Factors Effecting Regional Variation in Cardiovascular Mortality

A Study in the North of the Netherlands

Bachelor Thesis:
Vincent de Vegt
S2517094
Supervisors: O. Couwenberg & A. Imperiale
June 2017

Word Count: 6569

Table of Contents

Abstract

Cardiovascular disease (CVD) will be the leading cause of death in the Netherlands in the near future. 90% of CVD mortality can be attributed to nine modifiable risk factors. The mortality caused by CVD varies significantly throughout the North of the Netherlands. In order to reduce the CVD mortality rate in the North of the Netherlands it is important to know what the underlying cause is within the region. This study aims to determine whether or not the variation in CVD mortality can be explained by a selection of the modifiable risk factors. With the aid of multiple linear regression an attempt was to determine if the following factors contributed to the CVD mortality rate; high blood pressure, overweight, diabetes, cholesterol, smoking and physical inactivity. Only blood pressure resulted in a significant outcome, the remaining variables were insignificant. The results do however indicate that there is considerable regional variation in the factors.

1. Introduction

1.1 Background

Cardiovascular disease (CVD)[1] will be the leading cause of morbidity worldwide by 2030, accounting for nearly 25% of all deaths (Kreatsoulas & Anand, 2010). The Netherlands will be no exception and it is expected that CVD will become the leading cause of death in the Netherlands once again (Muller-Nordhorn et al, 2008). This relative increase will result in CVD becoming the disease with the highest burden on society also in terms of cost. The current CVD related costs for the Dutch healthcare already exceed 8 billion euros per year (CBS, 2014). There are a number of factors that contribute to the prevalence of CVD, categorized into two categories; genetic factors and population-attributable risk factors (Kreatsoulas & Anand, 2010). The INTERHEART study found that over 90% of the population-attributable risk factors could be divided into nine different modifiable risk factors (Yusuf et al 2004). As shown by the results of this study, these factors alone have enough influence on the prevalence of CVD that they should be key focus points in any government prevention program. The factors are commonly known as the "social-determinates of cardiovascular disease" these are; smoking, diabetes, hypertension, obesity, psychosocial factors, fruit/vegetable consumption, physical activity, alcohol consumption and apolipoprotein A/B ratio (Kreatsoulas and Anand, 2010, Yusuf et al., 2004).

The Netherlands is one of the countries with the lowest mortality rate related to CVD. However, concerning this issue, within the Netherlands significant differences on a regional level are still detectable. We believe that, in order to understand these regional differences, it is important to properly understand, region by region, which among the aforementioned social-determinates are most related to the prevalence of CVD. This paper will focus on the north of the Netherlands (Friesland, Groningen and Drenthe) and it will analyze, at regional level (municipal). A number of these factors and of their likely relationship to the prevalence of CVD. A selection has been made in the determinants, this is because factors such as alcohol consumption, fruit/vegetable consumption and psychosocial factors are difficult to accurately measure and relate to CVD at the scalar level that is needed. The specific factors that will be taken into consideration are; smoking, hypertension (blood pressure), obesity, physical activity, diabetes and the apolipoprotein A/B ratio (cholesterol). The results can indicate which key social determinate any government and municipal prevention policy and

[1] Cardiovascular disease is a class of diseases affecting the heart and/or the blood vessels.

program should focus on in order to further reduce the mortality rate of CVD in the North of the Netherlands.

1.2 Research Problem

With respect to CVD mortality there are many factors that contribute to the prevalence of CVD in a region. To decrease the mortality rate and reduce the overall cases of CVD it is important to understand why certain regions have either a higher or lower amount of CVD cases. The research question of this thesis is therefore: *Which of the chosen social-determinates (factors) of cardiovascular disease are responsible for the regional variation of the cardiovascular mortality rate in the North of the Netherlands?*

The sub-questions that will aid in the answering of the research question are as follows:

- How strong is the relation between the factors and the regional CVD mortality rate?
- Which factor relates the most to the CVD mortality rate?

1.3 Structure of the Thesis

The thesis will be structured as follows; first the relevant theories will be discussed in the Theoretical framework section of the thesis. Here the research topic supported with the aid of various similar studies conducted around the topic of CVD and regional variation in disease. The following section the Methodology for the collection and analysis of the data will be explained. Specifically, the choices that were made will be explained. This section is followed by the presentation of the Results. The results will be discussed in the Discussion where the possible relations will be explained and discussed. The final section will be the Conclusion where the results and discussion will be shortly summarized and the research question answered.

2. Theoretical framework

2.1 Theory related to regional variation in health.

The geographical spread of a disease is often related to the variation of health quality at a regional or local level (Curtis & Jones, 1998) (Williams, 2003). Regional variation of a disease within a country is often shown to illustrate health inequality. Within the Netherlands, nearly every disease will display a significant variation from the national average when regional totals are compared to one another.

When studying regional health variation there are two distinct methods through which the results can be interpreted; compositional and contextual. As stated by Curtis & Jones (1998) compositional effects are the result of the variation in distribution of certain types of people whose individual characteristics directly influence their health. People who only have a high-school diploma, for example, are generally associated with having below average health due to factors such a smoking, drinking, inactivity and an unhealthy diet (Marmot, 2005).

A compositional perspective would infer that people would have a similar health experience irrespective of where they live. Theories focusing on “contextual effects” state that the health experience of similar individuals depends partly on the social and physical environment where they live. This would mean that two individuals who are identical with respect to heath experience would have a different health status, (i.e. A person belonging to the healthiest 10% in one region and may only belong to the healthiest 75% of another region (Curtis & Jones, 1998).

Population composition is one factor that can explain the variation in CVD prevalence within the Netherlands (Krieger et. al, 2003). According to this perspective, broadly speaking, people with a poor socio-economic background are the most adversely affected by CVD. Variation in the quality of medical care can also affect the CVD mortality rate, and would be a logical example of a contextual effect (Williams, 2003). The theories presented clearly state that there are a number of underlying factors that can affect the regional variation of disease. This study, however, will not try to explain why there is a variation in CVD at a regional level but rather, what factor (Social Determinate) contributes most to the regional variation.

