## Smoking and incidence and mortality of colorectal cancer

## Ranjan Parajuli

A dissertation for the degree of Philosophiae Doctor-2014

Place picture here. Right click, choose Format shape and click "Fill". Choose a picture from your harddisk by clicking "Picture or texture fill" and then the "File" button. Then close. Be sure that the UiT logo in the right corner is visible. Place picture here. Right click, choose Format shape and click "Fill". Choose a picture from your harddisk by clicking "Picture or texture fill" and then the "File" button. Then close. Be sure that the UiT logo in the right corner is visible. Place picture here. Right click, choose Format shape and click "Fill". Choose a picture from your harddisk by clicking "Picture or texture fill" and then the "File" button. Then close. Be sure that the UiT logo in the right corner is visible. Place picture here. Right click, choose Format shape and click "Fill". Choose a picture from your harddisk by clicking "Picture or texture fill" and then the "File" button. Then close. Be sure that the UiT logo in the right corner is visible. Place picture here. Right click, choose Format shape and click "Fill". Choossmen, picture from your harddisk by clicking "Picture ture fill" and then the "File" button. Then close.

# Smoking and incidence and mortality of colorectal cancer 

## Ranjan Parajuli

A dissertation for the degree of Philosophiae Doctor

Department of Community Medicine
Faculty of Health Sciences
UiT, The Arctic University of Norway

Tromsø, Norway


NORWEGIAN CANCER SOCIETY

## Acknowledgements

This thesis is the part of PhD program offered by the Institute of Community Medicine, Faculty of Health Sciences, UiT The Arctic University of Norway.

First of all, I would like to dedicate my vote of thanks to Prof. Dr. Inger Torhild Gram, my main supervisor. Thank you very much for your supervision, encouragements, critics and comments due to which I have been able to accomplish my work. I would also like to thank to Prof. Dr. Elisabete Weiderpass, my co-supervisor for her constant support and guidance during my study period. Even though you were travelling in different part of the world, you always had a time to send us your feedback. It was a great pleasure working under you both!! I am truly indebt with Eivind Bjerkaas who has been my colleague and office partner for all these years. Thank you for the co-operation, understanding, and constant creative support throughout these years. You were always the source of inspiration and an ever-ready help whenever I needed. It was truly a great opportunity to meet a great friend like you, thank you very much!!!

I am very much thankful to my co-authors senior researchers Aage Tverdal and Randi Selmer from Norwegian Institute of Public Health and Prof. Loic Le Marchand from University of Hawaii for their help and support. My thanks go to Professor Anders Engeland for his suggestions and help during the merging of the datasets.

My sincere thanks go to Knut Hansen who dedicated his time to help us to organize the different health surveys datasets. Knut, without you it would have been very difficult to come up with final dataset file. Ilene Brill, many thanks to you as well that you took a chance to visit us here in Tromsø all the way from USA to help us further in the data merging process.

My thanks also go to my friend Ass. Prof. Kjell Halvorsen, a true inspiration and a happy man who taught me to handle the ups and downs of PhD period. Eivind and I, we both are very happy that we shared the office and the talks with you though it was for a short time period. I am also grateful to Prof. Dr. Ganesh Acharya who has always been supportive and helpful and yes, there are many things to learn from you!!!

My vote of thanks also goes to Prof. Dr. Purusotam Basnet for his support and words of encouragement during my learning period. I would also like to thank Dr. Sanat Wagle for his support.

Heartfelt thanks to Prof. Eilliv Lund and system epidemiology group. And many thanks Prof. Bjarne Jacobsen, Ass. Prof. Bjørn Straume, Prof. Maja-Lisa Løchen, Prof. Magritt Brustad, Prof. Tom Wilsgård, Ass. Prof. Tonje Braaten, Ass. Prof. Tormod Brenn and Dr. Guri Skeie for being very helpful and supportive during my study years!

Sincere thanks to Prof. Odd Mørkve and Dr. Jon Hardie from University of Bergen, my previous supervisors and source of inspiration in the research field. It is due to your faith that I have been able to pursue my research career after Master's degree from Centre for International Health, University of Bergen.

I would also like to express my sincere thanks to my friends Anita, Ammar, Bent-Martin Kristin Borch, Karina, Laila Toril for your kind support and friendship. Thank you very much for being available for me whenever I needed a help from you!!
'Thank you’ my EPINOR colleagues Annette, Line, Niklas, Oleg, Oxana, Trond, Unni, and for your support and friendship. And I am also grateful to all my friends who have contributed directly or indirectly and wished me good luck for my work.

I want to express my gratitude to administrative staffs Anne Fismen, Mona Ingebristen, Gerd Furumo, Trine Jensen for providing me with the necessary information and kind help. Thank you Mona, for your ever ready support and help.

I am very much in debt to Nawang Namgyal and Trine Andreassen. You have been always supportive and friendly towards me and my family. Thank you so much!! Many thanks to Trudy Pedrix-Thoma for her help during the final stage of my writing. It was really kind of you to help me up during the last hours.

I would also like to take an opportunity to thank Norwegian Cancer Society 'Kreftforeningen' who provided the grant for this PhD project. I would like to thank all the Norwegian health surveys participants.

And last but not the least, my heartfelt gratitude goes to my parents, my in-laws, my wife Madhu Wagle Parajuli, my brothers, family members, relatives and my friends for their encouragement and faith on me. My love and blessings are to my sons, Rehan and Riwaz for having a great patience to let me do my work. Thank you! You all are very special.

Finally thank you God for giving me enough strength, patience and courage to finish my work.

## Table of contents

Acknowledgements ..... 2
Summary ..... 6
List of papers ..... 8
Paper I ..... 8
Paper II ..... 8
Paper III ..... 8
List of figures and tables ..... 9
Abbreviations ..... 10
1 Introduction ..... 11
1.1 Definition and epidemiology of colorectal cancer ..... 11
1.2 Incidence of colorectal cancer in Nordic countries ..... 14
1.2 Colorectal cancer in Norway ..... 16
1.3 Prevalence of smoking ..... 19
1.3.1 Global prevalence ..... 19
1.3.2 Prevalence of smoking in the Nordic countries ..... 20
1.3.3 Prevalence of smoking in Norway ..... 21
1.4 Assessment of risk factors for colorectal cancer ..... 23
1.4.1 Non-modifiable risk factors ..... 23
1.4.2 Modifiable risk factors ..... 25
1.5 Smoking and colorectal cancer ..... 27
2 Aims of the thesis ..... 29
3 Materials and Methods ..... 30
3.1 Study population ..... 30
3.1 Pooling Datasets ..... 34
3.3 Exposure information ..... 35
3.4 Follow-up and endpoints ..... 36
3.5 Statistical analyses ..... 37
3.6 Ethical aspects ..... 38
4 Results - summary of papers ..... 40
4.1 Paper I ..... 40
4.2 Paper II ..... 41
4.3 Paper III ..... 42
5 Discussion ..... 44
5.1 Methodological issues ..... 44
5.1.1 Internal Validity ..... 44
Selection bias (Paper I-III) ..... 46
Information bias (Paper I-III) ..... 49
Confounding and statistical analyses (Paper I-III) ..... 49
5.1.2 External validity ..... 51
5.2 Discussion of the main results ..... 52
5.2.1 Gender differences in smoking related colon cancer and by location ..... 52
5.2.2 Smoking related risk of rectal cancer among women is same as in men ..... 54
5.2.3 Smoking increases the risk of CRC Mortality ..... 55
6 Conclusions ..... 57
$7 \quad$ Implications for public health practice and further perspectives ..... 58
8 Erratum ..... 59
9 References ..... 60

## Paper I

## Paper II

Paper III
Appendices

## Summary

Smoking is one of the most important causes of cancer and premature death worldwide. Two different reports, the most recent monograph published by International Agency for Research on Cancer (IARC) in 2012 and the Unites States Surgeon General's report of 2014, concluded that smoking is risk factor for both colon and rectal cancer. In addition to being one of the most common cancers in Norway, mortality from colorectal cancer (CRC) is also high. The main aim of this thesis was to examine the association between smoking and CRC incidence and mortality overall and by gender. We examined the association between smoking and colon cancer by location and gender (Paper I), rectal cancer by gender (Paper II) and CRC mortality by subsite and gender (Paper III).

The cohort included 652,792 Norwegians ( $49 \%$ men) recruited from four Norwegian health screening surveys. These surveys were conducted between 1972 and 2003: the Oslo study I (1972-1973), the Norwegian counties study (1974-1988), the 40 years cohort (1985-1999) and the Cohort of Norway (CONOR, 1994-2003). The participation rate for the different surveys varied from 56-88\%.

Women ever smokers had a $19 \%$ and men ever smokers had $8 \%$ increased risk of colon cancer. Furthermore, women ever smokers had an increased risk of proximal colon cancer compared to men ever smokers (Paper I). Ever smokers had an increased risk of rectal cancer at around $25 \%$ and the risk increased was similar for men and women (Paper II). Men and women ever smokers had a similar increased risk of CRC mortality of about $20 \%$. The risk of rectal and proximal colon cancer mortality was most pronounced among men and women smokers, respectively (Paper III).

In conclusion, smoking increased the risk of colon cancer, especially proximal colon cancer among women. Furthermore, smoking increased the risk of rectal cancer, with a similar risk
being observed among women as in men ever smokers. Smoking is associated with increased CRC mortality among both men and women. The risk of rectal and proximal cancer mortality was most pronounced among men and women smokers, respectively.

## List of papers

This thesis is based on the three papers listed below:

## Paper I

Parajuli R, Bjerkaas E, Tverdal A, Selmer R, Le Marchand L, Weiderpass E, Gram IT. The increased risk of colon cancer due to cigarette smoking may be greater in women than men. Cancer Epidemiol Biomarkers Prev.2013; 22(5), 862-71. PubMED:PMID 23632818)

## Paper II

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 (submitted)

## Paper III

Parajuli R, Bjerkaas E, Tverdal A, Le Marchand L, Weiderpass E, Gram IT. Cigarette smoking and colorectal cancer mortality among 602,242 Norwegian males and females. Clinical Epidemiology, Dovepress (Online)

## List of figures and tables

Figure 1: Worldwide estimated age standardized rates of CRC incidence and mortality rates per 100,000 by gender. (Globocan 2012, IARC) ..... 13
Figure 2: Estimated age standardized CRC incidence and mortality rates per 100,000 in Northern Europe by gender (Globocan 2012, IARC) ..... 14
Figure 3: Age standardized rate of CRC incidence and mortality rates per 100,000 in the Nordic countries 1972 and 2007 (NORDCAN) ..... 15
Figure 4: Age standardized incidence rate of colon cancer by gender in Norway (1972-2011) ..... 16
Figure 5: Age standardized incidence rate of rectal cancer by gender in Norway (1972-2011).. 17Figure 6: Age standardized incidence rate of CRC by gender in Norway (1972-2011)17
Figure 7: Estimates of age standardized incidence and mortality rate per 100,000 for different cancer sites in Norway by gender (Globocan 2012) ..... 18
Figure 8: A descriptive model of cigarette epidemic in developed countries (Lopez et al. 1994). ..... 20
Figure 9: Men and women current smokers aged 16-74 years old since 1973-2009 ..... 22
Figure 10: Pathway for causation of cancer by carcinogens in tobacco smoke ..... 28
Figure 11: Detailed flowchart of participants from the different surveys ..... 33
Figure 12: A systematic approach to bias ..... 45
Figure 13: The prevalence of current smokers included in surveys by gender ..... 48
Figure 14: The prevalence of current smokers aged 16-74 years from 1973-2003 in Norway by gender ..... 48
Table 1: List of different surveys included in the study ..... 32
Table2: Age at enrolment of participants included in different health surveys ..... 47

## Abbreviations

| ASR | Age standardized rates |
| :--- | :--- |
| BMI | Body mass index |
| CI | Confidence interval |
| CONOR | Cohort of Norway |
| CRC | Colorectal cancer |
| DNA | Deoxyribonucleic acid |
| EPIC | European Prospective Investigation into Cancer and Nutrition |
| FAP | Familial adenomatous polyposis |
| HNPCC | Hereditary non polyposis colorectal cancer |
| HR | Hazard ratio |
| HRT | Hormonal replacement therapy |
| IARC | International Agency for Research on Cancer |
| IBD | Inflammatory bowel disease |
| PAH | Polycyclic Aromatic Hydrocarbons |
| REK | Regional komité for medisinsk og helsefaglig forskningsetikk (Regional |
| WHO | Comealth Organization |

## 1 Introduction

This thesis describes the association between cigarette smoking and colorectal cancer (CRC) incidence and mortality overall and by subsite among Norwegian men and women who participated in four different Norwegian health surveys.

### 1.1 Definition and epidemiology of colorectal cancer

In 2012, there were around 14 million new cancer cases (all types combined), 8 million cancer deaths and around 32 million people were living with cancer worldwide. Fifty-seven percent (8 million) of these new cancer cases and 65\% (5.3 million) of cancer deaths occurred in low and medium income countries (1).

CRC is one of the major causes of morbidity and mortality around the world (2). CRC is confined to the main parts of large intestine, the colon and rectum. Adenocarcinoma is the predominant histological subtype and begins as adenomatous polyps before reaching the malignant stage. The progression from adenomatous polyps to carcinoma occurs with potential damage to DNA. Other histological subtypes of CRC include carcinoid tumors, gastrointestinal, stromal tumors, lymphomas and sarcomas. More than 95\% CRC are sporadic, originating in individual without significant genetic or hereditary risk factor (3). If the diagnosis is made early, CRC is highly treatable. CRC is known as disease of western world as it is more prevalent in high-income countries.

Globally, CRC is the third most common cancer in men and the second most common cancer in women representing about $9 \%$ and $10 \%$ of all incident cancer respectively (2). CRC incidence rates worldwide have changed with time, but usually men have higher rates compared to women (2). There is a wide variation in CRC incidence across the world population but the patterns of variation in men and women are similar. The CRC incidence rates vary tenfold, with the highest estimates in Australia and New Zealand (age-standardized
incidence rate, ASR 44.8 and 32.2 per 100,000 in men and women, respectively) and the lowest in Western Africa (4.5 and 3.8 per 100,000 in men and women, respectively) (1). There is also a geographical difference in the global occurrence of CRC. High-income countries usually have higher incidence rates and accounts for almost $55 \%$ of all incident cases CRC worldwide (4).

CRC incidence rates are decreasing in the United States, whereas in Northern and Western Europe CRC incidence rates are stabilizing. However, high income countries like Japan, Singapore, and some Eastern European countries are showing a substantial increase in CRC incidence (5;6).

CRC accounts for $8 \%$ of all cancers deaths, which makes it the fourth most common cause of death from cancer worldwide (7). It has been reported that about $12 \%$ of CRC deaths are attributed to smoking $(6 ; 8)$. CRC mortality rates are lower in women than men except in the Caribbean region (7). Worldwide, CRC mortality rates vary less than CRC incidence rates (six fold in men, and four fold in women). The highest mortality rates are observed in Central and Eastern Europe (20.3 and 11.7 per 100.000 among men and women, respectively) and lowest in western Africa (3.5 and 3.0 per 100,000 among men and women, respectively) (1). In the United States, it is the third most common cause of cancer death although the overall mortality rates have decreased by $2.8 \%$ and $2.6 \%$ per year in men and women, respectively since 1998 (9). CRC is the second most common cause of cancer deaths in Europe (1). Latest CRC incidence and mortality rates worldwide are shown in figure 1.


Figure 1: Worldwide estimated age standardized rates of CRC incidence and mortality rates per 100,000 by gender. (Globocan 2012, IARC)

### 1.2 Incidence of colorectal cancer in Nordic countries



Figure 2: Estimated age standardized CRC incidence and mortality rates per 100,000 in Northern Europe by gender (Globocan 2012, IARC)

Figure 2 illustrates the present CRC incidence and mortality rates in Northern Europe among men and women.

```
Colorectal (1972)
ASR (World) age 0-85+
```



```
Colorectal (2007)
ASR (World) age 0-85+
```



```
NORDCAN © Association of the Nordic Cancer Registries (15.3.2014)
```

Figure 3: Age standardized rate of CRC incidence and mortality rates per 100,000 in the Nordic countries 1972 and 2007 (NORDCAN)

Figure 3 shows the incidence and mortality rates in the Nordic countries during 1972 and 2007 that is the beginning and end of our study period, respectively. Denmark had the highest incidence rate back in 1972 both among men and women. By 2007, Norway and Denmark were observing almost similar CRC incidence rates. Norwegian women had slightly higher incidence rate compared to Danish women. However, regarding mortality rates, Icelandic men had the highest rates followed by Danish men during 1972 whereas by 2007 highest rates were observed in Denmark and Norway. Danish men had highest CRC mortality rate whereas the rates were highest among Norwegian women in 2007.

### 1.2 Colorectal cancer in Norway

Over the last half century, Norway has experienced one of the most rapid and steady rises in CRC incidence. In the late 1950s, the age standardized incidence rate for colon cancer was 10 per 100,000 for both men and women. The incidence rate of rectal cancer in the same period was approximately around 7 and 4 for per 100,000 for men and women, respectively. By the beginning of 1970s, the incidence rate of colon cancer was around 14 for both men and women; the incidence rate of rectal cancer was 11 and 8 per 100,000 for men and women, respectively. Current incidence rates of both colon and rectal cancer are more than double what they were 50 years ago for both men and women. The present age standardized five year incidence rate of CRC for year 2007-2011 is 43 for and 35 per 100,000 for men and women respectively. Among men, the incidence rate of colon and rectal cancer is 26 and 17 per 100,000 respectively. Similarly, for women, the incidence of colon and rectal cancer is 24 and 11 per 100,000 respectively (10). The corresponding figures for CRC incidence rate and by subsite in Norway by gender from 1972-2011 are presented in the figure below (Figure 4, 5, and 6).


Figure 4: Age standardized incidence rate of colon cancer by gender in Norway (1972-2011) (Norwegian Cancer Registry, 2013)


Figure 5: Age standardized incidence rate of rectal cancer by gender in Norway (1972-2011) (Norwegian Cancer Registry, 2013)


Figure 6: Age standardized incidence rate of CRC by gender in Norway (1972-2011) (Source: Norwegian Cancer Registry, 2013)

The colon cancer incidence rates among men and women are almost similar but men have higher incidence of rectal cancer than women. The gender difference in CRC incidence is due to men having more rectal but not colon cancer than women.


Figure 7: Estimates of age standardized incidence and mortality rate per 100,000 for different cancer sites in Norway by gender (Globocan 2012)

Figure 7 shows the ASR for different cancer in Norwegian by gender in 2012. In 2002, women in Norway had the highest CRC incidence rate in Europe and second highest incidence rate worldwide, only surpassed by women in New Zealand (11). In addition to being one of the most common cancers among Norwegian, CRC is also a cancer type with a high mortality. The latest report showed that in Norway, the CRC mortality rate is ranked
second after lung cancer among women and third after lung and prostate cancer among men (10).

### 1.3 Prevalence of smoking

### 1.3.1 Global prevalence

There are an estimated 1.3 billion smokers worldwide and that number is expected to increase to 1.6 billion by $2025(12 ; 13)$. Seventy-three percent of smokers are from low and medium income countries. Smoking is one of the major leading preventable causes of death in the world (13-15) and attributed to approximately 6 million premature deaths each year globally. If prevention measures are not implemented soon, the deaths toll could reach approximately 8 million by 2030. Recent report on tobacco from World Health Organization (WHO) reported that in the 20th century almost 100 million deaths have been caused by tobacco smoking and if this trend continues further, one billion smoking related deaths will occur in the 21st century (13).

A four stage model for describing the effects of smoking on mortality was purposed by Lopez and colleagues almost 2 decades ago (16). Women in high-income countries lagged behind men by 20-30 years in relation to smoking and its attributed mortality. This model was further reviewed in 2012 and the predictions matched recent trends in smoking and smoking related mortality (Figure 8). The authors concluded that the model reflected the situation of many high income countries reasonably well with a few exceptions in low and medium income countries (17).


Figure 8: A descriptive model of cigarette epidemic in developed countries (Lopez et al. 1994)
Stages of the cigarette epidemic on entering its second century (Thune et al 2012):
(Reprinted with permission from BMJ publisher group)

### 1.3.2 Prevalence of smoking in the Nordic countries

In 1920, Denmark had the highest prevalence of smoking in the Nordic countries. A report from 2006 showed the highest prevalence in Denmark and Norway (25 and 24, respectively), and the lowest prevalence in Sweden and Iceland (18). Direct comparisons of the smoking prevalence in Nordic countries are somewhat difficult as the data on smoking habits are collected in different age groups. However, in all of the Nordic countries a decreasing trend in the prevalence of smoking was associated with an increased level of education (19).

### 1.3.3 Prevalence of smoking in Norway

The trend of smoking prevalence for men current smokers has been different from that of women in Norway. The prevalence of smoking among men peaked at $65 \%$ in the late 1950 's; and then decreased to $50 \%$ in 1975 and $33 \%$ in 1999. This decrease continued through 2007, when the prevalence of smoking among men was $50 \%$ lower than that in the 1970s. This is quite different from the corresponding figures of smoking prevalence among women. In 1954, the prevalence of smoking among women which was $23 \%$ in 1954, peaked at $37 \%$ in 1970 and then stabilized to $32 \%$ for the rest of the century. After 2002, a decline in the prevalence of smoking was seen among women and by 2007 which is the end of our follow-up period; the prevalence was similar in both men and women $(18 ; 20 ; 21)$. By the year 2013, $15 \%$ Norwegian men and women were current smokers (22). This smoking pattern is in accordance with the tobacco epidemic stages model suggested by Lopez et al. almost 20 years ago (16) which suggested that the smoking-attributed mortality for women, will in the same way as the smoking prevalence, lagged behind that of men and both will peak at a lower level than that of men. The difference in smoking habits is one of the main explanations for social inequalities in health in Norway. Recently, it has been reported that Norway is one of the four countries along with Canada, Iceland and Mexico that are successful in achieving reductions of smoking prevalence in both men and women by more than $50 \%$ (23). Figure 9 shows the prevalence of current smokers by gender in Norway between years 1973-2013.


Figure 9: Men and women current smokers aged 16-74 years old since 1973-2009
Source: Statistics Norway

### 1.4 Assessment of risk factors for colorectal cancer

### 1.4.1 Non-modifiable risk factors

Age

Increased life span is one of the contributors for increasing number of cancer cases and CRC is no exception (24). CRC is common in older age groups: people aged 50 years and older accounting for more than $90 \%$ of cases and CRC incidence is low among people aged less than 50 years (25). However, recent trends show that CRC incidence is also increasing among those under 50 years of age $(26 ; 27)$.

## Gender

As previously mentioned, CRC incidence and mortality rates are generally higher among men than women (6) and this difference may reach 35-40\% higher in men compared to women(9). Differences by gender in CRC incidence are more obvious for rectal cancer which has a higher incidence among men. The reason for this difference is difficult to explain but may be partly due to exposures to different risk factors and hormones (28).

## Geographical variations and race

CRC prevalence varies according to geographical locations and race. The number of CRC cases is declining in the United States, and stabilizing in most of Northern and Western Europe (25;29). Although, rates are low in Asia and Africa, CRC incidence is increasing in countries like Japan, Singapore and most Eastern European countries.

## Adenomatous polyps

Adenomatous polyps are recognized precursor lesions of CRC and are common after 50 years of age. They represent almost two-thirds of colorectal adenomas and have a high potential to progress to malignancy. The majority of CRC develop from adenomatous polyps through a
series of genetic changes (30) but only around $10 \%$ of adenomatous polyps develop into cancer (31). An association between cigarette smoking and adenomatous polyps has been reported recently and it was suggested that smoking could play an important role in both the formation and aggressiveness of adenomatous polyps (32;33).

## Inflammatory bowel diseases

Inflammatory bowel diseases (IBD) such as ulcerating colitis and Crohn's disease might predispose to CRC development though these diseases account for very few cases of CRC in the general population and only around $15 \%$ of all CRC deaths occur among individuals with IBD $(34 ; 35)$. Factors such as early age at IBD diagnosis, longer duration of symptoms and severity of dysplasia and inflammation increase the risk of CRC.

## Family and personal history of colorectal cancer or adenomatous polyps

A family history of CRC is a well-established risk factor (28) and is associated with an increased risk of the CRC (36). Individuals with a family history of CRC and colorectal adenomas mainly adenomatous polyps have higher risk of CRC (37). The risk of CRC is increases when a first degree relative has one or more colorectal adenomas mainly adenomatous polyps (38) and the risk is doubled when a first degree relative is affected with CRC. Similarly, individuals with multiple relatives affected with CRC who were diagnosed at a young age have a risk of CRC that is three to six times than that of general population (39). Almost $20 \%$ of all CRC cases have a close relative who have been diagnosed with the same cancer (40). Person who had CRC are more likely to develop it again in other areas of colon and rectum. This can occur even when the first cancer is removed completely. The risk further increases if the first cancer is diagnosed at 60 years of age or younger (9). Furthermore, person with previous adenomatous polyps are in increased risk of CRC and this is more probable if the polyps were multiple and were of large sizes (41).

## Genetic risk factor

The risk of CRC associated with hereditary conditions is about 5 to $10 \%$ (42). The two types of hereditary conditions are familial adenomatous polyps (FAP) and lynch syndrome, which is also known as hereditary non-polyposis colorectal cancer (HNPCC). The genes that mutate and lead to carcinogenesis have been identified in both of these conditions. MLH1 and MLH2 are responsible for mutations in individuals with HNPCC (43) whereas APC genes are responsible for mutation in FAP (44). HNPCC is the most common of these genetic syndromes and accounts almost 2-4\% of CRC (45), whereas AFP accounts for less than $1 \%$ (46).

### 1.4.2 Modifiable risk factors

## Physical activity and obesity

The association between a high level of physical activity and decrease colon and rectal risk of cancer has been reported previously in a recent meta-analysis which included 52 cohort and case control studies (47). The study reported around a $20-30 \%$ decreased risk of colon cancer among physically active individuals compared with less active ones. Similarly, another metaanalysis concluded that physical activity is associated with reduced risk of both proximal and colon cancer which did not differ by location (48). Lack of physical activity can also lead to obesity, another major risk factor for CRC (49), but a high level of physical activity can lower the risk of CRC even without the significant weight loss (50). Nevertheless, many studies have supported the notion that obesity leads to the development of CRC, and have reported that obesity as an independent risk factor (51-56).

## Diet

Diet is a major modifiable risk factor for CRC. It has been reported that changes in dietary patterns can reduce the CRC burden by $70 \%(49 ; 57)$. Diets that are high in fat and high meat consumption have been implicated in the development of CRC (49;58;59). Diets consisting of large amounts of red meat and highly refined carbohydrates increase the risk of CRC as do diets low in vegetables and fruits (50;60-62).

## Alcohol consumption

The IARC has concluded that alcohol consumption is a potential risk factor for CRC (33). Indeed, alcohol consumption is one of the most important modifiable risk factors for all human cancers (63). Heavy alcohol consumption is linked to an increased CRC and could even give rise to CRC at younger age (8;64). Metabolic product of alcohol such as acetaldehyde is considered to be carcinogenic (65). Alcohol can also work as a solvent which could allow other carcinogenic molecules into the colon and rectum mucosa (66). Similarly, an individual with high alcohol consumption and a diet low in essential nutrients is more vulnerable to the carcinogenic effects of alcohol. Several meta-analysis and pooled studies carried out in different parts of the world reported an increased risk of CRC with high regular alcohol consumption (67-75).

## Medications, supplements and hormonal replacement therapy

There is growing evidence that COX inhibitors such as aspirin, calcium supplements and hormonal replacement therapy (HRT) may have preventive effects towards the CRC $(9 ; 76 ; 77)$ Calcium supplements have been shown to reduce the risk of recurrent polyps (78). The longterm use of aspirin has been shown to have preventive effects on CRC $(77 ; 79)$ but it is not prescribed routinely for this purpose because of its side effects which includes gastrointestinal bleeding (9). Although, HRT has shown protective effects against CRC, it can increase the
risk for breast and other cancers, and therefore is not presently used for CRC prevention (76;77;79;80).

### 1.5 Smoking and colorectal cancer

Smoking is a major contributing factor to human carcinogenesis and is one of the most important modifiable risk factors for cancer and premature death worldwide (24). The main hazards of smoking are related to exposures such as age at smoking initiation, numbers of cigarettes smoked per day, smoking inhalation or type of cigarettes such as either tar and nicotine, or content or filter type (81). Cigarette smoke contains more than 7000 chemical compounds majority of which are carcinogens such as polycyclic aromatic hydrocarbons (PAH) and nitrosamines in addition to other promoters. These mixtures contribute to complete carcinogenesis in the mucosa of the colon and rectum (82). The carcinogenic effects of smoking could be initiated through multiple pathways such as DNA binding and mutations, oxidative stress, epigenetic changes, or inflammation (14). Figure 10 shows the pathway for causation of cancer via the carcinogenic effects of smoking. In the most recent monograph published in 2012 (33), and the report from the Unites States Surgeon General (15), the conclusion was that there is a casual association between smoking and CRC. The association between smoking and CRC risk has been shown to be dose-related (83-85). A longer exposure to or duration of smoking (35-40 years) has been shown to be associated with increased risk of CRC $(86 ; 87)$. The association between smoking and colorectal adenomas which are precursor lesions for most CRC was confirmed in a recent meta-analysis (32).


Figure 10: Pathway for causation of cancer by carcinogens in tobacco smoke
(Reprinted from the United States Department of Health and Human Services (2004). The Health consequences of Smoking: A Report of the Surgeon General. Atlanta, GA: The United States Department of Health and Human Services, Center for Disease Control and Prevention, National Center for Chronic Disease)

## 2 Aims of the thesis

The main aim of this thesis was to examine the association between smoking and CRC incidence and mortality overall and by subsites and gender.

The specific objectives were:

1. To investigate the association between smoking and the risk of colon cancer overall, and by localization and gender.
2. To investigate the association between smoking and the risk of rectal cancer by gender.
3. To examine the association between smoking and CRC mortality overall, by subsites and gender.
4. To examine the association between different smoking exposures i.e., age at smoking initiation, numbers of cigarettes smoked per day, smoking duration and number of pack-years smoked and colon and rectal cancer by gender.
5. To examine the association between different smoking exposures i.e., age at smoking initiation, numbers of cigarettes smoked per day, smoking duration and number of pack-years smoked CRC mortality by gender.

## 3 Materials and Methods

### 3.1 Study population

The cohort included 652,792 Norwegians (49\% men) born between 1897 and 1975, recruited from several Norwegian health screening surveys initiated by the National Health Screening Service (now included in the Norwegian Institute of Public Health). These surveys were conducted between 1972 and 2003 and are as follows: the Oslo study I (1972-1973), the Norwegian counties study (1974-1988), the 40 years cohort (1985-1999) and the Cohort of Norway (CONOR, 1994-2003).

In all surveys included, information was gathered through questionnaires and a short health examination. The design and protocol of these surveys were very similar, but there were some modifications during different time periods, mainly in the questionnaires regarding questions on smoking, alcohol consumption, physical activity and other lifestyle factors. In most surveys, the attendees were given another supplementary questionnaire which they completed at home and mailed back in a pre-addressed stamped envelope. The participation rates for the different surveys varied from $56-88 \%$. A flow chart with a detailed description of study participants has been provided below (Figure 11).

## The Oslo study I

This survey was conducted in 1972-1973 among men living in the municipality of Oslo. Men aged 40-49 years in Oslo and a random sample of 7\% of the general male population aged 2039 years were invited to participate in screening for tuberculosis and cardiovascular disease. About 30,000 men were invited and almost 18,000 attended the screening (i.e., a participation rate of approximately 60\%). The participants answered one-page questionnaire which focused on symptoms of cardiovascular disease and diabetes, smoking habits and physical activity. This was one of the first large epidemiological studies of that period and became a model for
establishing other population based health studies in Norway later on. Height, weight and blood pressure were measured during screening using a standard procedure (88-91).

## The Norwegian counties study

These surveys included participants of cardiovascular disease screening in three Norwegian counties (Finnmark, Sogn og Fjordane and Oppland) during three different time periods: 1974-1978, 1977-1983 and 1985-1988. All residents aged 35-49 years as well as random sample of $10 \%$ of the general population aged 20-34 years were invited to a first screening. A second and third screening was carried out, and included a combination of previous cohort as well as new ones. Similar protocols and questionnaires were applied for these surveys. The attendance rates were $88 \%, 88 \%$ and $84 \%$ at the three screening rounds, respectively (91-93).

## The 40 years cohort

These surveys included about 420,000 Norwegian men and women, and were carried out in all of the 19 counties of Norway in 1985-1999 for cardiovascular disease screening. Men and women aged 40-42 years were the largest invited population. Individuals aged 65-67 years were also invited to the first round of surveys in some of the counties (Nord-Trøndelag, Møre and Romsdal and Hordaland). The participation rate was $69 \%(94 ; 95)$. Of all the surveys included in this thesis, the 40 years cohort had the largest number of participants.

## The Cohort of Norway

CONOR is a very large collaborative project including regional data from 10 epidemiological studies conducted in 1994-2003 which have been merged into a national database (please refer to Table 1 for details of surveys included in CONOR). Standardized protocols, procedures and questionnaires were used together with a short health examination. The questions used in CONOR have been validated previously. The response rate varies across the surveys. The average response rate for the 10 different surveys in the CONOR study was

56\%. Altogether, around 309,000 individuals were invited of which about 181,000 accepted to participate and provided consent (91;96;97).

