in this thesis were able to use follow-up data complete up to the end of March 2014. However, these data only became available in May 2016 and prior to this complete outcome data were only available up to the end of December 2012.

Also, UK Biobank remains a fairly young cohort in terms of follow-up; median follow-up in these analyses was approximately 5 years. With longer follow-up, more participants will be diagnosed with colorectal cancer and the associations investigated in this thesis can be evaluated with greater precision.

Bowel screening programmes in England, Scotland and Wales were introduced during recruitment of UK Biobank which may have impacted on results found in this thesis. First of all, it will have resulted in "misclassification" of a number of participants with respect to bowel screening since participants' responses will have depended on whether they had been invited yet; many participants who reported not having bowel screening at baseline may have been screened shortly after.

The introduction of bowel cancer screening programmes may also lead to an increase in colorectal cancer incidence in the short term through the increased detection of prevalent cases.[13, 22] This should not impact on the associations between risk factors and colorectal cancer if all participants are equally likely to accept the screening invitation. However, people who follow healthy lifestyle behaviours may be more likely to undergo screening than people with unhealthy lifestyles.[446] This would result in a greater increase of colorectal cancer incidence among people following a healthy lifestyle relative to people with an unhealthy lifestyle, making the incidence rates in the two groups more similar, at least in the short term.

A large number of analyses were presented in this thesis. The issue with performing large numbers of statistical tests is that the probability of finding a statistically significant result is greatly increased. Therefore, some of the results presented in this thesis may simply be due to chance and so results should be treated with a certain amount of caution. However, many of the results in this thesis are in agreement with results from other studies which provides some external validity to these results.

8.5 Further Research

UK Biobank included information on a wide range of exposures at the baseline assessment. Since baseline, UK Biobank has continued to broaden this information with further data on participants and some of these data could be used in future analyses. For example, approximately 20,000 participants were completely re-assessed between 2012 and 2013. The purpose of this re-assessment was in order to adjust analyses for regression dilution.[53, 447] During follow-up, participants' level of exposure to different risk factors will change and as a result the association between baseline exposure and disease risk generally underestimates the association between "usual" exposure and disease risk.[447] However, this can be addressed using data on changes in exposure in a subset of participants. These data were not used to adjust for regression dilution in this study since there was not sufficient follow-up data in relation to when the re-assessment was carried out. Participants were re-assessed approximately five years after the baseline assessment on average and it is recommended that the re-assessment should take place at approximately the midpoint of follow-up (i.e. approximately ten years of follow-up are required before these data can be used to adjust for regression dilution).[447]

UK Biobank is also completing MRI imaging scans of 100,000 participants.[53] This will allow precise measures of fat distribution throughout the body including visceral fat. Understanding why certain measures of adiposity are more strongly related to colorectal cancer than others and why these measures tend to be more strongly related to colorectal cancer for men than for women remain important questions. Fat distribution may be important for understanding how adiposity relates to colorectal cancer risk yet previous prospective studies have not been able to acquire imaging data on a large number of participants.

Genetic data will also soon be available for the entire UK Biobank cohort. The investigation of how genetic factors interact with lifestyle and environmental factors to cause disease was an important aim of UK Biobank.[53] These data could potentially be used to understand how an individual's genotype interacts with lifestyle factors to modify the risk of colorectal cancer and thus to improve understanding of the mechanisms relating lifestyle factors to colorectal cancer.

It may also be possible to obtain genetic data on colorectal cancers for UK Biobank participants in the future in order to classify tumours by molecular subtypes. These data could be important to improve the understanding of how smoking increases the risk of colorectal cancer since smoking appears to be strongly related to a particular molecular subtype of colorectal cancer.[272, 289] These data may also help to shed light on why smoking is more strongly related to colorectal adenomas than colorectal cancer.[296, 297]

There are also a number of other important remaining research questions that cannot be investigated using the UK Biobank cohort. For example, the relationship between alcohol intake and colorectal cancer may depend on the pattern of alcohol intake. There is evidence that the effect of alcohol intake on coronary heart disease depends on the pattern of alcohol intake i.e. binge drinking was associated with a greater risk compared to more regular intake.[412] However, no existing studies of alcohol intake and colorectal cancer have included detailed information on pattern of drinking. Therefore, future studies should include questions on the pattern of alcohol intake.

Another important research question is how exposure to alcohol intake, adiposity and smoking throughout life influence an individual's risk of colorectal cancer. Information

on exposure throughout life may be particularly important since there is a long induction period for colorectal cancer yet most cohort studies assess exposure only during late adulthood. Thus, few studies have analysed exposure throughout the lifetime in relation to colorectal cancer. Some studies have investigated the relationship between early adulthood BMI and colorectal cancer though these studies generally relied on participant's recall to evaluate past BMI which may be not be very accurate.[195] However, acquiring exposure data on participants periodically from early adulthood would represent a significant cost and significant numbers of disease cases would not accumulate for many years. Most cohort studies focus on recruiting participants in mid-late age (e.g. 40-70 years old) so that a large number of cases can accumulate in a relatively short period of time but that only a small number of participants will have already suffered major illnesses. It is possible that future cohort studies will use online questionnaires to recruit people and to obtain information on exposures at multiple times in life at a much lower cost.

Further research could also be carried out to evaluate different methods for capturing data on exposure. Self-reported measures are commonly used in epidemiological studies. Such measures are known to have limitations and tend to underestimate actual exposure yet most cohort studies continue to use very similar questionnaires.[48, 353, 378] For example, including questions on atypical alcohol intake may improve the accuracy of alcohol intake questionnaires.[357] The use of diaries may provide a method for obtaining more detailed information on exposure and could possibly become more widespread since they could be easily completed online by participants at different times. Also, other methods may be developed that are able to reliably measure exposure indirectly e.g. transdermal alcohol sensors (which measure the amount of alcohol that is excreted through the skin via perspiration)[448, 449] which could possibly be employed in large epidemiological studies in a similar way to physical activity monitors.

8.6 Awareness of the Effects of Lifestyle on Cancer

This thesis clearly showed that alcohol intake, adiposity and smoking are all associated with the risk of colorectal cancer. Furthermore there is convincing evidence that these risk factors are associated with multiple forms of cancer.[26, 450, 451] Hence, a significant proportion of cancers could be prevented by reducing exposure to these risk factors in

the general population.[30] However, despite this potential for prevention, the awareness of the relationship between lifestyle and cancer is alarmingly low (except for smoking).[452, 453] For example, a survey commissioned by Cancer Research UK found that only a third of people identified "drinking alcohol frequently" as a risk factor for cancer.[452] Also, a survey by the WCRF found that 41% of people were unaware that being overweight increases the risk of cancer.[453] Therefore, awareness of the effects of lifestyle on cancer should be promoted.

8.7 Public Health and the Prevention of NCDs

The prevention of NCDs represents a major public health challenge.[3, 39, 454] The WHO aims to achieve a 25% relative reduction worldwide in overall mortality from major NCDs (cardiovascular diseases, cancer, diabetes and chronic respiratory diseases) by 2025.[3] Alcohol intake, adiposity and smoking are all important modifiable risk factors for a wide range of NCDs.[3, 39] Hence, there is an urgent need to reduce the prevalence of these risk factors in the general population in order to prevent these diseases.

A key role of public health is to provide information. There have been numerous public health messages promoting healthy lifestyles in an attempt to modify people's behaviour but it seems that public health messages alone are largely ineffective at influencing behaviour. One reason may be due to the numerous conflicting messages about lifestyle and health regularly observed in the media. For example, there are many conflicting stories about the effects of alcohol on health.[455] It sometimes seems as though there are just as many stories reporting that alcohol prevents cancer than stories reporting that alcohol causes cancer.[456-458] These mixed messages clearly cause confusion and may cause people to become sceptical about epidemiological research and public health messages. Researchers themselves are to blame in some circumstances since press releases are generally written with the intention of maximising interest in their research (and so researchers may choose to highlight conflicting results). However, in other instances, it seems that a story is intentionally misinterpreted in order to create a more exciting headline before later publishing a criticism of such interpretation.[459, 460]

The position of public health in modifying people's behaviour presents a critical ethical issue, namely whose responsibility is it that people lead a healthy lifestyle?[461] The role

of public health is to promote health and healthy living in the population. Most people would agree that this includes a duty to provide accurate information on risk factors for health so that people can make informed decisions about their own lifestyle. However, what is the responsibility of public health if the public are informed of the risks and choose not to modify their behaviour? A libertarian view would suggest that the government has no further role beyond providing information since people should be allowed to make their own decisions about how they live. Furthermore, governments are generally wary of introducing more intrusive measures to modify behaviour lest they be accused of "nannying" and interfering excessively with an individual's personal choice.

However, these unhealthy lifestyle behaviours represent a substantial cost to the NHS.[462-464] Furthermore, there is an increasing understanding that, rather than being simply the result of informed decision making, our lifestyle choices are influenced by numerous factors in our environment that lie outside of our control.[461, 465] Therefore, in order to prevent NCDs, public health institutions will need to introduce a wide range of measures in order to create environments that promote healthy lifestyles (for example by controlling the promotion, availability and pricing of unhealthy products).[466-468] For example, the prevalence of smoking has decreased in Great Britain and many other countries as a result of public health interventions including increased tax and bans on advertising.[44, 469]

To conclude, alcohol intake, adiposity and smoking contribute significantly to the global burden of NCDs, including colorectal cancer, and represent a substantial cost to health services. Reducing the prevalence of these risk factors in the general population is a vital challenge for public health and will require the introduction and evaluation of comprehensive interventions.

Appendix

Table A-1 BMI and the Risk of Colon Cancer and Rectal Cancer in UK Biobank (Adjusting for Different Alcohol Intake Variables)

		Colon cancer			Rectal cancer		
BMI (kg/m^2)	**Person-years**	**Cases**	**HR (95% CI)*†**	**HR (95% CI)*‡**	**Cases**	**HR (95% CI)*†**	**HR (95% CI)*‡**
Men							
<24.45	175,939	92	1.00	1.00	76	1.00	1.00
24.45-<26.40	181,554	106	1.02 (0.77-1.35)	1.02 (0.77-1.35)	75	0.88 (0.64-1.22)	0.88 (0.64-1.22)
26.40-<28.28	172,880	148	1.45 (1.11-1.88)	1.44 (1.11-1.88)	91	1.09 (0.80-1.48)	1.09 (0.80-1.48)
28.28-<30.84	184,678	162	1.44 (1.11-1.86)	1.43 (1.11-1.85)	106	1.15 (0.86-1.55)	1.15 (0.85-1.55)
≥30.84	167,417	163	1.59 (1.23-2.06)	1.59 (1.22-2.06)	91	1.08 (0.79-1.47)	1.07 (0.79-1.46)
P-trend§			0.0002	0.0002		0.3133	0.3362
Women							
<22.89	200,985	86	1.00	1.00	31	1.00	1.00
22.89-<25.04	199,338	109	1.17 (0.88-1.55)	1.17 (0.88-1.55)	46	1.34 (0.85-2.11)	1.34 (0.85-2.12)
25.04-<27.34	191,320	112	1.18 (0.89-1.57)	1.18 (0.89-1.57)	52	1.49 (0.95-2.33)	1.49 (0.95-2.33)
27.34-<30.80	186,810	110	1.15 (0.87-1.53)	1.16 (0.87-1.54)	39	1.10 (0.68-1.76)	1.10 (0.68-1.77)
≥30.80	168,356	91	1.12 (0.83-1.51)	1.12 (0.83-1.51)	31	0.98 (0.59-1.62)	0.98 (0.59-1.62)
P-trend§			0.9472	0.9313		0.5048	0.5080

* Adjusted for smoking status (never, former, current), Townsend deprivation index (quintiles), red meat intake (≤1, >1-<3, ≥3 times/week), processed meat (<1, 1, >1 time/week), family history of colorectal cancer (yes/no) and history of colorectal screening (yes/no).

† Adjusted for original categorical alcohol intake variable.

‡ Adjusted for alternative alcohol intake variable.

§ P-values for test for trend were calculated by assigning participants the median value of their category of BMI and this variable was modelled as a continuous variable.

Table A-2 Overall Smoking Status and the Risk of Colorectal Cancer in UK Biobank (Adjusting for Different Alcohol Intake Variables)

Overall smoking status	Person-years	Cases	HR (95% CI)*†	HR (95% CI)*‡
Overall				
Never	1,003,205	800	1.00	1.00
Former	645,088	833	1.22 (1.10-1.35)	1.22 (1.10-1.35)
Current	183,986	173	1.10 (0.93-1.31)	1.10 (0.93-1.31)
Men				
Never	426,639	407	1.00	1.00
Former	352,938	577	1.19 (1.01-1.40)	1.19 (1.01-1.40)
Current	104,324	124	1.03 (0.76-1.39)	1.03 (0.76-1.40)
Women				
Never	576,565	393	1.00	1.00
Former	292,150	256	1.23 (1.08-1.41)	1.23 (1.08-1.40)
Current	79,662	49	1.15 (0.93-1.41)	1.14 (0.93-1.41)

* Adjusted for sex, BMI (<25, 25-<30, ≥30 kg/m^2), Townsend deprivation index (quintiles), red meat intake (≤1, >1-<3, ≥3 times/week), processed meat (<1, 1, >1 time/week), height (continuous), family history of colorectal cancer (yes/no) and history of colorectal screening (yes/no).

† Adjusted for original categorical alcohol intake variable.

‡ Adjusted for alternative alcohol intake variable.

Table A-3 Alcohol Intake and the Risk of Colorectal Cancer in UK Biobank (with and without Adjustment for Physical Activity)

	Former	Never	<5	5-<15	15-<30	30-<45	≥45	P-trend§
Overall								
Person-years	56,229	64,886	192,833	466,255	388,954	165,797	134,186	
Cases	43	65	131	387	384	197	206	
HR (95% CI)*	0.68 (0.46-1.01)	1.00	0.70 (0.52-0.94)	0.79 (0.61-1.04)	0.86 (0.65-1.12)	0.94 (0.70-1.25)	1.14 (0.85-1.53)	0.0001
HR (95% CI)*†	0.68 (0.46-1.01)	1.00	0.70 (0.52-0.94)	0.79 (0.61-1.04)	0.86 (0.66-1.13)	0.94 (0.70-1.26)	1.14 (0.85-1.53)	0.0001
Men								
Person-years	26,318	19,725	54,445	179,318	221,105	122,596	116,921	
Cases	18	15	47	196	265	165	196	
HR (95% CI)*	0.80 (0.40-1.60)	1.00	1.07 (0.60-1.92)	1.31 (0.77-2.22)	1.39 (0.82-2.35)	1.51 (0.88-2.57)	1.83 (1.08-3.12)	<0.0001
HR (95% CI)*†	0.80 (0.40-1.60)	1.00	1.07 (0.60-1.92)	1.31 (0.77-2.22)	1.39 (0.82-2.35)	1.51 (0.88-2.57)	1.84 (1.08-3.12)	<0.0001
Women								
Person-years	29,911	45,161	138,388	286,937	167,849	43,201	17,264	
Cases	25	50	84	191	119	32	10	
HR (95% CI)*	0.74 (0.46-1.21)	1.00	0.57 (0.40-0.82)	0.61 (0.44-0.84)	0.65 (0.46-0.92)	0.69 (0.44-1.09)	0.57 (0.29-1.15)	0.8837
HR (95% CI)*†	0.74 (0.46-1.21)	1.00	0.57 (0.40-0.82)	0.61 (0.44-0.84)	0.66 (0.47-0.93)	0.69 (0.44-1.09)	0.58 (0.29-1.15)	0.8877

*Adjusted for sex, BMI (<25, 25-<30, ≥30 kg/m^2), overall smoking status (never, former, current), Townsend deprivation index (quintiles), red meat intake (≤1, >1-<3, ≥3 times/week), processed meat (<1, 1, >1 time/week), height (continuous), family history of colorectal cancer (yes/no) and history of colorectal screening (yes/no).

† Also adjusted for IPAQ physical activity (low, moderate, high).

§ P-values for test for trend were calculated by assigning participants the median value of their category of alcohol intake and this variable was modelled as a continuous variable (former drinkers were excluded).

Table A-4 Overall Smoking Status and the Risk of Colorectal Cancer in UK Biobank (with and without Adjustment for Physical Activity)

Overall smoking status	Person-years	Cases	HR (95% CI)*	HR (95% CI)*†
Overall				
Never	966,603	767	1.00	1.00
Former	594,287	719	1.16 (1.05-1.30)	1.16 (1.05-1.30)
Current	171,148	151	1.06 (0.89-1.27)	1.06 (0.88-1.27)
Men				
Never	412,608	399	1.00	1.00
Former	322,807	485	1.11 (0.97-1.28)	1.12 (0.97-1.28)
Current	95,702	106	1.05 (0.84-1.31)	1.05 (0.84-1.31)
Women				
Never	553,995	368	1.00	1.00
Former	271,480	234	1.23 (1.04-1.46)	1.23 (1.04-1.46)
Current	75,446	45	1.05 (0.77-1.44)	1.05 (0.77-1.44)

† Adjusted for sex, alcohol intake (never, former, special occasions only, one to three times a month, <5 g/d, 5-<15 g/d, 15-<30 g/d, 30-<45 g/d, ≥45 g/d), BMI (<25, 25-<30, ≥30 kg/m^2), Townsend deprivation index (quintiles), red meat intake (1, >1-<3, ≥3 times/week), processed meat (<1, 1, >1 time/week), height (continuous), family history of colorectal cancer (yes/no) and history of colorectal screening (yes/no).