2.2 Social Determinates of Health

As mentioned previously, this study will focus on the regional variations in mortality rate related to CVD in relation to the prevalence of the indicators known as the social-determinates of health (Kreatsoulas & Anand, 2010, Yusuf et al, 2004). According to the study done by Yusuf et al. (2004) nearly 90% of all cases of CVD can be attributed to certain lifestyle factors, all of which are modifiable. The risk factors are: smoking, diabetes, hypertension, obesity, psychosocial factors, fruit/vegetable consumption, physical activity, alcohol consumption and the apolipoprotein A/B ratio (cholesterol). The social determinates of health reflect the health impact of the social environment on people living in a certain community (Wilkinson & Marmot 2003). Each factor represents the conditions in which people were born, grow, work, live and get old while being affected by the distribution of resources at local level (Kreatsoulas & Anand, 2010). Within the Netherlands it is important to understand the underlying causes of regional CVD variation. The first step in doing so is learning more about the "causes of causes" (Kreatsoulas & Anand, 2010). Regional variation in the underlying risk factors can help explain the regional variation in CVD. This study will investigate if the regional variation is specifically attributable to any of the selected social-determinates. The selected social-determinates will be further referred to as "factors"

2.3 *Regional variation in cardiovascular disease*

Various studies conducted in the last decades, (focused on explaining the regional variation in CVD, both at national and at international level. Two important studies we can mention here are *Regional variation and time trends in mortality from ischemic heart disease: East and West Germany 10 years after reunification* by Muller-Nordhorn et al. (2004) and *North-South Gradients in Britain for Stroke and CHD, Are They Explained by the Same Factors?* by Morris et al. (2003). Also *An update on regional variation in cardiovascular mortality within Europe* by Muller-Nordhorn et al. (2008), this is a study conducted at European level that also provides some interesting insights about the differences between regions within the same country. In this regard, Muller-Nordhorn et al. (2008) explored the Netherlands by also looking through its respective provinces.

Each of the aforementioned studies underlines that there can be significant variation within the same country and among its different regions. However, none of these studies further develops reflections on what the underlying causes may be, and call for further research to be conducted on the possible causes. The importance of knowing what the underlying causes are is mentioned in Muller-Nordhorn et al. (2008). In their work they emphasize that the underlying causes dictate the type of approach that must be taken to eliminate the health inequality.

This study will test the variation in the North of the Netherlands of the various social-determinates against the mortality rate of CVD. From the findings that will be provided by this research, one can conclude if different approaches need to be taken to further improve public health and the government prevention policy and program in the North of the Netherlands.

2.4 Conceptual Model

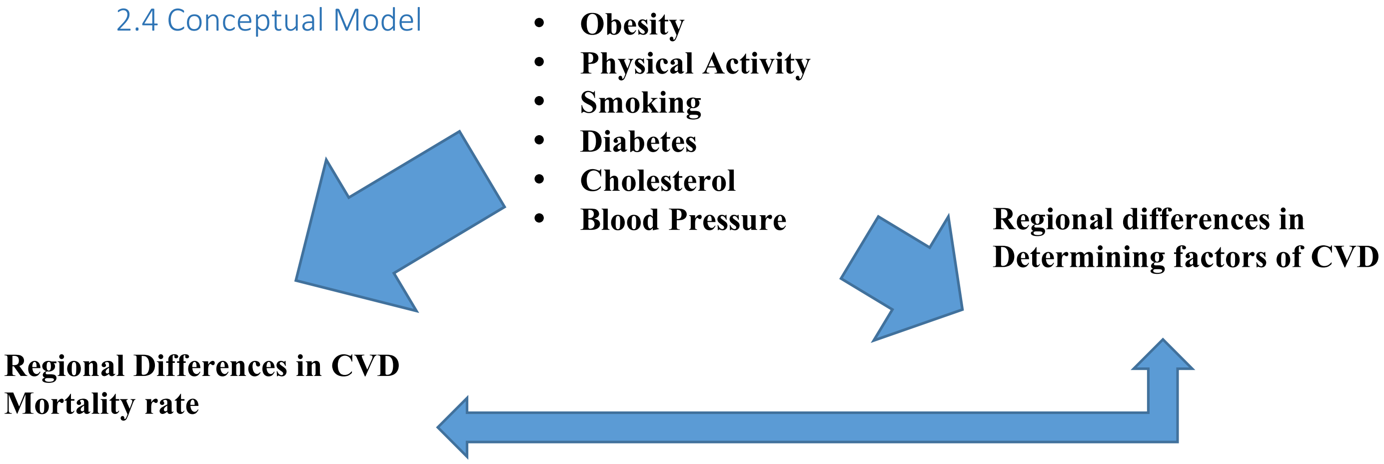


Figure 1: The Conceptual Model

The conceptual model above shows the relation between the regional difference in CVD and the regional difference in the determining factors (obesity, physical activity, smoking diabetes, cholesterol and blood pressure). This study aims to determine which of the factors show the same regional variation as the mortality rate of CVD. The factors that show similar variation over space as the mortality rate can be used to explain the underlying cause of the regional variation in CVD mortality. The model will be tested by measuring the relation between the individual factors and the mortality rate. From there one can conclude if the variation in factors is linked to the variation in CVD mortality.

2.5 Hypotheses

- There will be a significant correlation between smoking and CVD mortality
- There will be a significant correlation between Blood pressure and CVD mortality
- The remaining factors will all result in an insignificant outcome

3. Data & Methodology

3.1 Data Cardiovascular Disease & Factors

This research will focus on quantitative data. The data is secondary data, collecting primary data would not provide more information than currently available via secondary sources. The scale and detail of the secondary data is sufficient to be able to carry out the chosen analysis. The research will rely on data collected primarily by the National Statistical Bureau (CBS) and the National Public Health Institute (GGD). The focus of this research will lay on obesity, physical activity, smoking, diabetes, cholesterol and blood pressure since these are social determinates that can be measured through secondary data, and whose secondary data can be considered the most scientifically reliable. The data for each of these factors is from 2012 and are expressed in absolute figures and as percentages of the population. The calculations will use the 2012 population figures of each municipality. However only the adult population will be used (20 and older) this is because the data regarding the various factors is only about the adult population. All the collected data will be placed in a SPSS dataset for analysis. Table 1 shows the datasets that have been used.

Data	Years	Scalar level	Source
Cardiovascular mortality rate	2012	Municipality	CBS+RIVM
Obesity prevalence	2012	Municipality	CBS+RIVM
Smoking percentage	2012	Municipality	CBS+RIVM
Diabetes	2012	Municipality	CBS+RIVM+ Zorginstituut Nederland
Cholesterol data	2012	Municipality	CBS + Zorginstituut Nederland
Blood Pressure data	2012	Municipality	CBS+RIVM+ Zorginstituut Nederland
Physical Activity	2012	Municipality	CBS+RIVM

Table 1: Shows the various data sets used in the analysis and their respective sources

The CVD mortality figure has been published as the total amount of CVD related to deaths in a municipality in 2012. In order to be able to compare municipalities among each other, this figure must be recalculated. CVD mortality will be expressed as *the amount of CVD related deaths per 1000 inhabitants*. This figure was calculated by first determining the death rate of a municipality per 1000 inhabitants: *100/Inhabitants * total deaths * 10 = Deaths per 1000 Inhabitants.* Then the CVD specific death rate was calculated: *Deaths per 1000 inhabitants/100 * %CVD deaths.* This recalculated figure will be used in the analysis as the dependant variable[2]. The data has to be recalculated to ensure that the analysis will not be affected by the overall size of the municipality. For example; the city of Groningen has the most CVD related deaths, but it also has the largest population by a significant margin. The resulting figure used after the recalculation makes it possible to compare Groningen with a much smaller municipality. Furthermore the data from the four island municipalities has been combined. The *Wadden islands* are small with regard to population when compared to the other municipalities. As a result a single death more/less in a year has a significant impact on the death rate, the same applies to other variables. In order to make use of the data all four of the municipalities have been combined as one municipality. Due to the use of data from 2012 the now dissolved municipalities of *Gaasterlân-Sleat, Lemsterland, Skarsterlân* and *Boarnsterhim* have been included in the analysis. This means there are a total of 59 cases representing 63 municipalities in the dataset.