Table 1: List of different surveys included in the study

| Name of Survey | Year <br> Conducted | Populations from | Surveys |
| :--- | :--- | :--- | :--- |
| The Oslo study I | 1972 | Oslo (only men) | 1 |
| The Norwegian counties <br> study | $1974-88$ | Oppland, Sogn og Fjordane, Finnmark | 9 |
| 40 years cohort | $1985-99$ | All Norwegian counties included | 19 |
| CONOR | $1994-95$ | Tromsø | 1 |
| Tromsø Health Study IV | $1995-1997$ | Nord-Trøndelag | 1 |
| The second Nord-Trøndelag <br> Health study (HUNT 2) | $1997-99$ | Hordaland | 1 |
| Hordaland Health <br> Study(HUSK) | 2000 | Oslo | 1 |
| Oslo study II | $2000-2001$ | Oslo | 1 |
| HUBRO( The Oslo Health <br> Study) | Oppland and Hedmark | 1 |  |
| Oppland and Hedmark Health <br> Study (OPPHED) | $2000-1$ | Tromsø | 1 |
| Tromsø Health Study V | 2001 | Oslo | 1 |
| I-HUBRO(The Oslo <br> Immigrant Health Study) | 2002 | 2003 | 1 <br> Troms and Finnmark Health <br> Study (TROFINN) <br> MoRo II(The second part of <br> the Romsås in Motion Study <br> Total |
|  | 2002 |  |  |



Figure 11: Detailed flowchart of participants from the different surveys

### 3.1 Pooling Datasets

After obtaining specified variables from each survey's primary data using the unique key identifier for each participant, we created a standardized data base for the pooled analyses. There were total 833,871 registered observations including 181,079 doubles or more. For participants who took part in more than one survey, only the earliest survey was included. Variables common to all surveys were transformed to the same format. The variables in the CONOR study were adequately structured and this was taken as a reference for standardizing the questionnaires. All surveys had a baseline questionnaire, which included detailed assessments of smoking habits, physical activity, and other lifestyle factors. At the screening facility height and weight were measured in a standardized way by a trained person, which allowed us to calculate body mass index (BMI, $\mathrm{kg} / \mathrm{m}^{2}$ ). Question on smoking habits were similar but not identical across all surveys. The questions asked about current and former daily smoking habits, smoking duration, average number of cigarettes smoked per day and in few surveys former smokers were asked about time since cigarette quitting. Only the CONOR study asked about age at smoking initiation. In the other surveys, this variable was estimated for both current (age at enrolment minus duration of smoking in years) and former (age at enrolment minus years since quitting and duration of smoking) smokers. We also found common formats for other variables such as menopause, menarche, HRT and alcohol consumption which were available only in the latest surveys such as 40 years III and IV and CONOR. Due to large missing in these variables which reached more than $50 \%$, we were not able to use them in our main analysis. Detailed information on how the files were merged into single database is included in the appendix section (Appendix 3).

### 3.3 Exposure information

Participants who smoked daily were categorized as current smokers, and those who answered that they had smoked previously but not currently or if they answered the year since quitting smoking were categorized as former smokers. Current and former smokers were then combined into a single category called ever smokers. In Paper I, we further categorized ever smokers according to: age at smoking initiation ( $\leq 16,17-19,20-24, \geq 25$ ), numbers of cigarettes smoked per day (1-9, 10-19, $\geq 20$ ), smoking duration in years (1-19, 20-29, 30-39, $\geq 40$ ) and number of pack-years smoked (i.e., number of cigarettes smoked per day, divided by 20 , multiplied by the duration of smoking in years) ( $0-9,10-19, \geq 20$ ). In Paper II and III, we categorized ever smokers by different measures of smoking exposure: age at smoking initiation ( $\leq 19,20-24, \geq 25$ ), numbers of cigarettes smoked per day ( $1-9,10-19, \geq 20$ ), smoking duration in years $(1-19,20-29, \geq 30)$ and number of pack-years smoked ( $0-9,10-19, \geq 20$ ). In all three papers, participants who were neither current nor former smokers were classified as never smokers. Participants were categorized into three groups based on their level of physical activity at enrolment: sedentary (reading, watching television, sedentary activity, or walking, bicycling <4 hours per week), moderate (walking, bicycling, and/or similar activities $\geq 4$ hours per week), and heavy (light sports or heavy gardening $\geq 4$ hours per week, heavy exercise or daily competitive sports). The most recent information regarding duration of education was obtained from Statistics Norway and was used to assign subjects to one of three categories of duration of education ( $<10,10-12, \geq 13$ years).

### 3.4 Follow-up and endpoints

The study population comprised individuals who participated in of one of the four health surveys included in our thesis. We excluded participants who had emigrated or died before the start of follow-up $n=1,009$ ( $50 \%$ women) and those with prevalent cancer $n=11,476(62 \%$ women). We also excluded participants with missing information on either smoking exposure $n=6,299$ ( $45 \%$ women) or on any of the co-variates [BMI, physical activity, education $n=$ 31,766 (50\% women)]. Altogether 50,550 (48\% women) participants were excluded leaving 602,242 subjects ( $51 \%$ women) in the analytical cohort for all papers.

We followed all participants aged 19-67 years at enrolment through a linkage to the Cancer Registry of Norway and the Central Population Register, utilizing the unique 11-digit personal identification number to identify all cancer cases, emigrations and deaths. The participants were linked to the Cancer Registry of Norway, the Norwegian Cause of Death Registry and the Central Population Register. The national registries have accurate and detail information regarding cancer incidence and mortality (98). The national registries are both accurate and virtually complete ( $98 ; 99$ ). The start of follow-up was set at 1 January of the year after the baseline questionnaire was completed. In Paper I, person-years were calculated from the start of follow-up to the date of colon cancer diagnosis, the date of any incident cancer diagnosis (except skin basal cell carcinoma), emigration, death, or the end of follow-up, i.e., December 31, 2007, whichever occurred first. In Paper II, person-years was calculated from the start of follow-up to the date of rectal cancer diagnosis, the date of any incident cancer diagnosis (except skin basal cell carcinoma), emigration, death, or the end of follow-up, i.e. December 31, 2007, whichever occurred first. In paper III, follow-up ended at the time of death from primary CRC cancer, death from any other cancer (except basal cell carcinoma of the skin), emigration, death from other causes, or the end of follow up (December 31, 2007), whichever occurred first.

Colon and rectal cancer were classified according to the Seventh Revision of the International Statistical Classification of Diseases (ICD-7) (codes 153 and 154 respectively), and colon cancer was further categorized according to tumor location, i.e., proximal (codes 153.0/153.1) and distal (codes 153.2/153.3). Tumors that were overlapping (code 153.4), were specified as appendix (code 153.6), or were unspecified (code 153.9) were classified as "others" and were included in the analyses for the whole colon only. CRC mortality was classified according to ICD-9 and ICD-10.

### 3.5 Statistical analyses

We performed all analyses separately by gender. We used the t-test and $\chi^{2}$ test for investigating differences in the distribution of selected characteristics between cases and noncases and between ever and never smokers. The Cox proportional hazards model was used with age as the underlying time scale to estimate multivariate-adjusted hazard ratios (HR) with $95 \%$ confidence intervals (CIs) for the associations between different measures of smoking exposure age at smoking initiation, numbers of cigarettes smoked per day, smoking duration in years and number of pack-years smoked and colon cancer overall, and according to tumor location (Paper I), rectal cancer (Paper II) and CRC mortality (Paper III) with never smokers as the reference group. In Paper I, entry time was defined as age at enrolment and exit time was age at diagnosis of colon cancer, the date of any incident cancer diagnosis (except basal cell carcinoma), emigration, death, or the end of follow-up (31 December, 2007), whichever occurred first.

In Paper II, entry time was defined as age at enrolment and exit time was age at diagnosis of rectal cancer, the date of any incident cancer diagnosis (except basal cell carcinoma), emigration, death, or the end of follow-up (31 December, 2007), whichever occurred first.

In Paper III, entry time was defined as age at enrolment and exit time was age at death, emigration, or end of follow-up (31 December, 2007), whichever occurred first.

The possible confounders included in the final models in Paper I, II and III, selected a priori, were age at enrolment (continuous), level of physical activity (sedentary, moderate and heavy), BMI (continuous), all at enrolment and duration of education in years ( $<10,10-12$, $\geq 13$ ). Tests for linear trends were obtained by creating an ordinal exposure variable with equally spaced scores and including it in the models with never smokers as the reference group. Test of heterogeneity by gender and its effect on the association between smoking and the risk of colon cancer overall, and by location, rectal cancer and CRC mortality were tested using Wald $\chi^{2}$ statistics in Paper I, II and III, respectively. Two-sided p-values $<0.05$ were considered statistically significant. All analyses were conducted using STATA version 12.0 (Stata Corp., College Station, TX, USA).

In all the papers, the same methods of statistical analysis were used; only the outcome variable differed. Outcome for Paper I was colon cancer, Paper II was rectal cancer and Paper III was CRC mortality.

In all the papers, we re-analyzed the data excluding the 8,151 ( $99 \%$ men) participants who reported smoking only cigars or pipes. We had information on alcohol consumption for 37\% ( $\mathrm{n}=221,748$ ) of the participants. We did sensitivity analyses by gender for the main outcomes based on this population ( $49 \%$ men) with and without adjustment for alcohol consumption in all papers.

### 3.6 Ethical aspects

Oral or written informed consent was obtained from participants in the different surveys. Surveys carried out in 1995 and after had written consent. We also obtained approval from the respective steering committees to all the health surveys included. We obtained approvals from
the National Data Inspection Board, the Regional Committee for Medical Research Ethics (REK), the Norwegian Directorate of Health, Norwegian Tax Administration and Norwegian Public Health Institute. The data was handled in accordance with the permissions taken from the above mentioned governmental bodies.

## 4 Results - summary of papers

### 4.1 Paper I

The increased risk of colon cancer due to cigarette smoking may be greater in women than men.

In Paper I, we investigated the association between smoking and colon cancer overall, by location and gender. The study followed 602,242 Norwegian men and women and 3,998 colon cancer cases ( $46 \%$ of cases in women). Women ever smokers had a $19 \%$ (HR = 1.19, $95 \% \mathrm{CI}=1.09-1.32)$ and men ever smokers had $8 \%(\mathrm{HR}=1.08,95 \% \mathrm{CI}=0.97-1.19)$ increased risk of colon cancer compared with gender specific never smokers. For all four dose-response variables examined, women ever smokers in the most exposed category of age at smoking initiation, $(\mathrm{HR}=1.48,95 \% \mathrm{CI}=1.21-1.81)$, number of cigarettes smoked per day $(\mathrm{HR}=1.28,95 \% \mathrm{CI}=1.06-1.55)$, smoking duration $(\mathrm{HR}=1.47,95 \% \mathrm{CI}=1.11-1.95)$, and pack-years smoked $(\mathrm{HR}=1.33,95 \% \mathrm{CI}=1.11-1.57)$ had a significantly increased risk of more than $20 \%$ for colon cancer overall and of more than $40 \%$ for proximal colon cancer compared with never smokers. Women ever smokers had a higher risk of proximal colon cancer compared to men ever smokers (Wald $\chi^{2}, \mathrm{p}=0.02$ ).

Sensitivity analyses were carried out for participants with information on alcohol consumption which mainly included participants enrolled after 1995 (37\% of total analytical cohort, $\mathrm{n}=221,748$ ). The corresponding risk estimates for women ever smokers were $16 \%$ $(\mathrm{HR}=1.16,95 \% \mathrm{CI}=0.86-1.74), 27 \%(\mathrm{HR}=1.27 \%, 95 \% \mathrm{CI}=0.82-1.51)$ and $11 \%(\mathrm{HR}=$ 1.11, $95 \% \mathrm{CI}=0.78-1.59$ ) for colon, proximal colon and distal colon cancer, respectively. However, among men ever smokers risk estimates were (HR = 0.99, 95\% CI = 0.78-1.25), $(\mathrm{HR}=0.97,95 \% \mathrm{CI}=0.75-1.64),(\mathrm{HR}=0.82,95 \% \mathrm{CI}=0.68-1.15)$ for colon, proximal colon
and distal colon cancer, respectively. Risk estimates with and without alcohol adjustment did not differ significantly.

The conclusion was that women smokers may be more susceptible to colon cancer and especially to proximal colon cancer than men smokers.

### 4.2 Paper II

Smoking increases rectal cancer risk to the same extent in women as in men: Results from a Norwegian cohort study.

In Paper II, we examined the association between smoking and rectal cancer incidence by gender among 602,242 Norwegian men and women. During a mean follow-up of 14 years, 2,176 cases ( $61 \%$ cases in men) were diagnosed with invasive rectal cancer. Both men and women ever smokers had a significantly increased risk of rectal cancer of more than $25 \%$ for men $(\mathrm{HR}=1.27,95 \% \mathrm{CI}=1.11-1.45)$ and women $(\mathrm{HR}=1.28,95 \% \mathrm{CI}=1.11-1.48)$ compared with gender specific never smokers. Men smoking $\geq 20$ pack-years had an increased risk of rectal cancer of $35 \%(H R=1.35,95 \% C I=1.14-1.58)$, whereas women showed an increased risk of $47 \%(\mathrm{HR}=1.47,95 \% \mathrm{CI}=1.13-1.91)$ compared with gender specific never smokers. For both men and women, we observed significant dose-response associations with rectal cancer risk when looking at age at smoking initiation, number of cigarettes smoked per day, smoking duration and number of pack-years smoked and using never smokers as the reference group (p-trend<0.05). The test for heterogeneity by gender was not significant between smoking status and the risk of rectal cancer (Wald $\chi^{2}, \mathrm{p}$ value; current smokers $=$ 0.85 ; former smokers $=0.87$ and ever smokers $=1.00$ ).

In the sensitivity analyses for participants, mainly enrolled after 1995, with information on alcohol consumption, the risk estimate of rectal cancer incidence was $13 \% ~(H R=1.13,95 \%$
$\mathrm{CI}=0.83-1.55)$ with alcohol adjustment and $12 \%(\mathrm{HR}=1.12,95 \% \mathrm{CI}=0.82-1.54)$ without alcohol adjustment among men ever compared with men never smokers. The risk estimate was $37 \%(\mathrm{HR}=1.37,95 \% \mathrm{CI}=0.99-1.92)$ with alcohol adjustment and $39 \%(\mathrm{HR}=1.39$, $95 \% \mathrm{CI}=1.00-1.94)$ without alcohol adjustment among women ever compared with women never smokers.

In conclusion, increased risk of rectal cancer due to smoking is similar in women as in men.

### 4.3 Paper III

Cigarette smoking and colorectal cancer mortality among 602,242 Norwegian men and women.

In Paper III, we examined the association between different measures of smoking exposure and CRC mortality overall and by subsites among 602,242 Norwegian men and women and 2,333 CRC deaths ( $60 \%$ in men). There were 1,607 ( $57 \%$ in men) colon cancer and 726 ( $67 \%$ in men) rectal cancer deaths. Women ever smokers had a $22 \%$ (HR $=1.22,95 \% \mathrm{CI}=1.06-$ 1.40 ) increased risk CRC mortality compared with women never smokers. Men ever smokers had a CRC mortality risk of $23 \%(\mathrm{HR}=1.23,95 \% \mathrm{CI}=1.08-1.40)$ when compared with men never smokers. Women ever smokers had an almost $50 \%$ ( $\mathrm{HR}=1.49,95 \% \mathrm{CI}=1.20-1.87$ ) increased risk of mortality from proximal colon cancer compared with women never smokers. A test for heterogeneity by gender showed an increased risk of mortality from proximal colon cancer among women, which was statistically significant for ever smokers and former smokers (Wald $\chi^{2}=0.02$ and 0.04 , respectively). It was also significant for former smokers and the risk of rectal cancer showing increased risk among men (Wald $\chi^{2}=0.02$ ).

In the sensitivity analyses among participants with information on alcohol consumption (37\% of total analytical cohort), the risk estimates of CRC mortality was $(\mathrm{HR}=0.84,95 \% \mathrm{CI}=$
$0.60-1.18)$ and ( $\mathrm{HR}=1.25,95 \% \mathrm{CI}=0.89-1.74$ ) among men and women ever smokers respectively. Risk estimates with and without alcohol adjustment did not differ significantly. In conclusion, smoking is associated with increased CRC mortality both among men and women. The risk of rectal and proximal colon cancer mortality was more pronounced among men and women smokers, respectively.

## 5 Discussion

### 5.1 Methodological issues

A detailed discussion of the findings is presented separately in each paper. In the following chapter, discussions of those aspects which are applicable to this thesis in general are presented. Epidemiological studies primarily provide important information regarding the general population. The main purpose of such studies is to generalize the results to another target population and to establish the association between a risk factor and an outcome. In this regard, validity of the study is a very important issue. The validity of an epidemiological study can be divided into two groups: internal validity and external validity.

### 5.1.1 Internal Validity

Internal validity is defined as the true measure of the variable obtained for the study subjects and refers to the logical conclusions drawn from them. It deals mostly with the accuracy of observed results of the study. Internal validity is evaluated by determining whether the observed changes or outcomes can be attributed to the main exposure and not to other causes. Several factors can influence the validity of observed association between an exposure and an outcome (100;101). A major threat to internal validity could be lack of representativeness of the study population. The two major errors that can occur in epidemiological studies are random and systematic errors. Internal validity depends both on random error as well as systematic errors such as bias and confounding (100;101). Figure 12 shows the diagrammatic view of error and its classification which are often encountered in a large epidemiological study.


Figure 12: A systematic approach to bias
(Source: Appraising the evidence: what is selection bias? Henderson $M$ et al: Reprinted with permission)

Random error can arises due to sampling variability and can be addressed by appropriate statistical hypothesis testing. Random error may lead to non-reproducibility of study results which in turn could weaken or restrict the association between an exposure and an outcome (100). A large sample size gives more precision to a study. In our study, the large sample size minimized the sampling error and thus increased the precision (100). We have also addressed the issue of random error by applying the appropriate statistical procedures. Our hypothesis was tested at the $5 \%$ alpha level and $95 \%$ confidence intervals were calculated. The null hypothesis was rejected at a less than 5\% level. Another error encountered in epidemiological studies is systematic error. Epidemiological studies with a minimal systematic error have a high accuracy. These errors are independent of the size of the study and statistical significance
does not suggest the absence of any bias $(102 ; 103)$. We consider the discussion of selection and measurement bias relevant in relation to our study.

## Selection bias (Paper I-III)

Selection bias in cohort studies results from the process of selecting study participants and can arise due to systematic differences in selection criteria (100). The possibility of this bias arises when a study sample is not representative of the source population (104). However, it is also true that selection bias is less probable in cohort studies than other epidemiological studies as the outcome is not known at the time of enrolment (105). In our study, we had no possibility to control for differences between responders and non-responders as there was no information available for the non-responders.

In all of the surveys included in our study, age was a major criterion for enrolling participants. Most of the men and women enrolled were between 40-45 years of age and a large group of participants were included from the 40 years cohort. The detail description of the study participants categorized by age group during the time of enrolment in different surveys is shown in table 2. The overall participation rate ranged from 56-88\%. The attendance rate in CONOR was $56 \%$ (range $30-76 \%$ ) whereas in the Oslo study I, it was approximately $60 \%$. The participation rate for the Norwegian counties study remained between $78-90 \%$. In 40 years cohort, the overall response rate was $69 \%$ but during 1994-99, the participation rate went down to $62 \%$.

Table 2: Age at enrolment of participants included in different health surveys

| Age at <br> enrolment | Oslo study <br> I | Norwegian <br> counties study | 40 years <br> cohort | CONOR <br> study | Total (\%) |
| :--- | :--- | :--- | :--- | :--- | :--- |
| $16-30$ | 869 | 9,778 | 740 | 9,492 | $20,879(3.5)$ |
| $31-39$ | 689 | 20,216 | 652 | 23,873 | $45,430(7.5)$ |
| $40-44$ | 4,782 | 29,282 | 364,285 | 25,583 | $423,932(70.4)$ |
| $45-50$ | 9,506 | 23,458 | 5,281 | 16,675 | $54,920(9.1)$ |
| $\geq 50$ | 1,100 | 752 | 13,809 | 41,420 | $57,081(9.5)$ |
| Total | $\mathbf{1 6 , 9 4 6 ( 3 )}$ | $\mathbf{8 3 , 4 8 6 ( 1 4 )}$ | $\mathbf{3 8 4 , 7 6 7 ( 6 4 )}$ | $\mathbf{1 1 7 , 0 4 3 ( 1 9 )}$ | $\mathbf{6 0 2 , 2 4 2}$ |

Non-response bias is always a major issue in large longitudinal epidemiological studies like ours and declining participation rate is one of the major problems. However, low participation rates do not always indicate a high level of bias. Indeed, there has been very little evidence of substantial bias as a result of non-response and non-response introduces less influence on exposure-disease associations (106-108). Furthermore, we had a similar proportion of men and women participants in our study. A total of 50,550 participants excluded, $48 \%$ of which were women due to missing covariates. Thus, our study had a same proportion of men and women excluded due to the missing data. Those excluded group were similar to the analytical cohort in regards to their level of education and physical activity. Incidence rates for colon and rectal cancer among excluded group were also similar to the analytical cohort. Furthermore, smoking prevalence among participants from different health surveys in our cohort was comparable to the Norwegian general population during the same period (Fig 13 and 14).


Figure 13: The prevalence of current smokers included in surveys by gender


Figure 14: The prevalence of current smokers aged 16-74 years from 1973-2003 in Norway by gender

## Information bias (Paper I-III)

Information bias is also known as observation, classification or measurement bias and arises from incorrect determination of an exposure, an outcome, or both (109). Measurements bias occurs when exposures and outcome variables are incorrectly measured (100). In the different surveys included in our study, height and weight were measured according to the standard procedure to minimize the measurement errors. There were some differences in the measurement of exposures variable but we minimized these differences by finding a common format during the merging of the datasets. Smoking history was obtained at study enrolment, and so was not subject to recall bias. Furthermore, smoking habits change; current smokers could have stopped smoking whereas never smokers may have started smoking. Our analysis was based on ever and never smokers, thus only the status of never smokers could have changed during follow-up. In addition to this, very few Norwegians start to smoke after the age of 30 , and the mean age at enrolment for our study is more than 40 years, thus minimizing this type of bias. We assume that the possibility of information bias in our study is limited.

## Confounding and statistical analyses (Paper I-III)

Confounder is defined as a variable which is associated with main exposure variable but at the same time an independent risk factor for the dependent variable $(100 ; 101)$. As a confounding variable is associated with the exposure and also with outcome but does not stand in the intermediate pathway in the chain of causation between an exposure and an outcome (109), it leads to the mixing or blurring of effects. This is one of the major challenges of an observational study as it can either attenuate or inflate an association between an exposure and an outcome. In a way, confounder is similar to bias but it can be controlled by stratification and adjustment in multivariate models. The magnitude of confounding can be evaluated by comparing crude and adjusted effect measure. Age and gender are almost always potential confounders $(100 ; 101)$. Our analyses were stratified by gender and hazard ratios
(HRs) and 95\% CI were estimated by fitting Cox proportional hazard models where age was the primary time variable. In Papers I, II and III, age, BMI, physical activity at enrolment and duration of education were the confounders based on a priori, and were controlled for when estimating the association between smoking and colon and rectal cancer incidence and CRC mortality. The other important covariates that are established risk factors for CRC, such as alcohol consumption, HRT, diet such as red meat and COX inhibitors such as aspirin could not be adjusted for in the main analyses. Information on alcohol consumption was missing on more than $60 \%$ of the total participants whereas information on HRT was missing in more than $70 \%$ of total women. It has been reported that women could have protective hormonal effects until menopause from HRT which delay or protect them from development of CRC (76). The use of HRT declined after there was growing evidence that it could be risk factor for breast cancer and other cardiovascular disease (110). Similarly, we lacked information on molecular data and CRC screening, as it was not common in Norway when the surveys included in our study were conducted. In addition to this, the information on staging of CRC was also not available. Cigar and pipe smoking may have less potential to be confounders and this could be the reason our sensitivity analyses excluding those smoking only cigar and pipe did not materially change the estimates (33). We also performed the sensitivity analyses among participants who had information on alcohol consumption, with and without alcohol adjustment. Only $37 \%$ of the total cohort ( $48 \%$ men) had information on alcohol consumption. Our sensitivity analyses including only those with information on alcohol consumption, risk estimates increased among women and but decreased among men ever smokers for rectal cancer incidence as well as for CRC mortality compared to risk estimates for the main cohort. For colon cancer, the estimates were more or less similar for women but decreased among men compared to risk estimates for the main cohort. However, the results did not change materially with and without alcohol adjustment in this sub cohort either among
men or women indicating that the lack of alcohol intake in the main cohort might not be a major limitation. However, the interpretation of our sensitivity analyses should be done with caution as they included fewer cases, younger participants with less follow-up time than in the main cohort. We should be very cautious to interpret the results of our sensitivity analyses as we lost a large number of cases and follow up time period (>75\%). The studies such as Oslo study I, the Norwegian counties study and earlier rounds of 40 years cohort did not have the information on alcohol consumption. It is also true that the alcohol consumption is higher among men than women in Norway (111). Thus, the lack of adjustments of alcohol consumption in our main cohort analyses is likely to have inflated the estimates among men more than women and thus attenuated the gender difference.

The statistical approach to use Cox proportional hazards analysis with age as primary time variable to examine the association between smoking and CRC incidence and mortality was considered appropriate to answer the research questions in Papers I, II and III. Modelling the events using a proportional hazards model with age as the time scale has been recommended as an appropriate method in large health surveys with disease or death as outcome. Furthermore, it has been suggested that using age as a primary time variable is more meaningful and less restrictive than using time on study as the time scale (112).

### 5.1.2 External validity

External validity is the probability of generalizing the study results to a wider population. This can be also referred as the possibility, or the degree to which the results of the study is applicable to different population in other places and at different time periods (100;101;113). Internal validity is always a pre-requisite for external validity. Although, we had some issues with internal validity, we are convinced that it did not distort our results. Our study includes
very large participants from all over Norway. The separate health surveys included in our study have well-validated datasets. In general, it is difficult to generalize the study results to a wider population but we assume our study conclusion could be generalized to the Caucasian and Western population.

### 5.2 Discussion of the main results

The main findings are discussed in the respective papers (Papers I-III) in detail. Despite some methodological limitations in the three papers, they have contributed to further support the fact that smoking increases CRC incidence and mortality among both men and women. The discussion below is focused on the main messages of the three papers regarding the association between smoking and CRC.

### 5.2.1 Gender differences in smoking related colon cancer

The findings from Paper I is in agreement with IARC and United States Surgeon General's recent conclusion that cigarette smoking is associated with colon cancer (15;33). Incidence rates are more important and reliable indicator of trends in disease occurrence than mortality rates as incidence is not influenced by changes in treatment and survival (6). The main difference in CRC in general observed by gender is due to the higher incidence of rectal cancer in men than women. There is not much difference in incidence rates of colon cancer between men and women in Norway.

There are gender reported differences in incidence of colon cancer by location (i.e. proximal vs. distal colon cancer). Some studies have concluded in general that the risk of distal colon cancer is lower among women than in men (114-116). Previous knowledge regarding smoking and colon cancer incidence in general varies by gender. Some studies reported that
the association between smoking and colon cancer may be stronger in men as compared to women ( $75 ; 117 ; 118$ ). However, these reports could be attributed to the low prevalence of ever smoking women. On the other hand, the results of the studies among women only (119-122) reported findings which were more or less comparable to men for both colon as well as rectal cancer. A recent study from Europe which included men and women from ten European countries reported the risk estimates by subsites and indicated that the ever smokers have an increased risk of colon cancer, which was especially pronounced in the proximal than in the distal colon (123). However, this study did not report the risk estimates by gender. Another study of Norwegian women reported an increased risk of proximal than distal colon cancer among women ever smokers (119). A study among postmenopausal women in the United States aged 55-69 years at baseline also reported an increased risk of proximal than distal colon cancer (120). Furthermore, smoking has been shown to be associated with a higher incidence proximal colon cancer among Caucasian women in the United States as compared with distal colon cancer (124). A study from Japan which was conducted both among men and women and included around 400 colon cancer, reported the risk estimates by gender and the findings were insignificant increase risk of colon cancer among both men and women ever smokers (125). Increased risk of proximal colon cancer among women smokers has been reported to be related with epigenetic changes which are induced by tobacco related carcinogens (120). It has also been suggested that gender-related differences in hormonal factors (126) or susceptibility to tobacco related carcinogens (127) could have influenced the observed different associations for proximal and distal colon cancer by gender (120) which might explain the reason for increased risk of proximal colon cancer among women smokers compared to men smokers. There are not many prospective cohort studies examining the association between smoking and colon cancer by location and gender in detail. Our study is among the very few studies with a very large numbers of incidence cases as well as a large
proportions of ever and never smokers that examined the association between smoking and colon cancer incidence by location and gender. The findings from our study suggested that women smokers maybe more prone to colon cancer especially for proximal colon cancer than men smokers. Our findings could be a strong warning for the women smokers who could be more vulnerable to smoking related colon cancer than men. This may have important clinical and research implications if further confirmed by other large population based epidemiological studies.

### 5.2.2 Smoking related risk of rectal cancer among women is same as in men

The epidemiologic evidence supports that it takes decades before the increased risk of rectal cancer appears and that smoking plays an important role in early carcinogenesis both among men and women $(15 ; 86 ; 87)$. The incidence rate of rectal cancer is higher among Norwegian men compared to Norwegian women and as mentioned earlier this is the main reason for gender difference in CRC incidence rate in general. The difference in rectal cancer incidence rate was almost 1.5 fold higher among Norwegian men in the beginning of our study period and the situation remained similar until the end of our study period. In the latest report from Norwegian Cancer Registry, this difference is also valid for the present time period (10). Risk patterns were shown to be generally consistent for colon and rectal cancer (73;75). However, some studies reported a stronger dose response association between smoking and rectal rather than colon cancer ( $8 ; 118 ; 121 ; 122 ; 128)$. Recent meta-analyses also concluded that the ever smokers are in increased risk of rectal cancer ( $70 ; 83-85$ ), however these studies did not present the risk estimates by gender. Our findings are in accordance with findings of these meta-analyses regarding higher risk estimates for rectal than colon cancer. In a study done among women in the United States, an increased risk of rectal cancer but not colon cancer was observed among ever smokers (121). Another study done among Norwegian women
reported the higher risk for colon than rectal cancer among smokers (119). Furthermore, two recent studies, one from 10 European countries including almost half a million men and women and 950 rectal cancer cases (123) and another from Asia including 329 rectal cancer cases (64) are the largest cohort study done before ours examining the association between smoking and rectal cancer. The study from 10 European countries found a non-significant increase in rectal cancer; however the later study found a significant increased risk of rectal cancer among ever smokers. These studies did not report the risk estimates by gender. A few studies from Japan examined the association between smoking and rectal cancer, however they included 200 or less cases (73;74). Furthermore, these studies showed an insignificant increased risk of rectal cancer among men and women ever smokers. Our study is one of the few to examine the association between smoking and rectal cancer by gender in detail. Our findings indicated that there is a significant increased risk for rectal cancer among men and women ever smokers. Furthermore, the findings also concluded that the risk was similar for women as in men. This could be a very important finding as the impact of cigarette smoking could be reflected in future rectal cancer incidence among women as the smoking epidemic among women began later than men, and as for colon cancer, rectal cancer also has a long latent period.

### 5.2.3 Smoking increases the risk of CRC Mortality

In Paper III, we found increased risk of CRC mortality both among men and women ever smokers. We concluded that the risk of rectal cancer mortality was higher among men smokers and risk of proximal colon cancer mortality was higher among women smokers. Similarly, the increased mortality risk by subsites was slightly more pronounced among current smokers compared with the former smokers both among men and women. The higher risk of rectal cancer mortality among men ever smokers and increased proximal colon cancer mortality risk among women ever smokers could be a mere reflection of the colon and rectal
cancer incidence in our cohort. As mentioned earlier, smoking is one of the major preventable causes of death worldwide. Mortality from different diseases has been decreased in last decades due to early diagnosis and treatment; however current smokers have an increased risk of mortality compared to never smokers. Recently, two meta-analyses also reported that the risk of CRC mortality was higher among current than former smokers (83;84). Long term smoking is associated with an increased risk of CRC mortality both among men and women (15). Furthermore, increased mortality among current smokers could be due to possible differences in health behaviours. A recent report from the United States Surgeon General concluded that there is a sufficient evidence to infer a causal relationship between cigarette smoking and increased all-cause and cancer-specific mortality (15). Quitting smoking can decrease the mortality burden and CRC patients should be encouraged to quit smoking as smoking can lead to poorer response to cancer treatment (129). Furthermore, the relationship between smoking and mortality is stronger than before and recommendations encouraging smokers to quit is very important.

## 6 Conclusions

The main aim of this thesis was to examine the association between smoking and CRC incidence and mortality overall and by subsites and gender.

