Smoking and Mortality in the Netherlands:

The extent that variations in the COROP-regions for all-cause mortality can be attributed to smoking-related mortality in the period 2004-2008

Alette Sigrid Spriensma S1586467

Master thesis Population Studies

Supervisor: Dr. Fanny Janssen

Population Research Centre

Faculty of Spatial Sciences

University of Groningen

Groningen, March 2010

Acknowledgements

I would like to thank a number of people that have helped me in completing this Master thesis. First and foremost I would like express my respect to Dr. Fanny Janssen, who provided me with her good advice, assistance and support throughout my thesis and during the different courses she lectured. Furthermore I would like to thank Prof. Dr. Leo van Wissen, and Prof. Dr. Inge Hutter for the knowledge they shared during several courses of the Master Population Studies that helped me in writing my thesis. Also, I would like to express my gratitude to the Doodsoorzakenstatistiek of Statistics Netherlands for providing specific data on mortality, and Marinus de Bakker for his assistance in using ArcGIS. Finally I would like to thank my family, friends and classmates for their understanding and moral support.

Abstract

Objective: The aim of this research is to find out what the variations in mortality and smoking-related mortality are, and to explore to what extend regional differences in mortality can be attributed to smoking-related mortality in the different COROP-regions of the Netherlands in the period 2004-2008 for sexes. **Methods:** The cause-specific mortality data by age, sex, year, and region was provided by Doodsoorzakenstatistiek of Statistics Netherlands. The following methods were used in order to be able to get to the results: First, the smoking-related mortality was calculated from the lung cancer mortality, and then the age-standardization was applied for the cause-specific mortality rates. Following, the standardized data could be implemented into GIS. Thirdly the significance in proportional differences between a region and the average of the Netherlands were calculated. Also, spatial autocorrelation, the indexed variance, variance for rates, covariance, and correlation were calculated. **Results:** Oost-Groningen and Zuid-Limburg were the regions that very often belonged in the highest mortality rates for the different sexes and causes of mortality. The western part of the Netherlands overall showed very often lower rates in mortality for different causes and the southern and eastern part occasionally showed higher mortality rates. The smoking-related mortality rates for females showed a very distinct cluster of the low mortality rates that were located in the north of the Netherlands. The comparison of the all-cause mortality and the smoking-related mortality showed patterns that were the most alike for males. When smoking-related mortality was excluded from the indexed variance, the variance was substantially lower for males and females together, as well as separately. The correlation between smoking-related mortality and non-smoking related mortality for males and females, and males was positive, and significant, but for females there was no correlation to be found. **Conclusion:** Concluding it can be stated that smoking-related mortality has an influence on the variations in all-cause mortality. The results indicate that there is still a lot that can be done to reduce the smoking-related mortality in influencing the all-cause mortality, especially for males. There are also other causes in mortality that are not included in the smoking-related mortality rates that have a strong influence on the all-cause mortality rates.

Keywords: All-cause mortality, smoking-related mortality, COROP-regions, mortality rates, GIS

Table of contents

List of figures

List of tables

1. Introduction

1.1 Background

Within the Netherlands, an unhealthy lifestyle has a big influence on the total disease burden and on mortality. Smoking is the most important lifestyle that causes diseases and mortality, 13 % of the total disease burden in the Netherlands can be attributed to the use of tobacco. Overweight caused by too much food intake and/or not enough exercise accounts for 10 % of the total disease burden. Excessive alcohol use causes 4.5 % of the total disease burden (VTV, 2006). Smoking also scores weak when the Netherlands is compared to other countries in the European Union, the Netherlands belongs to the group of countries that have the highest percentage of daily smokers (Zantinge, 2009). In 2005, almost 20,000 people died in the Netherlands because of smoking (Gelder et al., 2007).

The epidemiologic transition theory of Omran is related to health as well as mortality. Omran (1998) describes five stages in the western transition model that countries go through in the process of modernisation. The stages of the epidemiologic transition are based on broad categories of cause-specific mortality. (Wolleswinkel-van den Bosch et. al., 1997). The first is the stage of "pestilence and famine; the second stage is the "age of receding pandemics; the third is the "age of degenerative and man-made diseases"; the fourth age is characterised by "declining cardiovascular mortality, ageing, lifestyles modification, emerging and resurgent diseases; The last stage is a futuristic age in which there is a "aspired quality of life, with paradoxal longevity and persistent inequities" (Omran, 1998, p.102). According to Wolleswinkel-van den Bosch et al. (1997) the Netherlands is at the fourth stage of the epidemiologic transition: "age of declining cardiovascular mortality, ageing, lifestyles modification, emerging and resurging diseases".

A very distinctive feature of this fourth stage is the levelling off, and then the decrease of cardiovascular mortality. Reasons for the decline in cardiovascular mortality are modifications in lifestyle. Even though there is a decline in mortality from cardiovascular diseases and cancers, they will continue to be the leading causes of death, because they have predominance over other diseases (Omran, 1998). This predominance of cardiovascular diseases and cancers means that there is still a lot of room for more modification in lifestyle. Lifestyle can be seen as an exogenous determinant. A determinant is neither positive nor negative; this depends on how it is used. Mostly the focus is on the negative health threatening factors. (Ruwaard and Kramers, 1993)

A study done by van der Wilk and Jansen (2004) showed that the gap in lifestyle-related risk factors in Europe has become smaller in the past 30 to 40 years, but that there should be alertness towards intranational variations in lifestyle. Regional differences in lifestyle are expected to overtake the international differences in lifestyle. According to Kunst et al. (1990a) lifestyles differ for people from different social classes, this for example means that people with a lower social-economic status consummate more tobacco. In the regions of The Netherlands there is a difference in social status, so this should lead to variations in smoking-related mortality. They also found that the total mortality in regions of The Netherlands were significantly higher in the regions with a lower socio-economic level. As expected from what was explained above, there is a variation in the standardized number of deaths per region within the Netherlands. The COROP-region of Oost-Groningen has the highest death rate for both males and females; this is 10 deaths per 1,000 of the population for males and 10.2 for females. The COROP-region of Het Gooi and Vechtstreek has the lowest death rate of 7.2 deaths per 1,000 of the population for males, for females the region of Noord-Drenthe has the lowest death rate of 7.5 deaths per 1,000 of the population (Statistics Netherlands, 2010).

What would be interesting is to find out what the variations and patterns in all-cause mortality and smoking-related mortality in the Netherlands are on a regional level, and to what extent the smoking-related mortality can be attributable to the all-cause mortality. Recent research has been done by Janssen et al. (2007) on mortality decline in seven European countries, which included the Netherlands. In this study they distinguished between smoking-related mortality and non-smoking-related mortality. On a regional level however, there has been little research. Kunst et al. have done research about determinants of regional differences in lung cancer mortality in The Netherlands, and Kunst and Mackenbach have also looked at regional lung cancer death rates in relationship to smoking; both these researches date from 1993.

1.2 Relevance

Societal: Of interest for the government of The Netherlands. Deaths separated to cause, can be used to discover if there is more or less of a relationship between smoking-related mortality and all-cause mortality in the different regions of the Netherlands. The government can undertake actions to lower the deaths in the regions that have more deaths in relationship to smoking. This is possible because smoking is seen as a lifestyle, and lifestyle in its turn is a possible modifiable effect.

Scientific: Research on smoking-related mortality has mostly been done on a national level. Two other researches in the Netherlands on a regional level have been done by Kunst et al., but this research dates from 1990 and 1993. The research of 1990 looked at socio-economic factors whereas this research will look at the lifestyle of smoking. The research of 1993 is about determinants of regional differences in lung cancer mortality in the Netherlands. This research will be related to these other researches, but will have newer data available and look at the factor of variations in smoking-related mortality that influences the variations in all-cause mortality. The reason why it is important to look at a regional level is that regions, just a much as countries, have their own identity. The smaller your classifications, the more differences you get (Pater et al., 2002).

1.3 Objective

The aim of this research is to find out what the variations in all-cause mortality and smoking-related mortality are, and to explore to what extent regional differences in all-cause mortality can be attributed to smoking-related mortality in the different COROP-regions of the Netherlands in the period 2004-2008 for sexes.

1.4 Research questions

The main question of this research is:
<u>What are the variations and patterns in all-cause mortality and smoking-related mortality in the different COROP-regions of the Netherlands for sexes in the period of 2004-2008, and to what extent can regional differences in mortality be attributed to smoking?</u>

The research questions that can be derived from the main question are:

1. What are the variations and patterns in all-cause mortality in different regions of the Netherlands, for sexes in the period of 2004-2008?

2.a What are the variations and patterns in the top three smoking-related causes of death in different regions of the Netherlands for sexes in the period of 2004-2008?

2.b What are the variations and patterns in smoking-related mortality in different regions of the Netherlands for sexes in the period of 2004-2008?

3. To what extent can regional differences in mortality in the Netherlands be attributed by smoking in the period of 2004-2008?

1.4 Approach

The all-cause mortality and the smoking-related mortality in the period of 2004-2008 by sex and region are studied in this research. The top three smoking-related causes of death are also looked at, because they show from what the smoking-related mortality exists. The reason for the period of five years is to get a more reliable image of the results. Because there are more males than females that smoke (Gelder et al., 2007), there also should be different variations. This research looks at the males and females together, and separately.

The most important theory used in this research is the epidemiologic transition theory of Omran. All the stages have broad indications on causes of mortality; but since the Netherlands is at the fourth stage according to Wolleswinkel-van den Bosch et al. (1997), this is the most important stage for this research. In this stage modifications in lifestyle are of a great influence in the decrease of cardiovascular mortality and deaths because of cancer (Omran 1998). The conceptual base model of public health contains information on health, prevention and care in a region within the Netherlands and therefore contains important information how the Netherlands deals with public health. Mackenbach et al. (1990) say that even though the mean number of deaths is low, there are distinct differences inside the Netherlands on a regional level. The results of the study done by Janssen et al. (2007) show that smoking seems to be more important than other factors originating earlier in life.

The cause-specific mortality data by age, sex, year, and region was provided by Doodsoorzaakstatistiek of the CBS. The 40 different COROP-regions should give an indication if there is a certain variation or pattern for all-cause mortality and smoking-related mortality within the Netherlands. The following methods are going to be used in order to be able to get to the results: First, the smoking-related mortality will be calculated from the lung cancer mortality, this can be done by means of a simpler version of the indirect Peto-Lopez method that was created by Janssen et al. (2007). Then, the age-standardization will be applied for the cause-specific mortality rates. Following, the standardized data could be implemented into GIS in order to find out if there are any variations or patterns. Fourthly the significance in proportional differences between a region and the average of the Netherlands is calculated. Also, spatial autocorrelation to show clustering or dispersing and the variance, covariance and correlation will be calculated.

1.5 Structure of thesis

In order to be able to answer the research questions that were stated above, this thesis is build up in different chapters. The following chapter consists from the theory that was used in this research, the theoretical framework consists from the main theories that were related to this research, followed by the Public Health Status and Forecasts, then there will be a short explanation on how smoking influences the mortality, related research on national levels and on a regional on smoking-related mortality, and finally the conceptual model. Chapter three firstly describes the data that was used and were it was obtained from; secondly the methods used to convert the data into the results were explained into detail. The chapter following this explains the results by means of maps, tables and analysis of the results. The conclusion gives a short overview of the results in chapter four with, followed by the discussion, finishing with recommendations.

2. Theory

In this chapter the already existing theories and research on smoking-related mortality will be reviewed. The demographic transition theory will be discussed shortly, because this theory can be seen as the base for the epidemiologic transition theory. The epidemiologic theory is the main theory that is used for this thesis and there already has been given a brief summary in the introduction. Following are the determinants in mortality, and the link between smoking and mortality is explained further. Finally, previous research on smoking-related mortality will also be explained, first on a national level and then followed by the regional level.

2.1 Theoretical framework

2.1.1 Demographic transition theory

The demographic transition theory can essentially be stated as: "Societies that experience modernisation progress from a pre-modern regime of high fertility and high mortality to a post-modern one in which both are low". (Kirk, 1996, p. 361). Warren Thompson was the first to make classifications in populations that have different population sizes because of mortality and fertility, and even though Notestein was not the first to make the essential classifications for the demographic transition theory, his are accepted as the classical theory. He made the following statements:

- The populations of Western and Central Europe would peak around 1950 and decline after that, for Southern Europe the date was set around 1970.
- A big decline in fertility
- The world population in 2000 will be 3.3 billion

Looking back on Notestein's statements, he has overestimated the decline in fertility and underestimated the world population. The greatest strength of the demographic transition theory is that the transition will occur in every country that is undergoing modernisation. The weakness is that it is hard to make estimates on the precise threshold for fertility to drop. (Kirk, 1996)

In the modern world, three stages in historical mortality decline can be distinguished. The first stage is that the most obvious; decrease was in the late part of the eighteenth century and in the first half of the nineteenth century. It is likely that the increasing incomes, better nutrition, improvements in hygiene have contributed to the decrease in mortality. The second stage occurred at the last third of the nineteenth century and lasted until the First World War In this time there were revolutionary discoveries in medicine by for example Pasteur and Koch. In this stage there was a lot of reduction of child mortality and infant mortality. The last and third stage started during World War Two. The discovery of penicillin by Fleming introduced a great use in antibiotics. Because of the use of antibiotics, there was a great reduction in epidemic and contagious diseases. (Kirk, 1996)

2.1.2 Epidemiologic transition theory

The original epidemiologic transition theory was created by Omran in 1971. This theory describes how epidemiologic changes in health can be attributable to certain determinants, or how health correlates in different societal settings. Mortality is a very important variable in the epidemiologic transition. The epidemiologic transition theory of 1971 first started of as having three different stages: The age of pestilence and famine; the age of receding pandemics; and the age of degenerative, stress, and man-made diseases. (Omran, 1998)

Because changes in the epidemiologic transition theory are not evenly distributed across time and between populations, the epidemiologic transition theory of Omran has been revised. On one side there is the western transition model, on the other side there is the non-western transition model. This research will focus on the classical western transition model, because the Netherlands can be classified under this model. Omran (1998, pp. 102) uses five different stages in the western transition model:

1. "age of pestilence and famine"
   This is characterised by high mortality, high fertility, and slow population growth
   Mortality had lots of fluctuations with peaks that correspond with the epidemics. Life expectancy was between 20 and 30+. Heath care was provided by local systems, just a few of them are still used today. Fertility was high, but with the young age at death, the consequence was a young population. There were low living standards and the environment was unsanitary.
2. "age of receding pandemics"
   Mortality began to decline, and life expectancy at birth increased 40 to 50 years. Infant mortality also declined. Communicable diseases were still the leading cause of death. Fertility remained high through most of this stage; this led to a rapid population growth. More in the end of this stage, the fertility also declined. Access to health care was not available for everyone. Better housing resulted in small improvements in living conditions.
3. "age of degenerative, stress, and man-made diseases"
   Is characterised by the increasing prevalence of heart diseases, strokes, cancer, diabetes, chronic obstructive pulmonary disease and metabolic disorders. Also man-made diseases increased, these diseases were introduced by man. Mortality was still declining and life expectancy rose to 50 to 75+ years. Health care became wide spread, organised on a local and national level. Living conditions and sanitation are significantly improved.
4. "age of declining cardiovascular mortality, ageing, lifestyles modification, emerging and resurging diseases"
   This stage is characterised by further increases in life expectancy. Also the levelling off, and then decline in cardiovascular deaths is a feature of this stage. There are three influences in this decline. First there is the deliberate modifications in lifestyle, second is the influence of medical breakthroughs, the third influence is the treatment of risk conditions. In spite of the decline in deaths from cardiovascular diseases and some cancers, they still are the leading causes of death. Also there was a turnout of new diseases and a comeback of old diseases. Health care systems continue to improve, fertility continues to be low, and the standard of living is high.
5. "age of aspired quality of life, with paradoxical longevity and (futuristic stage) persistent inequities"

The epidemiologic transition theory looks at different stages that occur over time. All the stages have broad indications on causes of mortality. The Netherlands is at the fourth stage according to Wolleswinkel-van den Bosch et al. (1997), in this stage modifications in lifestyle are of a great influence in the decrease of cardiovascular mortality and deaths because of cancer (Omran 1998). That the Netherlands is in the fourth stage means that only the fourth stage of the epidemiologic transition theory will be used because this research is focused on a period and thus cross-sectional in nature.

2.2 Health determinants and the relation between smoking and mortality

2.2.1 Health determinants

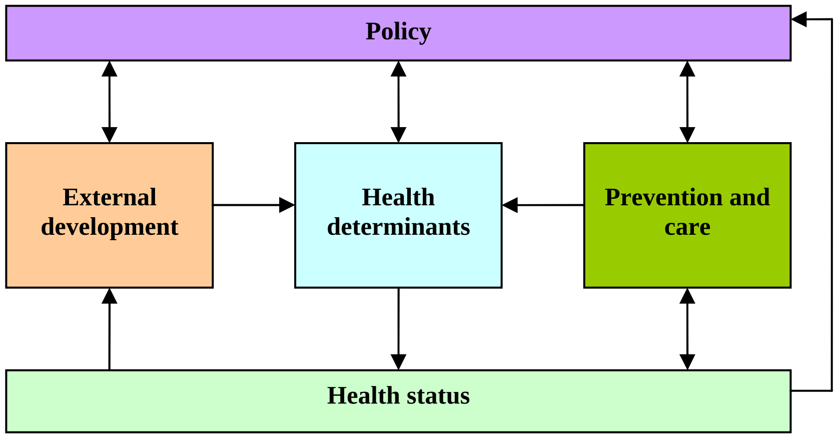
Young (1998) discussed different major health determinants that can influence each other in order to create health or diseases within individuals and populations. He distinguished between: Genetic susceptibility; Physical environment; Personal lifestyles and behaviours; and Social, cultural and economic factors. When it comes to genetic susceptibility, genes can be seen as the basis of heredity. There are different disease levels of the genetic involvement: Single-gene diseases, Chromosomal disorders, and multifactorial or polygenic diseases. Most diseases have some level of interaction between genetics and the environment. Mostly though, the genetic susceptibility determinant is discussed as an opposite to the environmental determinant. The physical environment includes everything that is outside the body of an individual. There are different factors in the environment that can have a negative influence on health: The nature of the hazard; their source; where it occurs; the site of exposure; and the route of the exposure. These factors can lead to either acute or chronic effects in health; the chronic effects are more difficult to reach, and thus of greater concern. There are many diseases or problems in health that can be associated with a lifestyle or behaviour. (Young, 1998) For example, a large part of the mortality caused by diseases such as cardiovascular diseases, cancers, chronic respiratory diseases and diabetes (that are the four non-communicable diseases that cause the highest mortality world wide), could be prevented by eliminating risk factors in lifestyle (WHO, 2008). Important examples of such lifestyles are: Smoking, food intake, alcohol and drug use, physical activity, sexual behaviour, and safety practices (Young, 1998). A lifestyle is an exogenous determinant. A determinant on itself is not a positive thing, nor a negative thing; this depends on how it is used, for example: food intake can be good on one hand if it is healthy food, but it can be bad when there is overconsumption of fat and unhealthy food. Mostly, the focus is on the negative health threatening factors. Smoking is the most important risk factor in terms of mortality and contributes strongly to lung cancer, coronary heart diseases, stroke, and respiratory diseases (Ruwaard and Kramers, 1993). Just as for the physical environment, there are some determinants in lifestyle that can take some time span to cause a (chronic) disease. This is for example the case for lung cancer and smoking; it takes on average 20 to 30 years to develop. The long time span means that the lifestyle pattern of 20 to 30 years ago is portrayed in the disease pattern of today (VTV, 2006). Differences in smoking patterns vary between countries; within a country however, there are also differences: between regions, ethnicity, age, sex, and socioeconomic status. Overall, smoking seems to be more present among the population groups that have a lower education as well as a low income. (Young 1998) The relationship between the social, cultural, and economic factors and health has been apparent for over a long period. The socioeconomic status can be referred to as 'a hierarchical continuum according to prestige and lifestyles based largely on educational and occupational achievements' (Young, 1998: p120). Social networks and social support can be seen as an independent outcome in diseases. Also, culture has an effect on health; this can be through different manners of cultural beliefs and different practices (Young, 1998).

Health determinants can be seen as an explanation as to why the health status is what it is. There have been different attempts to try and find a relationship between health determinants and health status/health care. Young (1998) states that the working document of the Canadian government 'A New perspective on the Health of Canadians' containing 'the health field concept' is still very useful for health care planners, practitioners, and administrators. The focus of the health field concept is mostly on issues in relation to the delivery of services in health care (Young, 1998). An example of using a model that is based on the health field concept is the one that the Netherlands uses. The Netherlands uses the Public Health Status and Forecasts (VTV) model in order to find out what the health situation on a regional level is. The model is thus a development of 'the health field concept model' that was created by the Canadian minister Marc Lalonde. The VTV model contains the following concepts: Policy; External development; Health determinants; Prevention and care; and Health status, as can be seen in box 2.1 (Schrijvers, 2007). The VTV model contains the concept of Health determinants. According to Schrijvers (2007) the health determinants within the VTV model can be divided into: personal factors, lifestyle, factors in the physical environment, and social factors.

The model does differ from the Canadian concept, where Human Biology, Environment, Lifestyle, and Health Care Organisation were the used concepts (Health and Welfare Canada, 1974).

Box 2.1: Public Health Status and Forecasts (VTV) model

The conceptual base model of public health



Source: VTV, 2006

2.2.2 The relation between smoking and mortality

On average, the life expectancy worldwide has been increasing over time. That the life expectancy is still increasing in developed countries is mostly due to the fact that mortality among adults can be reduced further (WHO, 2003). Janssen et al. (2003) found that the Netherlands does not follow this average pattern of life expectancy increasing over time. The mortality declines that were apparent in the 1970s did not follow in the 1980s and 1990s. Mortality decline stagnated in these periods and even showed some increases in mortality and thus a stagnation or reduction in life expectancy. Smoking-related diseases were found to be one of the contributing causes in the stagnation of mortality, but they could not fully explain the stagnation in mortality among the elderly in the Netherlands.

When looking at smoking more closely, there is one well known substance in tobacco smoke that has a bad influence on the health of a person: nicotine. Smoking is addictive because of nicotine (Young, 1998). In total, there are about 4000 different substances in tobacco smoke, 40 of these substances are known to cause cancer (US Department of health and Human services, 1989). People that smoke passively also have a higher risk on getting cancer. In tobacco smoke there are a lot of carcinogenic substances. The carcinogenic substances can create changes in the mucous membrane of the bronchial tubes, which leads to lung cancer. The chance on getting lung cancer increases with: the amount of cigarettes smoked; the amount of years that someone smokes; and the age when a person starts smoking (Zandwijk and Leeuwen, 2005).