The remaining independent variables have been collected by the CBS in a nation wide study whose aim is to collect data that are considered relevant to understand public health. The study known as the *Gezondheidsmonitor* (Health Monitor) is conducted every four years and is combined with the data from the *Gezondheidsenquete* (Health Survey). The two studies collected data about every factor that can be of influence on the health of the adult population, and data about health indicators. The aim of the data is to be representative for the adult population in a municipality, therefore it is suitable for this research. For this research the data regarding; overweight, smoking and (in)activity originate from this dataset. The data has been published as percentages of the population practicing/affected by the named factor. In order to make the data comparable with the CVD figure the data has been recalculated as a figure per 1000 inhabitants. For example 33% will become 330 adults per 1000 adults in a municipality. Data regarding physical activity has been based on a norm set by the Dutch government. If people achieve this norm they are considered to be physically active, if not they are considered to be inactive. The percentage of physically active people was subtracted from 100% to create a figure that represents the percentage of inactive people.

For diabetes, blood pressure and cholesterol data were expressed as the total amount of prescriptions for medication in each municipality. Here data was transformed to be expressed as a percentage. Using data on prescriptions gives an accurate impression of the distribution

[2] Map of the CVD mortality rate is locate in the Appendix

of the factor over the population. The data itself is reliable, it has been collected through the information provided by insurance companies about the types of medication used by their customers.

Various types of cholesterol medication are grouped by the CBS into a single category. For blood pressure medication, however, the choice was made to use the two most common types of medication. *Beta-Blockers* and *Diuretics* are generally not used in combination for extended periods with one another (Messerli et al., 2008) and are the first choice of medication given to relive high blood pressure. Their widespread use and universal acceptance makes them ideal indicators for blood pressure data. Although there are other types of medication as well, these are less suitable for our research purposes, since other existing medications are not so specific and can help treat other pathologies or can be used in combination with other medication or their use is extremely limited. The data regarding people with diabetes is also based on the use of diabetes medication. As with the cholesterol data the diabetes medication has been grouped into a single figure by the CBS. The figures of all three factors will be expressed in a figure per 1000 adults.

3.2 Methodology

This research will look for a potential relation between the CVD mortality rate of the municipalities and the various health factors known to contribute to CVD. With the collected data a multiple linear regression analysis will be conducted to find a relation. In a multiple linear regression analysis there are multiple independent variables, in this case we therefore consider as the independent variables the following: smoking rate, overweight rate, inactivity rate, cholesterol medication rate, blood pressure medication rate and diabetes medication rate. The dependant variable is the CVD mortality rate. The independent variables must be either interval or ratio in order to be able to complete such an analysis. Other methods such as Chi-square testing are not suitable as it is not possible to consider the variables completely independent of one another.

Prior to the analysis all data was first collected and transformed into an excel spreadsheet. From there, relevant data was selected and placed into a SPSS Statistics (IBM) dataset. The first step conducted was determining if the relationship between the dependant and independent variables was linear: only linear relationships can be studied with linear regression analysis. The resulting scatter plot matrix indicates that the relationships are linear.

Although some of the relationships are weak, such as the relation between CVD and smoking is only 0.022. An attempt was made to strengthen the linear relationship by calculating the natural logarithm of the smoking data. This however did not strengthen the relationship so the original values were used. Next, the distribution of the variables must be even, the Kolmogorov-Smirnov test indicates that the data of each variable is evenly distributed. The final test is to determine if there is any autocorrelation between the variables. The collinearity diagnostics and the Durbin-Watson test both indicated that this is not the case. The Durbin-Watson test must be under a score of 2.5. The second test to indicate that there is no autocorrelation are the VIF value and tolerance value. A tolerance of >0.1 and a VIF of <10 indicate that there is no autocorrelation that would be harmful to the model (O'Brian, 2007). None of the variables tested had a significant autocorrelation with one another.

Before conducting the multiple linear regression each variable was placed separately in a single variable linear regression. The variables were also placed in a correlation matrix[3], both of these calculations have been done in order to fully understand the relationship between the variables.

3.3 Ethical Considerations

The most relevant ethical issue is that of stereotyping and unfounded assumptions. It is important not to attribute the results specifically to a certain group within the population. Since this research will be conducted on a large scalar level it is important not to generalize the results and attribute them to a specific group, this would be unfounded. This research is also not aimed at studying individuals nor specific groups, it is simply aimed at investigating if there is a relation between the named variables and the regional variation in CVD. The results may have as implication that individuals reading this study automatically attribute the results to a certain group within the population. In order to avoid this, an effort will be made to clarify that data used have been obtained at a large scalar level and cannot be used to implicate the role of individuals.

[3] The correlation matrix is located in the Appendix

4. Results

This section will address the results of the analysis conducted, and illustrate the results with the aid of tables and accompanying statistics. The results and the implications of the results will be discussed in the next section.

As previously mentioned there are 59 cases is this analysis, representing 63 municipalities. All of the independent variables have data for all of the cases, there are no cases with missing data. First each variable was tested individually, the results can be seen in Appendix 1. A correlation matrix was also made and can be seen in table Appendix 2. Two regression analyses will be conducted. The first will force all of the variables into the model, the second will only make use of the variables that have a significant impact on the model.[4]

4.1 Model with all variables

The model will place all of the variables into the regression regardless of the fact if they benefit the result. When looking at tables in the Appendix it would appear that there are variables that will not contribute to the model: these are Smoking and Overweight. The most important aspect of the regression is if it can be considered to be valid. This is confirmed or unconfirmed by whether or not the test is significant. The other relevant figures are R, R^2 and the adjusted $R^{2.}$ In the table below the results of the regression are presented.