The conclusions to be drawn from the studies are:

1. Smoking increased the risk of colon cancer among both men and women. The increased risk of colon cancer especially proximal colon cancer due, to smoking may be greater in women than men.
2. Smoking increased the risk of rectal cancer among both men and women. The risk was similar for women as for men.
3. Smoking increased the risk of CRC mortality among both men and women. The risk of rectal and proximal colon cancer mortality was most pronounced among men and women ever smokers, respectively.
4. The observed smoking related increased risk in colon and rectal cancer was dependent on different smoking exposures such as age at smoking initiation, number of cigarettes smoked per day, duration of smoking and pack years smoked both among men and women.
5. The observed smoking related increased risk in CRC mortality was dependent on different smoking exposures such as age at smoking initiation, number of cigarettes smoked per day, duration of smoking and pack years smoked both among men and women.

## 7 Implications for public health practice and further perspectives

CRC is one of the major public health problems in Norway. Our findings are consistent with the latest report from the IARC (1) and the United States Surgeon General (15) regarding the association between smoking and CRC. Smoking is possibly the most important modifiable risk factor of CRC. Detailed knowledge about the adverse harmful effects of smoking is important for general public health and future strategy planning. Additional strict rules against tobacco companies and tobacco sales should be implemented. The general population should be made aware of the possible harmful effects of smoking on the risk of CRC and younger age groups should be given special attention regarding smoking cessation and encouraged not to start smoking. Since women may be more vulnerable to the carcinogenic effects of smoking in relation to CRC, women-oriented awareness of harmful effects of smoking should be initiated. Current smokers should be encouraged to quit since the comorbid situation is increased among current smokers. More emphasis should be placed on taxes and price policies in the control of tobacco use to improve public health. Furthermore, CRC screening programme could be very helpful for early diagnosis and treatment.

As there is a long latent period between smoking and risk of CRC, an investigation with a longer follow up period could reveal more exact risk estimates. Future studies should focus on the replication of the present findings and it will be very important to conduct these studies with detailed information on most available covariates in relation to smoking and CRC.

## 8 Erratum

In Paper I:

For the excluded men and women, the overall incidence of colon cancer was 53 and 59 per 100, 000 person-years, respectively.

The overall incidence of colon cancer among men and women was 49 and 44 per 100, 000 person-years, respectively.

Above presented overall incidence rates were for CRC and not only for colon cancer.

## 9 References

(1) Ferlay J, Soerjomataram I, Ervik M, Dikshit R, Eser S, Mathers C, et al. Globocan 2012 v1.0. Cancer Incidence and Mortality Worldwide: IARC CancerBase No.11[internt]. 2013. Lyon, France, International Agency for Reasearch on Cancer. IARC Cancer Base. Cited from: http://globocan.iarc.fr on 10 March 2014.
(2) Ferlay J, Steliarova-Foucher E, Lortet-Tieulent J, Rosso S, Coebergh JW, Comber H, et al. Cancer incidence and mortality patterns in Europe: estimates for 40 countries in 2012. Eur J Cancer 2013 Apr;49(6):1374-403.
(3) Watson AJ, Collins PD. Colon cancer: a civilization disorder. Dig Dis 2011;29(2):222-8.
(4) Janout V, Kollarova H. Epidemiology of colorectal cancer. Biomed Pap Med Fac Univ Palacky Olomouc Czech Repub 2001 Sep;145(1):5-10.
(5) Boyle P, Langman JS. ABC of colorectal cancer: Epidemiology. BMJ 2000 Sep 30;321(7264):805-8.
(6) Haggar FA, Boushey RP. Colorectal cancer epidemiology: incidence, mortality, survival, and risk factors. Clin Colon Rectal Surg 2009 Nov;22(4):191-7.
(7) Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. Int J Cancer 2010 Dec 15;127(12):2893-917.
(8) Zisman AL, Nickolov A, Brand RE, Gorchow A, Roy HK. Associations between the age at diagnosis and location of colorectal cancer and the use of alcohol and tobacco: implications for screening. Arch Intern Med 2006 Mar 27;166(6):629-34.
(9) American Cancer Society. Colorectal Cancer Facts and Figures 2011-2013. Atlanta: American Cancer Society; 2011.
(10) Cancer Registry of Norway. Cancer in Norway 2011 - Cancer incidence, mortality, survival and prevalence in Norway. Oslo: Cancer Registry of Norway; 2013.
(11) Ferlay J, Bray F, Pisani P, Parkin DM. GLOBOCAN 2002: Cancer Incidence, Mortality and Prevalence Worldwide. Lyon: IARC Press; 2004.
(12) Beaglehole R, Bonita R, Horton R, Adams C, Alleyne G, Asaria P, et al. Priority actions for the non-communicable disease crisis. Lancet 2011 Apr 23;377(9775):143847.
(13) World Health Organization. Tobacco. 2013. Geneva, The World Health Organization. 6-6-2013. Cited from: http://www.who.int/mediacentre/factsheets/fs339/en/ on 11 February 2014.
(14) International Agency for Research on Cancer. World Cancer Report 2014. Lyon, France: International Agency for Research on Cancer; 2014.
(15) U.S Deparment of Health and Human Services. 2014 Surgeon General's Report: The Health Consequences of Smoking-50 Years of Progress. Rockville, MD, USA: U.S Department of Health and Human Services; 2014.
(16) Lopez AD, Collishaw NE, Piha T. A descriptive model of the cigarette epidemic in developed countries. Tobacco Control 1994;3:242-7.
(17) Thun M, Peto R, Boreham J, Lopez AD. Stages of the cigarette epidemic on entering its second century. Tob Control 2012 Mar;21(2):96-101.
(18) Lund M, Lindback R. Norwegian Tobacco Statistics 1973-2006. SIRUS- Writings 3/2007. 2007.
(19) Pukkala E, Martinsen JI, Lynge E, Gunnarsdottir HK, Sparen P, Tryggvadottir L, et al. Occupation and cancer - follow-up of 15 million people in five Nordic countries. Acta Oncol 2009;48(5):646-790.
(20) Helleve A, Weisæth A, Lindbak R. Tall om tabakk 1973-2009 (Figures about tobacco 1973-2009). Oslo, Norway: Norwegian directorate of Health; 2010.
(21) Norges offentlige utredninger.Tobakksindustriens erstatningsansvar [Tobacco industryliability]. Norway's public reports. Oslo, Norway: Statens forvaltningstjeneste,Informasjonsforvaltning.NOU; 2000.
(22) Statistics Norway 2014. Smoking habits,2013. 2014. Statistics Norway. Cited from: http://ssb.no/en/helse/statistikker/royk on 11 February, 2014.
(23) Ng M, Freeman MK, Fleming TD, Robinson M, Dwyer-Lindgren L, Thomson B, et al. Smoking prevalence and cigarette consumption in 187 countries, 1980-2012. JAMA 2014 Jan 8;311(2):183-92.
(24) Thun MJ, DeLancey JO, Center MM, Jemal A, Ward EM. The global burden of cancer: priorities for prevention. Carcinogenesis 2010 Jan;31(1):100-10.
(25) Brenner H, Kloor M, Pox CP. Colorectal cancer. Lancet 2013 Nov 8.
(26) O'Connell JB, Maggard MA, Liu JH, Etzioni DA, Livingston EH, Ko CY. Rates of colon and rectal cancers are increasing in young adults. Am Surg 2003 Oct;69(10):866-72.
(27) O'Connell JB, Maggard MA, Livingston EH, Yo CK. Colorectal cancer in the young. Am J Surg 2004 Mar;187(3):343-8.
(28) Murphy G, Devesa SS, Cross AJ, Inskip PD, McGlynn KA, Cook MB. Sex disparities in colorectal cancer incidence by anatomic subsite, race and age. Int J Cancer 2011 Apr 1;128(7):1668-75.
(29) Jemal A, Thun MJ, Ries LA, Howe HL, Weir HK, Center MM, et al. Annual report to the nation on the status of cancer, 1975-2005, featuring trends in lung cancer, tobacco use, and tobacco control. J Natl Cancer Inst 2008 Dec 3;100(23):1672-94.
(30) Bretthauer M. Evidence for colorectal cancer screening. Best Pract Res Clin Gastroenterol 2010 Aug;24(4):417-25.
(31) Levine JS, Ahnen DJ. Clinical practice. Adenomatous polyps of the colon. N Engl J Med 2006 Dec 14;355(24):2551-7.
(32) Botteri E, Iodice S, Raimondi S, Maisonneuve P, Lowenfels AB. Cigarette smoking and adenomatous polyps: a meta-analysis. Gastroenterology 2008 Feb;134(2):388-95.
(33) International Agency for Research on Cancer. Monographs on the evaluation of carcinogenic risks to humans. Personal habits and indoor combustions.A review of human carcinogens. IARC Press,Lyon, France; 2012.
(34) Munkholm P. Review article: the incidence and prevalence of colorectal cancer in inflammatory bowel disease. Aliment Pharmacol Ther 2003 Sep;18 Suppl 2:1-5.
(35) Triantafillidis JK, Nasioulas G, Kosmidis PA. Colorectal cancer and inflammatory bowel disease: epidemiology, risk factors, mechanisms of carcinogenesis and prevention strategies. Anticancer Res 2009 Jul; 29(7):2727-37.
(36) Fuchs CS, Giovannucci EL, Colditz GA, Hunter DJ, Speizer FE, Willett WC. A prospective study of family history and the risk of colorectal cancer. N Engl J Med 1994 Dec 22;331(25):1669-74.
(37) Johns LE, Houlston RS. A systematic review and meta-analysis of familial colorectal cancer risk. Am J Gastroenterol 2001 Oct;96(10):2992-3003.
(38) Imperiale TF, Ransohoff DF. Risk for colorectal cancer in persons with a family history of adenomatous polyps: a systematic review. Ann Intern Med 2012 May 15;156(10):703-9.
(39) Butterworth AS, Higgins JP, Pharoah P. Relative and absolute risk of colorectal cancer for individuals with a family history: a meta-analysis. Eur J Cancer 2006 Jan;42(2):216-27.
(40) Lynch HT, de la Chapelle A. Hereditary colorectal cancer. N Engl J Med 2003 Mar 6;348(10):919-32.
(41) Schatzkin A, Freedman LS, Dawsey SM, Lanza E. Interpreting precursor studies: what polyp trials tell us about large-bowel cancer. J Natl Cancer Inst 1994 Jul 20;86(14):1053-7.
(42) Jackson-Thompson J, Ahmed F, German RR, Lai SM, Friedman C. Descriptive epidemiology of colorectal cancer in the United States, 1998-2001. Cancer 2006 Sep 1;107(5 Suppl):1103-11.
(43) Papadopoulos N, Nicolaides NC, Wei YF, Ruben SM, Carter KC, Rosen CA, et al. Mutation of a mutL homolog in hereditary colon cancer. Science 1994 Mar 18;263(5153):1625-9.
(44) Wilmink AB. Overview of the epidemiology of colorectal cancer. Dis Colon Rectum 1997 Apr;40(4):483-93.
(45) Jasperson KW, Tuohy TM, Neklason DW, Burt RW. Hereditary and familial colon cancer. Gastroenterology 2010 Jun;138(6):2044-58.
(46) Davies RJ, Miller R, Coleman N. Colorectal cancer screening: prospects for molecular stool analysis. Nat Rev Cancer 2005 Mar;5(3):199-209.
(47) Wolin KY, Yan Y, Colditz GA, Lee IM. Physical activity and colon cancer prevention: a meta-analysis. Br J Cancer 2009 Feb 24;100(4):611-6.
(48) Boyle T, Keegel T, Bull F, Heyworth J, Fritschi L. Physical activity and risks of proximal and distal colon cancers: a systematic review and meta-analysis. J Natl Cancer Inst 2012 Oct 17;104(20):1548-61.
(49) Campbell PT, Cotterchio M, Dicks E, Parfrey P, Gallinger S, McLaughlin JR. Excess body weight and colorectal cancer risk in Canada: associations in subgroups of clinically defined familial risk of cancer. Cancer Epidemiol Biomarkers Prev 2007 Sep;16(9):1735-44.
(50) Chan AT, Giovannucci EL. Primary prevention of colorectal cancer. Gastroenterology 2010 Jun;138(6):2029-43.
(51) Calle EE, Kaaks R. Overweight, obesity and cancer: epidemiological evidence and proposed mechanisms. Nat Rev Cancer 2004 Aug;4(8):579-91.
(52) Calle EE, Teras LR, Thun MJ. Obesity and mortality. N Engl J Med 2005 Nov 17;353(20):2197-9.
(53) Ceschi M, Gutzwiller F, Moch H, Eichholzer M, Probst-Hensch NM. Epidemiology and pathophysiology of obesity as cause of cancer. Swiss Med Wkly 2007 Jan 27;137(3-4):50-6.
(54) Moghaddam AA, Woodward M, Huxley R. Obesity and risk of colorectal cancer: a meta-analysis of 31 studies with 70,000 events. Cancer Epidemiol Biomarkers Prev 2007 Dec;16(12):2533-47.
(55) Pischon T, Nothlings U, Boeing H. Obesity and cancer. Proc Nutr Soc 2008 May;67(2):128-45.
(56) Vrieling A, Kampman E. The role of body mass index, physical activity, and diet in colorectal cancer recurrence and survival: a review of the literature. Am J Clin Nutr 2010 Sep;92(3):471-90.
(57) Willett WC. Diet and cancer. Oncologist 2000;5(5):393-404.
(58) Larsson SC, Wolk A. Meat consumption and risk of colorectal cancer: a meta-analysis of prospective studies. Int J Cancer 2006 Dec 1;119(11):2657-64.
(59) Santarelli RL, Pierre F, Corpet DE. Processed meat and colorectal cancer: a review of epidemiologic and experimental evidence. Nutr Cancer 2008;60(2):131-44.
(60) Dixon LB, Balder HF, Virtanen MJ, Rashidkhani B, Mannisto S, Krogh V, et al. Dietary patterns associated with colon and rectal cancer: results from the Dietary Patterns and Cancer (DIETSCAN) Project. Am J Clin Nutr 2004 Oct;80(4):1003-11.
(61) Slattery ML, Boucher KM, Caan BJ, Potter JD, Ma KN. Eating patterns and risk of colon cancer. Am J Epidemiol 1998 Jul 1;148(1):4-16.
(62) Williams CD, Satia JA, Adair LS, Stevens J, Galanko J, Keku TO, et al. Dietary patterns, food groups, and rectal cancer risk in Whites and African-Americans. Cancer Epidemiol Biomarkers Prev 2009 May;18(5):1552-61.
(63) Wiseman M. The second World Cancer Research Fund/American Institute for Cancer Research expert report. Food, nutrition, physical activity, and the prevention of cancer: a global perspective. Proc Nutr Soc 2008 Aug;67(3):253-6.
(64) Tsong WH, Koh WP, Yuan JM, Wang R, Sun CL, Yu MC. Cigarettes and alcohol in relation to colorectal cancer: the Singapore Chinese Health Study. Br J Cancer 2007 Mar 12;96(5):821-7.
(65) Seitz HK, Simanowski UA, Kommerell B. [Alcohol and cancer]. Z Gastroenterol 1988 Oct;26 Suppl 3:106-19.
(66) Poschl G, Seitz HK. Alcohol and cancer. Alcohol Alcohol 2004 May;39(3):155-65.
(67) Bagnardi V, Blangiardo M, La VC, Corrao G. A meta-analysis of alcohol drinking and cancer risk. Br J Cancer 2001 Nov 30;85(11):1700-5.
(68) 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. Ann Intern Med 2004 Apr 20;140(8):603-13.
(69) Corrao G, Bagnardi V, Zambon A, Arico S. Exploring the dose-response relationship between alcohol consumption and the risk of several alcohol-related conditions: a meta-analysis. Addiction 1999 Oct;94(10):1551-73.
(70) 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. Int J Cancer 2009 Jul 1;125(1):171-80.
(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. Am J Epidemiol 2008 Jun 15;167(12):1397-406.
(72) Moskal A, Norat T, Ferrari P, Riboli E. Alcohol intake and colorectal cancer risk: a dose-response meta-analysis of published cohort studies. Int J Cancer 2007 Feb 1;120(3):664-71.
(73) 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 Epidemiol Biomarkers Prev 2003 Dec;12(12):1492-500.
(74) 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. Br J Cancer 2003 Apr 7;88(7):1038-43.
(75) 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. J Epidemiol 2005 Jun; 15 Suppl 2:S173-S179.
(76) Beral V, Banks E, Reeves G. Evidence from randomised trials on the long-term effects of hormone replacement therapy. Lancet 2002 Sep 21;360(9337):942-4.
(77) Rothwell PM, Wilson M, Elwin CE, Norrving B, Algra A, Warlow CP, et al. Longterm effect of aspirin on colorectal cancer incidence and mortality: 20-year follow-up of five randomised trials. Lancet 2010 Nov 20;376(9754):1741-50.
(78) Baron JA, Beach M, Mandel JS, Van Stolk RU, Haile RW, Sandler RS, et al. Calcium supplements for the prevention of colorectal adenomas. Calcium Polyp Prevention Study Group. N Engl J Med 1999 Jan 14;340(2):101-7.
(79) Flossmann E, Rothwell PM. Effect of aspirin on long-term risk of colorectal cancer: consistent evidence from randomised and observational studies. Lancet 2007 May 12;369(9573):1603-13.
(80) Chan AT, Giovannucci EL. Primary prevention of colorectal cancer. Gastroenterology 2010 Jun;138(6):2029-43.
(81) Ezzati M, Lopez AD. Estimates of global mortality attributable to smoking in 2000. Lancet 2003 Sep 13;362(9387):847-52.
(82) Irigaray P, Newby JA, Clapp R, Hardell L, Howard V, Montagnier L, et al. Lifestylerelated factors and environmental agents causing cancer: an overview. Biomed Pharmacother 2007 Dec;61(10):640-58.
(83) Botteri E, Iodice S, Bagnardi V, Raimondi S, Lowenfels AB, Maisonneuve P. Smoking and colorectal cancer: a meta-analysis. JAMA 2008 Dec 17;300(23):276578.
(84) Liang PS, Chen TY, Giovannucci E. Cigarette smoking and colorectal cancer incidence and mortality: systematic review and meta-analysis. Int J Cancer 2009 May 15;124(10):2406-15.
(85) 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. Clin Gastroenterol Hepatol 2009 Jun;7(6):682-8.
(86) 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 U.S. women. J Natl Cancer Inst 1994 Feb 2;86(3):192-9.
(87) Giovannucci E, Martinez ME. Tobacco, colorectal cancer, and adenomas: a review of the evidence. J Natl Cancer Inst 1996 Dec 4;88(23):1717-30.
(88) Leren P, Askevold EM, Foss OP, Froili A, Grymyr D, Helgeland A, et al. The Oslo study. Cardiovascular disease in middle-aged and young Oslo men. Acta Med Scand Suppl 1975;588:1-38.
(89) Lund HL, Wisloff TF, Holme I, Nafstad P. Metabolic syndrome predicts prostate cancer in a cohort of middle-aged Norwegian men followed for 27 years. Am J Epidemiol 2006 Oct 15;164(8):769-74.
(90) Solberg LA, Strong JP, Holme I, Helgeland A, Hjermann I, Leren P, et al. Stenoses in the coronary arteries. Relation to atherosclerotic lesions, coronary heart disease, and risk factors. The Oslo Study. Lab Invest 1985 Dec;53(6):648-55.
(91) Stocks T, Borena W, Strohmaier S, Bjorge T, Manjer J, Engeland A, et al. Cohort Profile: The Metabolic syndrome and Cancer project (Me-Can). Int J Epidemiol 2010 Jun;39(3):660-7.
(92) Bjartveit K, Foss OP, Gjervig T, Lund-Larsen PG. The cardiovascular disease study in Norwegian counties. Background and organization. Acta Med Scand Suppl 1979;634:1-70.
(93) Tverdal A, Bjartveit K. Health consequences of reduced daily cigarette consumption. Tob Control 2006 Dec;15(6):472-80.
(94) Aires N, Selmer R, Thelle D. The validity of self-reported leisure time physical activity, and its relationship to serum cholesterol, blood pressure and body mass index. A population based study of 332,182 men and women aged $40-42$ years. Eur J Epidemiol 2003;18(6):479-85.
(95) Bjartveit K, Stensvold I, Lund-Larsen PG, Gjervig T, Kruger O, Urdal P. [Cardiovascular screenings in Norwegian counties. Background and implementation. Status of risk pattern during the period 1986-90 among persons aged 40-42 years in 14 counties]. Tidsskr Nor Laegeforen 1991 Jun 30;111(17):2063-72.
(96) Aamodt G, Sogaard AJ, Naess O, Beckstrom AC, Samuelsen SO. [The CONOR database--a little piece of Norway]. Tidsskr Nor Laegeforen 2010 Feb 11;130(3):2645.
(97) Naess O, Sogaard AJ, Arnesen E, Beckstrom AC, Bjertness E, Engeland A, et al. Cohort profile: cohort of Norway (CONOR). Int J Epidemiol 2008 Jun;37(3):481-5.
(98) Larsen IK, Smastuen M, Johannesen TB, Langmark F, Parkin DM, Bray F, et al. Data quality at the Cancer Registry of Norway: an overview of comparability, completeness, validity and timeliness. Eur J Cancer 2009 May;45(7):1218-31.
(99) Svensson E, Grotmol T, Hoff G, Langmark F, Norstein J, Tretli S. Trends in colorectal cancer incidence in Norway by gender and anatomic site: an age-period-cohort analysis. Eur J Cancer Prev 2002 Oct;11(5):489-95.
(100) R Bonita, R Beaglehole, T Kjellstrom. Basic Epidemiology. 2nd ed. World Health Organization; 2006.
(101) Rothman KJ. Epidemiology: An Introduction. 2nd ed. Oxford: Oxford University Press; 2012.
(102) R Bonita, R Beaglehole, T Kjellstrom. Types of studies. 2nd edition ed. World Health Organization; 2014.
(103) Rothman KJ, Greenland S. Causation and causal inference in epidemiology. Am J Public Health 2005;95 Suppl 1:S144-S150.
(104) Gerhard T. Bias: considerations for research practice. Am J Health Syst Pharm 2008 Nov 15;65(22):2159-68.
(105) Henderson M, Page L. Appraising the evidence: what is selection bias? Evid Based Ment Health 2007 Aug;10(3):67-8.
(106) Galea S, Tracy M. Participation rates in epidemiologic studies. Ann Epidemiol 2007 Sep;17(9):643-53.
(107) Gerrits MH, van den Oord EJ, Voogt R. An evaluation of nonresponse bias in peer, self, and teacher ratings of children's psychosocial adjustment. J Child Psychol Psychiatry $2001 \mathrm{Jul} ; 42(5): 593-602$.
(108) Kypri K, Stephenson S, Langley J. Assessment of nonresponse bias in an internet survey of alcohol use. Alcohol Clin Exp Res 2004 Apr;28(4):630-4.
(109) Grimes DA, Schulz KF. Bias and causal associations in observational research. Lancet 2002 Jan 19;359(9302):248-52.
(110) Rossouw JE, Anderson GL, Prentice RL, Lacroix AZ, Kooperberg C, Stefanick ML, et al. Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results From the Women's Health Initiative randomized controlled trial. JAMA 2002 Jul 17;288(3):321-33.
(111) Strand BH, Steiro A. [Alcohol consumption, income and education in Norway, 19932000]. Tidsskr Nor Laegeforen 2003 Oct 23;123(20):2849-53.
(112) Edwards LK, Barry I.G, Douglas M. Time-to-Event Analysis of Longitudinal Followup of a Survey: Choice of the Time-scale. American Journal of Epidemiology 1997;145:72-80.
(113) Raj Bhopal. Epidemiological study design and principles of data analysis: An integrated suite of methods. In: Oxford University Press, editor. Concepts of epidemiology;Integrating the ideas , theories, principles and methods of epidemiology. second edition ed. 2008.
(114) Bonithon-Kopp C, Benhamiche AM. Are there several colorectal cancers? Epidemiological data. Eur J Cancer Prev 1999 Dec;8 Suppl 1:S3-12.
(115) McCashland TM, Brand R, Lyden E, de GP. Gender differences in colorectal polyps and tumors. Am J Gastroenterol 2001 Mar;96(3):882-6.
(116) Stewart RJ, Stewart AW, Turnbull PR, Isbister WH. Sex differences in subsite incidence of large-bowel cancer. Dis Colon Rectum 1983 Oct;26(10):658-60.
(117) Colangelo LA, Gapstur SM, Gann PH, Dyer AR. Cigarette smoking and colorectal carcinoma mortality in a cohort with long-term follow-up. Cancer 2004 Jan 15;100(2):288-93.
(118) 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. Br J Cancer 2003 Apr 7;88(7):1038-43.
(119) Gram IT, Braaten T, Lund E, Le Marchand L, Weiderpass E. Cigarette smoking and risk of colorectal cancer among Norwegian women. Cancer Causes Control 2009 Aug;20(6):895-903.
(120) Limburg PJ, Vierkant RA, Cerhan JR, Yang P, Lazovich D, Potter JD, et al. Cigarette smoking and colorectal cancer: long-term, subsite-specific risks in a cohort study of postmenopausal women. Clin Gastroenterol Hepatol 2003 May;1(3):202-10.
(121) 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. J Natl Cancer Inst 2007 Nov 21;99(22):1729-35.
(122) Terry PD, Miller AB, Rohan TE. Prospective cohort study of cigarette smoking and colorectal cancer risk in women. Int J Cancer 2002 May 20;99(3):480-3.
(123) 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. Clin Gastroenterol Hepatol 2011 Feb;9(2):137-44.
(124) Cucino C, Buchner AM, Sonnenberg A. Continued rightward shift of colorectal cancer. Dis Colon Rectum 2002 Aug;45(8):1035-40.
(125) 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. J Epidemiol 2003 Nov;13(6):323-32.
(126) Yoo KY, Tajima K, Inoue M, Takezaki T, Hirose K, Hamajima N, et al. Reproductive factors related to the risk of colorectal cancer by subsite: a case-control analysis. Br J Cancer 1999 Apr;79(11-12):1901-6.
(127) Pope M, Ashley MJ, Ferrence R. The carcinogenic and toxic effects of tobacco smoke: are women particularly susceptible? J Gend Specif Med 1999 Nov;2(6):45-51.
(128) 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 Epidemiol Biomarkers Prev 2009 Oct;18(10):2745-50.
(129) Mai SK, Welzel G, Haegele V, Wenz F. The influence of smoking and other risk factors on the outcome after radiochemotherapy for anal cancer. Radiat Oncol 2007;2:30.

## PAPER I

Parajuli R, Bjerkaas E, Tverdal A, Selmer R, Le Marchand L, Weiderpass E, Gram IT
The increased risk of colon cancer due to cigarette smoking may be greater in women than men

Cancer Epidemiol Biomarkers Prev.2013; 22(5), 862-71
PubMED: PMID 23632818

## PAPER II

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
(Submitted to BMC Cancer)

## PAPER III

Parajuli R, Bjerkaas E, Tverdal A, Le Marchand L, Weiderpass E, Gram IT
Cigarette smoking and colorectal cancer mortality among 602,242
Norwegian males and females
Clinical Epidemiology, Dovepress (Online)

## Appendices



1. Surveys questionnaires
2. Description of methodology
3. Variable description
4. Summary of some of the prospective studies examining the association between smoking and colorectal cancer published between 2002-2013

Appendix 1

QUESTIONAIRE OSLO HEALTH STUDY I

## HJERTE - KARSYKDOMMER

SPøRRESK」EMA


## QUESTIONAIRE THREE COUNTIES <br> FINNMARK COUNTY, ROUND 1 AND 2 NORWEGIAN



MELDING OM SKJERMBILDEFOTOGRAFERING OG HJERTE-KARUNDERS®KELSE
(Gjelder bare den person brevet er adressert til)
$\Gamma$
$\llcorner$

Skjermbildefotograforingen kommer na til Deres distrikt.
Tid og sted for Deres frammote vil De finne Ogsa denne gangon vil on del av befolkningen à tilbud om hjerto-karundersokelse. Do tilhorer denne gruppe. En orientering om undersokelsen or gitt i vedlagte brosjyre.
Vennligst fyll ut sporreskjemaet pa baksidon og ta det med til undersokolson. Ta ogsa mod
tuberkulinkort eller helsebok, om De har. Fraveer bes eventuelt meldt pà vedlagte seddel. Med hilsen
$-1$
HELSERADET FYLKESLEGEN
STATENS SKJERMBILDEFOTOGRAFERING

Krotanr.
Klokkenlet


M: $\quad$ te



# QUESTIONNAIRE THREE COUNTIES STUDY, SOGN OG FJORDANE AND OPPLAND COUNTIES, ROUND 1 AND 2 



MELDING OM SKJERMBILDEFOTOGRAFERING OG HJERTE-KARUNDERS®KELSE
(Gjelder bare den person brevet er adressert til)
$\Gamma$
$\llcorner$

Skjermbildefotograforingen kommer na til Deres distrikt.
Tid og sted for Deres frammote vil De finne Ogsa denne gangon vil on del av befolkningen à tilbud om hjerto-karundersokelse. Do tilhorer denne gruppe. En orientering om undersokelsen or gitt i vedlagte brosjyre.
Vennligst fyll ut sporreskjemaet pa baksidon og ta det med til undersokolson. Ta ogsa mod
tuberkulinkort eller helsebok, om De har. Fraveer bes eventuelt meldt pà vedlagte seddel. Med hilsen
$-1$
HELSERADET FYLKESLEGEN
STATENS SKJERMBILDEFOTOGRAFERING

Krotanr.
Klokkenlet


M: $\quad$ te



## QUESTIONNAIRE THREE COUNTIES STUDY, ALL COUNTIES COUNTY, ROUND 3 <br> NORWEGIAN



## QUESTIONNAIRE 40 YEARS STUDY, ROUND 1



## QUESTIONNAIRE 40 YEARS STUDY, ROUND 2

Har en eller flere av foreldre eller sosken hatt hjerteinfarkt (sảr pả hjertet) efler angina pectoris (hjertekrampe)?

## EGEN SYKDOM

Har De, eller har De hatt
Hjerteinfarkt?
Angina pectorisihjertekrampe)? $\ldots \ldots .1^{13}$
Hjemeslag? 14

Sukkersyke? 16

Hvis De har sukkersyke, i hvilket àr
ble diagnosen stillet?
19
Er De under medikamentell behandling
for høyt blodtrykk?

## C SYMPTOMER

Fâr De smerter eller ubehag i brystet nár De: Gas i bakker, trapper eller
fort pá flat mark?
a flat mark? $\qquad$
Gár i vanlig takt pà flat mark?.
Dersom De fár smerter eller vondt
i brystet ved gange, pleier De da á: Stoppe?

22
Saktne farten?
Fortsette i samme takt?
....
$\qquad$

Dersom De stopper eller saktner farten,
forsvinner smertene da:
Etter mindre enn 10 minutter?
Etter mer enn 10 minutter?
Har De vanligvis:
Hoste om morgenen? . . . . . . . . . . . . . . . . . . . . . 24
Oppspylt fra brystet om morgenen? . . . . . . . . 25

Bevegelse og kroppslig anstrengelse i Deres fritid. Hvis aktiviteten varierer meget f.eks. mellom sommer og vinter. sa ta et gjennomsnitt. Sporsmalet gjelder bare det siste áret.
Sett kryas i den ruta hwor *JA× passer best
Leser, ser pá fjernsyn eller annen stillesittende beskjeftigelse? $\qquad$ . . 26

Spaserer, sykler eller beveger Dem pá annen mate minst 4 timer i uka?
(Her skal De ogsả regne med gang eller
sykling til arbeidsstedet, sondagsturer m.m.)
Criver mosjonsidrett, tyngre hagearbeid e.l.?
Merk at aktiviteten skal vare minst
4 timer i uka.)
Trener hardt eller driver konkurranseidrett regelmessig og flere ganger i uka?

## SALT/FETT

Hvor ofte bruker De salt kjott
eller salt fisk til middag?
Sell kryss i den ruta hwor ajAn passer best
Abri eller sjeldnere enn en gang
i mảneden.
ang i uka
Contil to ganger i uka
Mer enn to ganger i uka
....
$\qquad$

Hvor ofte pleier De stro ekstra salt pá middagsmaten?
Selt kryss i den ruta hwor ..JA." passer best
Sjelden eller aldri
......... Av og til eller ofte.
....

Alltid eller nesten allid.
Hva slags margarin eller smor bruker De til vanligy pả brod?
Selt kryss i den ruta hyor "Ja." passer best
Bruker ikke smor eller margarin pá bred .... 29 Smer
Hard margarin.
Myk (Soft) margarin
Smer/margarin blanding
Hva slags fett blir til vanlig brukt til
matlaging i Deres husholdning?
Sett kryss i den ruta hoor "JA. passer best
Smer eller hard margarin......
Myk (Soft) mangarin eller olje
Smor/margarin blanding


## QUESTIONNAIRE 40 YEARS STUDY, ROUND 3

EGEN HELSE
Hvordan er helsen din ná? Sett bare ett kryss.


Bruker du medisin mot hoyt blodtrykk?


Hvis ja, hvilket merke bruker du ná?
$\square$ Ikke skiv her
Har du i lopet av det siste aret vært plaget med smerter og/eller stivhet i muskier og ledd som 111 har vart minst 3 maneder sammenhongende?