† Also adjusted for IPAQ physical activity (low, moderate, high).

Appendix

Table A-5 Alcohol Intake and the Risk of Colorectal Cancer in UK Biobank (Including Missing Categories for Confounder Variables)

	Alcohol intake (grams/day)							
	Former	**Never**	**<5**	**5-<15**	**15-<30**	**30-<45**	**≥45**	**P-trend§**
Overall								
Person-years	83,544	104,815	268,576	617,013	498,902	209,384	170,994	
Cases	80	88	198	522	500	266	260	
HR (95% CI)*	1.00 (0.74-1.36)	1.00	0.89 (0.69-1.15)	0.95 (0.76-1.20)	1.04 (0.82-1.31)	1.20 (0.93-1.54)	1.36 (1.05-1.75)	<0.0001
Men								
Person-years	37,689	30,541	71,829	226,972	277,961	153,383	148,632	
Cases	36	22	63	241	341	218	250	
HR (95% CI)*	1.12 (0.66-1.91)	1.00	1.11 (0.68-1.81)	1.29 (0.83-2.01)	1.45 (0.94-2.24)	1.62 (1.04-2.52)	1.87 (1.20-2.92)	<0.0001
Women								
Person-years	45,855	74,274	196,746	390,041	220,941	56,000	22,363	
Cases	44	66	135	281	159	48	10	
HR (95% CI)*	1.05 (0.71-1.54)	1.00	0.80 (0.59-1.08)	0.81 (0.62-1.07)	0.83 (0.61-1.11)	1.00 (0.68-1.47)	0.56 (0.28-1.09)	0.7858

*Adjusted for sex, BMI (<25, 25-<30, ≥30 kg/m^2), overall smoking status (never, former, current), Townsend deprivation index (quintiles), red meat intake (≤1, >1-<3, ≥3 times/week), processed meat (<1, 1, >1 time/week), height (quintiles), family history of colorectal cancer (yes/no) and history of colorectal screening (yes/no).

§ P-values for test for trend were calculated by assigning participants the median value of their category of alcohol intake and this variable was modelled as a continuous variable (former drinkers were excluded).

Table A-6 BMI and the Risk of Colon Cancer and Rectal Cancer in UK Biobank (Including Missing Categories for Confounder Variables)

		Colon cancer		Rectal cancer	
BMI (kg/m^2)	**Person-years**	**Cases**	**HR (95% CI)***	**Cases**	**HR (95% CI)***
Men					
<24.45	215,916	113	1.00	91	1.00
24.45-<26.40	217,825	128	1.02 (0.79-1.31)	85	0.86 (0.64-1.15)
26.40-<28.28	208,260	168	1.35 (1.06-1.71)	100	1.03 (0.77-1.37)
28.28-<30.84	225,906	204	1.46 (1.16-1.84)	118	1.10 (0.83-1.45)
≥30.84	216,640	203	1.51 (1.20-1.91)	115	1.13 (0.85-1.50)
P-trend§			<0.0001		0.1140
Women					
<22.89	256,367	115	1.00	43	1.00
22.89-<25.04	255,419	139	1.09 (0.85-1.40)	63	1.32 (0.90-1.95)
25.04-<27.34	252,802	153	1.14 (0.90-1.46)	67	1.35 (0.92-1.98)
27.34-<30.80	256,081	156	1.10 (0.86-1.41)	55	1.05 (0.70-1.57)
≥30.80	254,237	142	1.05 (0.81-1.35)	42	0.81 (0.53-1.26)
P-trend§			0.9483		0.1841

***** Results adjusted for alcohol intake (never, former, special occasions only, one to three times a month, <5 g/d, 5-<15 g/d, 15-<30 g/d, 30-<45 g/d, ≥45 g/d), overall smoking status (never, former, current), Townsend deprivation index (quintiles), red meat intake (1, >1-<3, ≥3 times/week), processed meat (<1, 1, >1 time/week), family history of colorectal cancer (yes/no).

§ P-values for test for trend were calculated by assigning participants the median value of their category of BMI and this variable was modelled as a continuous variable.

Table A-7 Overall Smoking Status and the Risk of Colorectal Cancer in UK Biobank (Including Missing Categories for Confounder Variables)

	Overall smoking status		
	Never	**Former**	**Current**
Overall			
Person-years	1,296,524	787,565	248,495
Cases	1,034	998	231
HR (95% CI)*	1.00	1.20 (1.10-1.32)	1.11 (0.96-1.29)
Men			
Person-years	523,076	417,467	135,486
Cases	487	666	158
HR (95% CI)*	1.00	1.22 (1.08-1.37)	1.16 (0.97-1.40)
Women			
Person-years	773,448	370,098	113,009
Cases	547	332	73
HR (95% CI)*	1.00	1.18 (1.02-1.36)	1.03 (0.80-1.32)

* Adjusted for sex, alcohol intake (never, former, special occasions only, one to three times a month, <5 g/d, 5-<15 g/d, 15-<30 g/d, 30-<45 g/d, ≥45 g/d), BMI (<25, 25-<30, ≥30 kg/m^2), Townsend deprivation index (quintiles), red meat intake (1, >1-<3, ≥3 times/week), processed meat (<1, 1, >1 time/week), height (continuous), family history of colorectal cancer (yes/no) and history of colorectal screening (yes/no).

Table A-8 Alcohol Intake and the Risk of Colorectal Cancer in UK Biobank with Adjustment for Different Confounders

	Alcohol intake (grams/day)							
Confounder adjustment	**Former**	**Never**	**<5**	**5-<15**	**15-<30**	**30-<45**	**≥45**	**P-trend§**
Cases	70	79	187	499	479	248	244	
Sex	1.04 (0.76-1.44)	1.00	0.92 (0.71-1.20)	1.02 (0.81-1.30)	1.15 (0.90-1.46)	1.33 (1.03-1.72)	1.56 (1.20-2.02)	<0.0001
Sex, BMI	1.04 (0.76-1.44)	1.00	0.93 (0.72-1.21)	1.05 (0.82-1.33)	1.17 (0.92-1.48)	1.34 (1.03-1.73)	1.56 (1.20-2.03)	<0.0001
Sex, overall smoking status	0.97 (0.70-1.34)	1.00	0.89 (0.69-1.16)	0.98 (0.77-1.24)	1.07 (0.84-1.36)	1.22 (0.94-1.58)	1.42 (1.09-1.85)	<0.0001
Sex, Townsend index	1.04 (0.76-1.44)	1.00	0.92 (0.71-1.20)	1.03 (0.81-1.30)	1.15 (0.90-1.46)	1.33 (1.02-1.72)	1.56 (1.20-2.02)	<0.0001
Sex, red meat intake	1.04 (0.75-1.44)	1.00	0.92 (0.71-1.20)	1.01 (0.80-1.28)	1.13 (0.88-1.43)	1.29 (1.00-1.68)	1.51 (1.16-1.96)	<0.0001
Sex, processed meat intake	1.04 (0.76-1.44)	1.00	0.92 (0.71-1.20)	1.02 (0.80-1.29)	1.13 (0.89-1.44)	1.30 (1.01-1.69)	1.52 (1.17-1.97)	<0.0001
Sex, height	1.03 (0.75-1.42)	1.00	0.91 (0.70-1.18)	1.00 (0.79-1.27)	1.12 (0.88-1.42)	1.29 (1.00-1.68)	1.52 (1.17-1.98)	<0.0001
Sex, family history of colorectal cancer	1.04 (0.75-1.43)	1.00	0.92 (0.71-1.20)	1.02 (0.81-1.30)	1.14 (0.90-1.46)	1.32 (1.02-1.71)	1.55 (1.19-2.01)	<0.0001
Sex, bowel screening	1.05 (0.76-1.45)	1.00	0.94 (0.72-1.22)	1.03 (0.81-1.30)	1.15 (0.91-1.47)	1.33 (1.03-1.73)	1.56 (1.20-2.03)	<0.0001

§ P-values for test for trend were calculated by assigning participants the median value of their category of alcohol intake and this variable was modelled as a continuous variable (former drinkers were excluded).

Table A-9 BMI and the Risk of Colon Cancer for Men in UK Biobank with Adjustment for Different Confounders

	BMI (kg/m^2)					
Confounder adjustment	**<24.45**	**24.45-<26.40**	**26.40-<28.28**	**28.28-<30.84**	**≥30.84**	**P-trend§**
Cases	99	118	159	184	186	
None	1.00	1.11 (0.85-1.45)	1.56 (1.21-2.00)	1.67 (1.30-2.13)	1.81 (1.42-2.32)	<0.0001
Alcohol intake	1.00	1.09 (0.83-1.42)	1.51 (1.18-1.94)	1.61 (1.26-2.05)	1.76 (1.38-2.25)	<0.0001
Overall smoking status	1.00	1.09 (0.84-1.43)	1.52 (1.18-1.95)	1.61 (1.26-2.06)	1.73 (1.36-2.22)	<0.0001
Townsend index	1.00	1.11 (0.85-1.45)	1.56 (1.21-2.00)	1.66 (1.30-2.12)	1.81 (1.42-2.32)	<0.0001
Red meat intake	1.00	1.10 (0.84-1.44)	1.53 (1.19-1.97)	1.62 (1.27-2.08)	1.75 (1.37-2.23)	<0.0001
Processed meat intake	1.00	1.10 (0.85-1.44)	1.54 (1.20-1.98)	1.64 (1.29-2.10)	1.78 (1.39-2.27)	<0.0001
Family history of colorectal cancer	1.00	1.11 (0.85-1.45)	1.56 (1.21-1.99)	1.66 (1.30-2.11)	1.80 (1.41-2.30)	<0.0001
Bowel screening	1.00	1.11 (0.85-1.45)	1.56 (1.21-2.00)	1.67 (1.30-2.13)	1.81 (1.42-2.32)	<0.0001

§ P-values for test for trend were calculated by assigning participants the median value of their category of BMI and this variable was modelled as a continuous variable.

Table A-10 BMI and the Risk of Colon Cancer for Women in UK Biobank with Adjustment for Different Confounders

	BMI (kg/m^2)					
Confounder adjustment	**<24.45**	**24.45-<26.40**	**26.40-<28.28**	**28.28-<30.84**	**≥30.84**	**P-trend§**
Cases	106	134	137	143	127	
None	1.00	1.15 (0.89-1.49)	1.13 (0.88-1.46)	1.14 (0.88-1.46)	1.08 (0.84-1.40)	0.7827
Alcohol intake	1.00	1.15 (0.89-1.49)	1.13 (0.88-1.46)	1.14 (0.88-1.46)	1.08 (0.83-1.40)	0.8015
Overall smoking status	1.00	1.15 (0.89-1.48)	1.13 (0.88-1.46)	1.13 (0.88-1.46)	1.08 (0.83-1.40)	0.8073
Townsend index	1.00	1.15 (0.89-1.49)	1.13 (0.88-1.46)	1.14 (0.88-1.46)	1.08 (0.83-1.40)	0.7876
Red meat intake	1.00	1.15 (0.89-1.48)	1.13 (0.87-1.45)	1.13 (0.88-1.46)	1.08 (0.83-1.39)	0.8162
Processed meat intake	1.00	1.15 (0.89-1.49)	1.14 (0.88-1.47)	1.14 (0.89-1.47)	1.09 (0.84-1.41)	0.7447
Family history of colorectal cancer	1.00	1.15 (0.89-1.49)	1.13 (0.88-1.46)	1.14 (0.88-1.46)	1.08 (0.83-1.40)	0.7900
Bowel screening	1.00	1.15 (0.89-1.49)	1.13 (0.88-1.46)	1.14 (0.88-1.46)	1.08 (0.84-1.40)	0.7838

§ P-values for test for trend were calculated by assigning participants the median value of their category of BMI and this variable was modelled as a continuous variable.

Table A-11 BMI and the Risk of Rectal Cancer for Men in UK Biobank with Adjustment for Different Confounders

	BMI (kg/m^2)					
Confounder adjustment	**<24.45**	**24.45-<26.40**	**26.40-<28.28**	**28.28-<30.84**	**≥30.84**	**P-trend§**
Cases	83	81	98	113	106	
None	1.00	0.91 (0.67-1.24)	1.15 (0.86-1.54)	1.23 (0.92-1.63)	1.23 (0.93-1.65)	0.0377
Alcohol intake	1.00	0.89 (0.66-1.21)	1.11 (0.83-1.48)	1.17 (0.88-1.56)	1.19 (0.89-1.59)	0.0668
Overall smoking status	1.00	0.91 (0.67-1.24)	1.14 (0.85-1.53)	1.21 (0.91-1.60)	1.20 (0.90-1.61)	0.0638
Townsend index	1.00	0.92 (0.67-1.25)	1.15 (0.86-1.55)	1.23 (0.93-1.63)	1.23 (0.92-1.64)	0.0410
Red meat intake	1.00	0.91 (0.67-1.24)	1.14 (0.85-1.53)	1.21 (0.91-1.61)	1.21 (0.91-1.62)	0.0542
Processed meat intake	1.00	0.91 (0.67-1.23)	1.13 (0.85-1.52)	1.20 (0.90-1.59)	1.18 (0.88-1.57)	0.0862
Family history of colorectal cancer	1.00	0.91 (0.67-1.24)	1.15 (0.86-1.54)	1.22 (0.92-1.63)	1.23 (0.92-1.64)	0.0395
Bowel screening	1.00	0.91 (0.67-1.24)	1.15 (0.86-1.54)	1.23 (0.93-1.63)	1.23 (0.93-1.65)	0.0376

§ P-values for test for trend were calculated by assigning participants the median value of their category of BMI and this variable was modelled as a continuous variable.

Table A-12 BMI and the Risk of Rectal Cancer for Women in UK Biobank with Adjustment for Different Confounders

	BMI (kg/m^2)					
Confounder adjustment	**<24.45**	**24.45-<26.40**	**26.40-<28.28**	**28.28-<30.84**	**≥30.84**	**P-trend§**
Cases	38	59	60	53	39	
None	1.00	1.44 (0.96-2.16)	1.42 (0.95-2.14)	1.22 (0.80-1.85)	0.94 (0.60-1.47)	0.2995
Alcohol intake	1.00	1.45 (0.96-2.17)	1.42 (0.95-2.14)	1.21 (0.79-1.84)	0.92 (0.59-1.45)	0.2533
Overall smoking status	1.00	1.42 (0.94-2.13)	1.40 (0.93-2.10)	1.19 (0.78-1.81)	0.91 (0.58-1.43)	0.2327
Townsend index	1.00	1.44 (0.95-2.16)	1.43 (0.95-2.14)	1.23 (0.81-1.87)	0.96 (0.61-1.51)	0.3682
Red meat intake	1.00	1.43 (0.95-2.16)	1.42 (0.94-2.13)	1.21 (0.80-1.84)	0.94 (0.60-1.47)	0.2880
Processed meat intake	1.00	1.44 (0.96-2.17)	1.42 (0.95-2.14)	1.22 (0.80-1.85)	0.94 (0.60-1.48)	0.3005
Family history of colorectal cancer	1.00	1.44 (0.96-2.17)	1.42 (0.95-2.14)	1.22 (0.80-1.85)	0.94 (0.60-1.48)	0.3013
Bowel screening	1.00	1.44 (0.95-2.16)	1.42 (0.94-2.13)	1.21 (0.80-1.84)	0.94 (0.60-1.47)	0.2995

§ P-values for test for trend were calculated by assigning participants the median value of their category of BMI and this variable was modelled as a continuous variable.

Table A-13 Overall Smoking Status and the Risk of Colorectal Cancer in UK Biobank with Adjustment for Different Confounders

	Overall smoking status		
Confounder adjustment	**Never**	**Former**	**Current**
Cases	973	929	207
Sex	1.00	1.26 (1.15-1.38)	1.18 (1.02-1.37)
Sex, alcohol intake	1.00	1.21 (1.10-1.33)	1.11 (0.95-1.30)
Sex, BMI	1.00	1.25 (1.14-1.37)	1.18 (1.02-1.37)
Sex, Townsend index	1.00	1.26 (1.15-1.39)	1.18 (1.02-1.38)
Sex, red meat intake	1.00	1.26 (1.15-1.38)	1.17 (1.00-1.36)
Sex, processed meat intake	1.00	1.26 (1.15-1.38)	1.17 (1.00-1.36)
Sex, height	1.00	1.26 (1.15-1.38)	1.19 (1.02-1.38)
Sex, family history of colorectal cancer	1.00	1.26 (1.15-1.38)	1.18 (1.01-1.37)
Sex, bowel screening	1.00	1.27 (1.16-1.39)	1.18 (1.01-1.37)

Table A-14 BMI and the Risk of Colon Cancer and Rectal Cancer in UK Biobank (Using Multiply Imputed Alcohol Intake Data)

	Colon cancer		Rectal cancer	
BMI (kg/m^2)	**Cases**	**HR (95% CI)***	**Cases**	**HR (95% CI)***
Men				
<25	138	1.00	101	1.00
25-<30	371	1.22 (1.00-1.48)	249	1.15 (0.91-1.45)
≥30	237	1.50 (1.21-1.86)	133	1.18 (0.91-1.54)
P-trend§		0.0001		0.1368
Women				
<25	238	1.00	97	1.00
25-<30	254	1.04 (0.87-1.24)	105	1.05 (0.80-1.39)
≥30	152	1.01 (0.82-1.24)	48	0.76 (0.53-1.08)
P-trend§		0.9444		0.2483

* Results adjusted for alcohol intake (never, former, <5 g/d, 5-<15 g/d, 15-<30 g/d, 30-<45 g/d, ≥45 g/d), smoking status (never, former, current), Townsend deprivation index (quintiles), red meat intake (1, >1-<3, ≥3 times/week), processed meat (<1, 1, >1 time/week), family history of colorectal cancer (yes/no) and history of colorectal screening (yes/no).