R	R^2	Adjusted R^2	Significance
0.463	0.214	0.123	0.043

Table 2: Results of the multiple linear regression with all variables

The regression is significant; this can be seen by the fact that the significance level is just under the threshold of $p<0.05$. Because the result is significant the model can be used later in the discussion to aid in the answering of the research question. The remaining results indicate a few problems with the model. The R score indicates how linear the relationship is within the model. Whereby 0 indicates no linear relationship and 1 a perfect relationship. A score of 0.463 indicates a weak/moderate linear relationship. The R^2 score indicates how much of the variation in CVD mortality is explained by the variables in the model. In this case that would be 21.4% of the variation. However, the adjusted R^2 indicates a much lower figure, namely 12.3%. The fact that the adjusted R^2 score is far lower than that of the R^2 indicates that there are variables that do not contribute to the model, and only make it weaker. Table 3 indicating the B values and the significance levels proves this as well. Namely that all the independent

[4] The data table and the results can be found in the appendix

variables are insignificant, see table 3. This can be caused by the fact that the variables are more related to one another than the collinearity test would indicate.

	B	t	Significance
Constant	3.260	2.776	0.008
Diabetes	0.008	0.543	0.590
Overweight	-0.001	-0.294	0.770
Smoking	-0.002	-0.820	0.416
Blood Pressure	0.009	1.722	0.091
Cholesterol	-0.004	-0.452	0.653
Inactivity	-0.003	-1.985	0.052

Table 3: The B and t values for the regression with their respective P values ($P<0.05$ indicates a significant result)

4.2 Stepwise Regression

As there are variables in the model that do not contribute, the second analysis will be a stepwise regression. The difference with the previous regression is that variables are only included if they significantly contribute to the model. This is done by setting a criterion that SPSS uses to determine whether it includes a variable or not. Inclusion is based on whether the variable increases the F score. Variables are included if they increase F by at least 0.05 and then excluded again if they do not increase F by more than 0.10. The result will be multiple models whereby each time a new variable is added. Only the models with a significant outcome are left. The table below shows the results of the stepwise regression.

Model	Included Variables	R	R^2	Adjusted R^2	Significance
1	Blood Pressure	0.367	0.141	0.126	0.003
2	Blood Pressure and Inactivity	0.446	0.199	0.170	0.002

Table 4: The results of the stepwise multiple linear regression ($P<0.05$ indicates a significant result)

What becomes clear immediately is that only two of the seven variables were selected for the regression. The combination of blood pressure and inactivity produced the strongest model. The R score of model 2 is comparable to that of the first regression and this strengthens the assumption that the other five variables did not contribute to the model. The R^2 however is slightly smaller, this however is not a problem as the adjusted R^2 is much higher than that of the first model. The stepwise model shows that 17.0% of the variation in CVD mortality can be attributed to a combination of high blood pressure and inactivity. The other relevant figures are the B value and the significance of the variables. The B value allows the model to predict what effect a change in the value of an independent variable would have on the dependent variable. One can use the B value for blood pressure as it is significant, the B value indicates that a one-unit increase in blood pressure (1 person more per 1000 people) would result in a 0.008 increase in CVD deaths per 1000 people. Inactivity is also significant as p is not greater than 0.05. However, the B value is negative, something one would not expect.

	B	t	Significance
Constant	2.541	2.925	0.005
Blood Pressure	0.008	3.081	0.003
Inactivity	-0.003	-2.005	0.050

Table 5: The B and t values for the stepwise regression with their respective P values (P<0.05 indicates a significant result)

5. Discussion

The goal of this research was not to determine whether the factors chosen had a relation with CVD. Previously conducted research and a large majority of the medical community agrees that the factors discussed significantly influence the probability of getting CVD. The goal of this research was to determine whether or not the influencing factors were spread over the North of the Netherlands (i.e. Groningen, Friesland, Drenthe) in the same way CVD mortality was. This discussion will first address the factors that had no significant influence and then move to the consequences of the regression analysis.

5.1 Insignificant outcomes

What becomes clear immediately is that only two of the factors addressed namely, blood pressure and inactivity showed their meaningfulness in the regression analysis. Cholesterol and diabetes that showed their relevance in the single variable regression analysis. Overweight and smoking were not significant at any time (Multiple regression, single regression, correlation). In the INTERHEART (Yusuf et al, 2004) study, it has been stated that smoking influences CVD more than any other variable. Yet here there is absolutely no indication that smoking has any relation to CVD mortality. This is due to the fact that the CVD mortality figure was used. The CVD mortality figure indicates how many individuals died with CVD being the cause of death. This figure works perfectly well for factors such as blood pressure, because if a person dies due to an ailment related to high blood pressure the death is classified as a CVD related death. Lung cancer is the leading cause of death among smokers (U.S. Cancer Statistics Working Group, 2016). A hypothetical example of this is a follows: a smoker might be suffering from CVD however if their death is a result of lung cancer or any other disease except CVD, their cause of death will be attributed to something other than CVD. As a result, the illusion is created that smoking is unrelated to CVD mortality, purely based on the figures used this is true.

Being overweight was also insignificant. The most probable explanation for this is that the figure used was based on Body Mass Index (BMI), and that an average of 50.5% of the population is considered to be overweight. A person is considered overweight when they have a BMI of >25, and obese when they have a BMI of >30. Data regarding obesity if available, would have been more suitable than the data about people who are considered to be overweight. The INTERHEART study confirms this, as the use of BMI was also insignificant in their study. Even when using BMI >30. Instead abdominal obesity was used as a variable.

This differs from BMI as it only takes weight that is of negative influence on a person's health into consideration, namely weight caused by excess body fat. The combination of using BMI and the fact that the threshold to be considered overweight is rather low resulted in an insignificant outcome.

The outcome of the stepwise multiple regression analysis showed that among the selected variables only blood pressure and inactivity contributed to the model. Inactivity however does not seem to fit properly. This is because it is known that being inactive negatively effects CVD. Yet the regression would have one believe that more inactivity results in less CVD. Once again the flaw would appear to be the result in the definition of inactivity. The definition used by the CBS states that someone is physically active if they are moderately active for at least 30 minutes, six times per week. Moderate activity is not synonymous with a healthy lifestyle (Marmot, 2005). As a result, the data does not accurately represent the advantage being physically fit has on CVD. To be considered physically fit requires more physically demanding activity to be conducted (Yusuf et al, 2004).

