UNIVERSITY OF TWENTE

SPATIAL ANALYSIS OF OUT-OF-HOSPITAL CARDIAC ARREST INCIDENCES IN THE NETHERLANDS

IN COOPERATION WITH ARREST GROUP, ACADEMIC MEDICAL CENTER (AMC), AMSTERDAM, THE NETHERLANDS

Author:
G. Balistreri

Supervisors:
Dr. D. Demirtas
Dr. C.J.M. Doggen

A thesis submitted in fulfilment of the requirements for the degree of Master of Science in the field of Industrial Engineering and Management

February 12, 2017

MSc graduation assignment Industrial Engineering & Management
University of Twente

In cooperation with the Academic Medical Center (AMC) in Amsterdam / AmsteRdam REsuscitation STudies (ARREST)

"SPATIAL ANALYSIS OF OUT-OF-HOSPITAL CARDIAC ARREST INCIDENCES IN THE NETHERLANDS"

February 12, 2017

Author:
Gino Balistreri

Supervisory committee at the University of Twente:
Department Industrial Engineering and Business Information Systems (IEBIS):
Dr. Derya Demirtas
d.demirtas@utwente.nl

Department Health Technology and Services Research (HTSR):
Dr. Carine J.M. Doggen
c.j.m.doggen@utwente.nl

Project Initiator:
Dr. Derya Demirtas

VERSION: PUBLIC

UNIVERSITY OF TWENTE.

amc

In God we trust, all others bring data.

— William Edwards Deming (1900-1993)

MANAGEMENT SUMMARY

Background & motivation

Out-of-hospital cardiac arrest (OHCA) is a leading cause of death in the world. Incidence and outcome vary greatly between countries. Prior studies report geographic variation in incidence rates between countries, cities, as well as between neighbourhoods within the same city. Areas with a relatively high incidence of OHCA cases may be appropriate targets for community-based resuscitation training and awareness efforts. Since survival rates are dramatically lower for OHCAs that have a response time longer than several minutes, OHCA prevention and resuscitation can be improved effectively by investigating cardiac arrest incidence, underlying causes of high incidence as well as current resuscitation efforts at a detailed geographic level. Hence, it is fundamental to have a better understanding of both spatial and temporal OHCA variability. In terms of spatial variability, we focus on underlying differences in socio-economic status (SES) and ethnicity. Prior studies suggest that independently of an individual's SES, low community SES is associated with one's overall health condition, leading to poor cardiovascular and metabolic health. However, no prior studies examine whether districts with high SES in the Netherlands also have lower OHCA incidence. In addition, although some report that discrepancies in OHCA survival are inherent to race, numerous studies explain this phenomenon as a confounding factor of SES rather than genetic nature. Yet, no study has investigated the relation between OHCA incidence and the ethnic composition of geographic areas.

Objective

The intent of this study is to 1) identify high- and low-risk districts for OHCA, 2) determine whether these districts remain high- or low-risk over time, 3) find underlying demographic, socio-economic and ethnicity associations that may explain spatial disparities in OHCA incidence rates, and 4) explore the current state of resuscitation efforts in these districts to identify the most beneficial opportunities for short-term improvement, by focusing on communities at greatest risk. The emphasis is on determining the underlying associations between OHCA incidence and demographic, socio-economic and ethnicity factors at a district level, through the use of machine learning (ML) algorithms, which are proven to be effective in cardiovascular risk prediction.

Study setting

The study area comprises of the province of North Holland between 2006 and 2016 and the region of Twente between 2010 and 2016. Data on all consecutive suspected OHCAs with a cardiac cause is used in collaboration with the registry of the AmsteRdam REsuscitation STudies (ARREST). Publicly available, regional data is retrieved from the statistical database of Statistics Netherlands. All districts are classified as having either high or low OHCA incidence. The set of attributes is manually reduced to 16 attributes in North Holland and 13 attributes in Twente, including measures of basic demographics, housing, income, social protection and ethnicity. We find optimal attribute subsets automatically by constructing Genetic Algorithms (GAs).

Methods

We apply a generalized linear mixed model to assess overall temporal variation within all considered districts. Next to that, we apply four types of ML methods that show promising predictive performance in similar studies in the field of cardiovascular research: Logistic

Regression (LR), Support Vector Machine (SVM), Artificial Neural Networks (ANNs), and Adaptive Boosting (AdaBoost). Finally, bystander-initiated CPR (BCPR), ambulance defibrillator usage, AED usage, and the number of deployed public access AEDs per capita are compared with OHCA incidence at a district level by means of LR analyses.

Results
Spatial disparities in standard and age-adjusted OHCA incidence are observed in the Dutch regions of North Holland and Twente. A vast majority of areas with high incidence given the entire day also have high incidence during night-time. Although we did not find significantly higher overall rates for any of the included years, we did observe significant unexplained temporal variability within districts. Thus, areas that are considered to have high OHCA incidence rates throughout the study period do not necessarily report similar rates from year to year. Underlying causes of spatial disparities are considered complex and multifactorial. LR structurally provides higher predictive accuracy than ANNs, SVMs and AdaBoost in predicting areas with high and low incidence, based on demographic, socio-economic and ethnicity factors. The proportion of income receivers that belong to the national 40% with the lowest income (OR 1.16; 95% CI 1.06, 1.27), the proportion of at least 65 years of age, and the proportion of male inhabitants are significantly associated with increased odds of a district in North Holland having high OHCA incidence. Average household size is significantly associated with similar, decreased odds. In addition, the proportion of receivers of low income (OR 1.12; 95% CI 1.02, 1.22) and average household size are significant determinants of age-adjusted incidence rates. The average household size and the proportion of married inhabitants are significantly associated with the odds of a district in North Holland having high OHCA incidence during night-time. In Twente, we find the proportion of elderly, male, and married inhabitants to be significantly associated with the odds of a district having high incidence. After adjusting for age structure, we observe the proportion of recipients of labour disability acts (OR 1.42; 95% CI 1.08, 1.87) and the average property value (OR 1.17; 95% CI 1.03, 1.33) to be significantly associated with OHCA incidence. In both regions, ethnicity was not found to be significantly associated with OHCA incidence.

Disparities between districts are observed for all considered resuscitation efforts. In North Holland, we find a significantly higher proportion of applied AEDs (OR 1.02; 95% CI 1.00, 1.04) across rural districts that belong to the 50% with high OHCA incidence rates in the region, compared to the remaining 50% with low rates. In Twente, we find significantly higher values for BCPR, AED usage, and the number of AEDs per 10,000 population (OR 1.15; 95% CI 1.02, 1.30) across urban districts with relatively high incidence rates, compared to those with low rates.

Conclusion
Public health campaigns should target districts with a high proportion of receivers of low income, recipients of labour disability acts, elderly population, male inhabitants, a low proportion of married inhabitants as well as those with small household sizes. Public health campaigns that aim to improve resuscitation in areas with known high OHCA incidence are advised to prioritize on those resuscitation efforts for which unfavoured disparities are found: the number of AEDs per capita, AED usage, and, to a smaller extent, BCPR rates. These campaigns should take into account the presence of temporal variation within districts.

PREFACE

Writing this master thesis has been a very exciting and informative project, containing both ups and downs. The opportunity to combine and apply methods from statistics, data science, spatial analytics, and machine learning in the field of cardiac arrest, is not only unique, it is very informative, relevant and challenging as well.

I am greatly indebted to both of my supervisors, dr. Derya Demirtas and dr. Carine Doggen. I am especially grateful to dr. Derya Demirtas for her supervision, her insightful feedback, for making time throughout the thesis to hold many informative meetings, and for giving me the opportunity to work on this amazing project. I thank dr. Carine Doggen for her valuable contribution and the interesting discussions on the main topics of my thesis. I am very thankful that she wanted to get involved into the project herself, as she brings a valuable, epidemiologic perspective to the topics of this thesis. Finally, I would like to thank dr. Erwin Hans for making time to step in at the final stage, and for his valuable feedback during the remainder of the project.

We express our gratitude to all researchers and staff members of the ARREST study, whose research has proven to be invaluable in the field of Out-of-Hospital Cardiac Arrest. We are greatly indebted to the registry for sharing cardiac arrest data. We especially thank dr. Rudolph Koster, dr. Marieke Blom, dr. Steffie Beesems, and Remy Stieglis of the ARREST registry for their cooperation, support and contribution to this research and the collection of cardiac arrest data. We are also greatly indebted to HartslagNu for sharing valuable AED location data. Finally, we would like to thank all the EMS personnel, fire fighters, police officers, nurses, physicians, civilian responders, employees of HartslagNu, as well as the staff at the EMS dispatch centres that participate in the ARREST study, who work to save lives of cardiac arrest victims each and every day.

I would also like to thank Arthur Nazarian for the excellent collaboration during the initial stage of the project. Finally, I would like to thank my family, friends, and my girlfriend Estefanía, for their unconditional support throughout the entire project. Their support has been invaluable.

Gino Balistreri

Enschede, January 2018

TABLE OF CONTENTS

GLOSSARY OF ABBREVIATIONS

AdaBoost	Adaptive Boosting
AED	Automated External Defibrillator
ANN	Artificial Neural Network
ARREST	Amsterdam Resuscitation Studies
CBS	Statistics Netherlands (in Dutch: 'Centraal Bureau voor de Statistiek')
COROP	Coordination Commission Regional Research Programme (in Dutch: 'Coördinatie Commissie Regionaal Onderzoeks-Programma')
CPR	Cardiopulmonary resuscitation
DT	Decision Tree
EMS	Emergency Medical Services
ESRI	Environmental Systems Research Institute
GA	Genetic Algorithm
GIS	Geographic Information System
GLMM	Generalized Linear Mixed Model
ICC	Intraclass Correlation Coefficient
LR	Logistic Regression
ML	Machine learning
OAD	Surrounding address density (in Dutch: 'Omgevingsadressendichtheid')
OR	Odds ratio
OHCA	Out-of-Hospital Cardiac Arrest
SES	Socio-economic status
StatLine	Statistical Database of Statistics Netherlands
SVM	Support Vector Machine
RBF	Radial basis function
WW	Unemployment Benefits Act (in Dutch: 'Werkloosheidswet')
WWB	Work and Social Assistance Act (in Dutch: 'Wet Werk en Bijstand')

CHAPTER 1 – INTRODUCTION

This chapter contains a description of the project. Section 1.1 provides the problem description. Sections 1.2-1.4 provide the problem statement, the purpose and the contributions of this study, followed by a formulation of the research questions in Section 1.5.

1.1 Problem description

Out-of-hospital cardiac arrest (OHCA) is a leading cause of death in developed and many developing countries (Ma et al., 2007). However, OHCA incidence[1] and outcome varies greatly around the globe (Berdowski et al., 2010; Gräsner et al., 2016). The number of patients who suffer from an OHCA annually in Europe and the United States have been reported to be approximately 350,000-700,000 and 360,000 respectively (Mozaffarian et al., 2016; Perkins, Handley, et al., 2015). We first state the importance of targeted resuscitation efforts to enhance survival. We then discuss the spatial and temporal variability of OHCA. Subsequently, we discuss the literature gap in terms of socio-economic and ethnicity factors. We then briefly describe the potential of machine learning to examine underlying associations between these factors and OHCA incidence. Finally, we briefly mention the relevant literature gap in terms of OHCA incidence and resuscitation efforts.

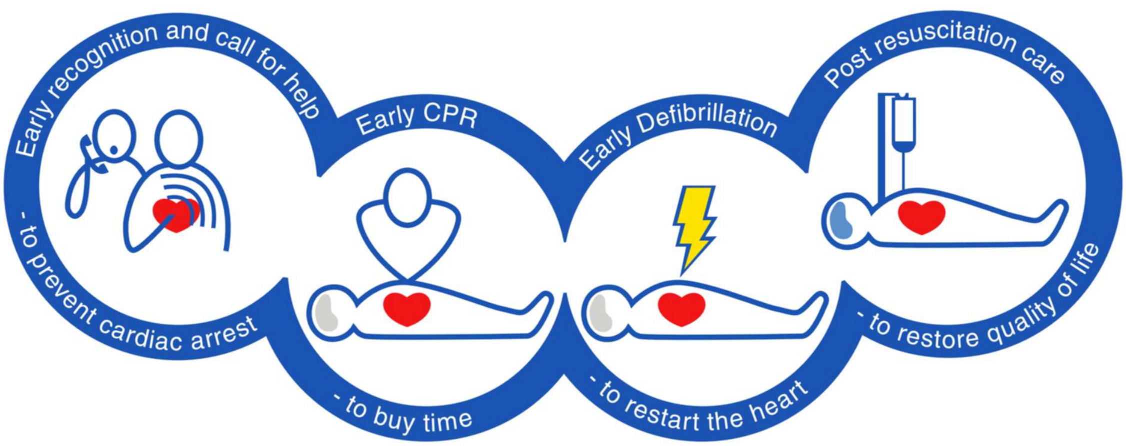


Figure 1.1. *Chain of survival in case of an out-of-hospital cardiac arrest (Perkins, Handley, et al., 2015).*

1.1.1 OHCA survival

Short response times are critical to survive from OHCA (Eisenberg et al., 1979; Roth et al., 1984; Larsen et al., 1993). Survival rates drop dramatically during the first minutes after the cardiac arrest event when no resuscitation is started (Larsen et al., 1993). According to Beesems et al. (2012), the survival rate of OHCA victims in the Dutch regions of North Holland and Twente was 23% in 2011. Although this rate has gradually increased with respect to previous years, the majority of victims still do not survive from cardiac arrest. The Netherlands Heart Foundation (in Dutch: 'De Nederlandse Hartstichting') therefore aims to improve resuscitation by transforming the Netherlands into one integrated 'six-minute' zone, in which every cardiac arrest victim can be offered the right aid within six minutes. This is achieved by calling Emergency Medical Services

[1] In epidemiology, incidence is defined as the number of new cases per unit of person-time at risk (Berdowski et al., 2010) and is in most cases reported as a rate (i.e. 'incidence rate') per 100,000 inhabitants.

(EMS), start resuscitation and defibrillate the victim by means of an automated external defibrillator (AED) that is deployed close to the victim (Zijlstra et al., 2014). Targeted AED deployment and CPR training programmes in areas with relatively high incidence rates are required for these zones to be successful. In addition, these programs are advised to consider the presence of temporal variability of OHCA incidence rates, as areas may not remain high-risk over time. Nevertheless, targeted public health efforts are more complex in areas for which OHCA data is not available. To this end, socio-demographic and socio-economic characteristics that are associated with significantly higher OHCA incidence rates at a detailed geographic level, can be used as an alternative guideline.

Figure 1.2. *Automated external defibrillator.*

1.1.2 Spatial variability of OHCA incidence rates

According to Berdowski et al. (2010), to improve OHCA prevention and resuscitation, it is fundamental to have a better understanding of the spatial variability of OHCA incidence rates. The AmsteRdam REsuscitation STudies (ARREST) registry collects EMS-assessed OHCA patient data since July 2005 in the province of North Holland, as well as in the region of Twente since 2010. The incidence of EMS-treated OHCAs with a cardiac cause per 100,000 population in 2014 was reported 36 in North Holland and 41 in Twente (Zijlstra et al., 2016). In the same year, EMS-treated OHCA incidence rates were 39 per 100,000 inhabitants in the Dutch province Limburg and 31 in Utrecht. Similar geographic variation in OHCA incidence rates is observed in the United States by Nichol et al. (2008) and in Taiwan by Kao et al. (2017). Studies show that OHCA incidence rates even vary significantly between cities (e.g. Berdowski et al., 2010; Moon et al., 2015) as well as between neighbourhoods within the same city (e.g. Reinier et al., 2011; Sasson et al., 2012; Root et al., 2013; Fosbøl et al., 2014; Demirtas et al., 2015). However, no prior studies have been performed in the Netherlands that examine OHCA incidence rates at a more detailed regional level rather than the entire study region. Therefore, it is unclear which areas in the Netherlands tend to be high-risk or low-risk based on OHCA incidence only.

1.1.3 Socio-economic factors

Research to the relation between socio-economic status (SES) and health status of individuals has been going on for more than 150 years (Macintyre, 1997). In addition, recent studies suggest that independently of an individual's SES, low community SES is associated with one's overall health condition (Diez-Roux et al., 1997; Schootman et al., 2007), leading to poor cardiovascular and metabolic health (Mirowsky et al., 2017). While numerous studies examine the relation between SES and risk factors of cardiovascular disease (CVD) such as type 2 diabetes mellitus (Krishnan et al., 2010; Saydah et al., 2013), only few studies investigate the relation between SES and OHCA incidence (e.g. Reinier et al., 2006) at a detailed geographic level. Moreover, a majority of these studies include a single measure of SES, rather than evaluating the influence of multiple indicators.

Inhabitants of Dutch districts with high SES are considered to have better living conditions than inhabitants of districts with a relatively low status (Bijl et al., 2011). However, it is unknown whether these areas also have significantly lower OHCA incidence rates. The most related study is performed at a province level by Hendrix et al. (2010), who associate low SES with significantly high incidence of sudden cardiac death (SCD) in patients ≤ 40 years of age. The authors define SCD as a sudden unexpected death with a cardiac cause within 24 hours after the onset of symptoms. To the extent of our knowledge, no prior research has been done in the Netherlands on the relation between socio-economic factors and OHCA incidence at a more detailed level. We emphasize that this analysis is performed at a district level, rather than at an individual level. Moreover, we examine districts where EMS-attended OHCA cases are reported, as opposed to the district where the victim resides. Our study may thus include victims of OHCA that are not present in the same district as their residence at the time of the cardiac arrest. In line with the findings of Reinier et al. (2006), we assume that a vast majority of cardiac arrest occur within the same district as the patient's residence.

1.1.4 Ethnicity

Studies that examine the relation between OHCA incidence rates and ethnicity is scarce and instead mostly focus on race. A majority of studies take place in the US and find significantly higher OHCA incidence rates among African Americans compared to whites or Hispanics. Researchers even claim significantly lower survival rates (Becker et al., 1993; Cowie et al., 1993), as well as bystander CPR rates for African American victims of OHCA (Brookoff et al., 1994; Moon et al., 2014; Wilde et al., 2012). Although some report that differences in mortality are inherent to race, other studies explain this phenomenon as a confounding factor of SES rather than genetic nature (Chu et al., 1998; Hallstrom et al., 1993; Kucharska-Newton et al., 2011). Yet, no study has investigated the relation between OHCA incidence and inhabitant's ethnic backgrounds. The most related study is performed in Victoria, Australia by Straney et al. (2016). The authors assess several socio-demographic factors but did not find any association between OHCA incidence rates and the proportion of inhabitants that is either born overseas and/or speaks a language other than English at home. In our study, we follow definitions from Statistics Netherlands, classifying inhabitants of having either a Western or non-Western background. We further subdivide the latter into the four main groups with non-Western backgrounds in the Netherlands: Antilleans and Arubans, Moroccan, Surinamese and Turkish inhabitants. Similar to the analysis between OHCA incidence and socio-economic factors, we emphasize that the analysis between OHCA incidence and ethnicity concerns the ethnic composition of districts, as opposed to the ethnicity background of individuals. Furthermore, we examine districts in which EMS-attended OHCA cases are reported, which may not necessarily be the same district as the patient's residence.

1.1.5 Predictive modelling

A vast majority of studies that examine the relation between SES and cardiovascular disease incidence or survival apply logistic regression analyses (Chu et al., 1998; Mirowsky et al., 2017; Soo et al., 2001). As in many other fields in health care, binary logistic regression is the statistical learning technique of choice to predict binary outcomes (e.g. low vs. high risk). However, in an era of advanced big data analytics, the field of machine learning offers new methods of predicting risk rather than regression-based methods alone (Banerjee, 2017; Goldstein et al., 2016). To the extent of our knowledge, we are the first to study the relation between OHCA incidence and socio-economic attributes by applying multiple machine learning algorithms. In addition, the 'No Free Lunch Theorem' in the field of machine learning tells us that it is unknown in advance whether a certain method will perform well in practice. In order to best examine this relation, we therefore apply several methods that have performed well in practice in the field of cardiovascular research.

1.1.6 Resuscitation efforts

Districts with a high incidence of OHCA events may be appropriate targets for community-based resuscitation training and awareness efforts (Sasson, Magid, et al., 2012). In addition, a higher proportion of AEDs should intuitively be deployed in so-called 'high-risk' areas. However, no studies in the Netherlands examine the relation between OHCA incidence and resuscitation efforts such as bystander CPR, AED usage and AED deployment at a detailed geographic level.

1.2 Problem statement

We state our research problem as follows:

"It is unknown whether (1) incidence rates of EMS-attended OHCAs vary significantly over space and time simultaneously, whether (2) there is a relation between incidence rates of EMS-attended OHCA and both socio-economic factors and ethnicity and whether (3) there is a relation between incidence rates of EMS-attended OHCA and resuscitation efforts at a detailed geographic level in the regions of Twente and North Holland in the Netherlands. A better understanding of all three topics can improve current public health efforts, such as targeted AED deployment and CPR training programmes."