§ P-values for test for trend were calculated by assigning participants the median value of their category of BMI and this variable was modelled as a continuous variable.

Table A-15 Overall Smoking Status and the Risk of Colorectal Cancer in UK Biobank (Using Multiply Imputed Alcohol Intake Data)

	Overall		Men		Women	
Overall smoking status	**Cases**	**HR (95% CI)***	**Cases**	**HR (95% CI)***	**Cases**	**HR (95% CI)***
Never	972	1.00	460	1.00	512	1.00
Former	926	1.19 (1.09-1.31)	620	1.21 (1.06-1.37)	306	1.16 (1.00-1.35)
Current	205	1.10 (0.94-1.28)	141	1.16 (0.95-1.40)	64	1.01 (0.77-1.32)

* Adjusted for sex, alcohol intake (never, former, <5 g/d, 5-<15 g/d, 15-<30 g/d, 30-<45 g/d, ≥45 g/d), BMI (<25, 25-<30, ≥30 kg/m^2), Townsend deprivation index (quintiles), red meat intake (1, >1-<3, ≥3 times/week), processed meat (<1, 1, >1 time/week), height (continuous), family history of colorectal cancer (yes/no) and history of colorectal screening (yes/no).

References

1. Lozano R, Naghavi M, Foreman K, Lim S, Shibuya K, Aboyans V, et al. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet. 2012;380(9859):2095-128.
2. Murray CJ, Vos T, Lozano R, Naghavi M, Flaxman AD, Michaud C, et al. Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet. 2012;380(9859):2197-223.
3. World Health Organization. Global Status Report on Noncommunicable Diseases 2014. Geneva: World Health Organization; 2014.
4. United Nations Department of Economic and Social Affairs Population Division. World Population Prospects: The 2015 Revision, Key Findings and Advance Tables. Working Paper No. ESA/P/WP.241; 2015.
5. Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A. Global cancer statistics, 2012. CA: A Cancer Journal for Clinicians. 2015;65(2):87-108.
6. Bray F, Jemal A, Grey N, Ferlay J, Forman D. Global cancer transitions according to the Human Development Index (2008-2030): a population-based study. Lancet Oncology. 2012;13(8):790-801.
7. Ahmad A, Ormiston-Smith N, Sasieni P. Trends in the lifetime risk of developing cancer in Great Britain: comparison of risk for those born from 1930 to 1960. British Journal of Cancer. 2015;112(5):943-7.
8. Cancer Research UK. Cancer statistics for the UK. Available from: http://www.cancerresearchuk.org/health-professional/cancer-statistics. [accessed 30th June 2016].
9. Fritz A, Percy C, Jack A, Shanmugaratnam K, Sobin L, Parkin DM, et al. International Classification of Diseases for Oncology, Third Edition. Geneva: World Health Organization; 2000.
10. McPhail S, Johnson S, Greenberg D, Peake M, Rous B. Stage at diagnosis and early mortality from cancer in England. British Journal of Cancer. 2015;112 Suppl 1:S108-15.
11. Maringe C, Walters S, Rachet B, Butler J, Fields T, Finan P, et al. Stage at diagnosis and colorectal cancer survival in six high-income countries: a population-based study of patients diagnosed during 2000-2007. Acta Oncologica. 2013;52(5):919-32.
12. Hewitson P, Glasziou P, Irwig L, Towler B, Watson E. Screening for colorectal cancer using the faecal occult blood test, Hemoccult. The Cochrane Database of Systematic Reviews. 2007;(1):CD001216.
13. McClements PL, Madurasinghe V, Thomson CS, Fraser CG, Carey FA, Steele RJ, et al. Impact of the UK colorectal cancer screening pilot studies on incidence, stage distribution and mortality trends. Cancer Epidemiology. 2012;36(4):e232-42.
14. Zauber AG. The impact of screening on colorectal cancer mortality and incidence: has it really made a difference? Digestive Diseases and Sciences. 2015;60(3):681-91.
15. Rees CJ, Bevan R. The National Health Service Bowel Cancer Screening Program: the early years. Expert Review of Gastroenterology & Hepatology. 2013;7(5):421-37.
16. ISD Scotland. Scottish Bowel Screening Programme. Available from: http://www.isdscotland.org/Health-Topics/Cancer/Bowel-Screening/. [accessed 28th July 2016].
17. Public Health Wales. First annual report from Bowel Screening Wales shows increase in uptake. Available from: http://www.wales.nhs.uk/sitesplus/888/news/36504. [accessed 28th July 2016].
18. Lambert R, Sauvaget C, Sankaranarayanan R. Mass screening for colorectal cancer is not justified in most developing countries. International Journal of Cancer. 2009;125(2):253-6.
19. Giovannucci E, Wu K. Cancers of the Colon and Rectum. In: D. S, Fraumeni JF, Jr., editors. Cancer Epidemiology and Prevention Third Edition. Oxford: Oxford University Press; 2006. p. 809-29.

20. Winawer SJ, Zauber AG. The advanced adenoma as the primary target of screening. Gastrointestinal Endoscopy Clinics of North America. 2002;12(1):1-9, v.
21. Levin B, Lieberman DA, McFarland B, Smith RA, Brooks D, Andrews KS, et al. Screening and surveillance for the early detection of colorectal cancer and adenomatous polyps, 2008: a joint guideline from the American Cancer Society, the US Multi-Society Task Force on Colorectal Cancer, and the American College of Radiology. CA: A Cancer Journal for Clinicians. 2008;58(3):130-60.
22. Center MM, Jemal A, Smith RA, Ward E. Worldwide variations in colorectal cancer. CA: A Cancer Journal for Clinicians. 2009;59(6):366-78.
23. Center MM, Jemal A, Ward E. International trends in colorectal cancer incidence rates. Cancer Epidemiology Biomarkers & Prevention. 2009;18(6):1688-94.
24. Parkin DM. International variation. Oncogene. 2004;23(38):6329-40.
25. World Cancer Research Fund / American Institute for Cancer Research. Food, Nutrition, Physical Activity, and the Prevention of Cancer: a Global Perspective. Washington DC: AICR; 2007.
26. International Agency for Research on Cancer. A Review of Human Carcinogens: Personal Habits and Indoor Combustions. Volume 100E. Lyon, France: International Agency for Research on Cancer; 2012. Report No.: 928321322X.
27. International Agency for Research on Cancer. IARC Handbooks of Cancer Prevention Volume 6. Weight Control and Physical Activty Lyon: IARC Press; 2002.
28. Bouvard V, Loomis D, Guyton KZ, Grosse Y, Ghissassi FE, Benbrahim-Tallaa L, et al. Carcinogenicity of consumption of red and processed meat. Lancet Oncology. 2015;16(16):1599-600.
29. U.S. Department of Health and Human Services. The Health Consequences of Smoking—50 Years of Progress. A Report of the Surgeon General. Atlanta, GA: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health; 2014.
30. Parkin DM, Boyd L, Walker L. 16. The fraction of cancer attributable to lifestyle and environmental factors in the UK in 2010. British Journal of Cancer. 2011;105:S77-S81.
31. Short E, Thomas LE, Hurley J, Jose S, Sampson JR. Inherited predisposition to colorectal cancer: towards a more complete picture. Journal of Medical Genetics. 2015;52(12):791-6.
32. Mishra N, Hall J. Identification of patients at risk for hereditary colorectal cancer. Clinics in Colon and Rectal Surgery. 2012;25(2):67-82.
33. Jasperson KW, Tuohy TM, Neklason DW, Burt RW. Hereditary and familial colon cancer. Gastroenterology. 2010;138(6):2044-58.
34. World Health Organization. Global Status Report on Alcohol and Health 2014. Luxembourg: World Health Organisation; 2014.
35. Office for National Statistics. Adult Drinking Habits in Great Britain, 2013. 2015 Available from: http://webarchive.nationalarchives.gov.uk/20160105160709/http://www.ons.gov.uk/ons/dcp171778_395191.pdf. [accessed 19th September 2016].
36. Swinburn BA, Sacks G, Hall KD, McPherson K, Finegood DT, Moodie ML, et al. The global obesity pandemic: shaped by global drivers and local environments. Lancet. 2011;378(9793):804-14.
37. Finucane MM, Stevens GA, Cowan MJ, Danaei G, Lin JK, Paciorek CJ, et al. National, regional, and global trends in body-mass index since 1980: systematic analysis of health examination surveys and epidemiological studies with 960 country-years and 9.1 million participants. Lancet. 2011;377(9765):557-67.
38. World Health Organization. Obesity and overweight. Available from: www.who.int/mediacentre/factsheets/fs311/en/. [accessed 8th December 2015].
39. Lim SS, Vos T, Flaxman AD, Danaei G, Shibuya K, Adair-Rohani H, et al. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet. 2012;380(9859):2224-60.

40. Health and Social Care Information Centre. Statistics on Obesity, Physical Activity and Diet: England 2015. 2015 Available from: http://digital.nhs.uk/catalogue/PUB16988/obes-phys-acti-diet-eng-2015.pdf. [accessed 19th September 2016].
41. World Health Organization. WHO Global Report: Mortality Attributable to Tobacco. Geneva: World Health Organisation; 2012.
42. Doll R, Peto R, Boreham J, Sutherland I. Mortality in relation to smoking: 50 years' observations on male British doctors. British Medical Journal. 2004;328(7455):1519.
43. Jha P. Avoidable global cancer deaths and total deaths from smoking. Nature Reviews: Cancer. 2009;9(9):655-64.
44. Office for National Statistics. Adult Smoking Habits in Great Britain, 2013. 2014 Available from: http://webarchive.nationalarchives.gov.uk/20160105160709/http://www.ons.gov.uk/ons/dcp171778_386291.pdf. [accessed 19th September 2016].
45. Cho E, Smith-Warner SA, Ritz J, van den Brandt PA, Colditz GA, Folsom AR, et al. Alcohol intake and colorectal cancer: a pooled analysis of 8 cohort studies. Annals of Internal Medicine. 2004;140(8):603-13.
46. Shaper AG, Wannamethee G, Walker M. Alcohol and mortality in British men: explaining the U-shaped curve. Lancet. 1988;332(8623):1267-73.
47. Pischon T, Lahmann PH, Boeing H, Friedenreich C, Norat T, Tjønneland A, et al. Body size and risk of colon and rectal cancer in the European Prospective Investigation Into Cancer and Nutrition (EPIC). Journal of the National Cancer Institute. 2006;98(13):920-31.
48. Connor Gorber S, Tremblay M, Moher D, Gorber B. A comparison of direct vs. self-report measures for assessing height, weight and body mass index: a systematic review. Obesity Reviews. 2007;8(4):307-26.
49. Tsoi KK, Pau CY, Wu WK, Chan FK, Griffiths S, Sung JJ. Cigarette smoking and the risk of colorectal cancer: a meta-analysis of prospective cohort studies. Clinical Gastroenterology and Hepatology. 2009;7(6):682-8. e5.
50. Cheng J, Chen Y, Wang X, Wang J, Yan Z, Gong G, et al. Meta-analysis of prospective cohort studies of cigarette smoking and the incidence of colon and rectal cancers. European Journal of Cancer Prevention. 2015;24(1):6-15.
51. Iacopetta B. Are there two sides to colorectal cancer? International Journal of Cancer. 2002;101(5):403-8.
52. Yamauchi M, Lochhead P, Morikawa T, Huttenhower C, Chan AT, Giovannucci E, et al. Colorectal cancer: a tale of two sides or a continuum? Gut. 2012;61(6):794-7.
53. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, et al. UK Biobank: an Open Access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Medicine. 2015;12(3):1-10.
54. Grimes DA, Schulz KF. Cohort studies: marching towards outcomes. Lancet. 2002;359(9303):341-5.
55. Schulz KF, Grimes DA. Case-control studies: research in reverse. Lancet. 2002;359(9304):431-4.
56. World Cancer Research Fund / American Institute for Cancer Research. Continuous Update Project Report. Food, Nutrition, Physical Activity, and the Prevention of Colorectal Cancer; 2011.
57. NHS Choices. Alcohol units. Available from: http://www.nhs.uk/Livewell/alcohol/Pages/alcohol-units.aspx. [accessed 12th December 2015].
58. Food Standards Agency. McCance and Widdowson's The Composition of Foods, 6th Summary Edition. Cambridge: The Royal Society of Chemistry; 2002.
59. International Agency for Research on Cancer. Alcohol Consumption and Ethyl Carbamate. Lyon, France: International Agency for Research on Cancer; 2010. Report No.: 1017-1606.
60. Moskal A, Norat T, Ferrari P, Riboli E. Alcohol intake and colorectal cancer risk: A dose–response meta-analysis of published cohort studies. International Journal of Cancer. 2007;120(3):664-71.

61. Huxley RR, Ansary-Moghaddam A, Clifton P, Czernichow S, Parr CL, Woodward M. The impact of dietary and lifestyle risk factors on risk of colorectal cancer: a quantitative overview of the epidemiological evidence. International Journal of Cancer. 2009;125(1):171-80.
62. Fedirko V, Tramacere I, Bagnardi V, Rota M, Scotti L, Islami F, et al. Alcohol drinking and colorectal cancer risk: an overall and dose–response meta-analysis of published studies. Annals of Oncology. 2011;22(9):1958-72.
63. Bagnardi V, Rota M, Botteri E, Tramacere I, Islami F, Fedirko V, et al. Alcohol consumption and site-specific cancer risk: a comprehensive dose–response meta-analysis. British Journal of Cancer. 2015;112(3):580-93.
64. Chan AT, Giovannucci EL. Primary prevention of colorectal cancer. Gastroenterology. 2010;138(6):2029-43. e10.
65. Bongaerts BW, van den Brandt PA, Goldbohm RA, de Goeij AF, Weijenberg MP. Alcohol consumption, type of alcoholic beverage and risk of colorectal cancer at specific subsites. International Journal of Cancer. 2008;123(10):2411-7.
66. Nan H, Lee JE, Rimm EB, Fuchs CS, Giovannucci EL, Cho E. Prospective study of alcohol consumption and the risk of colorectal cancer before and after folic acid fortification in the United States. Annals of Epidemiology. 2013;23(9):558-63.
67. Razzak AA, Oxentenko AS, Vierkant RA, Tillmans LS, Wang AH, Weisenberger DJ, et al. Alcohol intake and colorectal cancer risk by molecularly defined subtypes in a prospective study of older women. Cancer Prevention Research. 2011;4(12):2035-43.
68. Kabat G, Miller A, Jain M, Rohan T. Dietary intake of selected B vitamins in relation to risk of major cancers in women. British Journal of Cancer. 2008;99(5):816-21.
69. Ferrari P, Jenab M, Norat T, Moskal A, Slimani N, Olsen A, et al. Lifetime and baseline alcohol intake and risk of colon and rectal cancers in the European prospective investigation into cancer and nutrition (EPIC). International Journal of Cancer. 2007;121(9):2065-72.
70. Allen NE, Beral V, Casabonne D, Kan SW, Reeves GK, Brown A, et al. Moderate alcohol intake and cancer incidence in women. Journal of the National Cancer Institute. 2009;101(5):296-305.
71. Mizoue T, Inoue M, Wakai K, Nagata C, Shimazu T, Tsuji I, et al. Alcohol drinking and colorectal cancer in Japanese: a pooled analysis of results from five cohort studies. American Journal of Epidemiology. 2008;167(12):1397-406.
72. Butler L, Wang R, Koh W, Yu M. Prospective study of dietary patterns and colorectal cancer among Singapore Chinese. British Journal of Cancer. 2008;99(9):1511-6.
73. Pedersen A, Johansen C, Grønbæk M. Relations between amount and type of alcohol and colon and rectal cancer in a Danish population based cohort study. Gut. 2003;52(6):861-7.
74. Chyou P-H, Nomura AM, Stemmermann GN. A prospective study of colon and rectal cancer among Hawaii Japanese men. Annals of Epidemiology. 1996;6(4):276-82.
75. Klatsky AL, Armstrong MA, Friedman GD, Hiatt RA. The relations of alcoholic beverage use to colon and rectal cancer. American Journal of Epidemiology. 1988;128(5):1007-15.
76. Wu AH, Paganini-Hill A, Ross R, Henderson B. Alcohol, physical activity and other risk factors for colorectal cancer: a prospective study. British Journal of Cancer. 1987;55(6):687.
77. Tsong W, Koh W, Yuan J, Wang R, Sun C, Yu M. Cigarettes and alcohol in relation to colorectal cancer: the Singapore Chinese Health Study. British Journal of Cancer. 2007;96(5):821-7.
78. Wei EK, Giovannucci E, Wu K, Rosner B, Fuchs CS, Willett WC, et al. Comparison of risk factors for colon and rectal cancer. International Journal of Cancer. 2004;108(3):433-42.
79. Chen K, Jiang Q, Ma X, Li Q, Yao K, Yu W, et al. Alcohol drinking and colorectal cancer: a population-based prospective cohort study in China. European Journal of Epidemiology. 2005;20(2):149-54.
80. Thygesen LC, Wu K, Grønbæk M, Fuchs CS, Willett WC, Giovannucci E. Alcohol intake and colorectal cancer: a comparison of approaches for including repeated measures of alcohol consumption. Epidemiology. 2008;19(2):258-64.
81. Li FY, Lai MD. Colorectal cancer, one entity or three. Journal of Zhejiang University SCIENCE B. 2009;10(3):219-29.