Ochsner and DeBakey (1939) linked the increased cases of lung cancer in the first part of the twentieth century to smoking and were among the first to do so. They stated the following: “In our opinion the increase in smoking with the universal custom of inhaling is probably a responsible factor, as the inhaled smoke, constantly repeated over a long period of time undoubtedly is a source of chronic irritation to the bronchial mucosa” (Ochsner and DeBakey, 1939, p. 435). Further early research on smoking as an important determinant in lung cancer mortality and other mortality causes was conducted by Doll and Hill (1956). They send out questionnaires on smoking in the United Kingdom to all the people that had a medical profession at that time. From the questionnaires, different groups were created to indicate how much someone smoked. In each of the created groups, the mortality that took place was recorded as well. The analysis of the questionnaires showed that as the lung cancer mortality rate increased, so did the frequency in smoking. Also, there was a great reduction in mortality the longer someone had stopped smoking. No large differences were found between residences from cities or residences from smaller places; this means that air pollution was not a contributing factor in the lung cancer mortality. There was only one other form of cancer that showed association between mortality and smoking, that was cancer of the upper respiratory and upper digestive tracts. Coronary thrombosis also showed a significant relation to smoking, but this relation was small and was only apparent at the youngest ages of 35-54. There were three other causes of mortality that increased as the prevalence in smoking increased: Chronic bronchitis, peptic ulcer, and pulmonary tuberculosis. (Doll and Hill, 1956)

Smoking can thus been seen as the most important determinant of lung cancer. In about 85 percent of all lung cancer cases, smoking is the cause. Smoking also very much increases the risk on getting larynx cancer, chronic obstructive pulmonary disease (COPD), oral cavity cancer/cancer of the throat, and oesophagus cancer. Other diseases that smoking has a known influence on are: ischaemic heart disease, cerebro vascular accident (CVA), and heart failure. Table 2.1 shows the diseases that occur more often among smokers than non-smokers for males and females separately. The diseases are expressed through means of the Relative Risk (RR). The RR varies with age between the highest and the lowest RR-number. It can be seen that the 'comments column' that is written in the column next to the RR-numbers, has some diseases that have a certain age were the RR is highest, or that the RR lowers as age increases. When looking at the RR numbers for Males, the chance on getting lung cancer has a RR of 11.9-29.3; this means that the chance of getting lung cancer is 12 to 29 times higher for males that smoke, than for males that do not smoke. (Gelder et al., 2007)

Table 2.1 Diseases that occur more often for smokers than for non-smokers, expressed through relative risk, separately for males and females

Diseases for which smoking is a risk factor	**Males**		**Females**	
	Relative risk (bi)	Comments	Relative risk (bi)	Comments
Lung cancer	11,9 - 29,3	Highest RR at 60-64	7,9 - 16,3	Highest RR at 45-49
COPD	3,1 - 13,7	Highest RR at 70-74	2,3 - 9,1	Highest RR at 65-69
Oesophageal cancer	2,6 - 8,5	RR Lowers with age	2,6 - 8,5	RR Lowers with age
Laryngeal cancer	11,6		11,6	
Oral cavity cancer and cancer of the throat	3,9 - 7,4	Highest RR at 55-59	3,9 - 7,4	Highest RR at 55-59
Ischaemic heart disease	1,3 - 4,5	RR Lowers with age	1,1 - 4,6	RR Lowers with age
Heart failure	1,3 - 1,7	RR Lowers with age	1,3 - 1,7	RR Lowers with age
Stroke (Cerebro Vascular Accident)	1,1 - 3,5	RR Lowers with age	1,0 - 3,7	RR Lowers with age
Cancer of the bladder	1,7 - 2,7	Highest RR at 55-59	1,7 - 2,7	Highest RR at 55-59
Stomach cancer	1,0 - 1,5	Highest RR at 55-59	1,0 - 1,5	Highest RR at 55-59
Kidney cancer	1,5 - 1,6	Highest RR at 50-59	1,5 - 1,6	Highest RR at 50-59
Pancreatic cancer	1,2 - 2,5	RR Lowers with age	1,2 - 2,5	RR Lowers with age
Diabetes mellitus type 2 [a]	1,15		1,15	

Source: This table is based on Surgeon General, 2004. Revised by the RIVM; Van Baal et al., 2006c; Hoogenveen et al., 2007

[a] The conclusion of this relation is based on: Patja et al., 2005

The Population attributive risk (PAR) can calculate how much loss in health can be attributable to an unhealthy lifestyle, or in this specific case smoking. The PAR is based on the prevalence of the determinant in a population and in most cases the relative risk (RR); this is a measurement for the relation between the determinant and a disease. (VTV, 2002)

Some facts on smoking

Within the European Union there are differences in smoking between countries. Sweden has the lowest percentage in smoking (16%) and Greece the highest percentage smokers (38%). Younger people tend to smoke more than older generations. Also, males tend to smoke more than females, even though the differences between sexes are becoming smaller. The smoking-related mortality for males and females is declining in the western European countries. (Kaiser and Gommer, 2007)

The percentage of male smokers in The Netherlands has also been declining in the period of 1958-1991 from 90 percent to 38 percent. Since the eighties the decline started getting smaller. For females the percentage increased until in the seventies, after that a small decline started for females over the age of twenty. The percentage of people who have stopped smoking is higher in the older age groups. (Ruwaard and Kramers, 1993) Also for the Netherlands, there are more males than females that smoke. In 2005, almost 20,000 people died in The Netherlands because of smoking (Gelder et al., 2007). Table 2.1 shows the eight most important smoking-related mortality causes. Even though ischaemic heart disease, cerebro vascular disease, and heart failure have a smaller risk of being caused by smoking, than oesophagus cancer, larynx cancer, and oral cavity cancer/cancer of the throat, more people in the Netherlands die of these diseases.

Table 2.2 Smoking-related mortality in the Netherlands by the eight most important causes for people of 20 years and older in 2007, also by percentage

Mortality cause	**Males**		**Females**		**Total males and females**	
Lung cancer	5.830	9,0%	2.491	3,7%	8.320	6,3%
COPD	3.128	4,8%	1.864	2,7%	4.992	3,8%
Ischaemic Heart diseaese	1.876	2,9%	645	0,9%	2.521	1,9%
Oesophageal cancer	854	1,3%	250	0,4%	1.104	0,8%
Heart failure	765	1,2%	583	0,9%	1.348	1,0%
Cerebro Vascular Accident (CVA)	428	0,7%	216	0,3%	644	0,5%
Oral cavity cancer/cancer of the throat	319	0,5%	101	0,1%	419	0,3%
Laryngeal cancer	133	0,2%	39	0,1%	173	0,1%
Total smoking-related mortality	13.332	20,6%	6.189	9,1%	19.521	14,7%
All-cause mortality	64.797	100,0%	68.225	100,0%	133.022	100,0%

Source: CBS Doodsoorzakenstatistiek, revised by RIVM, 2007
Percentages were calculated by adding also the all-cause mortality of 2007 from Statline, 2010

2.3 Literature review

Now that the general overlapping theories, the determinants in health, and the relation between smoking and mortality have been explained, this subsection will show some of the newer research that already has been conducted on smoking in relation to mortality. First the related research on smoking and mortality on a national level will be discussed. Secondly, the related research on smoking and mortality on a regional level.

2.3.1 Related research on smoking and mortality on national levels

Using data on mortality and the consummation of tobacco

Previous research has mostly been done by using data on (cause-specific) mortality and comparing this to data on the consumption of tobacco. The following researches show the results that came from both mortality- and tobacco consumption data.

The relationship between smoking and adult mortality at the national level in the United States was examined by Rogers et al. (2005). By using data from a representative sample, detailed measures of smoking status and age, and information on mortality from all causes of death, they have estimated smoking–attributable deaths. The research also looked at factors that confound the relationship between smoking and mortality such as excess alcohol intake, lack in exercise, and a lower socio-economic status. The results showed that the influence of these other factors is modest. There are some limitations to their results because the reports on smoking status do not completely show the patterns in smoking of the individual through life. They concluded that is still a lot of room to lower the death rates that can be related to smoking, this is even more so the case for males. (Rogers et al., 2005)

Hummer et al. (1998), also did research on a national level in the United States on adult mortality differentials associated with smoking. They found that smoking behaviour is often measured by classifying individuals as 'smokers' or 'non-smokers'. These two categories are broad indications when looking at behaviour that can vary through life, that is why they have specified five categories in smoking behaviour. They concluded that the five-group smoking classification gave interesting results. As they had expected, smoking was related to higher mortality rates for both sexes, all age groups, and for most underlying causes of death such as lung cancer. Even though their expectations were true, there were still other patterns found within this general pattern. Long-term former smokers showed mortality rates that were much more similar to people that never smoked, than to current heavy smokers. Estimates also show that light smokers show more similarity to never smokers than to current heavy smokers. Smokers have shown a high circulatory disease mortality rate at relative young ages in comparison to never smokers when looking at the underlying cause of death. Smoking still is a key variable for understanding the mortality differences by gender. When the estimated mortality rates of females never smoking were compared to male heavy smoking, the differences in mortality were large. The comparison of estimated mortality for males never smoking to females heavy smoking also gave differences in mortality, but not as severe as the other way around. (Hummer et al. 1998)

A study designed to estimate the mortality and morbidity attributable to amongst others tobacco, was done by Single et al. (1999) in Canada. Smoking-related lung cancer accounted for the largest amount of deaths in 1992 when compared to other smoking-related mortality. When comparing the smoking-related mortality to former studies in Canada, the estimates in this research are much lower. The reasons for the lower estimates on mortality that is related to tobacco are largely due to use of pooled estimates of relative risk. Even though the estimates are lower in this research, they still indicate that tobacco use is still a big source of disease and mortality in Canada. (Single et al., 1999)

In Germany, John and Hanke (2003) calculated the burden on public health that is caused by tobacco smokers and alcohol use. Their research looks at mortality as well as potential life years lost. In order to calculate the smoking-related mortality, they used a formula that was developed by Schultz et al. The formula consists from calculating the tobacco-attributable fraction (TAF), the tobacco-attributable mortality (TAM) and the summed tobacco-attributable mortality (TAM). The limitations in the data have probably led to an underestimation in the impact of substance-use on mortality. Still, the results of this study show a clear overlap in mortality that can be attributable to tobacco and alcohol use. (John and Hanke, 2003)

<u>Using mortality data only</u>
All the studies mentioned above, have used data on mortality and data on the consummation of tobacco. The age-sex-specific smoking histories are very often hard to get by, or not available at all. The two researches below only use data on mortality. They have used the method that is called the Peto-Lopez method (Peto et al., 1992).

The research of Peto et al. (1992) provides estimates for early middle age, later middle age, and old middle age mortality that are caused by tobacco use in developed countries. They used the absolute lung cancer rate in a certain population to be able to indicate the proportions of the deaths from other diseases that attribute to smoking indirectly. The advantage of using this method is that only the national age-sex-specific mortality rates from various causes are needed. In all the populations that were researched when looking at sex, the highest smoking-related mortality could be attributed to males. The results in their study show that about half of the deaths that are caused by tobacco are in the middle age group. This shows that tobacco is a very important cause of premature death. (Peto et al., 1992)

Janssen et al. (2007) have done research on old-age mortality decline in seven European countries. Their focus is on rates of mortality change, instead of absolute mortality levels. In this study they distinguish between smoking-related mortality and non-smoking mortality. In order to estimate the smoking-related mortality a simpler version of the indirect Peto-Lopez method is used. By multiplying all-cause mortality with the etiological fraction, they estimated the level of smoking-related mortality. They found that old-age mortality decline was much bigger when the smoking-related mortality was not included. The results of this study show that there are variations between countries, periods and sexes, and that smoking seems to be more important in the pace of old-age mortality decline than other factors that originate earlier in life. (Janssen et al., 1997)

From national to a regional level

A study on lifestyle differences in Europe tried to explain whether the gap in lifestyle-related risk factors in Europe has become smaller in the previous 30 to 40 years. Their results indicated that the gap of lifestyle differences on a national level has become smaller over the past 30 to 40 years. Smoking has declined mostly for women and not for men. Although the trends of risk factors in variations in lifestyle are converging on a national level, there also needs to be alertness towards intranational variations in lifestyle. Regional differences are expected to overtake the international differences in lifestyle. The high prevalence in risk factors of lifestyle in combination with increasing differences in socio-economic health, indicate that prevention plans should be renewed in order to realize health benefits. (van der Wilk and Jansen, 2004)

2.3.2 Related research on regional mortality differences and smoking

The last part of the previous subparagraph showed that the gap on lifestyle-related risk factors between countries may become smaller, but that the regional differences may overtake the national ones. There are not many studies that have focused on differences between regions and mortality that can be ascribed to a certain lifestyle such as smoking, but there has been research done by Mackenbach et al. (1990), Balarajan and McDowall (1988), Kunst and Mackenbach (1993) and will be described further in this subparagraph.

Mackenbach et al. (1990) say that even though the mean number of deaths is low, there are distinct differences inside The Netherlands on a regional level. Especially the south of The Netherlands is dominated by high standardized mortality ratios. Even though the differences in mortality have gotten smaller over time, they are still big enough to do research about. In this research they say that there are three different possible causes that may explain differences in regional mortality, but there focus is mainly on the prevalence of causal determinants in mortality for the different regions. Religion, average income, and the degree of urbanisation seem to be population characteristics that are unbound to specific regions. The Netherlands differs from other countries because the most important factor in explaining the high mortality in the southern regions is through religion. This is mostly because of the higher smoking prevalence for Catholics. Socio-economic level is another factor that has an influence on the mortality differences between regions. Mackenbach et al. (1990) find that smoking is a very important reason for the regional mortality differences in the Netherlands. They say that when trying to reduce the differences in mortality between the regions, the prevention should be especially focussed on the regions that have a relative high amount of smokers.

In Great Britain, a study on regional socioeconomic differences in mortality among men aged from 20 to 64 years was done by Balarajan and McDowall (1988). They found that the smoking level habits increased from the south-east to the north-west, and the smoking-related mortality also showed a south-east to north-west gradient but the pattern was slightly different. The groupings of the regions were not into much detail, but the analysis did show that there is regional inequality in socio-economic mortality. (Balarajan and McDowall, 1988)

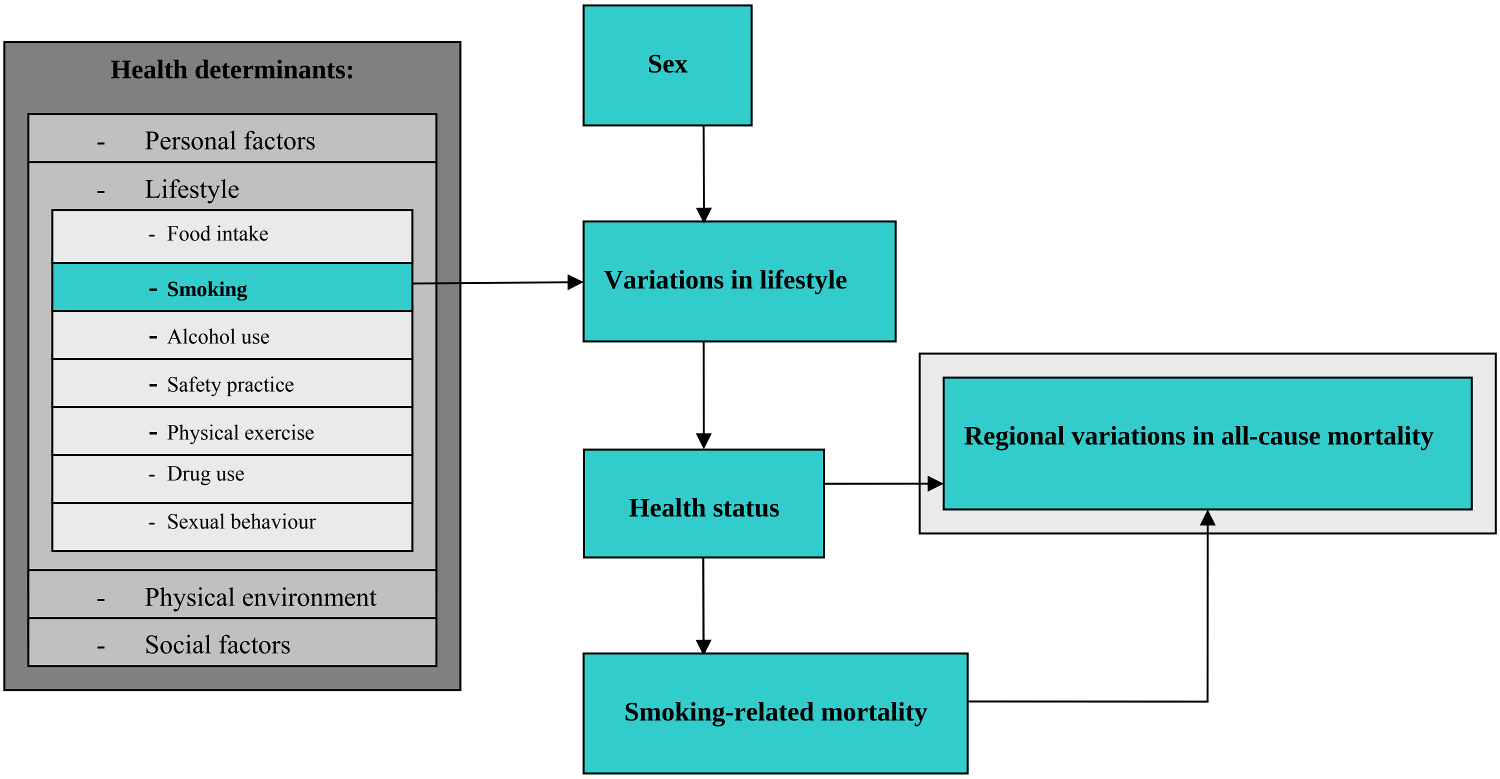
Kunst and Mackenbach (1993) tried to explain regional differences in lung cancer mortality in the period of 1980-1984 in the Netherlands by means of data on past smoking. They found that for women a large part of the regional variation in lung cancer mortality could be explained by the regional differences in cigarette consumption in 1972, for men this was not the case. When looking at the lung cancer mortality for men that were 75 years and above, there was a strong relationship with the 1930 tobacco consumption. Because of the limited availability of the data, it would be expected that with more precise data on regional tobacco consumption, a bigger part of the regional lung cancer mortality could be explained. Regional variations in lung cancer mortality may also be caused by other factors than tobacco consumption. Most areas that have high lung cancer mortality are often also heavily urbanised and industrialised. Nevertheless, not only current smoking, but also life-time exposure to smoking should be taken into account. (Kunst and Mackenbach, 1993)

Kunst et al. (1993) also have done research about determinants of regional differences in lung cancer mortality in The Netherlands. The reason for doing this study was that in various other countries on a regional scale there was only a weak relationship between smoking tobacco and getting lung cancer. In most other countries there is a regional variation in death rates by lung cancer. Differences in lung cancer mortality are likely to be caused by tobacco consumption, but a large share is still unexplained. The period of 1980 to 1984 is used to give an idea on how to explain the regional differences. 39 different COROP-regions were used in the data. There were five different age groups. Regional mortality patterns are strongly determined by cohort effects, they vary between birth cohorts. There is diffusion in low- and high income regions in lung cancer; this indicates that there is a relationship in cigarette smoking and regional differences.

2.4 Conceptual model

In this paragraph the conceptual model will be discussed. The conceptual model is needed in order to show the main concepts and their interrelations, this can be seen in Figure 2.1. The aim of this research is to find out what the variations in mortality and smoking-related mortality are, and to explore to what extend regional differences in mortality can be attributed to smoking in the different regions of the Netherlands in the period 2004-2008 between sexes.

Figure 2.1: Conceptual model



The conceptual model in figure 2.1 shows the main concepts that have been derived from the theory in the previous paragraphs. Health determinants are the most important factors that can influence health. The factors can be divided into: personal factors, lifestyle, factors in the physical environment, and social factors (Schrijvers, 2007). The epidemiologic transition theory of Omran (1998) is not present in the model, but functions as a background for the indication that the Netherlands is at the fourth stage in the epidemiologic transition, and therefore modifications in lifestyle are of great influence on health status and mortality (Wolleswinkel-van den et al., 1997; Omran, 1998). Lifestyle can be divided into: food intake, smoking, alcohol use, physical exercise, drug use, and sexual behaviour (VTV 2010). Of all the different lifestyles, smoking is the most important risk factor in terms of mortality (Ruwaard and Kramers, 1993) and influences the variations in lifestyle. Variations in lifestyle are also influenced by sex, because there are differences between males and females when it comes to smoking. Mostly the males show higher estimations of mortality in relation to smoking (Rogers et al., 2005; Hummer et al., 1998; Peto et al., 1992). Lifestyle can be seen as a behavioural factor and can have an influence on health status (Schrijvers, 2007). The concept of Health status influences the regional variations in all-cause mortality in a region either directly, or trough smoking-related mortality.

2.4.1 Hypothesis

As Mackenbach et al. (1990) have discovered in there research, there are distinct differences inside the Netherlands on a regional level and that the south of the Netherlands has higher mortality ratios. This leads to believe that there are variations to be found within the different COROP-regions of the Netherlands and also possible patterns. Hypothesis one to four are about the variations and patterns. If there is a relationship between smoking-related mortality and all-cause mortality, they should also show the same variations and patterns. The results of the study of Janssen et al. (2007) showed that smoking seemed to be an important factor in mortality and when it was excluded, the mortality decline was higher. The seventh hypothesis is expected to show that smoking has an influence on the all-cause mortality.

1. There are regional variations in all-cause mortality in the different regions of the Netherlands

2.a There are regional variations in the top three smoking-related causes of death in the different regions of the Netherlands

2.b There are regional variations in smoking-related mortality in the different regions of the Netherlands

3. There are significant variations in all-cause mortality, top three smoking-related causes of death, and smoking-related mortality in the different regions of the Netherlands
4. Patterns/clusters will be found in all-cause mortality, top three smoking-related causes of death, and smoking-related mortality
5. The same variations in smoking-related mortality rates and all-cause mortality rates will be found
6. Clustering and randomness have the same pattern for smoking-related mortality as for all-cause mortality
7. The variance for non-smoking-related mortality will be lower than the variance for all-cause mortality; in other words, smoking-related mortality contributes to the regional variations in mortality in the Netherlands
8. The covariance and the correlation should show that smoking-related mortality has an influence on all-cause mortality

3. Data and methodology

3.1 Study design

The main question of this research is:

What are the variations and patterns in all-cause mortality and smoking-related mortality in the different regions of the Netherlands between sexes in the period of 2004-2008, and to what extend can regional differences in mortality be attributed to smoking-related mortality?

The research questions that can be derived from the main question are:

1. What are the variations and patterns in all-cause mortality in different regions of the Netherlands, for sexes in the period of 2004-2008?

2.a What are the variations and patterns in mortality of the top three smoking-related causes of death in different regions of the Netherlands for sexes in the period of 2004-2008?

2.b What are the variations and patterns in smoking-related mortality in different regions of the Netherlands for sexes in the period of 2004-2008?

3. To what extent can regional differences in all-cause mortality in the Netherlands be attributed by smoking-related mortality in the period of 2004-2008?

Paragraph 3.2 of this chapter describes the data that was used for this research. It explains where the data came from, the accuracy, what data actually was needed to answer the research questions, and the used level of analysis. The methodology is explained in paragraph 3.3. First the data needed some adjusting. The smoking-related mortality was calculated from the lung cancer mortality, and then standardization was applied for all-cause mortality, the top three smoking-related causes of death, and smoking-related mortality. Secondly the standardized data could be implemented into GIS. Thirdly the significance in proportional differences between a region and the average of the Netherlands were calculated. Also, spatial autocorrelation, the variance, covariance and correlation were calculated.