1.3 Purpose

Based on the problem statement, the purpose of this research is thus threefold:

"The purpose of this research is to study the spatio-temporal distribution of EMS-attended OHCA, the relation between incidence of EMS-attended OHCA and socio-economic factors as well as ethnicity, and the relation between incidence of EMS-attended OHCA and resuscitation efforts, at a detailed geographic level in both public and private settings. We study these three relations using data for the regions of Twente and North Holland in the Netherlands, in order to improve current public health efforts."

1.4 Contributions

This is the first study in the world on the relation between OHCA incidence and the ethnic composition of geographic areas. In addition, this is the first study to examine the relation between OHCA incidence and demographic as well as socio-economic factors through the application of multiple machine learning algorithms.

Moreover, this is the first study in the Netherlands on (1) spatial and temporal variability of OHCA incidence, (2) the relation between OHCA incidence and socio-economic factors, and (3) the relation between OHCA incidence and resuscitation efforts (bystander CPR, defibrillator usage, and AED deployment) at a detailed geographic level.

1.5 Research questions

Main research question

“What is the relation between incidence of EMS-attended OHCA and both socio-economic factors and ethnicity backgrounds, as well as the relation between incidence of EMS-attended OHCA and resuscitation efforts at a detailed geographic level in both public and private settings in the Dutch regions of Twente and North Holland?

Sub questions

1. **What is the spatio-temporal variability of EMS-attended OHCA at a detailed geographic level?**
2. **What is the relation between incidence of EMS-attended OHCA and both socio-economic factors and ethnicity backgrounds, of which data is publicly available, at a detailed geographic level?**
   - 2.1 For which socio-economic factors is the relation between OHCA incidence rates and the corresponding factor(s) discussed in the literature?
   - 2.2 For which ethnicity backgrounds is the relation between OHCA incidence rates and the corresponding factor(s) discussed in the literature?
   - 2.3 For which of the identified socio-economic factors is data publicly available?
   - 2.4 Which machine learning methods are most suitable to examine this relation?
3. **What is the relation between incidence of EMS-attended OHCA and resuscitation efforts at a detailed geographic level?**
   - 3.1 What is the relation between OHCA incidence rates and rates of bystander CPR?
   - 3.2 What is the relation between OHCA incidence rates and defibrillator usage?
   - 3.3 What is the relation between OHCA counts and the number of deployed AEDs?

Chapter 2 provides the context analysis. Chapter 3 contains a review of the most relevant literature. Chapters 4 and 5 discuss the methods and results respectively. Finally, Chapter 6 states the discussion and conclusion of our study.

CHAPTER 2 – CONTEXT ANALYSIS

A cardiac arrest is defined as the cessation of cardiac mechanical activity as confirmed by the absence of signs of circulation (Perkins, Jacobs, et al., 2015). If the cardiac arrest did not occur inside a hospital, it is defined as an out-of-hospital cardiac arrest (OHCA). This chapter provides theoretical context concerning OHCA related to this study. Section 2.1 describes the emergency response system in case of an OHCA. Section 2.2 assesses the importance of early defibrillation, emphasizing the necessity of targeted resuscitation efforts. We then briefly describe the ARREST registry in Section 2.3, of which data on EMS-treated OHCAs is used in this study. Finally, we discuss the characteristics of the included OHCA cases in Section 2.4.

2.1 Cardiac arrest response system

In this section, we describe the current emergency response system in the Netherlands in case of a cardiac arrest.

2.1.1 Emergency call

In case of a medical emergency, people dial the Dutch national emergency services number '112'. These calls are immediately redirected to the regional EMS dispatch centre (Blom et al., 2013). EMS dispatchers are mostly nurses or employees in other areas of healthcare who completed additional dispatch training and follow a standard protocol during the phone call. First, the dispatcher establishes the exact location of the event and the phone number of the caller. Second, the caller is asked to describe the event and the patient's current condition. Based on the caller's description, the dispatcher might presume an OHCA (e.g. agonal breathing is a sign of cardiac arrest) (Berdowski et al., 2009). Since the EMS dispatcher did not witness the cardiac arrest, it is unlikely that at this stage, the occurrence of an OHCA can be confirmed (Perkins, Jacobs, et al., 2015). The dispatch system automatically generates and stores the times of the emergency call and the EMS dispatch to the established location (Berdowski et al., 2009).

2.1.2 EMS dispatch

In case the EMS dispatcher suspects an OHCA, two ambulances are sent to the given location from a single tier (Berdowski et al., 2009). The dispatched EMS teams are qualified to perform advanced cardiopulmonary life support and equipped with a manual defibrillator (Berdowski et al., 2011). If the dispatcher does not suspect an OHCA, only one ambulance is sent. However, if EMS personnel on site find the patient in cardiac arrest, a second ambulance is sent (Berdowski et al., 2009). The reason to send a second ambulance, is that performing cardiopulmonary resuscitation (CPR) of a patient in cardiac arrest can be very intensive. To avoid the same EMS personnel performing chest compressions consecutively for a considerable amount of time, additional EMS personnel is sent to assist in the resuscitation of the patient.

2.1.3 First responders

Immediately after sending two ambulances, the dispatcher sends local first responders (firefighters, police officers and/or general practitioners) equipped with an Automated External Defibrillator (AED) to the given location (Bardai et al., 2011; Berdowski et al., 2011).

2.1.4 Lay responders

Next to the EMS teams and first responders, the EMS dispatcher also sends local civilian responders (Blom et al., 2014). To this end, organized emergency response systems for registered lay responders are used. Initially at this stage, EMS dispatchers in the Twente region activated the 'AED-Alert' system. Introduced in 2008, this system sent a text message to automatically selected lay responders that either live or work within approximately 1000 metres of the OHCA event (Scholten et al. , 2011) or within 500 metres of the AEDs located in this area (Zijlstra et al., 2014). The system was also implemented in North Holland in 2013 (Zijlstra et al., 2014). AED-Alert has been replaced in 2015 by a similar app notification alert system, shown in Figure 2.1. By means of the notification, the selected lay responders are requested to perform early CPR and defibrillation. The notification contains both the address of the OHCA event and the location of the nearest AED (if applicable).

Studies show that the system effectively shortens time to early defibrillation and significantly improves the survival to hospital discharge (Pijls et al., 2016). Two app notification alert systems are currently being used throughout the Netherlands: 'HartslagNu' and 'Hartveilig Wonen'. HartslagNu is implemented by the EMS service in Twente ('Ambulance Oost') and by one of the three EMS services in North Holland ('Noord Holland Noord'). The EMS services in Kennemerland implemented the system of Hartveilig Wonen, whereas the EMS services in Amsterdam currently do not use any alert system to dispatch lay responders.

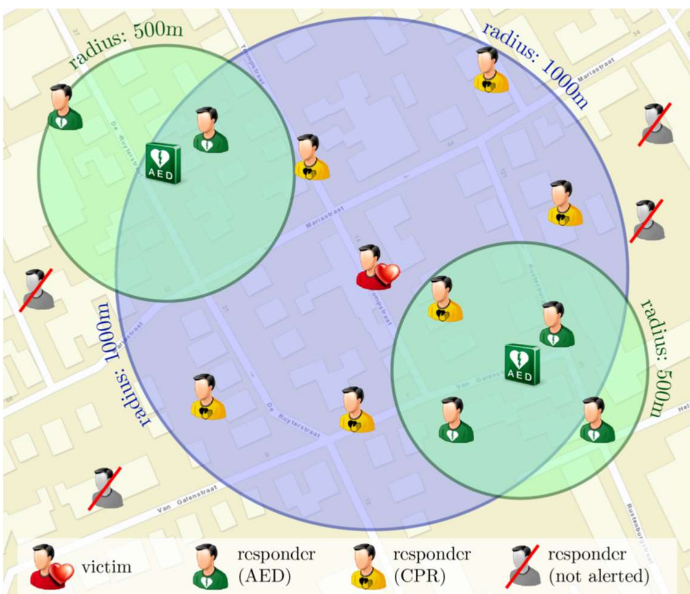


Figure 2.1. *Organized emergency response system for registered lay responders in the Netherlands.*

2.2 Importance of early defibrillation

The probability to survive an OHCA has been estimated to decrease by 10% for each minute that CPR is not initiated from the moment of collapse onwards (Cummins et al., 1991). Hence, defibrillation of cardiac arrest victims within 3 to 5 minutes after occurrence correspond to survival rates up to 50-70% (Perkins, Jacobs, et al., 2015). Such early defibrillation can only be achieved by organizing and dispatching lay responders to cardiac arrest victims within seconds. To this end, AEDs should be placed in areas with a high density of citizens, such as residential areas, airports, railway stations, bus terminals, sport facilities, shopping malls and offices. Unfortunately, the AED deployment in public areas by both public and private parties is not centrally controlled or directed (Berdowski et al., 2011). Historical data of the number of cardiac arrests in a particular area, together with regional characteristics, may help guide AED deployment (Perkins, Jacobs, et al., 2015). AED deployment is generally performed together with resuscitation training for local lay responders, in accordance with the guidelines of the European Resuscitation Council (Berdowski et al., 2011; Perkins, Jacobs, et al., 2015). The training includes Basic Life Support as well as AED usage (Zijlstra et al., 2014). Conducting training programmes in communities is crucial to shorten defibrillation time by increasing bystander CPR rates and the number of registered lay responders.

2.3 The ARREST registry

Since 2004, the number and scope of national and regional registries of OHCA cases and resuscitation attempts increased considerably (Perkins, Jacobs, et al., 2015). Major registries were established in North America (McNally et al., 2009; Morrison et al., 2008), Europe (Gräsner et al., 2011; Savastano et al., 2015), Asia (Kitamura et al., 2010; Ong, Shin, et al., 2011) and Oceania (Beck et al., 2016). This includes the registry of the AmsteRdam REsuscitation STudies (ARREST), the largest OHCA registry in the Netherlands in terms of study area, total population served and the number of EMS-assessed OHCAs (Zijlstra et al., 2016). The Academic Medical Center (AMC) in Amsterdam organizes the registry. ARREST is an ongoing prospective, population-based study focused on establishing the genetic, clinical, pharmacological and environmental determinants of OHCA in the general population as well as determinants of OHCA outcome (Berdowski et al., 2009; Bezzina et al., 2010; Blom et al., 2014). It contains records of all consecutive suspected cardiac arrests since July 2005. This includes bystander witnessed OHCAs, OHCAs without any witness of collapse as well as cases in which CPR attempts were made.

The study region consists of all COROP regions in the province of North Holland, except for 'Het Gooi en Vechtstreek'. The total population was approximately 2,536,000 in 2016 ("Statistics Netherlands StatLine database," n.d.). Three EMS services primarily serve the region and participate in the ARREST study (Blom et al., 2014). The population in the region of Twente was approximately 626,000 in 2016. One EMS primarily serves the region (Beesems et al., 2012). Urban and rural communities are present in both regions. Moreover, both regions use a tiered response system: in case of a suspected OHCA, the EMS service often dispatches two ambulances to one emergency call (Berdowski et al., 2009), as well as local first responders equipped with an AED (firefighters, police officers and/or general practitioners) (Bardai et al., 2011; Berdowski et al., 2011) and local civilian responders (Blom et al., 2014; Zijlstra et al., 2016). Ambulances are equipped with a manual defibrillator and qualified to perform Advanced Life Support (Blom et al., 2014). Furthermore, the AED deployment in public areas by both public and private parties is not centrally controlled or directed (Berdowski et al., 2011).

ARREST extensively collects data on all EMS-treated OHCAs within the study area to minimize selection bias. This is done in accordance with the Utstein guidelines (Perkins, Jacobs, et al., 2015). All OHCAs are considered to result from a cardiac arrest, unless an unambiguous non-cardiac cause (i.e. trauma, drowning or intoxication) was established (Blom et al., 2014). After each resuscitation attempt, information is gathered on among others the patient, location, presence of lay responders and witnesses of the collapse. Information on defibrillator usage is documented automatically, such as the connection time, the patient's initial rhythm and the time of the first shock. To read more about ARREST, please refer to Blom et al. (2014) and Bezzina et al. (2010).

2.4 Cardiac arrest data of ARREST

We received data of all EMS-treated OHCAs with a cardiac cause for which resuscitation was attempted. We consider 10,667 OHCAs in North Holland between January 1, 2006 and December 31, 2016. In Twente, we consider 1,788 cases between February 1, 2010 and December 31, 2016. Appendix A (Confidential) shows an overview of the characteristics of all considered cases. The average patient's age in both regions was around 66 years, with a standard deviation of 15 years. By far the most patients were male. Moreover, around two third of all events occurred during daytime. Almost the same fraction received bystander CPR, and in more than two out of five cases, an AED was used during the resuscitation attempt. Although results between both study regions are similar, we observe large differences compared to other studies in terms of bystander CPR and the applied AED usage. Demirtas et al. (2015) report characteristics of 24,605 EMS-treated, public OHCA patients in Toronto, Canada between 2007 and 2014. Less than one out of four (22.8%) of the patients received bystander CPR. In only 1.4% of the patients, an AED was applied before EMS services arrived. A similar trend is observed by the Resuscitation Outcomes Consortium (ROC) Epistry throughout multiple US sites. Although the reported overall bystander CPR rate in 2014 was 46.1%, bystanders only applied AEDs in 2.0% of the cases (Mozaffarian et al., 2016). As opposed to these studies however, our data includes the provision of CPR and the usage of AEDs by first responders. Hence, differences in CPR provision and AED usage by lay responders between these studies are likely to be smaller.

Figures 2.2 and 2.3 show the annual number of EMS-treated OHCAs versus the annual total population in North Holland and Twente respectively. The number of OHCAs in North Holland seems to have increased roughly in line with the annual study population, whereas the population and the number of OHCAs in Twente approximately remained similar.

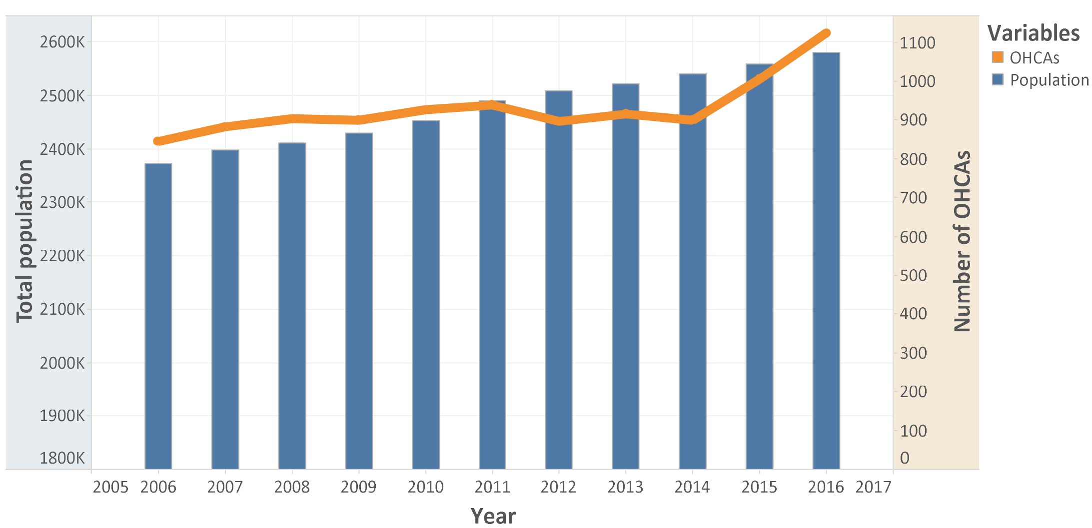


Figure 2.2. *Number of EMS-attended OHCAs (10,667 in total) versus population in North Holland (2006-2016) by year.*

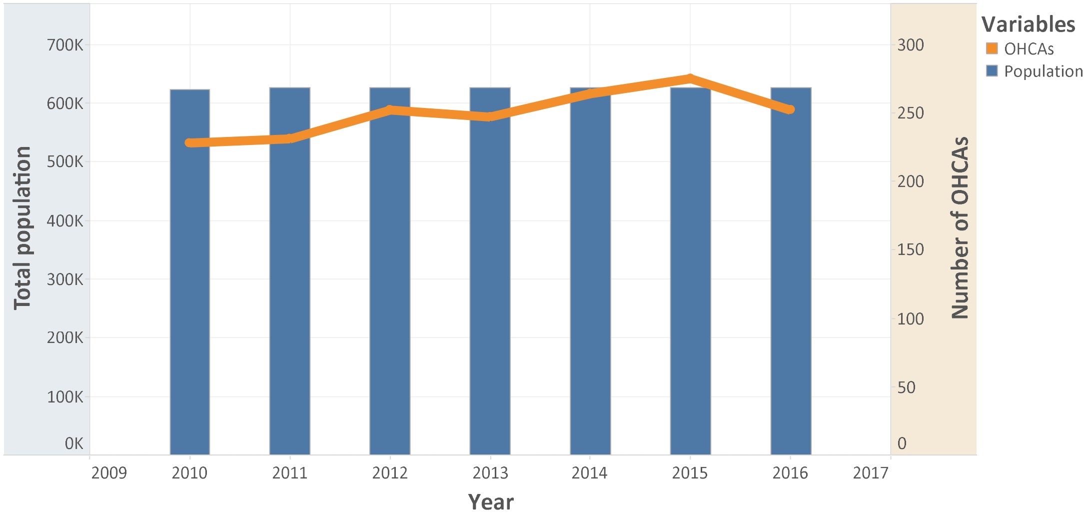


Figure 2.3. *Number of EMS-attended OHCAs (1,788 in total) versus the population in Twente (2010-2016) by year.*

CHAPTER 3 – LITERATURE REVIEW

This chapter provides a review of the literature relevant to this study. Section 3.1 states prior studies on spatial disparities of OHCAs. Section 3.2 first discusses prior studies concerning the spatio-temporal variation of OHCAs, followed by the identification of suitable alternative measures. Section 3.3 states the relation between health status and socio-economic status (SES), as well as those SES measures that are included in prior studies to OHCA incidence at a geographic level. Section 3.4 briefly describes prior studies that examine disparities in health status among migrant groups in the Netherlands with different ethnic backgrounds, the role of ethnicity as a confounding factor of SES, and those studies that examine the relation between OHCA incidence and ethnicity. Section 3.5 first gives a description of a predictive modelling approach through machine learning (ML) algorithms and challenges of this approach related to this study. The remainder of this section discusses similar studies that apply ML algorithms and the corresponding predictive performance of these algorithms. Finally, Section 3.6 discusses prior studies on the relation between OHCA incidence and resuscitation efforts at a geographic level.

3.1 Geographic distributions of OHCAs

Over the last years, the number of studies that examine geographic distributions of cardiac arrests increased considerably. Soo et al. (2001) analyse OHCA incidence rates in electoral areas in Nottinghamshire, United Kingdom. Lerner et al. (2005) create OHCA clusters based on a Kernel Density Estimation to identify cardiac arrest hot spots in Rochester, New York. Raun et al. (2013) identify census tracts with high incidence rates and low prevalence of bystander CPR in Houston, Texas. Straney et al. (2015) identify similar census tracts in Victoria, Australia. Moreover, Demirtas et al. (2015) determine OHCA incidence rates of EMS-treated OHCAs that occurred in public locations in the city of Toronto, Canada. The authors identify several high-risk neighbourhoods for OHCA. Sasson, Cudnik et al. (2012) identify census tracts in Columbus, Ohio with high OHCA incidence rates as well as low bystander CPR rates by combining three different spatial clustering methods. Similarly, Nassel et al. (2014) consider a census tract in Denver, Colorado to have an increased risk for OHCA when two out of three methods identify the census tract as high-risk. Finally, Kao et al. (2017) apply several spatial clustering methods to identify high-risk neighbourhoods in New Taipei City, Taiwan.

Sasson, Cudnik et al. (2012) attribute the increase in geographic studies to the availability of more sophisticated spatial analysis tools, due to rapid technology advancement of Geographic Information System (GIS) software. The studies mentioned above typically identify areas with significantly higher OHCA incidence rates successfully. These areas can serve as priority locations for resuscitation training programs and AED deployment to eventually enhance OHCA survival rates. Applied methods range from a comparison of OHCA counts between census tracts to complex cluster techniques. The main advantage of applying spatial clustering techniques is the ability to find patterns independent of any predefined geographic structure. The main difference with other studies mentioned above, is the fact that we will use regional data on demographic, socio-economic and ethnicity factors to further examine incidence rates, rather than identifying hot spots alone. Using a geographic structure thus allows us to assess spatial variation of OHCAs, and use regional data in subsequent steps.