82. Akhter M, Kuriyama S, Nakaya N, Shimazu T, Ohmori K, Nishino Y, et al. Alcohol consumption is associated with an increased risk of distal colon and rectal cancer in Japanese men: the Miyagi Cohort Study. European Journal of Cancer. 2007;43(2):383-90.
83. Park JY, Mitrou PN, Dahm CC, Luben RN, Wareham NJ, Khaw K-T, et al. Baseline alcohol consumption, type of alcoholic beverage and risk of colorectal cancer in the European Prospective Investigation into Cancer and Nutrition-Norfolk study. Cancer Epidemiology. 2009;33(5):347-54.
84. Zhao J, Zhu Y, Wang PP, West R, Buehler S, Sun Z, et al. Interaction between alcohol drinking and obesity in relation to colorectal cancer risk: a case-control study in Newfoundland and Labrador, Canada. BMC Public Health. 2012;12(1):94.
85. Kim YI. Folate and colorectal cancer: An evidence-based critical review. Molecular Nutrition & Food Research. 2007;51(3):267-92.
86. Bollheimer LC, Buettner R, Kullmann A, Kullmann F. Folate and its preventive potential in colorectal carcinogenesis.: How strong is the biological and epidemiological evidence? Critical Reviews in Oncology/Hematology. 2005;55(1):13-36.
87. Hubner R, Houlston R. Folate and colorectal cancer prevention. British Journal of Cancer. 2009;100(2):233-9.
88. Giovannucci E. Epidemiologic studies of folate and colorectal neoplasia: a review. The Journal of Nutrition. 2002;132(8):2350S-5S.
89. Mason JB, Choi S-W. Effects of alcohol on folate metabolism: implications for carcinogenesis. Alcohol. 2005;35(3):235-41.
90. Lee JE, Willett WC, Fuchs CS, Smith-Warner SA, Wu K, Ma J, et al. Folate intake and risk of colorectal cancer and adenoma: modification by time. American Journal of Clinical Nutrition. 2011;93(4):817-25.
91. Gibson TM, Weinstein SJ, Pfeiffer RM, Hollenbeck AR, Subar AF, Schatzkin A, et al. Pre-and postfortification intake of folate and risk of colorectal cancer in a large prospective cohort study in the United States. American Journal of Clinical Nutrition. 2011;94(4):1053-62.
92. Stevens VL, McCullough ML, Sun J, Jacobs EJ, Campbell PT, Gapstur SM. High levels of folate from supplements and fortification are not associated with increased risk of colorectal cancer. Gastroenterology. 2011;141(1):98-105. e1.
93. Roswall N, Olsen A, Christensen J, Dragsted LO, Overvad K, Tjønneland A. Micronutrient intake and risk of colon and rectal cancer in a Danish cohort. Cancer Epidemiology. 2010;34(1):40-6.
94. Larsson SC, Giovannucci E, Wolk A. A prospective study of dietary folate intake and risk of colorectal cancer: modification by caffeine intake and cigarette smoking. Cancer Epidemiology Biomarkers & Prevention. 2005;14(3):740-3.
95. Razzak AA, Oxentenko AS, Vierkant RA, Tillmans LS, Wang AH, Weisenberger DJ, et al. Associations between intake of folate and related micronutrients with molecularly defined colorectal cancer risks in the Iowa Women's Health Study. Nutrition and Cancer. 2012;64(7):899-910.
96. de Vogel S, Dindore V, van Engeland M, Goldbohm RA, van den Brandt PA, Weijenberg MP. Dietary folate, methionine, riboflavin, and vitamin B-6 and risk of sporadic colorectal cancer. The Journal of Nutrition. 2008;138(12):2372-8.
97. Zhang SM, Moore SC, Lin J, Cook NR, Manson JE, Lee I-M, et al. Folate, vitamin B6, multivitamin supplements, and colorectal cancer risk in women. American Journal of Epidemiology. 2006;163(2):108-15.
98. Flood A, Caprario L, Chaterjee N, Lacey Jr JV, Schairer C, Schatzkin A. Folate, methionine, alcohol, and colorectal cancer in a prospective study of women in the United States. Cancer Causes and Control. 2002;13(6):551-61.
99. Shrubsole MJ, Yang G, Gao Y-T, Chow WH, Shu XO, Cai Q, et al. Dietary B vitamin and methionine intakes and plasma folate are not associated with colorectal cancer risk in Chinese women. Cancer Epidemiology Biomarkers & Prevention. 2009;18(3):1003-6.
100. Ishihara J, Otani T, Inoue M, Iwasaki M, Sasazuki S, Tsugane S. Low intake of vitamin B-6 is associated with increased risk of colorectal cancer in Japanese men. The Journal of Nutrition. 2007;137(7):1808-14.

101. Glynn SA, Albanes D, Pietinen P, Brown CC, Rautalahti M, Tangrea JA, et al. Colorectal cancer and folate status: a nested case-control study among male smokers. Cancer Epidemiology Biomarkers & Prevention. 1996;5(7):487-94.
102. Hutcheon JA, Chiolero A, Hanley JA. Random measurement error and regression dilution bias. British Medical Journal. 2010;340(7761):1402-6.
103. Chuang S-C, Rota M, Gunter MJ, Zeleniuch-Jacquotte A, Eussen SJ, Vollset SE, et al. Quantifying the dose-response relationship between circulating folate concentrations and colorectal cancer in cohort studies: a meta-analysis based on a flexible meta-regression model. American Journal of Epidemiology. 2013;178(7):1028-37.
104. Lee JE, Wei EK, Fuchs CS, Hunter DJ, Lee I-M, Selhub J, et al. Plasma folate, methylenetetrahydrofolate reductase (MTHFR), and colorectal cancer risk in three large nested case–control studies. Cancer Causes and Control. 2012;23(4):537-45.
105. Van Guelpen B, Hultdin J, Johansson I, Hallmans G, Stenling R, Riboli E, et al. Low folate levels may protect against colorectal cancer. Gut. 2006;55(10):1461-6.
106. Takata Y, Shrubsole MJ, Li H, Cai Q, Gao J, Wagner C, et al. Plasma folate concentrations and colorectal cancer risk: A case-control study nested within the Shanghai Men's Health Study. International Journal of Cancer. 2014;135(9):2191-8.
107. Jaszewski R, Misra S, Tobi M, Ullah N, Naumoff JA, Kucuk O, et al. Folic acid supplementation inhibits recurrence of colorectal adenomas: a randomized chemoprevention trial. World Journal of Gastroenterology. 2008;14(28):4492.
108. Paspatis GA, Karamanolis DG. Folate supplementation and adenomatous colonic polyps. Diseases of the Colon and Rectum. 1994;37(12):1340-1.
109. Cole BF, Baron JA, Sandler RS, Haile RW, Ahnen DJ, Bresalier RS, et al. Folic acid for the prevention of colorectal adenomas: a randomized clinical trial. JAMA. 2007;297(21):2351-9.
110. Logan RF, Grainge MJ, Shepherd VC, Armitage NC, Muir KR, ukCAP Trial Group. Aspirin and folic acid for the prevention of recurrent colorectal adenomas. Gastroenterology. 2008;134(1):29-38.
111. Wu K, Platz EA, Willett WC, Fuchs CS, Selhub J, Rosner BA, et al. A randomized trial on folic acid supplementation and risk of recurrent colorectal adenoma. American Journal of Clinical Nutrition. 2009;90(6):1623-31.
112. Song Y, Manson JE, Lee I-M, Cook NR, Paul L, Selhub J, et al. Effect of combined folic acid, vitamin B6, and vitamin B12 on colorectal adenoma. Journal of the National Cancer Institute. 2012;104(20):1562-75.
113. Rothman KJ, Greenland S, Lash TL. Modern Epidemiology. 3rd ed. Philadelphia: Lippincott Williams and Wilkins; 2008.
114. Meyskens FL, Szabo E. Diet and cancer: the disconnect between epidemiology and randomized clinical trials. Cancer Epidemiology Biomarkers & Prevention. 2005;14(6):1366-9.
115. Gibson TM, Ferrucci LM, Tangrea JA, Schatzkin A. Epidemiological and clinical studies of nutrition. Seminars in Oncology. 2010;37(3):282-96.
116. Martínez ME, Marshall JR, Giovannucci E. Diet and cancer prevention: the roles of observation and experimentation. Nature Reviews: Cancer. 2008;8(9):694-703.
117. Zschabitz S, Cheng TY, Neuhouser ML, Zheng Y, Ray RM, Miller JW, et al. B vitamin intakes and incidence of colorectal cancer: results from the Women's Health Initiative Observational Study cohort. American Journal of Clinical Nutrition. 2013;97(2):332-43.
118. Su LJ, Arab L. Nutritional status of folate and colon cancer risk: evidence from NHANES I epidemiologic follow-up study. Annals of Epidemiology. 2001;11(1):65-72.
119. Eussen SJ, Vollset SE, Igland J, Meyer K, Fredriksen Å, Ueland PM, et al. Plasma folate, related genetic variants, and colorectal cancer risk in EPIC. Cancer Epidemiology Biomarkers & Prevention. 2010;19(5):1328-40.
120. Otani T, Iwasaki M, Yamamoto S, Sobue T, Hanaoka T, Inoue M, et al. Alcohol Consumption, Smoking, and Subsequent Risk of Colorectal Cancer in Middle-Aged and Elderly Japanese Men and Women Japan Public Health Center-based Prospective Study. Cancer Epidemiology Biomarkers & Prevention. 2003;12(12):1492-500.
121. Seitz HK, Stickel F. Molecular mechanisms of alcohol-mediated carcinogenesis. Nature Reviews: Cancer. 2007;7(8):599-612.

122. Druesne-Pecollo N, Tehard B, Mallet Y, Gerber M, Norat T, Hercberg S, et al. Alcohol and genetic polymorphisms: effect on risk of alcohol-related cancer. Lancet Oncology. 2009;10(2):173-80.
123. Hurley TD, Edenberg HJ. Genes encoding enzymes involved in ethanol metabolism. Alcohol Research: Current Reviews. 2012;34(3):339-44.
124. Seitz HK, Becker P. Alcohol metabolism and cancer risk. Alcohol Research & Health. 2007;30(1):38-41, 4-7.
125. Murata M, Tagawa M, Watanabe S, Kimura H, Takeshita T, Morimoto K. Genotype difference of aldehyde dehydrogenase 2 gene in alcohol drinkers influences the incidence of Japanese colorectal cancer patients. Cancer Science. 1999;90(7):711-9.
126. Matsuo K, Hamajima N, Hirai T, Kato T, Koike K, Inoue M, et al. Aldehyde dehydrogenase 2 (ALDH2) genotype affects rectal cancer susceptibility due to alcohol consumption. Journal of Epidemiology. 2002;12(2):70-6.
127. Yokoyama A, Muramatsu T, Ohmori T, Yokoyama T, Okuyama K, Takahashi H, et al. Alcohol-related cancers and aldehyde dehydrogenase-2 in Japanese alcoholics. Carcinogenesis. 1998;19(8):1383-7.
128. Ferrari P, McKay JD, Jenab M, Brennan P, Canzian F, Vogel U, et al. Alcohol dehydrogenase and aldehyde dehydrogenase gene polymorphisms, alcohol intake and the risk of colorectal cancer in the European Prospective Investigation into Cancer and Nutrition study. European Journal of Clinical Nutrition. 2012;66(12):1303-8.
129. Yin G, Kono S, Toyomura K, Moore MA, Nagano J, Mizoue T, et al. Alcohol dehydrogenase and aldehyde dehydrogenase polymorphisms and colorectal cancer: the Fukuoka Colorectal Cancer Study. Cancer Science. 2007;98(8):1248-53.
130. Hirose M, Kono S, Tabata S, Ogawa S, Yamaguchi K, Mineshita M, et al. Genetic polymorphisms of methylenetetrahydrofolate reductase and aldehyde dehydrogenase 2, alcohol use and risk of colorectal adenomas: Self-Defense Forces Health Study. Cancer Science. 2005;96(8):513-8.
131. Gao C-M, Takezaki T, Wu J-Z, Zhang X-M, Cao H-X, Ding J-H, et al. Polymorphisms of alcohol dehydrogenase 2 and aldehyde dehydrogenase 2 and colorectal cancer risk in Chinese males. World Journal of Gastroenterology. 2008;14(32):5078.
132. Giovannucci E, Chen J, Smith-Warner SA, Rimm EB, Fuchs CS, Palomeque C, et al. Methylenetetrahydrofolate reductase, alcohol dehydrogenase, diet, and risk of colorectal adenomas. Cancer Epidemiology Biomarkers & Prevention. 2003;12(10):970-9.
133. Bongaerts BW, de Goeij AF, Wouters KA, van Engeland M, Gottschalk RW, Van Schooten FJ, et al. Alcohol consumption, alcohol dehydrogenase 1C (ADH1C) genotype, and risk of colorectal cancer in the Netherlands Cohort Study on diet and cancer. Alcohol. 2011;45(3):217-25.
134. Jung AY, Poole EM, Bigler J, Whitton J, Potter JD, Ulrich CM. DNA methyltransferase and alcohol dehydrogenase: gene-nutrient interactions in relation to risk of colorectal polyps. Cancer Epidemiology Biomarkers & Prevention. 2008;17(2):330-8.
135. Chen J, Ma J, Stampfer MJ, Hines LM, Selhub J, Hunter DJ. Alcohol dehydrogenase 3 genotype is not predictive for risk of colorectal cancer. Cancer Epidemiology Biomarkers & Prevention. 2001;10(12):1303-4.
136. Tiemersma EW, Wark PA, Ocké MC, Bunschoten A, Otten MH, Kok FJ, et al. Alcohol consumption, alcohol dehydrogenase 3 polymorphism, and colorectal adenomas. Cancer Epidemiology Biomarkers & Prevention. 2003;12(5):419-25.
137. Homann N, Konig IR, Marks M, Benesova M, Stickel F, Millonig G, et al. Alcohol and colorectal cancer: the role of alcohol dehydrogenase 1C polymorphism. Alcoholism, Clinical and Experimental Research. 2009;33(3):551-6.
138. Shin A, Joo J, Bak J, Yang H-R, Kim J, Park S, et al. Site-specific risk factors for colorectal cancer in a Korean population. PloS One. 2011;6(8):e23196.
139. Park JY, Dahm CC, Keogh RH, Mitrou PN, Cairns BJ, Greenwood DC, et al. Alcohol intake and risk of colorectal cancer: results from the UK Dietary Cohort Consortium. British Journal of Cancer. 2010;103(5):747-56.

140. Le Marchand L, White KK, Nomura AM, Wilkens LR, Selhub JS, Tiirikainen M, et al. Plasma levels of B vitamins and colorectal cancer risk: the multiethnic cohort study. Cancer Epidemiology Biomarkers & Prevention. 2009;18(8):2195-201.
141. Toriola AT, Kurl S, Laukanen JA, Mazengo C, Kauhanen J. Alcohol consumption and risk of colorectal cancer: the Findrink study. European Journal of Epidemiology. 2008;23(6):395-401.
142. Yeh C-C, You S-L, Chen C-J, Sung F-C. Peanut consumption and reduced risk of colorectal cancer in women: a prospective study in Taiwan. World Journal of Gastroenterology. 2006;12(2):222.
143. Wakai K, Kojima M, Tamakoshi K, Watanabe Y, Hayakawa N, Suzuki K, et al. Alcohol consumption and colorectal cancer risk: findings from the JACC Study. Journal of Epidemiology. 2005;15(Supplement_II):S173-S9.
144. Sanjoaquin M, Appleby P, Thorogood M, Mann J, Key T. Nutrition, lifestyle and colorectal cancer incidence: a prospective investigation of 10 998 vegetarians and non-vegetarians in the United Kingdom. British Journal of Cancer. 2004;90(1):118-21.
145. Su LJ, Arab L. Alcohol consumption and risk of colon cancer: evidence from the national health and nutrition examination survey I epidemiologic follow-up study. Nutrition and Cancer. 2004;50(2):111-9.
146. Shimizu N, Nagata C, Shimizu H, Kametani M, Takeyama N, Ohnuma T, et al. Height, weight, and alcohol consumption in relation to the risk of colorectal cancer in Japan: a prospective study. British Journal of Cancer. 2003;88(7):1038-43.
147. Singh PN, Fraser GE. Dietary risk factors for colon cancer in a low-risk population. American Journal of Epidemiology. 1998;148(8):761-74.
148. Glynn SA, Albanes D, Pietinen P, Brown CC, Rautalahti M, Tangrea JA, et al. Alcohol consumption and risk of colorectal cancer in a cohort of Finnish men. Cancer Causes and Control. 1996;7(2):214-23.
149. Kreger BE, Anderson KM, Schatzkin A, Splansky GL. Serum cholesterol level, body mass index, and the risk of colon cancer. The Framingham Study. Cancer. 1992;70(5):1038-43.
150. Ma Y, Yang Y, Wang F, Zhang P, Shi C, Zou Y, et al. Obesity and risk of colorectal cancer: a systematic review of prospective studies. PloS One. 2013;8(1):e53916.
151. Robsahm TE, Aagnes B, Hjartåker A, Langseth H, Bray FI, Larsen IK. Body mass index, physical activity, and colorectal cancer by anatomical subsites: a systematic review and meta-analysis of cohort studies. European Journal of Cancer Prevention. 2013;22(6):492-505.
152. Ning Y, Wang L, Giovannucci E. A quantitative analysis of body mass index and colorectal cancer: findings from 56 observational studies. Obesity Reviews. 2010;11(1):19-30.
153. Harriss D, Atkinson G, George K, Tim Cable N, Reilly T, Haboubi N, et al. Lifestyle factors and colorectal cancer risk (1): systematic review and meta-analysis of associations with body mass index. Colorectal Disease. 2009;11(6):547-63.
154. Larsson SC, Wolk A. Obesity and colon and rectal cancer risk: a meta-analysis of prospective studies. American Journal of Clinical Nutrition. 2007;86(3):556-65.
155. Moghaddam AA, Woodward M, Huxley R. Obesity and risk of colorectal cancer: a meta-analysis of 31 studies with 70,000 events. Cancer Epidemiology, Biomarkers and Prevention. 2007;16(12):2533-47.
156. Dai Z, Xu YC, Niu L. Obesity and colorectal cancer risk: a meta-analysis of cohort studies. World Journal of Gastroenterology. 2007;13(31):4199-206.
157. Aleksandrova K, Nimptsch K, Pischon T. Obesity and colorectal cancer. Frontiers in Bioscience. 2013;5:61-77.
158. Bardou M, Barkun AN, Martel M. Obesity and colorectal cancer. Gut. 2013;62(6):933-47.
159. World Health Organization. Obesity: Preventing and Managing the Global Epidemic: Report of a WHO Consultation. Geneva: World Health Organisation; 2000.
160. Adams KF, Leitzmann MF, Albanes D, Kipnis V, Mouw T, Hollenbeck A, et al. Body mass and colorectal cancer risk in the NIH–AARP cohort. American Journal of Epidemiology. 2007;166(1):36-45.