5.2 Significant Outcomes

The only factor that can be considered to be related to CVD in this research is high blood pressure (hypertension)[5]. Along with high cholesterol, high blood pressure is a condition that if left unchecked would significantly increase the chance of a CVD related death. The results of the regression indicate that high blood pressure accounts for 12.6% of the variation in CVD mortality. Not a particularly large figure, but it is rather interesting that the relation is so clearly visible at the used scalar level. Comorbidity is an important factor that must be kept in mind when analyzing the results. The factors analyzed are related each other, this is evident when looking at the correlation matrix[6]. The matrix indicates the interrelatedness between the variables. For example the variables cholesterol and overweight have a correlation of .725. This indicates a strong and significant correlation between these two variables. The strong significant correlations indicate that the figures used are similar, as they are related to one another. The fact that the variables are interrelated likely led to the insignificant outcome for diabetes and cholesterol. Although both variables have a correlation with CVD and test significant in the single variable regression analysis. The use of blood

[5] Map showing the amount of hypertension patients is located in the Appendix
[6] The correlation matrix can be found in the Appendix

pressure medication therefore is a suitable indicator for the overall cardiovascular health of a municipality.

The B value calculated in the regression analysis provides the information necessary to fully understand the effect of high blood pressure on CVD mortality. The B value was 0.008 in the multiple linear regression and 0.009 in the single variable linear regression. The B value indicates that if the blood pressure variable were to increase by one unit the CVD mortality rate per 1000 inhabitants would increase by 0.008. This figure seems rather insignificant, but that is because the CVD mortality figure represents the amount of CVD deaths per 1000 inhabitants, not per 1000 deaths. To illustrate this point, the table below shows what would happen to the CVD mortality rate if the amount of people suffering from hypertension were to decrease by 15%. This could be achieved by reducing salt consumption as stated in He & MacGregor (2002)

Municipality	CVD mortality (2012)	CVD mortality (15% decrease hypertension)	Reduction in deaths (%)	Reduction in deaths (absolute)
Groningen	2.16	1.96	2.3%	7.6
Oldambt	4.78	4.37	2.8%	4.1
Loppersum	4.31	3.98	2.8%	1
Emmen	3.79	4.46	2.5%	8
Leeuwarden	3.35	3.10	2.2%	5.4
Tytsjerksteradiel	4.04	3.76	1.9%	1.9
Assen	3.76	3.48	2.3%	4.3
Heerenveen	3.94	3.65	2.3%	3
Gaasterlân-Sleat	3.73	3.44	1.9%	0.6
Harlingen	3.70	3.42	1.9%	1
Franekeradeel	3.84	3.56	2.3%	1.4
Dantumadiel	4.14	3.83	2.5%	1.5
Haren	6.35	6.05	1.6%	1.5
Wadden	4.47	4.20	2.0%	0.75
North-Netherlands	3.63	3.32	2.5%	112.2

*Table 6: A selection of municipalities, municipalities in **Bold** have an above average mortality rate and a below average amount of blood pressure patients*

The results above clearly show the relation between hypertension and CVD mortality. What must be kept in mind is that the results of the regression likely underrepresent reality. Due to the comorbidity of the factors and the fact that the amount of people with hypertension is greater than the figure used. The actual figure should be greater, also considering that the single variable regression indicated a 0.009 B value. Recalculating with 0.009 would reduce the amount of deaths by 126 or 2.75%[7], when there is a decrease of hypertension by 15%.

The model also gives valuable information about what is not yet known. There are nine municipalities[8] where the CVD mortality rate is above average, and the amount of people with high blood pressure is below average. However, it is not possible with the data at hand to provide a definitive explanation. We can conclude that it is due to factors other than or in combination with hypertension[9]. The most peculiar fact about these nine municipalities is that seven of the nine are located in Friesland. In total Friesland has nine municipalities that have an above average mortality rate. This means that a hypothetical decrease in the amount of people suffering from hypertension would have a smaller impact on the amount of CVD deaths in Friesland than in for example Groningen. This result implies that to reduce CVD deaths in Friesland as effectively as in Groningen a different approach is needed. For example, Tytsjerksteradiel would only see a decrease of 1.9%. Despite having an above average mortality rate. Pekela is a municipality in Groningen with a comparable CVD mortality rate but with a relativity greater amount of hypertension patients, here there would be a decrease of 2.6%. This indicates that a different approach would be more effective in reducing the mortality rate.

---

[7] Map of the effects of a 15% decrease in hypertension patients is located in the Appendix
[8] One of the municipalities is that of the Wadden islands, in reality there are four municipalities but for this analysis they have been combined into one.
[9] Map showing the nine municipalities is located in the Appendix

6. Conclusion

Based on the results of previous research and on theories accompanying this topic. It is more than likely that the regional variation in CVD is accompanied by similar variation in the determining factors. The study attempted to discover if the regional variation in CVD was accompanied by a similar variation in the various determining factors. As mentioned in the discussion the factors are related to one another and influence one another. As a result, the analysis only gives minimal information about the actual reality. Therefore, this research is unable to definitively conclude which factors are specifically responsible for the regional variation in the north of the Netherlands. Primarily due to the fact that the data used is inadequate to provide a model with a generally applicable result.

What can be made, however, is a conclusion about the general direction and relation between the factors and CVD mortality. For the 59 cases we can conclude that 17% of the variation can be attributed to the model created. Whereby blood pressure accounts for 12.6% of the variation. The figures that can be predicted from this result, may not provide a fully accurate picture of reality but they do indicate a general direction. The relation between CVD mortality and high blood pressure has been illustrated once more. The remaining variables also have a relation with CVD mortality, however purely based on the results of this research that conclusion cannot be made. The results however cannot be compared to the results of the INTERHEART study nor to Muller-Nordhorn (2008) as these studies were conducted with different data types and used a different methodology.

Yusuf et al (2004), Kreatsoulas and Anand (2010), Ezzati et al. (2002), Muller-Nordhorn et al (2004) (2008) and He and MacGregor (2002) all conclude that high blood pressure in a population has a significant influence on the CVD mortality rate. As mentioned in this research a reduction of hypertension by 15% in the population would result in a 2.5% decrease in CVD deaths. A hypothetical 15% reduction in hypertension patients can be achieved in multiple ways such as; promoting more physical activity, discouraging smoking and stimulating changes in diets (Kreatsoulas & Anand, 2010). Currently the Dutch government is attempting to achieve this through encouraging reduction in salt intake (RIVM, 2016). With the objective of reduction in CVD related deaths.

This research may have not used the most suitable data to achieve a representative outcome. The methodology paired with a more complex dataset could achieve what this research initially set out to determine. Using the database from the LIFELINES[10] research project would provide a far more detailed analysis, that could be able to definitively prove which factors are responsible for the regional variation in CVD. This database would allow for detailed information that can be analyzed at neighborhood level instead of municipal level, along with unique cases representing individuals. Using the database from the LIFELINES research project would better help understand more in detail the context of the north of the Netherlands. Future research should make use, indeed, of detailed data sets that can also show a change over time. As of finishing this research the 2016 figures from the *Gezondheidsmonitor* where published. These figures can be used to conduct a similar research whereby the results can be compared.