3.2 Spatio-temporal variation of OHCAs

Few studies have been done on both spatial and temporal variation of OHCA events. In Fulton County, Georgia, Sasson et al. (2010) are the first to investigate this variation by using an Intraclass Correlation Coefficient (ICC). Unfortunately, the specific type of ICC is not specified. Although the value of the coefficient is rather low (0.36 [95% CI, 0.26 to 0.50]), the authors conclude that OHCA incidence rates are stable within census tracts over the study period of 38 months. Based on these findings, further research in Columbus, Ohio by Sasson et al. (2012), Warden et al. (2012) and Semple et al. (2013) focuses on spatial variation of OHCA events only, assuming that clusters for OHCA incidence are stable over time. Demirtas et al. (2015) determine incidence rates and spatio-temporal variation of EMS-treated, public OHCAs that occurred in public locations in the city of Toronto, Canada. The authors use a study period of eight years, which is significantly larger than the approximate three-year study period used by Sasson et al. (2010). Demirtas et al. (2015) use the ICC(2,1) (Fisher, 1921; Shrout & Fleiss, 1979) to measure the relative annual variability of the number of OHCAs per neighbourhood. The authors conclude that the distribution of OHCAs in Toronto is stable over space and time simultaneously. Based on these studies, only forms of the ICC seem applicable to assess spatio-temporal variation. However, parametric ICCs assume multivariate Normal distributions and equal variances. These important assumptions cannot be violated (Costa-Santos et al., 2010; Müller & Büttner, 1994). As OHCAs can be considered as so-called 'rare' events, the assumption of normality may be valid only if in each year, OHCAs in the considered areas occur frequently. As opposed to Sasson et al. (2010) and Demirtas et al. (2015), we include both urban and rural areas, rather than urban areas alone. Hence, these assumptions most likely do not hold, as the distribution is expected to be highly skewed and to contain values of zero reported cardiac arrests. Next to that, ICC forms are in fact reliability coefficients, while agreement methods are suitable to detect changes over time.

When modelling different years as random raters (i.e., observers) that independently measure the number of occurrences in a set of subjects, inter-rater agreement study results are used to estimate the intrinsic measurement error. Measurements may be taken under different conditions, by different raters or by the same rater at different moments in time (De Vet, 1998). Agreement and reliability are different concepts, although used interchangeably in literature (De Vet et al., 2006; Kottner et al., 2011). Guyatt et al. (1987) emphasize that an analysis aiming to discriminate subjects (e.g., persons) requires a high level of reliability, whereas an analysis aiming to evaluate raters or changes over time requires a high level of agreement. In the first case, subjects can only be discriminated successfully when the measurement error is relatively small in comparison with the variability between subjects. In the second case, the variability between subjects is irrelevant: the likelihood of observing between-rater dissimilarities depends solely on the magnitude of the measurement error (Guyatt et al., 1987; De Vet et al., 2006).

Medical studies often concern the ability to detect changes over time. Hence, agreement coefficients are more favourable to evaluate possible changes. Yet, a majority of studies investigate the level of reliability rather than the level of agreement, usually resulting in drawing incorrect conclusions of agreement results (Berchtold, 2016; Costa-Santos et al., 2005; De Vet et al., 2006; Gow et al., 2008). By exclusively presenting study results using a reliability coefficient, one can conclude whether subjects can be discriminated within a certain sample. However, the capability of monitoring changes over time remains unclear (De Vet et al., 2006). However, methods to assess inter-rater agreement with more than two raters and non-Gaussian data are scarce. Popular agreement methods such as Lin's Concordance Correlation Coefficient (CCC) (Lin, 1989), the standard error of measurement (Schmidt & Hunter, 1989; Stratford & Goldsmith, 1997), Bland and Altman's limits of agreement, as well as extensions of the latter to multiple raters (Bland & Altman, 1983, 1986, 1999; Jones et al., 2011), cannot be used. We identify two potential alternatives: a generalized linear mixed model (Gu & Ma, 2005) as well as a non-parametric Information-Based Measure of Disagreement (Henriques et al., 2013).

3.3 Socio-economic factors

3.3.1 Socio-economic status in health studies

SES is defined as a construct of one's combined social and economic position within a society. Generally, SES is accepted to be a composite measure of education level, income and occupation. Other indicators such as financial wealth or neighbourhood disadvantage are occasionally included as well (Baker, 2014). Financial wealth refers to one's acquired assets, such as home and car ownership. Although SES is considered complex and multifactorial, Braveman et al. (2005) state that a majority of health studies measure SES by means of a single socio-economic indicator. Included measures depend on the study population and available indicators (Galobardes et al., 2006; Herd et al., 2007). Regardless of the measures used, studies consistently show that individual-level SES is related to mortality risk (Lantz et al., 2010; Mackenbach et al., 2008; Stringhini et al., 2017; Turrell et al., 2007) in all stages of life (Kunst, 2010). Throughout Europe, life expectancy of men associated with low SES is estimated to be four to six years lower compared to men with high SES. Similar striking differences are reported among women (about two to four years) (Kunst, 2010).

Several studies suggest that independently of an individual's SES, community SES is associated with one's overall health condition (Diez-Roux et al., 1997; Schootman et al., 2007). It is beyond the scope of this thesis to perform extensive explanatory research on the relation between low community SES and low health status. To explain why this relation may exist in the first place, we briefly mention the most important risk factors. Numerous studies aim to explain this phenomenon through neighbourhood differences in diet (Diez Roux et al., 1999; Morland et al., 2002), lack of physical activity (Moore et al., 2008; Yen & Kaplan, 1998), physical health (Gallo & Matthews, 2003; Lorant et al., 2003), lack of access to healthcare facilities (Stringhini et al., 2010), social stress levels caused by noise, financial problems or crime (Anderson et al., 1997; DeCarlo Santiago et al., 2011; Sundquist et al., 2006; Van Oort et al., 2005), housing characteristics (Haan, Kaplan, & Camacho, 1987) or environmental factors (Rosenthal et al., 2008). These factors between SES and health status can roughly be categorized as material, behavioural, and psychosocial factors, as shown in Figure 3.1. Moreover, these categories are interrelated. For example, communities associated with high material disadvantage may be highly associated with psychosocial stress and lack of physical exercise (Mackenbach, 2006). In addition, the highly educated population may in general be better in coping with psychosocial stress, in part due to high occupation and income level (Mackenbach et al., 1994). This may especially be valid in open communities, in which social selection continuously determines one's social position (Kunst, 2010). Finally, the complexity of explaining the relation between health outcomes and community-level SES is illustrated by Haan et al. (1987). The authors find an increased mortality risk in the low poverty area in Oakland, California, even after adjusting for a vast amount of risk factors (baseline health status, race, income, status of employment, access to healthcare facilities, health insurance coverage, smoking, alcohol consumption, physical activity, body mass index, sleep patterns, social isolation, marital status, depression and personal uncertainty).

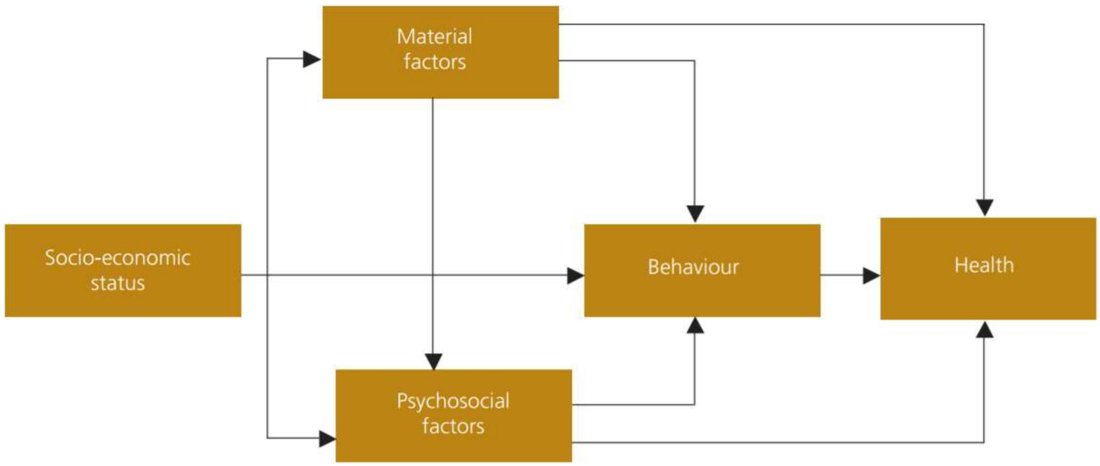


Figure 3.1. *Factors that are considered to mediate between SES and health status (from: Mackenbach, 2006).*

3.3.2 Relation between socio-economic factors and OHCA incidence

Approximately 80% of OHCA patients have signs of underlying coronary artery disease (CAD) (Chugh et al., 2008). CAD risk factors such as smoking, obesity, diabetes mellitus, hypertension, hyperlipidaemia and lack of physical activity are associated with low individual-level SES (Jaglal & Goel, 1994; Mensah et al., 2005; Rozanski et al., 2005; Saydah et al., 2013). However, studies that examine the relation between OHCA incidence and community-level SES are limited. A majority of studies that examine this relation, include OHCAs over a relatively short study period of no more than two years. In Portland, Oregon, Feero et al. (1995) find that age- and gender-adjusted incidence rates are only associated with median income levels, as opposed to population density and African-American race. The analysis is limited to seven separate regions. In a nation-wide study in Sweden, Strömsöe et al. (2011) do not find any significant association between age- and gender-adjusted OHCA incidence rates and population density among 21 large regions.

Reinier et al. (2006) examine OHCA incidence rates in Multnomah County, Oregon at a more detailed, census tract level (170 in total), although the number of OHCAs is limited to 697 cases. The authors independently assign census tracts to quartiles of each included SES indicator: (1) median property value, (2) median household income, (3) the proportion of inhabitants with income below the poverty level, and (4) the proportion with at least a Bachelor's degree. Incidence rates are 30-80% higher in the census tracts in the lowest SES quartile compared to the highest. The authors conclude that all four indicators are independently associated with higher OHCA incidence. In an extensive follow-up study in seven regions in North America, Reinier et al. (2011) find significant associations between standard and age-adjusted incidence rates and median household income at a census tract level. The latter is included as a single measure of SES.

Only two studies examine OHCA incidence rates over more than two years. Over a four-year study period in Nottinghamshire, Soo et al. (2001) conclude that socially deprived areas, measured based on 1) the degree of unemployment, 2) car ownership, 3) housing tenure, and 4) household overcrowding, are indeed associated with higher OHCA incidence. Finally, Straney et al. (2016) investigate areas with high OHCA incidence in Victoria, Australia over a three-year study period, and include both demographic and cardiovascular risk factors. A predefined index of socio-economic disadvantage is included as a single measure of SES. The proportion of the population over 65 years of age, SES, smoking prevalence and education level are significant predictors of incidence, collectively explaining most of the observed spatial variation (93.9%) of incidence rates. Both studies do not adjust incidence rates for age or gender. Moreover, these studies do not identify which included SES measures are most associated to high OHCA incidence.

3.3.3 Relation between socio-economic factors and cardiovascular disease

Due to the limited number of studies on OHCA incidence and SES, we slightly extend our literature review to include the most important factors found in the field of cardiovascular disease (CVD). A recent study by Mirowsky et al. (2017) in North Carolina includes twelve demographic factors and investigates whether spatial disparities exist in terms of metabolic and cardiovascular health. After controlling for age, gender, race, and smoking status, differences in the prevalence of obesity, diabetes, congestive heart failure and hypertension are found in a cluster that is highly urban, and has significantly larger proportions of African Americans, unemployed inhabitants, inhabitants below the poverty level, and households on public assistance. The latter three factors seem to be highly correlated, although intercorrelations are not discussed.

Category	**Factor**	**Studies that include corresponding factor**	**Studies that report a significant association**
Basic demographics	Population of at least 65 years	Straney et al. (2016*)	100% (1/1)
	Urbanisation	Mirowsky et al. (2017*)	100% (1/1)
	Population density	Feero et al. (1995) Mirowsky et al. (2017) Straney et al. (2016); Strömsöe et al. (2011)	0% (0/4)
Income	Median income	Feero et al. (1995*) Reinier et al. (2006*, 2011*)	100% (3/3)
	Population with income below the poverty level	Mirowsky et al. (2017*) Reinier et al. (2006*)	100% (2/2)
Wealth	Median property value	Reinier et al. (2006*)	100% (1/1)
Social protection	Households receiving social assistance	Mirowsky et al. (2017*)	100% (1/1)
Education	Population with at least a Bachelor's degree	Mirowsky et al. (2017) Reinier et al. (2006*) Straney et al. (2016*)	67% (2/3)
Occupation	Unemployment	Mirowsky et al. (2017*)	100% (1/1)
	Non-managerial positions	Mirowsky et al. (2017)	0% (0/1)
Housing	Single parent housing	Mirowsky et al. (2017)	0% (0/1)
	Private properties	Mirowsky et al. (2017)	0% (0/1)
Other	Socio-economic disadvantage index	Straney et al. (2016*)	100% (1/1)
	Townsend index	Soo et al. (2001*)	100% (1/1)

**Study finds significant association with OHCA or CVD incidence.*

Table 3.1. *Socio-economic factors included in prior studies.*

In summary, the number of studies that examine the association of community-level SES and OHCA incidence is scarce. We therefore expanded our review with one major study in the field of CVD. As a result, there seems to be a lack of consensus in literature on the most important factors to represent community-level SES. We show the presence of this phenomenon in studies that address the association between SES and OHCA incidence. Table 3.1 gives an overview of the factors included in prior studies. Results between these few studies seem to be largely consistent. Education level is the only factor for which inconsistency is observed.

3.4 Ethnicity

Migrant groups across Europe are associated with higher rates of several communicable and non-communicable diseases, poor mental health, poor living conditions, and poor occupational health compared to the native European population (Agudelo-Suárez et al., 2011; Ahonen et al., 2007; Gushulak et al., 2010; Rechel et al., 2011, 2013; Schenker, 2008, 2010).

3.4.1 Ethnicity and cardiovascular disease

In the Netherlands, considerable differences are reported among CVD and CVD risk factors in Dutch natives and inhabitants with either a Turkish or Moroccan background (Kunst et al., 2011). Figure 3.2 shows self-reported prevalence of CVDs (stroke, myocardial infarction) and CVD risk factors (diabetes mellitus, being overweight). Studies on CVD mortality rates among Turkish and Moroccan inhabitants are however inconclusive (Bos et al., 2005; Ujcic-Voortman et al., 2012). Bos et al. (2005) do not find large inequalities in CVD mortality rates among Turkish, Moroccan, Antillean/Aruban, and Surinamese between neighbourhoods with high and low SES. By contrast, the authors do observe inequalities in mortality rates among Dutch inhabitants. Furthermore, in line with Kunst et al. (2011), Ujcic-Voortman et al. (2012) show that obesity levels are higher among those with a Turkish or Moroccan background compared to Dutch natives.

Discrepancies in cardiovascular health between natives and migrants may be due to inequalities among health care use, that may be caused by fees, language barriers, unfamiliarity with the health care system, or by obstacles in the administrative process (Kraler & Reichel, 2010; Padilla & Miguel, 2009; Rafnsson & Bhopal, 2008). After controlling for SES, these discrepancies in health tend to dissolve (WHO, 2010). Several studies therefore consider low SES to be an outcome of migrant status and ethnic background, as a result of processes of social exclusion (Davies et al., 2009; Ingleby, 2009). This shows the importance of ethnicity as a social determinant of health.

3.4.2 Ethnicity as a confounding factor of SES

Inhabitants with a non-Western background mostly live in Dutch districts with low SES. In the four largest Dutch cities (including Amsterdam), almost half of the population in low SES districts is considered to have a non-Western background (Knol, 2012). Native Dutch live more often in high income neighbourhoods than inhabitants with a background from any of the four main non-Western ethnic minorities. In addition, inhabitants with a background from Suriname or the Netherlands Antilles/Aruba more often live in high income neighbourhoods than those with a background from Turkey or Morocco (Bos et al., 2005). In contrast, immigrants with a Western background mostly live in high SES districts within those cities, and are roughly equally divided between low and high SES districts in the rest of the country. An increase in the number of Western immigrants is associated with an increase in SES, possibly due to an increase in the number of expats with Western backgrounds and high individual SES (Knol, 2012).

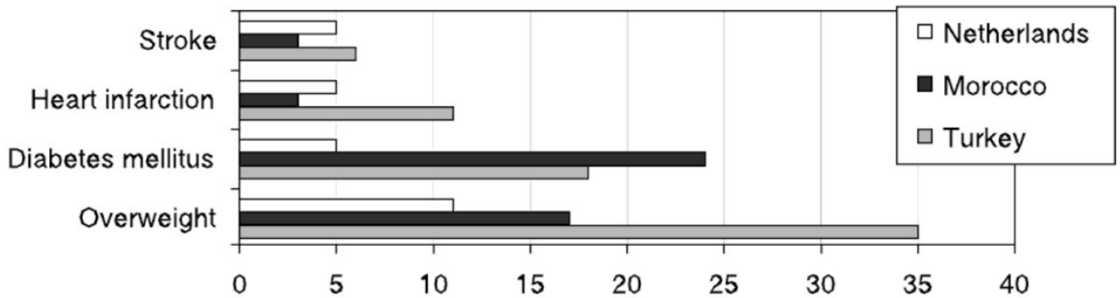


Figure 3.2. *Self-reported prevalence (%) of several non-communicable diseases in Amsterdam among Dutch natives and inhabitants born in Turkey and Morocco, 1999-2000, men and women combined (From: Kunst et al. (2011), edited).*

3.4.3 Relation between ethnicity and OHCA incidence

Studies that examine the relation between OHCA incidence and different inhabitants' backgrounds mainly focus on race. These studies distinguish whites, African Americans, and less frequently, Hispanics. The same is observed in survival (e.g. Becker et al., 1993; Groeneveld et al., 2003; Wilde et al., 2012) and resuscitation studies (e.g. Ghobrial et al., 2016; Moon et al., 2014).

According to Bhopal (2004), race and ethnicity are increasingly being used as synonyms in scientific research. The authors define race as "*the group a person belongs to as a result of a mix of physical features such as skin colour and hair texture, which reflect ancestry and geographic origins, as identified by others or as self-identified.*" Consequently, ethnicity is defined as "*the social group a person belongs to, and either identifies with or is identified with by others, as a result of a mix of cultural and other factors including language, diet, religion, ancestry, and physical features traditionally associated with race*" (Bhopal, 2004). As opposed to race, ethnicity includes social factors, making it more useful in studies concerning factors influencing OHCA incidence due to a lack of evidence that differences in incidence between racial groups are of genetic nature. If ethnicity is used, one needs to consider that wide ethnic categorisation may contain considerable unnoticed within-group heterogeneity. This is likely to result in a loss of information to comprehend variations among different ethnic groups. Moreover, ethnic minorities may not identify themselves with the defined, generalized groups (Bhopal, 2004).

Studies that examine the relation between OHCA incidence and race are inconclusive. Becker et al. (1993), Cowie et al. (1993), and Galea et al. (2007) associate higher OHCA incidence rates with African Americans in Chicago, Seattle, and New York City respectively. In addition, Galea et al. (2007) find that white OHCA patients are approximately four times less likely to live in a census tract with low median household income. By contrast, Feero et al. (1995) find African-American race not to be associated with age- and gender-adjusted rates in Portland, Oregon. In a US nationwide study, Gillum (1997) find that age-adjusted incidence rates are lower among Hispanics compared to non-Hispanic whites and African Americans. As far as we know, the only study to OHCA incidence that includes ethnic rather than racial groups is done by Straney et al. (2016). The authors do not find any significant association between OHCA incidence and the proportion of inhabitants that (1) is born overseas and (2) speaks a language at home other than English. Being born in a particular country however does not necessarily imply that one has a background from that same country. The study might still include native inhabitants with non-Australian backgrounds by including the proportion that does not speak English at home. Insufficient information is provided to confirm which ethnic groups are considered.

3.5 Predictive modelling

Risk prediction is an important tool in cardiovascular research (Goldstein et al. 2016). Traditional approaches in this field use models based on regression. As in many other fields in health care, binary logistic regression is the statistical learning technique of choice to predict binary outcomes. However, in an era of advanced big data analytics, ML offers new methods of predicting risk rather than regression-based methods alone (Banerjee, 2017; Goldstein et al., 2016). ML algorithms can be applied to find patterns without being explicitly programmed how to do so. These algorithms do not rely on a predetermined model equation, but 'learn' patterns from the data directly. In most cases, ML algorithms are applied on heterogeneous data sets (i.e., data combined from different sources) (Libbrecht & Noble, 2015). ML developed from the field of pattern recognition and computational learning, also known as Artificial Intelligence (AI). AI also covers other subfields such as robotics, computer vision and speech recognition.