161. Bhaskaran K, Douglas I, Forbes H, dos-Santos-Silva I, Leon DA, Smeeth L. Body-mass index and risk of 22 specific cancers: a population-based cohort study of 5· 24 million UK adults. Lancet. 2014;384(9945):755-65.
162. WHO Expert Consultation. Appropriate body-mass index for Asian populations and its implications for policy and intervention strategies. Lancet. 2004;363(9403):157-63.
163. Matsuo K, Mizoue T, Tanaka K, Tsuji I, Sugawara Y, Sasazuki S, et al. Association between body mass index and the colorectal cancer risk in Japan: pooled analysis of population-based cohort studies in Japan. Annals of Oncology. 2012;23(2):479-90.
164. Li H, Yang G, Xiang Y-B, Gao J, Zhang X, Zheng W, et al. Body weight, fat distribution and colorectal cancer risk: a report from cohort studies of 134 255 Chinese men and women. International Journal of Obesity. 2013;37(6):783-9.
165. Jee SH, Yun JE, Park EJ, Cho ER, Park IS, Sull JW, et al. Body mass index and cancer risk in Korean men and women. International Journal of Cancer. 2008;123(8):1892-6.
166. Kitahara CM, Berndt SI, de Gonzalez AB, Coleman HG, Schoen RE, Hayes RB, et al. Prospective investigation of body mass index, colorectal adenoma, and colorectal cancer in the prostate, lung, colorectal, and ovarian cancer screening trial. Journal of Clinical Oncology. 2013;31(19):2450-9.
167. Renehan AG, Flood A, Adams KF, Olden M, Hollenbeck AR, Cross AJ, et al. Body mass index at different adult ages, weight change, and colorectal cancer risk in the National Institutes of Health-AARP Cohort. American Journal of Epidemiology. 2012;176(12):1130-40.
168. Hughes LA, Simons CC, van den Brandt PA, Goldbohm RA, van Engeland M, Weijenberg MP. Body size and colorectal cancer risk after 16.3 years of follow-up: an analysis from the Netherlands Cohort Study. American Journal of Epidemiology. 2011;174(10):1127-39.
169. Laake I, Thune I, Selmer R, Tretli S, Slattery ML, Veierød MB. A prospective study of body mass index, weight change, and risk of cancer in the proximal and distal colon. Cancer Epidemiology Biomarkers & Prevention. 2010;19(6):1511-22.
170. Larsson SC, Rutegård J, Bergkvist L, Wolk A. Physical activity, obesity, and risk of colon and rectal cancer in a cohort of Swedish men. European Journal of Cancer. 2006;42(15):2590-7.
171. Samanic C, Chow WH, Gridley G, Jarvholm B, Fraumeni JF, Jr. Relation of body mass index to cancer risk in 362,552 Swedish men. Cancer Causes and Control. 2006;17(7):901-9.
172. MacInnis RJ, English DR, Hopper JL, Gertig DM, Haydon AM, Giles GG. Body size and composition and colon cancer risk in women. International Journal of Cancer. 2006;118(6):1496-500.
173. MacInnis RJ, English DR, Hopper JL, Haydon AM, Gertig DM, Giles GG. Body size and composition and colon cancer risk in men. Cancer Epidemiology Biomarkers & Prevention. 2004;13(4):553-9.
174. Lin J, Zhang SM, Cook NR, Rexrode KM, Lee IM, Buring JE. Body mass index and risk of colorectal cancer in women (United States). Cancer Causes and Control. 2004;15(6):581-9.
175. Moore LL, Bradlee ML, Singer MR, Splansky GL, Proctor MH, Ellison RC, et al. BMI and waist circumference as predictors of lifetime colon cancer risk in Framingham Study adults. International Journal of Obesity and Related Metabolic Disorders. 2004;28(4):559-67.
176. Saydah SH, Platz EA, Rifai N, Pollak MN, Brancati FL, Helzlsouer KJ. Association of markers of insulin and glucose control with subsequent colorectal cancer risk. Cancer Epidemiology, Biomarkers and Prevention. 2003;12(5):412-8.
177. Terry P, Miller A, Rohan T. Obesity and colorectal cancer risk in women. Gut. 2002;51(2):191-4.
178. Terry P, Giovannucci E, Bergkvist L, Holmberg L, Wolk A. Body weight and colorectal cancer risk in a cohort of Swedish women: relation varies by age and cancer site. British Journal of Cancer. 2001;85(3):346-9.
179. Robsahm TE, Tretli S. Height, weight and gastrointestinal cancer: a follow-up study in Norway. European Journal of Cancer Prevention. 1999;8(2):105-13.
180. Martinez ME, Giovannucci E, Spiegelman D, Hunter DJ, Willett WC, Colditz GA. Leisure-time physical activity, body size, and colon cancer in women. Nurses' Health Study Research Group. Journal of the National Cancer Institute. 1997;89(13):948-55.

181. Spencer EA, Appleby PN, Davey GK, Key TJ. Validity of self-reported height and weight in 4808 EPIC–Oxford participants. Public Health Nutrition. 2002;5(04):561-5.
182. Nyholm M, Gullberg B, Merlo J, Lundqvist-Persson C, Rastam L, Lindblad U. The validity of obesity based on self-reported weight and height: Implications for population studies. Obesity. 2007;15(1):197-208.
183. Merrill R, Richardson J. Validity of self-reported height, weight, and body mass index: findings from the National Health and Nutrition Examination Survey, 2001-2006. Preventing Chronic Disease. 2009;6(4):A121.
184. Reeves GK, Pirie K, Beral V, Green J, Spencer E, Bull D. Cancer incidence and mortality in relation to body mass index in the Million Women Study: cohort study. British Medical Journal. 2007;335(7630):1134.
185. Pischon T, Nöthlings U, Boeing H. Obesity and cancer. Proceedings of the Nutrition Society. 2008;67(02):128-45.
186. Keimling M, Renehan AG, Behrens G, Fischer B, Hollenbeck AR, Cross AJ, et al. Comparison of associations of body mass index, abdominal adiposity, and risk of colorectal cancer in a large prospective cohort study. Cancer Epidemiology Biomarkers & Prevention. 2013;22(8):1383-94.
187. Wang Y, Jacobs EJ, Patel AV, Rodríguez C, McCullough ML, Thun MJ, et al. A prospective study of waist circumference and body mass index in relation to colorectal cancer incidence. Cancer Causes and Control. 2008;19(7):783-92.
188. Oxentenko AS, Bardia A, Vierkant RA, Wang AH, Anderson KE, Campbell PT, et al. Body size and incident colorectal cancer: a prospective study of older women. Cancer Prevention Research. 2010;3(12):1608-20.
189. Levi Z, Kark JD, Barchana M, Liphshitz I, Zavdi O, Tzur D, et al. Measured body mass index in adolescence and the incidence of colorectal cancer in a cohort of 1.1 million males. Cancer Epidemiology Biomarkers & Prevention. 2011;20(12):2524-31.
190. Burton A, Martin R, Galobardes B, Smith GD, Jeffreys M. Young adulthood body mass index and risk of cancer in later adulthood: historical cohort study. Cancer Causes and Control. 2010;21(12):2069-77.
191. Le Marchand L, Wilkens LR, Mi M-P. Obesity in youth and middle age and risk of colorectal cancer in men. Cancer Causes and Control. 1992;3(4):349-54.
192. Lee I-M, Paffenbarger RS. Quetelet's index and risk of colon cancer in college alumni. Journal of the National Cancer Institute. 1992;84(17):1326-31.
193. Kovalchik S. Validity of adult lifetime self-reported body weight. Public Health Nutrition. 2009;12(08):1072-7.
194. Tamakoshi K, Yatsuya H, Kondo T, Hirano T, Hori Y, Yoshida T, et al. The accuracy of long-term recall of past body weight in Japanese adult men. International Journal of Obesity. 2003;27(2):247-52.
195. Dahl AK, Reynolds CA. Accuracy of recalled body weight—a study with 20-years of follow-up. Obesity. 2013;21(6):1293-8.
196. Bassett JK, Severi G, English DR, Baglietto L, Krishnan K, Hopper JL, et al. Body size, weight change, and risk of colon cancer. Cancer Epidemiology Biomarkers & Prevention. 2010;19(11):2978-86.
197. Han X, Stevens J, Truesdale KP, Bradshaw PT, Kucharska-Newton A, Prizment AE, et al. Body mass index at early adulthood, subsequent weight change and cancer incidence and mortality. International Journal of Cancer. 2014;135(12):2900-9.
198. Zhang X, Wu K, Giovannucci EL, Ma J, Colditz GA, Fuchs CS, et al. Early life body fatness and risk of colorectal cancer in u.s. Women and men-results from two large cohort studies. Cancer Epidemiology, Biomarkers and Prevention. 2015;24(4):690-7.
199. Keum N, Greenwood DC, Lee DH, Kim R, Aune D, Ju W, et al. Adult weight gain and adiposity-related cancers: a dose-response meta-analysis of prospective observational studies. Journal of the National Cancer Institute. 2015;107(3):dju428.
200. Song M, Hu FB, Spiegelman D, Chan AT, Wu K, Ogino S, et al. Adulthood weight change and risk of colorectal cancer in the Nurses' Health Study and Health Professionals Follow-up Study. Cancer Prevention Research. 2015;8(7):620-7.

201. Aleksandrova K, Pischon T, Buijsse B, May AM, Peeters PH, Bueno-de-Mesquita HB, et al. Adult weight change and risk of colorectal cancer in the European Prospective Investigation into Cancer and Nutrition. European Journal of Cancer. 2013;49(16):3526-36.
202. Steins Bisschop CN, van Gils CH, Emaus MJ, Bueno-de-Mesquita HB, Monninkhof EM, Boeing H, et al. Weight change later in life and colon and rectal cancer risk in participants in the EPIC-PANACEA study. American Journal of Clinical Nutrition. 2014;99(1):139-47.
203. Rapp K, Klenk J, Ulmer H, Concin H, Diem G, Oberaigner W, et al. Weight change and cancer risk in a cohort of more than 65 000 adults in Austria. Annals of Oncology. 2008;19(4):641-8.
204. Byers T, Sedjo RL. Does intentional weight loss reduce cancer risk? Diabetes, Obesity & Metabolism. 2011;13(12):1063-72.
205. Chen Q, Wang J, Yang J, Jin Z, Shi W, Qin Y, et al. Association between adult weight gain and colorectal cancer: A dose–response meta-analysis of observational studies. International Journal of Cancer. 2015;136(12):2880-9.
206. Giovannucci E. Obesity, gender, and colon cancer. Gut. 2002;51(2):147.
207. Chlebowski RT, Wactawski-Wende J, Ritenbaugh C, Hubbell FA, Ascensao J, Rodabough RJ, et al. Estrogen plus progestin and colorectal cancer in postmenopausal women. New England Journal of Medicine. 2004;350(10):991-1004.
208. Simon MS, Chlebowski RT, Wactawski-Wende J, Johnson KC, Muskovitz A, Kato I, et al. Estrogen plus progestin and colorectal cancer incidence and mortality. Journal of Clinical Oncology. 2012;30(32):3983-90.
209. Ritenbaugh C, Stanford JL, Wu L, Shikany JM, Schoen RE, Stefanick ML, et al. Conjugated equine estrogens and colorectal cancer incidence and survival: the Women's Health Initiative randomized clinical trial. Cancer Epidemiology, Biomarkers and Prevention. 2008;17(10):2609-18.
210. Johnson JR, Lacey JV, Jr., Lazovich D, Geller MA, Schairer C, Schatzkin A, et al. Menopausal hormone therapy and risk of colorectal cancer. Cancer Epidemiology, Biomarkers and Prevention. 2009;18(1):196-203.
211. Hildebrand JS, Jacobs EJ, Campbell PT, McCullough ML, Teras LR, Thun MJ, et al. Colorectal cancer incidence and postmenopausal hormone use by type, recency, and duration in cancer prevention study II. Cancer Epidemiology, Biomarkers and Prevention. 2009;18(11):2835-41.
212. Delellis Henderson K, Duan L, Sullivan-Halley J, Ma H, Clarke CA, Neuhausen SL, et al. Menopausal hormone therapy use and risk of invasive colon cancer: the California Teachers Study. American Journal of Epidemiology. 2010;171(4):415-25.
213. Limsui D, Vierkant RA, Tillmans LS, Wang AH, Weisenberger DJ, Laird PW, et al. Postmenopausal hormone therapy and colorectal cancer risk by molecularly defined subtypes among older women. Gut. 2012;61(9):1299-305.
214. Lin JH, Morikawa T, Chan AT, Kuchiba A, Shima K, Nosho K, et al. Postmenopausal hormone therapy is associated with a reduced risk of colorectal cancer lacking CDKN1A expression. Cancer Research. 2012;72(12):3020-8.
215. Tsilidis KK, Allen NE, Key TJ, Sanjoaquin MA, Bakken K, Berrino F, et al. Menopausal hormone therapy and risk of colorectal cancer in the European Prospective Investigation into Cancer and Nutrition. International Journal of Cancer. 2011;128(8):1881-9.
216. Lin KJ, Cheung WY, Lai JY, Giovannucci EL. The effect of estrogen vs. combined estrogen-progestogen therapy on the risk of colorectal cancer. International Journal of Cancer. 2012;130(2):419-30.
217. Gunter MJ, Hoover DR, Yu H, Wassertheil-Smoller S, Rohan TE, Manson JE, et al. Insulin, insulin-like growth factor-I, endogenous estradiol, and risk of colorectal cancer in postmenopausal women. Cancer Research. 2008;68(1):329-37.
218. Clendenen TV, Koenig KL, Shore RE, Levitz M, Arslan AA, Zeleniuch-Jacquotte A. Postmenopausal levels of endogenous sex hormones and risk of colorectal cancer. Cancer Epidemiology, Biomarkers and Prevention. 2009;18(1):275-81.
219. Lin JH, Zhang SM, Rexrode KM, Manson JE, Chan AT, Wu K, et al. Association between sex hormones and colorectal cancer risk in men and women. Clinical Gastroenterology and Hepatology. 2013;11(4):419-24.e1.