3.2 Data

In this paragraph, the data will be described. Doodsoorzakenstatistiek of Statistics Netherlands has provided the data that was used for this research. The Doodsoorzakenstatitiek collects data on causes of mortality for all the inhabitants of the Netherlands that have past away since 1901. Their main goal is to obtain information on causes of death for all the inhabitants of the Netherlands. The information collected is used to make CBS-publications, and to give information to researchers under strict circumstances. The population taken into account is the group that has past away and was registered in the Gemeentelijke Basisadministratie Persoonsgegevens (GBA). Everybody that legally stays in the Netherlands for over four months will be taken into the GBA of their municipality. (CBS, 2010)

The accuracy of the data collection in the Netherlands is high; in 2008, 98.6 percent of the deaths in the Netherlands could be linked to a certain cause of death. A big part of the missing percentage is because of the Dutch population dying in other countries. The causes of death get certain codes that originate from the "International Statistical Classification of Diseases and Related Health Problems. The physician fills in a declaration for the cause of death. When this declaration is unclear or incomplete, contact is made with the physician to find out what the reason is behind it. (CBS, 2010)

In order to be able to answer all the research questions, the data on all-cause mortality, and the top three diseases that can lead to smoking-related mortality for age and sex by COROP-region in the period of 2004 to 2008 for the Netherlands were needed. As was shown in table 2.2, the top three diseases that cause the most smoking-related mortality in the Netherlands are: lung cancer, chronic obstructive pulmonary disease and ischaemic heart disease. Table 3.1 shows the codes of the International Classification of Diseases and Related Health Problems of the World Health Organization (WHO) for the different causes of mortality in the 10th revision (WHO, 2007).

Table 3.1 Codes of the International Classification of Diseases and Related Health Problems for different causes of mortality

Mortality cause	ICD 10
• All-cause mortality	A00-Z99
• Lung cancer mortality	C33-C34
• Chronic obstructive pulmonary disease mortality	J40-J47
• Coronary/Ischaemic heart disease mortality	I20-I25

Source: World Health Organization, 2007

Because the short list of cause-specific mortality data on Statline of Statistics Netherlands does not contain all the age categories that were needed to be able to make the smoking-related mortality calculations for this research, the all-cause mortality and the lung cancer mortality were provided by Doodsoorzakenstatistiek of Statistics Netherlands in the different age categories: 0-39, 40-44, 45-49, 50-54, 55-59, 60-64, 65-69, 70-74, 75-79, 80-84, 85-90, 90 years and older. The other data from Statline Statistics Netherlands was available in the age categories: 0-49, 50-59, 60-64, 65-69, 70-74, 75-79, 80-84, 85-89, 90 years and older. The reason for the use of data in five years as a period is because of the small amount of mortality that can occur on the regional COROP level. Calculations will lead to a more reliable result by taking a five-year period as the average.

Level of analysis

In the Netherlands there are three different levels of classification that are part of the hierarchal Nomenclature of Territorial Units for Statistics (NUTS):

- NUTS I Country parts
- NUTS II Provinces
- NUTS III COROP-regions

NUTS Region: Article 1: "The purpose of this Regulation is to establish a common statistical classification of territorial units, hereinafter referred to as "NUTS", in order to enable the collection, compilation and dissemination of harmonised regional statistics in the Community" (European Commission 2003, pp. 131). The smallest possible level for the cause-specific mortality data of the Netherlands was NUTS level three. The COROP classification was designed around the 1970's by the Coördinatie Commissie Regionaal Onderzoeksprogramma. The Netherlands has been divided into 40 different areas, and is a classification between the municipalities and the provinces (Giesbers, 2005). In table 3.2 all the different COROP-regions can be seen, whereas figure 3.1 shows where the specific regions are located.

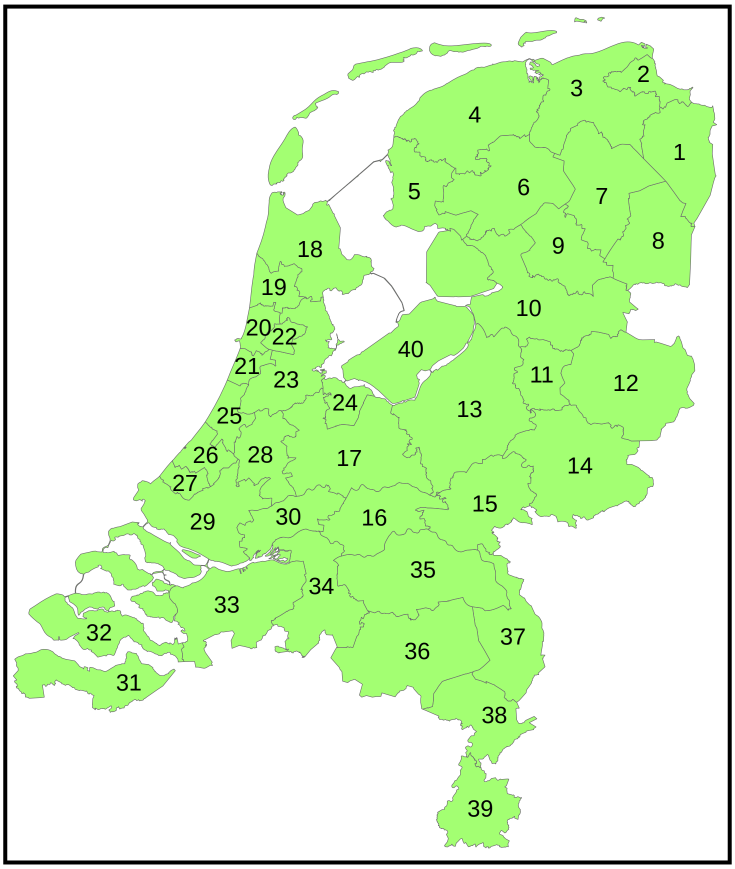
Table 3.2 The 40 COROP-regions of the Netherlands

1	Oost-Groningen (CR)	21	Agglomeratie Haarlem (CR)
2	Delfzijl en omgeving (CR)	22	Zaanstreek (CR)
3	Overig Groningen (CR)	23	Groot-Amsterdam (CR)
4	Noord-Friesland (CR)	24	Het Gooi en Vechtstreek (CR)
5	Zuidwest-Friesland (CR)	25	Agglomeratie Leiden en Bollenstreek (CR)
6	Zuidoost-Friesland (CR)	26	Agglomeratie 's-Gravenhage (CR)
7	Noord-Drenthe (CR)	27	Delft en Westland (CR)
8	Zuidoost-Drenthe (CR)	28	Oost-Zuid-Holland (CR)
9	Zuidwest-Drenthe (CR)	29	Groot-Rijnmond (CR)
10	Noord-Overijssel (CR)	30	Zuidoost-Zuid-Holland (CR)
11	Zuidwest-Overijssel (CR)	31	Zeeuwsch-Vlaanderen (CR)
12	Twente (CR)	32	Overig Zeeland (CR)
13	Veluwe (CR)	33	West-Noord-Brabant (CR)
14	Achterhoek (CR)	34	Midden-Noord-Brabant (CR)
15	Arnhem / Nijmegen (CR)	35	Noordoost-Noord-Brabant (CR)

16	Zuidwest-Gelderland (CR)	36	Zuidoost-Noord-Brabant (CR)
17	Utrecht (CR)	37	Noord-Limburg (CR)
18	Kop van Noord-Holland (CR)	38	Midden-Limburg (CR)
19	Alkmaar en omgeving (CR)	39	Zuid-Limburg (CR)
20	IJmond (CR)	40	Flevoland (CR)

Doodsoorzakenstatistiek Statistics Netherlands from obtained data (2009)

Figure 3.1 The 40 different COROP-regions of the Netherlands by number



3.2.1 Ethical issues in relation to the obtained data

Specific data on Lung cancer mortality and the all-cause mortality were obtained from Doodsoorzakenstatistiek of Statistics Netherlands. The data should be handled with care, because of the confidentiality. This research is conducted on a regional level, and needed different age groups. The mortality numbers in certain regions could be low and in that way traceable. This research shows the mortality data in a period over five years and for all the ages together in order to not harm confidentiality of the data.

3.3 Methodology

Smoking-related- and non-smoking-related mortality
To calculate the smoking-related mortality, a slightly different version of the indirect Peto-lopez technique was used here; this different version was created by Janssen et al. (2007). The method uses the observed lung cancer mortality rates in a certain time by age and sex to estimate the percentage of the population that has smoked at one point in time. For this research the lung cancer mortality rates by age and sex in the period of 2004-2008 for the different COROP-regions where used. It is taken into account that some people that die of lung cancer have never smoked. With smoothed age- and sex specific relative risks (RR) of mortality that have as result the exposure to smoking, the etiologic fractions where calculated next. The relative risks originally came from the "American Cancer Society cancer prevention study-II and they also have been corrected because smoking is related to other dangerous ways of life. The etiologic fraction is the share of the all-cause mortality that is caused by smoking-related mortality. By multiplying the etiologic fraction with the all-cause mortality, you get the smoking-related mortality. Non-smoking-related mortality was calculated by means of the difference between the all-cause mortality and the smoking-related mortality. The etiologic fractions of 80 years and older were applied to al the different five-year age categories that were above 80 years. With the estimation of smoking-related mortality, the percentage of exposure of above 100% is equal to 100%. With the estimation of the etiologic fraction, the maximum level of the relative risks is applied only in the end, when the exposure is set to 100%

Standardization
Because populations have compositional differences, the crude rates cannot be compared to each other without some form of standardization. There are two different methods that can be applied to standardize, these are the direct- and the indirect standardization methods. When using indirect standardization, the total number of events (D^A), and the age-specific mid-year population (N_i^A) are known. The age-specific rates for the standard population (M_i^S) are needed. The requirements for direct standardization data are: Age-specific mid-year population (N_i^A) and age-specific events (D_i^A). (Preston, 2001)

The direct standardization method has been used in this research, since both the age-specific mid-year population and the age-specific events are known. Figure 3.2 shows the calculation of the Crude Death Rate and the Age-Standardized Crude Death Rate. When looking at the second equation, M stands for death rate, C for proportion in age category, and $_i$ for the age-group. The third equation shows two new letters, the ^{j} stands for region, and the ^{s} for the standard population. By multiplying the observed age-specific rates for different subpopulations by the age-structure of a standard population you get the age-standardised crude death rate. Preston et al. (2001) state that when many populations are compared to each other, a standard should be used that is close to the mean or median of population structures in the population that is being researched. In this case the average population by age for males and females, males, and females of the Netherlands was used as the standard population, in order compare between the COROP-regions.

Figure 3.2 Calculation of age-standardized death rate

$$CDR = \frac{D}{N} = \frac{\sum_{x=0}^{\infty} {}_nD_x}{N} = \frac{\sum\limits_{x=0}^{\infty} \frac{{}_nD_x}{{}_nN_x} {}_nN_x}{N} = \sum_{x=0}^{\infty} \frac{{}_nD_x}{{}_nN_x} \cdot \frac{{}_nN_x}{N} = \sum_{x=0}^{\infty} {}_nM_x \cdot {}_nC_x$$

$$CDR = \sum_{i=1}^{\infty} M_i \cdot C_i$$

$$ASCDR^j = \sum_{i=1}^{\infty} M_i^j \cdot C_i^s$$

Source: Preston et al. 2001 pp. 23

The significance that is different from the average

In order to find out if the mortality in a COROP region is significantly different from the average of the Netherlands, the difference between proportions first needed to be calculated. The equation can be seen in figure 3.3, where P1 and P2 were the standardised mortality rates for the Netherlands and a specific region and n1 and n2 the populations of the Netherlands and a specific region.

Figure 3.3 Difference between proportions

$$\sigma_{p1\ -\ p2} = \text{sqrt}\{\ [P_1 * (1 - P_1) / n_1] + [P_2 * (1 - P_2) / n_2]\ \}$$

Source: StatTrek, 2009

In order to find out if there are significant differences in the variations of mortality, a standard score, or z-score had be calculated. According to Norusis (2002) you can determine the position of a specific case in the distribution of observed values by calculating the standard score. The standard score shows you how many standard deviation units a specific case is above or below the mean. From the standard deviation, the Z-score was looked up. A Z-score of above 1.962 has a confidence interval of 95 % and the z-score larger than 2.578 has a confidence interval of 99 %.

Mapping the results

The results of the calculations have been mapped in the program called ArcGIS (Geographical information systems) in order to discover variations and patterns. The projected coordinate system used is RD_new, and the geographic coordinate system name is GCS_Amersfoort. To make the maps, the calculations that were made were joined with a layer that had the boundaries of all the different COROP regions. The variations were mapped in five different classes of categories. With variations, the numbers that are different from the norm are meant. The classes were created manually because the features in the maps are compared to a specific meaningful value. This meaningful value is the average for the total of the Netherlands. The average value of the Netherlands was thus taken as a middle value and from there the other classification groups were created that fitted the specific range of the data.

To join table data to the layers data of COROP regions:

- Go to layer properties > click joins & relates > in the join box click add
- Choose join attributes from a table
- Field the join in layer will be based on > COROP
- Upload the table containing the data on mortality
- Field the join in table will be based on > COROP
- Keep all records
- Click OK

To make a map with different classifications:

- Go to layer properties
- Choose symbology
- In the fields box set the value to the cause of mortality by sex you want to know
- Create classes manually
- Click OK

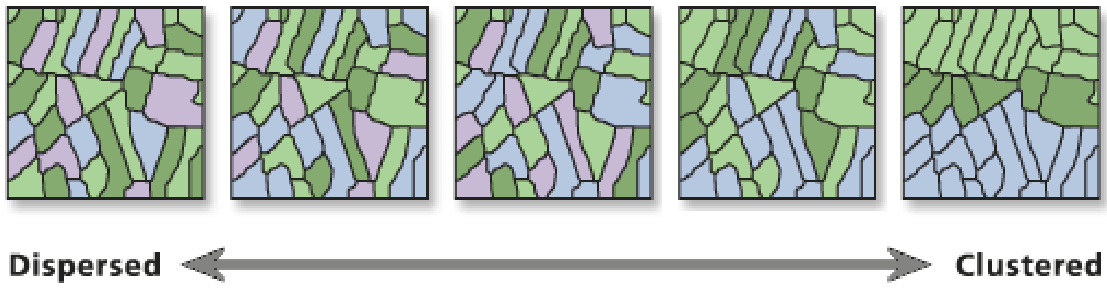
Analyzing a pattern

There are two statistics that are commonly used to find patterns for features that have continuous data: One is called the Geary's contiguity ratio and the other is the Moran's Index. The Geary's c uses the difference in values between two neighbouring features; the Moran's I calculates the difference between the target feature and the mean for all the features, and also the difference between each neighbour and the mean. The differences are then compared for the target feature and each neighbour in turn. (Mitchell, 2005)

This research has used the Moran's I to find patterns for the different causes of mortality. The values of the Moran's I can run from -1 to 1. If the index has a value of -1, the values are distributed dispersedly; high and low values are located next to each other. An index value of 1 indicates that there is complete clustering of the values; Low values are located next to each other and high values are located next to one another. The middle value (0) means that the pattern is totally random. (Mitchell, 2005)

In order to discover if regions of high incidence cluster together or if they are randomly distributed, a spatial autocorrelation can be used (Gatrell, 2002). Spatial autocorrelation, otherwise known as global Moran's I, indicates the clustering both by location and value. The Moran's I index shows if data is random, clustered, or dispersed, as can be seen in figure 3.3. Where the global Moran's I proved to be clustered a local Moran's was also obtained by means of maps showing were the clustering was located. (Allen, 2009)

Figure 3.4 Spatial autocorrelation: from dispersed to clustered



Source: ArcGIS Desktop 9.3 Help

In order to run the spatial autocorrelation tool (Global Moran's I), the following steps were taken:

1. In ArcMap open ArcToolbox, expand Spatial Statistics Tools, also expand Analyzing Patterns and double-click the Spatial Autocorrelation tool
2. Input Feature Class = Spatial autocorrelation moransi
3. Input Field = The sex- and mortality specific field
4. Check Display Output Geographically box = Yes
5. Conceptualization of spatial relationships = Inverse Distance
6. Distance Method = Euclidean
7. Standardization = None
8. Distance Band or Threshold Distance (optional) = none
9. Weights Matrix File (optional) = none

In step 5, the Inverse Distance was chosen as the conceptualization of spatial relationships, this means that the spatial relationships decay over distance. Every feature has an influence on another feature, but the more distance between the features, the less the impact. Combining the Inverse distance with the Euclidean distance method in step 6 is appropriate for modelling continuous data like mortality. Standardization has already been done, so it is not needed again in step 7. Because no distance band was filled in for step 8, a default distance will be computed. The default distance is the minimum distance to make sure that at least one neighbour is assigned to every feature. (ArcGIS Desktop 9.3 Help, 2009)

In order to run the local moran's I, the following steps were taken:

1. In ArcMap open ArcToolbox, expand Spatial Statistics Tools, also expand Mapping Clusters and double-click the Cluster and Outlier Analysis (Anselin Local Moran's I) tool
2. Input Feature Class = Spatial autocorrelation moransi
3. Input Field = The sex- and mortality specific field
4. Output Feature Class = localmoransi: sex- and mortality specified
5. Conceptualization of spatial relationships = Inverse Distance
6. Distance Method = Euclidean

7. Standardization = None
8. Distance Band or Threshold Distance (optional) = none
9. Weights Matrix File (optional) = none

<u>The variance, covariance, and correlation</u>
The measure of variability that is used most commonly is the variance. Figure 3.5 shows the formula for computing the variance. The variance is based on the squared distances between the values of the individual cases (X) and the mean (μ), where the different COROP-regions are the individual cases. The squared distance can be calculated by subtracting the mean from the value after which the difference can be squared. To calculate the variance, the sum of the squared distances has to be divided by the number of cases (N). When the variance has a value of zero, all the cases have the same value. The higher the variance, the more the values differ from each other. (Norusis, 2002)

Figure 3.5 Formula for computing the variance

$$\sigma^2 = \frac{\sum (X - \mu)^2}{N}$$

Source: Victoria University, 2010

Another way of calculating the variance is through the command "VARP" in excel =VARP( ) The data were the variance needs to be calculated over has to be selected between the brackets. This can also be done when calculating the covariance = COVAR( ) and the correlation = CORREL ( ), for these two functions two datasets have to be selected between the brackets in order to get results.

The Pearson's correlation coefficient is a frequently used method in order to find out to which extent the increase in one variable result in a proportional increase of the second variable. For this research the correlation between smoking-related mortality and non-smoking-related mortality were calculated. The covariance shows if a positive, negative, or no relation (0) exists between the two variables. The value of the correlation coefficient can vary between -1 and 1. A negative correlation value shows that as one variable increases the other decreases, positive correlation on the other hand means that as one variable increases, so does the other. (Mitchell, 2005)

In order to be able to compare the variances of all-cause mortality, smoking-related mortality, non-smoking-related mortality, lung cancer mortality, COPD mortality, and ischaemic heart disease mortality to each other, a measure is needed to give them the same average. This average can be accomplished by calculating the Comparative Mortality Figures (CMF's) from the mortality rates. The CMF's shows how the mortality in a total population is in proportion to a subpopulation. The mortality in the total population has a CMF of 100. A higher CMF indicates that the mortality in a subpopulation is higher than the total population, and a lower CMF indicates a lower mortality in the subpopulation (Zwakhals et al., 2009).

4. Results

In this chapter, the results will be discussed trough showing different maps that were created by means of a GIS (Geographical Information system). First the variations and patterns in all-cause mortality in the different COROP-regions of the Netherlands will be shown, secondly the mortality variations and patterns in the top-three smoking-related causes of death, third is the variation and patterns in smoking-related mortality, and last, the extent that regional mortality differences are attributed by smoking.

The maps for mortality rates per 10,000 inhabitants are shown in the first part of paragraph 4.1, 4.2, and 4.3. The second part of paragraph 4.1, 4.2, and 4.3 show the significance that a region can differ from the average of the Netherlands. In order to find out if the patterns that were found in the previous maps are significantly clustered or dispersed, the results of the Moran's I will be discussed in the last parts of the just mentioned paragraphs. The global Moran's I can indicate if there is a clustering, randomness, or dispersing of the regions. The expected index value of the Moran's I for a random distribution is -0.025641. The null hypothesis for this analysis is: The mortality rate of a COROP-region is randomly distributed across the Netherlands.

4.1 All-cause mortality

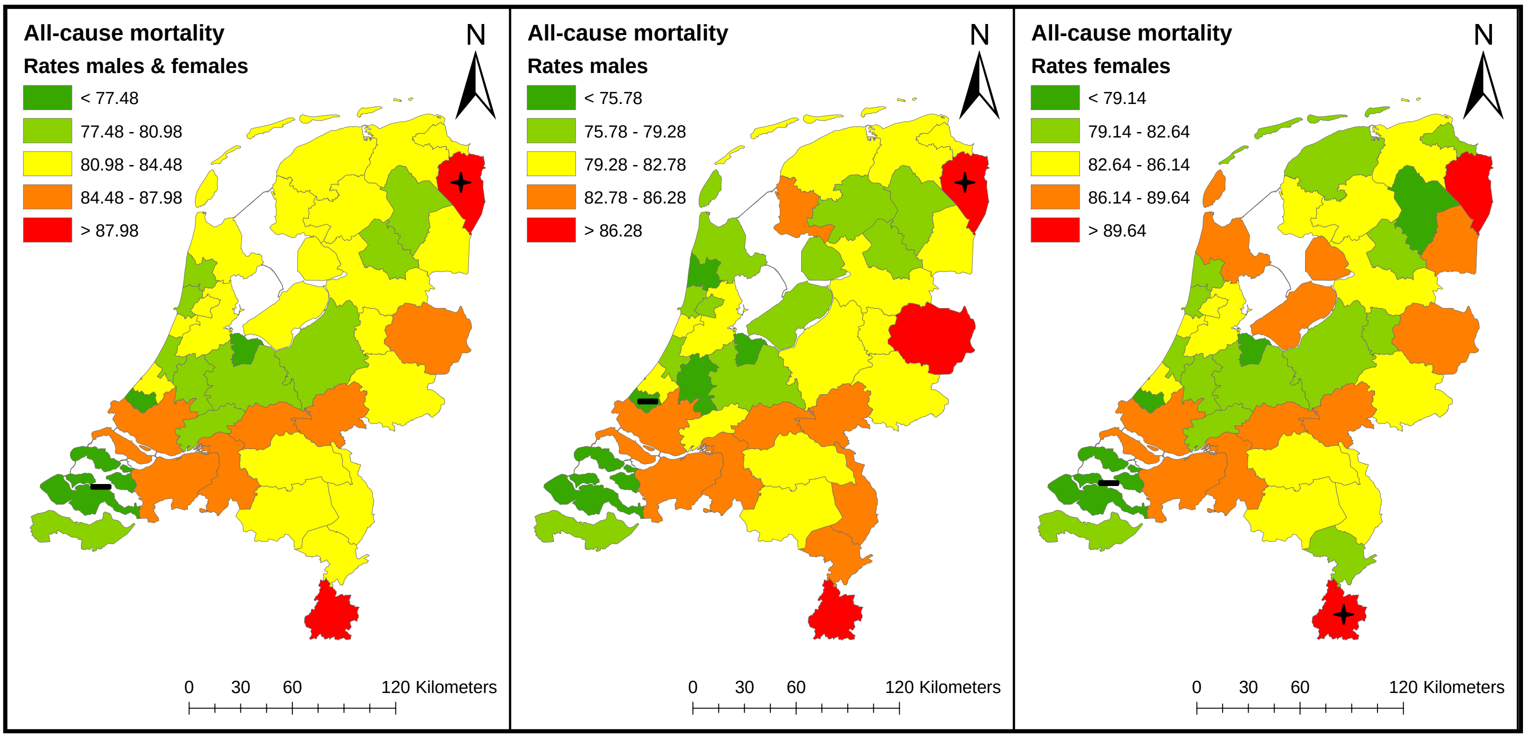
Figure 4.1 shows three separate maps that consist from the age-standardized all-cause mortality rates for males and females together and separately per 10,000 of the population.