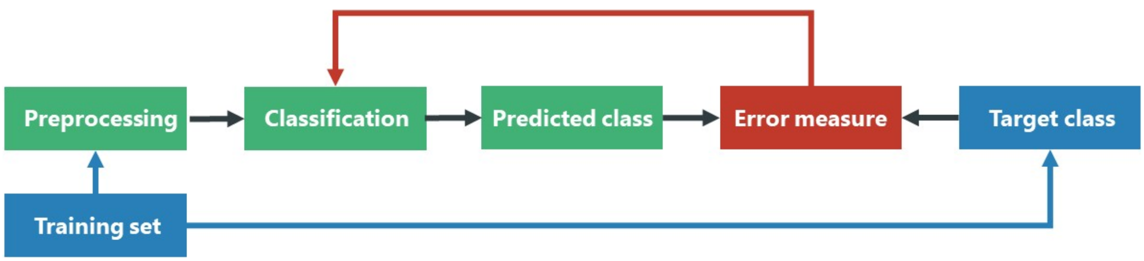


Figure 3.3. *Schematic view of supervised learning (classification).*

Two types of ML exist: *supervised* and *unsupervised* learning. In supervised learning, the predictive values are known (e.g. patient's health status after treatment). In this case, the ML method can be guided towards the right answer, by minimizing the number of misclassified instances. In unsupervised learning, the predictive class is not known, and ML algorithms are deployed to identify clusters of instances that share similar characteristics. In the field of cardiovascular risk prediction, supervised learning is most commonly applied to predict classes that are previously known. The model is trained towards the values of a given target vector $\mathbf{t}$, which corresponds to the subject values that the model needs to predict as accurate as possible. A training set of subjects is used to tune the parameters of the (adaptive) algorithm. If $\mathbf{t}$ consists of continuous measures, the required task to predict each numeric value is called *regression*. If $\mathbf{t}$ consists of classes, a class value needs to be predicted. Hence, this task is known as *classification*. The prediction is expressed as a function $y(x)$, and has the same value type as $\mathbf{t}$. The error measure is the distance between the predicted value and the target value. After calculating the error measure in one iteration for each trained instance, the algorithm aims to improve predictions of these instances in the next iteration. The exact form of $y(x)$ is thus defined during the training phase (Bishop, 2006). After successfully training a model, its predictive ability is tested on unseen data, commonly referred to as a test set. In this way, a reliable estimate of the overall error estimate can be found.

Researchers who deploy ML algorithms need to be aware of all the challenges and pitfalls. Challenges related to this study are discussed in this report. More information on the potential and challenges of ML can be found in e.g. Libbrecht & Noble (2015) and Goldstein et al. (2016).

3.5.1 Bias-variance trade-off

For all ML algorithms, a balance needs to be found between bias (i.e., accuracy) and variance (i.e., generalizability) through the use of suitable parameter settings (Goldstein et al., 2016). Often when complex models are built during the training phase on few instances, the model is not generalizable to other instances, resulting in poor performance on the test set (Libbrecht & Noble, 2015). The latter is called *overfitting*, and is caused by a considerable gap between the resubstitution error (i.e., error measure in training phase) and the generalization error (i.e., overall error measure). By contrast, underfitted models do not learn the actual pattern of the data.

3.5.2 Model selection

Over the years, a large number of ML algorithms has been developed. The Law of Conservation for Generalization Performance, commonly referred to as the No Free Lunch Theorem, states that all ML algorithms have similar overall performance when deployed across all learning tasks (Schaffer, 1994). Hence, there is no universally best ML algorithm, and each of these algorithms tackles classification tasks differently. As demonstrated by Caruana & Niculescu-Mizil (2006), even the best algorithms that yield high overall performance, result in average to poor performance on several problems. This presents researchers with the challenge of defining which algorithms are most likely to build an accurate predictive model, given the particular problem.

3.5.3 Attribute subset selection

In many situations, a high-dimensional input space of attributes serves as input for either classification or regression tasks. On the one hand, the objective of predictive modelling is to find the best performance by training machine learning algorithms on this input space. On the other hand, the objective is to come up with the smallest set of attributes to shorten computation time, improve interpretability by reducing complexity and gain insights in which (combination of) attributes contribute the most in improving predictive performance. Although this might sound as two separate objectives, this is often not the case. Based on experiments, Witten et al. (2016) argue that adding useless attributes reduces performance of many learning algorithms (i.e. decision trees, rule induction, linear regression and instance-based algorithms). The underlying problem is referred to as the fragmentation problem, as at some point during model training, only a small fragment of the original training set is left. At this point, a useless attribute might look good just by chance, leading to good performance on the training data but poor performance on test data (i.e. overfitting). Even for ML algorithms that take the possible presence of irrelevant attributes into account, an attribute preselection phase frequently improves predictive performance (Witten et al., 2016). Next to that, analyses that include high-dimensional input spaces suffer from the curse of dimensionality: the complexity of the analysis increases with the number of input dimensions, making those computationally very intensive as the number of input dimensions grows large (Keogh & Mueen, 2011). Hence, the goal of attribute selection is to select the most valuable subset of attributes given a particular setting that maximizes predictive performance by minimizing the generalization error. It is also referred to as ‘feature selection’ or ‘variable selection’. More importantly, attribute selection leads to a simplified representation by focusing on the most relevant attributes without incurring much loss of information (Witten et al., 2016). The most accurate way to estimate the relevance of an attribute is manually, based on a deep understanding of the target concept and the meaning of the particular attribute. In addition, automated attribute subset selection methods can also be useful (Witten et al., 2016).

The identification of an optimal attribute subset itself also suffers from the curse of dimensionality, since the number of possible subsets grows exponentially with the number of attributes. In a setting with k attributes, a brute-force (i.e. exhaustive) search evaluates all possible 2^k subsets. Moreover, exhaustive search may lead to overfitting if the algorithms aim to correct a misclassified instance that was in fact caused by noise (Witten et al., 2016). Alternatively, two general approaches exist: the filter approach and the wrapper approach. The filter approach is applied prior to the learning algorithm. Attributes are selected with regard to a predefined relevance measure which does not consider any generalization performance (i.e., error estimate on unseen data) (Kohavi & John, 1997). Consequently, the filter method generally requires much shorter computation times, making them useful on high dimensional data sets (Hall, 1999). The wrapper approach uses a given attribute subset to train a learning algorithm and returns the predictive performance on the subset. This process is repeated for all subsets that serve as input. In contrast to the filter approach, the wrapper approach hereby does take the generalization error of the algorithm into account. An optimal subset is therefore more likely to be found by means of a wrapper approach, although it comes at the cost of increased computation time.

Standard greedy search operators, such as forward selection, backward elimination and bi-directional search, guarantee to find a local optimum. After each iteration, evaluated candidate solutions are never reconsidered (Hall, 1999). However, this does not have to be a global optimum, which corresponds with the optimal set of attributes. More sophisticated methods with a stochastic nature exist, such as Genetic Algorithms.

3.5.4 Performance evaluation

Several performance evaluation measures are available when evaluating classification models. The most common method is the model accuracy, which refers to the rate of correctly classified instances. Sensitivity refers to the rate of correctly classified positive instances, whereas specificity refers to the rate of correctly classified negative instances.

To reduce overfitting of the model, the generalization error is commonly estimated by using k-fold cross-validation (CV) (Fröhlich et al., 2004). An input space is partitioned into k subsets, in which $k-1$ subsets are used to train the model and the remaining subset serves to evaluate the error. This step is repeated k times, in which each of the k subsets is used at least once for evaluation of the error function. A gap between the generalization error on $k-1$ subsets and the test error on the remaining subset indicates overfitting. The final test error estimate is the average out of all k test error measures. In general, a k-value of 10 is recommended, as is shown that the inductive algorithm with ten folds tends to be approximately stable by reducing variance and increasing bias. A smaller value of k leads to an increase in variance due to instability of the training sets (Kohavi, 1995). Finally, stratified CV is generally recommended (Kohavi, 1995), in which each of the k folds has the right proportion of class values.

3.5.5 Relevant applications of ML

In the field of cardiology, researchers recently started to apply more ML. The most commonly used algorithms are Logistic Regression (LR), Artificial Neural Networks (ANNs), Support Vector Machines (SVMs), and Decision Trees (DTs) (Kadi et al., 2017). In the field of OHCA survival, LR is used by a vast majority of studies, whether or not accounting for SES, race or ethnicity (e.g. Becker et al., 1993; Buick et al., 2016; Clarke et al., 2005; Rajan et al., 2015; Sayegh et al., 1999; Wells et al., 2016; White et al., 2016). The same trend is observed in studies that investigate bystander CPR among communities. In the few studies that examine the relation between OHCA incidence and community-level SES measures, applied methods vary from the Spearman's rank correlation (Feero et al., 1995) to incidence rate ratios (Reinier et al., 2006, 2011). However, no ML approaches are applied. Consequently, prior studies do not provide us any reference of the predictive performance of any ML algorithm in a similar setting.

Few studies that consider OHCA use predictive modelling at an individual level. Krizmaric et al. (2009) deploy six algorithms to predict OHCA outcomes in Slovenia, and find varying results for each classifier between eight distinct analyses. The only exception is Random Forest (RF), which outperforms all other methods, including ANNs, SVMs, and DTs. Insufficient information is provided to conclude whether observed differences are statistically significant. inherent to the method, factors that contribute most to the high accuracy obtained by RF are not identified. Other studies that are not discussed here, predict a cardiac arrest within 72 hours (Liu et al., 2015; Ong et al., 2012), or defibrillation outcome of OHCA victims (Chicote et al., 2016; Howe et al., 2014).

Recently, a study by Weng et al. (2017) received much attention in both the public and scientific world. The authors successfully improve the prediction of myocardial infarction among 378,256 patients of an established algorithm based on guidelines of the American College of Cardiology. Four methods are deployed: LR, ANN, RF, and Gradient Boosted Trees (GBTs). All 30 attributes are included in each model, of which a majority is clinical, and two attributes concern the patient's ethnicity and SES (measured by Townsend deprivation index), respectively. Hence, no attribute selection is performed, which may negatively affect predictive results as discussed earlier, especially in the case of LR. Attribute ranking was achieved by the coefficient effect size (i.e. odds ratio) for LR, the selection frequency as a decision node for RF and GBT, and the overall weighting in ANNs. This approach is likely to be oversimplified, since attribute importance in tree-based models highly depend on the level in which the particular attribute is included. Similarly, using weights of attribute importance in ANNs ignores possible dependencies on other attributes (Leray & Gallinari, 1999). Nevertheless, although the identification of the most important attributes highly varied between methods, all four achieved significantly higher results than the baseline algorithm. Moreover, RF was outperformed by the three alternative methods.

Dominic et al. (2015) use four clinical datasets of 303 patients to predict CVD health status. A GA is applied to reduce the 75 attributes to a subset of with low between-attribute correlation and high correlation with the target attribute. The number of classes is set to five, and the number of patients between classes vary considerably. Six algorithms are applied: DT, Naïve Bayes, SVM, LR, ANN, and Adaptive Boosting. Highest accuracy is found by the Adaptive Boosting algorithm. The authors do not state the use of any parameter tuning, which can significantly reduce performance of several classifiers. In a comparable study, Austin et al. (2013) classify 4,515 patients suffering from heart failure according to one of two heart failure subtypes. The authors use 34 clinical attributes and apply LR, SVM, DT, RF, bagging, and AdaBoost. Slightly higher performance is achieved by AdaBoost after tuning the maximum tree depth parameter. Parameters of other methods are not tuned. Nevertheless, performance of traditional tree-based algorithms was found significantly worse than the more complex tree-based alternatives included in the study. In contrast with these studies, Das et al. (2009) use a supervised ensemble method of three ANNs to predict which patients suffer from CAD. Using typical risk factors, the authors obtain a classification accuracy of 89.01%. The use of multiple ANNs within the same classification process however most likely results in relatively low computation speed.

Finally, ML methods are also widely used in many other medical fields, such as the prediction and prognosis of several types of cancer and diabetes. Based on an extensive literature review, Kourou et al. (2015) show that ANNs, Bayesian algorithms, SVMs, and DTs are widely and successfully applied to predict cancer susceptibility, recurrence, as well as survival. Highest classification accuracies are generally achieved with ANNs and SVMs. Similarly, SVMs obtain high accuracies in the field of diabetes research (i.e., biomarker discovery, prediction of diabetes diagnosis outcomes, and prediction of diabetic complications) (Kavakiotis et al., 2017).

3.6 Resuscitation efforts

3.6.1 Relation between OHCA incidence and bystander CPR

Numerous studies examine the relation between OHCAs and bystander-initiated cardiopulmonary resuscitation (BCPR). We limit ourselves to those studies that examine spatial disparities in both CPR provision and OHCA incidence. Raun et al. (2013) identify census tracts with a low BCPR rate relative to a high OHCA incidence rate in Houston, Texas. Fosbøl et al. (2014) identify census tracts in North Carolina with high incidence and low BCPR based on median rates. In total, 21,8% of the considered OHCAs occurred in areas with low incidence and high BCPR rates. Finally, Straney et al. (2015) separately examine disparities in OHCA incidence and BCPR rates in Victoria, Australia. Incidence rates are about three times as high in the area with the highest rate, compared to that with the lowest. In addition, a 25.1% absolute difference in BCPR rates is observed between the areas with the highest and lowest rates. Based on the results of these studies, we conclude that considerable geographic differences in CPR provision do seem to exist.

3.6.2 Relation between OHCAs and AED deployment

Several studies show that structured public access defibrillation programs in communities improve response to victims of OHCAs (Chan et al., 2016; Hallstrom et al., 2004; Kitamura et al., 2010; Tsai et al., 2012). However, we find only one study that examines the relation between the number of deployed public access AEDs and the number of OHCAs among communities. Moon et al. (2015) examine this relation in Phoenix, Arizona. By combining a cardiac arrest heat map with AED counts in each area, the authors visually identify regions with a lack of AEDs. Although more OHCAs occurred in communities with high population density, only a small number of AEDs was deployed in these areas. A Spearman's rank correlation test shows a weak correlation (rho = 0.283; p = 0.002) between OHCA counts and the number of AEDs in city areas. The authors conclude that future policy and public access defibrillation programs need to be revised.

CHAPTER 4 – METHODS

This chapter describes the methodology that is used in this study. Sections 4.1-4.5 state the standard elements of this study. Section 4.6 describes which software is used during the data pre-processing and subsequent data analysis phase. Section 4.7 then gives a detailed description of each pre-processing step taken in this study. Section 4.8 briefly states the geocoding procedure of OHCA cases. Section 4.9 then provides the computations of OHCA incidence rates that serve as input to our three main analyses. The methodology of the first main analysis, spatio-temporal variation of OHCA, is discussed in Section 4.10. The methodology of the second analysis, the association between OHCA incidence and demographic, socio-economic, and ethnicity factors, is extensively discussed in Section 4.11. Finally, Section 4.12 provides the methodology to investigate the current state of resuscitation efforts.

4.1 Study design

This is a retrospective population-based cohort study. Anonymous data is used from the cardiac arrest registry of the AmsteRdam REsuscitation STudies (ARREST). The study design of the ARREST registry is described in more detail in Section 2.3.

4.2 Study setting

The North Holland region covers an area of 2,390 km^2 and has a population of approximately 2,5 million. In February 2010, the study area was expanded to the region of Twente, which covers an area of 1,489 km^2 and has a population of about 625,000.

4.3 Study population

We consider all EMS-treated OHCA events with a (presumed) cardiac cause that occurred between January 1, 2006 and December 31, 2016 in North Holland, and those between February 1, 2010 and December 31, 2016 in Twente. Cardiac arrest victims of any age are identified as having an OHCA. Furthermore, both public and private locations are included.

4.4 Data sources

4.4.1 Cardiac arrests

This study uses anonymous cardiac arrest data from the ARREST registry.

4.4.2 Regional definitions

In the Netherlands, municipalities (“gemeenten”) are divided into districts (“wijken”) and neighbourhoods (“buurten”). Neighbourhoods are the lowest regional level and are delineated along lines of planning or socio-economic structure. Neighbourhoods are mostly dominated by a given type of land use or buildings, such as an industrial area or a residential area with mainly high-rise or low-rise buildings (Statistics Netherlands, 2016). Districts are composed of one or more contiguous neighbourhoods. Furthermore, districts generally have a homogenous socio-economic structure or planning, such as a residential, industrial or recreational structure (Statistics Netherlands, 2016). Municipalities determine the division of their area into districts and neighbourhoods. Statistics Netherlands coordinates this division at a national level (Statistics Netherlands, 2016). Based on visualizations of the regional boundaries, we find the neighbourhood level too detailed to successfully perform analyses of a ‘rare’ event such as OHCA. We therefore perform our analysis at a district level.

4.4.3 Digital geometry

The digital geometry of districts is made available annually by Statistics Netherlands as ESRI™ shapefiles. Borders of municipalities are based on the Basic Registration Kadaster, whereas the borders of districts are based on reports of municipalities (Statistics Netherlands, 2016).

4.4.4 Socio-economic and ethnicity data

Statistics Netherlands, the Dutch Ministry of the Interior and Kingdom Relations (BZK) and the Netherlands Institute for Social Research (SCP) offer publicly available data at a regional level. We aim to include socio-economic factors separately rather than through a predefined definition of socio-economic status (SES). Therefore, we only consider data from Statistics Netherlands, as it is the only source that offers individual factors rather than a single socio-economic measure. We do not include the quality of living index from BZK or the SES scores from SCP to avoid interaction.

Statistics Netherlands gathers and stores key socio-demographic and socio-economic data in the publicly accessible database ‘StatLine’ (“Statistics Netherlands StatLine database,” n.d.). Data can be found for COROP regions, municipalities, postal codes, districts and neighbourhoods. Indicators have been made available in the categories of, among others, basic demographics, housing, income, social protection, occupation, and ethnicity. A majority of attributes is measured per capita or per private household. Unfortunately, not all indicators are measured every year. In addition, data on several attributes has not been made available for some areas because the values are either unknown, unreliable, or classified to prevent that information can be traced back to individuals, companies or institutions (Statistics Netherlands, 2016).

4.4.5 AED locations

Location data of public access AEDs has been made available by HartslagNu.

4.5 Data collection

4.5.1 Cardiac arrest data

We use collected data of all emergency calls in which the EMS dispatcher suspects an OHCA. We then exclude all cases for which no resuscitation was started. Next, we exclude all OHCAs with an obvious non-cardiac (i.e., traumatic) cause. Data of the remaining EMS-treated OHCA events with a (presumed) cardiac cause in both Twente and North Holland is used in this study. The data contains the following information per case: the patient’s age, the patient’s gender, the resuscitation date, the resuscitation location (town, postal code, street name, house number, type of location), whether CPR was given before ambulance arrival, and whether an AED was connected before ambulance arrival.

4.5.2 Digital geometry

We import annual ESRI™ shapefiles of districts from Statistics Netherlands into ArcGIS 10.5 software. We remove the digital geometry in the area of Amsterdam Schiphol Airport as well as those outside of our study regions. Table 4.1 shows the resulting number of districts per year.

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
North Holland	272	273	273	273	274	267	272	272	272	270	373
Twente	-	-	-	-	104	104	104	119	119	119	119

Table 4.1. *Number of districts per year in both study regions.*

Along with changes in structures over time, several districts were assigned a different distinct code. Due to our multi-year study period, a fraction of these regions was assigned a different code more than once. Moreover, attribute field names differ considerably between years. This increases the complexity of joining annual data sets into one larger set.

4.5.3 Socio-economic and ethnicity data

We obtain available key socio-economic ethnicity data at the district and municipality level from Statistics Netherlands. We retrieve data files corresponding to each year of our study period.

4.5.4 AED location data

Latitude-longitude coordinates are retrieved for 1,135 AEDs in North Holland and 1,036 AEDs in Twente. Locations are included for all public access AEDs that were registered into the system at the time of the collection (April 2017). AED location data is gathered from municipalities in which HartslagNu is used at the time of the collection. We therefore consider all public access AED locations in 25 out of 41 municipalities in North Holland and in all 14 municipalities in Twente.

4.6 Software

ArcGIS 10.5 software (Environmental Systems Research Institute (ESRI) Inc., Redlands, CA) is used to create maps of geographic distributions of OHCA events and OHCA risk assessments at a regional level. We pre-process collected data on cardiac arrests, AED locations, socio-economic factors and ethnicity by using Extraction, Transformation and Loading (ETL) tools from Alteryx Designer 11.0. Alteryx is an advanced self-service analytics platform. Spatio-temporal variation is assessed by using Statistical Package for the Social Sciences (SPSS) 24 (IBM Corporation, Chicago, Illinois, United States of America). We then use RapidMiner to perform our machine learning (ML) analyses, which is a GUI-based data science platform that offers a wide range of attribute selection and ML algorithms. Moreover, RapidMiner is one of the current market leaders of data science platforms (Linden et al., 2017). Finally, Microsoft Excel 2016 (Microsoft, n.d.) and Tableau Desktop 10.4 (Tableau Software, n.d.) are used to visualise the results from our data analyses. Tableau Desktop is Business Intelligence data visualisation software to visualise data in sophisticated, interactive graphs and tables.