[10] A long-term study with using the collected information of nearly 100000 people living in the North of the Netherlands.

References

CBS (2012). *Doodsoorzakenstatistiek. Heerlen:* Centraal Bureau voor de Statistiek

CBS (2012). *Gezondheidsmonitor Volwassenen 2012*. Heerlen: Centraal Bureau voor de Statistiek, Gemeentelijke Gezondheidsdienst & Rijks Instituut voor Volksgezondheid en Milieu

CBS (2012). *Personen met verstrekte geneesmiddelen; regio (gemeente).* Heerlen: Centraal Bureau voor de Statistiek

Curtis, S. & Jones, I.R. (1998). Is there a place for geography in the analysis of health inequality? *Sociology of Health & Illness.* 20(5), 645-672

Ezzati, M., Lopez, A.D., Rodgers, A., Vander-Hoorn, S., Murray, C.J.L (2002) Selected major risk factors and global and regional burden of disease. *The Lancet,* 360(9343), 1347-1360

He, F.J., & MacGregor G.A., (2002) Effect of modest salt reduction on blood pressure: a meta-analysis of randomized trials. Implications for public health, *Journal of Human Hypertension*, 16, 761-770

Kreatsoulas, C., & Anand, S. S. (2010). The impact of social determinants on cardiovascular disease. *The Canadian Journal of Cardiology*, *26*, 8C–13C.

Krieger, N., Chen, J.T., Waterman, P.D., Rehkopf, D.H., Subramanian S.V. (2003). Race/Ethnicity, Gender, and Monitoring Socio-economic Gradients in Health: A Comparison of Area-Based Socio-economic Measures; The Public Health Disparities Geocoding Project. *Am J Public Health,* 93(10), 1655-1671

Mackenbach, J.P., Kunst, A.E., Looman, C.W.N., Habbema, J.D.F., Maas van der, P.J. (1988) Regional differences in mortality from conditions amenable to medical intervention in The Netherlands: a comparison of four time periods. *Journal of epidemiology and Community Health,* 42(4), 325-332

Marmot, M. (2005) Social Determinants of Health Inequalities. *The Lancet,* 365(9464), 1099-1104

Messerli F.H., Bangalore S., Julius S. (2008) Risk/Benefit Assessment of β-Blockers and Diuretics Precludes Their Use for First-Line Therapy in Hypertension. *Controversies in Cardiovascular Medicine.* 117(20), 2706-2715

Morris, R.W., Whincup, P.H., Emberson, J.R., Lampe, F.C., Walker, M., Shaper, A.G. (2003) North-South Gradients in Britain for Stroke and CHD, Are They Explained by the Same Factors? *Stroke*, 34(11), 2604-2611

Muller-Nordhorn, J., Rossnagel, K., Mey, W., Willich, S.N. (2004) Regional variation and time trends in mortality from ischemic heart disease: East and West Germany 10 years after reunification. *J Epidemiol Community Health,* 58, 481-485

Muller-Nordhorn, J., Binting, S., Roll, S., Willich, S.N. (2008) An update on regional variation in cardiovascular mortality within Europe. *European Heart Journal,* 29(10), 1316-1326

O'Brian, R.M., (2007) A Caution Regarding Rules of Thumb for Variance Inflation Factors. *Quality & Quantity.* 41, 673-690

RIVM (2016). *Minder zout eten kan tienduizenden hartinfarcten en beroertes voorkomen. Bilthoven*: Rijksinstituut voor Volksgezondheid en Mileu

U.S. Cancer Statistics Working Group (2016). *United States Cancer Statistics: 1999–2013 Incidence and Mortality Web-based Report.* Atlanta: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention and National Cancer Institute

Williams, G.H. (2003) The determinates of Health: Structure, Context and Agency. *Sociology of Health and Illness.* 25(3), 131-154

Wilkinson, R. & Marmot, M. (2003) *Social Determinants of Health: The Solid Facts.* 2nd edition. Denmark: World Health Organization

Yusuf, S., Hawken, S., Ôunpuu, S., Dans, T., Avezum, A., Lanas, F., McQueen, M., Budaj A., Pais, P., Varigos, J., Lisheng, L. (2004) Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study): case-control study. *The Lancet,* 364(9348), 937-952

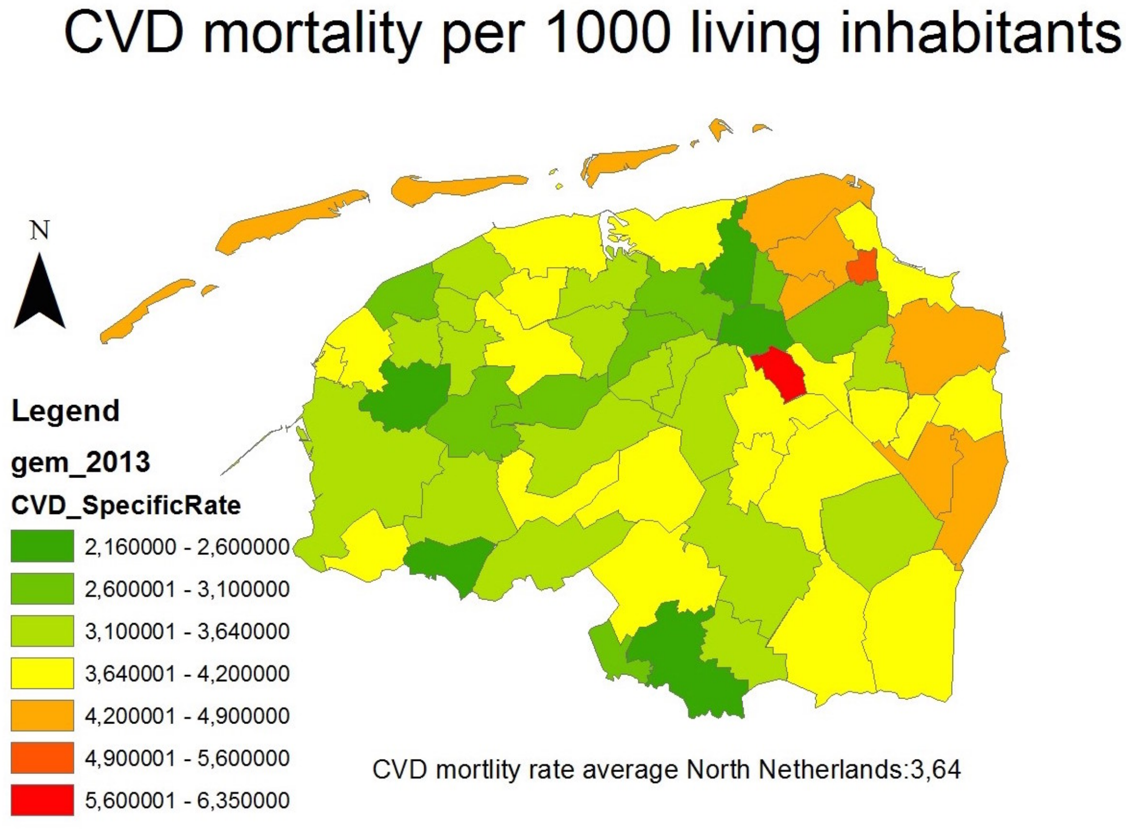
Appendix

Variable	R square	Adjusted R square	Significance
Inactivity	0.063	0.047	0.055
Smoking	0.000	-0.017	0.936
Overweight	0.039	0.022	0.135
Cholesterol	0.088	0.072	***0.022***
Blood Pressure	0.141	0.126	***0.003***
Diabetes	0.099	0.084	***0.015***