The first map in figure 4.1 shows the age standardized all-cause mortality rates for males and females per 10,000 of the population. The average for the total of the Netherlands for males and females is 82.73. Most regions are close to the average of the Netherlands. The regions with the highest mortality rates are the regions of Oost-Groningen and Zuid-Limburg. As can be seen from figure 4.1, Oost-Groningen shows the highest rate, with a value of 89.63. What is remarkable is that the two regions with the highest mortality rates have neighbouring regions that are equal or lower than the average mortality rate of the Netherlands. The regions that have the lowest mortality rates are Het Gooi en Vechtstreek, Delft en Westland, and Overig Zeeland. The first two regions are in a part with more regions that have lower rates. As is shown in the first map of figure 4.1, Overig Zeeland has the lowest rate, that of 75.27.

The age standardized all-cause mortality rates for males per 10,000 of the population are depicted in the second map of figure 4.1. The average all-cause mortality rate of males for the total of the Netherlands is 81.03. The regions with the highest rates are Oost-Groningen, Twente, and Zuid-Limburg. Figure 4.1 shows that Oost-Groningen again has the highest rate of 89.56. All of the regions that have the highest mortality rates border to Germany, Zuid-Limburg also borders to Belgium. The regions with the lowest rates are Alkmaar en omgeving, Het Gooi en Vechtstreek, Oost-Zuid-Holland, Delft en Westland, and Overig Zeeland. These regions are located in the western part of the Netherlands. Figure 4.1 shows that the lowest mortality rate of 72.77 is located in the region of Delft en Westland.

In the last map of figure 4.1 the age standardized all-cause mortality rates for females per 10,000 of the population are shown. The average all-cause mortality rate of females for the total of the Netherlands is 84.39. The regions with the highest mortality rates are Oost-Groningen, and Zuid-Limburg. From figure 4.1 can be seen that Zuid-Limburg has the highest rate, with a value of 90.44. The regions with the lowest rates are Noord-Drenthe, Het Gooi en Vechtstreek, Delft en Westland, and Overig Zeeland. Overig Zeeland has the lowest rate, with a value of 75.58.

Figure 4.1 Age standardized all-cause mortality rates for males & females, males, and females per 10,000 inhabitants 2004-2008

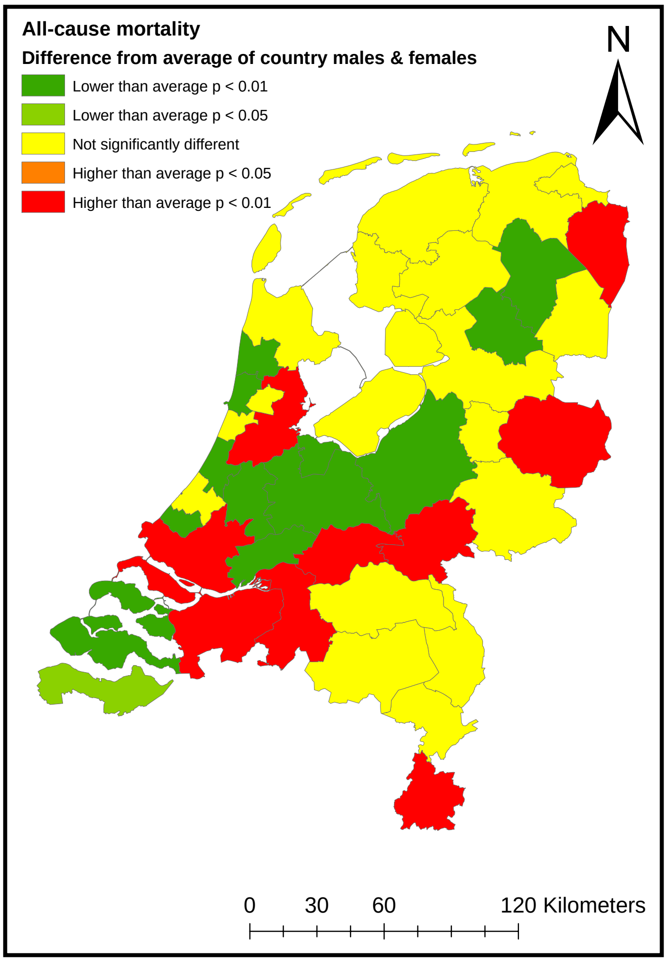


▬ Is region with lowest mortality rate per 10,000 inhabitants

✚ Is region with highest mortality rate per 10,000 inhabitants

Figure 4.2 shows significant difference of the all-cause mortality rates for regions when compared to the average of the Netherlands for males and females. Most regions are not significantly different from the average. The regions that are significantly different, are highly significant, they fall in the 99% confidence interval above or below the average of the Netherlands. Only Zeeuwsch-Vlaanderen falls in the 95% confidence interval below the average. The middle part of the Netherlands shows a cluster of regions that are lower than the average where $p < 0.01$: Veluwe, Het Gooi en Vechtstreek, Utrecht, Oost-Zuid-Holland, Agglomeratie Leiden en Bollenstreek, and Zuidoost-Zuid-Holland. Beneath this cluster, there are is a line of regions that have a higher significance than the average $p < 0.01$: Arnhem / Nijmegen, Zuidwest-Gelderland, Midden-Noord-Brabant, Noordoost-Noord-Brabant, and Groot-Rijnmond.

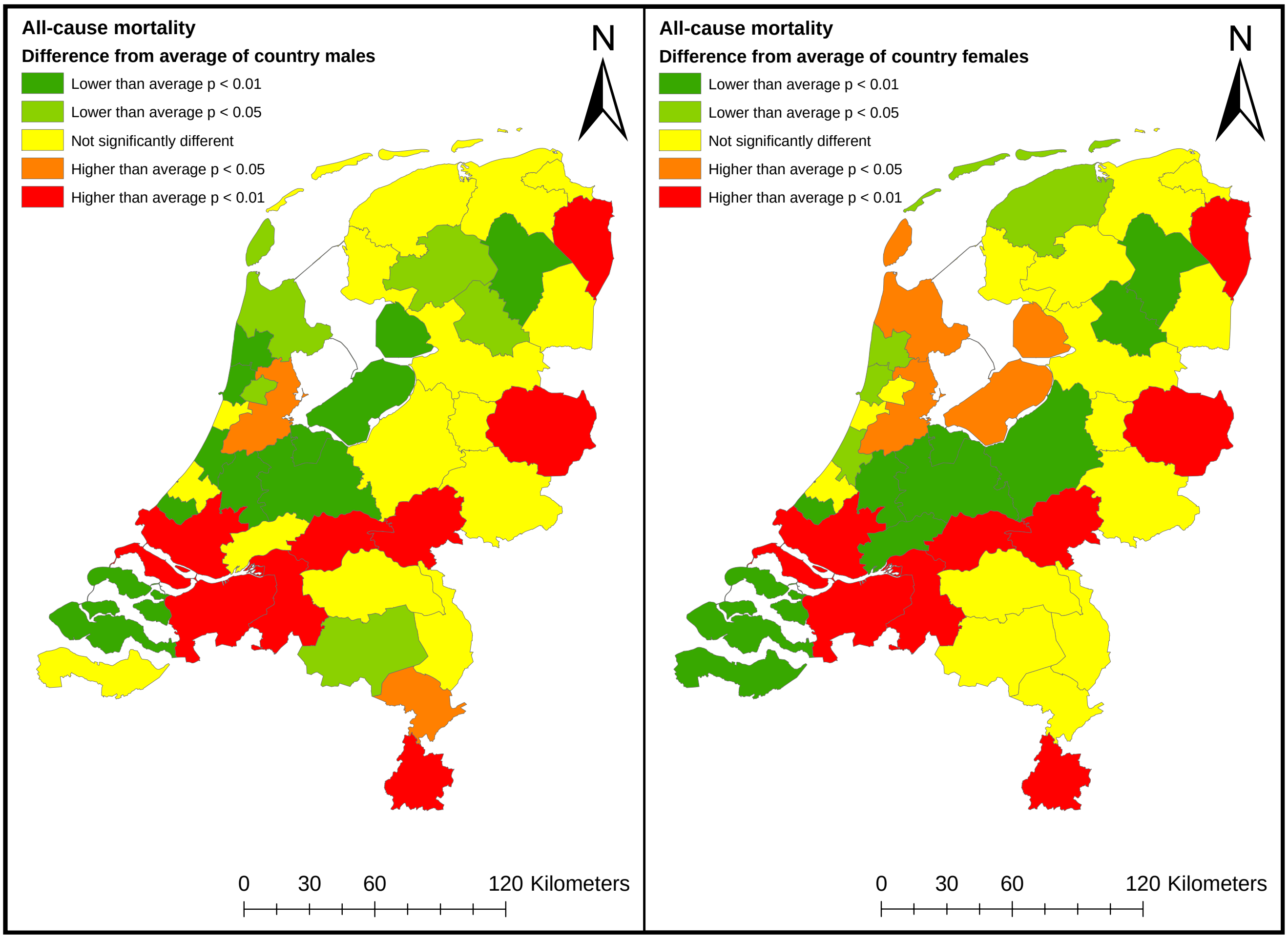
Figure 4.2 All-cause mortality significant differences from the average of the Netherlands for males and females 2004-2008



The first map of figure 4.3 shows how significantly different the all-cause mortality rates for regions are from the average of the Netherlands for males. Only fourteen regions are not significantly different from the average of the Netherlands. The mid-west part of the Netherlands shows a cluster of regions that are lower than the average where $p < 0.01$. This area is slightly smaller than the one in figure 4.2. Beneath this cluster, there are regions that are higher than the average $p < 0.01$. These regions are the same as mentioned with figure 4.2.

The second map in figure 4.3 shows the significant difference in all-cause mortality rates for regions when compared to the average of the Netherlands for females. The variations in this figure are almost the same as were discussed for males and females, and females. One difference is that the regions that lie in the cluster of the middle part of the Netherlands and that are lower than the average where $p < 0.01$ are now also bordered above with regions that have a significant higher mortality than the average, where $p < 0.05$. These are the regions of Flevoland and Groot-Amsterdam.

Figure 4.3 All-cause mortality significant differences from the average of the Netherlands for males and females separately 2004-2008



When looking at table 4.1, it can be seen that none of the Moran's Indexes for all-cause mortality are significantly different; this means that the mortality rates are randomly distributed across the different COROP-regions. The highest Moran's I value is that of the males: 0.143644. The all-cause mortality for males is not significantly different, but ArcGIS indicated here that: 'While somewhat clustered, the pattern may be due to random chance'.

Table 4.1 results of the global Moran's I of all-cause mortality for males and females, males, and females in the period of 2004-2008

Global Moran's I Summary			
All-cause mortality	**Males and females**	**Males**	**Females**
Expected Index	-0.025641	0.025641	-0.025641
Moran's Index	0.079895	0.143644	-0.063499
Z Score	0.863677	1.384788	-0.308527
p-value	0.387765	0.166117	0.757681

4.2 Top three smoking-related causes of death

In this paragraph, the top three smoking-related causes of death will be discussed; these are the following: COPD, IHD, and Lung cancer. This paragraph should give more understanding about smoking-related mortality and what it consists from.

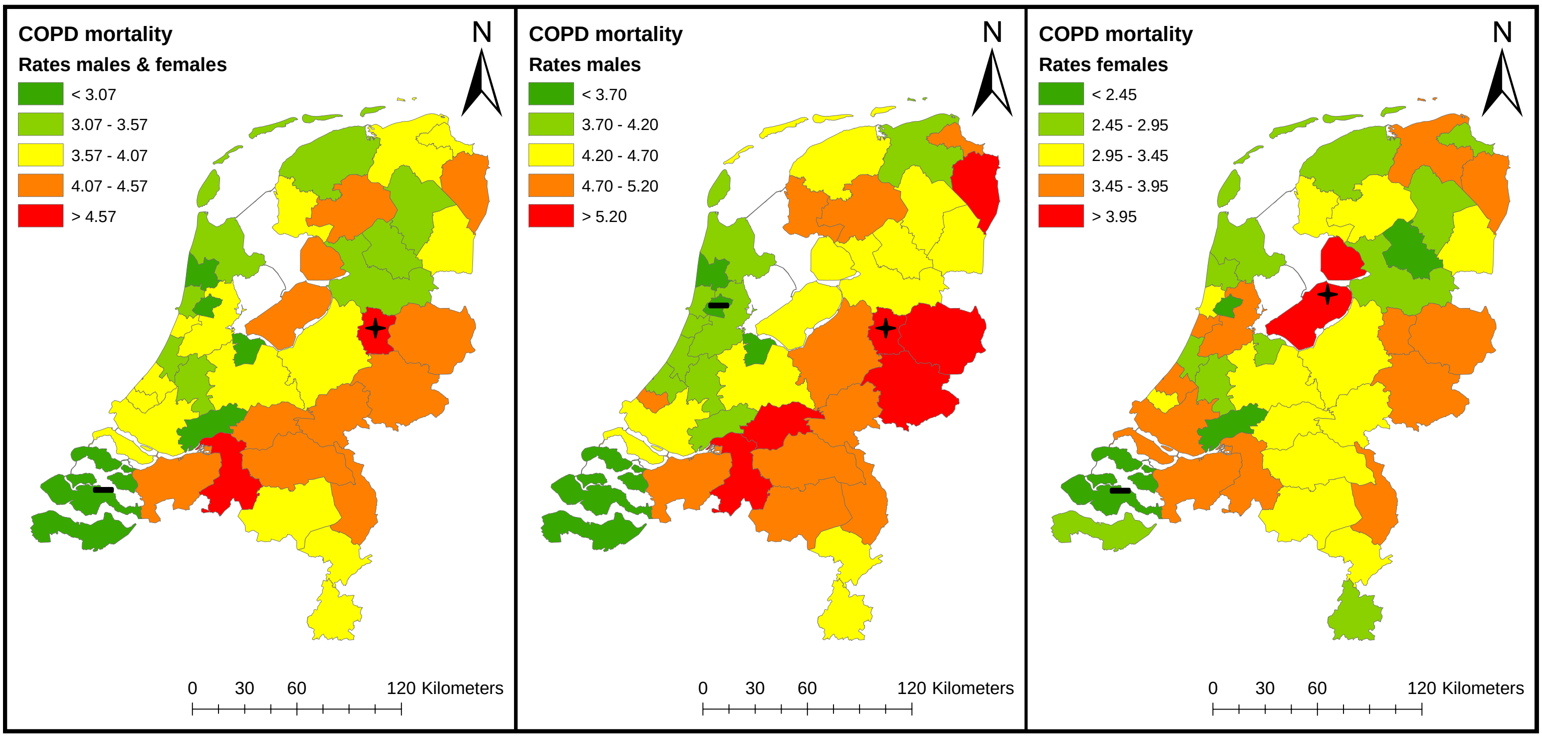
4.2.1 COPD mortality

Figure 4.4 shows the age standardized COPD mortality rates for males and females, males, and females per 10,000 of the population for the period of 2004-2008. The first map in figure 4.4 shows the variation in COPD mortality rates for males and females. The COPD mortality rate for the average of the Netherlands for males and females is 3.82. The regions of Zuidwest-Overijssel and Noordoost-Noord-Brabant have the highest mortality rates. Zuidwest-Overijssel has the highest mortality rate: 4.90. The regions of Alkmaar en omgeving, Zaanstreek, Het Gooi en Vechtstreek, Zuidoost-Zuid-Holland, Zeeuwsch-Vlaanderen, and Overig Zeeland have the lowest mortality rates. Overig Zeeland has the very lowest rate, with the value of 2.79.

The second map in figure 4.4 shows the COPD mortality rates for males. The average for the whole of the Netherlands is 4.45. This map has more regions that fall in the highest mortality rates category than in the first map of figure 4.4. For males, Oost-Groningen, Twente, Achterhoek, and Zuidwest-Gelderland are also included. The highest COPD mortality rate can be found in the region of zuidwest-Overijssel, with a value of 5.95. The only region that does not fall in the lowest mortality rate, when compared to the first map is the region of Zuidoost-Zuid-Holland. The region of the Zaanstreek has the lowest value of COPD mortality rate: 3.17. When looking at the map of the males, it seems there is a sort of cluster of higher mortality rates in the eastern and southern part of the Netherlands when compared to the average rate of the whole of the Netherlands. The western part of the Netherlands shows more a cluster of lower mortality rates.

The third map that shows the COPD mortality rates of the females has a very different pattern when looking at the highest mortality rates. The average of the Netherlands is 3.20. Only the region of Flevoland falls in the highest mortality rate group, with a value of 4.24. There are also fewer regions that fall in the lowest mortality rates range: Zuidwest-Drenthe, Zaanstreek, Zuidoost-Zuid-Holland, and overig Zeeland. The last region has the lowest COPD mortality rate of 2.12. The region of Zuid-Limburg has a lower mortality rate than the average of the Netherlands. This is different from the other causes of mortality, because Zuid-Limburg has higher mortality rates than average for the other mortality causes in this research.

Figure 4.4 Age standardized chronic obstructive pulmonary disease mortality rates for males & females, males, and females per 10,000 inhabitants 2004-2008



▬ Is region with lowest mortality rate per 10,000 inhabitants

✚ Is region with highest mortality rate per 10,000 inhabitants

Figure 4.5 shows how significantly different the COPD mortality rates for regions are from the average of the Netherlands for males and females. There is a very distinct clustering of regions that have higher a mortality rate that is significantly different from the Netherlands, $p < 0.01$. The regions are: Zuidwest-Overijssel, Twente, Achterhoek, Arnhem/Nijmegen, Noordoost-Noord-Brabant, Noord-Limburg, Midden-Noord-Brabant, and West-Noord-Brabant. A small cluster of regions that have a lower mortality rate that is significantly from the average of the Netherlands are: Het Gooi en Vechtstreek, Utrecht, Zuidoost-Zuid-Holland, Oost-Zuid-Holland, and Agglomeratie Leiden en Bollenstreek. Only Oost-Zuid Holland falls in the significance where $p < 0.05$, the other regions that where just mentioned fall in the significance where $p < 0.01$.

Figure 4.5 COPD mortality significant differences from the average of the Netherlands for males and females 2004-2008

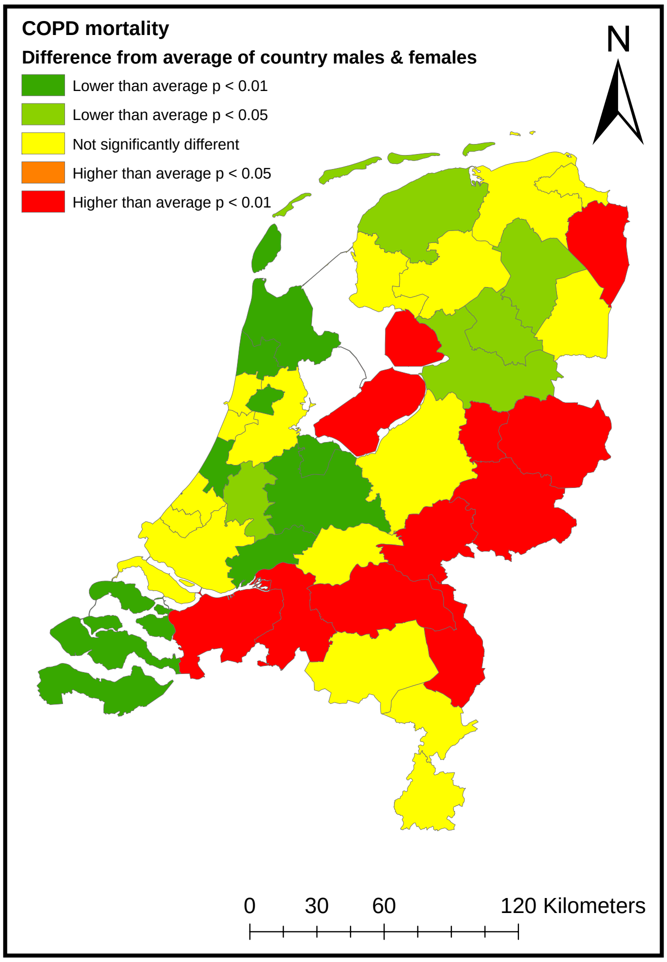


Figure 4.6 shows firstly what the significant difference the COPD mortality rates for regions are, when compared to the average of the Netherlands for males. Again, there is a very distinct clustering of regions that have higher a mortality rate that is significantly different from the Netherlands, $p < 0.01$. The difference in the cluster of figure 4.6 when looking at figure 4.5 is that West-Noord-Brabant is not included and Zuidwest-Gelderland is included. The regions of Zuidoost-Noord-Brabant and Noord-Limburg, which also belong to the cluster, are significantly different from the Netherlands with a p value < 0.05. This cluster can be found in the east-southern part of the Netherlands. A cluster of regions that have a lower mortality rate that is significantly different from the average of the Netherlands are the regions that are situated across the western (coastline) part of the Netherlands.

The second map in Figure 4.6 shows how significantly different the COPD mortality rates for regions are from the average of the Netherlands for females. In the south-eastern part of the Netherlands, the regions are not really significantly different from the Netherlands. In the rest of the country, there are significant differences that vary from higher mortality rates to lower mortality rates than the average of the Netherlands.