220. Murphy N, Strickler HD, Stanczyk FZ, Xue X, Wassertheil-Smoller S, Rohan TE, et al. A prospective evaluation of endogenous sex hormone levels and colorectal cancer risk in postmenopausal women. Journal of the National Cancer Institute. 2015;107(10):djv210.
221. Bjornerem A, Straume B, Midtby M, Fonnebo V, Sundsfjord J, Svartberg J, et al. Endogenous sex hormones in relation to age, sex, lifestyle factors, and chronic diseases in a general population: the Tromso Study. Journal of Clinical Endocrinology and Metabolism. 2004;89(12):6039-47.
222. Lukanova A, Lundin E, Zeleniuch-Jacquotte A, Muti P, Mure A, Rinaldi S, et al. Body mass index, circulating levels of sex-steroid hormones, IGF-I and IGF-binding protein-3: a cross-sectional study in healthy women. European Journal of Endocrinology of the European Federation of Endocrine Societies. 2004;150(2):161-71.
223. Song Y-M, Sung J, Ha M. Obesity and risk of cancer in postmenopausal Korean women. Journal of Clinical Oncology. 2008;26(20):3395-402.
224. Wang Y, Jacobs EJ, Teras LR, Pavluck AL, Rodriguez C, Thun MJ, et al. Lack of evidence for effect modification by estrogen of association between body mass index and colorectal cancer risk among postmenopausal women. Cancer Causes and Control. 2007;18(8):793-9.
225. Wolin K, Yan Y, Colditz G, Lee I. Physical activity and colon cancer prevention: a meta-analysis. British Journal of Cancer. 2009;100(4):611-6.
226. Botteri E, Iodice S, Bagnardi V, Raimondi S, Lowenfels AB, Maisonneuve P. Smoking and colorectal cancer: a meta-analysis. JAMA. 2008;300(23):2765-78.
227. Chiolero A, Faeh D, Paccaud F, Cornuz J. Consequences of smoking for body weight, body fat distribution, and insulin resistance. American Journal of Clinical Nutrition. 2008;87(4):801-9.
228. Odegaard AO, Koh WP, Yu MC, Yuan JM. Body mass index and risk of colorectal cancer in Chinese Singaporeans. Cancer. 2011;117(16):3841-9.
229. Calle EE, Kaaks R. Overweight, obesity and cancer: epidemiological evidence and proposed mechanisms. Nature Reviews: Cancer. 2004;4(8):579-91.
230. Khalili H, Chan AT. Is diabetes a risk factor for colorectal cancer? Digestive Diseases and Sciences. 2012;57(6):1427-9.
231. Cnattingius S, Lundberg F, Iliadou A. Birth characteristics and risk of colorectal cancer: a study among Swedish twins. British Journal of Cancer. 2009;100(5):803-6.
232. Driver JA, Gaziano JM, Gelber RP, Lee IM, Buring JE, Kurth T. Development of a risk score for colorectal cancer in men. American Journal of Medicine. 2007;120(3):257-63.
233. Lundqvist E, Kaprio J, Verkasalo PK, Pukkala E, Koskenvuo M, Soderberg KC, et al. Co-twin control and cohort analyses of body mass index and height in relation to breast, prostate, ovarian, corpus uteri, colon and rectal cancer among Swedish and Finnish twins. International Journal of Cancer. 2007;121(4):810-8.
234. Ahmed RL, Schmitz KH, Anderson KE, Rosamond WD, Folsom AR. The metabolic syndrome and risk of incident colorectal cancer. Cancer. 2006;107(1):28-36.
235. Bowers K, Albanes D, Limburg P, Pietinen P, Taylor PR, Virtamo J, et al. A prospective study of anthropometric and clinical measurements associated with insulin resistance syndrome and colorectal cancer in male smokers. American Journal of Epidemiology. 2006;164(7):652-64.
236. Lukanova A, Bjor O, Kaaks R, Lenner P, Lindahl B, Hallmans G, et al. Body mass index and cancer: results from the Northern Sweden Health and Disease Cohort. International Journal of Cancer. 2006;118(2):458-66.
237. MacInnis RJ, English DR, Haydon AM, Hopper JL, Gertig DM, Giles GG. Body size and composition and risk of rectal cancer (Australia). Cancer Causes and Control. 2006;17(10):1291-7.
238. Engeland A, Tretli S, Austad G, Bjorge T. Height and body mass index in relation to colorectal and gallbladder cancer in two million Norwegian men and women. Cancer Causes and Control. 2005;16(8):987-96.
239. Otani T, Iwasaki M, Inoue M. Body mass index, body height, and subsequent risk of colorectal cancer in middle-aged and elderly Japanese men and women: Japan public health center-based prospective study. Cancer Causes and Control. 2005;16(7):839-50.

240. Rapp K, Schroeder J, Klenk J, Stoehr S, Ulmer H, Concin H, et al. Obesity and incidence of cancer: a large cohort study of over 145 000 adults in Austria. British Journal of Cancer. 2005;93(9):1062-7.
241. Nilsen TL, Vatten LJ. Prospective study of colorectal cancer risk and physical activity, diabetes, blood glucose and BMI: exploring the hyperinsulinaemia hypothesis. British Journal of Cancer. 2001;84(3):417.
242. Kaaks R, Toniolo P, Akhmedkhanov A, Lukanova A, Biessy C, Dechaud H, et al. Serum C-peptide, insulin-like growth factor (IGF)-I, IGF-binding proteins, and colorectal cancer risk in women. Journal of the National Cancer Institute. 2000;92(19):1592-600.
243. Ford ES. Body mass index and colon cancer in a national sample of adult US men and women. American Journal of Epidemiology. 1999;150(4):390-8.
244. Schoen RE, Tangen CM, Kuller LH, Burke GL, Cushman M, Tracy RP, et al. Increased blood glucose and insulin, body size, and incident colorectal cancer. Journal of the National Cancer Institute. 1999;91(13):1147-54.
245. Gaard M, Tretli S, Urdal P. Blood lipid and lipoprotein levels and the risk of cancer of the colon and rectum. A prospective study of 62,173 Norwegian men and women. Scandinavian Journal of Gastroenterology. 1997;32(2):162-8.
246. Tulinius H, Sigfusson N, Sigvaldason H, Bjarnadottir K, Tryggvadottir L. Risk factors for malignant diseases: a cohort study on a population of 22,946 Icelanders. Cancer Epidemiology, Biomarkers and Prevention. 1997;6(11):863-73.
247. Thune I, Lund E. Physical activity and risk of colorectal cancer in men and women. British Journal of Cancer. 1996;73(9):1134-40.
248. Giovannucci E, Ascherio A, Rimm EB, Colditz GA, Stampfer MJ, Willett WC. Physical activity, obesity, and risk for colon cancer and adenoma in men. Annals of Internal Medicine. 1995;122(5):327-34.
249. Giovannucci E. An updated review of the epidemiological evidence that cigarette smoking increases risk of colorectal cancer. Cancer Epidemiology Biomarkers & Prevention. 2001;10(7):725-31.
250. International Agency for Research on Cancer. Tobacco Smoke and Involuntary Smoking. Lyon, France: International Agency for Research on Cancer; 2004. Report No.: 1017-1606.
251. U.S. Department of Health and Human Services. The Health Consequences of Smoking: A Report of the Surgeon General. Atlanta, GA: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health; 2004.
252. Liang PS, Chen TY, Giovannucci E. Cigarette smoking and colorectal cancer incidence and mortality: systematic review and meta-analysis. International Journal of Cancer. 2009;124(10):2406-15.
253. Leufkens AM, Van Duijnhoven FJ, Siersema PD, Boshuizen HC, Vrieling A, Agudo A, et al. Cigarette smoking and colorectal cancer risk in the European Prospective Investigation into Cancer and Nutrition study. Clinical Gastroenterology and Hepatology. 2011;9(2):137-44.
254. Hannan LM, Jacobs EJ, Thun MJ. The association between cigarette smoking and risk of colorectal cancer in a large prospective cohort from the United States. Cancer Epidemiology Biomarkers & Prevention. 2009;18(12):3362-7.
255. Parajuli R, Bjerkaas E, Tverdal A, Selmer R, Le Marchand L, Weiderpass E, et al. The increased risk of colon cancer due to cigarette smoking may be greater in women than men. Cancer Epidemiology Biomarkers & Prevention. 2013;22(5):862-71.
256. Parajuli R, Bjerkaas E, Tverdal A, Le Marchand L, Weiderpass E, Gram IT. Smoking increases rectal cancer risk to the same extent in women as in men: results from a Norwegian cohort study. BMC Cancer. 2014;14(1):321.
257. Paskett ED, Reeves KW, Rohan TE, Allison MA, Williams CD, Messina CR, et al. Association between cigarette smoking and colorectal cancer in the Women's Health Initiative. Journal of the National Cancer Institute. 2007;99(22):1729-35.
258. Hurley S, Goldberg D, Nelson DO, Lu Y, Henderson K, Bernstein L, et al. Risk of colorectal cancer associated with active smoking among female teachers. Cancer Causes and Control. 2013;24(7):1291-304.

259. Gram IT, Braaten T, Lund E, Le Marchand L, Weiderpass E. Cigarette smoking and risk of colorectal cancer among Norwegian women. Cancer Causes and Control. 2009;20(6):895-903.
260. Gong J, Hutter C, Baron JA, Berndt S, Caan B, Campbell PT, et al. A pooled analysis of smoking and colorectal cancer: timing of exposure and interactions with environmental factors. Cancer Epidemiology Biomarkers & Prevention. 2012;21(11):1974-85.
261. Vangeli E, Stapleton J, Smit ES, Borland R, West R. Predictors of attempts to stop smoking and their success in adult general population samples: a systematic review. Addiction. 2011;106(12):2110-21.
262. Hyland A, Li Q, Bauer JE, Giovino GA, Steger C, Cummings KM. Predictors of cessation in a cohort of current and former smokers followed over 13 years. Nicotine Tob Res. 2004;6 Suppl 3:S363-9.
263. Osler M, Prescott E. Psychosocial, behavioural, and health determinants of successful smoking cessation: a longitudinal study of Danish adults. Tobacco Control. 1998;7(3):262-7.
264. Hyland A, Borland R, Li Q, Yong HH, McNeill A, Fong GT, et al. Individual-level predictors of cessation behaviours among participants in the International Tobacco Control (ITC) Four Country Survey. Tobacco Control. 2006;15 Suppl 3:iii83-94.
265. Hymowitz N, Cummings KM, Hyland A, Lynn WR, Pechacek TF, Hartwell TD. Predictors of smoking cessation in a cohort of adult smokers followed for five years. Tobacco Control. 1997;6 Suppl 2:S57-62.
266. West R, McEwen A, Bolling K, Owen L. Smoking cessation and smoking patterns in the general population: a 1-year follow-up. Addiction. 2001;96(6):891-902.
267. Cox DR. Regression Models and Life-Tables. Journal of the Royal Statistical Society Series B (Methodological). 1972;34(2):187-220.
268. Filozof C, Fernandez Pinilla MC, Fernandez-Cruz A. Smoking cessation and weight gain. Obesity Reviews. 2004;5(2):95-103.
269. Giovannucci E, Rimm EB, Stampfer MJ, Colditz GA, Ascherio A, Kearney J, et al. A prospective study of cigarette smoking and risk of colorectal adenoma and colorectal cancer in US men. Journal of the National Cancer Institute. 1994;86(3):183-91.
270. Giovannucci E, Colditz GA, Stampfer MJ, Hunter D, Rosner BA, Willett WC, et al. A prospective study of cigarette smoking and risk of colorectal adenoma and colorectal cancer in US women. Journal of the National Cancer Institute. 1994;86(3):192-9.
271. Giovannucci E, Martinez ME. Tobacco, colorectal cancer, and adenomas: a review of the evidence. Journal of the National Cancer Institute. 1996;88(23):1717-30.
272. Limsui D, Vierkant RA, Tillmans LS, Wang AH, Weisenberger DJ, Laird PW, et al. Cigarette smoking and colorectal cancer risk by molecularly defined subtypes. Journal of the National Cancer Institute. 2010;102(14):1012-22.
273. Hu MC, Davies M, Kandel DB. Epidemiology and correlates of daily smoking and nicotine dependence among young adults in the United States. American Journal of Public Health. 2006;96(2):299-308.
274. Fernandez E, Schiaffino A, La Vecchia C, Borras JM, Nebot M, Salto E, et al. Age at starting smoking and number of cigarettes smoked in Catalonia, Spain. Preventive Medicine. 1999;28(4):361-6.
275. Everett SA, Warren CW, Sharp D, Kann L, Husten CG, Crossett LS. Initiation of cigarette smoking and subsequent smoking behavior among U.S. high school students. Preventive Medicine. 1999;29(5):327-33.
276. Peto J. That the effects of smoking should be measured in pack-years: misconceptions 4. British Journal of Cancer. 2012;107(3):406-7.
277. Leffondré K, Abrahamowicz M, Siemiatycki J, Rachet B. Modeling smoking history: a comparison of different approaches. American Journal of Epidemiology. 2002;156(9):813-23.
278. Thomas DC. Invited commentary: is it time to retire the "pack-years" variable? Maybe not! American Journal of Epidemiology. 2013;179(3):299-302.
279. Rupprecht LE, Donny EC, Sved AF. Obese Smokers as a Potential Subpopulation of Risk in Tobacco Reduction Policy. Yale Journal of Biology and Medicine. 2015;88(3):289-94.
280. Fearon ER, Vogelstein B. A genetic model for colorectal tumorigenesis. Cell. 1990;61(5):759-67.

281. Botteri E, Iodice S, Raimondi S, Maisonneuve P, Lowenfels AB. Cigarette smoking and adenomatous polyps: a meta-analysis. Gastroenterology. 2008;134(2):388-95. e3.
282. Abrams JA, Terry MB, Neugut AI. Cigarette smoking and the colorectal adenoma-carcinoma sequence. Gastroenterology. 2008;134(2):617-9.
283. Terry MB, Neugut AI. Cigarette smoking and the colorectal adenoma-carcinoma sequence: a hypothesis to explain the paradox. American Journal of Epidemiology. 1998;147(10):903-10.
284. Poole C. Controls who experienced hypothetical causal intermediates should not be excluded from case-control studies. American Journal of Epidemiology. 1999;150(6):547-51.
285. Zhu JZ, Wang YM, Zhou QY, Zhu KF, Yu CH, Li YM. Systematic review with meta-analysis: alcohol consumption and the risk of colorectal adenoma. Alimentary Pharmacology and Therapeutics. 2014;40(4):325-37.
286. Ben Q, Wang L, Liu J, Qian A, Wang Q, Yuan Y. Alcohol drinking and the risk of colorectal adenoma: a dose-response meta-analysis. European Journal of Cancer Prevention. 2015;24(4):286-95.
287. Ben Q, An W, Jiang Y, Zhan X, Du Y, Cai QC, et al. Body mass index increases risk for colorectal adenomas based on meta-analysis. Gastroenterology. 2012;142(4):762-72.
288. Okabayashi K, Ashrafian H, Hasegawa H, Yoo JH, Patel VM, Harling L, et al. Body mass index category as a risk factor for colorectal adenomas: a systematic review and meta-analysis. American Journal of Gastroenterology. 2012;107(8):1175-85.
289. Nishihara R, Morikawa T, Kuchiba A, Lochhead P, Yamauchi M, Liao X, et al. A prospective study of duration of smoking cessation and colorectal cancer risk by epigenetics-related tumor classification. American Journal of Epidemiology. 2013;178(1):84-100.
290. Slattery ML, Curtin K, Anderson K, Ma KN, Ballard L, Edwards S, et al. Associations between cigarette smoking, lifestyle factors, and microsatellite instability in colon tumors. Journal of the National Cancer Institute. 2000;92(22):1831-6.
291. Samowitz WS, Albertsen H, Sweeney C, Herrick J, Caan BJ, Anderson KE, et al. Association of smoking, CpG island methylator phenotype, and V600E BRAF mutations in colon cancer. Journal of the National Cancer Institute. 2006;98(23):1731-8.
292. Curtin K, Samowitz WS, Wolff RK, Herrick J, Caan BJ, Slattery ML. Somatic alterations, metabolizing genes and smoking in rectal cancer. International Journal of Cancer. 2009;125(1):158-64.
293. Poynter JN, Haile RW, Siegmund KD, Campbell PT, Figueiredo JC, Limburg P, et al. Associations between smoking, alcohol consumption, and colorectal cancer, overall and by tumor microsatellite instability status. Cancer Epidemiology, Biomarkers and Prevention. 2009;18(10):2745-50.
294. Chia VM, Newcomb PA, Bigler J, Morimoto LM, Thibodeau SN, Potter JD. Risk of microsatellite-unstable colorectal cancer is associated jointly with smoking and nonsteroidal anti-inflammatory drug use. Cancer Research. 2006;66(13):6877-83.
295. Yamauchi M, Morikawa T, Kuchiba A, Imamura Y, Qian ZR, Nishihara R, et al. Assessment of colorectal cancer molecular features along bowel subsites challenges the conception of distinct dichotomy of proximal versus distal colorectum. Gut. 2012;61(6):847-54.
296. Jass JR. Classification of colorectal cancer based on correlation of clinical, morphological and molecular features. Histopathology. 2007;50(1):113-30.
297. Snover DC. Update on the serrated pathway to colorectal carcinoma. Human Pathology. 2011;42(1):1-10.
298. Rex DK, Ahnen DJ, Baron JA, Batts KP, Burke CA, Burt RW, et al. Serrated lesions of the colorectum: review and recommendations from an expert panel. American Journal of Gastroenterology. 2012;107(9):1315-29.
299. Bettington M, Walker N, Clouston A, Brown I, Leggett B, Whitehall V. The serrated pathway to colorectal carcinoma: current concepts and challenges. Histopathology. 2013;62(3):367-86.
300. Hetzel JT, Huang CS, Coukos JA, Omstead K, Cerda SR, Yang S, et al. Variation in the detection of serrated polyps in an average risk colorectal cancer screening cohort. American Journal of Gastroenterology. 2010;105(12):2656-64.