4.7 Data pre-processing

4.7.1 Cardiac arrest data

We exclude all OHCA events that occurred at Amsterdam Airport Schiphol prior to the geographic data analysis, due to several reasons. First of all, the airport is served by an autonomous emergency system and emergency dispatch centre. Second, patients suffered from cardiac arrest in the area are considered not to be representative of neighbourhood demographics. Third, airports generally are priority locations in terms of AED deployment and EMS personnel and are typically heavily monitored (Chan et al., 2016; Sasson, Magid, et al., 2012).

4.7.2 Structural reforms

Regional reforms in structure in the Netherlands lead to new encodings, reforms at lower regional levels, and require recalculations to track developments over time (Keuning, 2012). Therefore, we cope with changes in regional structures over time to allow longitudinal studies. Next to changes in structures, several regions were assigned a different unique code (existing or non-existing) over time. Due to the multi-year study period, a fraction of these regions was assigned a different distinct code more than once. Moreover, attribute names differ significantly between years. These changes increase the complexity of joining the annual data sets into one larger set that represents the entire period. We distinguish four types of regional changes:

1. A change of the district's unique code.
2. A split-up of a single districts into two or more regions.
3. A merge of at least two districts into one region.
4. Restructuring of at least two districts in at least two new regions.

In ArcGIS, we identify all districts located in both study areas for each year. We then identify any changes by comparing ID codes between consecutive years. All districts that are not matched based on this code, are used as input to be matched based on the district's name. Afterwards, all remaining districts are explored visually in ArcGIS for both years. We then combine the districts to each year of Statistics Netherlands data. We start by adding a column to each file containing the year that the statistical information represents. We cleanse the data by removing unwanted characters (e.g. Unicode characters and whitespaces) and negative values. We assign each attribute with a unique, consistent field name and set the appropriate corresponding data type. After that, we merge the annual data sets into one large set. Irrelevant data (e.g., districts in Twente between 2006-2009) in the resulting set is removed. For each district in which any change occurred, we then do the following:

1. If any code change, merge or restructuring occurred, we change all ID codes *prior* to the final code change, to the final ID code.
2. If any split-up occurred, we change all ID codes *after* the final code change, to the ID code before the split-up.

We give special attention to the few districts with more than one change, to ensure that the new sets of districts are valid. We then combine all districts that are changed over time by summarizing those based on similar ID codes. Regional data is recalculated based on appropriate weighted averages. For instance, combining the annual proportion of men of two districts that are merged into one, is done by using the number of inhabitants in the particular year as weights. To determine the degree of urbanisation, we first estimate the surrounding address density (OAD) in our established sets of districts. We add up the average number of residences and businesses (all types) throughout the study period, and divide the result by the total land surface (in km^2). Based on these values, the degree of urbanisation can then be determined (see Table 4.2).

Degree	Criterion	Category
1	$\mathrm{OAD_{est}} \geq 2500$	Extremely urbanised
2	$1500 \leq \mathrm{OAD_{est}} < 2500$	Strongly urbanised
3	$1000 \leq \mathrm{OAD_{est}} < 1500$	Moderately urbanised
4	$500 \leq \mathrm{OAD_{est}} < 1000$	Hardly urbanised
5	$\mathrm{OAD_{est}} < 1500$	Not urbanised

Table 4.2. *Degree of urbanisation and corresponding classification, based on criteria of Statistics Netherlands (Statistics Netherlands, 2016).*

4.7.3 Region exclusion

We exclude regions that do not meet eligibility criteria in any year within our study period. Until 2013, income statistics have not been made available if the respective region or constructed set of regions comprises less than 200 inhabitants. Since 2013, this threshold is set on 100 income receivers within private households. Similarly, the minimum number of households for publishing statistics on the proportion of individual recipients of the WWB is set on 50 until 2014. Hence, we exclude regions in the following order:

1. The number of inhabitants needs to be at least 200.
2. The number of private households needs to be at least 50.
3. The number of income receivers from private households need to be at least 100.

In addition, we consider the community-wide SES of regions that are dominated by industry not to be representative for the particular region. We include two criteria to remove industrial sites that have not yet been excluded due to the threshold population:

4. The name of the region cannot contain the Dutch word for either 'company' ('bedrijf'), companies ('bedrijven'), industry ('industrie') or harbour ('haven').
5. We calculate a company/population ratio by simply dividing the number of listed companies in a particular region by the total population. The ratio cannot exceed 1, implying that the particular area comprises more companies than inhabitants.

Finally, data on any of the attributes needs to be available for at least one year.

4.8 Geographic data analysis

Based upon the address information (street, house number, postal code, city) of registered OHCA locations, the events are geocoded and plotted in ArcGIS 10.5 software. We exclude all events that did not occur in our study area. Cases of which the highest level of accuracy is either a city or a 4-digit postal code level, are also excluded from further analysis.

4.9 OHCA incidence rates

Prior studies that examine the relation between OHCA incidence and socio-economic factors mainly analyse standard incidence rates, and, to a smaller extent, age-adjusted rates. By contrast, we additionally examine standard and age-adjusted night-time rates. The purpose of including night-time OHCAs is to examine whether excluding victims that arrested during office hours significantly changes the observed relation between OHCA incidence and community-level SES.

We determine the time of day that each OHCA event occurred as either daytime or night-time. Definitions of night-time in studies focusing on OHCAs differ considerably within the range of 06:00 pm and 07:59 am (Brooks et al., 2010; De Vreede-Swagemakers et al., 1998; Demirtas et al., 2015; Hansen et al., 2013; Ong et al., 2008; Ro et al., 2015; Søholm et al., 2014; Sun et al., 2016). Since office hours in the Netherlands vary roughly between 08:00 am and 06:00 pm, we define night-time OHCAs as those occurred between 08:00 pm and 06:59 am. Consequently, daytime is defined as the interval between 07:00 am and 07:59 pm. We calculate the crude incidence rate per 100,000 person-years of observed population over the entire study period in each district by adding up the number of OHCAs, divide the sum by the total population over time and multiply the result by 100,000. Similarly, we calculate the crude incidence rate limited to the OHCAs that occurred during night-time only. The number of OHCA victims in Twente in 2010 are adjusted for only including cases on 11 of the 12 calendar months, since the region was not yet included in the ARREST study in the first month of the particular year.

We adjust both population-based incidence rates for age structure by using the direct standardization method (Fleiss, 1981). We directly adjust for the proportion of inhabitants of less than 65 years, as well as the proportion of at least 65 years of age. First introduced by Neison (1844), this method enables us to compare incidence rates across districts after accounting for differences in age distributions. This may in turn result in new insights in terms of SES and ethnicity measures. The entire population in the considered districts in North Holland and Twente respectively is used as the reference population. In each district, we calculate the incidence rate per 100,000 population over the entire study period by adding up the number of OHCA victims that belong to each age group and divide the amount by the sum of the exposed population. We multiply the result with the fraction of the reference population that belongs to the exposed age group, summarize the rates for both age groups, and multiply the resulting rate with 100,000.

4.10 Spatio-temporal variation of OHCAs

We investigate whether there is any systematic difference between the number of cardiac arrests in each observed district among multiple years. Modelling annual OHCAs at a regional level most often concerns modelling of 'rare' events. Especially for districts in rural areas, reporting no cardiac arrests in any given year is not uncommon. This however does not tell us whether the number of OHCAs remains low throughout the entire study period. To prevent a majority of annual measurements to be equal to zero, we consider only those districts that report an average of at least two cardiac arrests per year. Therefore, we focus on the most important subset of districts in which differences in public health programs can be made.

4.10.1 Target attribute

Suppose that the annual number of OHCAs in a particular district with approximately 1,000 inhabitants varies between zero and three throughout a two-year study period. If the number of OHCAs equals zero in exactly one year, a significant gap appears between the two reported annual incidence rates per 100,000 population (i.e., 0 vs. 300). In contrast, the absolute difference in the number of OHCAs between years equals no more than three. To avoid the presence of such large gaps, we examine the spatio-temporal variation by using the number of OHCAs rather than corresponding incidence rates, and include the annual population as offset. The offset is an additional predictor of which the coefficient is always assumed to be 1. This way, we expect more OHCAs in highly populated areas.

4.10.2 Model type

We model 'rare' OHCA events in relatively small communities. Hence, our count data is expected to be non-Gaussian. Not only are data transformations not always feasible in situations with highly skewed data to obtain normality of residuals and homogeneity of variance (Zivin & Bartko, 1976), transformations may also induce loss of information. Changes over time may be made smaller and consequently more difficult to detect (Zuur et al., 2010). Rather than transforming data to be used in a traditional statistical framework, with the risk of losing important properties in the process, we choose a generalized linear mixed model (GLMM) as our statistical method. GLMMs are a class of regression models that accounts for multiple measurements within the same subject. The values of these repeated measurements do not necessarily have to identical, and can be compared relative to a certain grouping attribute (e.g. "before treatment" vs. "after treatment"). The data of the target attribute does not have to follow a normal distribution. GLMMs are common in longitudinal studies in which repeated measurements are correlated and errors are not normally distributed (Bolker et al., 2009; Gu & Ma, 2005).

4.10.3 Objective

Our objective of applying a GLMM is to test how much of the variance in the number of OHCAs within districts is explained by the random effect of time. In other words, we test whether the number of OHCAs, taking into account the yearly district population, significantly differs for any of the included study years. To this end, we first need to model the non-normal distribution. In general, counts of events can be modelled by means of a Poisson or a negative binomial distribution. As the number of OHCAs are expected to be relatively low in rural districts, we take into account the presence of overdispersion, that is: more variance is observed than predicted by the statistical model (Bolker et al., 2009). Standard errors may be underestimated if the presence of overdispersion is not accounted for (Heck et al., 2013). Compared to the Poisson model, the negative binomial model contains an additional parameter that accounts for random variance. Hence, we assume a negative binomial distribution to overcome overdispersion.

4.10.4 Model design

GLMMs distinguish two types of effects. Fixed effects are constant factors within the analysis, while random effects represent a sample from a larger population. As we consider all feasible districts in our study area (i.e., the entire district 'population'), we define districts as fixed effects. In addition, to make results representative for all years, including those that are not considered in this study, we include each year within our study period as a random effect. We then specify the district level to be the level at which the random effect of time may vary. The GLMM also requires a certain link function, which models the relation between the linear predictor and the independent variables. Few link functions are suitable to be used in the negative binomial mixed model (Heck et al., 2013). We use the log link function, which is commonly used in modelling count events. In addition, an offset can be added to the model. The offset is an additional predictor, of which the coefficient is always assumed to be 1. Offset values are therefore directly added to the model. To account for differences in population, we add the annual number of inhabitants (x 10,000 in Twente; x 100,000 in North Holland) as offset. The resulting mixed-effects model is defined as follows:

$$\log\left(E\left(Y_{ij}\right)\right) = \log(P_i) + X_i\beta + Z_jb + \varepsilon \tag{4.1}$$

In this formula:

- Y_{ij} refers to the response variable; the number of OHCAs in district i in the j^{th} year.
- P_i is the offset; the population in district i.
- β is the vector of fixed effects of districts X_i.
- b is the vector of K random effects for the sample of years Z_j.

Finally, we need to assume a certain covariance structure. The covariance matrix specifies the variance and covariance parameters of the random effects within one district. By choosing a certain covariance structure, we specify the form and complexity of our random effects within each district. No full consensus among statisticians exists concerning which covariance matrix to use in mixed-effects models (Kincaid, 2005). We consider two relatively simple covariance structures, namely 'variance components' (VC) and 'compound symmetry' (CS). Both structures assume equal variances among years between the predicted and the observed number of OHCAs within each district. In this way, we can test how much of the variance within districts is explained by the random effect of time. We define the variance within each district between two included years as σ_Y^2. If significant, we reject the null hypothesis H_0: $\sigma_Y^2 = 0$ and accept the alternative H_1: $\sigma_Y^2 > 0$. VC also assumes that the covariances of all pairs of random effects are equal to zero. Residuals observed in one year are independent of those observed in the remaining years, regardless of the distance in time. CS however assumes that within-subject errors between years are in fact correlated.

4.10.5 Model application

We apply our GLMM in SPSS Statistics 24.0 separately for North Holland and Twente. Statistical significance is based on the z-test, and 95% CIs of σ_Y^2 are computed.

4.11 Predictive modelling

We apply ML algorithms to find relations between OHCA incidence rates and demographic, socio-economic, and ethnicity attributes. To this end, given a set of attributes, we predict whether each district belongs to the set of approximate 50% of districts with relatively high incidence rates, or to the approximate 50% of districts with relatively low incidence rates. The distinction between these classes is based on the median incidence rate. Our objective is to find the values of district characteristics that separates the instances of both classes in our training data as good as possible. This can be done by applying an ML algorithm that separates the data by means of a decision boundary, as demonstrated in Figures 4.1 and 4.2. We interpret the models that show the highest predictive accuracy on unseen data.

We separately consider four different incidence rates as target attributes in our analyses: incidence rates based on OHCAs that occurred during the entire day or during night-time only, as well as standard and age-adjusted incidence rates. All manually selected attributes other than age are considered in each of these configurations. Furthermore, each configuration is evaluated by all selected ML algorithms. Analyses are performed separately for North Holland and Twente.

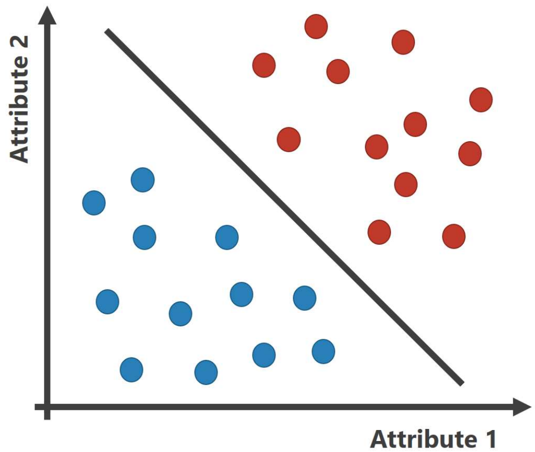


***Figure 4.1**. A decision boundary that separates a linearly separable data set.*

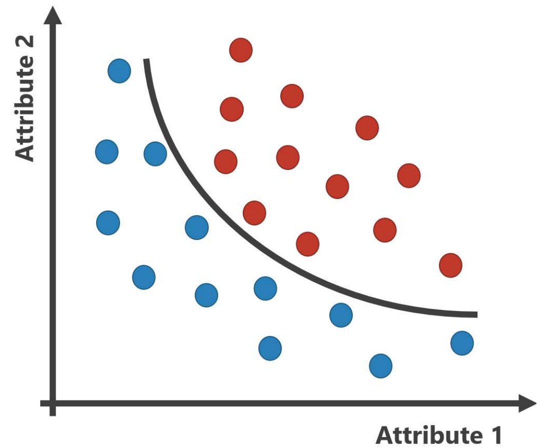


***Figure 4.2**. A non-linear decision boundary that separates instances from two classes.*

Settings	*C1*	*C2*	*C3*	*C4*
Time of day	Full	Full	Night	Night
Age-adjusted incidence rates	No	Yes	No	Yes

***Table 4.3.** Different configurations included in this study.*

Four configurations and two study regions implies that our attribute subset space consists of eight subsets. Hence, eight subsets serve as input to our dimensionality reduction. For each configuration, we take a closer look at the algorithm that yields the highest performance.

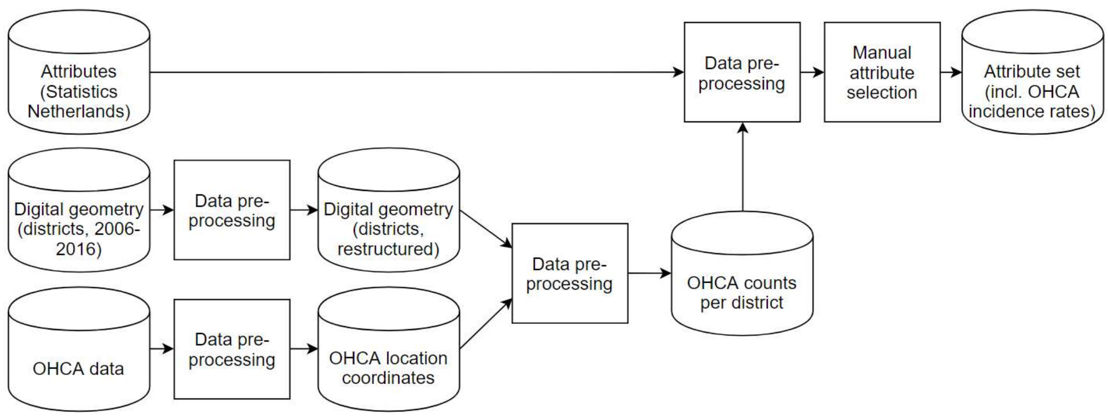
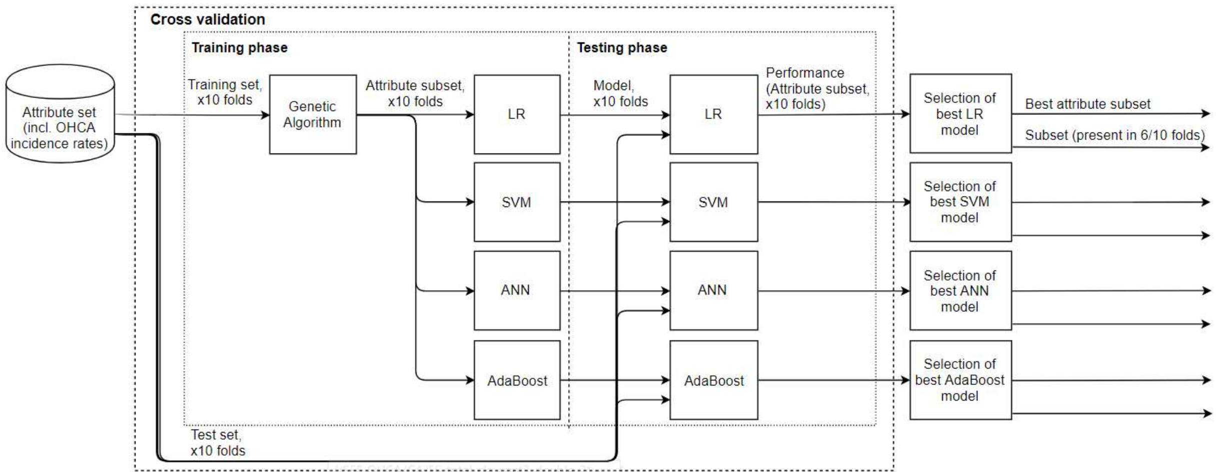


***Figure 4.3.** Schematic overview of data-processing phase.*

A schematic overview of the pre-processing phase is given in Figure 4.3. We merge the pre-processed demographic, socio-economic and ethnicity data from Statistics Netherlands with the OHCA count data into one heterogeneous data set. We then manually narrow down the number of attributes. To further reduce the set of attributes to the most relevant subsets, we construct Genetic Algorithms (GAs). GA is an adaptive search heuristic that can be used to solve complex optimization problems. The subset of attributes found by the GA are then used by a classification algorithm to train a model. We test the model on a separate test set that has not been used to train the model. Four different classifiers are used that show promising predictive performance in similar studies: Logistic Regression (LR), Artificial Neural Networks (ANNs), Support Vector Machine (SVM), and Adaptive Boosting (AdaBoost). For each classifier, the model that yields the highest predictive accuracy is selected. Predictive accuracy is estimated with ten-fold cross validation (CV), in which the test error estimate is the average out of all ten test error measures. An overview of our learning approach is given in Figure 4.4. This process is repeated in both regions for four different OHCA incidence rates and two different GAs. Subsets are evaluated with a separate test through four different classification algorithms.

First, we discuss the attribute selection phase in more detail. We then discuss the performance measurement and the methodology of hyperparameter tuning in this study. In the remainder of this chapter, we describe the selected models.



Note: OHCA = Out-of-Hospital Cardiac Arrest, LR = Logistic Regression, SVM = Support Vector Machine, ANN = Artificial Neural Network, AdaBoost = Adaptive Boosting.

Figure 4.4. *Schematic overview of automated dimensionality reduction and subsequent learning phase for each configuration evaluated in this study. The process is repeated for four different OHCA incidence rates and two different Genetic Algorithms.*

4.11.1 Manual attribute selection

We manually reduce the number of attributes to be collected to the most relevant subset. We do not consider those attributes for which data is only available prior to 2010. We then remove all mutually exclusive attributes, followed by removing highly dependent attributes, that may negatively influence the results. Finally, we remove measures of similar information. Ideally, we would include education levels among our districts. This data is unfortunately not available. Hence, we focus on a wide range of other measures of SES. In 0, we describe the manual attribute selection phase in detail. As shown in Table 4.4, 16 attributes remain that will serve as input to the automated dimensionality reduction. This corresponds to 2^{16} possible subsets. Attribute categories include measures of basic demographics, housing, income, social protection, and ethnicity. To reduce within-group heterogeneity of non-western ethnic groups, we only consider each non-western ethnic group separately.