Figure 4.6 COPD mortality significant differences from the average of the Netherlands for males and females separately 2004-2008

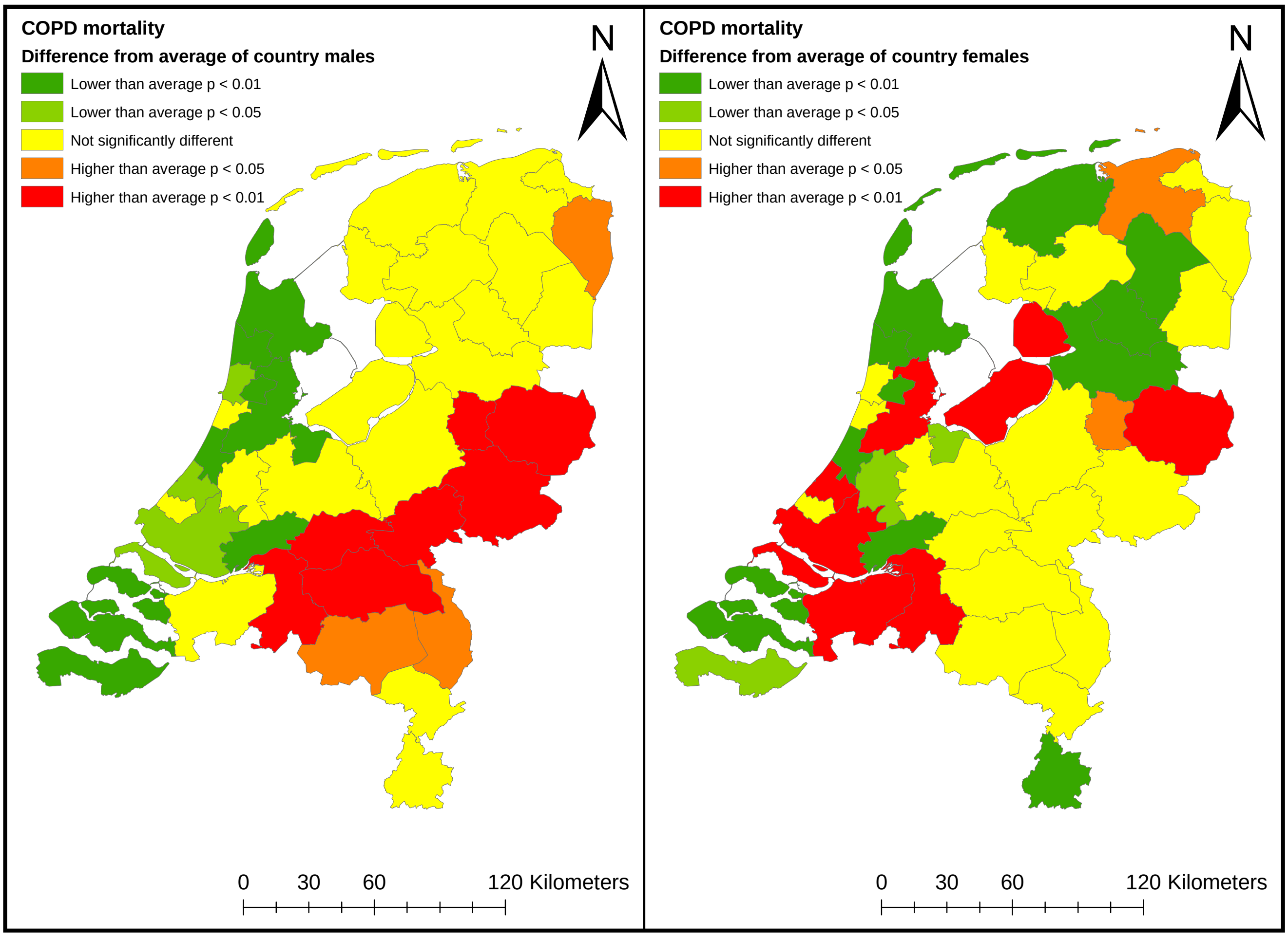


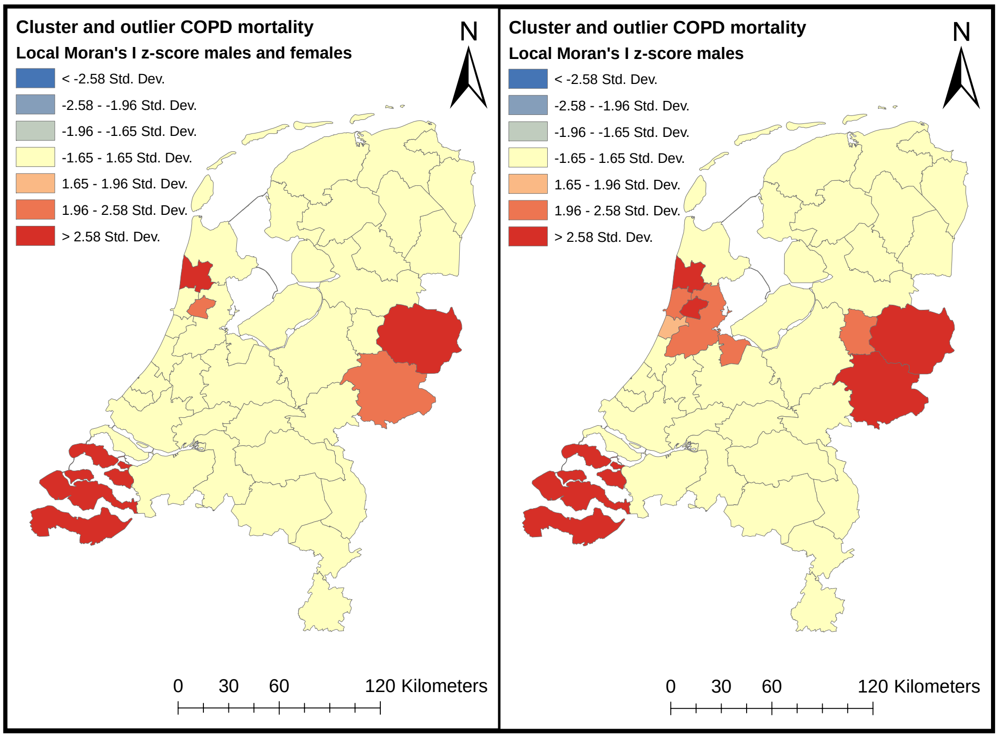
Table 4.2 shows the results of the global Moran's I for COPD mortality by sex. The COPD mortality rates for males and females, and males are significantly clustered and fall in the 99% confidence interval level. For males and females the Moran's I is 0.311125 and for males the Moran's I is higher, with a value of 0.588261. The COPD mortality rates for females are completely random distributed, as can be seen when comparing the expected index to the Moran's Index.

Table 4.2 results of the global Moran's I of COPD mortality for males and females, males, and females in the period of 2004-2008

Global Moran's I Summary			
COPD mortality	Males and females	Males	Females
Expected Index	-0.025641	0.025641	-0.025641
Moran's Index	0.311125	0.588361	-0.025753
Z Score	2.738458	4.999718	-0.000911
P-value	0.006173	0.000001	0.999273

Figure 4.7 shows were the clusters are located by using the local Moran's I that were discovered when the global Moran's I was calculated for males and females, and for males. There are no significant negative values to be found, and thus there are no regions surrounded by other regions with dissimilar values. However, there are regions that are surrounded by other regions that show significant similarity. Both for males and females, as for males, the clusters can be found in the same areas; in the mid-west, east-west, and in the south-west.

Figure 4.7 Clusters and outliers of the local Moran's I of COPD mortality for males and females, and males in 2004-2008



4.2.2 Ischaemic heart disease mortality

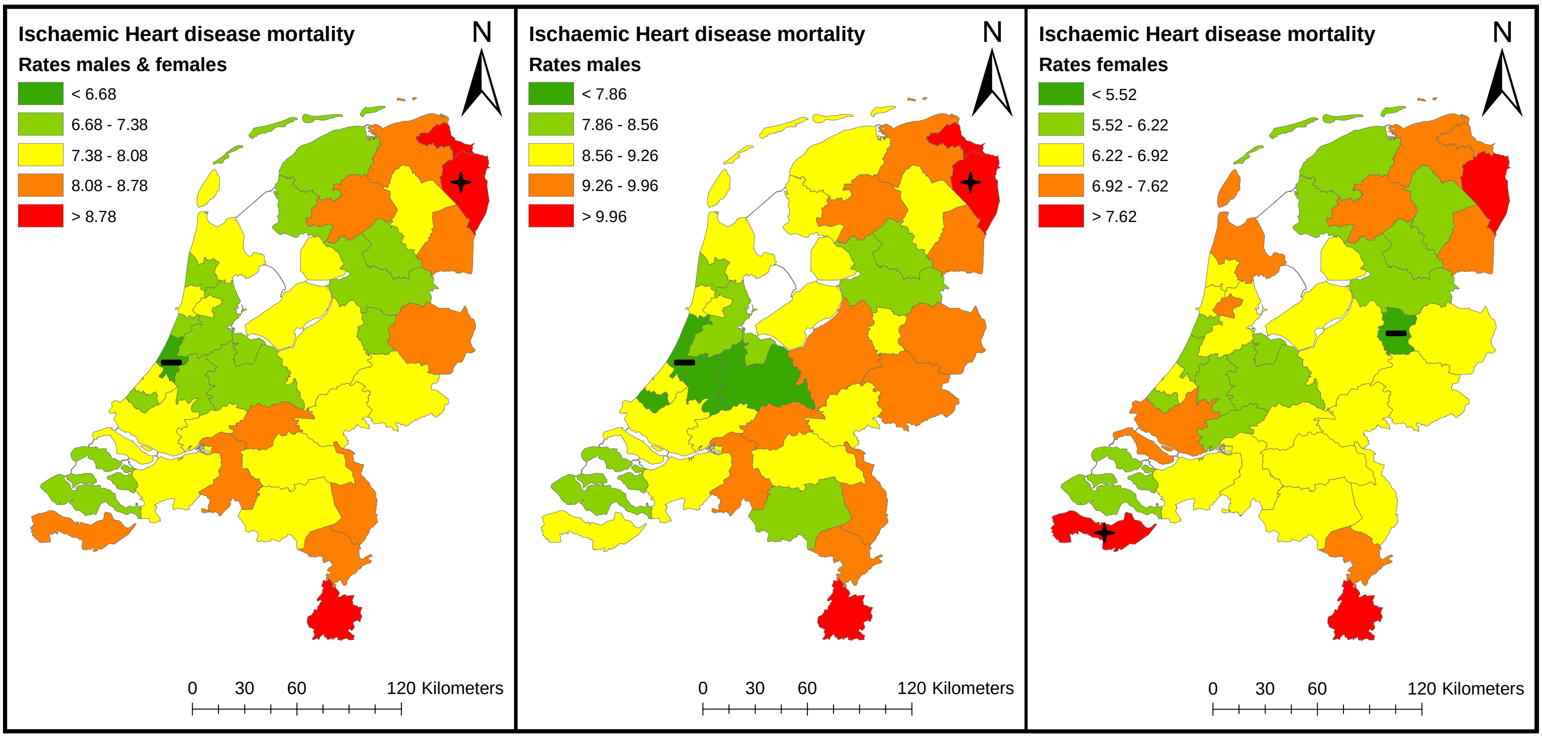
Figure 4.8 shows the age-standardized ischaemic heart disease mortality rates for males and females, males, and females per 10,000 of the population for the period of 2004-2008.

The first map in figure 4.8 shows the variation in IHD mortality rates for males and females. The average mortality rate for the total of the Netherlands is 7.73. The regions of Delfzijl en omgeving, Oost-Groningen, and Zuid-Limburg have the highest mortality rates. Oost-Groningen has the highest mortality rate of 9.77. Although there is a sort of concentration in the west of the Netherlands that have a lower IHD mortality rate than the average, the only region that falls in the lowest mortality rate category is Agglomeratie Leiden en Bollenstreek, with a mortality rate of 6.59.

It can be seen from the second map in figure 4.8 that the average IHD mortality rate for males is 8.91 for the total of the Netherlands. The map for males has the same regions that fall in the highest mortality rates category as for the males and females. The region with the highest mortality rate is Oost-Groningen, with a value of 11.57. The regions that fall in the lowest mortality rates category are more for males than in the first map for the males and females. Again, there is a sort of clustering in the west. Agglomeratie Leiden en Bollenstreek has the lowest mortality rate: 7.57.

The last map that shows the IHD mortality rates of the females. The average mortality rate for the total of the Netherlands for females is 6.57. The only change that can be seen in the map for females is that the region of Delfzijl en omgeving does not belong to the highest mortality rate category anymore; instead Zeeuwsch Vlaanderen now belongs to this group. The highest mortality rate belongs to the region of Zeeuwsch-Vlaanderen: 8.21. There is only one region that falls in the lowest rates range: Zuidwest-Overijssel, with a value of 5.36.

Figure 4.8 Age standardized ischaemic heart disease mortality rates for males & females, males, and females per 10,000 inhabitants 2004-2008



— Is region with lowest mortality rate per 10,000 inhabitants
+ Is region with highest mortality rate per 10,000 inhabitants

Figure 4.9 shows how significantly different the IHD mortality rates for regions are from the average of the Netherlands for males and females. Most regions are not significantly different from the average. There is a small cluster of regions in the mid-west of the Netherlands that have a significantly lower mortality rate where $p < 0.01$. In the north-east, a very small cluster can be found of regions that have a significantly higher mortality rate than the average of the Netherlands.

Figure 4.9 Ischaemic heart disease mortality significant differences from the average of the Netherlands for males and females 2004-2008

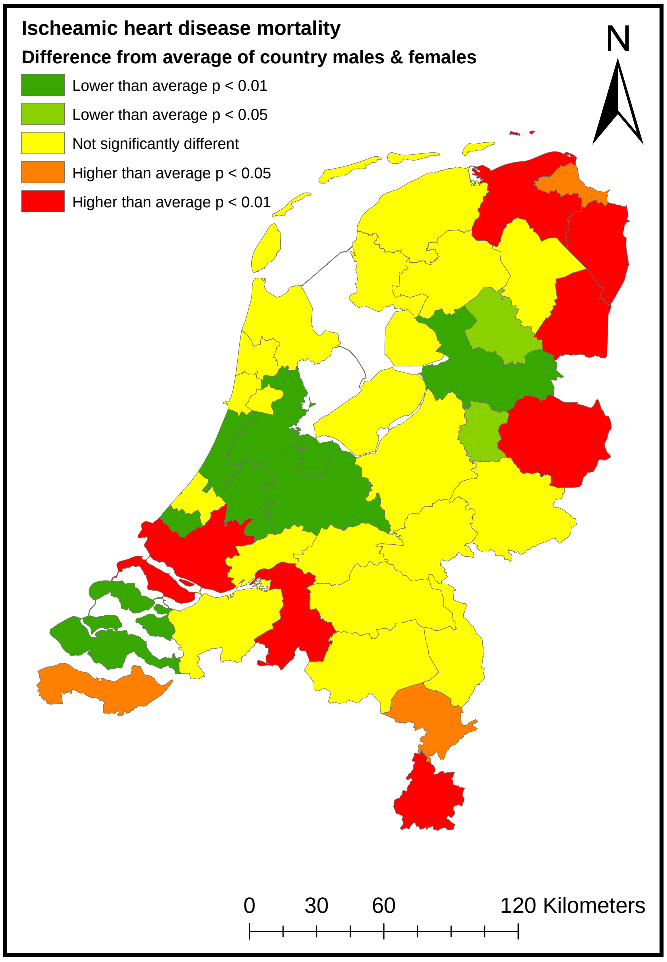


Figure 4.10 shows in the first map how significantly different the IHD mortality rates for regions are from the average of the Netherlands for males. Most regions are not significantly different from the average. Again, as was the case for the males and females, a small cluster of regions in the mid-west of the Netherlands can be found that have a significantly lower mortality rate where p < 0.01. The rest of the regions that differ significantly from the average of the Netherlands is scattered across the map.

In the second map, figure 4.10 shows the significant difference in IHD mortality rates for regions from the average of the Netherlands for females. Most regions are not significantly different from the average. The cluster that could be found for males and females, and males in the mid-west of the country is now almost gone. The rest of the regions that differ significantly from the average of the Netherlands are scattered across the map, there is not a certain pattern that can be found.

Figure 4.10 Ischaemic heart disease mortality significant differences from the average of the Netherlands for males and females separately 2004-2008

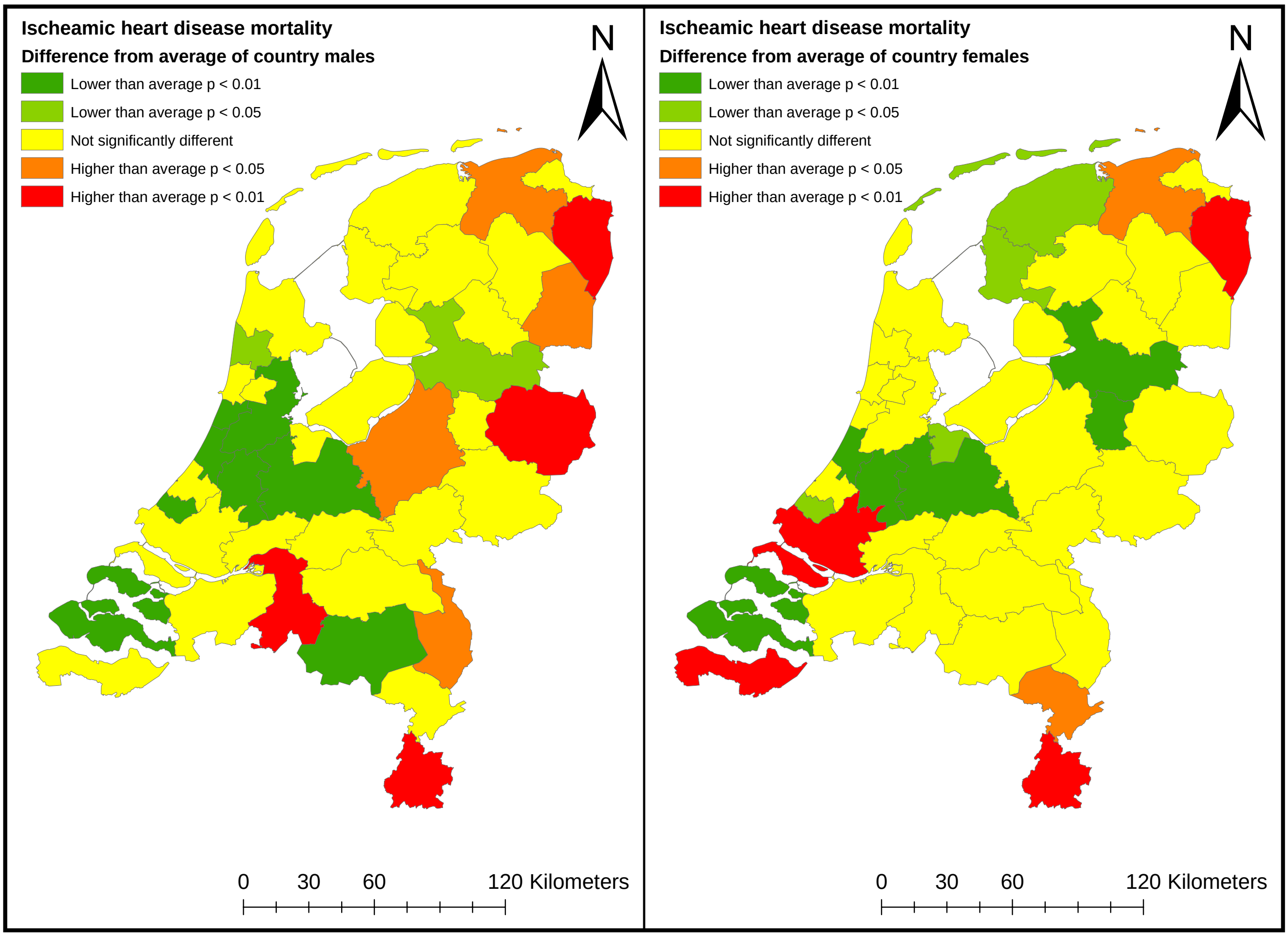


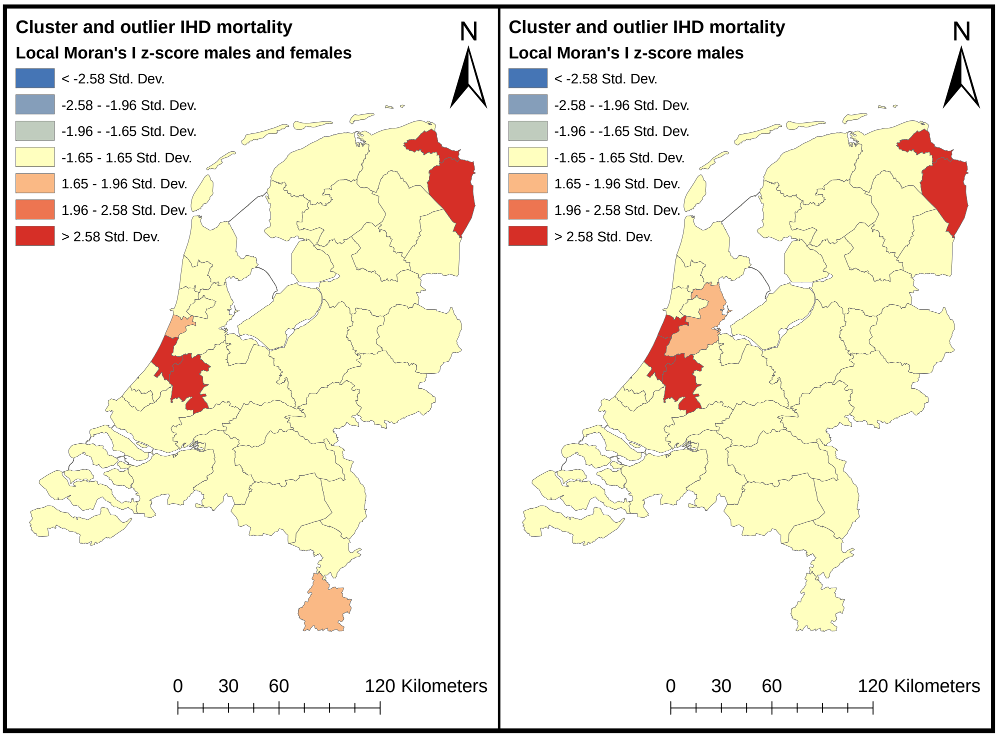
Table 4.3 shows the results of the global Moran's I of ischaemic heart disease by sex. For males and females, and males the IHD mortality rates are significantly clustered. They both fall in the 99% confidence interval level. For males and females the Moran's I is 0.319336, for males the Moran's I is 0.462547, and thus more clustered. The Moran's I of IHD mortality for females is to close to the expected value and randomly distributed.

Table 4.3 results of the global Moran's I of IHD mortality for males and females, males, and females in the period of 2004-2008

Global Moran's I Summary			
IHD mortality	**Males and females**	**Males**	**Females**
Expected Index	-0.025641	0.025641	-0.025641
Moran's Index	0.319336	0.462547	0.108396
Z Score	2.835624	4.040946	1.094361
p-value	0.004574	0.000053	0.273797

The clusters to be found for ischaemic heart disease for males and females, and males are depicted in figure 4.11. Both for males and females, as for males the clusters are located in somewhat the same areas. The first cluster can be found in the mid-west, and the second cluster can be found in the north-east of the Netherlands. For males and females the region of Zuid-Limburg indicates that it is surrounded by a similar value, only this is weakly significant, it falls in the 90% confidence interval level.

Figure 4.11 Clusters and outliers of the local Moran's I of IHD mortality for males and females, and males in 2004-2008



4.2.3 Lung cancer mortality

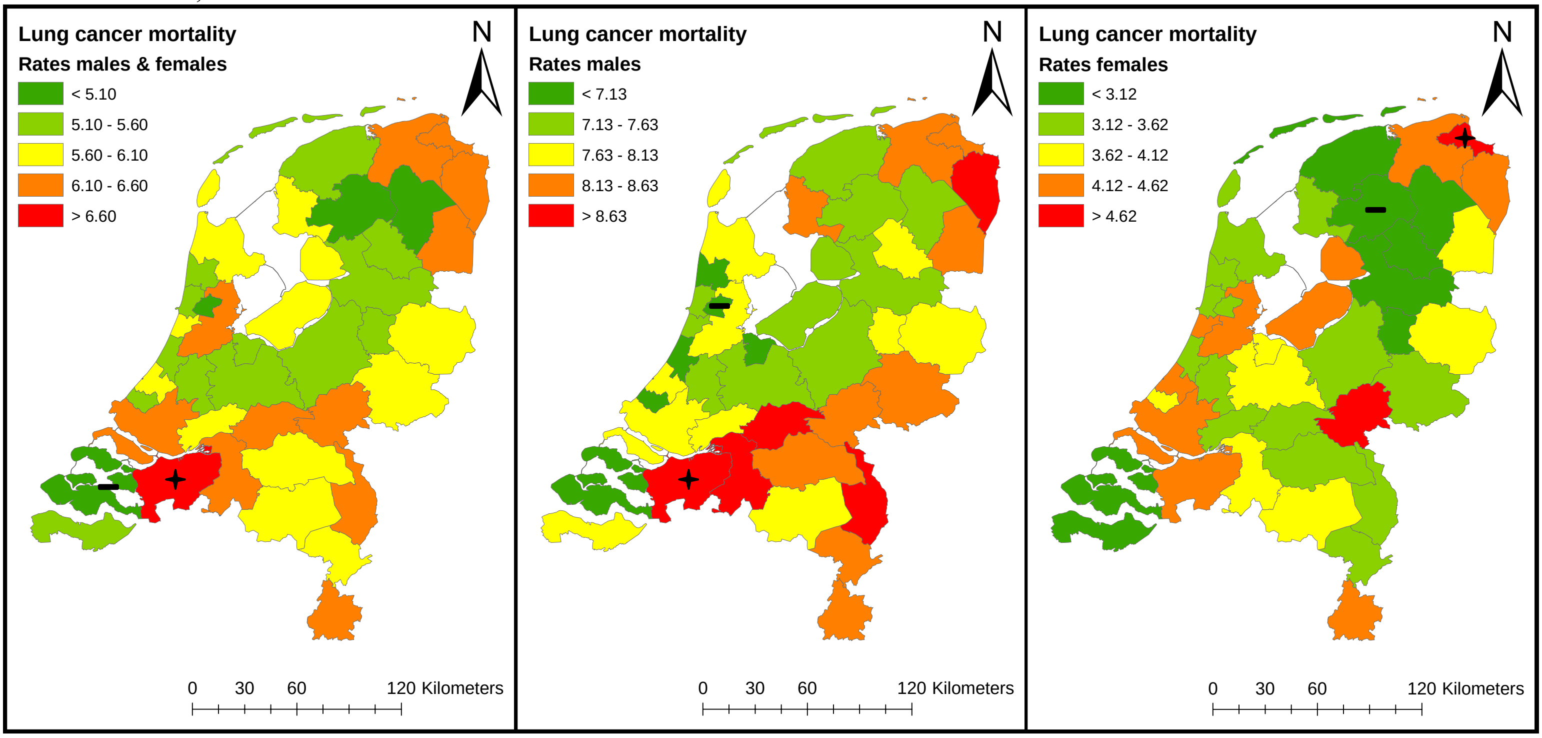
Figure 4.12 shows the age-standardized lung cancer mortality rates for males and females, males, and females per 10,000 of the population for the period of 2004-2008.