301. Kahi CJ, Hewett DG, Norton DL, Eckert GJ, Rex DK. Prevalence and variable detection of proximal colon serrated polyps during screening colonoscopy. Clinical Gastroenterology and Hepatology. 2011;9(1):42-6.
302. Figueiredo JC, Crockett SD, Snover DC, Morris CB, McKeown-Eyssen G, Sandler RS, et al. Smoking-associated risks of conventional adenomas and serrated polyps in the colorectum. Cancer Causes and Control. 2015;26(3):377-86.
303. Burnett-Hartman AN, Passarelli MN, Adams SV, Upton MP, Zhu LC, Potter JD, et al. Differences in epidemiologic risk factors for colorectal adenomas and serrated polyps by lesion severity and anatomical site. American Journal of Epidemiology. 2013;177(7):625-37.
304. Fu Z, Shrubsole MJ, Smalley WE, Wu H, Chen Z, Shyr Y, et al. Lifestyle factors and their combined impact on the risk of colorectal polyps. American Journal of Epidemiology. 2012;176(9):766-76.
305. Ji BT, Weissfeld JL, Chow WH, Huang WY, Schoen RE, Hayes RB. Tobacco smoking and colorectal hyperplastic and adenomatous polyps. Cancer Epidemiology, Biomarkers and Prevention. 2006;15(5):897-901.
306. Hecht SS. Tobacco carcinogens, their biomarkers and tobacco-induced cancer. Nature Reviews: Cancer. 2003;3(10):733-44.
307. Hoffmann D, Hoffmann I. The changing cigarette, 1950-1995. Journal of Toxicology and Environmental Health. 1997;50(4):307-64.
308. U.S. Department of Health and Human Services. How Tobacco Smoke Causes Disease: The Biology and Behavioral Basis for Smoking-Attributable Disease: A Report of the Surgeon General. Atlanta, GA: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health; 2010.
309. Cross AJ, Boca S, Freedman ND, Caporaso NE, Huang WY, Sinha R, et al. Metabolites of tobacco smoking and colorectal cancer risk. Carcinogenesis. 2014;35(7):1516-22.
310. Hansen RD, Albieri V, Tjønneland A, Overvad K, Andersen KK, Raaschou–Nielsen O. Effects of smoking and antioxidant micronutrients on risk of colorectal cancer. Clinical Gastroenterology and Hepatology. 2013;11(4):406-15. e3.
311. Doubeni CA, Major JM, Laiyemo AO, Schootman M, Zauber AG, Hollenbeck AR, et al. Contribution of behavioral risk factors and obesity to socioeconomic differences in colorectal cancer incidence. Journal of the National Cancer Institute. 2012;104(18):1353-62.
312. Nordenvall C, Nilsson PJ, Ye W, Nyren O. Smoking, snus use and risk of right- and left-sided colon, rectal and anal cancer: a 37-year follow-up study. International Journal of Cancer. 2011;128(1):157-65.
313. Nöthlings U, Yamamoto JF, Wilkens LR, Murphy SP, Park SY, Henderson BE, et al. Meat and heterocyclic amine intake, smoking, NAT1 and NAT2 polymorphisms, and colorectal cancer risk in the multiethnic cohort study. Cancer Epidemiology, Biomarkers and Prevention. 2009;18(7):2098-106.
314. Hooker CM, Gallicchio L, Genkinger JM, Comstock GW, Alberg AJ. A prospective cohort study of rectal cancer risk in relation to active cigarette smoking and passive smoke exposure. Annals of Epidemiology. 2008;18(1):28-35.
315. Shankar A, Yuan JM, Koh WP, Lee HP, Yu MC. Morbidity and mortality in relation to smoking among women and men of Chinese ethnicity: the Singapore Chinese Health Study. European Journal of Cancer. 2008;44(1):100-9.
316. Weijenberg MP, Aardening PW, de Kok TM, de Goeij AF, van den Brandt PA. Cigarette smoking and KRAS oncogene mutations in sporadic colorectal cancer: results from the Netherlands Cohort Study. Mutation Research. 2008;652(1):54-64.
317. Akhter M, Nishino Y, Nakaya N, Kurashima K, Sato Y, Kuriyama S, et al. Cigarette smoking and the risk of colorectal cancer among men: a prospective study in Japan. European Journal of Cancer Prevention. 2007;16(2):102-7.
318. Yun YH, Jung KW, Bae JM, Lee JS, Shin SA, Min Park S, et al. Cigarette smoking and cancer incidence risk in adult men: National Health Insurance Corporation Study. Cancer Detection and Prevention. 2005;29(1):15-24.
319. Jee SH, Samet JM, Ohrr H, Kim JH, Kim IS. Smoking and cancer risk in Korean men and women. Cancer Causes and Control. 2004;15(4):341-8.

320. Wakai K, Hayakawa N, Kojima M, Tamakoshi K, Watanabe Y, Suzuki K, et al. Smoking and colorectal cancer in a non-Western population: a prospective cohort study in Japan. Journal of Epidemiology. 2003;13(6):323-32.
321. Terry PD, Miller AB, Rohan TE. Prospective cohort study of cigarette smoking and colorectal cancer risk in women. International Journal of Cancer. 2002;99(3):480-3.
322. Tiemersma EW, Kampman E, Bueno de Mesquita HB, Bunschoten A, van Schothorst EM, Kok FJ, et al. Meat consumption, cigarette smoking, and genetic susceptibility in the etiology of colorectal cancer: results from a Dutch prospective study. Cancer Causes and Control. 2002;13(4):383-93.
323. Terry P, Ekbom A, Lichtenstein P, Feychting M, Wolk A. Long-term tobacco smoking and colorectal cancer in a prospective cohort study. International Journal of Cancer. 2001;91(4):585-7.
324. Stürmer T, Glynn RJ, Lee I-M, Christen WG, Hennekens CH. Lifetime cigarette smoking and colorectal cancer incidence in the Physicians' Health Study I. Journal of the National Cancer Institute. 2000;92(14):1178-81.
325. Knekt P, Hakama M, Jarvinen R, Pukkala E, Heliovaara M. Smoking and risk of colorectal cancer. British Journal of Cancer. 1998;78(1):136-9.
326. Kato I, Akhmedkhanov A, Koenig K, Toniolo PG, Shore RE, Riboli E. Prospective study of diet and female colorectal cancer: the New York University Women's Health Study. Nutrition and Cancer. 1997;28(3):276-81.
327. Nordlund LA, Carstensen JM, Pershagen G. Cancer incidence in female smokers: a 26-year follow-up. International Journal of Cancer. 1997;73(5):625-8.
328. Engeland A, Andersen A, Haldorsen T, Tretli S. Smoking habits and risk of cancers other than lung cancer: 28 years' follow-up of 26,000 Norwegian men and women. Cancer Causes and Control. 1996;7(5):497-506.
329. Nyren O, Bergstrom R, Nystrom L, Engholm G, Ekbom A, Adami HO, et al. Smoking and colorectal cancer: a 20-year follow-up study of Swedish construction workers. Journal of the National Cancer Institute. 1996;88(18):1302-7.
330. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, et al. Initial sequencing and analysis of the human genome. Nature. 2001;409(6822):860-921.
331. Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, et al. The sequence of the human genome. Science. 2001;291(5507):1304-51.
332. Chakravarti A, Little P. Nature, nurture and human disease. Nature. 2003;421(6921):412-4.
333. Manolio TA, Bailey-Wilson JE, Collins FS. Genes, environment and the value of prospective cohort studies. Nature Reviews Genetics. 2006;7(10):812-20.
334. Allen N, Sudlow C, Downey P, Peakman T, Danesh J, Elliott P, et al. UK Biobank: Current status and what it means for epidemiology. Health Policy and Technology. 2012;1(3):123-6.
335. Chen Z, Chen J, Collins R, Guo Y, Peto R, Wu F, et al. China Kadoorie Biobank of 0.5 million people: survey methods, baseline characteristics and long-term follow-up. International Journal of Epidemiology. 2011;40(6):1652-66.
336. Riboli E, Hunt KJ, Slimani N, Ferrari P, Norat T, Fahey M, et al. European Prospective Investigation into Cancer and Nutrition (EPIC): study populations and data collection. Public Health Nutrition. 2002;5(6b):1113-24.
337. Tapia-Conyer R, Kuri-Morales P, Alegre-Diaz J, Whitlock G, Emberson J, Clark S, et al. Cohort profile: the Mexico City Prospective Study. International Journal of Epidemiology. 2006;35(2):243-9.
338. UK Biobank. UK Biobank: Protocol for a large-scale prospective epidemiological resource. 2007 Available from: http://www.ukbiobank.ac.uk/wp-content/uploads/2011/11/UK-Biobank-Protocol.pdf?phpMyAdmin=trmKQlYdjjnQIgJ%2CfAzikMhEnx6. [accessed 15th July 2016].
339. Barbour V. UK Biobank: a project in search of a protocol? Lancet. 2003;361(9370):1734-8.
340. UK Biobank. UK Biobank: Report of the integrated pilot phase. 2006 Available from: http://www.ukbiobank.ac.uk/wp-

content/uploads/2011/06/Pilot_report.pdf?phpMyAdmin=trmKQlYdjjnQIgJ%2CfAzikMhEnx6. [accessed 15th July 2016].
341. EurekAlert! UK Biobank gets unanimous backing from international experts after piloting phase. 2006 Available from: http://www.eurekalert.org/pub_releases/2006-08/ub-ubg082006.php. [accessed 1st December 2015].
342. Galea S, Tracy M. Participation rates in epidemiologic studies. Annals of Epidemiology. 2007;17(9):643-53.
343. UK Biobank Ethics and Governance Council. 15th Meeting. Available from: http://egcukbiobank.org.uk/sites/default/files/meetings/EGC15%20agenda%20and%20report.pdf. [accessed 12th March 2016].
344. University of Glasgow. Glasgow contributes to world's biggest medical study. Available from: http://www.gla.ac.uk/news/archiveofnews/2009/january/headline_107470_en.html. [accessed 12th March 2016].
345. UK Biobank Ethics and Governance Council. 16th Meeting. Available from: http://egcukbiobank.org.uk/sites/default/files/meetings/EGC16%20agenda%20and%20report.pdf. [accessed 12th March 2016].
346. Birmingham K. Event Review: Should Genes be Public? Available from: www.bionews.org.uk/page_401802.asp. [accessed 12th March 2016].
347. Rothman KJ, Gallacher JE, Hatch EE. Why representativeness should be avoided. International Journal of Epidemiology. 2013;42(4):1012-4.
348. The Telegraph. Opinion polls failure at 2015 election 'due to unrepresentative samples'. 2016 Available from: http://www.telegraph.co.uk/news/general-election-2015/12107167/Opinion-polls-failure-at-2015-election-due-to-unrepresentative-samples.html. [accessed 18th April 2016].
349. UK Biobank. Repeat assessment data. 2013 Available from: http://biobank.ctsu.ox.ac.uk/~bbdatan/Repeat_assessment_doc_v1.0.pdf. [accessed 15th July 2016].
350. Office for National Statistics. Ethnicity and national identity in England and Wales 2011. 2012 Available from: http://www.ons.gov.uk/peoplepopulationandcommunity/culturalidentity/ethnicity/articles/ethnicityandnationalidentityinenglandandwales/2012-12-11/pdf. [accessed 19th September 2016].
351. Craig R, Mindell J, Hirani V. Health Survey for England 2008. Volume 1: Physical activity and fitness. 2009 Available from: http://digital.nhs.uk/catalogue/PUB00430/heal-surv-phys-acti-fitn-eng-2008-rep-v2.pdf. [accessed 19th Spetember 2016].
352. Bellis M, Hughes K, Cook P, Morleo M. Off measure: how we underestimate the amount we drink. London: Alcohol Concern; 2009.
353. Greenfield TK, Kerr WC. Alcohol measurement methodology in epidemiology: recent advances and opportunities. Addiction. 2008;103(7):1082-99.
354. Gmel G, Rehm J. Measuring alcohol consumption. Contemporary Drug Problems. 2004;31:467.
355. Greenfield TK, Kerr WC, Bond J, Ye Y, Stockwell T. Improving graduated frequencies alcohol measures for monitoring consumption patterns: results from an Australian national survey and a US diary validity study. Contemporary Drug Problems. 2009;36(3-4):705-33.
356. Gmel G, Graham K, Kuendig H, Kuntsche S. Measuring alcohol consumption—should the ‘graduated frequency’approach become the norm in survey research? Addiction. 2006;101(1):16-30.
357. Bellis MA, Hughes K, Jones L, Morleo M, Nicholls J, McCoy E, et al. Holidays, celebrations, and commiserations: measuring drinking during feasting and fasting to improve national and individual estimates of alcohol consumption. BMC Medicine. 2015;13(1):113.
358. Stockwell T, Zhao J, Chikritzhs T, Greenfield TK. What did you drink yesterday? Public health relevance of a recent recall method used in the 2004 Australian National Drug Strategy Household Survey. Addiction. 2008;103(6):919-28.
359. Kerr WC, Stockwell T. Understanding standard drinks and drinking guidelines. Drug and Alcohol Review. 2012;31(2):200-5.

360. Di Castelnuovo A, Costanzo S, Bagnardi V, Donati MB, Iacoviello L, de Gaetano G. Alcohol dosing and total mortality in men and women: an updated meta-analysis of 34 prospective studies. Archives of Internal Medicine. 2006;166(22):2437-45.
361. Corrao G, Rubbiati L, Bagnardi V, Zambon A, Poikolainen K. Alcohol and coronary heart disease: a meta-analysis. Addiction. 2000;95(10):1505-23.
362. Wannamethee G, Shaper AG. Men who do not drink: a report from the British Regional Heart Study. International Journal of Epidemiology. 1988;17(2):307-16.
363. Fillmore KM, Golding JM, Graves KL, Kniep S, Leino EV, Romelsjo A, et al. Alcohol consumption and mortality. I. Characteristics of drinking groups. Addiction. 1998;93(2):183-203.
364. Green CA, Polen MR. The health and health behaviors of people who do not drink alcohol. American Journal of Preventive Medicine. 2001;21(4):298-305.
365. Ng Fat L, Cable N, Shelton N. Worsening of health and a cessation or reduction in alcohol consumption to special occasion drinking across three decades of the life course. Alcoholism, Clinical and Experimental Research. 2015;39(1):166-74.
366. Rehm J, Irving H, Ye Y, Kerr WC, Bond J, Greenfield TK. Are lifetime abstainers the best control group in alcohol epidemiology? On the stability and validity of reported lifetime abstention. American Journal of Epidemiology. 2008;168(8):866-71.
367. Emberson JR, Bennett DA. Effect of alcohol on risk of coronary heart disease and stroke: causality, bias, or a bit of both? Vascular Health and Risk Management. 2006;2(3):239.
368. Fekjaer HO. Alcohol-a universal preventive agent? A critical analysis. Addiction. 2013;108(12):2051-7.
369. Ng Fat L, Cable N, Marmot MG, Shelton N. Persistent long-standing illness and non-drinking over time, implications for the use of lifetime abstainers as a control group. Journal of Epidemiology and Community Health. 2014;68(1):71-7.
370. Caldwell T, Rodgers B, Power C, Clark C, Stansfeld S. Drinking histories of self-identified lifetime abstainers and occasional drinkers: findings from the 1958 British Birth Cohort Study. Alcohol and Alcoholism. 2006;41(6):650-4.
371. Goddard E. Obtaining information about drinking through surveys of the general population. 2001 Available from: http://webarchive.nationalarchives.gov.uk/20160105160709/http://www.ons.gov.uk/ons/guide-method/method-quality/specific/gss-methodology-series/gss-methodology-series--24--obtaining-information-about-drinking-through-surveys-of-the-general-population.pdf. [accessed 19th September 2016].
372. Klatsky AL. Invited commentary: never, or hardly ever? It could make a difference. American Journal of Epidemiology. 2008;168(8):872-5.
373. Liu B, Young H, Crowe FL, Benson VS, Spencer EA, Key TJ, et al. Development and evaluation of the Oxford WebQ, a low-cost, web-based method for assessment of previous 24 h dietary intakes in large-scale prospective studies. Public Health Nutrition. 2011;14(11):1998-2005.
374. Chan W, Brown J, Buss D. Miscellaneous Foods. Fourth Supplement to McCance and Widdowson's The Composition of Foods. Cambridge: Royal Society of Chemistry. 1994.
375. White IR, Royston P, Wood AM. Multiple imputation using chained equations: issues and guidance for practice. Statistics in Medicine. 2011;30(4):377-99.
376. Sterne JA, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. British Medical Journal. 2009;338:b2393.
377. Kenward MG, Carpenter J. Multiple imputation: current perspectives. Statistical Methods in Medical Research. 2007;16(3):199-218.
378. Connor Gorber S, Schofield-Hurwitz S, Hardt J, Levasseur G, Tremblay M. The accuracy of self-reported smoking: a systematic review of the relationship between self-reported and cotinine-assessed smoking status. Nicotine and Tobacco Research. 2009;11(1):12-24.
379. Bondy SJ, Victor JC, Diemert LM. Origin and use of the 100 cigarette criterion in tobacco surveys. Tobacco Control. 2009;18(4):317-23.