Har du de siste to ukene folt deg:

|  | Nei | Litt | En god | Svart |
| :---: | :---: | :---: | :---: | :---: |
| Nervas og urolig? ....... 34 |  |  |  |  |
| Plaget av angst? .......... 35 |  |  |  |  |
| Trygg og rolig? ............ 36 |  |  |  |  |
| Irritabel?..................... 37 |  |  |  |  |
| Glad og optimistisk? .... 38 |  |  |  |  |
| Nedfor/deprimert? ........ 39 |  |  |  |  |
| Ensom? ..................... 40 |  |  |  |  |

Fâr du smerter eller ubehag i brystet nár du:
Gâr i bakker, trapper eller fort på flat mark?
Hvis du fár slike smerter, pleier du da à: Stoppe?
Saktne farten?
Fortsette i samme takt?
Dersom du stopper, forsvinner smertene da etter mindre enn 10 minutter?
Kan slike smerter like gjerne opptre
Mottar du ná noen av folgende ytelser?
Syketrygd (sykmeldt)
$\qquad$
Attforingspenger $\qquad$
. ........................................................
Ald $\qquad$ Arbeidsledighetstrygd SEVANER
Dette gjelder din interesse for á endre
helsevaner. Roykesporsmàlet
besvares bare av dem som noyker.
Har du de siste 12 mnd. forsøkt å:


Om 5 ar, tror du at du har endret vaner pà noen av disse omrảdene?

Anslà din hoyeste og laveste vekt i loppet av de siste 5 ar.
(Se bort fra vekt under svangerskap)


## SYKDOM I FAMILIEN

Har en eller flere av foreldre eller sasken hatt hjerteinfarkt (sár pá hjertet) eller angina pectoris (hjertekrampe)?

| IA | NEI | VEKE |
| :---: | :---: | :---: |
|  |  |  |

Har én eller flere foreldre/søsken hatt: Hjerteinfarkt for de fylte 60 àr? $\qquad$ .. 62
Hjerneslag for de fylte 70 àr? $\qquad$

## RØYKING

Hvor lenge er du vanligvis daglig
til stede i roykfylt rom?................................... Sett O hvis du ikke oppholder deg I reykfylt som.
Royker du selv?
Sigaretter daglig?
Sigarer/sigarillos daglig?67

Pipe daglig? .................................................. 68
Hvis du har roykt daglig tidligere, hvor lenge er det siden du sluttet? $\qquad$ .69

Huis du røyker daglig ná eller har roykt tidligere:

Hoor mange sigaretter royker eller
roykte du vanligvis daglig?
.............
Hvor gammel
royke daglig? $\qquad$
Hivor mange àr tilsammen har du roykt daglig?


## MOSJON

Hvordan har din fysiske aktivitet i fritiden vsert det siste âret? Tenk deg et ukentlig gjennomsnitt for áret.


KAFFE/TE/ALKOHOL
Hvor mange kopper kafte/te drikker du daglig?
Sett 0 hvis ou ikke orikker kaffe/te daglig.
Kokekaffe. $\qquad$
Annen kaffe $\qquad$
$\qquad$
Er du total avholdsmann/-kvinne? $\qquad$ 87

| Amail kopper |
| :--- |
| Antall kopper |
| Antall kopper  <br> JA NEI <br>   |

Hvor mange ganger i måneden drikker du vanlig vis alkohol? Regn ikke med lettol.
Sett 0 hvis mindre enn 1 gang i mnd $\qquad$
Hvor mange glass al, vin eller brennevin drikker du vanligvis i lopet av to uker? 90
Regn ikke med lettøl.
Sett o hvis du ikke drikker alkohol.


## FETT

Hva slags margarin eller smør bruker du vanligvis pá bradet? Sett ett kryss.

Bruker ikke smør/margarin ........................................ 96
Meierismar..
Hard margarin
Blot (soft) margarin
Smerimargarin blanding $\qquad$
Lettmargarin


## UTDANNING

Hvilken utdanning er den høyeste du har fullfart? Grunnskole 7-10 âr, framhaldsskole,
folkehøgskole. $\qquad$
Realskole, middelskole, yrkesskole, 1-2 árig
videregáende skole
$\qquad$
Artium, øk.gymnas, allmennfaglig retning
i videregâende skole.
$\qquad$
Hagskole/universitet, mindre enn 4 ár
Hegskole/universitet, 4 àr eller mer $\qquad$


## ETTERUNDERSGKELSE

Hvis denne helseundersakelsen visor at du bor undersokes næermere, hvilken allmennpraktiserende lege/kommunelege onsker du da â bli henvist til? Oppgi legens navn:

## QUESTIONNAIRE 40 YEARS STUDY, ROUND 4



## 6. UTDANNING

Hvilken utdanning er den høyeste du har fullført? Sett bare ett kryss.

Mindre enn 7 ár grunnskole $\qquad$ $\square$
Grunnskole 7-10 ár, framhaldsskole,
folkehogskole.
................................
Realskole, middelskole, yrkesskole,
1-2 árig videregàende skole $\qquad$
Artium, ok.gymnas, allmennfaglig retning
i videregáende skole $\qquad$ $\square 3$

Høgskole/universitet, mindre enn 4 ár $\qquad$ $\square 4$

Hogskole/universitet, 4 ár eller mer. $\qquad$ $\square 5$

## 7. KOST

Hvor ofte bruker du disse matvarene?
Sett kryss i de rutene som beskriver dift forbruk best.

| $\begin{aligned} & \text { Flareg } \\ & \text { wigig } \end{aligned}$ | $\mathrm{Dagig}_{0}$ | 1-59. | $\begin{aligned} & 1-3 \mathrm{l}, \\ & \text { prand } \end{aligned}$ |  |
| :---: | :---: | :---: | :---: | :---: |
| Fisk (middag, paslegg) ..... $\square$ | $\square$ | $\square$ | $\square$ |  |
| Frukt/grent..................... $\square$ | $\square$ | $\square$ | $\square$ |  |
| Helmelk, kefir, yoghurt .... $\square$ | $\square$ | $\square$ | $\square$ | $\square$ |
| Lettmelk, lettyoghurt ....... $\square$ | $\square$ | $\square$ | $\square$ | $\square$ |
| Skummet melk (surisat) .. $\square_{1}$ | $\square$ | $\square$ | $\square$ | $\square$ |

Hva slags smor eller margarin bruker du vanligvis PA BRøDET?
Sett kryss i den ruta som passer best.

| Bruker ikke smar/margarin. |  |
| :---: | :---: |
| Meierismar. | 2 |
| Hard margarin | 3 |
| Blat (soft) margarin. | 4 |
| Smør/margarin blanding | 5 |
| Lettmargarin/iettsmor (Brele | $\square 6$ |

Hva slags fett bruker du/dere vanligvis TIL MATLAGING? Seft kryss iden ruta som passer best.

| Smrr/margarin. | 1 |
| :---: | :---: |
| Myk (soft) margariniolje | 2 |
| Bare olje | 3 |
| Vet ikke. | 4 |

## 8. KAFFE /TE/ALKOHOL

Hvor mange kopper kaffeite drikker du daglig? Sett O hvis du ikke drikker kaffe/te daglig.
Antall kopper daglig
$\mathrm{Il}_{\mathrm{Te}}$
Kokekaffe Annen kaffe

Er du total avholdsmannf-kvinne?. $\qquad$
Hvor mange ganger I màneden drikker du vanligvis alkohol? Regn ikke med lettor. Sett O hvis mindre enn 1 gang i mnd. Antall ganger


Hvor mange glass øl, vin eller brennevin drikker du VANLIGVIS I lopet av to uker? Regn ikke med lettol. Sett 0 hvis du ikke drikker alkohol.
9. RØYKING

Hvor lenge er du vanligvis daglig tilstede i røykfylt rom? $\qquad$ Antall hele timer $\square$ Sett 0 hvis du jkke oppholder deg i roykfylt rom.

Rayker du selv:
Sigaretter daglig?
Sigarer/sigarillos daglig? $\qquad$


NEI

Pipe daglig?
Aldri røykt daglig $\qquad$ (Sett kryss)

Hvis du har roykt daglig tidligere, hvor lenge er det siden du sluttet? $\qquad$ Antall ár $\qquad$
Hvis du royker daglig ná eller har roykt tidligere:

| Hvor mange sigaretter røyker eller raykte du vanligvis daglig? <br> Antall sigaretter |  |
| :---: | :---: |
| Hvor gammel var du da du begynte à royke daglig? $\qquad$ Alder iár |  |
| Hvor mange àr til sammen har du reykt daglig? $\qquad$ Antall àr |  |

## 10. MOSJON

Hvordan har din fysiske aktivitet i fritiden vært det siste áret?
Tenk deg et ukentlig gjennomsnitt for àret.
Arbeiơsvei regnes som fritid. Besvar begge spørsmálene.

| Timer pr, uke |  |  |
| :---: | :---: | :---: |
| Under 1 | $\mathbf{1 - 2}$ | 3ogmer |
| $\square$ | $\square$ | $\square$ |
| $\square$ | $\square$ | $\square$ |

Bevegelse og kroppslig anstrengelse I din fritid. Hvis aktiviteten varierer meget f.eks. mellom sommer og vinter, sẩ ta et gjennomsnitt. Sporsmálet gjelder bare det siste âret.
Sett kryss iden ruta som passer best.
Leser, ser pâ fjernsyn eller annen
stillesittende beskjeftigelse?. $\qquad$
$\qquad$

Spaserer, sykler eller beveger deg pả
annen mâte minst 4 timer i uka? $\qquad$ ........................... $\square 2$
(Her skal du ogsá regne med gang eiller
sykling til arbeidsstedet, sandagsturer m.m.)
Driver mosjonsidrett, tyngre hagearbeid e.l.? $\qquad$
(Merk at aktiviteten skal vare minst 4 timer i uka)
Trener hardt eller driver konkurranseidrett regelmessig og flere ganger i uka? $\square 4$

## 11. ENDRING AV HELSEVANER



12. MEDISIN MOT HOYT BLODTRYKK

Bruker du medisin mot hoyt blodtrykk?

| Nà | Far, men ikke nà | Aldri brukt |
| :---: | :---: | :---: |
| $\square 1$ | $\square 2$ | $\square 3$ |

Hvis du bruker medisin nà, hvilke(t) merke(r) bruker du?

ikke skriv idisse rutene

## 13. MEDISIN MOT HOYT KOLESTEROL

Bruker du kolesterolsenkende medisiner NÅ?
Hvis NEI, gâ til 14. ETTERUNDERSøKELSE. $\qquad$

Hvor gammel var du da du begynte med kolesterolsenkende medisiner? Alder iáar
$\square$

Hvis du bruker kolesterolsenkende medisiner, hva var grunnen til at du begynte med slik medisin?


Hvilke kolesterolsenkende medisiner bruker du NÅ og hvilken dose bruker du?
Hvilke(t) merke(r) bruker du?

$\qquad$
$\square$


## 14. ETTERUNDERSOKELSE

Hvis denne helseundersokelsen viser at du bor undersøkes nærmere, hvilken allmennpraktiserende lege/kommunelege onsker du da á bli henvist til?
Oppgi legens navn:
15. TIL KVINNER SOM DELTAR I HELSEUNDERSOKELSEN

Hvor gammel var du da du fikk menstruasjon aller forste gang?

Alder ià àr
Har du for tiden regelmessig menstruasjon?
Regn den for regelmessig hvis den ikke har vært
borte mer enn 3 mnd. sammenhengende siste ar.
JA NEI

Til deg som svarte JA: Omtrent hvor mange dager etter starten pá siste menstruasjon

T
skjer helseundersokelsen? (Sett bare ett kryss)


Hvis du for tiden ikke har regelmessig menstruasjon, ber vi deg fylle ut nedenfor' (Sett bare eft kryss)


Til deg som bruker p-pille, hormonspiral (ikke vanlig spiral) eller hormoner i overgangsalderen $N \AA$ A:
Hvilke(t) merke(r) bruker du?


Take far utfyllingen!
Nok en gang:
Velkammen til undersakelsen!


|  | svett/andpusten) <br> Hard fysisk aktivitet (svett/andpusten) |
| :---: | :---: |
| a6_3 | 5 b. Angi bevegelse og kroppslig anstrengelse i din fritid. Hvis aktiviteten varierer meget f.eks mellom sommer og vinter, så ta et gjennomsnitt. <br> Spørsmålet gjelder bare det siste året. <br> (Sett ett kryss i den ruta som passer best) <br> Lese, ser på fjernsyn eller annen stillesittende beskjeftigelse? <br> Spaserer, sykler eller beveger deg på annen måte minst 4 timer i uka? <br> (Her skal du regne med gang eller sykling til arbeidsstedet, søndagsturer m.m) <br> Driver mosjonsidrett, tyngre hagearbeid e.l? <br> (Merk at aktiviteten skal vare minst 4 timer i uka) <br> Trener hardt eller driver konkurranseidrett regelmessig og flere ganger i uka |
|  | RøYKING |
| a7_2 | 6. Hvor lenge er du vanligvis daglig til stede i røykfylt rom? Sett 0 hvis du ikke oppholder deg i røykfylt rom. Antall timer. |
| a7_3 | 7. Røkte noen av de voksne hjemme da du vokste opp? Ja Nei |
| a7_4 | 8. Bor du/har du bodd sammen med noen daglig-røykere etter fylte 20 år? Ja Nei |
| a8_0 to a8_3 | 9. Røyker du selv ? |
| a9 | 10. Hvis du har røykt daglig tidligere, hvor lenge er det siden du sluttet? |
| a10 | 11. Hvis du røker daglig nå eller har røykt tidligere: Hvor mange sigaretter røyker eller røykte du vanligvis daglig? Antall sigaretter. |
| a11 | 12. Hvor gammel var du da du begnte å reyke? |
| a12_1 | 13. Hvor mange år til sammen har du roykt daglig ? |


|  | ..år |
| :---: | :---: |
|  | KAFFE, TE OG ALKOHOL |
| $\begin{aligned} & \text { a13_1 to a13_2 } \\ & \text { a13_4 } \end{aligned}$ | 14.a Hvor mange kopper kaffe drikker du daglig? <br> Sett 0 hvis du ikke drikker kaffe daglig <br> Kokekaffe, antall kopper. $\qquad$ <br> Annen kaffe, antall kopper. $\qquad$ |
| a13_5 to a13_8 | 14.b Hva slags kaffe drikke du vanligvis? <br> Sett kryss <br> Filter-/pulverkaffe <br> Kokekaffe/trykkanne <br> Annen kaffe (espresso og lignende) <br> Drikker ikke kaffe |
| a13_9 to a13_10 | 14c. Hvor mange kopper kaffe/te drikker du daglig? Sett 0 hvis du ikke drikker kaffe/te daglig Antall kopper kaffe. $\qquad$ Antall kopper te. $\qquad$ |
| a14_1 and a14_1_2 (a14_1 made of 14_1_1 and 14_1_2) | 15 a. Hvor mange ganger i måneden drikker du vanligvis alkohol? Regn ikke med lettøl. Sett 0 hvis mindre enn 1 gang i måneden. Antall ganger.............. |
| a14_1 and a14_1_1 <br> (a14_1 made of 14_1_1 and 14_1_2) | 15 b. Omtrent hvor ofte har du i løpet av det siste året drukket alkohol? <br> (Lettøl og alkoholfritt øl regnes ikke med) <br> 4-7 ganger i uka <br> 2-3 ganger i uka <br> Ca 1.gang i uka <br> 2-3 ganger pr måned <br> Omtrent1 gang i mnd. <br> Noen få ganger siste år <br> Har ikke drukket alkohol siste år <br> Har aldri drukket alkohol |
| a14_4_1, a14_5_1 | 16 a. Hvor mange glass $\boldsymbol{\varnothing}$, vin eller brennevin drikker du vanligvis i løpet av to uker? Regn ikke med lettøl. Sett 0 hvis du ikke drikker alkohol. <br> Øl.....glass Vin.....glass Brennevin.....glass |
| a14_2 | Til dem som har drukket siste år <br> 16 b. Når du har drukket alkohol, hvor mange glass/og eller drinker har du vanligvis drukket? <br> Antall. |
| a14_3 | 16 c . Omtrent hvor mange ganger i løpet av det siste året har du drukket så mye som minst 5 glass og/eller drinker $i$ løpet av et døgn? <br> Antall ganger........... |


| a14_4, a14_5, | 16 d. Når du drikker alkohol, drikker du da vanligvis: (Sett ett eller flere kryss). |
| :---: | :---: |
| a14_7 | 17. Er du total avholdsmann/-kvinne? <br> Ja <br> Nei |
|  | SKOLEGANG |
| a15, a15_2 <br> (made of a15_1 and a15_2) | 18 a. Hvilken utdanning er den høyeste du har fullført? <br> Mindre enn 7 år grunnskole <br> Grunnskole 7-10 år, framhaldsskole, folkehøyskole <br> Realskole, middelskole, yrkesskole, 1-2 årig videregående skole <br> Artium, økonomisk gymnas, allmennfaglig retning i videregående skole <br> Høgskole/universitet, mindre enn 4 år <br> Høgskole/universitet, 4 år eller mer |
| a15, a15_1 <br> (made of a15_1 and a15_2) | 18 b. Hvor mange års skolegang har du gjennomført? (Ta med alle år du har gått på skole eller studert) Antall år.............. |
|  | SYKDOM I FAMILIEN |
| a16 | 19. Har en eller flere av foreldre eller søsken hatt hjerteinfarkt (sår på hjertet) eller angina pectoris (hjertekrampe)? <br> Ja <br> Nei <br> Vet ikke |
| b15_1 to b15_30 | 20. Kryss for de slektninger som har eller har hatt noen av sykdommene: <br> Mor Far Bror Søster Barn <br> Hjerneslag eller <br> hjerneblødning <br> Hjerteinfarkt før 60 <br> års alder <br> Astma <br> Kreftsykdom <br> Sukkersyke (diabetes) <br> Alder da de fikk sukkersyke |
|  | LOKALMILJØ OG BOLIG |
| b1 | 21. I hvilken kommune bodde du da du fylte 1 år? <br> Hvis du ikke bodde i Norge, oppgi hvilket land i stedet for fylke. $\qquad$ |
| b2 | 22. Hvilken type bolig bor du i? <br> Enebolig/ villa <br> Gårdsbruk <br> Blokk/terrasseleilighet |


|  | Rekkehus/2-4mannsbolig Annen bolig/institusjon/omsorgsbolig |
| :---: | :---: |
| b3 | 23. Hvor stor er din boenhet? .........m2 |
| b29 | 24. Er det heldekkende tepper i stua? Nei |
| b30 | 25. Er det katt i boligen? |
|  | FAMILIE OG VENNER |
| Sjekke | 26a. Hvem bor du sammen med? Sett ett kryss for hert spørsmål og angi antall.   <br>  Ja Nei Antall <br> Ektefelle/samboer    <br> Andre personer over 18 år    <br> Personer under 18 år    |
| b4_1 to b4_6 | 26 b. Bor du sammen med noen? <br> Ja <br> Nei <br> Hvis JA: <br> Ektefelle/samboer <br> Andre personer, 18 år og eldre <br> Personer under 18 å |
| b4_7 and b4_8 | 26 c (kun på eldreskjema) <br> Bor du ? Sett kryss <br> Hjemme <br> Institusjon/bofellesskap <br> Bor du sammen med? <br> Ektefelle/samboer? <br> Andre personer? |
| b31 | 27. Hvor mange av barna har plass i barnehage? |
| b5 | 28. Hvor mange gode venner har du? Regn med de du kan snakke fortrolig med og som kan gi deg hjelp når du trenger det? (Tell ike med de du bor sammen med, men ta med andre slektninger) …................... |
| b6 | 29. Foler du at du har nok gode venner? |


|  | Nei |
| :---: | :---: |
| b | 30. Hvor ofte tar du vanligvis del i foreningsvirksomhet som for eksempel syklubb, idrettslag, politiske lag, religiøse eller andre foreninger? Aldri, eller noen få ganger i året <br> 1-2 ganger i måneden (før år 1996), 1-3 ganger i måneden (etter år 1996) Omtrent 1 gang i uken <br> Mer enn en gang i uken |
|  | ARBEID |
| b8_1 to b8_4 | 31. Hva slags arbeidssituasjon har du nå? Lønnet arbeid Heltids husarbeid Utdanning, militærtjeneste Arbeidsledig, permittert |
| b9 and bs_1 | 32a. Hvor mange timer lønnet arbeid har du i uka? .....................timer |
| b9 | 32 b . Er du i inntektsgivende arbeid? <br> Ja, full tid <br> Ja, deltid <br> Nei |
| b10_1, b10_2, b10_3 b10_4, b10_5, b10_6 b10_7 | 33. Mottar du noen av følgende ytelser? <br> Sykepenger (er sykemeldt) Alderstrygd, førtidspensjon (AFP) eller etterlattepensjon Rehabiliterings-/attføringspenger Uførepensjon (helt eller delvis) Dagpenger under arbeidsledighet So1sialhjelp/stønad Overgangsstønad for enslige forsørgere |
| b11 | 34. Har du skiftarbeid, nattarbeid eller går vakter? <br> Ja <br> Nei |
| ${ }^{612}$ | 35. Hvis du er i lønnet eller ulønnet arbeid, hvordan vil du beskrive arbeidet ditt? For det meste stillesittende arbeid? <br> (f.eks1 skrivebordsarbeid, montering) <br> Arbeid som krever at du går mye? <br> (f.eks ekspeditørarbeid, lett industriarbeid, undervisning) <br> Arbeid der du går og løfter mye? <br> (f.eks postbud, pleier, bygningsarbeider) <br> Tungt kroppsarbeid?(f.eks skogsarbeid, tungt jordbruksarbeid, tungt bygningsarbeid) |
| b32 | 36. Kan du selv bestemme hvordan arbeidet ditt skal legges opp? (Sett bare ett kryss) |


|  | Nei, ikke i det hele tatt <br> I liten grad <br> Ja, stort sett <br> Ja, det bestemmer jeg selv |
| :---: | :---: |
| b33_1, b33_2, b33_3 | 37a. Har du noen av følgende yrker ? <br> (heltid eller deltid) Sett kryss for hvert spørsmål <br> Sjåfør <br> Bonde/gårdbruker <br> Fisker |
| b33_4, b33_5 | 37b. Hvilket yrke/tittel har eller hadde du på dette arbeidsstedet? <br> (spørsmålet henviser til et mellomliggende spørsmål (ikke CONOR)om den virksomhet man har arbeidet i lengst tid siste 12 mnd ) (For eksempel; sekretcr, lcerer, industriarbeider, barnepleier, møbelsnekker, avdelingsleder, selger sjåfør e.l) <br> Yrke. |
|  | SYKDOM OG SKADER |
| $\begin{aligned} & \text { b13_1, b13_2, b13_3 } \\ & \text { b13_4, b13_5, b13_6 } \\ & \text { b13_7, b13_8 } \end{aligned}$ | 38. Har du noen gang hatt: <br> Sett et kryss for hvert spørsmål. Oppgi også alder ved hendelsen. <br> Hvis det har skjedd flere ganger, hvor gammel var du siste gang. |
| $\begin{aligned} & \text { b14_1, b14_2, b14_3 } \\ & \text { b14_4, b14_5 } \end{aligned}$ | 39. Har du eller har du hatt? <br> Kryss av ja eller nei for hvert spørsmål <br> Høysnue <br> Kronisk bronkitt/emfysem <br> Benskjørhet (osteoporose) <br> Fibromyalgi/fibrositt/kronisk)smertesykdom <br> Psykiske plager som du har søkt hjelp for |
| b17 | 40. Hoster du omtrent daglig i perioder av året? Ja Nei |
| b18 | 41. Hvis ja: <br> Er hosten vanligvis ledsaget av oppspytt? <br> Ja Nei |
| b19 | 42. Har du hatt slik hoste så lenge som $i$ en 3 måneders periode $i$ begge de to siste år? <br> Ja <br> Nei |


| b20 | 43. Hvor ofte er du plaget av søvnløshet? <br> Aldri, eller noen få ganger i året <br> 1-2 ganger i måneden (før år 2000), 1-3 ganger i måneden (etter år 2000) <br> Omtrent 1 gang i uken <br> Mer enn 1 gang i uken |
| :---: | :---: |
| b21 | 44. Har du siste året vært plaget av søvnløshet som har gått utover arbeidsevnen? |
|  | BRUK AV MEDISINER |
| b16_1, b16_2 | 45. Bruker du? |
| b16_19 to b16_24 | 46a. Har du i løpet av det siste året brukt noen av følgende <br> midler daglig eller nesten daglig? <br> Angi hvor mange måneder du brukte dem. Sett 0 hvis du ikke har brukt noen av midlene. <br> Legemidler <br> Smertestillende $\qquad$ mnd. <br> Sovemedisin <br> .........mnd. <br> Beroligende midler <br> ........mnd. <br> Midler mot depresjon $\qquad$ mnd. <br> Allergimedisin <br> Astmamedisin $\qquad$ mnd. <br> Med medisiner mener vi her medisiner som er kjøpt på apotek. <br> Kosttilskudd og vitaminer regnes ikke med. |
| b16_3 to b16_8 | 46 b . Hvor ofte har du i løpet av de siste 4 ukene <br> brukt følgende medisiner? <br> (Sett ett kryss per linje) <br> Smertestillende uten resept <br> Smertestillende på resept <br> Sovemedisin <br> Beroligende medisin <br> Antidepressiva <br> Annen medisin på resept |
| b16_9_1 to b16_18_3 | 46c. Fyll inn navn på medisin, årsak til bruk og tiden den ble brukt fra sp 46b |


|  | $\begin{aligned} & 2 . \\ & 3 . \\ & 4 . \\ & 5 . \\ & 5 . \\ & 6 . \end{aligned}$ |
| :---: | :---: |
|  | KOSTTILSKUDD |
| b16_25 to b16_27 | 47 a . Har du i løpet av det siste året brukt noen av følgende midler <br> daglig eller nesten daglig? <br> Angi hvor mange måneder du brukte dem. Sett 0 hvis du ikke har brukt noen av midlene. Jerntabletter <br> .........mnd. <br> Vitamin D-tilskudd <br> .........mnd. <br> Andre vitamintilskudd <br> .........mnd. <br> Tran <br> .........mnd. |
| b16_28, b16_29 | 47 b. Bruker du følgende kosttilskudd? <br> Tran, trankapsler, <br> Fiskeoljekapsler <br> Vitamin- og/eller <br> mineraltilskudd |
|  | RESTEN AV SKJEMAET SKAL BARE BESVARES AV KVINNER |
| b22 | 48. Hvor gammel var du da du fikk menstruasjon første gang? ...........år |
| b23 | 49. Hvis du ikke lenger har menstruasjon, hvor gammel var du da den sluttet? ...........år |
| b24 | 50. Er du gravid nå?   <br> Ja Nei Usikker Over fruktbar alder |
| b25 | 51. Hvor mange barn har du født tidligere? $\square$ .barn |
| b26_1 to b26_12 | 52. Hvis du har født, fyll ut for hvert barn barnets fødselsår og omtrent antall måneder du ammet hvert barn. <br> Fødselsår <br> Antall måneder med amming <br> 1. <br> 2. <br> 3. <br> 4. <br> 5. <br> 6. |
| b27_1 to b27_4 | 53. Bruker du eller har du brukt: |







| Terraced/semi-detached house Other/institution/care home |
| :---: |
| 23. How large is your home? ..........m2 |
| 24. Do you have wall-to-wall carpets in the living-room? Yes No |
| 25. Is there a cat in your home? <br> Yes No |
| FAMILY AND FRIENDS |
| 26 a. With whom do you live? Tick one for each question and write the number |
| 26 b. Do you live with anyone? <br> Yes <br> No <br> If YES: <br> Spouse/Partner <br> Other persons older than 18 years <br> Persons younger than 18 years |
| 26 c (only at the questionary for the elderly) Where do you live? Please tick <br> Home <br> Institution <br> Do you live with? Yes No <br> Spouse/Partner? <br> Other persones? |
| 27. How many of the children attend day care/kindergarten/nursery school? ......... |
| 28. How many good friends do you have with whom you can talk confidentially and who can provide help if you need it? <br> (Do not count people you live with, but do include other relatives) $\qquad$ |
| 29. Do you feel that you have enough good friends? Yes |


| No |
| :--- |
| 30. How often do you usually take part in organised activities, e.g. |
| sewing circles, sports clubs, political meetings, religious or other organizations? |
| Never, or just a few times a year |
| 1-2 times a month (before year 1996), 1-3 times a month (after year 1996) |
| Approximately once a week |
| More than once a week |$|$| WORK |
| :--- |
| 31. What is your current work situation? <br> Paid work <br> Full-time housework <br> Under education, military service <br> Unemployed, on leave without payment |
| 32 a. How many hours of paid work do you have per week? |
| 3..............number of hours |
| 32 b. What is your current work situation - paid work? <br> Yes, full-time <br> Yes, part time <br> No |
| 33. Do you receive any of the following? <br> Sickness benefit? <br> Old-age pension? <br> Rehabilitation benefit? <br> Disability pension? <br> Unemploment benefits? <br> Social welfare benefits? <br> Social benefit-single parent? |
| 34. Do you work shifts or nights? <br> Yes <br> No |
| 35. If you have paid or unpaid work, which statement describes your work best? |
| Mostly sedentary work? |
| (e.g. office work, mounting) |
| Work that requires a lot of walking? |
| (e.g. shop assistant, light industrial work, teaching) |
| Work that requires a lot of walking and lifting? |
| (e.g. postman, nursing, construction) |
| Heavy manual labour? (e.g. forestry, heavy farmwork, heavy construction) |
| 36. Do you_decide yourself how your work will be done? (Tick one only) |


| Not at all <br> Very little <br> Yes, sometimes <br> Yes, my own decision |
| :---: |
| 37 a. Do you have any of the following occupations? (full time or part time) Tick one for each question Yes <br> No <br> Driver <br> Farmer <br> Fisherman |
| 37 b . What occupation/title did you have at this work? <br> (the question refers to another question (not CONOR) about the occupation where they worked the longest period during the past year) <br> Ex secretary, teacher, industrial worker, nursing, carpenter, l <br> eader, salesman, driver etc) <br> Occupation:. |
| YOUR OWN ILLNESS and INJURIES |
| 38. Have you ever had: <br> Tick one for each question. State age at event. If it has happened several times, write age at the last event. Yes No Age at last time <br> Hip fracture <br> Wrist/forearm fracture <br> Whiplash <br> Injury requiring hospital <br> admission |
| 39. Do you have or have you ever had? <br> Tick yes or no for each question <br> Hay fever <br> Chronic bronchitis/emphysema <br> Osteoporosis <br> Fibromyalgia/fibrositis/chronic pain syndrome <br> Psychological problems for which you have sought help |
| 40. Do you cough almost daily for some periods of the year? Yes No |
| 41. If yes, <br> do you bring up phlegm? <br> Yes No |
| 42. If you cough almost daily for some periods of the year, have you had this kind of cough for as long as 3 months in each of the last two years? <br> Yes No |




|  | Previously |  |
| :---: | :---: | :---: |
| Contraceptive pills (OC) (incl. minipill) <br> Contraceptive injections <br> Hormonal intrauterine device <br> Estrogen (tablets or patches) <br> Estrogen (cream or suppositories) |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
| 54. If you use contraceptive pills, hormona what brand do you currently use? | trauterine | ice, or es |

Appendix 2

## METHODS DESCRIPTION NORWEGIAN HELATH STUDIES

## Randi Selmer 30 Nov 2007. Updated 23 June 2008. Measurements in Health Surveys 1972-2003.

## Blood pressure

1. 1972-84: Systolic and diastolic blood pressure were measured twice with a standard mercury sphygmomanometer after 4 minutes rest. The second measurement has usually been used in follow up studies. The interval between first and second measurement was 1 minute. Diastolic blood pressure was recorded at the disappearance of the Korotkoff sounds (phase V). When phase V was absent, phase IV was used. Standard size cuffs were used throughout. The blood pressure was measured on the right upper arm with the person sitting on a chair.
2. 1985-2003: Pulse recordings, systolic and diastolic blood pressures were measured by an automatic device (DINAMAP, Criticon, Tampa, USA), which measured the blood pressure in mm Hg automatically by an oscillometric method. After 2 minutes preceding rest, three recordings were made at one-minute intervals. The values of the mean of the second and third systolic blood pressure measurements were used in calculating the cardiovascular risk score (CVD risk score). Arm circumference of right upper arm was measured 10 cm above fossa cubiti. From these measurements small, medium or large cuff was chosen. The blood pressure was measured on the right upper arm with the person sitting on a chair.
The two methods have been compared (PG Lund-Larsen: Blodtrykk målt med kvikksølvmanometer og med Dinamap under feltforhold- en sammenligning. Norsk epidemiologi 1997; 7 (2): 235-41)

## Serum analyses

Sera from the screenings were sent to the Department of Clinical Chemistry, Ullevål University Hospital, Oslo, Norway

## Serum lipids

Non-enzymatic methods: Total cholesterol and triglycerides
Non enzymatic methods were used in Oslo 1972-73, first screening in Finnmark, Oppland and Sogn og Fjordane 1974-78 and second screening in Finnmark 1977-78. Enzymatic methods were used from second screening in Sogn og Fjordane 1980.

Stensvold et al. BMJ 1993:
"A blood sample was taken from non-fasting subjects and analysed for serum concentrations of total cholesterol and triglycerides, both components being measured non-enzymatically on a Tchnicon AutoAnalyzer. On later comparison with enzymatic methods, the non-enzymatic methods used gave on average $10 \%$ higher triglyceride values and $8 \%$ higher cholesterol values. The participants reported the time since last meal."

The triglyceride values included in the data set are corrected values compatible with enzymatic methods according to the formula:
$($ New method $)=0.90 \times($ Old method $)-0.11$
The cholesterol values included in the data set are corrected values compatible with enzymatic methods according to the formula:
$($ New method $)=0.92 \times($ Old method $)+0.03$
The formula was evolved after extensive test program comparing new and old method.