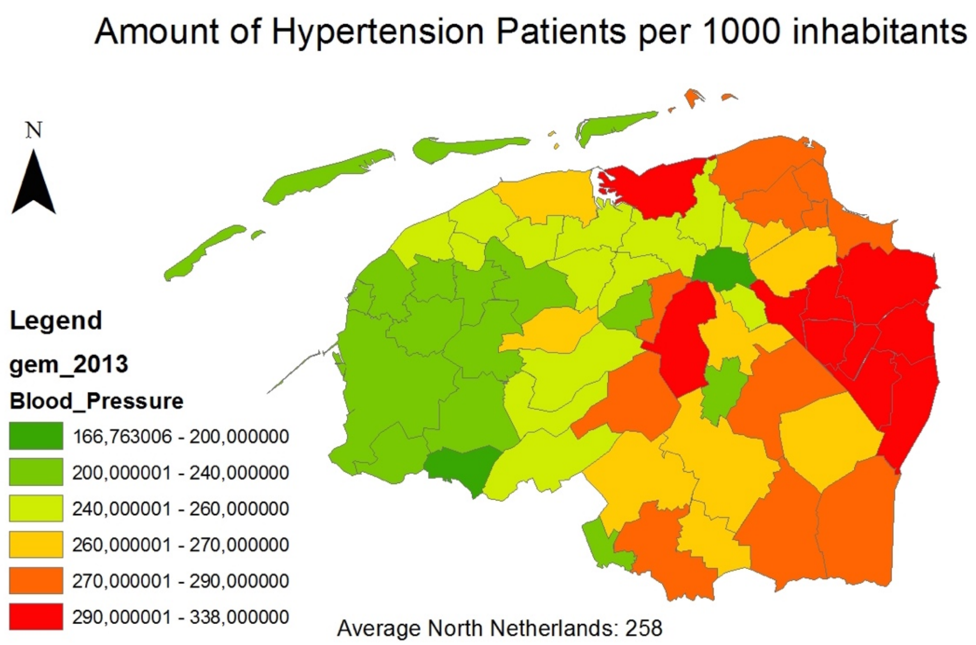
Appendix 1: Single variable linear regression results table (Number in bold indicate significant results)

	CVD	Smoking	Overweight	Cholesterol	Blood Pressure	Inactivity	Diabetes
CVD	1	.011	.197	***.297***	***.376***	***-.251***	***.315***
Smoking	.011	1	***.280***	***.357***	.175	-.169	***.375***
Overweight	.197	***.280***	1	***.725***	***.646***	.047	***.725***
Cholesterol	***.297***	***.254***	***.633***	1	***.787***	-.156	***.813***
Blood Pressure	***.367***	.175	***.646***	***.787***	1	-.031	***.793***
Inactivity	***-.251***	-.169	.047	-.156	-.031	1	-.102
Diabetes	***.315***	***.375***	***.725***	***.813***	***.793***	-.102	1

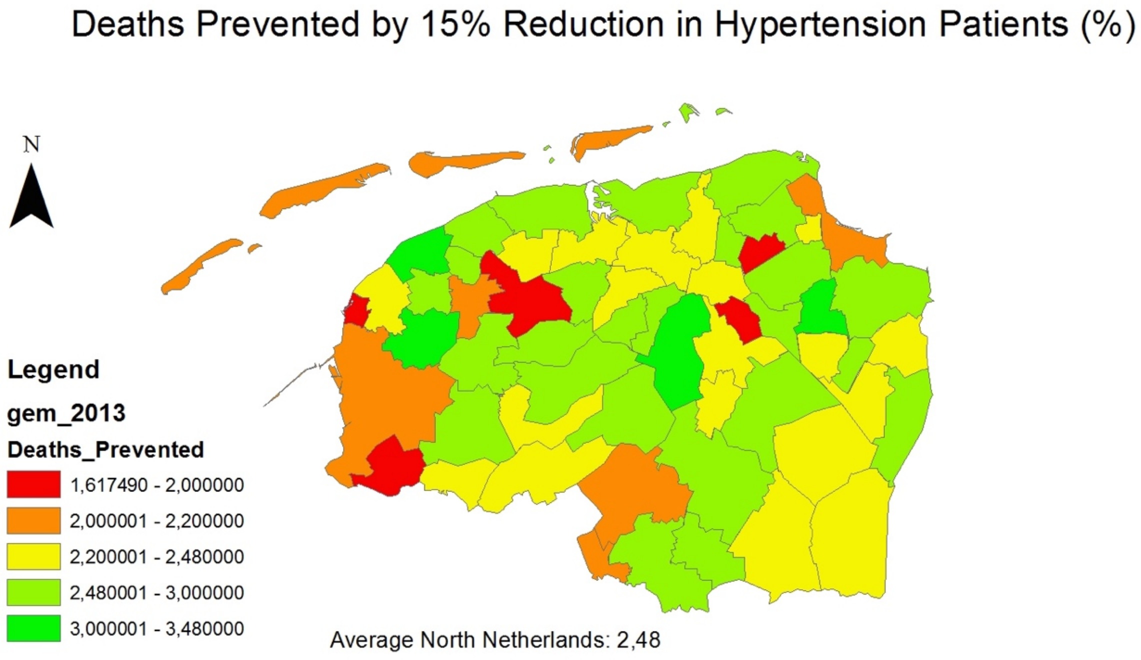
Appendix 2: Correlation Matrix (Number in bold indicate significant correlations)