4.11.2 Automated dimensionality reduction

We apply GAs for automated attribute subset selection. GA is an adaptive search technique based on the foundations of natural selection and genetics (Holland, 1975). The algorithm can be applied to search the space of attribute subsets to find globally optimal subsets (Fröhlich et al., 2004; Vafaie & De Jong, 1992; Yang & Honavar, 1998). More importantly, GAs can efficiently search high-dimensional input spaces of which little is known beforehand. In comparison with other attribute subset selection techniques, GAs are relatively insensitive to 'noisy' attributes (De Jong, 1988), increasing the probability to find a global optimum rather than a local optimum. To effectively explore the input space to identify a global optimum, GAs require randomization (Forrest, 1993). The improved performance of GAs comes with the cost of considerably higher computation time compared to traditional attribute selection techniques. Moreover, GAs take relatively long to converge and may cause some overfitting. A high-level overview of all steps taken during the GA procedure is shown in Figure 4.5. A description of the genetic operators and an overview of the GA's hyperparameter settings are stated in Appendix D.

Category	*Attribute*	*Value*
Demographics	Degree of urbanisation[1]	#
	Inhabitants of at least 65 years old	%
	Male inhabitants	%
	Married inhabitants	%
Housing	Average household size	#.#
	Average property value	€
Income	Income receivers that belong to the national 40% with the lowest income	%
Social protection	Recipients of social benefits based on the Work and Social Assistance Act	%[2]
	Recipients of social benefits based on labour disability acts[3]	%[2]
	Recipients of social benefits based on the Unemployment Benefits Act	%[2]
Ethnicity	Inhabitants with a Western background[4]	%
	Inhabitants with a background from the Netherlands Antilles or Aruba	%
	Inhabitants with a background from Morocco	%
	Inhabitants with a background from Suriname	%
	Inhabitants with a background from Turkey	%
	Remaining inhabitants with a non-Western background[5]	%

[1]*Estimated by adding up the number of residences and the number of businesses.*
[2]*Out of the total working age population (15-65 years old).*
[3]*WAO, WAZ, WIA, Wajong.*
[4]*An individual's background is considered to be Western if he or she, or at least one of his or her parents is born in Europe (excluding the Netherlands and Turkey), North America, Oceania, Indonesia or Japan.*
[5]*Other than from the Netherlands Antilles, Aruba, Morocco, Suriname and Turkey.*

Table 4.4. *Overview of the attributes that serve as input to our automated dimensionality reduction.*

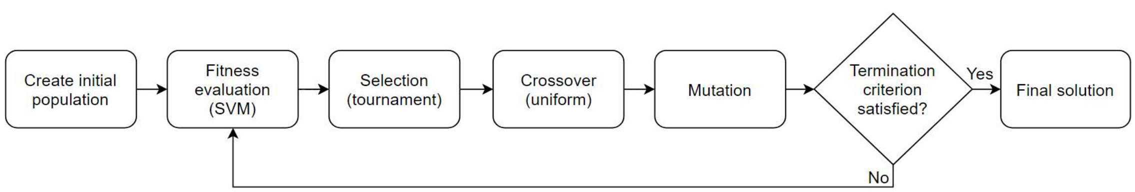


Figure 4.5. *High-level overview of the steps in our Genetic Algorithm, with an SVM as classifier.*

We represent the space of all possible attribute subsets by means of a binary presentation in which attributes are modelled as genes and attribute subset as individuals. In a setting with n attributes, individuals are modelled as binary strings of length n. In each of these strings, $n_i = 0$ indicates exclusion and $n_i = 1$ corresponds to inclusion of the i^{th} attribute (Vafaie & De Jong, 1992). We initialize our GA in generation $j = 1$ using a population size of 25 individuals. For each of these individuals, we set the probability that the i^{th} attribute is included ($n_i := 1$) equal to $p_{init} = 0.5$. The algorithm terminates when a maximum number of 50 generations is reached, or when no improvement has been made for 15 consecutive generations.

The fitness of an individual in any generation is defined as the predictive performance after stratified ten-fold cross-validation (CV). For each configuration, we choose to apply GAs with different fitness functions. The method used to evaluate fitness levels need to generally show high predictive classification accuracy within short computation time. Based on that, we choose Logistic Regression (LR) and a Support Vector Machine (SVM) as our fitness functions. Especially the SVM is commonly used for fitness evaluation purposes, showing high predictive ability in both linear and non-linear classification tasks. In case underlying relations are linear, LR may outperform alternative methods. The specified type of kernel of the SVM is the radial basis function (RBF), also known as the Gaussian kernel. Hence, we seek to select an optimal attribute subset, given optimal values of our regularization parameter C and scale parameter γ (Fröhlich et al., 2004). Different values for parameters C and γ change the allowed degree of misclassified instances near the decision boundary. Although this may lead to an increase in the training error, the intent of changing these values is to decrease the error on the test data. This in turn is achieved by being less dependent on points along the decision boundary, which significantly affects performance when these are not meaningful.

4.11.3 Performance measures

We include a stratified ten-fold CV to estimate the generalization error (i.e., error estimate on unseen data). The average error estimate is computed out of all ten test error measures. We measure the predictive accuracy, sensitivity and specificity as our performance metrics. Accuracy refers to the proportion of correctly classified instances, out of all predicted instances. Sensitivity refers to the proportion of correctly classified high-risk areas out of all high-risk areas. Finally, specificity refers to the proportion of correctly classified low-risk areas out of all low-risk areas. We construct 95% confidence intervals (CIs) of these performance measures by making use of the normal approximation to the binomial distribution.

4.11.4 Hyperparameter tuning

Hyperparameters are those model parameters of which the values are set prior to the learning stage. Performance can vary considerably between different hyperparameter values. For each applied model, we aim to find the most suitable hyperparameters that maximizes classification performance. To this end, we apply grid search, in which every combination of hyperparameter values out of a specified set of values is evaluated.

We implement an inner stratified ten-fold CV for any hyperparameter tuning during the training phase, in addition to the outer CV for the generalization error estimation. In other words, 90 percent of the training data in each of the ten CV folds is used to identify the most promising hyperparameters, which in turn are also found through ten-fold CV. These hyperparameters are then used in classifying the entire training set. Hence, we implement any hyperparameter tuning as an integral part of our algorithm training process (Witten et al., 2016).

4.11.5 Model selection

The majority of studies to OHCA incidence that are discussed in Section 3.3.2 perform univariate analyses, i.e., the number of attributes to be analysed is restricted to one. Two of these studies include a pre-defined socio-economic index as a single attribute (Soo et al., 2001; Straney et al., 2016). An incidence rate ratio (IRR) is computed by several of these studies, to assess the relative risk of the occurrence of OHCA across the population in areas associated with low SES, versus the relative risk across areas with high SES. The objective of our analysis is to find those models that show the strongest association between OHCA incidence and demographic, socio-economic, and ethnicity factors. Interpreting these models gives us information which district characteristics are most associated with high OHCA incidence. Rather than computing a wide range of univariate analyses, we propose an alternative, ML approach that includes multiple multivariate analyses. This approach allows us, among others, to evaluate the relative importance of different attributes within the same model. Moreover, different types of ML algorithms are applied and evaluated given the same data sets, in order to find the most strong association with OHCA incidence. One of these ML algorithms is Logistic Regression (LR). LR is a regression model in which the target attribute only takes binary values. It is used in few OHCA incidence studies that investigate the association between OHCA incidence and population density (Strömsöe et al., 2011) or urbanisation (Soo et al., 2001) across geographic areas. Since the incidence of OHCA can be considered rare (the common guideline is below 10% of the considered population (Zhang & Yu, 1998)), the odds ratio obtained by LR approximates the relative risk. Moreover, LR is also used by many studies that examine the indirect association between neighbourhood characteristics and OHCA outcome (e.g. Ahn et al., 2011; Buick et al., 2016). We include LR as a baseline for our other ML methods. The performance of LR varies between medical settings, although high accuracy is still reported. In general, LR performs well when no interactions between attributes exist, and data can be clearly separated by a single, linear boundary (Goldstein et al., 2016). Therefore, we deploy different types of ML algorithms that perform well in medical studies and can handle non-linearities and interactions between attributes.

Studies that predict risk in the field of cardiovascular disease (CVD) through multiple ML algorithms are discussed in Section 3.5.5. ANNs and SVMs are frequently applied across these studies and consistently return relatively high performance. In addition, tree-based algorithms are frequently included as well. Traditional tree-based learners such as the Decision Tree (DT) are however mostly outperformed by more complex tree-based alternative learners. With relatively small data sets, Dominic et al. (2015) and Austin et al. (2013) both include six different classifiers. In both studies, Adaptive Boosting (AdaBoost) yields highest classification performance. We therefore choose to apply ANNs, SVMs and AdaBoost. Both ANNs and AdaBoost can handle both non-linearities as well as interactions between attributes. Especially ANNs can exploit complex interactions, making the algorithm likely to outperform LR when important interactions exist (Tu, 1996). SVMs can also handle non-linearities, but are unable to specifically exploit the presence of any interactions. Nevertheless, SVMs have shown excellent generalization performance in classification studies. As a result, SVMs also perform well when only few instances are available to train upon. In addition, a key advantage of SVMs is the use of linearly constrained convex quadratic programming, which avoids to be trapped in local minima and provides only one global minimum (Cristianini & Shawe-Taylor, 2000). By contrast, the non-convex optimization problem of DTs and ANNs allow for the presence of many local optima. Moreover, computation time of SVMs and most tree-based algorithms are considerably low, especially compared to ANNs. A final drawback of ANNs is the interpretability of the resulting network, a problem which can partially be solved by limiting the size of the network.

We conclude that each of the methods mentioned above has various strengths and limitations. The predictive power of the three algorithms other than LR is shown by, among others, Caruana & Niculescu-Mizil (2006). In an extensive comparison study of ML algorithms, the authors find SVMs, ANNs and boosted trees to be part of the strongest classification models. In contrast, LR shows poor performance when averaged over all metrics included in the study.

4.11.6 Logistic Regression

LR is a probabilistic generative model, and was first introduced by Cox (1958). It is commonly used to predict binary outcomes, and is considered to be part of both statistics and ML. LR restricts posterior probability values to the interval $[0,1]$ by means of a sigmoid ('logistic') activation function. A probability estimate is computed for each district. This estimate tells us how likely each district is considered to have relatively high OHCA incidence. If this value exceeds 0.5, we predict the district to have high incidence. If this value is lower than 0.5, we predict a district that corresponds with a relatively low incidence. We implement this standard version of LR in RapidMiner within a ten-fold cross-validation.

4.11.7 Artificial Neural Networks

Artificial Neural Networks (ANNs) consist of nodes and edges and are inspired by neurons and synapses of the human brain. In its most basic form, it consists of a single neuron and is called a perceptron (Rosenblatt, 1962). The perceptron handles multiple input values and yields a single output value. It is based on the (possible) transmission of signals between neurons of the brain, conditional to the activation of input signals. Similar to LR, perceptrons are linear classification models. However, rather than computing probability estimations, a linear perceptron learns to separate instances that belong to different classes by means of hyperplanes (i.e., decision boundaries). This type of classification algorithm is called a linear discriminant function, which is described in detail in Appendix O.

A non-linear classification process requires several adaptations to the perceptron algorithm. This includes the use of multiple neurons, divided over more than one layer. The multilayer perceptron (MLP) is the most basic and common type of an ANN. An example of an MLP is shown in Figure 4.6. The values on the edges in the network are the weights that the algorithm learns during the training of the network.

The training phase of an MLP starts with random weights. After passing forward through the network, the overall classification error in the output neuron is measured. This step is known as forward propagation. A backward pass, also known as backward propagation, is then applied. In backward propagation, the partial derivative of the error with respect to the weights in each node of the network is calculated, using a clever application of the chain rule (not discussed here). This step is required to compute how the error function changes as the weights change. Given a predefined learning rate, all weights in the network are then updated before the next iteration is initiated. A small value of the learning rate results in small incremental improvements and thus requires a large number of improvement steps to converge. Conversely, a large value results in large incremental improvements, resulting in few required steps. However, the latter comes at the price of risking too large improvement steps, which would prohibit the algorithm to converge to the global optimum in the convex space.

The process of forward and propagation is repeated multiple times. The training phase terminates when the weights of the network have converged, resulting in a non-decreasing error function. Using weight decay decreases the learning rate during the training phase. This allows for large changes to the weights at the beginning of the training phase, and small changes later on. Finally, adding a momentum term improves performance by allowing the next weight change to be proportional to the previous weight change. This speeds up the algorithm and helps to avoid local minima (Tu, 1996). A more detailed description on backward propagation can be found in Bishop (2006).

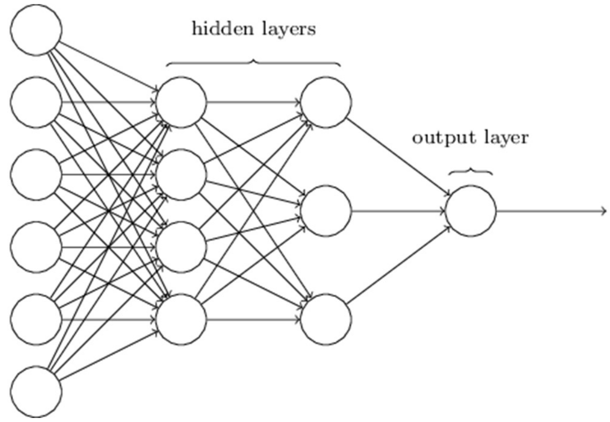


Figure 4.6. *An example of a simple Artificial Neural Network (Witten et al., 2016). The network contains six input nodes, two hidden layers with three nodes each and one output layer with a single node.*

The size of our input layer matches the number of attributes in the particular subset that is evaluated, as each attribute is modelled as a single neuron. In addition, we restrict ourselves to the use of one hidden layer, since large ANNs generally require large amounts of data to tune parameters during the training phase. More specifically, we restrict the number of layers in the network to three: one input layer, one hidden layer and one output layer. We add a bias node to each layer, which is a constant and comparable to the intercept term in regression models.

We implement RapidMiner's standard ANN algorithm. To avoid overfitting of hyperparameters, we implement a similar, separate ANN within an inner ten-fold cross-validation (CV). This particular CV is referred to as 'inner', since the CV is located within the CV of the first ANN. To this end, hyperparameters are selected based on 90% of the available training data. We let both the learning rate and the momentum vary between 0 and 1. To limit the number of options, we vary both hyperparameters on the following grid: [0.3, 0.5, 0.7, 0.9]. The combination of values for the learning rate and momentum that yields the highest accuracy with 90% of the data is used as hyperparameter settings within the ANN that is in turn trained based on 100% of the data. In addition, we activate weight decay to decrease the learning rate during the training algorithm. Appendix N provides an overview of all hyperparameters of the ANN in this study.

Our weights and our predicted classes for each district are thus updated throughout the training phase. Note that a change in the weights may imply that other previously correctly classified instances are now misclassified. Hence, the algorithm does not necessarily reduce overall error in each step. Moreover, training an ANN takes relatively long compared to other algorithms. More than one solution may exist, and the solution to be found depends on the initialization of the network's parameters (Bishop, 2006).

4.11.8 Support Vector Machines

Support Vector Machines (SVMs) algorithms were first introduced by Vapnik (1995). As opposed to many ML algorithms that construct any feasible decision boundary, the SVM constructs the optimal decision boundary by maximizing the distance between the nearest instances (the ‘support vectors’) of two different classes. The difference between feasible decision boundaries and the optimal decision boundary found by SVM is shown in Figures 4.7 and 4.8 respectively.

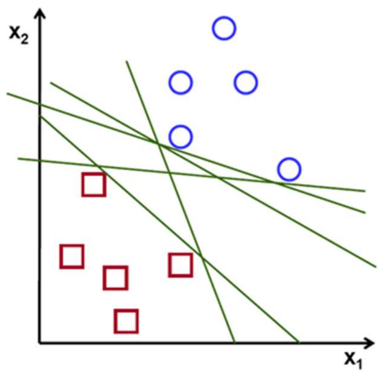


Figure 4.7. *Feasible hyperplanes, given training instances of two classes.*

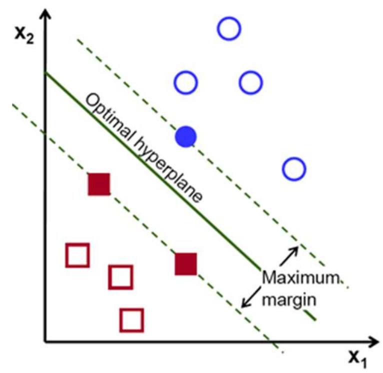


Figure 4.8. *Optimal hyperplane, given support vectors of two classes.*

The use of the optimal plane arises from the structural risk minimization principle and thus aims to minimize overfitting. The algorithm is based on the fact that with a suitable function, instances of different classes are always linearly separable (Duda et al., 2012). The SVM uses kernel functions to map instances into a new higher dimensional space, and linearly separate the instances of both classes by means of an optimal hyperplane (shown in Figure 4.9). In favour of computation speed, the SVM limits the number of training instances that are used to construct this plane to the support vectors (Witten et al., 2016).

LIBSVM is a popular open source ML library to apply SVMs. We implement the library’s SVM algorithm within RapidMiner’s predefined LIBSVM function. Similar to the GA discussed in Section 4.11.2, we apply hyperparameter tuning to find optimal values for regularization parameter C and scale parameter γ to improve predictive accuracy. Using suitable values for these hyperparameters within a given problem can considerably increase predictive accuracy. We implement an inner cross-validation that uses 90% of the available training data to find optimal hyperparameters of C and γ, each time that an SVM is trained. Values for both hyperparameters are commonly tuned along the following logarithmic grid (Fröhlich et al., 2004): $[10^{-3}, 10^{2}, \dots, 10^{2}, 10^{3}]$. Therefore, we implement this hyperparameter tuning scheme as well, leaving 49 options to be evaluated by a similar, separate SVM located within the inner cross-validation. The combination of values for C and γ that yields the highest accuracy with 90% of the data is used as hyperparameter settings within the SVM that subsequently is trained based on 100% of the data.

Although attributes can be ranked based on their absolute values in a linear SVM, this is not the case if a kernel function is included. However, the sign of the weights still tells us whether a positive or negative relation is observed.

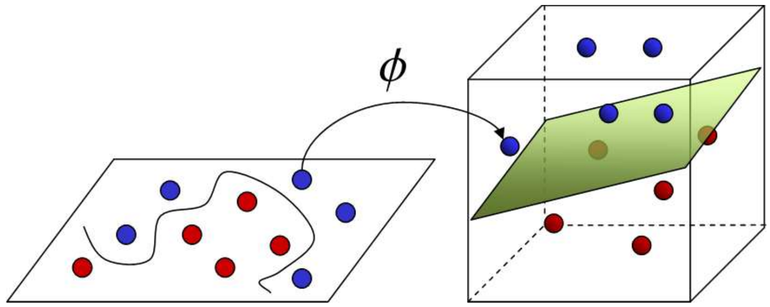


Figure 4.9. *Simplified representation of mapping instances from a 2D space into a new 3D space by a kernel SVM, in which the hyperplane is constructed.*

4.11.9 Adaptive Boosting

Ensemble learning algorithms aim to combine many weak learners into one strong learner. A weak learner produces error rates that mostly are only slightly better than random guessing (Hastie et al., 2009). The combination of many different models can intuitively be explained by considering a committee consisting of W specialists who all excel in a limited domain. A decision made by the ensemble of specialists who all excel in a limited domain may outperform decisions made by an individual expert (Witten et al., 2016). The type of ensemble learner that we apply in this study is Adaptive Boosting (AdaBoost) (Freund & Schapire, 1996). Boosting is an iterative algorithm, and in each step, a weak classifier is applied to a sample taken with replacement. In the first step, weights of all instances are equal to 1. After each iteration, instances that were incorrectly classified are assigned a relatively higher weight than the weights of correctly classified instances. This process is repeated until no improvement has been made, or until the maximum number of iterations is achieved. Predictions are then made through weighted voting of the individual predictions of weak learners. More specifically, the class prediction $C^{(i)}$ of district i corresponds to the prediction made by the majority of W learners, times the weight assigned to that learner. We will use DTs as the base learner of our AdaBoost algorithm. According to Breiman, using a DT in a boosting algorithm is one of the best classifiers available (Hastie et al., 2009). In addition, even the simplest AdaBoost algorithm with a DT as a base learner generally shows substantial improvement of classification accuracy over extensive DTs (Hastie et al., 2009).