## Enzymatic methods:

All measurements of HDL cholesterol were enzymatic. (Stensvold I, Urdal P, Thürmer H, Tverdal A, Lund-Larsen PG, Foss OP. High-density lipoprotein cholesterol and coronary, cardiovascular and all cause mortality among middle-aged Norwegian men and women.Eur Heart J. 1992 Sep;13(9):1155-63.)

Non-fasting serum total cholesterol, serum HDL cholesterol, glucose and serum triglycerides were measured directly by an enzymatic method (Technicon or Hitachi autoanalyzer). Seronorm Lipoprotein was used as internal quality control material for the lipid analyses and Autonorm Human Liquid for the glucose. The control material was done at the start and for every $30^{\text {th }}$ sample.

Stability of cholesterol measurements from 1972 has been documented ( OP Foss and P Urdal: Kolesterol gjennom mer enn 25 år: kan svarene sammenliknes over så lang tid? Norsk epidemiologi 2003; 13 (1): 85-88) )

## Glucose

Serum glucose was measured in first screening in Finnmark, Oppland and Sogn og Fjordane 1974-78 and second screening in Finnmark 1977-78 and in a sample in second screening in Oppland 1981-83 by a non enzymatic method by Brown ( ME Brown: Ultra-micro sugar determinations using 2, 9-dimethyl-1, 10-phenanthroline hydrochloride (Neocuproine). Diebetes 10:60, 1961.) The same method was used in Oslo 1972-73. The results obtained with this method were about $0.8-1.1 \mathrm{mmol} / \mathrm{l}$ higher than the true concentration defined as the value found with a specific enzymatic method.

From 1994 non fasting serum glucose was measured by enzymatic method described above. The old glucose values have not been adjusted to levels comparable with enzymatic methods.

## Weight and height

Body weight (in kilograms, one decimal) and height (in centimetres, one decimal) was measured according to standard protocol with the participants wearing light clothing without shoes (manually recorded until 2000 and after that with an electronic Height and Weight scale)

## Waist and hip

Waist and hip were measured from Finnmark and Akershus 1996/97 and onwards. Waist circumference was measured at the umbilicus to the nearest cm with the subject standing and breathing normally. In obese individuals, waist circumference was defined as the midpoint between the iliac crest and lower margin of ribs. Hip circumference was measured as the maximum circumference around the buttocks. Both waist and hip were measured with a measuring tape of steel - which was emphasized to be horizontal. Waist and hip circumference were used to calculate the waist-hip ratio using the formula waist (cm)/ hip circumference (cm).

Measurements of lipids in three counties 1974-1988

|  | Finnmark | Sogn og Fjordane | Oppland |
| :---: | :---: | :---: | :---: |
| Name |  |  |  |
| Screening 1 |  |  |  |
| u1kol_mg | total cholesterol $\mathrm{mg} / \mathrm{dl}$ old method | total cholesterol mg/dl old method | total cholesterol mg/dl old method |
| u1kolest | total cholesterol old method converted to $\mathrm{mmol} / \mathrm{l}$ by factor 0.02586 | total cholesterol old method converted to mmol/l by factor 0.02586 | total cholesterol old method converted to $\mathrm{mmol} / \mathrm{l}$ by factor 0.02586 |
| u1kolenz | total cholesterol mmol/l converted to enzymatic values from u1kolest by formulae | total cholesterol mmol/I converted to enzymatic values from u1kolest by formulae | total cholesterol mmol/I converted to enzymatic values from u1kolest by formulae |
| No HDL measurements |  |  |  |
| u1trigly | triglycerides mmol/l old method | triglycerides mmol/l old method | triglycerides mmol/l old method |
| u1trienz | triglycerides mmol/l converted to enzymatic values from u1trigly by formulae | triglycerides mmol// converted to enzymatic values from u1trigly by formulae | triglycerides mmol/l converted to enzymatic values from u1trigly by formulae |
| Screening 2 |  |  |  |
| u2kol_mg | total cholesterol $\mathrm{mg} / \mathrm{dl}$ old method | total cholesterol mg/dl enzymatic method | total cholesterol mg/dl enzymatic method |
| u2kolest | total cholesterol old method converted to $\mathrm{mmol} / \mathrm{l}$ by factor $0.02586$ | total cholesterol enzymaticmethod converted to mmol/l by factor 0.02586 | total cholesterol enzymaticmethod converted to $\mathrm{mmol} / \mathrm{l}$ by factor 0.02586 |
| u2kolenz | total cholesterol mmol/l converted to enzymatic values from u2kolest by formulae | u2kolenz=u2kolest | u2kolenz=u2kolest |
| u2hdlkol | mg/dl, enzymatic* | mg/dl, enzymatic* | mg/dl, enzymatic* |
| u2hdlkl | converted to $\mathrm{mmol} / \mathrm{l}$ by factor 0.02586 | converted to mmol// by factor 0.02586 | converted to $\mathrm{mmol} / \mathrm{l}$ by factor 0.02586 |
| u2trigly | triglycerides mmol/l old method | triglycerides mmol/l enzymaticmethod | triglycerides mmol/l enzymatic method |
| u2trienz | triglycerides mmol/l converted to enzymatic values from u1trigly by formulae | u2trienz=u2trigly | u2trienz=u2trigly |
| Screening 3 |  |  |  |
| u3kolest/u3kolenz | All values enzymatic mmol/l. Sometimes renamed u3kolest to u3kolenz to indicate that these are enzymatic values. |  |  |
| u3hdlkl | Nomeasurements | All values enzymatic mmol/I* | All values enzymatic mmol//* |
| u3trigly/u3trienz | All values enzymatic mmol/I. Sometimes renamed u3trigly to u3trienz to indicate that these are enzymatic values. |  |  |
|  |  |  |  |

*Eur Heart J. 1992 Sep; 13(9):1155-63.
High-density lipoprotein cholesterol and coronary, cardiovascular and all-cause mortality among middleaged Norwegian men and women. Stensvold I, Urdal P, Thürmer H, Tverdal A, Lund-Larsen PG, Foss OP.

## SUMMARY THREE COUNTIES STUDY

The cardiovascular surveys in Finnmark, Sogn og Fjordane and Oppland 1974-78, 1977-83 and 1985-88. Sources: Final reports from each survey in each county

| County | Period | Age groups invited | Number invited | Number attending | \% attendance, fully invited ages |
| :---: | :---: | :---: | :---: | :---: | :---: |
| Finnmark | 1974-75 | All residents in age 35-49 by Dec 1974 (born 25-39). Age 20-34: 10\% random samples | 17401 | 14340 | 82.4 <br> Men: 78.8, women: 86.2 |
|  | 1977-78 | All residents born 1925-42, samples in youngerages from 20 years. | 20647 | 17145 | 83.0 <br> Men: 79.2 <br> women: 87.3 |
|  | 1987-88 | All residents in age 40-62 by Dec 1987 (born 1925-47) + those aged 30-39 and invited in 1977-78 + $10 \%$ of non-invited in age 20-39. All residents 18 years or older in Bugøynes. | 22994 | 17852 | 77.6 <br> Men: 73.4, <br> women: 82.6 |
| Sogn og Fjordane | 1975-76 | All residents in age 35-49 by Dec 1975 (born 1926-40) $+10 \%$ random sample in age 20-39. | 16603 | 14966 | 90.1 <br> Men: 87.4, women:93.1 |
|  | 1980-81 | All residents born 1926-40 + samples in youngerages from 17 years. | 19506 | 17473 | 89.6 <br> Men: 86.8, women:92.6 |
|  | 1985-86 | All residents in age 40-54 by Dec 311985 (born 1931-45) + those younger than 40 years and invited in 1980-81 $+5-\%$ sample of those in age 20-39 not invited in 1980-81 +10 \% sample of invited in 198081 in age 55-59. A few older subjects in a hypertension register. | 21423 | 18669 | 87.1 <br> Men: 83.9, <br> women: 90.7 |
| Oppland | 1976-78 | All in age 35-49 by Dec 1976 (born 192741) $+10-\%$ random sample in age 20-39. | 31620 | 28399 | 89.8 <br> Men: 87.8, <br> women: 91.8 |
|  | 1981-83 | All residents born 1927-41 + samples in younger ages from 20 years. | 31581 | 28437 | 90.0 <br> Men: 88.1, <br> women: 91.9 |
|  | 1986-88 | All residents aged 40-54 on Dec 1986 (born 1932-46) + all residents below 40 years and a $10 \%$ sample in age $55-$ 59 if invited in 1981-83 + 5-\% of not invited in 1981-83 in age 20-39. A few older subjects in a hypertension register. | 37270 | 32124 | 86.2 <br> Men: 83.5, <br> women: 88.9 |

## CONOR STUDY <br> MATERIALS AND METHODS <br> DESCPRIPTION

# Cohort Norway (CONOR): Materials and methods 

Anne Johanne Søgaard, Norwegian Institute of Public Health, April 2006
CONOR (COhort NORway) is a large collaborative project between epidemiological centres at the University of Tromsø, the Norwegian University of Science and Technology in Trondheim, the University of Bergen, the University of Oslo, and the Norwegian Institute of Public Health.

## Data from 10 regional studies

In CONOR, regional data from 10 different epidemiological studies have been merged into a national database, which is more representative of the Norwegian population than each of the individual sites.

The database consists of information obtained from questionnaires, a simple physical examination, analyses of blood samples, and frozen stored blood and/or DNA. The main purpose of CONOR is to study the aetiology of rare diseases by testing environmental, inheritable, cultural and social factors in order to describe the dispersion of diseases and risk factors by time, place and socio-demographic factors.

CONOR is particularly suitable for studying gene-environment interactions and for linkages to various national registers (eg. cancer-, cause of death-, hospital- and medical birth registers).

## Invitation and procedures

Altogether 309,832 individuals were invited in the 10 studies based on addresses from the Population registry of Norway (Hammer, 2002). Some of the individual studies invited all subjects above a specific age (for example all above 19 years in HUNT II), whereas others invited all subjects in selected age groups (for example all 30-, 40-, 45-, 60 and 75 years in OPPHED and TROFINN). The web site for each study contains more detailed information (see Table 1).

In all CONOR surveys, the data collection followed a standard procedure. Letters of invitation were mailed about 2 weeks before the time of appointment and included a
questionnaire and a booklet with the aims of the study and information about the examinations and procedures. At the screening, the main questionnaire was collected from the attendees, they went through a physical examination and a non-fasting blood sample was drawn for analyses in fresh serum. Another sample was stored at minus 80 degrees. In most studies, the participants were given one or two supplementary questionnaires, which they were instructed to fill in at home and to return by mail in pre-addressed envelopes.

About four weeks after attending the examination, a letter with some results from the examination and blood tests was sent to all participants. Those with the highest scores of cardiovascular risk were offered a new clinical examination at the regional University Hospital - or, in some of the studies, were asked to visit their own general practitioner.

## Measures

All surveys have been carried out in collaboration with the National Health Screening Service, Oslo (now Norwegian Institute of Public Health). Experienced and trained personnel conducted all procedures. Non-fasting serum total and HDL cholesterol, glucose and triglycerides were measured directly by an enzymatic method (Boehringer 148393, Boehringer-Mannheim, Federal Republic of Germany - from 2000 Hitachi 917 auto analyzer, Roche Diagnostic, Switzerland).

The Department of Clinical Chemistry, Ullevål University Hospital, Oslo, performed all laboratory assessments except for HUNT II where the analyses were performed at the Department of Clinical Chemistry, Innherad Hospital, Levanger. Comparisons of blood-samples were performed between the laboratories, and small differences were found (Tverdal A et al 1997). Calibration procedures were carried out between these laboratories in connection with the surveys (Dr. Lund-Larsen PG, National Health Screening Service, personal communication). An acceptable stability of the laboratory analyses over time in the population surveys has been reported (Foss \& Urdal, 2003).

Heart rate, systolic and diastolic blood pressures were measured by an automatic device (DINAMAP, Criticon, Tampa, USA), which measured the blood pressure in
mm Hg automatically by an oscillometric method. After 2 minutes of preceding rest, three recordings were made at one-minute intervals. Mean values of the second and third systolic blood pressure measurements were used in calculating the cardiovascular risk score (CVD risk score) (Tverdal et al., 1989). The stability of the blood-pressure measures have been evaluated and deemed acceptable (Lund-Larsen, 1997).

Body weight (in kilograms, one decimal) and height (in cm, one decimal) was measured according to a standard protocol with the participants wearing light clothing without shoes (manually recorded until 2000 and after that with an electronic Height and Weight Scale). Body mass index (BMI) was calculated as $\mathrm{kg} / \mathrm{m}^{2}$. Waist circumference was measured at the umbilicus to the nearest cm and with the subject standing and breathing normally. In obese individuals, waist circumference was defined as the midpoint between the iliac crest and lower margin of ribs. Hip circumference was measured as the maximum circumference around the buttocks. Both waist and hip were measured with a measuring tape of steel - which was emphasized to be horizontal. Waist and hip circumference were used to calculate the waist-hip ratio using the formula waist (cm)/ hip circumference (cm).

Most of the studies consist of a central core and several supplementary projects - for example extra samples of blood, ECG, ultrasonographic examination of carotid artery and abdominal aorta, and bone mineral densitometry (BMD). The web site for each study contains more detailed information (see Table 1). Only a limited and mutual core of each study constitutes CONOR. Most of the studies have published reference papers with more detailed information about their own study (Table 2).

## The CONOR-questions

All surveys used 50 common CONOR-questions agreed upon before the first CONOR survey in Tromsø in 1994. The exact wording of the questions is available at the CONOR web site (http://www.fhi.no/dav/CA11310499.doc). Some of these questions were placed on the second questionnaire handed out at the screening station - and thus have lower response rate.

The CONOR-questions cover the following main topics: Self-reported health and diseases such as diabetes, asthma, coronary heart disease, stroke and mental distress, musculo-skeletal pains, family history of disease, risk factors and lifestyle, environment while growing up, social network and social support, education, work and housing, some types of occupation, use of medications and reproductive history (women).

Several of these questions have been evaluated or validated previously and were deemed acceptable (Tretli et al., 1982; Jacobsen \& Thelle, 1987; Løchen \& Rasmussen, 1992; Thune et al., 1997, Joakimsen et al., 1998; Saltin \& Grimsby, 1968; Derogatis et al., 1974; Ainsworth et al., 1996; Brugha et al., 1985; Strand et al., 2003; Søgaard et al 2003). The Population registry of Norway, which was used for invitation, contains information about gender, birth date, marital status, address and country of birth.

## Participation in the CONOR studies

Altogether 181,891 subjects accepted to participate and provided a declaration of consent $-7,460$ of these participated in more than one survey. The age distributing of these 174430 participants is shown in table 3 . The participation rate varied among the surveys. The participation was slightly reduced throughout the study-period 19942003 - and was higher in rural as compared to urban areas.

## Ethics and approvals

All participants of the studies included in CONOR, have given their written consent. The participant's names and personal ID numbers are omitted when data are used for research purposes. The Norwegian Data Inspectorate has approved - and the Regional Committees for Medical Research Ethics has evaluated each individual study. The studies have been conducted in full accordance with the World Medical Association Declaration of Helsinki.

## References

Ainsworth BE, Montoye HJ, Leon AS. Methods of assessing physical activity during leisure and work. In: Bouchard C, Shephard RJ, Stephens T (red.) Physical activity, fitness and health. Champaign, IL: Human Kinetics, 1994: 146-59.

Brugha, T, Bebbington, P, Tennant, C, Hurry, J. The list of threatening experiences: a subset of 12 life event categories with considerable long-term contextual threat. Psychol Med, 1985;15:189-194.

Derogatis LR, Lipman RS, Rickels K, Uhlenhuth EH, Covi L. The Hopkins Symptom Checklist (HSCL): a self-report symptom inventory. Behav Sci 1974;1:1-15.

Engeland A, Søgaard AJ. CONOR (COhort NORway) - en oversikt over en unik forskningsdatabank. Norsk Epidemiologi 2003; 13 (1): 73-77 73.

Jacobsen BK, Thelle DS. The Tromso Heart Study: food habits, serum total cholesterol, HDL cholesterol, and triglycerides. Am J Epidemiol 1987; 125:622-630.

Foss O, Urdal P. Cholesterol for more than 25 years: Could the results be compared throughout all this time [Kolesterol gjennom mer enn 25 år: Kan svarene sammenliknes over så lang tid?] Nor J Epidemiol 2003;13:85-8.

Hammer H . The central population registry in medical research [Article in Norwegian]. Det sentrale folkeregisteret i medisinsk forskning. Tidskr Nor Laegeforen 2002;122:2550.

Joakimsen RM, Fønnebø V, Magnus JH, Størmer J, Tollan A, Søgaard AJ. The Tromsø study: Physical activity and the incidence of fractures in a middle-aged population. J Bone Min Res 1998; 13: 1149-57.

Lund-Larsen PG. Blood pressure measured with sphygmomanometer and with Dinamap under field conditions - a comparison [Article in Norwegian]. Blodtrykk målt med kvikksølvmanometer og Dinamap under feltforhold - en sammenligning. Nor J Epidemiol 1997;7:235-41.

Løchen MJ, Rasmussen K. The Tromsø study: physical fitness, self reported physical activity, and their relationship to other coronary risk factors. J Epidemiol Com Health 1992; 46: 103-7.

Magnus, P, Arnesen E, Holmen J, Stoltenberg C, Søgaard AJ, Tell G. CONORCOhort NORway: historie, formål og potensiale. Norsk Epidemiologi 2003; 13 (1): 79-82 79.

Saltin B, Grimsby G. Physiological analysis of middle-aged and old former athletes. Circulation 1968; 38: 1104-15.

Strand BH, Dalgard OS, Tambs K, Rognerud M. Measuring the mental Health Status of the Norwegian Population: a comparison of the instruments SCL-25 SCL-10, SCL5 and MHI-5 (SF-36). Nord J Pschychiatry 2003;57:113-8.

Søgaard AJ, Bjelland I,Tell, GS, Røysamb E. A comparison of the CONOR Mental Health. Index to the HSCL-10 and HADS. Measuring mental health status in The Oslo Health Study and the Nord-Trøndelag Health Study. Nor J Epidemiol 2003; 13 (2): 279-284 279.

Thune I, Brenn T, Lund E, Gaard M. Physical activity and the risk of breast cancer. N Engl J Med 1997; 336: 1269-75.

Tretli S, Lund-Larsen PG, Foss OP. Reliability of questionnaire information on cardiovascular disease and diabetes: cardiovascular disease study in Finnmark county. J Epidemiol Community Health 1982; 36:269-273.

Tverdal A, Foss OP, Leren P, Holme I, Lund-Larsen PG, Bjartveit K. Serum triglycerides as an independent risk factor for death from coronary heart disease in middle-aged Norwegian men. Am J Epidemiol 1989;129:458-65.

Tverdal A, Åsberg A, Bønaa KH, Stensvold I, Grudt G, Ingebretsen OC, Urdal P. Sammenligning av to blodprøver fra samme person analysert ved to laboratorier. Norsk Epidemiologi 1997; 7 (2): 283-286.

TABLE 1. Number of invited and participating subjects in Cohort Norway (CONOR) 1994-2003.

|  |  |  |  | Number of participants |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Name of the study | Year of survey | Number invited ${ }^{\dagger}$ | Invited agegroups in years ${ }^{\ddagger}$ | Men | Women | Total | Web address |
| Tromsø IV (The fourth Tromsø Study) | 1994-1995 | 37,558 | 25 + | 12,797 | 14,128 | 26,925 | http://uit.no/tromsoundersokels en/tromso4/2 |
| HUNT II (The second North-Trøndelag Health Study) | 1995-1997 | 94,196 | $20+$ | 30,442 | 34,576 | 65,018 | http://www.hunt.ntnu.no/ |
| HUSK (The Hordaland Health Study) | 1997-1999 | 38,587 | $\begin{gathered} 40-44,46-47,70- \\ 72 \end{gathered}$ | 11,678 | 13,852 | 25,530 | http://www.uib.no/isf/husk/ |
| Oslo II (The second Oslo Study) | 2000 | $14,209{ }^{\text {§ }}$ | 48-77 | 6,919 |  | 6,919 | http://www.fhi.no/artikler/?id=54 $\underline{685}$ |
| HUBRO (The Oslo Health Study) | 2000-2001 | 58,660 ${ }^{\text {\# }}$ | $\begin{gathered} 30,31,40,45, \\ 46,59 / 60, \\ 75 / 76 \end{gathered}$ | 9,751 | 12,264 | 22,015 | $\begin{aligned} & \text { http://www.fhi.no/artikler/?id=5 } \\ & \hline 4464 \end{aligned}$ |
| OPPHED (The Oppland and Hedmark Health Study) | 2000-2001 | 22,327 | 30, 40, 45, 60, 75 | 5,650 | 6,752 | 12,402 | $\frac{\text { http://www.fhi.no/artikler/?id=2 }}{8233}$ |
| Tromsø V (The fifth Tromsø Study) | 2001 | 10,353 | $30+$ | 3,491 | 4,586 | 8,077** | http://uit.no/tromsoundersokels en/tromso5/2 |
| I-HUBRO (The Oslo Immigrant Health Study) | 2002 | $12,088^{\text {t }}$ | 20-60 | 1,915 | 1,768 | 3,683 | $\frac{\text { http://www.fhi.no/artikler/?id=2 }}{\underline{8217}}$ |
| TROFINN (The Troms and Finnmark Health Study) ${ }^{\text {\# }}$ | 2002 | 16,229 | 30-77 | 4,318 | 5,009 | 9,327 | $\frac{\text { http://www.fhi.no/artikler/?id=2 }}{\underline{8261}}$ |
| MoRo II (The second part of the Romsås in Motion Study) | 2003 | 5,535 | 34-70 | 899 | 1,096 | 1,995 | $\begin{aligned} & \text { http://www.fhi.no/artikler/?id=2 } \\ & 8254 \end{aligned}$ |
| CONOR (Cohort Norway) | 1994-2003 | 309,742 | 20-103 | 87,157 | 92,928 | 181,891 ${ }^{*}$ | $\underline{\underline{\text { http://www.fhi.no/artikler/?id=2 }}}$ |

[^0]I) born 1933-1969 - except those participating in HUBRO; TROFINN: All 30, 40, 45, 60, 75 years and all those participating in three Finnmark studies in the period 1974-1988 which included: All born 1925-1947, all born 1948-1968 invited to Finnmark I, II or III
§ 2,515 more men who belonged to the Oslo II cohort, also belonged to the HUBRO cohort, and were only invited to HUBRO. Of these 1,320 men participated. They are only counted as invited to HUBRO. 50 more men belonged to the MoRo-cohort, and are only counted as invited there.
\# Include 17,308 invitees ( 31 and 46 years - additional cohorts) who were not reminded. The attendance-rate of these was low.
** 7,166 of these participated also in Tromsø IV.
$\dagger$ Include 4,116 persons (20-30 years - additional cohort) who were not reminded. The attendance-rate of these was very low.
\# Include 18 of 25 municipalities in Troms and 10 of 19 municipalities in Finnmark. The other municipalities participated in Troms $ø \mathrm{~V}$ and in SAMINOR, i.e. a health survey in communities with Sámi and Norwegian population, at the same time.

Table 2. Reference papers to the 10 participating CONOR studies.

Tromsø IV: Wilsgard T. Longitudinal analyses of cardiovascular risk factors. The Tromsø study 1974-1995. ISM skriftserie nr. 65. Tromsø, Norway: Institute of Community Medicine, University of Tromsø, 2002.

HUNT II: Holmen J, Midthjell K, Krüger Ø, Langhammer A, Lingaas Holmen T, Bratberg GH, Vatten L, Lund-Larsen PG. The Nord-Trøndelag Health Study 1995-97 (HUNT 2): Objectives, contents, methods and participation. Nor J Epidemiol 2003; 13: 19-32.

HUSK: Bjelland I, Tell GS, Vollset SE, Refsum H, Ueland PM. Folate, vitamin B12, homocysteine, and the MTHFR 677C->T polymorphism in anxiety and depression: the Hordaland Homocysteine Study. Arch Gen Psychiatry 2003 Jun;60(6):618-26 - and Sanne B, Mykletun A, Dahl AA, Moen BE, Tell GS; Hordaland Health Study. Occupational differences in levels of anxiety and depression: the Hordaland Health Study. J Occup Environ Med 2003;45:628-38.
Oslo II: Lund Håheim L, Holme I, Hjermann I, Søgaard AJ, Lund-Larsen PG, Leren P. Resultater fra Oslo-undersøkelser blant de samme menn i 1972/3 og i år 2000. Endring i risikofaktorer for hjerte- og karsykdom. Tidskr Nor Laegefor (Cond accepted)

HUBRO: Søgaard AJ, Selmer R, Bjertness E, Thelle D. The Oslo Health Study. The impact of self-selection in a large, population-based survey. Int J Equity Health 2004:3: 1-24. Online: http://www.equityhealthj.com/content/3/1/3

OPPHED: Only web-site - http//www.fhi.no/artikler/? id=28233
Tromsø V: Johnsen SH, Fosse E, Joakimsen O, Mathiesen EB, Stensland-Bugge E, Njølstad I, Arnesen E. Monocyte count is a predictor of novel plaque formation: a 7 -year follow-up study of 2610 persons without carotid plaque at baseline the Tromso Study. Stroke. 2005;36(4):715-9.

I-HUBRO: Holvik K, Meyer HE, Haug E, Brunvand L.Prevalence and predictors of vitamin D deficiency in five immigrant groups living in Oslo, Norway: the Oslo Immigrant Health Study. Eur J Clin Nutr. 2005;59:57-63.

TROFINN: Only web-site - http///www.fhi.no/artikler/?id=28260

MoRo II: Jenum AK,. Anderssen SA, Birkeland KI, Holme I, Graff-Iversen S, Lorentzen C, Ommundsen Y, Raastad T, Ødegaard AK, Bahr R. Promoting physical activity in a low-income multi-ethnic district: behavioural, psychological and biological effects of a pseudo-experimental community intervention study to reduce risk factors for diabetes and cardiovascular disease (submitted)
CONOR: Engeland A, Søgaard AJ. CONOR (Cohort NORway) - en oversikt over en unik forskningsdatabank. Nor J Epidemiol 2003;13:73-7 - and Magnus P, Arnesen E, Holmen J, Stoltenberg C, Søgaard AJ, Tell GS. CONOR (Cohort NORway): historie, formål og potensiale. Nor J Epidemiol 2003;13:79-82.

Table 3 Number of participants in Cohort Norway (1994-2003) according to gender and age-groups (at the time they attended the screening station). If participating in more than one study, only the last one is counted.

| Age | Men | Women | Total |
| :---: | :---: | :---: | :---: |
|  | N | N | N |
| <20 | 116 | 148 | 264 |
| 20-29 | 5884 | 7236 | 13120 |
| 30-39 | 13322 | 15547 | 28869 |
| 40-49 | 27969 | 32148 | 60117 |
| 50-59 | 10517 | 10176 | 20693 |
| 60-69 | 12229 | 10373 | 22602 |
| 70-79 | 13119 | 11883 | 25002 |
| 80+ | 1460 | 2303 | 3763 |
| Total | 84616 | 89814 | 174430 |

## COHORT PROFILE

# Cohort Profile: Cohort of Norway (CONOR) 

Øyvind Næss, ${ }^{1,2,3 *}$ Anne Johanne Søgaard, ${ }^{1,2}$ Egil Arnesen, ${ }^{4}$ Anne Cathrine Beckstrøm, ${ }^{2}$ Espen Bjertness, ${ }^{2}$ Anders Engeland, ${ }^{5}$ Peter F Hjort, ${ }^{2}$ Jostein Holmen, ${ }^{6,7}$ Per Magnus, ${ }^{1,2}$ Inger Njølstad, ${ }^{4}$ Grethe S Tell, ${ }^{5}$ Lars Vatten, ${ }^{7}$ Stein Emil Vollset ${ }^{5}$ and Geir Aamodt ${ }^{1}$

## Accepted 1 October 2007

## How did the study come about?

A number of large population-based cardiovascular surveys have been conducted in Norway since the beginning of the 1970s. The surveys were carried out by the National Health Screening Service in cooperation with the universities and local health authorities. All surveys comprised a common set of questions, standardized anthropometric and blood pressure measurements and non-fasting blood samples that were analysed for serum lipids at the Ullevål Hospital Laboratory. These surveys provided considerable experience in conducting large-scale population-based surveys, thus an important background for the Cohort of Norway (CONOR). In the late 1980s the Research Council of Norway established a programme in epidemiology. This also gave stimulus to the idea of establishing a cohort including both core survey data and stored blood samples. In the early 1990s, all universities, the National Health Screening Service, The National Institute of Public Health and the Cancer Registry discussed the possibility of a national representative cohort. ${ }^{1}$ The issue of storing blood samples for future analyses raised some concern and it was discussed in the parliament. In 1994, the Ministry of Health appointed the Steering Committee for the CONOR collaboration. In 1994-95, the fourth round of the Tromsø Study was conducted, and became the first survey to provide data and blood samples for CONOR. During the years 1994-2003, a number of health

[^1]surveys that were carried out in other counties and cities also provided similar data for the network. So far, 10 different surveys have provided data and blood samples for CONOR (Figure 1). The administrative responsibility for CONOR was given to the Norwegian Institute of Public Health (NIPH) in 2002. The CONOR collaboration is currently a research collaboration between the NIPH and the Universities of Bergen, Oslo, Tromsø and Trondheim.

## The purpose of CONOR

The CONOR cohort has not been established on the basis of any single hypothesis but is rather a multipurpose study. The ambition was to set up a sufficiently large enough cohort to study aetiological factors for a wide range of diseases. Additionally, this cohort should make it possible to describe Norwegian men and women in terms of distribution of exposures and health status according to time, place and socio-economic factors.
In 2002, CONOR and the Norwegian Mother and Child study (MoBa), ${ }^{2}$ received a 5 -year grant from the Norwegian Research Council to build a technology platform under the Functional Genomics programme (FUGE), called the Biobanks for Health in Norway (Biohealth) platform. ${ }^{3}$ The overall aim was to investigate separate and combined effects of genes and environment on the risk of disease.

## Who is in the sample?

Altogether 309742 individuals were invited to the 10 surveys based on the 11-digit personal identifier and addresses from the Population Registry of Norway. ${ }^{4}$ The goal is to include 200000 participants. We defined those who attended the survey and/or answered at least one questionnaire and signed a written informed consent as participants. The numbers in Table 1 include individuals who participated and had given their written consent for research and linkage to health registries. A total of 7309 persons participated in two CONOR surveys, and one person participated in three. Thus, the total number of


Figure 1 Map of Norwegian counties with location of each sub-study included in cohort of Norway (CONOR)
individuals in the CONOR cohort is 173236 . The distribution of age at the first examination and the number of deaths during follow-up through 2003 is given in Table 2. The individual surveys may have published papers with slightly different total numbers. Sampling procedures differed somewhat between the individual studies. The web site for each study contains more detailed information (Table l).

## What has been measured?

In all the CONOR surveys, the data collection followed a standard procedure. Letters of invitation were mailed about 2 weeks before the time of appointment and included a questionnaire and a brochure with the aims of the study and information about the examinations and procedures. At the screening, this initial questionnaire was collected from the attendees, participants underwent a physical examination and a non-fasting blood sample was drawn. In most studies, the participants were given one or two supplementary questionnaires, which they were instructed to fill in at home and return by mail in pre-addressed stamped envelopes.
About 4 weeks after attending the examination, a letter with selected results from the examination and blood tests was sent to all participants. Those with the highest scores of cardiovascular risk (a modified Framingham risk score based on multiplying the relative risks attributable to the subject's gender, serum cholesterol, systolic blood pressure the number of cigarettes currently smoked per day and family history of

Table 1 Number of invited and participating subjects in cohort of Norway (CONOR) 1994-2003

| Name of the study | Year of survey | Number invited | Invited age-groups in years | Number of participants ${ }^{\text {a }}$ |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  |  | Men | Women | Total | Web address |
| Tromsø IV (The fourth Tromsø Study) | 1994-1995 | 37558 | 25+ | 12797 | 14128 | 26925 | http://uit.no/tromsoundersokelsen/tromso4/2 |
| HUNT II (The second North-Trøndelag Study) | 1995-1997 | 94196 | 20+ | 30441 | 34576 | 65017 | http://www.hunt.ntnu.no/ |
| HUSK (The Hordaland Health Study) | 1997-1999 | 38587 | $\begin{array}{r} 40-44,46-47, \\ 70-72 \end{array}$ | 11678 | 13851 | 25529 | http://www.uib.no/isf/husk/ |
| Oslo II (The second Oslo Study) | 2000 | 14209 | 48-77 | 6919 |  | 6919 | http://www.fhi.no/artikler/?id=54685 |
| HUBRO (The Oslo Health Study) | 2000-2001 | 58660 | $\begin{array}{r} 30,31,40,45 \\ 46,59 / 60 \\ 75 / 76 \end{array}$ | 9509 | 11852 | 21361 | http://www.fhi.no/artikler/?id=54464 |
| OPPHED (The Oppland and Hedmark Health Study) | 2000-2001 | 22327 | $\begin{array}{r} 30,40,45, \\ 60,75 \end{array}$ | 5602 | 6661 | 12263 | http://www.fhi.no/artikler/?id=28233 |
| Tromsø V (The fifth Tromsø Study) | 2001 | 10353 | $30+$ | 3440 | 4457 | 7897 | http://uit.no/tromsoundersokelsen/tromso5/2 |
| I-HUBRO (The Oslo Immigrant Health Study) | 2002 | 12088 | 20-60 | 1877 | 1737 | 3614 | http://www.fhi.no/artikler/?id=28217 |
| TROFINN (The Troms and Finnmark Health Study) | 2002 | 16229 | 30-77 | 4196 | 4836 | 9032 | http://www.fhi.no/artikler/?id=28261 |
| MoRo II (The second part of the Romsås in Motion Study) | 2003 | 5535 | 34-70 | 896 | 1093 | 1989 | http://www.fhi.no/artikler/?id=28254 |
| CONOR (Cohort Norway) ${ }^{\text {a }}$ | 1994-2003 | 309742 | 20-103 |  |  |  |  |
| Sum of participants |  |  |  | 87355 | 93191 | 180546 | http://www.fhi.no/artikler/?id=28138 |
| Sum of individuals |  |  |  | 84153 | 89083 | 173236 |  |

[^2]coronary heart disease) were advised to visit their own general practitioner, and in some cases offered a follow-up examination at the local hospital. ${ }^{5}$

## Measures

Only a restricted core set of measurements and questionnaire responses constitute the CONOR data. Most individual studies that contribute to CONOR have more detailed measurements and questionnaire data. In the following section we describe the key core measurements that all studies contribute to CONOR; at the end we briefly describe some of the additional measurements that are in some of the contributing individual studies. All surveys were carried out in collaboration with the National Health Screening Service, Oslo (now the NIPH). Experienced and trained personnel conducted all procedures. Non-fasting serum totaland HDL-cholesterol, glucose and triglycerides were measured directly by an enzymatic method (Boehringer 148393, BoehringerMannheim, Federal Republic of Germany—from 2000 Hitachi 917 auto analyzer, Roche Diagnostic, Switzerland).