The first map in figure 4.12 shows the variation in lung cancer mortality rates for males and females. The average mortality rate for the total of the Netherlands is 5.85. The only region that falls in the highest mortality rates category is West-Noord-Brabant, with a value of 6.87. The regions that are in the lowest mortality rate category are: Zuidoost-Friesland, Noord-Drenthe, and Overig Zeeland. The last region is the lowest, with a mortality rate of 4.91.

The average mortality rate of the Netherlands when looking at the second map for males is 7.88. The regions that fall in the highest mortality rate category are: Oost-Groningen, Zuidwest-Gelderland, Midden-Noord-Brabant, West-Noord-Brabant, and Noord-Limburg. West-Noord-Brabant has the highest mortality rate, that of 9.45. The regions that fall in the lowest mortality rate category are: Alkmaar en omgeving, Zaanstreek, Agglomeratie Leiden en Bollenstreek, het Gooi en Vechtstreek, Delft en Westland, and overig Zeeland. Of all the regions that where just mentioned, the Zaanstreek has the lowest mortality rate of 6.52.

The third map that shows the lung cancer mortality rates of the females. The average mortality rate for the total of the Netherlands for females is 3.87. Only the region of Arnhem/Nijmegen falls in the highest mortality rate category, with a value of 4.65. The regions that fall in the lowest category of mortality rates are: Noord-Friesland, Zuidoost-Friesland, Noord-Drenthe, Zuidwest-Drenthe, Noord-Overijssel, Overig Zeeland, and Zeeuwsch Vlaanderen. The first six regions that were just mentioned are located in a very distinct cluster in the north of the Netherlands. Zuidoost-Friesland has the lowest mortality rate, with a value of 2.60.

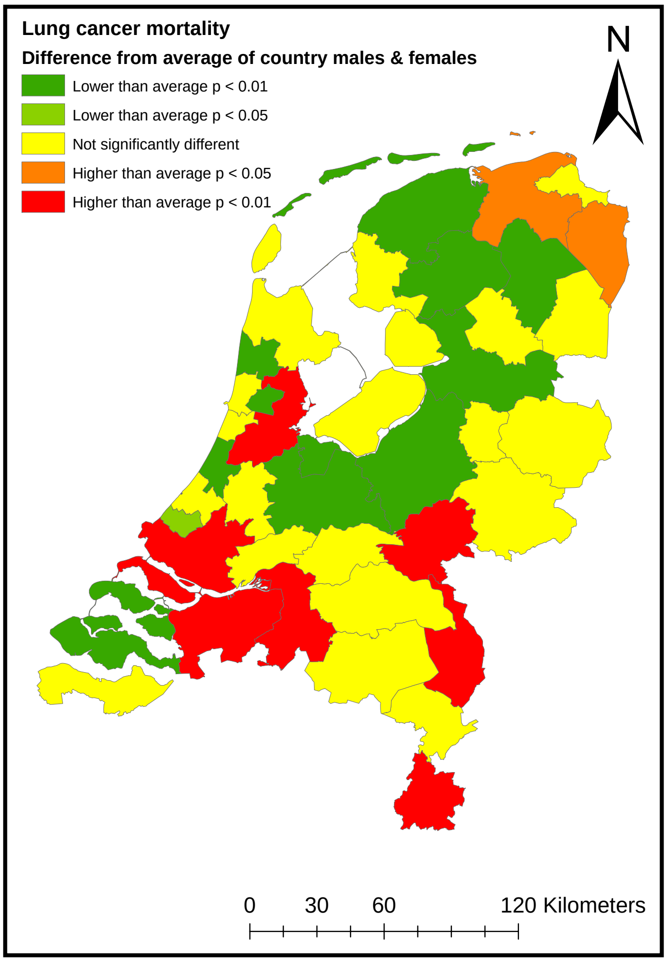
Figure 4.12 Age standardized lung cancer mortality rates for males & females, males, and females per 10,000 inhabitants 2004-2008



— Is region with lowest mortality rate per 10,000 inhabitants
✦ Is region with highest mortality rate per 10,000 inhabitants

Figure 4.13 shows how significantly different the lung cancer mortality rates for regions are from the average of the Netherlands for males and females. About half of the regions are not significantly different from the average. There is a sort of cluster of regions in that runs from the north of the Netherlands to the middle of the Netherlands that have a significantly lower mortality rate where $p < 0.01$. The mortality rates significantly higher than the average of the Netherlands where $p < 0.01$ can be found spread around in the lower half of the Netherlands.

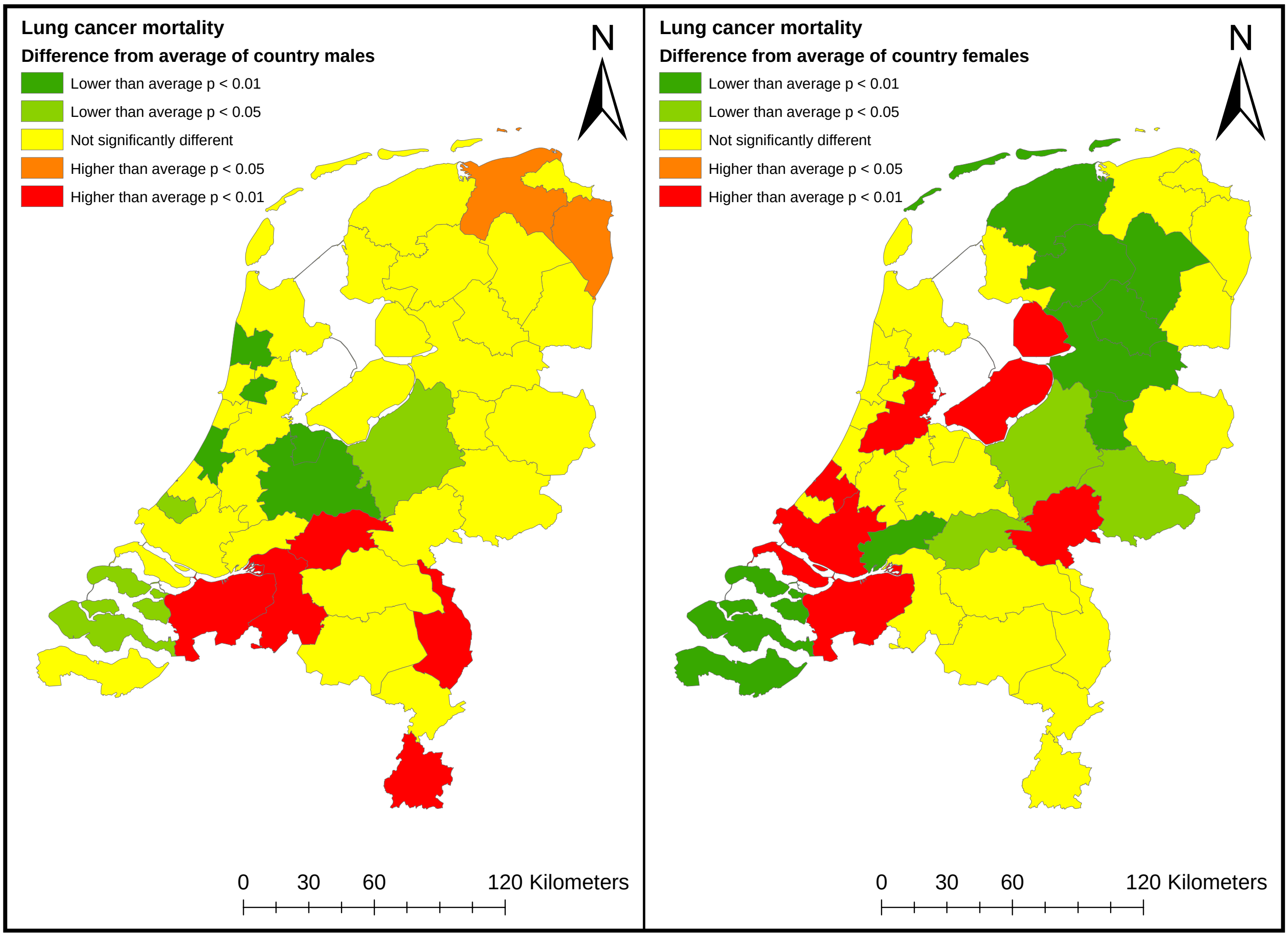
Figure 4.13 Lung cancer mortality significant differences from the average of the Netherlands for males and females 2004-2008



The first map of figure 4.14 shows how significantly different the lung cancer mortality rates for regions are from the average of the Netherlands for males. Most of the regions are not significantly different from the average of the Netherlands. The regions that have a significantly lower mortality rate where $p < 0.01$ or 0.05 are situated more in the middle part of the Netherlands. The mortality rates that are significantly higher than the average of the Netherlands where $p < 0.01$ or 0.05 can be found spread around the south and the northeast part of the Netherlands.

Figure 4.14 shows the significantly different the lung cancer mortality rates for regions when compared to the average of the Netherlands for females in the second map. About half of the regions are not significantly different from the average. There is a very distinct cluster of regions in situated in the north of the Netherlands that have a significantly lower mortality rate where $p < 0.01$. The mortality rates significantly higher than the average of the Netherlands where $p < 0.01$ can be found spread around in the middle part of the Netherlands.

Figure 4.14 Lung cancer mortality significant differences from the average of the Netherlands for males and females separately 2004-2008



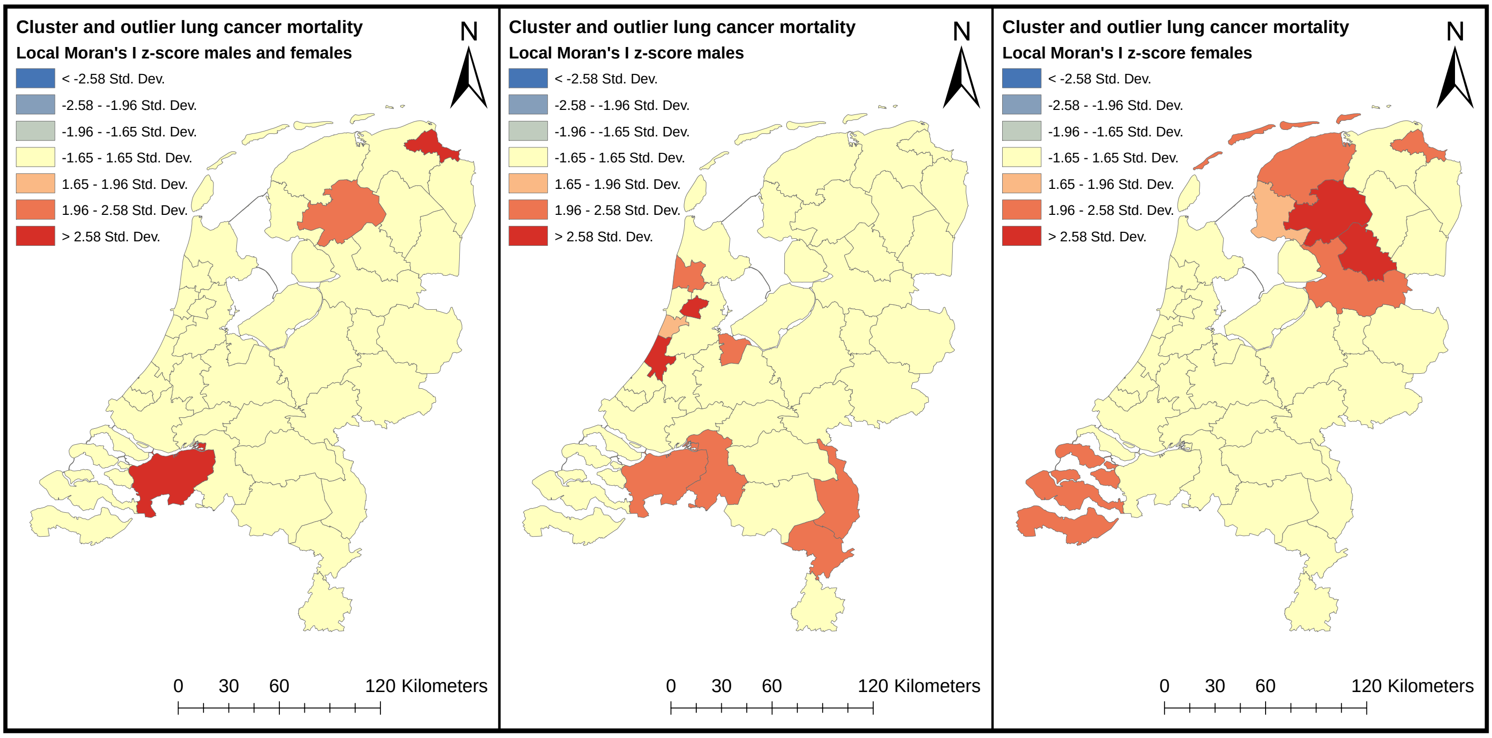
The results of the Global Moran's I for lung cancer mortality by sex are shown in table 4.4. The lung cancer mortality rates for males and females, males, and females are all significantly clustered. The clustering of the males falls in the 99% confidence interval level with a p-value of 0.000102, females in the 95% confidence interval level, and males and females fall in the 90 % confidence interval level.

Table 4.4 results of the global Moran's I summary lung cancer mortality for males and females, males, and females in the period of 2004-2008

Global Moran's I Summary			
Lung cancer mortality	**Males and females**	**Males**	**Females**
	-0.025641	0.025641	-0.025641
Moran's Index	0.184275	0.451103	0.241321
Z Score	1.702228	3.884624	2.167297
p-value	0.088713	0.000102	0.030212

Figure 4.15 shows where the clusters of lung cancer mortality by sex are located. The first map, for males and females, shows that the regions of West-Noord-Brabant in the south of the Netherlands and Delfzijl en omgeving in the north are significantly surrounded by regions with similar values with a confidence interval of 99%. The region of Zuid-West-Friesland falls in the 95% confidence interval. For males, the clustering is mostly located in the mid-west and south of the Netherlands. For females, this is very different, In the north of the Netherlands a very distinct cluster can be found.

Figure 4.15 Clusters and outliers of the local Moran's I of lung cancer mortality for males and females, males, and females in 2004-2008



4.3 Smoking-related mortality

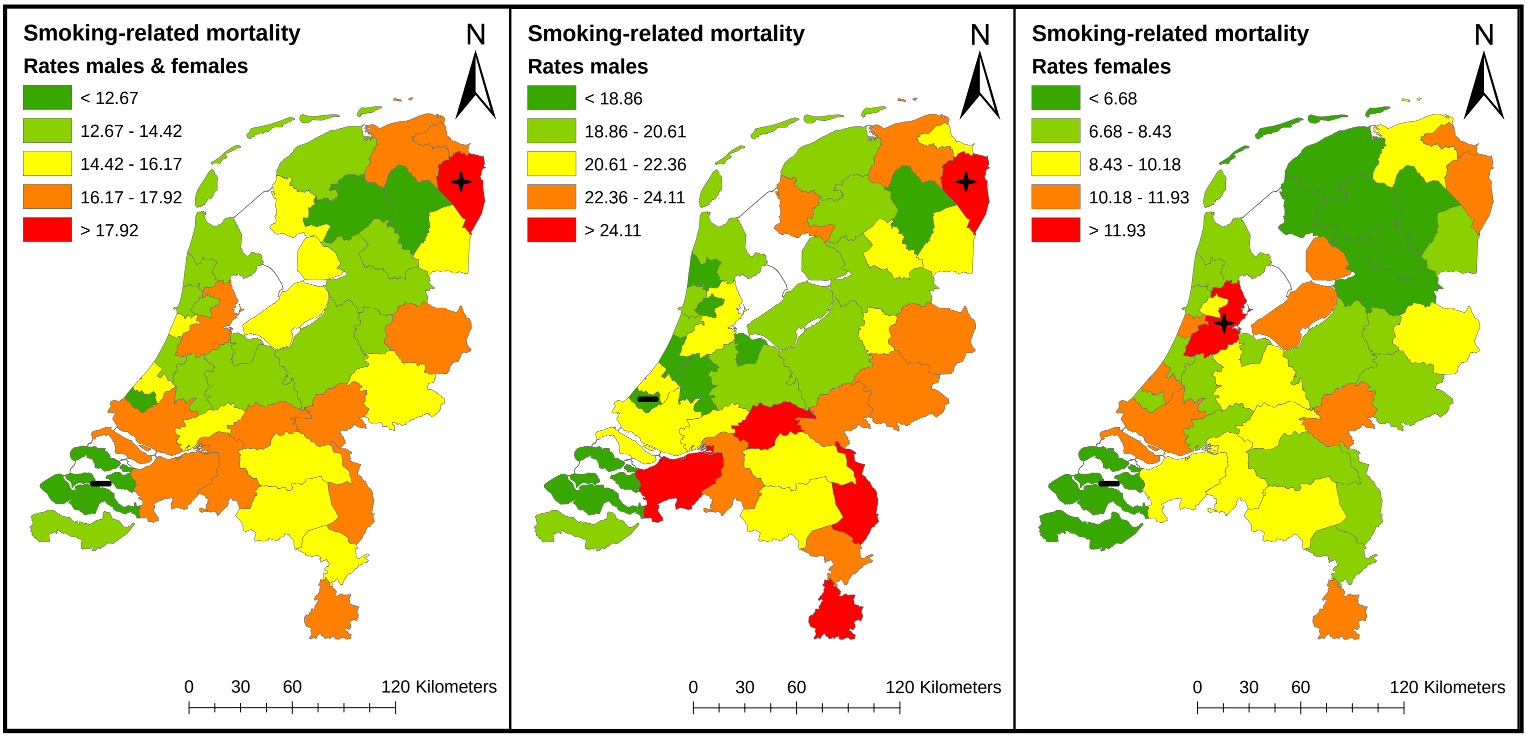
Figure 4.16 shows the age-standardized smoking-related mortality rates for males and females, males, and females per 10,000 of the population for the period of 2004-2008.

As can be seen in the first map of figure 4.16, the highest rate is located in the region of Oost-Groningen, with a value of 18.35. The average mortality rate for the Netherlands is 15.29. The lowest rates are in the regions of Zuidoost-Friesland, Noord-Drenthe, Delft en Westland, and Overig Zeeland. The lowest mortality rate can be found in the region of overig Zeeland, with a value of 11.59.

The second map of figure 4.16 shows the smoking-related mortality for males in the period of 2004-2008. The average mortality rate for the total of the Netherlands is 21.48. The high rates are in the regions of Oost-Groningen, Zuidwest-Gelderland, West-Noord-Brabant, Noord-Limburg, and Zuid-Limburg. Oost-Groningen has the highest rate of 26.19. The regions that fall in the lowest mortality rate category are Noord-Drenthe, Alkmaar en omgeving, Zaanstreek, Het Gooi en Vechtstreek, Agglomeratie Leiden en Bollenstreek, Oost-Zuid-Holland, Delft en Westland, and Overig Zeeland. Delft en Westland has the lowest mortality rate of 17.08.

In figure 4.16 the third map shows the highest smoking-related mortality rate in the region of Groot-Amsterdam, with a value of 12.29. The average lung cancer mortality rate for the total of the Netherlands is 9.30. The regions that are in the lowest category of the smoking-related mortality rates are: Noord-Friesland, Zuidwest-Friesland, Zuidoost-Friesland, Noord-Drenthe, Zuidwest-Drenthe, Noord-Overijssel, Zuidwest-Overijssel, Overig Zeeland, and Zeeuwsch-Vlaanderen. Overig Zeeland has the Lowest smoking-related mortality rate, with a value of 5.15. In the northern part of the Netherlands there is a very distinct cluster that can be found for low rates in smoking-related mortality.

Figure 4.16 Age standardized smoking-related mortality rates for males and females, males, and females per 10,000 inhabitants 2004-2008

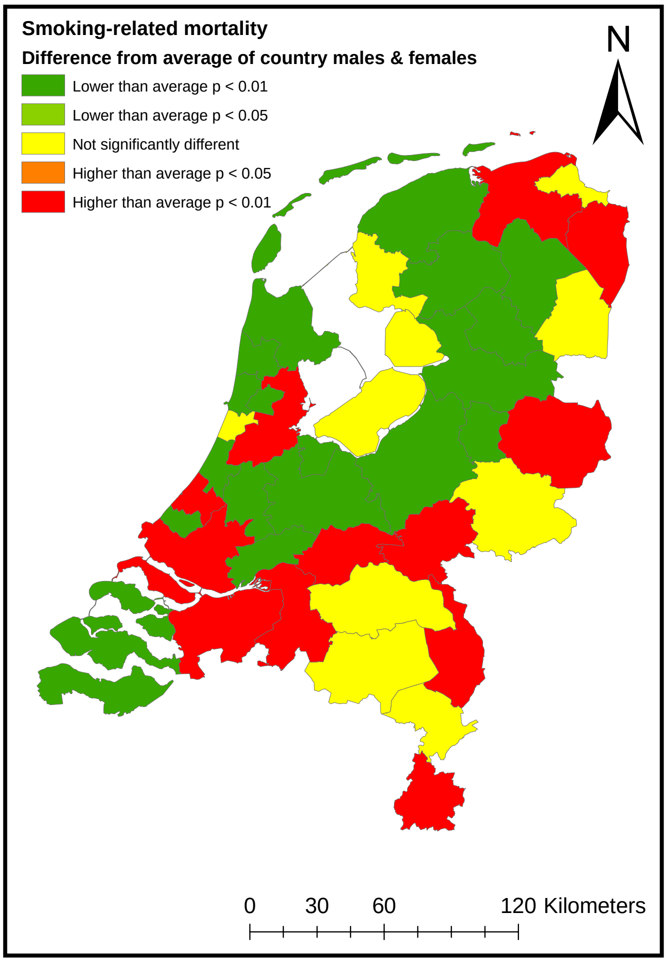


▬ Is region with lowest mortality rate per 10,000 inhabitants

✚ Is region with highest mortality rate per 10,000 inhabitants

Figure 4.17 shows how significantly different the smoking-related mortality rates for regions are from the average of the Netherlands for males and females. There are very little regions that are not significantly different from the average. There is a sort of cluster of regions in that runs from the north of the Netherlands to the middle of the Netherlands that have a significantly lower mortality rate where $p < 0.01$. The mortality rates significantly higher than the average of the Netherlands where $p < 0.01$ can be found as a strip on the lower side of the cluster with the low mortality rates.