The Decision Tree (DT) algorithm categorizes instances through a set of rules in a tree like representation (Breiman et al., 1984). An example of a DT model in the context of this study is given in Figure 4.10. The nodes of the tree correspond to decisions to come up with subsets given a certain cut-off value. Both the attribute and cut-off value that best discriminate a set of instances based on their class value $C^{(i)}$ are used. DT is a greedy heuristic, as subsets are created based on locally optimal decisions. DTs help us to understand the hierarchy within the set of attributes. The algorithm consists of two phases: the construction of the tree, followed by pruning of the resulting tree. In the construction phase, a single attribute is branched upon in each step, resulting in two new nodes at the next level of the tree. In each iteration, the attribute that corresponds to the highest information gain is chosen. Information gain is a concept from the field of information theory and measured in bits, by computing the difference in entropy values $H(a)$ before and after a potential split on attribute a. The entropy $H(x)$ of a random variable x is defined as follows (Bishop, 2006):

$$H(x) = -\sum_{x} p(x) \log p(x) \tag{4.2}$$

The negative sign in this formula guarantees $H(x) \geq 0$. Low probability values correspond to a high amount of information. The algorithm proceeds until one of the following criteria is satisfied:

1. $H_{prior}(x) - H_{after}(x) = \nabla H(x) < b$, for all x. In this formula, b is the minimal gain that needs to be achieved through splitting.
2. The maximum number of instances in the nodes at the lowest level is lower than the allowed number of instances to produce a split (i.e., minimal size for split).
3. The minimum number of instances as a result of a possible split is lower than the allowed number of instances in any node in the tree (i.e., the minimal leaf size).
4. The predefined maximum tree depth (i.e., the maximum number of splits) is reached. This includes the split at the lowest level that defines the decision boundary.

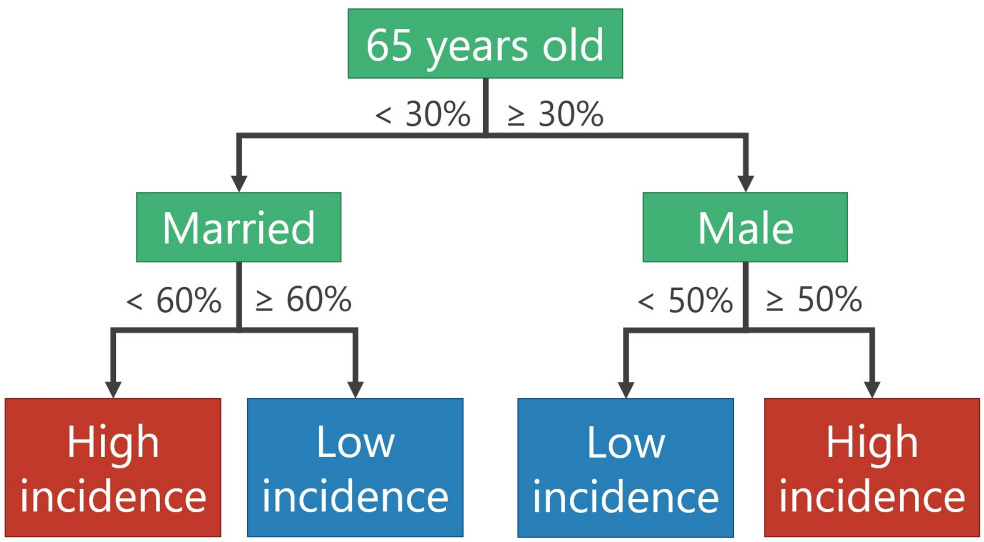


Figure 4.10. *Example of a Decision Tree within the context of this study. The attribute 'proportion of inhabitants of 65 years of age' was first split upon at a cut-off value of 30%, as it (in this example) provides the most 'pure' split in terms of entropy. The resulting two subsets are further split upon by making use of the attribute 'the proportion of married inhabitants' and 'the proportion of male inhabitants' respectively. The subsequent splits provide the risk prediction that will be made in terms of OHCA incidence for all corresponding districts.*

Since the nodes at the lowest level are not further split upon, these nodes are commonly referred to as terminal nodes, or the leaves of the tree. One of two class values (i.e. high or low risk) is then assigned to all instances in the same leaf. As a result, DTs construct a collection of decision boundaries orthogonal to one of the axes. The number of correctly classified instances is calculated in each node. After constructing a tree, the algorithm recursively tests and removes branches that do not provide added value in the classification process. This step is called pruning and reduces overfitting. Pruning is discussed in more detail by e.g. Hastie et al. (2009). Appendix N gives an overview of the settings of AdaBoost in this study.

The algorithms of AdaBoost and DT are both available within RapidMiner. We implement the DT within our AdaBoost, and set the maximum number of iterations in AdaBoost to 10. Similar to the other ML algorithms, we install AdaBoost within our CV mechanism. To prevent overfitting and improve interpretability of our results, we set the maximal depth of any DT to 6. Since we have considerably more instances in North Holland compared to Twente, we set the minimal leaf size 50% higher in North Holland (6 vs. 4).

4.12 Resuscitation efforts

We separately analyse the relation between each considered resuscitation effort and (standard) OHCA incidence rates. In line with our predictive modelling approach, we classify districts as either high- or low-risk based on the median incidence rate across all districts.

For each resuscitation effort, we first analyse the entire set of districts. We then run separate analyses of urban and rural districts, as large differences may be found. For example, a higher population density is likely to increase the probability of receiving CPR prior to arrival of EMS services. The type of urbanisation of each district is based on definitions of Statistics Netherlands (shown in Table 4.2). Districts with an urbanisation degree of 1-3 are defined as urban, whereas those with a degree of 4 or 5 are defined as rural. We apply LR algorithms in each of the settings. Odds ratios and corresponding 95% CIs are calculated. Moreover, p-values of less than 0.05 indicate statistical significance.

We first determine the number of OHCA cases in each district for which cardiopulmonary resuscitation (CPR) was given prior to arrival of EMS services. Consequently, we compute the bystander CPR rate in each district as the fraction of all OHCA victims that received CPR. Similarly, we compute the proportion of applied AEDs in each district. As a reference, we compute the usage of ambulance defibrillators. We then compute the number of deployed public access AEDs per capita in each of the districts.

Finally, for each resuscitation effort, we visually observe whether spatial disparities exist among the 50% of districts with relatively high OHCA incidence rates in terms of either bystander-initiated CPR rates, AED usage, and the number of allocated AEDs per capita. If applicable, we identify those districts for which the most unfavoured disparities in resuscitation efforts are observed.

CHAPTER 5 – RESULTS

5.1 Data processing

We find 11,963 geocoded events eligible for further analysis. The patient inclusion and geocoding procedure are described in Appendix D. After correcting for changes in geographic structures, we find 256 districts in North Holland and 97 in Twente. Flowcharts and maps of the structural reforming process are shown in Appendix F. Ten districts (3.9%) in North Holland and seven (7.2%) in Twente do not meet our eligibility criteria (see Appendix G). At least one year of data on each of the attributes is present for the remaining sets of districts. The fraction of missing data for each attribute is given in Appendix H. Characteristics of the remaining districts are given in Table 5.1. The total population, urbanisation, the average property value, and the proportion of low income receivers are higher in North Holland compared to Twente. Moreover, the proportion with a (non-)Western ethnic background is higher in North Holland for all but the Turkish background. The proportion of inhabitants of at least 65 years of age, the average household size and the proportion without occupation (receiving WW or WWB) is in turn higher in Twente.

	North Holland (n = 246)			Twente (n = 90)		
	$\bar{X}$	*SD*	*Interval*	$\bar{X}$	*SD*	*Interval*
Cardiac arrests						
Number of OHCAs	3.7	7.1	[0, 53.5]	2.7	2.7	[0, 13.4]
Night-time (8:00 pm–6:59 am)	1.2	2.4	[0, 18.6]	0.9	1.0	[0, 5.1]
OHCA incidence (per 100,000)	44.3	28.8	[0, 251.0]	43.6	19.2	[0, 112.8]
Night-time	14.6	10.3	[0, 77.0]	13.7	10.0	[0, 47.9]
OHCA incidence, age-adjusted	41.9	24.6	[0, 205.6]	43.2	18.1	[0, 106.7]
Night-time	13.8	9.4	[0, 68.3]	13.8	11.1	[0, 75.0]
Demographic & socio-economic factors						
Population, x 1,000	9.9	19.7	[2.4, 138.3]	7.0	7.5	[0.2, 34.5]
Urbanisation, degree	3.2	1.4	[1, 5]	3.9	1.2	[1, 5]
At least 65 years old, %	16.8	5.8	[2, 53]	17.8	4.0	[7, 29]
Male inhabitants, %	49.8	1.6	[42, 54]	50.6	1.4	[47, 54]
Married inhabitants, %	42.6	5.8	[21, 57]	44.7	5.4	[26, 54]
Average household size	2.3	0.3	[1.5, 3.1]	2.5	0.4	[1.6, 3.2]
Average property value, x €100,000	2.6	1.0	[1.3, 9.8]	2.4	0.7	[1.1, 3.9]
Receivers of low income*, %	37.2	4.2	[23, 49]	43.7	4.0	[32, 60]
Recipients of the WWB, %	1.4	1.8	[0, 9]	2.0	2.7	[0, 13]
Recipients of labour disability acts, %	7.5	3.2	[2, 31]	7.0	2.7	[2, 17]
Recipients of the WW, %	2.0	0.6	[0, 4]	3.0	0.9	[0, 6]
Ethnicity background						
Western, %	8.5	3.1	[2, 24]	6.8	4.0	[2, 18]
Non-Western, %	8.2	9.2	[0, 20]	5.5	7.2	[0, 33]
The Netherlands Antilles / Aruba**, %	0.5	0.8	[0, 6]	0.1	0.4	[0, 1]
Morocco**, %	1.0	2.1	[0, 20]	0.2	0.5	[0, 2]
Suriname**, %	1.5	2.6	[0, 32]	0.2	0.5	[0, 2]
Turkey, %	1.6	3.7	[0, 36]	2.5	4.5	[0, 23]
Remaining non-Western, %	3.4	2.7	[1, 23]	2.3	2.5	[0, 14]

Abbreviations: $\bar{X}$ = average, n = number of districts; SD = Standard deviation; OHCA = Out-of-Hospital Cardiac Arrest.
*Income receivers that belong to the national 40% of people with the lowest income. **Excluded in Twente.

Table 5.1. *Descriptive statistics of districts in North Holland (2006-16) and Twente (2010-16). The average, standard deviation, and the range of values across all districts are shown for each attribute.*

5.2 OHCA incidence rates

5.2.1 North Holland

Figure 5.1 shows a map of average annual OHCA incidence rates across our included districts in North Holland. Figure 5.2 illustrates these rates sorted from high to low. The median annual standard incidence rate throughout the study period across the considered districts is 38.7. In 37 districts (15.0%), this rate equals at least 60. In total, 12 districts (4.9%) had OHCA incidence rates above 80, which corresponds to approximately twice the median rate. Appendix J highlights the districts that correspond with the highest incidence rates throughout the study region. Eight out of ten districts for which the incidence rate given the entire time of day was highest, are also listed as having the highest night-time rates. This is in line with overall differences in classifications, as 185 districts (75.2%) are either high or low risk for both rates (see Table 5.2). Table 5.3 shows the computed age distributions, which are used to adjust crude incidence rates per 100,000 person-years for age structure. Table 5.4 shows median rates for each configuration. Geographic distributions are illustrated in Appendix I.

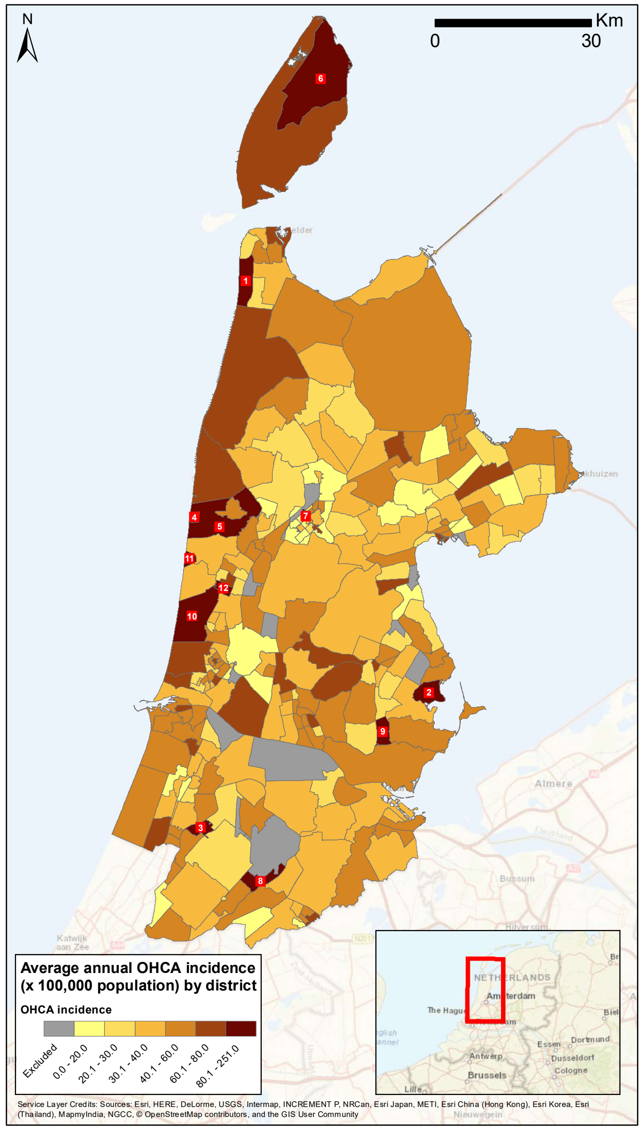


Figure 5.1. *Average annual OHCA incidence rates (entire day) by district in North Holland (256 districts, 2006-16). The numbers indicate the rank (from highest to lowest incidence rate) of the particular district, across all considered districts with OHCA incidence rates above 80.*

IR_1	IR_2	NH (*n* = 246)	TW (*n* = 90)
$IR_{standard,full}$	$IR_{standard,night}$	24.8	22.2
$IR_{standard,full}$	$IR_{adjusted,full}$	14.6	12.2
$IR_{standard,night}$	$IR_{adjusted,night}$	12.2	7.8

Abbreviations: n = number of districts, NH = North Holland, TW = Twente, IR = incidence rate.

Table 5.2. *Difference (in %) between binary classifications (i.e., high or low risk) of districts.*

	NH (*n* = 246)	TW (*n* = 90)
Younger than 65 years of age (%)	85.2	82.6
At least 65 years of age (%)	14.8	17.4

Abbreviations: n = number of districts, NH = North Holland, TW = Twente.

Table 5.3. *Computed age distributions in both study regions, based on the set of included districts in North Holland (246 districts, 2006-16) and Twente (90 districts, 2010-16).*

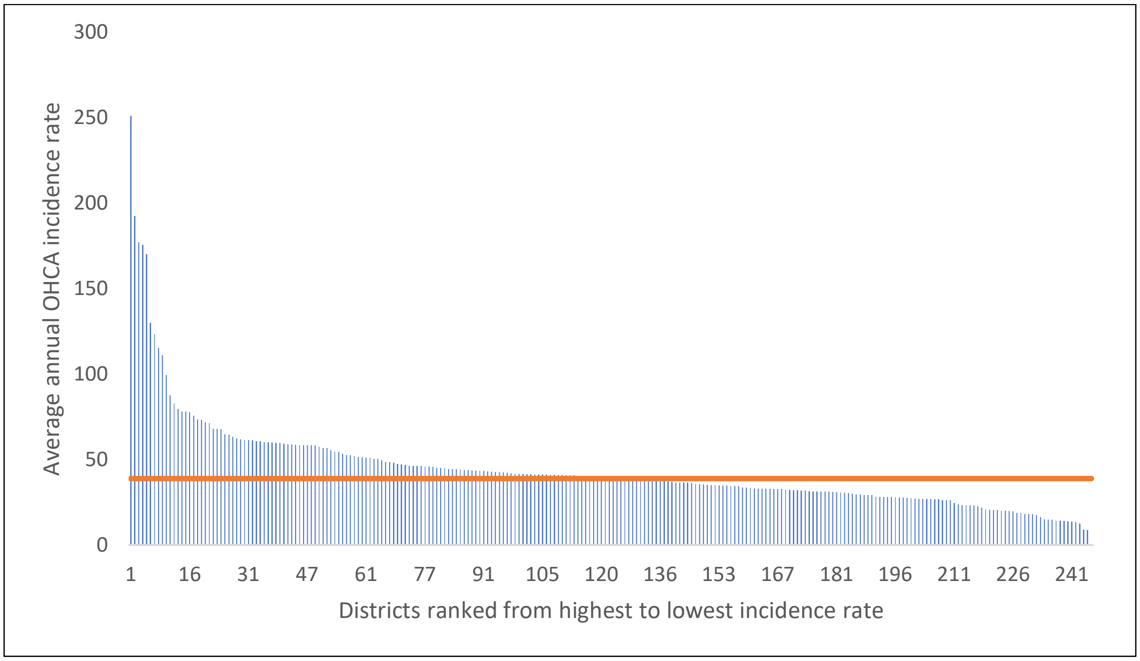


Figure 5.2. *Average annual standard OHCA incidence rates (entire day) across districts in North Holland (246 districts, 2006-16). The orange line indicates the median OHCA incidence across districts (m = 38.7).*

5.2.2 Twente

The median OHCA incidence rate across the districts in Twente throughout the study period is 40.6. We observe 16 districts in which the incidence rate is 60 or higher (17.8%). Moreover, OHCA incidence rates are above 80 in four districts (4.4%). Six out of ten districts that correspond with the highest incidence rates given the entire time of day, also correspond with the highest night-time incidence rates. In total, 70 districts (77.8%) have the same classification for both rates (see Table 5.2). No districts are observed without any OHCA patients.

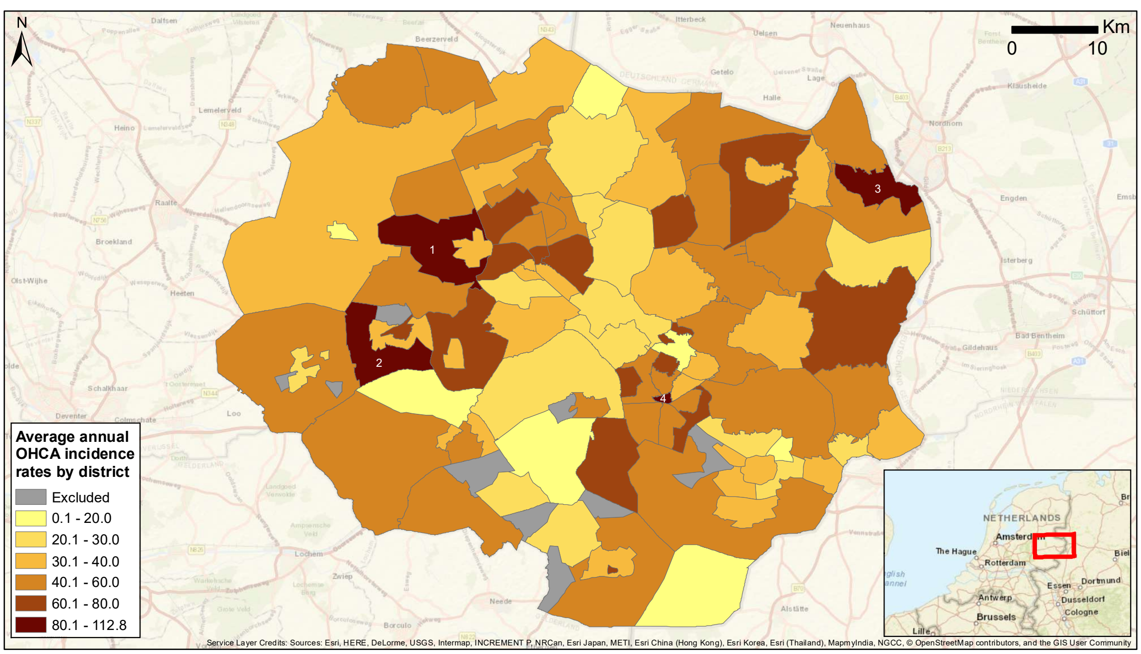


Figure 5.3. *Average annual OHCA incidence rates (entire day) by district in Twente (97 districts, 2010-16).*

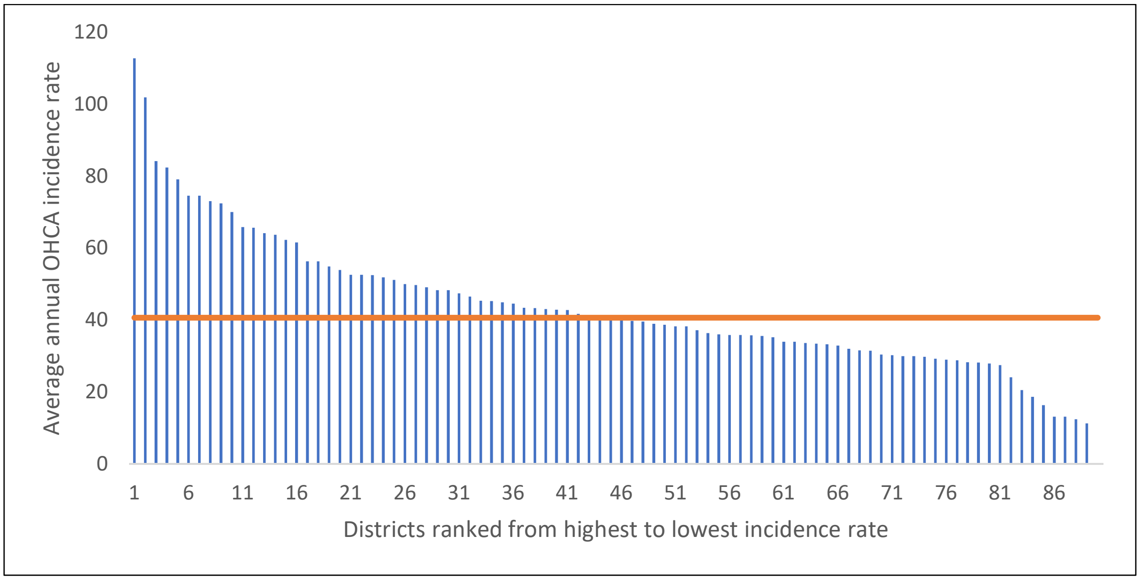


Figure 5.4. *Average annual standard OHCA incidence rates (entire day) across districts in Twente (90 districts, 2010-16). The orange line indicates the median OHCA incidence across districts (m = 40.6).*

	North Holland (n = 246)	Twente (n = 90)
OHCA incidence (per 100,000)	38.7	40.6
Night-time (8:00 pm–6:59 am)	12.9	13.3
OHCA incidence, age-adjusted	37.0	40.3
Night-time	12.5	13.0

Abbreviations: n = number of districts.