The Department of Clinical Chemistry, Ullevål University Hospital, Oslo, performed all laboratory assessments except for HUNT II (The second North-Trøndelag Study) where the analyses were performed at the Department of Clinical Chemistry, Levanger Hospital, Levanger. In Tromsø IV and V, cholesterol and triglycerides were measured at the Department of Clinical Chemistry, University Hospital North-Norway, Tromsø. Calibration procedures were carried out between these laboratories in connection with the surveys (Dr P.G. Lund-Larsen, National Health Screening Service, personal communication). An acceptable stability of the laboratory analyses over time in the population surveys has been reported. ${ }^{6}$

Heart rate, systolic and diastolic blood pressures were measured by an automatic device (DINAMAP, Criticon, Tampa, FL,USA). After 2 min of seated resting, three recordings were made at l-min intervals. Mean values of the second and third systolic blood pressure measurements were used in calculating the cardiovascular risk score (CVD risk score) (Tverdal, 1989 5/id). The stability of the blood pressure measures has been evaluated and deemed acceptable. ${ }^{7}$

Body weight (in kilograms, one decimal) and height (in centimetres, one decimal) was measured according to a standard protocol with the participants wearing light clothing without shoes (manually recorded until 2000 and after that with an electronic Height and Weight Scale). Body mass index (BMI) was calculated as kilograms per square metre. Waist circumference was measured at the umbilicus to the nearest centimetre and with the subject standing and breathing normally. In obese individuals, waist circumference was defined as the midpoint between the iliac crest and lower margin of ribs. Hip circumference was measured as the maximum circumference around the buttocks. Both waist and hip were measured with a measuring tape of steel-which was emphasized to be placed horizontally. The waist-hip circumferences were used to calculate the waist-hip ratio.

Most individual studies that contribute to CONOR have several additional measurements-for example, extra samples of blood, ECG and ultrasonographic examination of carotid artery and abdominal aorta. Four of the study sites measured bone mineral density (DEXA and/or SXA) and have established a research group called Norwegian Epidemiologic Osteoporosis Studies (NOREPOS). ${ }^{8}$ Altogether, around 28000 individuals
have had their bone mineral density measured and currently a number of collaborative studies are carried out.

## The CONOR questions

All surveys used about 50 core CONOR questions agreed upon before the first CONOR survey in Tromsø in 1994. The exact wording of the questions is available at the CONOR website (http://www.fhi.no/dav/CAl1310499.doc). Some questions have been slightly modified over the years.
The CONOR questions cover the following main topics: selfreported health and diseases such as diabetes, asthma, coronary heart disease, stroke and mental distress, musculo-skeletal pains, family history of disease, risk factors and lifestyle, social network and social support, education, work and housing, some types of occupation, use of medications and reproductive history (women).

Several of the questions have been evaluated or validated and deemed acceptable. ${ }^{9-18}$ The Population Registry of Norway that was used to identify eligible subjects, contains information about gender, date of birth, marital status, address and country of birth.

## Blood samples

Blood samples were drawn from the CONOR participants. EDTA blood for CONOR and the other sub-surveys have normally been collected in 7 or 5 ml vacutainers. These vacutainers were made by different manufacturers but were normally made of polypropylene. DNA has been extracted from more than 90000 specimens to medio 2007, and Biohealth intends to extract DNA from all samples by Spring 2008. The extracted DNA and an additional sample of 1.25 ml EDTA-blood will be stored at a national biobank storage site at HUNT/NTNU biobank in Levanger (Mid-Norway).

## What has been found?

Although a number of analyses from each participating study have been conducted, the CONOR file has only recently been compiled and made available for research. The first CONOR project was anchored in NOREPOS describing urban-rural differences in forearm fractures. ${ }^{19}$ Other methodological and validation studies have been completed as described above.

## What are the main strengths and weaknesses?

The CONOR database has several strengths: it is population based including populations from various parts of Norway, both rural and urban. The 11-digit personal identification number makes it possible to link cohort participants to national health registries. At present, several large linkages to other registers have been or are in the process of being conducted. These include linkages with census-based data for the whole population and the Medical Birth Registry of Norway, Disability Registry, Cancer Registry of Norway. Tables 2 and 3 present number of deaths and new cases of cancer in CONOR since date of examination by linkage to the death and cancer registries. Other large linkages include data from the Norwegian Drug Prescription Database and information from

Table 2 Number of participants ( $n$ ) and number of deaths until December 31, 2003 in the cohort of Norway (CONOR) by age at inclusion in the surveys

|  | Men |  |  | Women |  |
| :--- | ---: | :---: | :---: | :---: | :---: |
| Age (years) | $n$ | Deaths |  | $n$ |  |
| $<25$ | 2037 | 15 | 2512 | Deaths |  |
| $25-34$ | 12028 | 56 | 14658 | 22 |  |
| $35-44$ | 21544 | 158 | 24399 | 123 |  |
| $45-54$ | 17009 | 296 | 18474 | 218 |  |
| $55-64$ | 11698 | 604 | 11903 | 325 |  |
| $65-74$ | 13654 | 2008 | 9399 | 991 |  |
| $\geqslant 75$ | 6183 | 2138 | 7738 | 2141 |  |
| Total | 84153 | 5279 | 89083 | 3826 |  |

Table 3 Follow-up 1994-2006 ${ }^{\text {a }}$ of the CONOR cohort members. Number of cases of first cancer diagnosis in the Norwegian Cancer Registry after initial CONOR examination

|  | Men |  | Women |  |
| :---: | :---: | :---: | :---: | :---: |
|  | $<70$ years | $\geqslant 70$ years | <70years | $\geqslant 70$ years |
| Cancer site (ICD-7) |  |  |  |  |
| Colorectal cancer (152-4) | 582 | 631 | 528 | 476 |
| Trachea, bronchus and lung (162) | 191 | 300 | 133 | 110 |
| Breast (170) | 1 | 4 | 936 | 271 |
| Prostate (177) | 607 | 995 | 0 | 0 |
| Bladder and other urinary organs (181) | 102 | 235 | 33 | 51 |
| Melanoma of skin (190) | 170 | 89 | 238 | 82 |
| All sites (including basal cell carcinoma of skin) | 3180 | 3971 | 5411 | 2515 |

${ }^{\text {a }}$ Follow-up approximately through March 2006.
health surveys in several counties in the 1970s. There are also a number of disease registers that may be linked to the CONOR database. Earlier this year, the government passed a new legislation to make the national hospital discharge register personal identifiable, which would be possible to link to CONOR in the near future.
A major strength of CONOR is its sample size that means it would be able to make a unique contribution to establish main genetic effects and gene-environmental interactions, since precise and robust estimation of these effects requires very large sample sizes. ${ }^{20,21}$ Our aim is to reach 200000 individuals with blood samples and extracted DNA and we anticipate reaching this sample size by Spring 2008. For some hypotheses, it would be most efficient to employ a nested case control study design within CONOR, and we anticipate several such studies in the future. This comparatively large sample size means cases for a number of common and less common diseases may be identified from various sources.
There are some important weaknesses: the overall participation rate is $58 \%$ and is lowest in the surveys in Oslo and other
urban areas and became lower throughout the study period. However, the overall participation rate is influenced by low participation rate in those aged $\leqslant 30$ years. The study population is somewhat heterogeneous as it includes sampling from 10 geographical areas with various age groups included over a 10 -year period. The number of core variables is limited, and in some cases the wording of questions is slightly changed over the years.

## Can I get hold of the data? Where can I find out more?

Guidelines have been developed for projects using data from CONOR (www.fhi.no). These shall ensure that projects will have a high scientific quality, facilitate quick publication of results from CONOR and make the data accessible for research. Research groups may apply for access. A project leader must be appointed. Researchers not residing in Norway are advised to seek contact with Norwegian counterparts. The study objectives should be within the broader aims of CONOR. Further details of these guidelines are provided at the CONOR website.
Applications and enquiries can be sent electronically to the Norwegian Public Health Institute (email: conor@fhi.no). Applications will be evaluated by the CONOR Steering Committee.

## Acknowledgements

The authors would like to thank Kjell Bjartveit, Yngve Haugstveit, Per G Lund Larsen and Arve Sjølingstad for coordination of the blood samples, preparation and management of the CONOR cohort data file. The authors would also like to thank all those who have been involved in collecting the data at the Universities of Bergen, Oslo, Tromsø and Trondheim and the Norwegian Institute of Public Health.

## References

${ }^{1}$ Magnus P, Arnesen E, Holmen J et al. CONOR-Cohort NORway: historie, formål og potensiale. Nor J Epidemiol 2003;13:79-82.
${ }^{2}$ Magnus P, Irgens LM, Haug K et al. Cohort profile: the Norwegian Mother and Child Cohort Study (MoBa). Int $J$ Epidemiol 2006;35:1146-50.
${ }^{3}$ Norwegian Institute of Public Health. Administration and Handling of Applications for Data and Biological Materials from Biohealth Norway. 2007. Available at: http://www.fhi.no/eway/default.aspx?pid=2388 $\operatorname{trg}=$ MainLeft 5853\&MainArea $5811=5853: 0: 15,3467: 1: 0: 0::: 0: 08$ MainLeft_5853=5825:56736::1:5857:1:::0:0.
${ }^{4}$ Hammer H . The central population registry in medical research. Tidsskr Nor Laegeforen 2002;122:2550.
${ }^{5}$ Tverdal A, Foss OP, Leren P et al. Serum triglycerides as an independent risk factor for death from coronary heart disease in middle-aged Norwegian men. Am J Epidemiol 1989;129:458-65.
${ }^{6}$ Foss O, Urdal P. Cholesterol for more than 25 years: could the results be compared throughout all this time? Nor $J$ Epidemiol 2003;13:85-88.
${ }^{7}$ Lund-Larsen PG. Blood pressure measured with sphygmomanometer and with Dinamap under field conditions - a comparison. Nor J Epidemiol 2007;7:235-41.
${ }^{8}$ Meyer HE, Berntsen GK, Søgaard AJ et al. Higher bone mineral density in rural compared with urban dwellers: the NOREPOS study. Am J Epidemiol 2004;160:1039-46.
${ }^{9}$ Ainsworth BE, Montoye HJ, Leon AS. Methods of assessing physical activity during leisure and work. In: Bouchard C, Shephard RJ, Stephens T (eds). Physical Activity, Fitness and Health. Champaign, IL: Human Kinetics, 1994;146-59.
${ }^{10}$ Brugha T, Bebbington P, Tennant C et al. The list of threatening experiences: a subset of 12 life event categories with considerable long-term contextual threat. Psychol Med 1985;15:189-94.
${ }^{11}$ Derogatis LR, Lipman RS, Rickels K et al. The Hopkins Symptom Checklist (HSCL): a self-report symptom inventory. Behav Sci 1974;1:1-15.
12 Joakimsen RM, Fonnebo V, Magnus JH et al. The Tromsø Study: physical activity and the incidence of fractures in a middle-aged population. J Bone Miner Res 1998;13:1149-57.
${ }^{13}$ Løchen ML, Rasmussen K. The Tromsø Study: physical fitness, self reported physical activity, and their relationship to other coronary risk factors. J Epidemiol Community Health 1992;46:103-7.
${ }^{14}$ Saltin B, Grimsby G. Physiological analysis of middle-aged and old former athletes. Circulation 1968;38:1104-15.
${ }^{15}$ Søgaard AJ, Bjelland I, Tell GS et al. A comparison of the CONOR Mental Health Index to the HSCL-10 and HADS. Measuring mental health status in the Oslo Health Study and the Nord-Trøndelag Health Study. Nor J Epidemiol 2003;13:279-84.
${ }^{16}$ Strand BH, Dalgard OS, Tambs K et al. Measuring the mental health status of the Norwegian population: a comparison of the instruments SCL-25, SCL-10, SCL-5 and MHI-5 (SF-36). Nord J Psychiatry 2003;57:113-18.
${ }^{17}$ Thune I, Brenn T, Lund E et al. Physical activity and the risk of breast cancer. N Engl J Med 1997;336:1269-75.
${ }^{18}$ Tretli S, Lund-Larsen PG, Foss OP. Reliability of questionnaire information on cardiovascular disease and diabetes: cardiovascular disease study in Finnmark county. J Epidemiol Community Health 1982;36:269-73.
${ }^{19}$ Søgaard AJ, Gustad TK, Bjertness E et al. Urban-rural differences in distal forearm fractures: Cohort Norway. Osteoporos Int 2007;18:1063-72.
${ }^{20}$ Risch N. . Evolving methods in genetic epidemiology. II. Genetic linkage from an epidemiologic perspective. Epidemiol Rev 1997;19:24-32.
${ }^{21}$ Clayton MA, McKeigue PM. Epidemiological methods for studying genes and environmental factors in complex diseases. Lancet 2001;358:1356-60.

## Appendix 3

| Project name | The role of smoking and socio-economy in <br> explaining health disparities in breast <br> cancer and colorectal cancer incidence and <br> mortality |
| :---: | :---: |
| Authors | Variables Description |
| Finalized |  |
| Date of masterfile | 16 March 2012 |
| Name of masterfile | master_sc_v_112.zip |

## Inclusions selected on survey from data manager:

| 3 Counties I | 62220 |
| :--- | ---: |
| 3 Counties II | 9188 |
| 3 Counties III | 22538 |
| CONOR | 137182 |
| 40 Years (total) | 403691 |
| Oslo I | 17973 |
| Sum | 652,792 |

## Analytical cohort: 602, 242( m=299,376, f=302,866)

Cancer cases in cohort by smoking status

|  | Never-smokers | Former-smokers | Current-smokers | Total |
| :--- | :--- | :--- | :--- | :--- |
| Breast cancer | 3,028 | 1,581 | 2,881 | $7,490^{*}$ |
| Colon cancer | 1,368 | 1,099 | 1,531 | 3,998 |
| Rectal cancer | 648 | 602 | 926 | 2,176 |

*Only among women

Cancer Mortality in cohort by smoking

|  | Never-smokers | Former-smokers | Current-smokers | Total |
| :--- | :--- | :--- | :--- | :--- |
| Breast cancer | 459 | 216 | 431 | $1,106^{*}$ |
| Colon cancer | 1,607 | 443 | 642 | 1,607 |
| Rectal cancer | 202 | 181 | 343 | 726 |

*Only among women

## Daily smokers

The daily-smokers variable in CONOR was based on question "Do you smoke daily?" (In CONOR, this question includes cigarettes, pipe and cigar daily smokers, according to CONOR documentation (variable a8_0)).

In Oslo health study I, the question "Do you smoke daily?" is used for current smokers. Answering "yes" to this question will be current smokers.

In the Norwegian counties study (I, II and III), this was based on the question "Do you smoke daily now?" A positive answer will give a categorization of daily smoker. (We do not consider other answers regarding smoking to classify the current smokers.)

40 years I was based on the question "Do you smoke daily now?" Answering "Yes" will be current smokers.

40 years II was based on the questions "Do you smoke cigarettes daily? Or "Do you smoke cigar daily?" "Do you smoke pipe daily?" answering "Yes" to any of these questions gives daily-smokers.

The 40 years III and IV was based on "Do you smoke cigarettes daily?" or "Do you smoke cigar daily?" or "Do you smoke pipe daily?" If participants have answered "Yes" on any of the above questions, then they are categorized as current smokers.

## Former smokers

After we got all current smokers, then we categorized remaining participants in the formersmokers category as below:

In CONOR if participants have valid answer (greater than 0 ) in questions "How long time since quit smoking (a_9)?" or numbers of cigarettes smoking daily (a_10) or "How old were you when you start smoking (a_11)? or "How many years of smoking in total(a_12_1)??" ,then categorized as former- smokers.

Oslo study I: Those who answered "Yes" to the question "Have you smoked cigarettes daily previously" (tidlrok) in Oslo health study were classified as former smokers. In addition, we check if a valid value on (tidsidsl) "How long since quitting?!", if there is a valid value then we categorized them as former smokers.

In the Norwegian counties those answering "Yes" to the questions "Have you smoked cigarettes daily previously?" were categorized as former-smokers. If answering any value (except zero) to the question "How long since you quit smoking?", and "How many years have you smoked daily?" and "how many cigarettes do you or did you smoke daily?", and not a current smoker, then categorized as a former smoker.

40 years I and II is done similar as the Norwegian Counties. Those answering "Yes" to the questions "Have you smoked cigarettes daily previously?" were categorized as formersmokers. If answering any value (except zero) to the question "How long since you quit
smoking?", and "How many years have you smoked daily?" and "how many cigarettes do you or did you smoke daily?", and not a current smoker, then categorized as a former smoker.
(Please note the comment from Randi about classification this question in 40 years II.)
40 years III and IV: any answer more than zero in the question "if you have smoked previously, how long since you quit?" then a former smoker. (As answering option is in years, we might misclassify those answering zero because they have quit less than 1 year ago.) Also, answering any value more than zero to the questions "how many cigarettes do you smoke or did you smoke daily", "how old were you when you started to smoke daily?" or "how many years have you smoked daily?", then classified as former smoker, if not already classified as a current smoker.

After we have categorized current and former-smokers, from the remaining group of participants, we categorized never-smokers in the following ways:

## Never smokers

CONOR: Answering "No" to the question "Do you smoke daily (a8_0)?" then never smokers.

In the Norwegian counties study, participants answering "No" in the questions "Do you smoke cigarettes daily?" or Do you smoke cigars daily?" or Do you smoke pipes daily?" and if answering "No" to the question "Have you smoked cigarettes daily previously?" were categorized as never smokers.

In the 40 years I and II we did the same in the Norwegian counties. Participants answering "No" in the questions "Do you smoke cigarettes daily?" or "Do you smoke cigars daily?" or "Do you smoke pipes daily?" and if answering "No" to the question "Have you smoked cigarettes daily previously?" were categorized as never smokers.

40 years III: Participants answering "No" to the question "Do you smoke cigarettes daily?" Do you smoke cigars daily?" or "Do you smoke pipes daily?" and not answering the question "if you have smoked previously, how long since you quit?", then categorized as never smoker.

40 years IV: Participants answering "No" to the questions "Do you smoke cigarettes daily?" or "Do you smoke cigars daily?" or "Do you smoke pipes daily ?" and not answering the question "if you have smoked previously, how long since you quit?", then they are categorized as a never smoker. In addition we include the question unique for IV: "Never smoked daily?", then a never smoker. (Brings any records from missing to never, not from daily or former.)

Oslo: Those answering "No" to the both questions "Do you smoke daily?" and answering "No" to the question "Have you smoked cigarettes daily previously?" were categorized as never-smokers.

## Ever-smokers (daily+ former- smokers)

## Duration of smoking

The duration of smoking variable was based on two questions. In the CONOR and the Oslo health study I, daily and former smokers answered the questions "Numbers of years smoked?" In the Norwegian counties study and the 40 years cohort, subjects answering that they were ever smokers were asked "How many years all together have you smoked daily?" Duration of smoking will be further categorized into three groups (1-29, 30-39 and $>40$ )(Ref: Cigarette smoking and risk of colorectal cancer among Norwegian women). Suggestion: Look in EPIC article for different categories which can be appropriate to use in our cohort)

## Age at smoking initiation

The age at smoking initiation variable in CONOR and 40 years III+IV was based on question "How old were you when you started smoking"?

In the Norwegian counties study, 40 years I and II cohort and Oslo health study I, this variable is constructed. We subtracted total years of smoking from age at enrollment to construct the age at smoking initiation. This variable was available for both daily and former smokers.

## Numbers of cigarettes

The numbers of cigarettes variable was based on question "Numbers of cigarettes smoked daily?" in CONOR and Oslo health study I. In the Norwegian counties study(I, II and III) and 40 years cohort(I,II,III and IV), ever-smokers were asked "How many cigarettes do you smoke/smoked daily?" to extract information on numbers of cigarettes. We will further categorized it into three groups (1-9, 10-14 and > 15) (Ref: Gram et al: Cigarette smoking and risk of colorectal cancer among Norwegian women). This can be modified during the analysis by other categorizations if more groups needed.

## Time since quitting smoking (former smokers only)

The time since quitting smoking variable was based on question "How long since you have quit smoking?" in CONOR, 40 years III and IV.

Answering option in CONOR and 40 years III and IV was "time in years" continuous variable. (rokslutp3 roykslutp4)

In the Norwegian counties study, Oslo health study I and 40 years I there were four different answering options:
a. Quit since 3 months
b. Quit since 3 months to 1 year
c. Quit since 1 to 5 years
d. Quit for more than 5 years

In 40 years II the question was "If you have smoked previously, how long since you quit" with answering options "less than one year" and "more than one year". (roykslutp2)

Answers > 60 years is set to missing as outlier ( $\mathrm{n}=4$ ).

## Conclusion:

- For current smokers "time since quitting smoking" can be handled ok.
- For former smokers it is a problem for 40 years II because we can only differ between <1 year and > 1 year.
- We decide that former smokers from Norwegian Counties, 40 years I and II and Oslo I will be called missing in the continuous variable, but can still be handled as categorical variable with four options.


## Latency

We have used information from several variables (see below.). For current smokers the information is good. For former smokers, we have information from CONOR and 40 years III and IV. The others are set to missing.
Latency is a constructed variable
Latency for current smokers:
a. Years between smoking initiation and cohort enrollment(latency 1)
or
b. Years between smoking initiation and censoring/failures(latency 2)

For former-smokers
a. Years between smoking initiation and time since quitting

In some of the surveys, like in the Norwegian counties study 40 years I+II and Oslo health study I, we have "time since quitting" variable which was used for constructing latency for former-smokers was available only in four different options as:

1. Less than three months
2. Three months to 1 year
3. 1 year to 5 years
4. 5 years to more

Our main goal was to create a continuous latency variable which was not possible for former-smokers in these surveys.
a. Latency

Latency 1 (Total years from smoking initiation and quitting or cohort enrollment current smokers only)
b. Latency 2 (Total years between smoking initiation to failure/censoring - current smokers only)
c. Latency 3 (Total years between smoking initiation and quitting or cohort enrollment- former smokers only)
"Only for CONOR, 40 years III and IV"
\# missing here includes if participants are from other surveys rather than CONOR, 40 years III and IV".
d. Latency 4 (Total years between smoking initiation to failure/censuring - former smokers only)
"Only for CONOR, 40 years III and IV"

## Pack- years of smoking

This is calculated as number of cigarettes smoked per day, divided by 20 and multiplied by the number of years smoked.

## Pipe smokers

The "pipe_smoker_sc" variable yes/no comes from all our surveys.
The amount of pipe smoking ( packs pr week ) will come from 3C I, II, III, 40Y I, II, and Oslo I. Variable name "number_pipetobacco_sc".

In Oslo 1 they only ask about nr of packs in 3 categories. We have estimated that if answering $0-0,5$ pack will be 0,25 pack, $1-2$ packs will be 1,25 and 2 packs will be 2 packs. Then they are categorized in the variable "number_pipetobacco_sc".

Further, if any answer then considered "yes", if no answer then considered "no", in the "pipe_smoker_sc" variable.
(For BC analysis pipe smokers are disregarded due to very low number of female pipe smokers.)

## Alcohol Variables

The alcohol variables are from the CONOR and the 40 years study III and IV. The 40 years study I and II, the Oslo study and the Norwegian county study has no alcohol information.

## Teetotalers

In CONOR and 40 years study III and IV the question was "are you a teetotaler?" and there was a "yes/no" answering option.
We have added the persons who are light/moderate/heavy drinker from the "alcohol frequency" variable into the non-teetotalers group, to increase the numbers of non-teetotallers.

## Alcohol frequency

Our alcohol frequency variable is constructed to become a light, moderate and heavy ( $\mathrm{n}=42$, drinker as categorical variable. In general, we have considered a heavy drinker to drink more than once a week, a moderate drinker once a week, and a light drinker to drink less than once a week.

## CONOR

In the CONOR study the variable "drinking pattern" is a 1 to 5 categorical variable: 1 . Drinking more than once a week 2. Drinking once a week. 3. 2-3 times pr month 4. Once a month. 5. Less than once a month. The following categorization has been made: if answering 1 in CONOR, then categorized as heavy drinker. If answering 2 in Conor, then categorized as a moderate drinker. In answering 3,4 or 5 in CONOR, then categorized as a light drinker.

## 40 years

There is no information about alcohol consumption in 40 years I and II. In 40 years III and IV the question was "how many times pr month do you drink alcohol?". If drinking 5 times or
more pr week, then categorized as a heavy drinker. If drinking 4 times pr month (once a week) then categorized as a moderate drinker. If drinking less, then categorized as a light drinker.

The Norwegian counties study and Oslo health study I No information.

## Alcohol grams pr day

This variable has been constructed from information about drinking frequency and type of drink. According to the (ref: www.fhi.no), one glass of wine equals 14,4 grams of pure alcohol, one glass of beer equals 11,9 grams of pure alcohol, and one glass of spirits equals 12,8 grams of pure alcohol. Values larger than 100 grams pr day has been considered extreme, and have been set to missing $(\mathrm{n}=12)$.

## CONOR

In CONOR the question was "how many glasses of wine / beer / spirits do you drink in a two weeks period?" The calculated amount of grams was divided on 14 , to get the alcohol consumption per day.

40 years
In 40 years III and IV the question was "how many glasses of wine / beer / spirits do you drink in a two weeks period?" (Calculation as above).

## BMI

Height and weight were recorded at the health station for all participants, and body mass index (BMI) was calculated by standard formula (ref). Observations with extreme values for height and weight were set to missing as follows: height <100 or >250 cm, weight <35 or $>250 \mathrm{~kg}, \mathrm{BMI}<15$ or $>60 \mathrm{~kg} / \mathrm{m} 2$. (Ref: T Stocks Me-Can Cohort Profile 2009).

BMI is categorized in 4 different groups according to WHO classifications in following order:

1. <18.5
2. 18.5-24.9
3. $25-29.9$
4. $>30$

In the analysis we will collapse category 1 and 2 due to low number in category $1(1.17 \%)$ giving BMI as a 1-3 category.

## Other variables

Menopause assessment (women only)

Women were categorized as pre-, peri- or postmenopausal. Only 10 per cent of our cohort was equal to, or older than 48 years old at inclusion, therefore most in our cohort was premenopausal at inclusion.

Questions about menopause were present in CONOR and 40 years III and IV as a continuous variable "age at menopause". In the County Study and in 40 years I and II, this was a question with 6 options: "

$$
\begin{aligned}
& 1=\text { Ja, menopause inntrådt } \\
& 2=\text { Nei, menopause ikke inntrådt } \\
& 3=\text { Usikker om menopause } \\
& 4=\text { Gravid } \\
& 5=\text { primar amenorrhoe } \\
& 6=\text { Hysterectomy }
\end{aligned}
$$

Answering 1 and 6 were classified as postmenopausal, 2 and 4 were premenopausal, 3 and 5 were uncertain and classified as the other missing according to age (see below):

If missing information, women were classified as premenopausal if they were less than 46 years of age. If they were older than 55 years of age, they were classified as postmenopausal. Women who were between 46 and 55 years of age were classified as perimenopausal / unknown. (Ref: EPIC).

## Oral contraceptive use (woman only)

We made the variable "oral contraceptive use" a binary variable (ever / never). In CONOR it was reported in questionnaires as current, former or never user, and the current and former category were collapsed into ever user by us. There is no information about OC in the County Study.
In the 40 years study, this information was initially collected through interviews, later from questionnaires. Due to inconsistent information from several of these studies, we have only used information from 40 year III in our study. This is in accordance with advice from tex. Anders.

## Post- menopausal hormonal therapy (PMHT) (women only)

Post-menopausal hormonal therapy (PMHT) in CONOR was 5 category options, with different answering options for never users, former users, and for users of PHT with or without prescriptions. In the 40 years study, the answering options were ever, former, never. There is no information about PHT in the Norwegian counties study.

## Menarche (women only)

Age at menarche was categorized as a continuous variable. Information about menarche is in CONOR and 40 years III and IV.

Comment from Anders: use average age for menarche?

Women reporting menarche at age 6 years old or less $(\mathrm{n}=9)$, or 22 years old or more $(\mathrm{n}=31)$, were set to missing.

## Parity (women only)

Information about parity was provided by the Statistics Norway, and is the reported number of live born children at 31 . December 2001. This is the official data and is more updated than the questionnaire.

## Age at first childbirth (women only)

Variable created from information provided by the SSB, which provided the year for the persons first child, and birth year.
Year first childbirth - year born $=$ age at first childbirth

## Smoking exposure before first childbirth (woman only)

Year at first childbirth was given by the SSB.
Age at smoking initiation is a continuous variable in CONOR and 40 years III and IV.
The age at smoking initiation variable in CONOR and 40 years III+IV was based on question "How old were you when you started smoking"?

In the Norwegian counties study, 40 years I and II cohort and Oslo health study I, this variable is constructed. We subtracted total years of smoking from age at enrollment to construct the age at smoking initiation. This variable was available for both daily and former smokers.

We therefore have good information about smoking exposure before first childbirth, for both former and current smokers.

Formulas:

1. Year of survey assessment - total years of smoking $=$ year of smoking initiation

Year of smoking initiation - year of birth $=$ age at smoking initiation
2. Age at enrollment - total years of smoking $=$ age at smoking initiation

## Total: Age at smoking intiation

Year first childbirth - year smoking initiation $=$ years of smoking before first childbirth
Excluded:

- Male sex
- Non-smokers
- Smokers initiating after first childbirth
- No parity

In the variable exposure_before_first_childbirth are those with negative number (ie those initiating after first childbirth) not included.

## Physical activity

The physical activity variable was created as a 1 to 4 categorical variable, with the variable description from CONOR as a reference: 1. Reading, watch TV, other sedentary activity, etc. 2. Walking, bicycling, etc. 3. Light sports, heavy gardening > 4 hours pr week. 4. Hard exercise, competitive sports regularly. In all the included studies except 40 years III, there were a 1 to 4 categorical variable.

In the 40 years III, there were two questions for physical activity: "how much light activity do you do pr week?", and "how much heavy activity do you do pr week", with a 1 to 4 answering option for both questions.

If answering 1 or 2 to I aktiv then 1
3 or 4 to Iaktiv then 2
1 or 2 to h_aktiv then 3
3 or 4 to h_aktiv then 4

Group 1: Light physical
Group 2: Mild physical activity
Group 3: Moderate physical activity
Group 4: Hard physical activity

## Education

We have information about education level from SSB, and the 1970, 1980 and 1990 census. By consensus, we decide to use the highest level of education from the 1980 or 1990 census. If the information is missing, then we use the 1970 census. If no information from any census, then real missing.

Educational level was given in 1-8 categorical variables from SSB. Value 9 is not answered or unknown level of education:

1. 7 years primary school
2. 9-10 years primary/secondary school
3. Technical school, middle school, vocational school, 1-2 years senior school
4. University or university college level 1
5. University or university college level 2
6. University or university college level 3
7. University researcher level
8. Not answered or unknown level of education

These were merged into four levels of education as follows:
1: 1 and 2 low education level
2: 3 and 4 low/medium education level
3:5 and 6 medium/high education level
4: 7 and 8 high education level
This made four education categories (new_ses4groups_NEW).

## Income

As for education, information provided by SSB from the 1970, 1980, 1990. Information about income was categorized in different ways in the different census, which makes it difficult to compare the different time periods.

Income was categorized as follows: Distribution of all incomes at one census was categorized in quartiles. The first quartile was given value 1 , the second quartile was given value 2 , the third quartile was given 3, and the fourth quartile was given 4 . This was done for all three census independently.

The highest quartile registered at either census counted for that individual. The income files were organized by Knut Hansen in the master file (income_max_quart).