380. Klesges RC, Debon M, Ray JW. Are self-reports of smoking rate biased? Evidence from the Second National Health and Nutrition Examination Survey. Journal of Clinical Epidemiology. 1995;48(10):1225-33.
381. Shiffman S. How many cigarettes did you smoke? Assessing cigarette consumption by global report, Time-Line Follow-Back, and ecological momentary assessment. Health Psychology. 2009;28(5):519-26.
382. U.S. Department of Health and Human Services. Preventing Tobacco Use Among Youth and Young Adults: A Report of the Surgeon General. Atlanta, GA: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health; 2012.
383. Mayhew KP, Flay BR, Mott JA. Stages in the development of adolescent smoking. Drug and Alcohol Dependence. 2000;59 Suppl 1:S61-81.
384. Aarts MJ, Lemmens VE, Louwman MW, Kunst AE, Coebergh JW. Socioeconomic status and changing inequalities in colorectal cancer? A review of the associations with risk, treatment and outcome. European Journal of Cancer. 2010;46(15):2681-95.
385. Butterworth AS, Higgins J, Pharoah P. Relative and absolute risk of colorectal cancer for individuals with a family history: a meta-analysis. European Journal of Cancer. 2006;42(2):216-27.
386. Walsh JM, Terdiman JP. Colorectal cancer screening: scientific review. JAMA. 2003;289(10):1288-96.
387. Townsend P, Phillimore P, Beattie A. Health and deprivation: inequality and the North. London: Croom Helm; 1988.
388. Logan RF, Patnick J, Nickerson C, Coleman L, Rutter MD, von Wagner C. Outcomes of the Bowel Cancer Screening Programme (BCSP) in England after the first 1 million tests. Gut. 2012;61(10):1439-46.
389. Craig CL, Marshall AL, Sjostrom M, Bauman AE, Booth ML, Ainsworth BE, et al. International physical activity questionnaire: 12-country reliability and validity. Medicine and Science in Sports and Exercise. 2003;35(8):1381-95.
390. Greenland S, Finkle WD. A critical look at methods for handling missing covariates in epidemiologic regression analyses. American Journal of Epidemiology. 1995;142(12):1255-64.
391. Bray F, Parkin DM. Evaluation of data quality in the cancer registry: principles and methods. Part I: comparability, validity and timeliness. European Journal of Cancer. 2009;45(5):747-55.
392. Parkin DM, Bray F. Evaluation of data quality in the cancer registry: principles and methods Part II. Completeness. European Journal of Cancer. 2009;45(5):756-64.
393. Office for National Statistics. Cancer Registration Statistics, England, 2012. 2014 Available from: http://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/bulletins/cancerregistrationstatisticsengland/2014-06-19/pdf. [accessed 19th September 2016].
394. World Health Organization. International Statistical Classification of Diseases and Related Health Problems. 10th Revision. Geneva, Switzerland: World Health Organization; 1992.
395. Breslow N. Covariance analysis of censored survival data. Biometrics. 1974;30(1):89-99.
396. Machin D, Cheung YB, Parmar MKB. Survival analysis: a practical approach. 2nd ed. Chichester, UK: John Wiley and Sons Ltd; 2006.
397. Royston P, Altman DG. Regression using fractional polynomials of continuous covariates: parsimonious parametric modelling. Applied Statistics. 1994:429-67.
398. Royston P, Ambler G, Sauerbrei W. The use of fractional polynomials to model continuous risk variables in epidemiology. International Journal of Epidemiology. 1999;28(5):964-74.
399. Rockhill B, Newman B, Weinberg C. Use and misuse of population attributable fractions. American Journal of Public Health. 1998;88(1):15-9.
400. Newsom RB. Attributable and unattributable risks and fractions and other scenario comparisons. Stata J. 2013;13:672-98.

401. NHS Choices. New alcohol advice issued. 2016 Available from: http://www.nhs.uk/news/2016/01January/Pages/New-alcohol-advice-issued.aspx. [accessed 17th February 2017].
402. NHS Choices. Physical activity guidelines for adults. Available from: http://www.nhs.uk/Livewell/fitness/Pages/physical-activity-guidelines-for-adults.aspx. [accessed 17th February 2017].
403. Little RJ, Rubin DB. Statistical analysis with missing data. 2nd ed. New York: Wiley; 2002.
404. Little RJA. Missing-Data Adjustments in Large Surveys. Journal of Business & Economic Statistics. 1988;6(3):287-96.
405. Morris TP, White IR, Royston P. Tuning multiple imputation by predictive mean matching and local residual draws. BMC Medical Research Methodology. 2014;14:75.
406. Rubin DB. Multiple imputation for nonresponse in surveys. New York: Wiley; 1987.
407. Moons KG, Donders RA, Stijnen T, Harrell FE, Jr. Using the outcome for imputation of missing predictor values was preferred. Journal of Clinical Epidemiology. 2006;59(10):1092-101.
408. White IR, Royston P. Imputing missing covariate values for the Cox model. Statistics in Medicine. 2009;28(15):1982-98.
409. Klarich DS, Brasser SM, Hong MY. Moderate Alcohol Consumption and Colorectal Cancer Risk. Alcoholism, Clinical and Experimental Research. 2015;39(8):1280-91.
410. Lachenmeier DW, Godelmann R, Witt B, Riedel K, Rehm J. Can resveratrol in wine protect against the carcinogenicity of ethanol? A probabilistic dose-response assessment. International Journal of Cancer. 2014;134(1):144-53.
411. Eigenbrodt ML, Fuchs FD, Hutchinson RG, Paton CC, Goff DC, Jr., Couper DJ. Health-associated changes in drinking: a period prevalence study of the Atherosclerosis Risk In Communities (ARIC) cohort (1987-1995). Preventive Medicine. 2000;31(1):81-9.
412. Roerecke M, Rehm J. Irregular heavy drinking occasions and risk of ischemic heart disease: a systematic review and meta-analysis. American Journal of Epidemiology. 2010;171(6):633-44.
413. Trevisan M, Schisterman E, Mennotti A, Farchi G, Conti S. Drinking pattern and mortality: the Italian Risk Factor and Life Expectancy pooling project. Annals of Epidemiology. 2001;11(5):312-9.
414. Giovannucci E. Alcohol, one-carbon metabolism, and colorectal cancer: recent insights from molecular studies. The Journal of Nutrition. 2004;134(9):2475S-81S.
415. Deng L, Gui Z, Zhao L, Wang J, Shen L. Diabetes mellitus and the incidence of colorectal cancer: an updated systematic review and meta-analysis. Digestive Diseases and Sciences. 2012;57(6):1576-85.
416. Sun L, Yu S. Diabetes mellitus is an independent risk factor for colorectal cancer. Digestive Diseases and Sciences. 2012;57(6):1586-97.
417. Power ML, Schulkin J. Sex differences in fat storage, fat metabolism, and the health risks from obesity: possible evolutionary origins. British Journal of Nutrition. 2008;99(5):931-40.
418. Ibrahim MM. Subcutaneous and visceral adipose tissue: structural and functional differences. Obesity Reviews. 2010;11(1):11-8.
419. Donohoe CL, Doyle SL, Reynolds JV. Visceral adiposity, insulin resistance and cancer risk. Diabetology & Metabolic Syndrome. 2011;3:12.
420. Grundy SM, Neeland IJ, Turer AT, Vega GL. Waist circumference as measure of abdominal fat compartments. Journal of Obesity. 2013;2013:454285.
421. Yamaji T, Iwasaki M, Sasazuki S, Kurahashi N, Mutoh M, Yamamoto S, et al. Visceral fat volume and the prevalence of colorectal adenoma. American Journal of Epidemiology. 2009;170(12):1502-11.
422. Nam SY, Kim BC, Han KS, Ryu KH, Park BJ, Kim HB, et al. Abdominal visceral adipose tissue predicts risk of colorectal adenoma in both sexes. Clinical Gastroenterology and Hepatology. 2010;8(5):443-50.e1-2.

423. Nagata N, Sakamoto K, Arai T, Niikura R, Shimbo T, Shinozaki M, et al. Visceral abdominal fat measured by computed tomography is associated with an increased risk of colorectal adenoma. International Journal of Cancer. 2014;135(10):2273-81.
424. Kang HW, Kim D, Kim HJ, Kim CH, Kim YS, Park MJ, et al. Visceral obesity and insulin resistance as risk factors for colorectal adenoma: a cross-sectional, case-control study. American Journal of Gastroenterology. 2010;105(1):178-87.
425. Keum N, Lee DH, Kim R, Greenwood DC, Giovannucci EL. Visceral adiposity and colorectal adenomas: dose-response meta-analysis of observational studies. Annals of Oncology. 2015;26(6):1101-9.
426. Lee JY, Lee HS, Lee DC, Chu SH, Jeon JY, Kim NK, et al. Visceral fat accumulation is associated with colorectal cancer in postmenopausal women. PloS One. 2014;9(11):e110587.
427. Ashwell M, Hsieh SD. Six reasons why the waist-to-height ratio is a rapid and effective global indicator for health risks of obesity and how its use could simplify the international public health message on obesity. International Journal of Food Sciences and Nutrition. 2005;56(5):303-7.
428. Ashwell M, Gunn P, Gibson S. Waist-to-height ratio is a better screening tool than waist circumference and BMI for adult cardiometabolic risk factors: systematic review and meta-analysis. Obesity Reviews. 2012;13(3):275-86.
429. Lee CM, Huxley RR, Wildman RP, Woodward M. Indices of abdominal obesity are better discriminators of cardiovascular risk factors than BMI: a meta-analysis. Journal of Clinical Epidemiology. 2008;61(7):646-53.
430. Davis SR, Castelo-Branco C, Chedraui P, Lumsden MA, Nappi RE, Shah D, et al. Understanding weight gain at menopause. Climacteric. 2012;15(5):419-29.
431. Shephard RJ. Limits to the measurement of habitual physical activity by questionnaires. British Journal of Sports Medicine. 2003;37(3):197-206.
432. Prince SA, Adamo KB, Hamel ME, Hardt J, Connor Gorber S, Tremblay M. A comparison of direct versus self-report measures for assessing physical activity in adults: a systematic review. The International Journal of Behavioral Nutrition and Physical Activity. 2008;5:56.
433. Troiano RP, Berrigan D, Dodd KW, Masse LC, Tilert T, McDowell M. Physical activity in the United States measured by accelerometer. Medicine and Science in Sports and Exercise. 2008;40(1):181-8.
434. Borghouts LB, Keizer HA. Exercise and insulin sensitivity: a review. International Journal of Sports Medicine. 2000;21(1):1-12.
435. Goodyear LJ, Kahn BB. Exercise, glucose transport, and insulin sensitivity. Annual Review of Medicine. 1998;49:235-61.
436. Triantafillidis JK, Nasioulas G, Kosmidis PA. Colorectal cancer and inflammatory bowel disease: epidemiology, risk factors, mechanisms of carcinogenesis and prevention strategies. Anticancer Research. 2009;29(7):2727-37.
437. Ullman TA, Itzkowitz SH. Intestinal inflammation and cancer. Gastroenterology. 2011;140(6):1807-16.
438. Thomas GA, Rhodes J, Green JT. Inflammatory bowel disease and smoking--a review. American Journal of Gastroenterology. 1998;93(2):144-9.
439. Birrenbach T, Bocker U. Inflammatory bowel disease and smoking: a review of epidemiology, pathophysiology, and therapeutic implications. Inflammatory Bowel Diseases. 2004;10(6):848-59.
440. Thygesen LC, Ersboll AK. When the entire population is the sample: strengths and limitations in register-based epidemiology. European Journal of Epidemiology. 2014;29(8):551-8.
441. Gavrielov-Yusim N, Friger M. Use of administrative medical databases in population-based research. Journal of Epidemiology and Community Health. 2014;68(3):283-7.
442. The NHS Information Centre. Statistics on Smoking: England, 2010. 2010 Available from: http://digital.nhs.uk/catalogue/PUB00684/smok-eng-2010-rep.pdf. [accessed 19th September 2016].
443. Seltzer CC, Bosse R, Garvey AJ. Mail survey response by smoking status. American Journal of Epidemiology. 1974;100(6):453-7.

444. Tolonen H, Dobson A, Kulathinal S. Effect on trend estimates of the difference between survey respondents and non-respondents: results from 27 populations in the WHO MONICA Project. European Journal of Epidemiology. 2005;20(11):887-98.
445. Jackson R, Chambless LE, Yang K, Byrne T, Watson R, Folsom A, et al. Differences between respondents and nonrespondents in a multicenter community-based study vary by gender ethnicity. The Atherosclerosis Risk in Communities (ARIC) Study Investigators. Journal of Clinical Epidemiology. 1996;49(12):1441-46.
446. Blanks RG, Benson VS, Alison R, Brown A, Reeves GK, Beral V, et al. Nationwide bowel cancer screening programme in England: cohort study of lifestyle factors affecting participation and outcomes in women. British Journal of Cancer. 2015;112(9):1562-7.
447. Clarke R, Shipley M, Lewington S, Youngman L, Collins R, Marmot M, et al. Underestimation of risk associations due to regression dilution in long-term follow-up of prospective studies. American Journal of Epidemiology. 1999;150(4):341-53.
448. Greenfield TK, Bond J, Kerr WC. Biomonitoring for Improving Alcohol Consumption Surveys: The New Gold Standard? Alcohol Research: Current Reviews. 2014;36(1):39-45.
449. Leffingwell TR, Cooney NJ, Murphy JG, Luczak S, Rosen G, Dougherty DM, et al. Continuous objective monitoring of alcohol use: twenty-first century measurement using transdermal sensors. Alcoholism, Clinical and Experimental Research. 2013;37(1):16-22.
450. Connor J. Alcohol consumption as a cause of cancer. Addiction. 2016 Jul 21. doi: 10.1111/add.13477. [Epub ahead of print].
451. Lauby-Secretan B, Scoccianti C, Loomis D, Grosse Y, Bianchini F, Straif K. Body Fatness and Cancer — Viewpoint of the IARC Working Group. New England Journal of Medicine. 2016;375(8):794-8.
452. Redeker C, Wardle J, Wilder D, Hiom S, Miles A. The launch of Cancer Research UK's 'Reduce the Risk' campaign: baseline measurements of public awareness of cancer risk factors in 2004. European Journal of Cancer. 2009;45(5):827-36.
453. World Cancer Research Fund. Around 40% of Brits not aware that being overweight increases cancer risk. 2016 Available from: http://www.wcrf-uk.org/uk/media-centre/press-releases/around-40-brits-not-aware-being-overweight-increases-cancer-risk. [accessed 18th May 2016].
454. British Medical Association. Behaviour change, public health and the role of the state – BMA Position Statement. London: BMA; 2012.
455. NHS Choices. What's your poison? A sober analysis of alcohol and heath in the media. 2011 Available from: http://www.nhs.uk/news/2011/10October/Documents/whats_your_poison_1.0.pdf. [accessed 19th September 2016].
456. Daily Mail. Drinking alcohol DOES increase your chance of cancer - and even moderate drinkers are at greater risk. 2016 Available from: http://www.dailymail.co.uk/health/article-3662811/Drinking-alcohol-DOES-increase-chance-cancer-moderate-drinkers-greater-risk.html. [accessed 15th August 2016].
457. Daily Mail. Could BEER help prevent cancer? Hops extract 'activates a chemical in cells that prevents breast tumors'. 2016 Available from: http://www.dailymail.co.uk/health/article-3686678/Could-BEER-help-prevent-cancer-Hops-extract-activates-chemical-cells-prevents-breast-tumors.html. [accessed 15th August 2016].
458. The Telegraph. Red wine could help prevent breast cancer. 2008 Available from: http://www.telegraph.co.uk/news/uknews/2262150/Red-wine-could-help-prevent-breast-cancer.html. [accessed 15th August 2016].
459. The Telegraph. Bacon, ham and sausages 'as big a cancer threat as smoking', WHO to warn. 2015 Available from: http://www.telegraph.co.uk/news/health/news/11950018/Bacon-ham-and-sausages-as-big-a-cancer-threat-as-smoking-WHO-to-warn.html. [accessed 25th August 2016].
460. The Telegraph. Experts attack claims that bacon is 'as big a cancer threat as smoking'. 2015 Available from: http://www.telegraph.co.uk/journalists/laura-donnelly/11951047/Experts-attack-claims-that-bacon-is-as-big-a-cancer-threat-as-smoking.html. [accessed 25th August 2016].

461. Nuffield Council on Bioethics. Public health: ethical issues. London: Nuffield Council on Bioethics; 2007.
462. Public Health England. Alcohol treatment in England 2013-14. 2014 Available from: http://www.nta.nhs.uk/uploads/adult-alcohol-statistics-2013-14-commentary.pdf. [accessed 19th September 2016].
463. Public Health England. Adult obesity and type 2 diabetes. 2014 Available from: http://www.gov.uk/government/uploads/system/uploads/attachment_data/file/338934/Adult_obesity_and_type_2_diabetes_.pdf. [accessed 19th September 2016].
464. Action on Smoking and Health. The economics of tobacco. 2015 Available from: www.ash.org.uk/files/documents/ASH_121.pdf. [accessed 28th August 2016].
465. Marteau TM, Hollands GJ, Fletcher PC. Changing human behavior to prevent disease: the importance of targeting automatic processes. Science. 2012;337(6101):1492-5.
466. World Health Organization Europe. Evidence for the effectiveness and cost–effectiveness of interventions to reduce alcohol-related harm. Geneva: World Health Organisation; 2009.
467. Roberto CA, Swinburn B, Hawkes C, Huang TT, Costa SA, Ashe M, et al. Patchy progress on obesity prevention: emerging examples, entrenched barriers, and new thinking. Lancet. 2015;385(9985):2400-9.
468. Ezzati M, Riboli E. Can noncommunicable diseases be prevented? Lessons from studies of populations and individuals. Science. 2012;337(6101):1482-7.
469. Jha P, Peto R. Global effects of smoking, of quitting, and of taxing tobacco. New England Journal of Medicine. 2014;370(1):60-8.