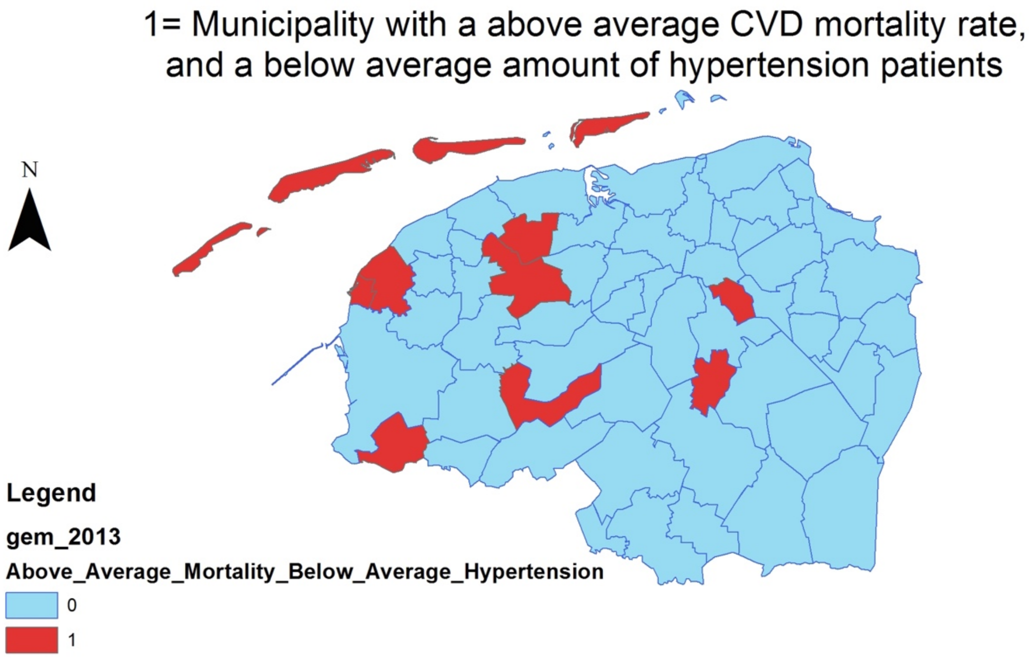
Map 1: Showing the CVD mortality per 1000 living inhabitants (Made by author)



Map 2: Showing the amount of hypertension patients per 1000 inhabitants (Made by author)



Map 3: Showing the effect of a 15% reduction in the amount of people suffering from hypertension. The figures represent the percentage of deaths reduced as a result of the 15% reduction in hypertension patients. (Made by author)



Map 4: Shows the municipalities with a CVD mortality rate that is above average, and with a below average amount of hypertension patients.(Made by author)

Data Table

All numbers that are accompanied by *per 1000* indicate the amount of people affected by the variable per 1000 inhabitants.

Municipality	CVD mortality 2012	Overweight per 1000	Smoking per 1000	Cholesterol per 1000	Blood Pressure per 1000	Diabetes per 1000	Inactive per 1000	CVD deaths per 1000
Aa en Hunze	78	490	190	160	273	67	360	3,89
Achtkarspelen	69	510	270	143	258	73	370	3,29
Appingedam	50	570	270	169	276	90	330	5,26
Assen	190	470	230	134	235	62	370	3,76
Bedum	22	450	200	156	260	58	230	2,77
Bellingwedde	30	550	290	170	304	89	270	4,02
het Bildt	25	550	260	137	256	61	420	3,10
Ten Boer	25	510	130	127	263	64	290	4,57
Borger-Odoorn	66	510	220	160	262	74	400	3,30
Coevorden	109	480	200	157	281	78	400	4,04
Dantumadiel	60	520	140	154	256	73	440	4,14
Delfzijl	103	610	230	177	281	101	350	4,97
Dongeradeel	71	510	190	163	270	62	490	3,89
Eemsmond	53	540	300	180	277	79	280	4,29
Emmen	320	570	260	152	274	74	410	3,79
Ferwerderadiel	22	420	180	124	247	54	390	3,34
Franekeradeel	60	430	250	139	233	62	440	3,84
Groningen (gemeente)	337	370	260	102	167	47	260	2,16
Grootegast	26	480	170	170	248	70	310	2,91
Haren	90	410	160	146	250	53	240	6,35
Harlingen	45	540	280	159	235	63	240	3,70
Heerenveen	132	480	220	128	245	63	360	3,94
Hoogeveen	130	470	250	181	269	75	380	3,13
Hoogezand-Sappemeer	100	550	250	185	314	94	300	3,70
Kollumerland en Nieuwkruisland	33	530	270	164	255	61	450	3,40
Leek	48	580	230	169	274	71	290	3,27
Leeuwarden	250	450	300	124	205	57	350	3,35
Leeuwarderadeel	25	510	160	161	253	58	390	3,22
Littenseradiel	18	490	230	128	225	58	430	2,26
Loppersum	34	550	200	191	272	80	280	4,31
De Marne	32	470	330	178	290	75	310	3,94
Marum	25	490	290	159	238	75	270	3,23
Menameradiel	35	450	240	127	230	60	400	3,40
Menterwolde	31	610	210	174	308	83	450	3,24
Meppel	75	480	260	135	229	64	390	3,05
Midden-Drenthe	82	510	170	151	264	68	380	3,19
Noordenveld	83	450	250	167	302	72	360	3,47

Oldambt	148	630	290	185	337	87	340	4,78
Ooststellingwerf	79	500	200	147	283	71	440	3,97
Opsterland	78	510	230	124	247	60	360	3,48
Pekela	40	650	320	197	319	103	400	3,99
Slochteren	36	590	200	152	266	74	370	3,07
Smallingerland	126	460	220	142	263	67	410	3,00
Stadskanaal	111	630	290	182	308	101	320	4,32
Súdwest-Fryslân	206	460	240	126	220	61	320	3,31
Tynaarlo	95	460	230	156	267	62	350	3,84
Tytsjerksteradiel	98	460	220	142	229	61	370	4,04
Veendam	84	490	240	178	294	86	310	3,87
Vlagtwedde	56	580	290	182	338	89	300	4,37
Westerveld	61	450	240	134	263	61	380	4,03
Weststellingwerf	70	490	200	141	253	67	370	3,50
Winsum	27	500	190	140	246	63	370	2,60
De Wolden	52	500	220	150	284	68	390	2,87
Zuidhorn	40	500	250	142	244	59	310	2,96
Wadden	36	530	267	127	223	56	340	4,47
Boarnsterhim	42	480	250	137	213	57	450	2,90
Gaasterlân-Sleat	29	450	210	154	235	66	340	3,73
Lemsterland	24	500	230	146	198	62	380	2,35
Skarsterlân	73	470	200	131	226	53	330	3,55
North of the Netherlands	4595	506	233	153	260	70	356	3,63
Municipality	CVD mortality 2012	Overweight per 1000	Smoking per 1000	Cholesterol per 1000	Blood Pressure per 1000	Diabetes per 1000	Inactive per 1000	CVD deaths per 1000