## SES

To create four groups for socioeconomic status (SES), income and education categories were added. The sum classified the individuals as follows:
A) 2 score $=$ SES group 1
B) 3 and 4 score $=$ SES group 2
C) 5 and 6 score $=$ SES group 3
D) 7 and 8 score $=$ SES group 4

Comment: we suggest creating 3 SES groups instead of 4 . The reason for this is that the groups 2 and 3 will be very homogenous, if we create 4 categories.

If we create 3 categories, we will have a low, middle and high SES category, which is a common way of classifying social groups. It probably gives a more correct picture of the data, as the most important issue about SES will be to differ between low and high SES. We therefor also create a variable (ses3groups_NEW), where the above group 2 and 3 is merged.
eb
rp

Appendix 4

Table: Prospective studies published in the period 2002-2013 examining the association between smoking and risk of colorectal cancer

\begin{tabular}{|c|c|c|c|c|c|c|c|c|}
\hline \multirow[t]{2}{*}{Reference, location, name of study} \& \multirow[t]{2}{*}{Cohort description} \& \multirow[t]{2}{*}{No. of subjects} \& \multirow[t]{2}{*}{Smoking categories} \& \multirow[t]{2}{*}{No of cases} \& \multicolumn{3}{|l|}{Relative risks (95\% CI or \(p\) value)} \& \multirow[t]{2}{*}{Adjustment factors/ comments} \\
\hline \& \& \& \& \& Colon cancer \& Rectal cancer \& Colorectal cancer \& \\
\hline Terry et al. (2002), USA, Canadian National Breast Screening Study (NBSS) \& Multicenter randomized controlled trial of mammography screening. 89835 women aged 40-59 years. Follow-up 19821993. Incident colorectal cancer or death was ascertained by computerized record linkage to the National Mortality Database and the Canadian Cancer Database. Participants completed a selfadministered questionnaire. \& 363 colon, 164 rectal incident cases \& \begin{tabular}{l}
Never smokers Ex-smokers Current smokers Cigarettes/d 1-9 \\
10-19 \\
20-29 \\
30-39 \\
40+ \\
p trend \\
Years smoked \\
1-9 \\
10-19 \\
20-29 \\
30-39 \\
40+ \\
p trend \\
Years since smoking commended 1-9 \\
10-19 \\
20-29 \\
30-39 \\
40+ \\
p trend
\end{tabular} \& \[
\begin{gathered}
\hline 274 \\
145 \\
108 \\
56 \\
78 \\
93 \\
12 \\
8 \\
\\
\\
42 \\
53 \\
83 \\
61 \\
12 \\
\\
\\
\\
12 \\
24 \\
85 \\
105 \\
22
\end{gathered}
\] \& 1.0
\(1.03(0.80-1.33)\)
\(0.93(0.71-1.24)\)
\(0.89(0.61-1.28)\)
\(0.94(0.67-1.32)\)
\(1.16(0.87-1.53)\)
\(0.87(0.44-1.69)\)
\(0.63(0.26-1.52)\)
0.99
\(0.93(0.61-1.40)\)
\(0.90(0.62-1.30)\)
\(1.04(0.77-1.42)\)
\(1.16(0.83-1.63)\)
\(0.68(0.25-1.86)\)
0.66

$1.50(0.74-3.05)$
$0.84(0.50-1.40)$
$0.91(0.67-1.24)$
$1.05(0.79-1.39$
$1.12(0.62-2.04)$

0.98 \& | 1.0 |
| :--- |
| 1.44 (1.00-2.06) |
| 1.17 (0.78-1.75) |
| 1.31 (0.80-2.14) |
| 1.98 (1.32-2.96) |
| 0.97 (0.61-1.56) |
| 0.72 (0.23-2.29) |
| 0.90 (0.28-2.85) |
| 0.82 |
| 1.31 (0.75-2.28) |
| 1.24 (0.75-2.05) |
| 1.37 (0.89-2.11) |
| 1.12 (0.65-1.94) |
| 3.14 (1.33-7.42) |
| 0.07 |
| 1.76 (0.64-4.82) |
| 0.97 (0.47-2.02) |
| 1.11 (0.72-1.73) |
| 1.52 (1.01-1.26) |
| 2.27 (1.06-4.87) |
| 0.03 | \& \& Adjusted for age (in 5year age groups), BMI (quartiles), educational level, vigorous physical activity, hormone replacement therapy, menopausal status and alcohol intake <br>

\hline | Tiemersma et al. (2002), |
| :--- |
| Netherlands, |
| Monitoring |
| Project on Cardiovascular Disease Risk Factors | \& Nested case-control study, controls frequency matched for age and gender. Using data from the prospective \& 102 incident cases, 537 controls \& Never smokers Former smokers Current smokers p trend \& \[

$$
\begin{aligned}
& 30 \\
& 43 \\
& 29
\end{aligned}
$$

\] \& \& \& \[

$$
\begin{aligned}
& 1.0 \\
& 1.01 .4(08-2.5) \\
& 0.9(0.5-1.7) \\
& 0.27
\end{aligned}
$$
\] \& Adjusted for age, sex, center, coffee and alcohol consumption and body mass index. <br>

\hline
\end{tabular}

| Reference, location, name of study | Cohort description | No. of subjects | Smoking categories | No of cases | Relative risks (95\% CI or $\boldsymbol{p}$ value) |  |  | Adjustment factors/ comments |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  |  |  | Colon cancer | Rectal cancer | Colorectal cancer |  |
| Tiemersma et al. (2002), <br> (contd) | Monitoring Project on Cardiovascular Disease Risk Factors conducted in Amsterdam, Maastricht and Doetinchen. Including persons aged 20-59 years. Follow-up 19871998. Incident cancer was obtained by record linkage with the Netherlands Cancer Registry and with the three regional cancer registries. Participants completed a selfadministered questionnaire. |  | Smoking duration (years) Former smokers $1-15$ $16-30$ $>30$ p trend Current smokers $1-15$ $16-30$ $>30$ p trend Cigarettes/d Former smokers $1-10$ $11-20$ $>20$ p trend Current smokers $1-10$ $11-20$ $>20$ p trend Time since quit smoking $>18$ years $9-18$ years $0-8$ years p trend | 13 <br> 23 <br> 7 <br> 3 <br> 7 <br> 19 <br> 12 <br> 21 <br> 10 <br> 10 <br> 14 <br> 5 <br> 18 <br> 16 <br> 9 |  |  | $\begin{aligned} & 1 \text { (ref.) } \\ & 2.7 \text { (1.03-7.4) } \\ & 3.2 \text { (1.04-9.8) } \\ & 0.04 \\ & \\ & 1 \text { (ref.) } \\ & 0.4 \text { (0.1-1.9) } \\ & 1.9(0.5-8.2) \\ & 0.28 \\ & \\ & \\ & 1 \text { (ref) } \\ & 2.1 \text { (0.9-5.0) } \\ & 1.7 \text { (0.6-4.6) } \\ & 0.15 \\ & \\ & 1 \text { (ref.) } \\ & 1.1 \text { (0.4-2.8) } \\ & 1.2 \text { (0.3-4.0) } \\ & 0.75 \\ & \\ & 1 \text { (ref.) } \\ & 2.6 \text { (1.0-6.5) } \\ & 2.2 \text { (0.8-5.5) } \\ & 0.10 \end{aligned}$ | Adjusted for age, sex, center, coffee and alcohol consumption and body mass index. |
| Limburg et al. (2003); <br> USA; <br> Iowa Women's <br> Health <br> Study | Baseline questionnaire was mailed in January 1986 to randomly selected women aged 55-69 years, 41836 (42,7\%) responded. Incident CRC cases | 869 incident CRC cases and 249 fatal CRC cases | CRC incidence Never smokers Current smokers Former smokers p trend | $\begin{aligned} & 558 \\ & 122 \\ & 189 \end{aligned}$ |  |  | $\begin{aligned} & 1.0 \\ & 1.10(0.89-1.37) \\ & 1.21(1.01-1.45) \\ & 0.14 \end{aligned}$ | Adjusted for age, BMI, waist-hip ratio, physical activity level, hormone replacement therapy, alcohol consumption, intake of methionine, total calories, fat, |


| Reference, | Cohort | No. of | Smoking | No of | Relative risks (95\% CI or $p$ value) |  |  | Adjustment factors/ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  |  |  | Colon cancer | Rectal cancer | Colorectal cancer |  |
| Limburg et al. (2003); (cont) | were identified through the IOWA Cancer Registry, Follow-up continued through December 1999 |  | Age at initiation <br> $\leq 30$ years <br> $>30$ years <br> p trend <br> Total duration <br> (yrs) <br> 1-19 <br> 20-39 <br> $\geq 40$ <br> p trend <br> Cigarettes/d <br> 1-19 <br> 20 <br> $>20$ <br> p trend <br> Pack-years <br> 1-19 <br> 20-39 <br> $\geq 40$ <br> p trend <br> CRC Mortality <br> Never <br> Current smokers <br> Former smokers <br> p trend <br> Age at initiation <br> $\leq 30$ years <br> $>30$ years <br> p trend | 129 <br> 111 <br> 163 <br> 99 <br> 49 <br> 123 <br> 105 <br> 83 <br> 158 <br> 45 <br> 46 <br> 81 <br> 10 |  |  | $\begin{aligned} & 1.20(1.02-1.40) \\ & 0.90(0.59-1.39) \\ & 0.03 \\ & \\ & 1.16 \text { (0.89-1.52) } \\ & 1.08 \text { (0.88-1.32) } \\ & 1.30(1.04-1.63) \\ & 0.03 \\ & \\ & 1.15(0.95-1.38) \\ & 1.23(0.97-1.54) \\ & 1.12(0.82-1.54) \\ & 0.08 \\ & \\ & 1.15(0.93-1.41) \\ & 1.16(0.92-1.45) \\ & 1.21(0.94-1.56) \\ & 0.06 \\ & \\ & 1.0 \\ & 1.58(1.09-2.29) \\ & 1.14(0.80-1.62) \\ & 0.02 \\ & \\ & 1.34(1.00-1.80) \\ & 1.04(0.48-2.22) \\ & 0.14 \end{aligned}$ | sucrose, red meat, calcium, folate, and vitamin E. |


| Reference, location, name of study | Cohort description | No. of subjects | Smoking categories | No of cases | Relative risks (95\% CI or $p$ value) |  |  | Adjustment factors/ comments |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  |  |  | Colon cancer | Rectal cancer | Colorectal cancer |  |
| $\begin{aligned} & \hline \text { Limburg et al. } \\ & \text { (2003); } \\ & \text { (cont) } \end{aligned}$ |  |  | Total <br> duration (yrs) <br> 1-19 <br> 20-39 <br> $\geq 40$ <br> p trend <br> Cigarettes/d <br> 1-19 <br> 20 <br> $>20$ <br> p trend <br> Pack-years <br> 1-19 <br> 20-39 <br> $\geq 40$ <br> p trend | $\begin{aligned} & 24 \\ & 32 \\ & 35 \end{aligned}$ $47$ $33$ $11$ $36$ $27$ $28$ |  |  | $\begin{aligned} & 1.53(0.96-2.43) \\ & 1.02(0.67-1.53) \\ & 1.55(1.04-2.31) \\ & 0.07 \\ & \\ & 1.27(0.89-1.80) \\ & 1.50(0.99-2.28) \\ & 1.07(0.57-2.00) \\ & 0.14 \\ & \\ & 1.30(0.88-1.91) \\ & 1.08(0.69-1.70) \\ & 1.63(1.05-2.49) \\ & 0.05 \end{aligned}$ | Adjusted for age, BMI, waist-hip ratio, physical activity level, hormone replacement therapy, alcohol consumption, intake of methionine, total calories, fat, sucrose, red meat, calcium, folate, and vitamin E. |


| Reference, location, name of study | Cohort description | No. of subjects | Smoking categories | No of cases |  |  | Relative risks (95\% CI or $p$ value) |  |  | Adjustment factors/comments |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  |  | C | R | $\begin{aligned} & \mathrm{CR} \\ & \mathrm{C} \\ & \hline \end{aligned}$ | Colon cancer | Rectal cancer | Colorectal cancer |  |
| Otani et al. (2003), <br> Japan, The Japan <br> Public Health <br> Center-based prospective study on cancer and cardiovascular disease (JPHC study) | Cohort I started 1990 in 5 areas in 5 prefectures (Iwate, Akita, Nagano, Okinawa, Tokyo) and covered all residents aged 40-59. Cohort II started 1993 in 6 areas in 6 prefectures (Ibaraki, Niigata, Kochi, Nagasaki, Okinawa, Osaka) and covered all residents aged 40-69. 57591 men and 59103 women. Active follow-up 1990-1999, 1993-1999 using data of Ministry of Health, Labor and Welfare for deaths and the JPHC cancer registry for incidence. Participants completed a selfadministered questionnaire. | 447 <br> incident <br> cases (299 <br> colon <br> cancers, <br> 148 rectal cancers) | Never smokers <br> Former smokers <br> Current smokers <br> Pack years <br> <20 <br> 20-29 <br> 30-39 <br> 40+ <br> p trend | $\begin{gathered} \hline 53 \\ 86 \\ 160 \\ \\ 17 \\ 31 \\ 55 \\ 54 \end{gathered}$ | $\begin{aligned} & 25 \\ & 38 \\ & 85 \\ & 16 \\ & 19 \\ & 18 \\ & 29 \end{aligned}$ | $\begin{gathered} 78 \\ 124 \\ 245 \\ \\ 33 \\ 50 \\ 73 \\ 83 \end{gathered}$ | 1.0 $1.4(0.96-1.9)$ $1.4(0.99-1.9)$ $0.9(0.5-1.5)$ $1.2(0.8-2.0)$ $1.7(1.1-2.4)$ $1.3(0.9-2.0)$ 0.16 | $\begin{aligned} & 1.0 \\ & 1.2(0.7-2.0) \\ & 1.4(0.9-2.3) \\ & 1.6(0.9-3.0) \\ & 1.5(0.8-2.7) \\ & 1.0(0.6-1.9) \\ & 1.4(0.8-2.3) \\ & 0.48 \end{aligned}$ | $\begin{aligned} & 1.0 \\ & 1.3(0.98-1.7) \\ & 1.4(1.1-1.8) \\ & \\ & 1.1(0.8-1.7) \\ & 1.3(0.9-1.9) \\ & 1.4(1.05-2.0) \\ & 1.4(0.99-1.8) \\ & 0.47 \end{aligned}$ | Adjusted for age (5 year groups), family history of colorectal cancer, BMI, alcohol consumption, physical exercise and 9 Public Health Center areas |



| Reference, location, name of study | Cohort description | No. of subjects | Smoking categories | No of cases |  | Relative risks (95\% CI or $p$ value) |  |  | Adjustment |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  |  |  |  | Colon cancer | Rectal cancer | Colorectal cancer |  |
| Wakai et al. (2003), Japan, <br> Japan Collaborative Cohort (JACC) | 110792 inhabitants aged 40-79 completed a baseline questionnaire. They were enrolled from 45 study areas throughout Japan, total 59879 eligible subjects from 24 study areas with cancer registries Follow-up 19881997 by cancer registries. | 408 colon cancer (219 men, 189 women) and 204 rectal cancer cases (147 men and 57 women) | Men <br> Never smoker <br> Former smokers <br> Current smokers <br> Women <br> Never <br> Former smokers <br> Current smokers <br> Men <br> Cigarettes/d <br> 0-19 <br> 20-39 <br> 40+ <br> p trend <br> Age at starting smoking (yrs) <br> 26+ <br> 23-25 <br> 20-22 <br> <20 <br> p trend <br> Years of smoking <br> 0-19 <br> 20-39 <br> 40+ <br> p trend <br> Pack-years <br> 0-19 <br> 20-39 <br> 40-59 <br> 60+ <br> p trend | 39 <br> 67 <br> 113 <br> 175 <br> 4 <br> 10 <br> 59 <br> 102 <br> 9 <br> 18 <br> 34 <br> 97 <br> 24 <br> 13 <br> 92 <br> 67 <br> 26 <br> 89 <br> 41 <br> 10 | $\begin{gathered} 34 \\ 44 \\ 69 \\ 55 \\ 55 \\ 1 \\ 1 \\ \\ 44 \\ 55 \\ 9 \\ \hline \end{gathered}$ | $\begin{aligned} & 1.0 \\ & 1.07 \text { (072-159) } \\ & 1.23(085-178) \\ & \\ & 1.0 \\ & 1.07 \text { (0.39-2.92) } \\ & 1.06 \text { (055-2.02) } \\ & \\ & \\ & 1.05(0.70-1.58) \\ & 1.30(0.89-1.89) \\ & 0.69(0.33-1.43) \\ & 0.56 \\ & \\ & \\ & 1.10(0.62-1.93) \\ & 1.54(0.97-2.44) \\ & 1.13(0.78-1.64) \\ & 1.04(0.62-1.74) \\ & 0.76 \\ & \\ & 0.99(0.53-1.87) \\ & 1.31(0.89-1.92) \\ & 1.07(0.71-1.61) \\ & 0.52 \\ & \\ & 0.92(0.56-1.52) \\ & 1.43(0.98-2.10) \\ & 1.11(0.71-1.73) \\ & 0.68(0.34-1.37) \\ & 0.90 \end{aligned}$ | $\begin{aligned} & 1.0 \\ & 0.88 \text { (0.56-1.39) } \\ & 0.83(0.55-1.26) \\ & \\ & 1.0 \\ & 1.05(0.14-7.69) \\ & 0.36 \text { (0.05-2.65) } \\ & \\ & \\ & 0.95(0.60-1.50) \\ & 0.79(0.51-1.22) \\ & 0.80(0.38-1.69) \\ & 0.26 \\ & \\ & \\ & 0.73(0.36-1.49) \\ & 0.84(0.46-1.53) \\ & 0.77(0.50-1.18) \\ & 1.18(0.69-1.99) \\ & 0.91 \\ & \\ & 0.58(0.24-1.39) \\ & 1.01(0.65-1.56) \\ & 0.72(0.45-1.16) \\ & 0.35 \\ & \\ & 0.96(0.56-1.66) \\ & 0.89(0.57-1.40) \\ & 0.72(0.42-1.22) \\ & 0.78(0.38-1.59) \\ & 0.23 \end{aligned}$ |  | Adjusted for age, area, education, family history of colorectal cancer, BMI, alcohol drinking, walking time, sedentary work and consumption of green leafy vegetables and beef. |


| Reference, location, name of study | Cohort description | No. of subjects | Smoking categories | No of cases |  | Relative risks (95\% CI or $p$ value) |  |  | Adjustment factors/comments |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  |  |  |  | Colon cancer | Rectal cancer | Colorectal cancer |  |
| Wakai et al. (2003); (cont) |  |  | Years since smoking cessation (yrs) $0-9$ $10-19$ $20+$ p trend | $\begin{aligned} & 31 \\ & 23 \\ & 12 \end{aligned}$ | $\begin{gathered} 16 \\ 20 \\ 6 \end{gathered}$ | $\begin{aligned} & 1.09(0.68-1.75) \\ & 1.29(0.77-2.17) \\ & 0.79(0.41-1.52) \\ & 0.29 \end{aligned}$ | $\begin{aligned} & 0.68(0.37-1.24) \\ & 1.47(0.84-2.57) \\ & 0.53(0.22-1.28) \\ & 0.80 \end{aligned}$ |  | Adjusted for age, area, education, family history of colorectal cancer, BMI, alcohol drinking, walking time, sedentary work and consumption of green leafy vegetables and beef. |
| Colangelo et al. (2004), USA, <br> The Chicago Heart Association Detection Project in Industry (CHA) | Screening program on cardiovascular disease. The CHA cohort was screened between 1967 and 1973. 39522 men and women from 84 cooperating companies and organizations in the Chicago area underwent baseline assessment. 22295 men and 17004 women remained for analyses. Active follow-up until 1979, after 1979 follow-up until 1997 by the National Death Index. | $\begin{aligned} & \hline 349 \mathrm{CRC} \\ & \text { deaths, } 208 \\ & \text { among } \\ & \text { men, } 141 \\ & \text { among } \\ & \text { women } \end{aligned}$ | Men <br> Never smoked <br> Past smoker <br> Cigarettes/d <br> 1-10 cigs/d <br> $11-20 \mathrm{cigs} / \mathrm{d}$ <br> $>20 \mathrm{cigs} / \mathrm{d}$ <br> p trend <br> Women <br> Never smoked <br> Past smoker <br> Cigarettes/d <br> 1-10 cigs/d <br> $11-20 \mathrm{cigs} / \mathrm{d}$ <br> $>20 \mathrm{cigs} / \mathrm{d}$ <br> p trend |  | $\begin{gathered} \hline \text { CRC } \\ 56 \\ 74 \\ 10 \\ 35 \\ 33 \\ \\ 70 \\ 18 \\ \\ 17 \\ 28 \\ 8 \end{gathered}$ |  |  | $\begin{aligned} & 1.0 \\ & 0.96(0.68-1.36) \\ & 0.75(0.38-1.48) \\ & \\ & 1.09(0.71-1.68) \\ & 1.36(0.88-2.11) \\ & 0.19 \\ & 1.0 \\ & \\ & 0.91(0.54-1.53) \\ & 1.23(0.72-2.09) \\ & \\ & 1.43(0.91-2.23) \\ & 1.25(0.59-2.62) \\ & 0.13 \end{aligned}$ | Adjusted for age, race, categories of education, body mass index, gender, and height. |


| Reference, location, name of study | Cohort description | No. of subjects | Smoking categories | No of cases | Relative risks (95\% CI or $p$ value) |  |  | Adjustment factors/comments |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  |  |  | Colon cancer | Rectal cancer | Colorectal cancer |  |
| Jee et al. (2004), Korea, The Korean Cancer Prevention Study (KCPS) | 1307275 Koreans from 30 to 95 years who received health insurance from the Korean Medical Insurance Corporation and who had biennial medical evaluations in 1992-1995. 1212209 participants were the final sample. For information on cancer mortality a Computerized search for death certificate data from the National Statistical Office in Korea was performed. Active follow up 19932001. | 511 colon Cancer cases in men | Men <br> Never smoker Former smokers Current smoker | $\begin{gathered} \text { Colon } \\ 91 \\ 139 \\ 281 \end{gathered}$ | $\begin{aligned} & 1.1(0.9-1.4) \\ & 1.1(0.8-1.4) \end{aligned}$ |  |  | Adjusted for age. |
| Sanjoaquin et al. (2004), United Kingdom, The Oxford Vegetarian Study | 11140 vegetarians and non-vegetarians who were recruited in the UK between 1980 and 1984 completed a questionnaire. Each participant was flagged at the UK National health Service central | 95 incident colorectal cancer cases | Never smoker <br> Former smokers <br> Current smokers | $\begin{gathered} \text { CRC } \\ 36 \\ 43 \\ 16 \end{gathered}$ |  |  | $\begin{aligned} & 1.0 \\ & 1.80(1.13-2.85) \\ & 1.70(0.92-3.15) \end{aligned}$ | Adjusted for age, sex, and alcohol. |

\begin{tabular}{|c|c|c|c|c|c|c|c|c|c|c|}
\hline \multirow[t]{2}{*}{Reference, location, name of study} \& \multirow[t]{2}{*}{Cohort description} \& \multirow[t]{2}{*}{No. of subjects} \& \multirow[t]{2}{*}{Smoking categories} \& \multicolumn{3}{|l|}{No of cases} \& \multicolumn{3}{|l|}{Relative risks (95\% CI or \(p\) value)} \& \multirow[t]{2}{*}{Adjustment factors/comments} \\
\hline \& \& \& \& \& \& \& Colon cancer \& Rectal cancer \& Colorectal cancer \& \\
\hline Sanjoaquin et al. (2004); (cont) \& register for information on cancer registration and death. A total of 10998 participants were included in the analysis. Follow up 1980-1999 \& \& \& \& \& \& \& \& \& Adjusted for age, sex, and alcohol. \\
\hline \begin{tabular}{l}
Doll et al. (2005) \\
United Kingdom
\end{tabular} \& 34439 male British doctors, who reported their smoking habits in November 1951 were follow-up periodically through mailed questionnaire; 50 year for mortality 1951-2001; 272 deaths from pancreatic cancer \& \& \begin{tabular}{l}
Never smoker \\
Cigarette \\
smokers \\
Former \\
Current \\
Current \\
cigarettes/d \\
1-14 \\
15-24 \\
\(\geq 25\) \\
Other smokers \\
Former \\
Current
\end{tabular} \& \& \& \& \[
\begin{aligned}
\& \hline 1.0 \\
\& \\
\& 1.43 \\
\& 1.33 \\
\& \\
\& 1.39 \\
\& 1.13 \\
\& 1.52 \\
\& \\
\& 1.54 \\
\& 1.27
\end{aligned}
\] \& 1.0

1.55
2.39
1.44
1.76
4.73
1.62
2.25 \& 1.0
1.45
1.56
1.39
1.29
2.22
1.55
1.48 \& Standardized indirectly for age and study year <br>

\hline Yun et al. (2005), Republic of Korea, National Health Insurance Corporation Study \& 733134 Korean men, 30 years old or older who were insured by the National Health Insurance Corporation and had a medical evaluation in 1996. Follow-up through 2000. Incident cancer cases were identified from the Korean Central Cancer Registry \& 417 colon, 453 rectum cancer cases \& | Never |
| :--- |
| Former smokers Current smokers Cigarettes/d |
| 1-9 |
| 10-19 |
| $\geq 20$ |
| Current smokers Years of smoking 1-9 |
| 10-29 |
| $\geq 30$ |
| p trend | \& $C$

99
148
170
36
102
32

59
45

66 \& $$
\begin{gathered}
\hline \mathrm{R} \\
106 \\
131 \\
216 \\
\\
38 \\
131 \\
47 \\
\\
\hline 62 \\
76 \\
78
\end{gathered}
$$ \& \[

$$
\begin{gathered}
\mathrm{CR} \\
\mathrm{C}
\end{gathered}
$$

\] \& \[

$$
\begin{aligned}
& 1.0 \\
& 1.37(1.06-1.77) \\
& 0.81(0.63-1.05) \\
& \\
& 0.97(0.66-1.43) \\
& 0.78 \text { (0.59-1.04) } \\
& 0.76(0.51-1.15) \\
& \\
& \\
& 0.87(0.62-1.23) \\
& 0.61(0.42-0.88) \\
& 0.96(0.69-1.33)
\end{aligned}
$$

\] \& \[

$$
\begin{aligned}
& 1.0 \\
& 1.17(0.91-1.52) \\
& 0.97(0.76-1.24) \\
& \\
& 0.95(0.65-1.39) \\
& 0.95(0.73-1.24) \\
& 1.05(0.74-1.50) \\
& \\
& \\
& 0.80(0.57-1.13) \\
& 1.00(0.74-1.36) \\
& 1.12(0.82-1.52) \\
& <0.01
\end{aligned}
$$
\] \& \& Adjusted for age, place of residence, BMI, alcohol drinking, leisure time physical activity frequency, meat consumption, preference for vegetables and meats. <br>

\hline
\end{tabular}

\begin{tabular}{|c|c|c|c|c|c|c|c|c|c|}
\hline \multirow[t]{2}{*}{Reference, location, name of study} \& \multirow[t]{2}{*}{Cohort description} \& \multirow[t]{2}{*}{No. of subjects} \& \multirow[t]{2}{*}{Smoking categories} \& \multicolumn{2}{|l|}{No of cases} \& \multicolumn{3}{|l|}{Relative risks (95\% CI or \(p\) value)} \& \multirow[t]{2}{*}{Adjustment factors/comments} \\
\hline \& \& \& \& \& \& Colon cancer \& Rectal cancer \& Colorectal cancer \& \\
\hline \[
\begin{aligned}
\& \text { Yun et al. (2005); } \\
\& \text { (cont) }
\end{aligned}
\] \& (KCCR) and six regional cancer registries (RCRs). \& \& \[
\begin{aligned}
\& \text { Former smokers } \\
\& \text { Years of smoking } \\
\& 1-19 \\
\& 20-29 \\
\& \geq 30 \\
\& \text { p trend }
\end{aligned}
\] \& \[
\begin{aligned}
\& 95 \\
\& 23 \\
\& 21
\end{aligned}
\] \& \[
\begin{gathered}
89 \\
24 \\
6
\end{gathered}
\] \& \[
\begin{aligned}
\& 1.36(1.02-1.80) \\
\& 1.15(0.72-1.83) \\
\& 2.08(1.29-3.37) \\
\& 0.06
\end{aligned}
\] \& \[
\begin{aligned}
\& 1.21(0.91-1.61) \\
\& 1.23(0.78-1.92) \\
\& 0.61(0.27-1.41)
\end{aligned}
\] \& \& Adjusted for age, place of residence, BMI, alcohol drinking, leisure time physical activity frequency, meat consumption, preference for vegetables and meats. \\
\hline \begin{tabular}{l}
Lüchtenborg et al. (2005), \\
Netherlands, The Netherlands Cohort Study on Diet and Cancer
\end{tabular} \& A total of 58279 men and 62573 women between the ages of 55 and 69 years from 204 municipal population registries completed a self-administered questionnaire in 1986. Incident cancer cases are identified through annual record linkage to the Netherlands Cancer Registry and the Pathologisch Anatomisch Landelijk Geautomatiseerd Archief (PALGA). The vital status of a sub cohort of 3,500 men and women \& 2948 sub cohort members, 661 colorectal cancer cases \& \begin{tabular}{l}
Never smoked \\
Former smokers \\
Current smokers Cigarettes/day <5 \\
5-<10 \\
\(10-<15\) \\
\(15-<20\) \\
20-<25 \\
\(\geq 25\) \\
p trend \\
Duration (yrs) \\
\(<10\) \\
\(10-<20\) \\
\(20-<30\) \\
\(30-<40\) \\
40-<50 \\
\(\geq 50\) \\
p trend
\end{tabular} \& \& CRC
206
298
146
47
50
84
61
76
95

17
53
92
128
109

38 \& \& \& $$
\begin{aligned}
& 1.0 \\
& 1.30(1.03-1.65) \\
& 0.91(0.71-1.18) \\
& \\
& 1.02(0.71-1.46) \\
& 0.91(0.59-1.30) \\
& 1.10(0.80-1.52) \\
& 1.16(0.82-1.64) \\
& 1.15(0.83-1.59) \\
& 1.59(1.16-2.17) \\
& 0.01 \\
& \\
& 1.02(0.59-1.78) \\
& 1.16(0.81-1.64) \\
& 1.15(0.85-1.55) \\
& 1.32(1.00-1.73) \\
& 0.90(0.67-1.20) \\
& 1.45(0.93-2.28) \\
& 0.49
\end{aligned}
$$ \& Adjusted for age (years), sex, family history of colorectal cancer, and BMI. <br>