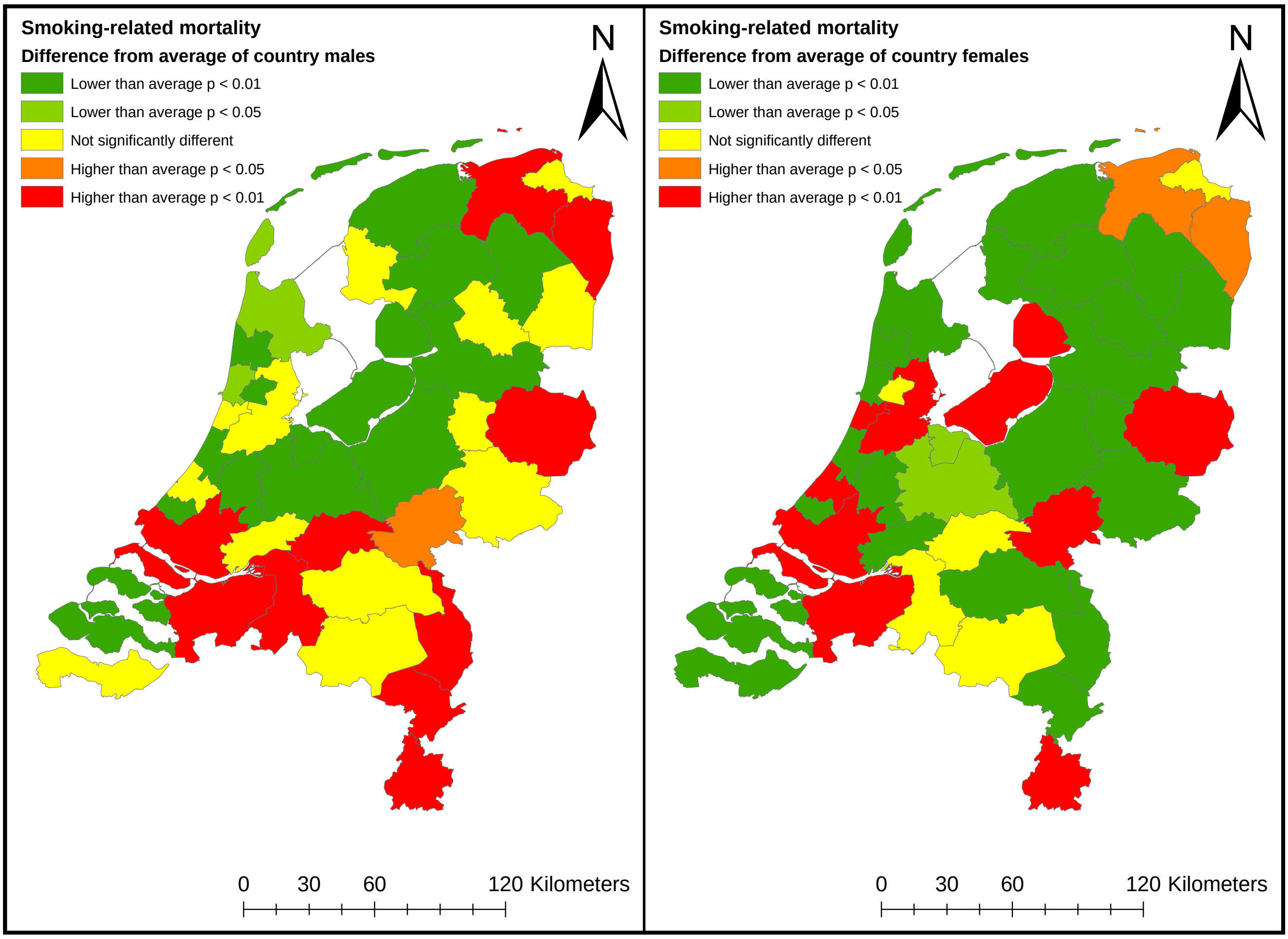
Figure 4.17 Smoking-related mortality significant differences from the average of the Netherlands for males and females 2004-2008



The first map of figure 4.18 shows the significant difference of the smoking-related mortality rates when regions are compared to the average of the Netherlands for males. There are very little regions that are not significantly different from the average of the Netherlands. The regions that have a significantly lower mortality rate where $p < 0.01$ or 0.05 are clustered in the middle and northern part of the Netherlands. The mortality rates that are significantly higher than the average of the Netherlands where $p < 0.01$ or 0.05, can be found spread around the south, east, and the northeast part of the Netherlands.

The second map of figure 4.18 shows how significantly different the smoking-related mortality rates for regions are from the average of the Netherlands for females. Only five regions are not significantly different from the average of the Netherlands. Most regions are significantly lower than the average of the Netherlands where $p > 0.01$. There is a very distinct cluster of regions in situated in the north of the Netherlands that runs to the middle of the Netherlands that have a significantly lower mortality rate where $p < 0.01$. The mortality rates significantly higher than the average of the Netherlands where $p < 0.01$ can be found spread around in the Netherlands.

Figure 4.18 Smoking-related mortality significant differences from the average of the Netherlands for males and females separately 2004-2008



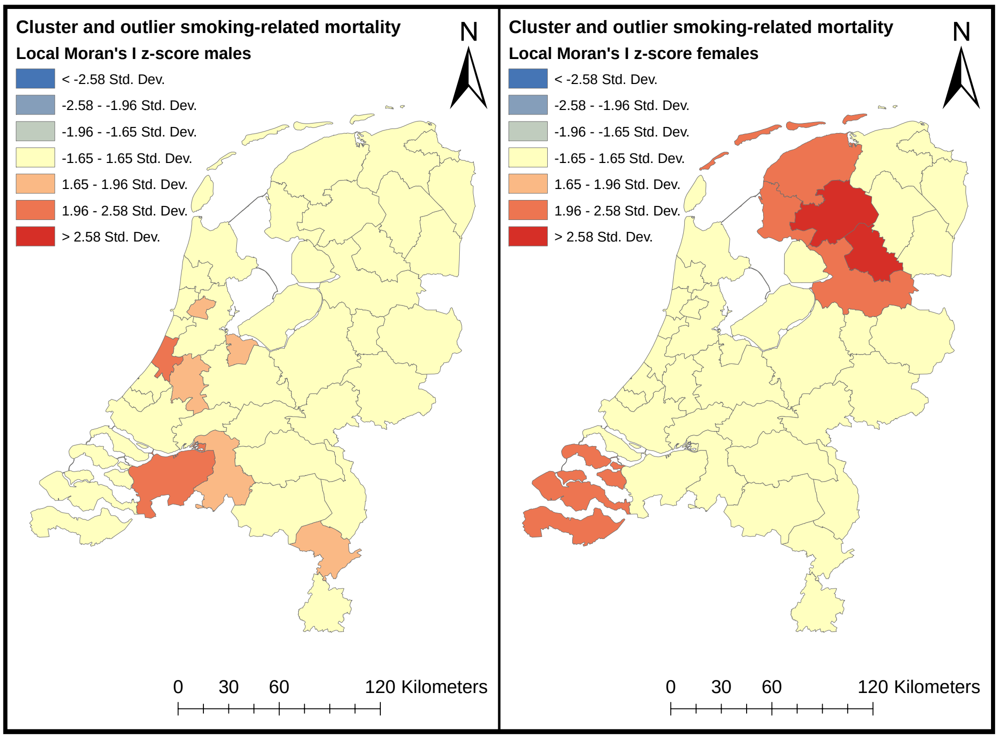
The mortality rates of smoking-related mortality for males and females separately are significantly clustered, as can be seen in table 4.5. The smoking-related mortality for males has a confidence interval level of 95%, although it almost falls in the 99% confidence interval level. The smoking-related mortality for females, on the other hand, falls in the 90% confidence interval and is not very significantly clustered.

Table 4.5 results of the global Moran's I summary smoking-related mortality for males and females, males, and females in the period of 2004-2008

Global Moran's I Summary			
Smoking-related mortality	**Males and females**	**Males**	**Females**
Expected Index	-0.025641	0.025641	-0.025641
Moran's Index	0.053480	0.290664	0.184691
Z Score	0.640226	2.569589	1.706814
p-value	0.522026	0.010182	0.087857

When looking at figure 4.19, it can bee seen that the specific clustering of the individual regions for males is not very significantly clustered. This is not very expected, because the global Moran's I indicated a clustering that almost fell in the 99% confidence interval level. For females there is a very distinct cluster to be found in the northern regions of the Netherlands, this is also not expected when looking at the p-value of the females in table 4.5, but when comparing it to third map in figure 4.16 there are similarities to be found.

Figure 4.19 Clusters and outliers of the local Moran's I of smoking-related mortality for males and females separately in 2004-2008



4.5 Regional differences in all-cause mortality attributed by smoking-related mortality

4.5.1 A comparison of all-cause mortality rates to smoking-related mortality rates

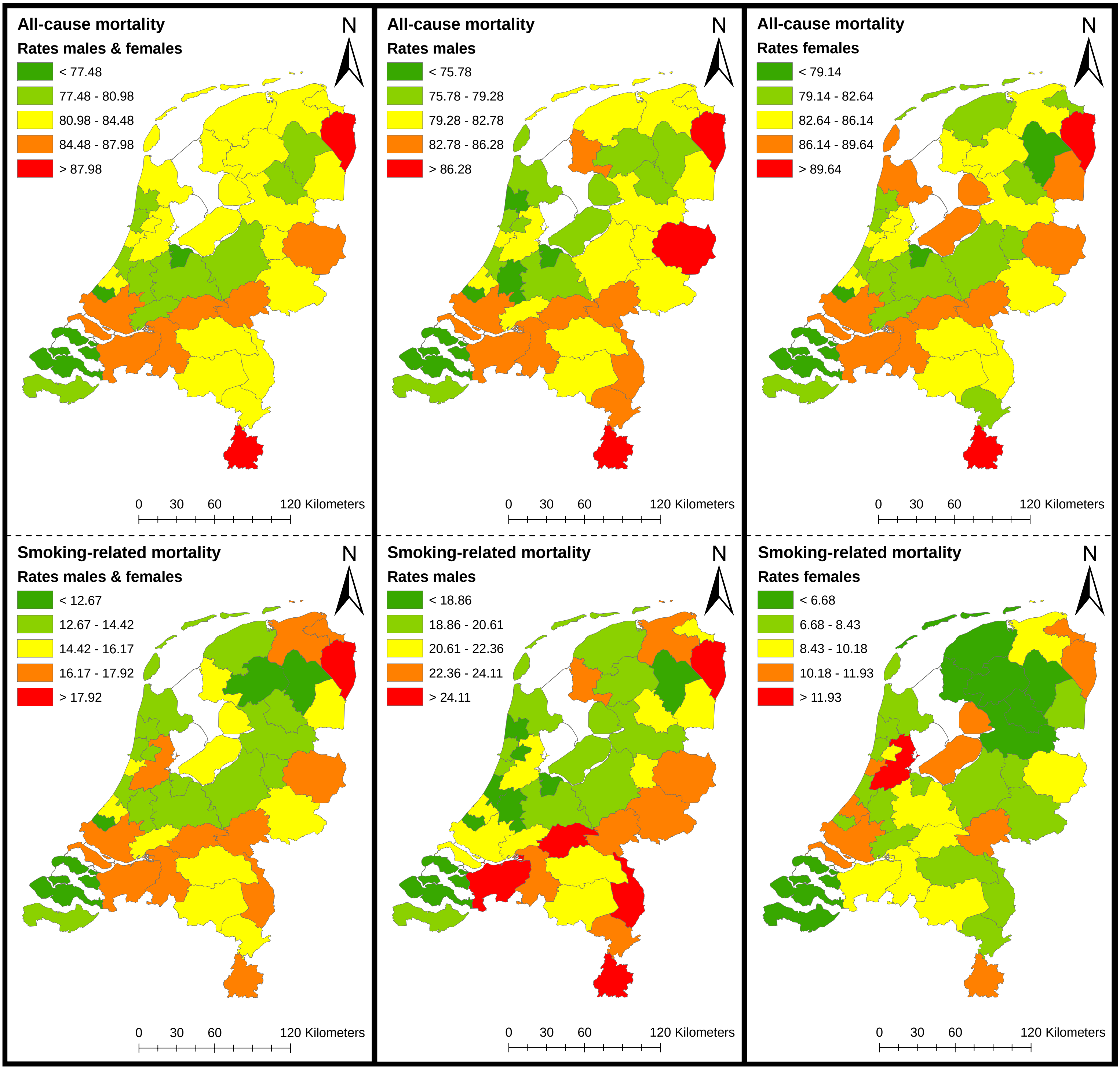
Figure 5.21 shows the standardized all-cause mortality rates and smoking-related mortality rates for males & females, males, and females per 10,000 inhabitants 2004-2008.

When comparing the mortality rates of all-cause mortality and smoking-related mortality for the males and females, it can be seen that there is some resemblance. For both causes of mortality, Oost-Groningen has the highest mortality rate. This is also the case for the lowest mortality rate, where Overig Zeeland is the lowest. The variations in all-cause mortality rates for the north of the Netherlands are less similar than the south of the Netherlands when it is compared to the smoking-related mortality rates. When looking at the map of all-cause mortality, Zuidoost-Friesland falls in the category that is similar to the average of the Netherlands. This is not the case for the same region in the map of smoking-related mortality. Zuidoost-Friesland falls here in the category of the lowest mortality rates.

For males, the all-cause mortality rates and smoking-related mortality rates also have the same highest and lowest mortality rates. The region with the highest mortality rate is Oost-Groningen; the region with the lowest mortality rate is Delft en Westland. There are some differences when comparing the maps of all-cause mortality rates to smoking-related mortality rates, but these differences are not very large and could be explained by the way the maps were classified.

When looking at the maps for females, Zuid-Limburg has the highest mortality rate for all-cause mortality, whilst Groot-Amsterdam has the highest rate for smoking-related mortality. This is a very big difference, because when looking at the all-cause mortality rates, the region of Groot-Amsterdam falls in the category that is similar to the average of the Netherlands. For the lowest mortality rates, both all-cause mortality and smoking-related morality have their lowest rate located in the region of Overig Zeeland. For all-cause mortality, the region of Kop van Noord-Holland is in the category that is just higher than the average of the Netherlands. When looking at smoking-related mortality, the Kop van Noord-Holland falls in the category that is just lower than the average of the Netherlands. For the region of Delfzijl en omgeving, this is the other way around. Another very striking difference is that there is a cluster of regions in the north of the Netherlands that fall in the category of the lowest smoking-related mortality rates when compared to the average of the Netherlands. There is no such cluster for all-cause mortality.

Figure 4.20 Age standardized all-cause mortality rates and smoking-related mortality rates for males and females, males, and females per 10,000 inhabitants 2004-2008



4.5.2 comparison of the percentages to rates

Figure 4.22 shows the percentage of smoking-related mortality, and the all-cause mortality rates per 10,000 of the population for males & females, males, and females in the period of 2004-2008.

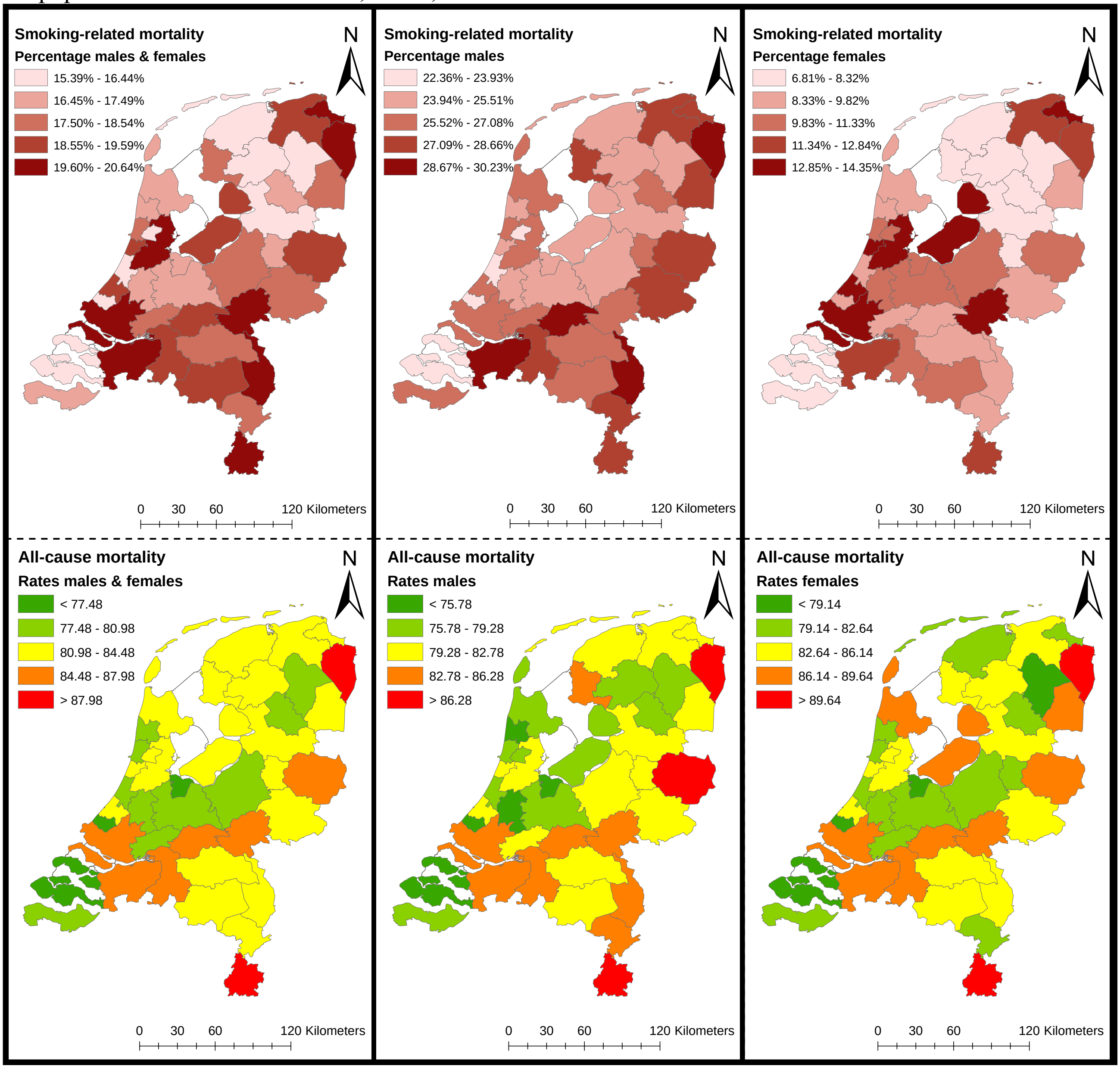
When looking at figure 4.22 for males and females, it can be seen that approximately 15 to 20 % of the all-cause mortality is caused by smoking-related mortality. The regions of Oost-Groningen and Zuid-Limburg fall both for the percentage as for the rates in the highest categories. When looking at the regions of Delfzijl en omgeving, Groot-Amsterdam, and Noord-Limburg it can be seen that the smoking-related mortality percentage falls in the highest mortality category. For the all-cause mortality rates, the regions are close to the average. The regions of Overig Zeeland and Delft en Westland fall both for the percentage as for the rates in the lowest mortality categories. Noord-Friesland, Zuidoost-Friesland, Zuidwest-Drenthe, Zaanstreek fall in the lowest percentage category for smoking-related mortality, whilst for the all-cause mortality the regions are close to the average of the Netherlands.

For males, in figure 4.22, the region of Oost-Groningen is the only region that falls for both the smoking-related mortality percentage, as for the all-cause mortality rate in the highest mortality category. The regions of Overig Zeeland and Delft en Westland fall both for the percentage as for the rates in the lowest mortality categories. There are no striking differences to be found that cannot be explained by the way the categories were classified.

The regions for females show a very different pattern when comparing the percentage of the smoking-related mortality to the all-cause mortality rates. A difference is that the regions of Agglomeratie Haarlem, and Agglomeratie 's-Gravenhage fall in the highest percentage category of the smoking-related mortality and are close to the average of the Netherlands when looking at the all-cause mortality rate. Overig Zeeland, and Noord-Drenthe fall both for the percentage as for the rates in the lowest mortality categories. Zuidwest-Friesland, Zuidoost-Friesland, and Noord-Overijssel fall in the lowest percentage category for smoking-related mortality, whilst for the all-cause mortality the regions are close to the average of the Netherlands. For the regions of Delfzijl en omgeving, the difference is even more apparent. For the percentage of smoking-related mortality the regions is in the highest mortality category, for the all-cause mortality rates the regions is in one of the lower mortality rate categories.

Where regions fall in the highest percentage mortality category for smoking and in the middle category for the all-cause mortality rates, it could mean that the smoking-related mortality has a big influence, but that this it is not seen back in the all-cause mortality rates, because of lower mortality rates from non-smoking mortality. It could also be the other way around that regions fall in the lowest percentage of smoking-related mortality category, and in the category that is close to the average of the Netherlands. Here, the non-smoking-related mortality could be higher.

Figure 4.21 Percentage of smoking-related mortality and the all-cause mortality rates per 10,000 population for males & females, males, and females 2004-2008



4.5.3 comparing spatial autocorrelation to the rates

When comparing the results of the spatial autocorrelation to the variations and patterns that were described in paragraph 4.1, 4.2, and 4.3, there were some similarities: The all-cause mortality for males and females, males, and females did not show very distinct clustering in both the map and the Global Moran's I. The COPD mortality did show a distinct clustering in the map for males and also in the spatial autocorrelation. The IHD mortality for males also showed a clustering in the map as well as in the Moran's I. For females, the lung cancer mortality showed a very distinct cluster on the map and also in the spatial auto correlation, this was also the case for females in smoking-related mortality, only the p-value is in the 90% confidence interval.

There also were some differences when comparing the maps and the spatial autocorrelation: Were there was no very distinct cluster found in the COPD mortality map for males and females, there was a high significance to be found in the spatial autocorrelation. This is also the case for the IHD mortality of males and females. In lung cancer mortality for males, as well as for males and females, there is not a very distinct clustering to be found on the map, but there are significant clusters to be found in the Global Moran's I. This is also very much the case for males when looking at smoking-related mortality. When comparing the Moran's I of all-cause mortality to the smoking-related mortality, it would be expected that they are both significantly clustered, or both randomly distributed. This is not the case for males. The Moran's I showed a very significant clustering in smoking-related mortality, but not for all-cause mortality.

4.5.4 Indexed variances

The CMF's were calculated in order to find out what the variances for the different causes of mortality were. The reason why the CMF's were used, and not the mortality rates is because it is hard to compare them to each other, they do not have an equal average value. The CMF Does have this, the average for the total population is set to 100, and the subpopulations are above or below this number in percentages.

When smoking-related mortality was excluded from the variance, the variance was substantially lower for males and females, males, and females as can be seen in the table 4.6. This can be seen by comparing the non-smoking-related mortality to the all-cause mortality. For example: When looking at the males and females, the all-cause mortality has a variance of 16. The non-smoking-related mortality has a variance of 9; this lower variance means that the values are closer together. The values of the all-cause mortality are thus more spread out because of the influence of the smoking-related mortality.

Table 4.6 The variance of the different kinds of mortality for males & females, males, and females 2004-2008

Males & females	**Minimum**	**Maximum**	**Mean**	**Variance**
Smoking-related mortality	76	120	97	128
Non-smoking-related mortality	94	107	100	9
All-cause mortality	91	108	99	16
Chronic obstructive pulmonary disease mortality	73	128	99	192
Ischaemic heart disease mortality	85	127	100	91
Lung cancer mortality	84	117	99	72
Males	**Minimum**	**Maximum**	**Mean**	**Variance**
Smoking-related mortality	80	122	99	128
Non-smoking-related mortality	92	107	99	11
All-cause mortality	90	111	99	23
Chronic obstructive pulmonary disease mortality	71	134	100	233
Ischaemic heart disease mortality	85	130	100	98
Lung cancer mortality	83	120	100	82
Females	**Minimum**	**Maximum**	**Mean**	**Variance**
Smoking-related mortality	55	132	92	396
Non-smoking-related mortality	93	107	100	13
All-cause mortality	90	107	99	16
Chronic obstructive pulmonary disease mortality	66	133	97	265
Ischaemic heart disease mortality	82	125	100	113
Lung cancer mortality	67	122	95	221

4.5.5 Variance, covariance, and correlation

Tabel 4.7 also shows the variance of smoking-related mortality, non-smoking-related mortality, and all-cause mortality for males and females, males, and females. The variance with the normal rates cannot be compared the way the indexed variance can. But what can be seen is that the variance for female- and male and female non-smoking-related mortality is lower than the variance of female- and male and female non-smoking-related mortality. This would not be expected when looking at the indexed variance in table 4.6. For males it is the other way around, the variance of smoking-related mortality is higher and the non-smoking-related mortality is lower, this is also the case for the indexed variance for males.

The covariance for smoking-related mortality and non-smoking-related mortality in table 1 indicates that for males, females, and males and females the correlation is positive. For males the covariance has the highest value of 2.82. Males also have a very high positive correlation of 0.599 and when looking at de significance level it shows that the correlation is significant at the 0.01 confidence level. For males and females together, the covariance is 1.99. The correlation does not differ very much from the correlation value for males with 0.561 and is also significant at the 0.01 confidence level. Females have a very low covariance value of 0.47, and the correlation value for females is 0.097. From table 4.7 it can be seen that there is no significant correlation between female smoking-related mortality and female non-smoking-related mortality. For males and females together and males separately, the positive correlation means that as there is an increase in smoking-related mortality, there also is an increase in non-smoking-related mortality.

Table 4.7 The variance, covariance, and correlation for males & females, males, and females 2004-2008 by rates of 10,000 of the population

		Males and females	**Males**	**Females**
Variance of population	Smoking-related mortality	3.00	5.92	3.43
	Non-smoking-related mortality	4.20	3.75	7.11
	All-cause mortality	11.18	15.32	11.49
Covariance of population between smoking-related and non-smoking-related mortality		1.99	2.82	0.48
Pearson Correlation smoking-related mortality and non-smoking-related mortality		**0.561****	**0.599****	0.097
Sig. (2-tailed)		0.000	0.000	0.552

** Correlation is significant at the 0.01 level (2-tailed)
* Correlation is significant at the 0.05 level (2-tailed)

5. Conclusion

5.1 Summary of the results

<u>All-cause mortality</u>

Oost-Groningen and Zuid-Limburg are the regions with high all-cause mortality rates for males and females, males, and females. For the males and females, the two regions with the highest mortality rates are neighboured by regions with lower all-cause mortality rates. For males, all of the regions that have the highest mortality rates border to Germany or Belgium. When looking at the lowest all-cause morality rate, the region of Overig Zeeland has the lowest rate for males and females, and for females. Delft en Westland has the lowest all-cause mortality rate for males. For males and females together, and separately, the middle part of the Netherlands shows a cluster of regions that are lower than the average where $p < 0.01$. Around this cluster, there are is a line of regions that have a higher significance than the average $p < 0.01$. None of the Moran's Indexes for all-cause mortality are significantly different; this means that the mortality rates are randomly distributed across the different COROP-regions.

<u>COPD mortality</u>

The highest COPD mortality rate for males and females, and males is Zuidwest-Overijssel. For females the region of Flevoland has the highest COPD mortality rate. For males, there is a sort of cluster of higher mortality rates in the eastern and southern part of the Netherlands, this also shows in the map showing the significance. Overig Zeeland has the lowest mortality rate for males and females, and for females. The region of Zaanstreek has the lowest mortality rate for males. The western part of the Netherlands shows a cluster of lower mortality rates for males. For females, the region of Zuid-Limburg is shows a low mortality rate when the COPD mortality rate is compared to other mortality causes. The COPD mortality rates for males and females, and males are significantly clustered. The COPD mortality rates for females are completely random distributed. Regions surrounded by other regions in the local Moran's Index show similarity for both males and females together.

<u>IHD mortality</u>

The highest IHD mortality rate for males and females is the region of Oost-Groningen; this is also the case for males. For females, the highest IHD mortality rate can be found in Zeeuwsch-Vlaanderen. The region with the lowest IHD mortality rate for males and females, and males is the Agglomeratie Leiden en Bollenstreek. Females have the lowest IHD mortality rate situated in Zuidwest-Overijssel. The regions that are located in the mid-west of the Netherlands show a sort of cluster. This is the case for males and females together, and separately. For males and females, and males the IHD mortality rates are significantly clustered. The Moran's Index of IHD mortality for females is randomly distributed. Both for males and females, as for males in the local Moran's I shows clusters that are located in somewhat the same areas: In the mid-west, and in the north-east of the Netherlands.

<u>Lung cancer mortality</u>

The highest lung cancer mortality rate can be found in the region of West-Noord-Brabant for males and females, and males. For females the highest lung cancer mortality rate can be found in Arnhem/Nijmegen. Overig Zeeland has the lowest lung cancer mortality rate for males and females, Zaanstreek for the males, and Zuidoost-Friesland for the females. For females, a very distinct cluster for mortality rates that fall in the lowest mortality category could be found in the north of the Netherlands, this could also be seen when looking at the significance map. The lung cancer mortality rates for males and females, males, and females are all significantly clustered. For males and females together it shows that the regions in the south and north of the Netherlands are significantly surrounded by regions with similar values. For males, the clustering is mostly located in the mid-west and south of the Netherlands. For females this is very different, in the north of the Netherlands a very distinct cluster can be found.

Smoking-related mortality

The highest smoking-related mortality rates for males and females, and males can be found in Oost-Groningen. For females, the highest smoking related mortality rate can be found in Groot-Amsterdam. The lowest smoking-related mortality rates for males and females, and females is located in Overig Zeeland. For males, the lowest smoking-related mortality can be found in the region of Delft and Westland. For females, there is a very distinct cluster up north for low rates in smoking-related mortality; this also shows in the significance maps and the local Moran's Index. The mortality rates of smoking-related mortality for males and females separately are significantly clustered.

Regional differences in mortality in the Netherlands attributed by smoking

When comparing the mortality rates of all-cause mortality and smoking-related mortality for the males and females there is some resemblance in the variations of the maps. The variations in all-cause mortality rates for the north of the Netherlands are less similar than the south when compared to smoking-related mortality. For males, the all-cause mortality rates and smoking-related mortality rates have the same highest and lowest mortality rates. There are some differences between the all-cause mortality maps and the smoking-related mortality rates, but they are not very large. When looking at the maps for females, Zuid-Limburg has the highest mortality rate for all-cause mortality, whilst Groot-Amsterdam has the highest rate for smoking-related mortality. For the lowest mortality rates, both all-cause mortality and smoking-related morality have their lowest rate located in the region of Overig Zeeland. A very striking difference is that there is a cluster of regions with low mortality rates in the north of the Netherlands, There is no such cluster for all-cause mortality.

For males and females, the regions of Oost-Groningen and Zuid-Limburg fall both for the percentage on smoking-related mortality as for the all-cause mortality rates in the highest categories. The regions of Overig Zeeland and Delft en Westland fall both for the percentage as for the rates in the lowest mortality categories. For males there are no very striking differences. The regions for females show a very different pattern when comparing the percentage of the smoking-related mortality to the all-cause mortality rates.

When smoking-related mortality was excluded from the indexed variance, the variance was substantially lower for males and females, males, and females. The covariance between smoking-related mortality and non-smoking-related mortality showed a positive correlation for males, females, and males and females. The correlation for males and females, and males was high and significant, but for females there was no correlation to be found.

5.2 Discussion of findings

Regional variations

For all-cause mortality, smoking-related mortality, and the top three smoking-related causes of death, for the different sexes it can be said that regional variations have been found. The regions of Zuid-Limburg and Oost-Groningen show high mortality rates for multiple causes of mortality and also for the different sexes. One exception is the COPD mortality for females; Zuid-Limburg has a lower mortality rate than the average of the Netherlands here. The high mortality in these two regions when looking at the different causes in mortality could be due to the fact that certain types of industry were located in these regions. There is a possibility this could have caused higher rates in all-cause mortality as well as in lung cancer mortality/smoking-related mortality. This being said, because the lung cancer mortality was used in order to calculate the smoking-related mortality. Kunst et al. (1993) stated that differences in lung cancer mortality are likely to be caused by tobacco consumption, but a large share is still unexplained.

Regional patterns

Regional patterns were found for the different causes of mortality and sexes. For males and females, The COPD- IHD- and lung cancer mortality were significantly clustered. For males, the smoking-related mortality, the COPD-, IHD-, and lung cancer mortality are significantly clustered. The only significant values for females are the ones for smoking-related mortality and for lung cancer mortality. One thing that should be kept in mind when looking at the results of spatial autocorrelation, is that the

Moran's I is affected by the geographical scale and the extent of the study area. When there are few features, areas that are not similar could be close to each other. Also, the distance method and distance band that are used have a big influence on the results of the spatial autocorrelation (Mitchell, 2005).

The extent that smoking-related mortality contributes to all-cause mortality

When comparing the mortality rates of all-cause mortality and smoking-related mortality for the males and females, the variations between all-cause mortality and smoking-related mortality rates for the north of the Netherlands are less similar than the south of the Netherlands. For males, there are some minor differences when comparing the maps of all-cause mortality rates to smoking-related mortality rates. These differences could be explained by the way the maps have been classified, this is because of differences in rates; all-cause mortality rates are much higher than those of smoking related rates, the classification groups have different ranges. Also, the average of the Netherlands has an influence on the range. When looking at the maps for females, it seems the smoking-related mortality has a much smaller impact. The maps show a lot of difference. A very striking difference is that there is a cluster of regions with low mortality rates in the north of the Netherlands, There is no such cluster for all-cause mortality.

When comparing the results of the spatial autocorrelation to the variations and patterns that were shown in the specific mortality rate maps, there were some similarities, but also differences. Sometimes a significant clustering was shown in the Moran's I, which could not be seen on the map. This difference could also be due to the way the maps were classified, as mentioned before. When comparing the Moran's I of all-cause mortality to the smoking-related mortality, it was expected that they were both significantly clustered, or both randomly distributed. This was very much not the case for males. The Moran's I showed a very significant clustering in smoking-related mortality, but not for all-cause mortality.

When smoking-related mortality was excluded from the variance, the variance was substantially lower for males and females, males, and females. This can be seen by comparing the non-smoking-related mortality to the all-cause mortality. The results of the covariance and correlation for females were to be expected, because from the maps it already showed that the all-cause mortality had different patterns than smoking-related mortality. For males on the other hand, the smoking-related mortality rates showed similar patterns to all-cause mortality rates. The smoking-related mortality thus seemed to have an influence on the all-cause mortality patterns. The results of the correlation between smoking-related mortality and non-smoking-related mortality showed that as one increased, the other also increases. Thus, there are other causes in mortality that are not included in the smoking-related mortality rates that have a strong influence on the all-cause mortality rates in the different COROP-regions of the Netherlands.

Concluding it can be stated that smoking-related mortality has an influence on the variations in all-cause mortality, this can be concluded both from comparing the maps and the indexed variance. The covariance and the correlation show somewhat different results. It seems that when looking at the maps, the most similarities could be found for males, and for males and females in a lesser degree. The results indicate that there is still a lot that can be done to reduce the smoking-related mortality influencing the all-cause mortality, especially for males, but there are also other causes in mortality that are not included in the smoking-related mortality rates that have a strong influence on the all-cause mortality rates.

5.3 Recommendations

The regions in relationship to the highest smoking-related mortality rates and all-cause mortality rates should be watched in order to find out what changes can be made to bring the high mortality rate down to a more acceptable level. Though, it must not be the case that the focus in lowering the smoking-related mortality is only on the regions that have both high all-cause mortality rates and high smoking-related mortality rates. There should even be more of a focus on the regions that fall in the highest mortality category for smoking-related mortality, and in the middle or lower category for the all-cause mortality rates; This means that the smoking-related mortality can have a big influence when compared to other surrounding regions and that there is still a lot of room for improvement. It could also be that regions fall in the lowest smoking-related mortality category percentage, and in the category that is close to the average of the Netherlands for all-cause mortality. Here, the non-smoking-related mortality is higher, and other solutions could be found.

In order to reduce smoking-related mortality in a region, they could take the neighbouring regions with low mortality rates as their example and cooperate to find out how to make improvements. To get even a better image of the variations between regions, it would be interesting for new research to find out if the variations in smoking-related mortality between regions are becoming smaller, or if the variations are becoming bigger; in other words convergence or divergence? Such a research has been done on a national level by van der Wilk and Jansen (2004). A comparison over time would be needed to get to these results. This result would make even clearer on how the policy making could be implemented and get the high mortality rates in certain regions down.

Literature references

Allen, D.W. (2009), *GIS tutorial two: Spatial analysis workbook.* ESRI Press: Redlands, California.

ArcGIS Desktop 9.3 Help (2009), 'Spatial Autocorrelation'. Internet: http://webhelp.esri.com/arcgisdesktop/9.3/index.cfm?id=2150&pid=2146&topicname=Spatial_Autocorrelation_(Morans_I)_(Spatial_Statistics). Last visited on Februari 10, 2010.

Doll, R. and A. Hill (1956), 'Lung cancer and other causes of death in relation to smoking: a second report on the mortality of British doctors'. *British Medical Journal*. 2 (5001), pp. 1071-81.

European Commission (2003), 'Regions - Nomenclature of territorial units for statistics - NUTS - 2003/EU25'. Office for Official Publications of the European Communities, Luxembourg.

Gatrell, A.C., (2002), *Geographies of health: An Introduction*. Blackwell Publishers Ltd, Oxford, UK.

Gelder, B.M., R.T. Hoogenveen and H.M.J.A. Leent-Loenen (2010), 'Wat zijn de mogelijke gezondheidsgevolgen van roken?/What are the possible health consequenses of smoking?' Internet: http://www.nationaalkompas.nl/gezondheidsdeterminanten/leefstijl/roken/wat-zijn-de-mogelijke-gezondheidsgevolgen-van-roken/. Last visited on March 20, 2010.

Giesbers, H. (2008), 'COROP-gebieden 2008/COROP-areas 2008'. Internet: http://www.rivm.nl/vtv/object_map/o2354n22203.html. Last visited on March, 2010.

Hummer, R.A., C.B. Nam and R.G. Rogers (1998), 'Adult mortality differentials associated with cigarette smoking in the USA'. *Population Research and Policy Review* 17 (3), pp. 285-304.

Hunter, G. (Smoke picture processed in front page) (2009), Internet: http://cloudcycle.net/log/tag/oumres/. Last visited on March 27, 2010.

Janssen, F., A. Kunst and J. Mackenbach (2007), 'Variations in the pace of old-age mortality decline in seven European countries, 1950–1999: the role of smoking and other factors earlier in life'. *European Journal of Population* 23 (2), pp. 171-88.

Janssen, F., W.J. Nusselder, C.W.N Looman, J.P. Mackenbach and A.E. Kunst (2003), 'Stagnation in Mortality Decline Among Elders in The Netherlands', *Gerontologist,* 43 (5), pp. 722-34.

John, U. and M. Hanke (2003), 'Tobacco- and alcohol-attributable mortality and years of potential life lost in Germany'. *European Journal of Public Health* 13, pp. 275-77.

Kaiser S, Gommer AM (2007), 'Smoking – Summary'. EUPHIX, EUphact. RIVM: Bilthoven. Internet: http://www.euphix.org/object_document/o4757n27423.html. Last visited on March 20, 2010.

Kirk, D. (1996), Demographic transition theory, Population Studies, 50 pp. 361-87.

Kunst, A.E., C.W.N. Looman, and J.P. Mackenbach (1993), 'Determinants of regional differences in lung cancer mortality in the Netherlands'. *Social science & medicine* 37 (5), pp. 623-32.

Kunst, A.E., C.W.N. Looman, and J.P. Mackenbach (1988), 'Medical care and regional mortality differences within the countries of the European community'. *European Journal of Population* 4 (3), pp. 223-45.

Kunst, A.E., C.W.N. Looman, and J.P. Mackenbach (1990a), *Regionale sterfteverschillen en sociaal-economische indicatoren/Regional differences in mortality and socio-economic indicators.* Instituut Maatschappelijke Gezondheidszorg, Erasmus Universiteit Rotterdam: Rotterdam.

Mackenbach, J.P., A.E. Kunst, C.N.W. Looman, and E.F. Beek, van (1990), 'Regionale sterfteverschillen: Een bundel ten behoeve van het symposium 'Regionale sterfteverschillen'/Regional differences in mortality: a bundle for the symposium 'Regional mortality differences''. november 9th, Rotterdam.

Mitchell, A. (2005), *The ESRI guide to GIS analysis: Volume 2: Spatial measurements and statistics.* ESRI Press: Redlands, California.

Norušis, M. J. (2002), *SPSS 11.0 Guide to data analysis*. Prentice Hall: Upper Saddle River, New Jersey.

Ochsner, A. DeBakey, M. (1939), 'Treatment by total pneumonectomy: analysis of 79 collected cases and presentation of 7 personal cases, symposium on cancer primary pulmonary malignancy'. *Surgery, gynaecology and obstretics*. 68 pp. 435-51

Omran, A.R. (1998), 'The epidemiologic transition theory revisited thirty years later', *World Health statistics Quarterly,* 51, pp. 99-119.

Peto R, A.D. Lopez, J. Boreham, M. Thun and C. Heath (1992), 'Mortality from tobacco in developed countries: indirect estimation from national vital statistics'. *Lancet* 339 (8804), pp. 268-78.

Preston, S. H., P. Heuveline, and M Guillot (2001), *Demography: Measuring and modelling population processes*. Blackwell: Oxford.

RIVM (2006), 'Volksgezondheid toekomst verkenning 2006/ Public Health Status and Forecasts'. Rijksinstituut voor Volksgezondheid en Milieu: Bilthoven. Internet: http://www.rivm.nl/vtv/object_class/sr_vtv06_nationale_onderneming.html. Last visited on Februari 10, 2010.

Rogers R. G., R.A. Hummer, P.M. Krueger and F.C. Pampel (2005), 'Mortality Attributable to Cigarette Smoking in the United States'. *Population and Development Review* 31(2), pp. 259-92.

Ruwaard, D., Kramers, P.G.N. (1993), *Volksgezondheid toekomst verkenning : de gezondheidstoestand van de Nederlandse bevolking in de periode1950-2010/ Public Health Status and Forecasts: the health status of the Dutch population in the period of 1950-2010.* SDU: 's-Gravenhage.

Schrijvers, C.T.M. (2007), Wat is een regionale VTV?/What is a regional VTV? Volksgezondheid Toekomst Verkenning/Public Health Status and Forecasts. Bilthoven: RIVM, Internet: http://www.toolkitvtv.nl. Last visited on November 16, 2010.

Single, E., L. Robson, J. Rehm and X. Xi (1999), 'Morbidity and mortality attributable to alcohol, tobacco, and illicit drug use in Canada'. *American Journal of Public Health* 89 (3), pp. 385-90.

Statistics Netherlands (2010), 'Doodsoorzakenstatistiek/Statistics in causes of death'. Internet: http://www.cbs.nl/nl-NL/menu/methoden/dataverzameling/doodsoorzakenstatistiek.htm. Last Visited on March 26, 2010.

Statistics Netherlands (2010), 'Mortality; sex, age, marital status and region'. Statline. Internet: http://statline.cbs.nl/StatWeb/publication/?DM=SLNL&PA=03747&D1=6-7&D2=17-56&D3=l&VW=T. Last visited on March 26, 2010.

Stattrek (2009), AP* Statistics Tutorial: Difference between proportions. Internet: http://stattrek.com/AP-Statistics-4/Difference-Proportions.aspx?Tutorial=AP. Last visited on March 20, 2010.

U.S. Department of Health and Human Services (1989), 'Reducing the health consequences of smoking: 25 years of progress'. A report of the surgeon general. Centres for Disease Control, Centre for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health, DHHS Publication No. (CDC) 89-8411.

Victoria University (2010), 'Describing variability'. Internet: http://tls.vu.edu.au/vucollege/learninghub/degreestudysupport/psychology_statistics/variability.html. Last visited on March 30, 2010.

Volksgezondheid Toekomst Verkenning (2002), 'Gezondheid op koers?/Health on the right track?'. Rijksinstituut voor volksgezondheid en milieu, Bilthoven.

Volksgezondheid Toekomst Verkenning (2006), 'Zorg voor gezondheid/Care for health care'. Rijksinstituut voor volksgezondheid en milieu, Bilthoven.

Volksgezondheid Toekomst Verkenning (2010), 'Toolkit Regional VTV'. Internet: http://www.rivm.nl/vtv/object_class/tkit_toolkit_ind_determinanten.html. Last visited on March 16, 2010

Willemsen, M.C., Leent-Loenen, H.M.J.A. (2007), 'Hoeveel mensen roken/how many people smoke?' Internet: http://www.rivm.nl/vtv/object_document/o1209n19085.html. Last visited on November 1, 2008.

Wolleswinkel-van den Bosch, J.H., Looman, C.W.N., Poppel, van F.W.A. & Mackenbach, J.P. (1997), 'Cause-Specific Mortality Trends in The Netherlands, 1875-1992: A Formal Analysis of the Epidemiologic Transition', *International journal of epidemiology : official journal of the International Epidemiological Association,* 26 (4), pp. 772-781.

World Health Organization (2008), '2008-2013 Action Plan for the Global Strategy for the Prevention and Control of Noncommunicable Diseases'. Internet: http://www.who.int/nmh/publications/ncd_action_plan_en.pdf. Last visited on March 16, 2010.

World Health Organization (2007), 'International Statistical Classification of Diseases and Related Health Problems 10th Revision'. http://apps.who.int/classifications/apps/icd/icd10online/. Last visited on Februari 10, 2010.

Young, T.K. (1998), *Population health: concepts and methods*. New York/Oxford: Oxford University press. Chapter 4: Modeling determinants of population health. pp. 95-125.

Zandwijk, N., Leeuwen F.E. (2005), 'Welke factoren beïnvloeden de kans op longkanker?/What factors influence the chance on getting lung cancer?' Internet: http://www.rivm.nl/vtv/object_document/o1069n17274.html. Last visited on November 1, 2008.

Zwakhals, S.L.N., H. Giesbers, F.R.J. Hertog, den (2009), Achtergronddocument bij de totale sterfte en de sterfte naar doodsoorzaak per GGD-regio 2003-2006/Background document on the Total mortality and mortality by cause per GGD-region 2003-2006. Internet: http://www.rivm.nl/vtv/object_map/o2704n21682.html. Last visited on Februari 10, 2010.

Zantinge, E.M. (2009) 'Zijn er binnen Nederland verschillen naar regio?/Are there differences within the Netherlands between regions?' Volksgezondheid Toekomst Verkenning, Nationaal Kompas

Volksgezondheid. Bilthoven: RIVM. Internet: http://www.nationaalkompas.nl. Last visited on Januari, 2010.