\hline
\end{tabular}

| Reference, location, name of study | Cohort description | No. of subjects | Smoking categories | No of cases | Relative risks (95\% CI or $p$ value) |  |  | Adjustment factors/comments |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  |  |  | Colon cancer | Rectal cancer | Colorectal cancer |  |
| Lüchtenborg et al. (2005); (cont) | was biannually examined. Follow up 1989-1994. |  | Age at first <br> exposure$<15$$15-<17$$17-<19$$19-<21$$21-<25$$\geq 25$p trendYears sincecessation$<1$$1-<10$$10-<20$$20-<30$$\geq 30$$p$ | 71 <br> 116 <br> 103 <br> 65 <br> 35 <br> 44 <br> 155 <br> 101 <br> 104 <br> 65 <br> 17 |  |  | $\begin{aligned} & 1.14(0.81-1.62) \\ & 1.41(1.04-1.92) \\ & 1.11(0.83-1.50) \\ & 1.26(0.90-1.77) \\ & 1.17(0.78-1.77) \\ & 0.87(0.61-1.25) \\ & 0.32 \\ & \\ & \\ & 0.94(0.73-1.22) \\ & 1.39(1.03-1.86) \\ & 1.38(1.03-1.86) \\ & 1.25(0.88-1.77) \\ & 0.75(0.43-1.29) \\ & 0.27 \end{aligned}$ | Adjusted for age (years), sex, family history of colorectal cancer, and BMI. |
| Kim et al (2006), Korea, Korea Elderly Phamacepidemiolo gic Cohort (KEPEC) | Population-based dynamic cohort with 14103 cohort members aged 65 years or more and living in Busan Metropolitan City from 1993-1998. The participants were beneficiaries of the Korean Medical Insurance Corporation (KIMIC). Baseline information was surveyed by a selfadministered | 100 <br> incident colorectal cancer cases | Non-smoker <br> Former smokers <br> Current smokers <br> p trend <br> Daily smoking <br> amount (packs) <br> $\leq 0.5$ <br> 0.5-1 <br> $>1$ <br> p trend <br> Smoking duration <br> $\leq 20$ pack-yrs <br> 20-40 pack-yrs <br> >40 pack-yrs <br> p trend | CRC <br> 57 <br> 14 <br> 26 <br> 4 <br> 20 <br> 16 <br> 6 <br> 32 2 |  |  | $\begin{aligned} & 1.0 \\ & 2.03(1.02-4.03) \\ & 1.36(0.80-2.32) \\ & 0.26 \\ & \\ & \\ & 1.56(0.56-4.35) \\ & 1.77(1.03-3.05) \\ & 0.95(0.51-1.76) \\ & 0.28 \\ & \\ & 1.29(0.52-3.22) \\ & 1.63(0.97-2.74) \\ & 0.96(0.27-3.24) \\ & 0.15 \end{aligned}$ | Adjusted for age at baseline, gender precancerous lesion of CRC, medication history of NSAID \& antibiotics, alcohol drinking and BMI. |


| Reference, location, name of study | Cohort description | No. of subjects | Smoking categories | No of cases <br> CRC | Relative risks (95\% CI or $p$ value) |  |  | Adjustment factors/comments |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  |  |  | Colon cancer | Rectal cancer | Colorectal cancer |  |
| $\begin{aligned} & \text { Kim et al (2006); } \\ & \text { (cont) } \end{aligned}$ | questionnaire. Follow up for a mean of 8.7 person years. |  | Smoking duration $\leq 45 \mathrm{yrs}$ $>45 \mathrm{yrs}$ <br> P trend <br> Age started smoking <br> $\geq$ age 20 <br> < age 20 <br> p trend | $\begin{gathered} 31 \\ 9 \end{gathered}$ $27$ $13$ |  |  | $\begin{aligned} & 1.51(0.97-2.34) \\ & 2.35(1.16-4.74) \\ & <0.01 \\ & \\ & 1.03(0.60-1.79) \\ & 2.15(1.17-3.93) \\ & 0.06 \end{aligned}$ | Adjusted for age at baseline, gender precancerous lesion of CRC, medication history of NSAID \& antibiotics, alcohol drinking and BMI. |
| Akhter et al. (2007), Japan | Prospective cohort study in 14 municipalities of Miyagi Prefecture in rural northern Japan. 47605 Participants aged 40-64 years (22836 men and 24769 women). <br> Information was obtained by selfadministered questionnaire. Follow-up 19901997. Record linkage with the Miyagi Prefectural Cancer Registry for information on incident cases. Analysis was limited to 21,695 men due to small | 188 incident colorectal cancer cases | Never smoker <br> Former smokers <br> Current smokers <br> Cigarettes/d <br> 1-19 <br> $\geq 20$ <br> p trend <br> Age started <br> smoking <br> $>22$ <br> 19-22 <br> $\leq 18$ <br> p trend <br> Smoking duration <br> (yrs) <br> 1-29 <br> 30-39 <br> $\geq 40$ <br> p trend | $\begin{gathered} 22 \\ 50 \\ 116 \\ 29 \\ 82 \\ \\ \\ 37 \\ 60 \\ 16 \\ \\ \\ 33 \\ 50 \\ 30 \end{gathered}$ |  |  | $\begin{aligned} & 1.0 \\ & 1.73(1.04-2.87) \\ & 1.74(0.93-2.34) \\ & \\ & 1.32(0.75-2.31) \\ & 1.60(0.99-2.58) \\ & 0.04 \\ & \\ & \\ & 1.36(0.80-2.32) \\ & 1.56(0.95-2.55) \\ & 1.86(0.97-3.58) \\ & 0.03 \\ & \\ & \\ & 1.46(0.82-2.60) \\ & 1.52(0.91-2.53) \\ & 1.59(0.89-2.86) \\ & 0.08 \end{aligned}$ | Adjusted for age in years, family history of colorectal cancer; education level, BMI, walking time, alcohol drinking and current drinkers, consumption frequencies of meat, green-yellow vegetables and fruits. |


| Reference, location, name of study | Cohort description | No. of subjects | Smoking categories | No of casesCRC | Relative risks (95\% CI or $\boldsymbol{p}$ value) |  |  | Adjustment factors/comments |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  |  |  | Colon cancer | Rectal cancer | Colorectal cancer |  |
| Akhter et al. (2007); (cont) | prevalence of smoking in women. |  | Smoking duration $\leq 45 \mathrm{yrs}$ <br> > 45 yrs <br> P trend <br> Age started <br> smoking <br> $\geq$ age 20 <br> < age 20 <br> p trend | $\begin{gathered} 31 \\ 9 \end{gathered}$ $27$ $13$ |  |  | $\begin{aligned} & 1.51(0.97-2.34) \\ & 2.35(1.16-4.74) \\ & <0.01 \\ & \\ & 1.03(0.60-1.79) \\ & 2.15(1.17-3.93) \\ & 0.06 \end{aligned}$ | Adjusted for age in years, family history of colorectal cancer; education level, BMI, walking time, alcohol drinking and current drinkers, consumption frequencies of meat, green-yellow vegetables and fruits |
| Huxley (2007) Asia <br> Pacific Cohort <br> Studies <br> Collaboration, <br> Asia Pacific Cohort <br> Studies <br> Collaboration <br> (APCSC) | Collaboration of 33 cohort studies in the region. 539201 participants (35\% female, $65 \%$ male). Studies were included if they had continued followup for at least 5000 person-years and had recorded vital status at the end of follow-up. Data on cigarette smoking were based on selfreport. | $751$ <br> colorectal cancers (454 men, 297 women) | Cigarette smoking <br> (Yes/No) <br> p value <br> Cigarette smoking (5/day) $p$ value |  |  |  | $\begin{aligned} & 1.43(1.09-1.88) \\ & 0.01 \\ & \\ & 1.00(0.92-1.09) \\ & 0.99 \end{aligned}$ | Adjusted for diabetes, BMI, and alcohol. |

\begin{tabular}{|c|c|c|c|c|c|c|c|c|}
\hline Reference, \& Cohort \& No. of \& Smoking \& No of cases \& \multicolumn{3}{|l|}{Relative risks (95\%CI or \(p\) value)} \& Adjustment \\
\hline \& \& \& \& \& Colon cancer \& Rectal cancer \& Colorectal cancer \& \\
\hline Paskett et al. (2007), USA, Women's Health Initiative (WHI) \& The WHI includes an observational study and three clinical trials. 146877 women. Clinical outcomes were reported semiannually for the clinical files and annually for the observation study. Follow-up 19982005. \& 1075 colon, 176 rectal cancer cases (461 right-sided and 296 left-sided) \& \begin{tabular}{l}
Former smokers Current smokers p trend \\
Age at smoking initiation \\
<20 \\
\(\geq 20\) \\
p value for trend Cigarettes/d \\
<25 \\
\(\geq 25\) \\
p trend \\
Duration of smoking \\
<20 \\
20-29 \\
30-39 \\
\(\geq 20\) \\
p trend \\
Age at smoking cessation \\
<30 \\
30-39 \\
40-49 \\
\(\geq 50\) \\
Current smoker \\
p trend \\
Time since \\
cessation \\
Current smoker \\
<10 \\
10-19 \\
20-29 \\
30-39 \\
\(\geq 50\) \\
p trend
\end{tabular} \& \& \(1.12(0.97-1.29)\)
\(1.03(0.77-1.38)\)
0.28
\(1.13(0.96-1.33)\)
\(1.08(0.91-1.29)\)
0.27
\(1.05(0.90-1.21)\)
\(1.47(1.16-1.85)\)
0.01

$0.95(0.79-1.15)$
$1.27(1.02-1.58)$
$1.18(0.93-1.50)$
$1.19(0.93-1.54)$
0.03
$0.95(0.72-1.27)$
$0.87(0.67-1.14)$
$1.24(0.98-1.56)$
$1.24(1.02-1.52)$
$1.04(0.78-1.39)$

0.06 \& $$
\begin{aligned}
& 1.15 \text { (0.80-1.67) } \\
& 1.95(1.10-3.47) \\
& 0.05 \\
& \\
& 1.14 \text { (0.75-1.75) } \\
& 1.39(0.91-2.10) \\
& 0.13 \\
& \\
& 1.29(0.90-1.86) \\
& 1.14 \text { (0.59-2.18) } \\
& 0.31 \\
& \\
& \\
& 0.87(0.52-1.43) \\
& 1.95(1.20-3.17) \\
& 1.24(0.68-2.27) \\
& 1.53(0.83-2.83) \\
& 0.05 \\
& \\
& \\
& 0.79(0.36-1.73) \\
& 0.84(0.42-1.70) \\
& 1.39(0.78-2.46) \\
& 1.53(0.93-2.52) \\
& 1.93(1.08-3.44) \\
& 0.01 \\
& \\
& 1.98(1.11-3.52) \\
& 1.81(0.77-4.26) \\
& 1.45(0.84-2.50) \\
& 1.27(0.71-2.28) \\
& 1.10(0.59-2.06) \\
& 0.53(0.19-1.46) \\
& 0.90 \\
& \hline
\end{aligned}
$$ \& \& Adjusted for age ethnicity, study arm, family history of colorectal cancer, total physical activity metabolic equivalents, duration of nonsteroidal anti-inflammatory drug use, alcohol, hormone therapy use, colonoscopy, history of diabetes, total dietary calcium, total dietary fibre, percent energy from fat, hemoglobin, waist circumference, red meat intake, and stratified by study (observational study, clinical trial nonhormone trial, hormone trial treatment assignment). <br>

\hline
\end{tabular}

\begin{tabular}{|c|c|c|c|c|c|c|c|c|c|c|}
\hline Reference, \& Cohort \& No. of \& Smoking \& \multicolumn{3}{|l|}{No of cases} \& \multicolumn{3}{|l|}{Relative risks (95\% CI or \(p\) value)} \& \multirow[t]{2}{*}{Adjustment factors/comments} \\
\hline \& \& \& \& C \& R \& \[
\begin{aligned}
\& \mathrm{CR} \\
\& \mathrm{C}
\end{aligned}
\] \& Colon cancer \& Rectal cancer \& Colorectal cancer \& \\
\hline \begin{tabular}{l}
Tsong et al. (2007), \\
Singapore, \\
Singapore Chinese \\
Health Study
\end{tabular} \& Citizens of Singapore who resided in government-built housing estates, 4574 years old, Hokkiens and Cantonese. 61,321 subjects. Baseline information collection by inperson interview. Linkage with the Singapore Cancer Registry and Singapore Registry of Births and Deaths. Follow-up 1993-2004 \& 845 incident cases (516 colon and 329 rectal) \& \begin{tabular}{l}
Never smokers \\
Former smokers \\
Current smokers \\
Cigarettes/day \\
<13 \\
13+ \\
p trend \\
Age at starting to smoke \\
15+ years \\
<15 years \\
p trend \\
Duration of smoking \\
<40 years \\
40+ years \\
p trend
\end{tabular} \& \[
\begin{array}{|c}
\hline 338 \\
75 \\
103 \\
\\
68 \\
110 \\
\\
\\
148 \\
30 \\
\\
\\
94 \\
84
\end{array}
\] \& \[
\begin{gathered}
157 \\
63 \\
109 \\
58 \\
114 \\
\\
\\
126 \\
46 \\
\\
\\
83 \\
89
\end{gathered}
\] \& \& 1.0
\(0.96(0.73-1.27)\)
\(0.83(0.64-1.06)\)
\(0.84(0.64-1.11)\)
\(0.91(0.71-1.17)\)
0.38

$0.89(0.71-1.12)$
$0.80(0.54-1.18)$
0.19
$0.88(0.68-1.14)$
$0.87(0.66-1.14)$

0.27 \& $$
\begin{aligned}
& 1.0 \\
& 1.45(1.04-2.01) \\
& 1.63(1.23-2.17) \\
& \\
& 1.38(0.99-1.90) \\
& 1.71(1.28-2.28) \\
& 0.0003 \\
& \\
& \\
& 1.40(1.07-1.84) \\
& 2.34(1.63-3.36) \\
& <0.0001 \\
& \\
& 1.37(1.01-1.84) \\
& 1.85(1.36-2.52) \\
& <0.0001
\end{aligned}
$$ \& \& Adjusted for age, gender, dialect group, year of recruitment, level of education, BMI, history of diabetes, family history of colorectal cancer, alcohol consumption, and physical exercise <br>

\hline Batty et al. (2008), UK, The Whitehall study \& 17322 London based government employees, aged 40-69 years, participated in a medical examination in the 1960s (response rate $74 \%$ ). $74 \%$ response Cancer mortality ascertained by using procedures of the National Health Service Central Registry until 2005 \& 52 colon cancer deaths, 16 rectum cancer deaths \& | Never |
| :--- |
| Former smokers |
| Current smokers |
| Former smokers |
| Effect per 10 |
| cigarettes/d |
| Effect per 10 |
| years of |
| smoking |
| Current smokers |
| Effect per 10 |
| cigarettes/d |
| Effect per 10 |
| years of |
| smoking | \& \[

$$
\begin{array}{|c|}
\hline 52 \\
118 \\
129
\end{array}
$$

\] \& \[

$$
\begin{aligned}
& 16 \\
& 58 \\
& 40
\end{aligned}
$$

\] \& \& \[

$$
\begin{aligned}
& 1.0 \\
& 1.11(0.80-1.55) \\
& 1.33(0.96-1.86) \\
& \\
& 1.03(0.88-1.22) \\
& 1.04(0.88-1.23) \\
& 1.05(0.87-1.27) \\
& 1.09(0.83-1.43)
\end{aligned}
$$

\] \& \[

$$
\begin{aligned}
& 1.0 \\
& 1.94(1.11-3.39) \\
& 1.51(0.84-2.74) \\
& \\
& 1.31(1.08-1.38) \\
& 1.13(0.87-1.45) \\
& \\
& 1.25(0.93-1.70) \\
& 1.26(0.77-2.05)
\end{aligned}
$$
\] \& \& Adjusted for age, employment grade, physical activity, BMI, marital status, systolic and diastolic blood pressure, cholesterol forced expiratory volume in 1s, height, impaired glucose tolerance, diabetes and disease at entry. <br>

\hline
\end{tabular}

\begin{tabular}{|c|c|c|c|c|c|c|c|c|}
\hline \multirow[t]{2}{*}{Reference, location, name of study} \& \multirow[t]{2}{*}{Cohort description} \& \multirow[t]{2}{*}{No. of subjects} \& \multirow[t]{2}{*}{Smoking categories} \& \multirow[t]{2}{*}{No of cases} \& \multicolumn{3}{|l|}{Relative risks (95\% CI or \(p\) value)} \& \multirow[t]{2}{*}{Adjustment factors/comments} \\
\hline \& \& \& \& \& Colon cancer \& Rectal cancer \& Colorectal cancer \& \\
\hline \begin{tabular}{l}
Kenfield et al. (2008), USA, \\
The Nurses Health Study
\end{tabular} \& Established 1976, 121700 female US registered nurses aged 30 up to 55 years, residing in 11 states. Baseline information obtained by mailed questionnaire. Deaths were usually reported by families and deaths among nonrespondents were identified by searching the National Death Index. Follow up 1980-2004. \& \begin{tabular}{l}
\[
578
\] \\
colorectal cancers deaths
\end{tabular} \& \begin{tabular}{l}
Never \\
Former smokers \\
Current smokers \\
Cigarettes/d \\
smoked by \\
current smokers \\
1-14 \\
15-24 \\
\(\geq 24\) \\
p trend \\
Starting age \\
among current \\
smokers, \\
\(\leq 35\) \\
18-21 \\
\(\geq 21\) \\
p trend \\
Years since \\
quitting in former \\
smokers, \\
<5 \\
5-<10 \\
\(10-<15\) \\
\(15-<20\) \\
\(\geq 26\) \\
p trend
\end{tabular} \& CRC
238
214
126

36
55

19
83
32

32
22
32
33

95 \& \& \& $$
\begin{aligned}
& 1.0 \\
& 1.23(1.02-1.49) \\
& 1.63(1.29-2.05) \\
& \\
& \\
& 1.37(0.95-1.96) \\
& 1.73(1.27-2.35) \\
& 1.83(1.26-2.64) \\
& 0.23 \\
& \\
& \\
& \\
& 1.25(0.77-2.02) \\
& 1.73(1.32-2.27) \\
& 1.55(1.01-2.39) \\
& 0.95 \\
& \\
& \\
& 0.87(0.59-1.29) \\
& 0.64(0.40-1.01) \\
& 0.96(0.65-1.43) \\
& 0.93(0.63-1.38) \\
& 0.70(0.53-0.93) \\
& 0.40
\end{aligned}
$$ \& Adjusted for age (months), follow-up period, history of hypertension, diabetes, high cholesterol levels, BMI, change in weight from age 18 years to baseline (1980), alcohol intake, physical activity, previous use of oral contraceptives, postmenopausal estrogen therapy use and menopausal status, parental history of myocardial infarction at age 65 years or younger and age at starting smoking, servings of beef, pork, lamb or processed meat, total calcium and folate intake, and duration of aspirin use. <br>

\hline
\end{tabular}

\begin{tabular}{|c|c|c|c|c|c|c|c|c|c|c|}
\hline Reference, \& Cohort \& No. of \& Smoking \& \multicolumn{3}{|l|}{No of cases} \& \multicolumn{3}{|l|}{Relative risks (95\% CI or \(p\) value)} \& Adjustment \\
\hline \& \& \& \& C \& R \& CRC \& Colon cancer \& Rectal cancer \& Colorectal cancer \& \\
\hline \begin{tabular}{l}
Gram et al. (2009), \\
Norway, The Norwegian Women and Cancer Study
\end{tabular} \& 68160 women aged 30-69 years who completed a questionnaire in 1996 or 1998. Follow-up by linkages to national registers through 2005. \& \begin{tabular}{l}
425 \\
incident \\
cases of histological confirmed primary invasive colorectal cancers, 284 colon (137 proximal, 108 distal) and 141 rectal cancer cases
\end{tabular} \& \begin{tabular}{l}
Never \\
Former smokers \\
Current smokers \\
Ever smokers \\
Smoking \\
initiation \\
\(\geq 20\) \\
\(<20\) \\
p trend \\
Cigarettes/d \\
1-9 \\
10-14 \\
\(\geq 15\) \\
p trend \\
Years smoked \\
1-19 \\
20-29 \\
\(\geq 30\) \\
p trend \\
Pack-years \\
smoked \\
0-9 \\
10-19 \\
\(\geq 20\) \\
p trend \\
Time since \\
quitting \\
smoking (years) \\
\(\geq 20\) \\
10-19 \\
1-9 \\
0 \\
p trend
\end{tabular} \& \[
\begin{gathered}
\hline 97 \\
107 \\
80 \\
187 \\
\\
98 \\
89 \\
\\
114 \\
53 \\
20 \\
\hline \\
55 \\
47 \\
85 \\
\hline
\end{gathered}
\] \& \begin{tabular}{l}
53 \\
40 \\
48 \\
88 \\
42 \\
46 \\
51 \\
28 \\
9 \\
23 \\
19 \\
46 \\
35 \\
28 \\
25 \\
16 \\
5 \\
13 \\
4
\end{tabular} \& \& 1.0
\(1.4(1.1-1.9)\)
\(1.1(0.8-1.6)\)
\(1.3(1.0-1.7)\)
\(1.3(1.0-1.7)\)
\(1.3(1.0-1.8)\)
0.05
\(1.3(1.0-1.7)\)
\(1.4(1.0-1.9)\)
\(1.2(0.7-2.0)\)
0.11
\(1.2(0.9-1.7)\)
\(1.4(1.0-2.0)\)
\(1.3(1.0-1.8)\)
0.07
\(1.1(0.8-1.5)\)
\(1.7(1.2-2.3)\)
\(1.2(0.8-1.8)\)
0.03 \& 1.0
\(0.9(0.6-1.4)\)
\(1.2(0.8-1.8)\)
\(1.1(0.7-1.5)\)

$1.0(0.6-1.5)$
$1.2(0.8-1.8)$
0.5
$1.0(0.7-1.5)$
$1.2(0.8-2.0)$
$0.9(0.4-1.9)$
0.7
$0.9(0.5-1.5)$
$0.9(0.5-1.6)$
$1.3(0.8-1.9)$
0.3
$0.9(0.6-1.3)$
$1.1(0.7-1.7)$
$1.5(0.9-2.5)$
0.13
$0.9(0.5-1.6)$
$1.1(0.6-2.1)$
$0.5(0.2-1.3)$
$1.2(0.8-1.8)$
0.5 \& \& Adjusted for age, menopausal status, hormonal contraceptive and postmenopausal hormonal therapy use, BMI and alcohol consumption, all at enrolment. <br>
\hline
\end{tabular}

|  | Cohort | No. of | Smoking | No of cases |  |  | Relative risks (95\% CI or $\boldsymbol{p}$ value) |  |  | Adjustment |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  |  | C | R | CRC | Colon cancer | Rectal cancer | Colorectal cancer |  |
| $\begin{aligned} & \text { Gram et al. } \\ & \text { (2009); } \\ & \text { (cont) } \end{aligned}$ |  |  | Proximal colon canc. <br> Never <br> Former smokers <br> Current smokers <br> Ever smokers <br> Distal colon <br> cancer <br> Never <br> Former smokers <br> Current smokers <br> Ever smokers | $\begin{aligned} & 44 \\ & 53 \\ & 40 \\ & 93 \\ & \\ & \hline 36 \\ & 46 \\ & 26 \\ & 72 \\ & \hline \end{aligned}$ |  |  | $\begin{aligned} & 1.0 \\ & 1.6(1.1-2.4) \\ & 1.4(0.9-2.1) \\ & 1.5(1.0-2.2) \\ & \\ & 1.0 \\ & 1.7(1.1-2.7) \\ & 1.0(0.6-1.6) \\ & 1.3(0.9-2.0) \end{aligned}$ |  |  | Adjusted for age, menopausal status, hormonal contraceptive and postmenopausal hormonal therapy use, BMI and alcohol consumption, all at enrolment. |
| Hannan et al. (2009), U.S. | Participants were drawn from the Cancer Prevention Study II Nutrition Cohort, a sub cohort of the CPS II mortality cohort, including residents in 21 states with population based state cancer registries and 50 to 74 years of age in 1992. Participants completed a mailed questionnaire. Follow-up questionnaires were sent in 1997, 1999, 2001, 2003, and 2005, with response rate among living | Selfreported cases were verified from medical records (n $=1227$ ) or through linkage to state cancer registries ( n $=422$ ). <br> Additional cases ( $\mathrm{n}=$ 313) were identified through linkage with the National Death | Never smokers Former smokers Current smokers Former smokers Age at cessation Before age 40 $40-49$ yrs of age $50-59$ yrs of age Age 60 or elder p trend Years since cessation $\geq 31$ yrs ago $21-30$ yrs ago $11-20$ yrs ago $1-10$ yrs ago p trend |  |  |  |  |  | 1.0 <br> 1.23(1.11-1.36) <br> 1.27(1.06-1.52) <br> 1.05(0.91-1.22) <br> 1.31(1.13-1.52) <br> 1.44(1.24-1.66) <br> 1.29(1.08-1.54) <br> 0.0014 <br> 1.03(0.89-1.19) <br> 1.28(1.10-1.49) <br> 1.33(1.14-1.55) <br> 1.48(1.27-1.73) 0.0003 | Adjusted for age , BMI, education, family history of colorectal cancer, physical activity, race, aspirin use, alcohol intake, Vegetable consumption, fibre/whole grain consumption, red and processed meat consumption, history of endoscopy |


| Reference, location, name of study | Cohort description | No. of subjects | Smoking categories | No of cases |  |  | Relative risks (95\% CI or $\boldsymbol{p}$ value) |  |  | Adjustment factors/comments |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  |  |  | CRC |  | Colon cancer | Rectal cancer | Colorectal cancer |  |
| Hannan et al. (2009); (cont) | participants at least $89 \%$. The follow up period ended on June 30, 2005. <br> 51365 men and 73386 women were included in the analysis. <br> Incident cases of colorectal cancer were identified by ICD-9 codes 153154.1 or by ICD 10 codes C18-C20 | Index (NDI). | Current smokers <br> Duration of smoking < 40 years 40-49 years 50+ years p trend |  | $\begin{aligned} & 29 \\ & 71 \\ & 56 \end{aligned}$ |  |  |  | $\begin{aligned} & 1.02(0.69-1.49) \\ & 1.32(1.02-1.72) \\ & 1.38(1.04-1.84) \\ & 0.052 \end{aligned}$ | Adjusted for age, BMI, education, family history of colorectal cancer, physical activity, race, aspirin use, alcohol intake, Vegetable consumption, fibre/whole grain consumption, red and processed meat consumption, history of endoscopy |
| Limsui et al.(2010) US | Among 37399 participants in a population-based cohort study (the Iowa Women's Health Study) | 1233 CRC | Never smokers <br> Former smokers Current smokers Ever smokers | C | R | CRC |  |  | $\begin{aligned} & 1.00 \\ & 1.16(1.00-1.35) \\ & 1.23(1.04-1.46) \\ & 1.19(1.05-1.35) \end{aligned}$ | Adjusted for age, BMI, waist-hip, physical activity, alcohol, exogenous estrogen, daily intake of total calories, fat, sucrose, red meat, calcium, vitamin E and methionine. |
| Leufkens et al.(2011), European Prospective Investigation into cancer and Nutrition (EPIC) | 465,879 <br> participants in the <br> European <br> Prospective <br> Investigation into <br> Cancer and <br> Nutrition (EPIC) <br> study. Mean <br> follow-up time was <br> 8.7 years | 1,791 colon 766 <br> proximal <br> tumors <br> 772 distal <br> tumors <br> 253 colon <br> tumors unspecified in location. 950 rectum cancer. <br> 2741 CRC | Never smokers <br> Former smokers Current smokers Ever smokers <br> Proximal colon cancer <br> Never smokers <br> Former smokers <br> Current smokers <br> Ever smokers | $\begin{aligned} & 746 \\ & 841 \\ & 556 \\ & 285 \\ & \\ & \\ & 303 \\ & 370 \\ & 239 \\ & 131 \end{aligned}$ | $\begin{aligned} & 378 \\ & 464 \\ & 306 \\ & 158 \end{aligned}$ | $\begin{aligned} & 1124 \\ & 1305 \\ & 862 \\ & 443 \end{aligned}$ | $\begin{aligned} & 1.0 \\ & 1.18(1.06-1.32) \\ & 1.21(1.08-1.36) \\ & 1.13(0.98-1.31) \\ & \\ & \\ & 1.0 \\ & 1.27(1.08-1.50) \\ & 1.25(1.04-1.50) \\ & 1.31(1.06-1.64) \end{aligned}$ | $\begin{aligned} & 1.0 \\ & 1.06(0.91-1.23) \\ & 1.10(0.94-1.30) \\ & 0.98(0.80-1.19) \end{aligned}$ | $\begin{aligned} & 1.0 \\ & 1.06(0.91-1.23) \\ & 1.10(0.94-1.30) \\ & 0.98(0.80-1.19) \end{aligned}$ | Adjusted for weight, height, physical activity, education, dietary intake of energy from fat, energy from nonfat, fiber, fruit, vegetables, red meat, processed meat, alcohol, and fish |


| Reference, location, name of study | Cohort description | No. of subjects | Smoking categories | No of cases |  |  | Relative risks (95\% CI or $\boldsymbol{p}$ value) |  |  | Adjustment factors/comments |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  |  | C | R | CRC | Colon cancer | Rectal cancer | Colorectal cancer |  |
| Leufkens et al.(2011) (cont) |  | cases. | Distal colon cancer Never smokers Former smokers Current smokers Ever smokers | $\begin{aligned} & 323 \\ & 358 \\ & 243 \\ & 115 \end{aligned}$ |  |  | $\begin{aligned} & 1.0 \\ & 1.05(0.89-1.24) \\ & 1.13(0.95-1.36) \\ & 0.91(0.73-1.14) \end{aligned}$ |  |  | Adjusted for weight, height, physical activity, education, dietary intake of energy from fat, energy from nonfat, fiber, fruit, vegetables, red meat, processed meat, alcohol, and fish |
| Parajuli et al. (2013), <br> Norway | 602, 242 men and women from four different Norwegian health surveys were followed. Mean follow up of 14 years. | 3998 colon <br> cancer <br> ( $46 \%$ in <br> women). <br> 2072 <br> proximal <br> colon <br> cancer <br> (47\% in <br> women) <br> and 1520 <br> distal colon <br> cancer <br> (43\% in <br> women) | Men <br> Never smokers <br> Former smokers <br> Current smokers <br> Ever smokers <br> Proximal colon <br> cancer <br> Never smokers <br> Former smokers <br> Current smokers <br> Ever smokers <br> Distal colon <br> cancer <br> Never smokers <br> Former smokers <br> Current smokers <br> Ever smokers <br> Women <br> Colon cancer <br> Never smokers <br> Former smokers <br> Current smokers <br> Ever smokers | $\begin{gathered} 834 \\ 355 \\ 657 \\ 1012 \end{gathered}$ |  |  | 1.0 <br> 1.14(1.02-1.27) <br> 1.03(0.92-1.15) <br> 1.08(0.97-1.19) <br> 1.0 <br> 1.06(0.90-1.24) <br> 1.02(0.86-1.19) <br> 1.03(0.90-1.19) <br> 1.0 <br> 1.24(1.03-1.47) <br> 0.95(0.79-1.13) <br> 1.08(0.92-1.26) <br> 1.0 <br> 1.16(1.02.1.31) <br> 1.22(1.10-1.36) <br> 1.19(1.09-1.32) |  |  | Adjusted for age, physical activity, BMI and education. |


| Reference, location, name of study | Cohort description | No. of subjects | Smoking categories | No of cases |  |  | Relative risks (95\%CI or $p$ value) |  |  | Adjustment factors/comments |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  |  | C | R | CRC | Colon cancer | Rectal cancer | Colorectal cancer |  |
| Parajuli et al. (2013) (cont) |  |  | Proximal colon cancer <br> Never smokers <br> Former smokers <br> Current smokers <br> Ever smokers <br> Distal colon <br> cancer <br> Never smokers <br> Former smokers <br> Current smokers <br> Ever smokers <br> Men <br> Colon cancer <br> Never smokers <br> Former smokers <br> Current smokers <br> Ever smokers <br> Proximal colon <br> cancer <br> Never smokers <br> Former smokers <br> Current smokers <br> Ever smokers <br> Distal colon <br> cancer <br> Never smokers <br> Former smokers <br> Current smokers <br> Ever smokers | $\begin{gathered} 438 \\ 186 \\ 362 \\ 548 \\ \\ 295 \\ 132 \\ 227 \\ 359 \\ \\ \hline 534 \\ 744 \\ 874 \\ 1,618 \\ \\ \hline 267 \\ 350 \\ 431 \\ 781 \\ \hline \end{gathered}$ |  |  | $\begin{aligned} & 1.0 \\ & 1.22(1.02-1.45) \\ & 1.37(1.18-1.59) \\ & 1.31(1.15-1.49) \\ & \\ & 1.0 \\ & 1.15(0.94-1.41) \\ & 1.12(0.93-1.34) \\ & 1.13(0.96-1.32) \\ & \\ & \\ & 1.0 \\ & 1.14(1.02-1.27) \\ & 1.03(0.92-1.15) \\ & 1.08(0.97-1.19) \\ & \\ & 1.0 \\ & 1.06(0.90-1.24) \\ & 1.02(0.86-1.19) \\ & 1.03(0.90-1.19) \\ & \\ & 1.0 \\ & 1.24(1.03-1.47) \\ & 0.95(0.79-1.13) \\ & 1.08(0.92-1.26) \end{aligned}$ |  |  | Adjusted for age, physical activity, BMI and education |

Source: International Agency on Research on Cancer (IARC) monograph 100 E 2012 for the cohorts until 2009


[^0]:    * Number of participants equals those who attended the survey and/or answered at least one questionnaire and signed a written consent. 7,460 persons participated in a second

    CONOR survey and 1 person participated in a third. Thus, the total numbers of participants with consent were 174,430.
    $\dagger$ The numbers include all individuals invited. The individual surveys could have published papers with slightly different total numbers.
    $\ddagger$ HUSK: All 40-44 years and those participating in a study in 1992-93 born 1950-51 and 1925-27; Oslo II: All those invited to the Oslo Study 1972-73, except those invited to HUBRO and MoRo I (Invited in 1972/73: all men born 1923-32 and 7\% random sample of those born 1933-52); Troms $\varnothing$ V: All 30, 40, 45, 60, 75 years and all those participating in phase II in Troms $ø$ IV - which included: all born 1920-1939, $5-10 \%$ sample of other age groups attending phase I, all women born 1940-44; I-HUBRO: 30\% random sample of people born in Pakistan, all born in Turkey, Sri Lanka, Iran, Vietnam - except those invited to HUBRO; MoRo II: All those participating in a study in 2 local districts in Oslo in 2000 (MoRo

[^1]:    ${ }^{1}$ Division of Epidemiology, Norwegian Institute of Public Health, Oslo, Norway.
    ${ }^{2}$ Institute of General Practice and Community Medicine, Medical Faculty, University of Oslo, Oslo, Norway.
    ${ }^{3}$ Institute of Health Management and Health Economics, Medical Faculty, University of Oslo, Oslo, Norway.
    ${ }^{4}$ Institute of Community Medicine, University of Tromsø, Tromsø, Norway.
    ${ }^{5}$ Department of Public Health and Primary Health Care, University of Bergen, Bergen, Norway.
    ${ }^{6}$ HUNT Research Centre, Institute of Community Medicine, NTNU, Verdal, Norway.
    ${ }^{7}$ Institute of Community Medicine, NTNU, Verdal,Norway.

    * Corresponding author. Norwegian Institute of Public Health, P.O. Box 4404 Nydalen, 0403 Oslo, Norway. E-mail: oena@fhi.no

[^2]:    ${ }^{\text {a }}$ Number of participants equals those who attended the survey and agreed that information from the CONOR survey and blood samples can be linked to other registers and used in research. A total of 7310 individuals participated in more than one survey. Thus, the total number of individuals equals 173236.