Table 5.4. *Median OHCA incidence rate per configuration in North Holland (2006-16) and Twente (2010-16).*

5.3 Spatio-temporal variation of OHCAs

We find 127 districts with an average of at least two cardiac arrests per year in North Holland, and 44 districts in Twente. Generalized linear mixed models (GLMMs) for which the covariance structure of the random effects were modelled as compound symmetry did not yield a positive-definite Hessian matrix. In contrast, the variance components (VC) structure was fit successfully to both sets of districts. From this, we can conclude that assuming no particular covariance between any two of the considered years does yield a feasible model. Characteristics of the successfully fitted models are shown in Tables 5.5 and 5.6.

Values predicted by the GLMMs are shown in Figures 5.5 and 5.6. In both regions, most predicted values correspond to less than five reported OHCAs. Table 5.7 shows the estimated variance σ_Y^2 due to the overall random effect of time based on the VC covariance structure. In both regions, these estimates are small and non-significant. Hence, we do not reject the null hypothesis H_0: $\sigma_Y^2 = 0$. Thus, in both regions, we do not find any significant trend of variation over time.

Residual random effects for each year are shown in Table 5.8. These effects are all found to be significant and considerably larger than σ_Y^2. Hence, a residual random effect seems to exist for each considered year in the particular study regions. This tells us that for each year, measurement variation exists that is not explained by either the modelled fixed effect of the particular districts or the overall random effect of time within our negative binomial mixed-effects regression model.

Characteristics	
Fixed effects, n	127 (districts)
Random effects, n	11 (2006-2016)
Measurements points, n	1397
Covariance structure	VC
Residual effects (covariance), n	11 (2006-2016)
Random effects (covariance), n	1 (time in years)

Note: VC = Variance components.

Table 5.5. *Characteristics of the GLMM in North Holland, considering those districts with an average of at least two OHCAs per year.*

Characteristics	
Fixed effects, n	44 (districts)
Random effects, n	7 (2010-2016)
Measurements points, n	308
Covariance structure	VC
Residual effects (covariance), n	7 (2010-2016)
Random effects (covariance), n	1 (time in years)

Note: VC = Variance components.

Table 5.6. *Characteristics of the GLMM in Twente, considering those districts with an average of at least two OHCAs per year.*

Study region	Random effect	Estimate	SE	95% CI	Z	p-value
North Holland	σ_Y^2	0.012	0.007	(0.004, 0.037)	1.748	0.080
Twente	σ_Y^2	0.054	0.053	(0.008, 0.361)	1.036	0.300

Note: Y = years, SE = Standard Error, CI = Confidence Interval.

Table 5.7. *Time as a random effect in both study regions according to the variance components covariance structure, and adjusted for district variation. Analysis is performed for all districts with an average of at least two cardiac arrests per year.*

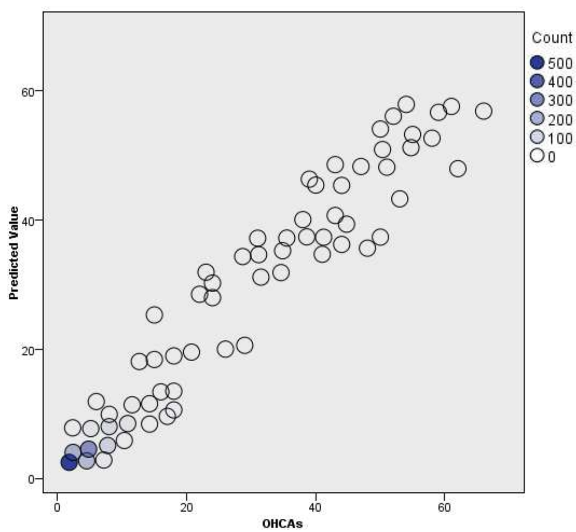


Figure 5.5. *The predicted number of cardiac arrests versus the actual number of cardiac arrests for each measurement point (n = 1397) in the fitted GLMM in North Holland between 2006 and 2016. The included set of districts (n = 127) corresponds with those districts with an average of at least two cardiac arrests per year.*

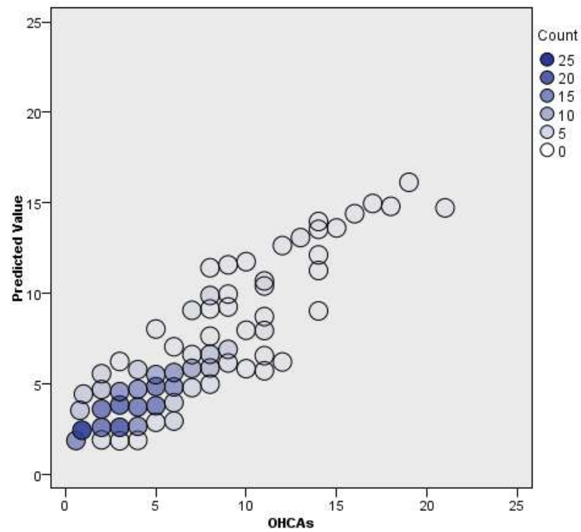


Figure 5.6. *The predicted number of cardiac arrests versus the actual number of cardiac arrests per measurement point (n = 308) in the fitted GLMM, in Twente between 2010 and 2016. The evaluated districts (n = 44) have an average of at least two cardiac arrests per year.*

5.4 Predictive modelling

The median OHCA incidence rate for each configuration (shown in Table 5.4) are used to subdivide the districts into sets of low and high risk respectively. As can be seen, all median rates are found to be slightly lower after adjusting for age. Moreover, as shown in Table 5.1, three out of four attributes in Twente that correspond with the proportion of inhabitants with non-Western backgrounds have very low average ($\bar{X} \leq 0.2\%$) and maximum values ($max \leq 2\%$). Therefore, we only include the proportion with a Turkish background in Twente. Hence, we consider 16 attributes in North Holland, compared to 13 attributes in Twente.

Study region	Residual effect	Estimate	SE	95% CI	Z	p-value
North Holland	σ^2_{2006}	0.736	0.120	(0.535, 1.013)	6.144	0.000
	σ^2_{2007}	0.964	0.150	(0.710, 1.308)	6.416	0.000
	σ^2_{2008}	0.754	0.117	(0.557, 1.021)	6.466	0.000
	σ^2_{2009}	1.229	0.176	(0.928, 1.628)	6.975	0.000
	σ^2_{2010}	0.839	0.127	(0.624, 1.128)	6.625	0.000
	σ^2_{2011}	1.235	0.175	(0.936, 1.629)	7.076	0.000
	σ^2_{2012}	0.628	0.110	(0.446, 0.884)	5.722	0.000
	σ^2_{2013}	0.998	0.150	(0.743, 1.340)	6.656	0.000
	σ^2_{2014}	0.835	0.128	(0.618, 1.127)	6.525	0.000
	σ^2_{2015}	1.021	0.154	(0.760, 1.373)	6.633	0.000
	σ^2_{2016}	1.200	0.178	(0.897, 1.606)	6.733	0.000
Twente	σ^2_{2010}	1.201	0.408	(0.617, 2.338)	2.944	0.003
	σ^2_{2011}	0.806	0.334	(0.358, 1.814)	2.418	0.016
	σ^2_{2012}	0.782	0.351	(0.324, 1.887)	2.224	0.026
	σ^2_{2013}	0.875	0.382	(0.372, 2.059)	2.290	0.022
	σ^2_{2014}	0.958	0.356	(0.463, 1.984)	2.692	0.007
	σ^2_{2015}	0.803	0.291	(0.395, 1.632)	2.763	0.006
	σ^2_{2016}	0.685	0.327	(0.269, 1.746)	2.094	0.036

Note: SE = Standard Error, CI = Confidence Interval.

Table 5.8. *Residual effects according to the diagonal covariance structure, given all districts with an average of at least two cardiac arrests per year. The corresponding random effect is modelled according to a variance components covariance structure, and adjusted for district variation.*

5.4.1 Automated dimensionality reduction

Average and highest predictive generational accuracy found by the Genetic Algorithm (GA) based on standard incidence rates in North Holland is shown in Appendix K. Attribute rankings of those attributes that are found in at least six out of the ten cross-validation (CV) folds are shown in Table 5.9 for North Holland and Table 5.10 for Twente. The relation of these attributes with the considered OHCA incidence rate can be either positive or negative. The exact number of times that each considered attribute was found in each configuration, is shown in Appendix L.

In North Holland, outcomes of both GAs show similarities. Most attributes found by the GA with LR as classifier are also found by the GA-SVM. In three out of four configurations, more attributes are found by the GA-SVM. Household size was found in all eight subsets. The proportion of low income receivers was found in seven subsets, whereas the proportion of married inhabitants was included in five of the subsets. The proportion of inhabitants of at least 65 years old is found for all standard incidence rates. The proportion of WW recipients is found in four out of eight subsets. In addition, the proportion of inhabitants with a background from the Netherlands Antilles or Aruba is included in three subsets. Remaining ethnicity backgrounds are only marginally included.

Also in Twente, more attributes are found through the GA-SVM procedure, compared to GA-LR. The attributes found by each GA procedure are however considerably less similar. For example, average household size was found in all configurations of the GA-SVM, in comparison with only one configuration of the GA-LR. Moreover, age was again included in all configurations concerning standard rates. The proportion of recipients of disability benefits was found in five rankings. The proportion of men as well as the proportion of married inhabitants are found in four and three subsets respectively. The proportion of low income receivers and the average property value are included in two subsets. Similarly, attributes reflecting ethnicity backgrounds are included in no more than two of the rankings.

GA Time IR	*LR Full Standard*	*Full Adjusted*	*Night Standard*	*Night Adjusted*	*SVM Full Standard*	*Full Adjusted*	*Night Standard*	*Night Adjusted*
1	HH size[a]	HH size	Income	Men	HH size[a]	Urban[a]	HH size	Men
2	Income[a]	Income[a]	HH size	HH size	Urban[a]	Antilleans[a]	Married	HH size
3	Age ≥ 65	Urban[a]	Married	WW	Income[b]	HH size[b]	Age ≥ 65[a]	Married
4	Men	WW	Age ≥ 65	Income	Age ≥ 65[b]	Western[b]	WW[a]	Income
5	Married	Antilleans	Moroccan		Married[b]	Moroccan[b]		WW
6	AO				Antilleans[b]	Income[c]		
7					Men[c]	Turkish[c]		
8					*Western*[c]			
9					*Other nW*[c]			

Abbreviations: IR = Incidence rate, GA = Genetic Algorithm, HH = Household, AO = labour disability, nW = non-Western.
[a, b, c]*Ranks within the particular configuration are equal.*

Table 5.9. *Attribute overall ranking of the automated dimensionality reduction phase in North Holland for each configuration. Only the attributes that are found in at least six out of ten cross-validation folds are considered.*

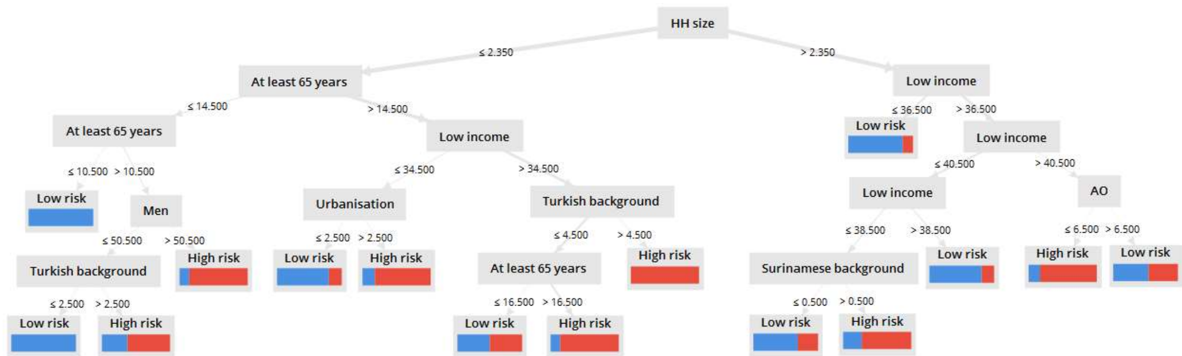
GA Time IR	*LR Full Standard*	*Full Adjusted*	*Night Standard*	*Night Adjusted*	*SVM Full Standard*	*Full Adjusted*	*Night Standard*	*Night Adjusted*
1	Age ≥ 65	WOZ[a]	Married	AO	Age ≥ 65[a]	HH size	Age ≥ 65	Income[a]
2	AO	AO[a]	Men	HH size[a]	WW[a]	Married[a]	Men	WOZ[a]
3		Urban	Age ≥ 65	Western[a]	Income[b]	WW[a]	AO	Other nW[a]
4		Men	Other nW	Turkish[a]	Urban[b]		HH size[a]	AO
5				Married	Men[b]		WW[a]	HH size[b]
6					Western[b]			Urban[b]
7					HH size			

Abbreviations: IR = Incidence rate, GA = Genetic Algorithm, HH = Household, AO = labour disability, nW = non-Western.
[a, b, c]*Ranks within the particular configuration are equal.*

Table 5.10. *Attribute overall ranking of the automated dimensionality reduction phase in Twente for each configuration. Only the attributes that are found in at least six out of ten cross-validation folds are considered.*

5.4.2 Classification results

Classification results for each method and configuration are shown in Appendix M. In both study regions, LR structurally provides higher predictive accuracy than Artificial Neural Networks (ANNs), SVMs and Adaptive Boosting (AdaBoost) in predicting areas with high and low incidence. LR is only outperformed by SVM when analysing age-adjusted night-time rates. The best LR models are based on attribute subsets obtained from the GA-LR procedure, whereas the best SVM model given night-time rates uses input from the GA-SVM procedure. These four attribute subsets correspond with the best fold of the particular GA. In fact, subsets that comprise of those attributes present in more than half of the GA's CV folds structurally result in lower predictive accuracy compared to accuracy based on subsets found in the best fold. Furthermore, AdaBoost and ANN show worse predictive accuracy in all configurations. In North Holland, both methods are even outperformed by LR and SVM on the basis of all three performance measures. The highest predictive accuracy obtained with AdaBoost in North Holland is (mainly) achieved by means of the Decision Tree shown in Figure 5.7.



HH = household, AO = labour disability acts.

Figure 5.7. *Decision tree with the highest weight in AdaBoost, given the best performing weight combination (GA-LR) in North Holland (246 districts, 2006-16, full day, non-adjusted OHCA incidence rates). All numbers except household size are given in proportions. The thickness of the arrows correspond with the number of districts passed through the branches of the tree. In the leaves of the tree, the red color refers to the proportion of actual high risk areas, whereas the blue color refers to the proportion of actual low risk areas. Corresponding accuracy is 73.2%. The Decision Tree is constructed by using RapidMiner software.*

North Holland

Attributes and corresponding odds ratios (ORs) of the best LR models are shown in Table 5.11. Significant predictors for standard OHCA incidence rates given the entire time of day are average household size, the proportion of income receivers that belong to the national 40% with the lowest income, the proportion of at least 65 years, and the proportion of men. The proportion of married inhabitants as well as the proportion with other non-Western backgrounds are non-significantly and positively related to standard rates. The final insignificant predictor is the proportion of recipients of labour disability acts. After adjusting incidence rates for age structure and separately searching the attribute input space for the optimal subset, we again find the proportion of low income receivers and the average household size as the most important and significant determinants in the model. Insignificant predictors positively related to age-adjusted incidence are the proportion of inhabitants with a Western background, urbanisation (a higher value implies a more rural area), the proportion of WW recipients and the average property value. Given standard night-time rates, we find the average household size and the proportion of married inhabitants as significant predictors. The proportion with a Western background is non-significant and negatively related to night-time OHCA incidence, whereas the proportion with other non-Western backgrounds has an non-significant, positive relation with night-time incidence. Finally, the best (SVM) model that was found given age-adjusted night-time rates (shown in Table 5.12) yields rather different predictor attributes compared to those found given standard night-time rates. As we cannot directly compare weights obtained by means of a non-linear SVM, we can only interpret the signs of the predictors. We observe a positive relation between the incidence and urbanisation, the proportion of men, and the proportion with a Western background. In addition, we find the proportion with a Turkish background to be negatively related to age-adjusted night-time incidence.

Time of day	*Incidence rate*	*Accuracy*	*Attribute*	*OR (95% CI)*	*p-value*	*p < 0.05?*
Full	*Standard*	77.2%	Average household size	0.58 (0.44, 0.77)	0.000	Yes
			Low income receivers	1.16 (1.06, 1.27)	0.002	Yes
			Inhabitants of at least 65 years	1.18 (1.04, 1.34)	0.009	Yes
			Male inhabitants	1.45 (1.06, 1.97)	0.019	Yes
			Married	1.10 (0.97, 1.24)	0.126	
			Social protection (labour disability)	0.92 (0.83, 1.03)	0.148	
			Other non-Western backgrounds	1.02 (0.87, 1.20)	0.765	
Full	*Adjusted*	68.4%	Low income receivers	1.12 (1.02, 1.22)	0.012	Yes
			Average household size	0.82 (0.90, 0.74)	0.029	Yes
			Western background	1.11 (0.59, 2.10)	0.177	
			Urbanisation	1.17 (0.76, 1.80)	0.264	
			Social protection (WW)	1.22 (0.97, 1.55)	0.491	
			Average property value	1.13 (0.89, 1.44)	0.526	
Night	*Standard*	65.0%	Average household size	0.65 (0.54, 0.78)	0.000	Yes
			Married	1.13 (1.05, 1.22)	0.001	Yes
			Western background	0.93 (0.83, 1.05)	0.264	
			Other non-Western backgrounds	1.04 (0.91, 1.19)	0.537	

Table 5.11. *Odds ratios for each configuration in North Holland (2006-16) in which LR reported highest accuracy.*

Time of day	*Incidence rate*	*Accuracy*	*Attribute*	*Sign*	*Weight*
Night	*Adjusted*	66.6%	Urbanisation	Positive	5.72
			Male inhabitants	Positive	15.01
			Western background	Positive	6.60
			Turkish background	Negative	-14.25

Table 5.12. *Results for the only configuration in North Holland (2006-16) in which the SVM with a radial basis function reported the highest accuracy.*

Twente

The best LR models based on the districts in Twente are shown in Table 5.13. The singular significant predictor for standard rates given the entire time of day is the proportion of at least 65 years of age. Non-significant predictors are average household size, the proportion of male inhabitants, and the degree of urbanisation. Given age-adjusted rates, we find recipients of labour disability acts as well as the average property value as significant predictors. The proportion of married inhabitants and the proportion of other non-Western backgrounds are found to have a non-significant, negative relation with the target. Furthermore, the proportion of married inhabitants as well as male inhabitants are significantly related to standard night-time incidence rates. The proportion of elderly, the average household size, the proportion of inhabitants with a Turkish background as well as those with other non-Western backgrounds are found to be non-significantly related with standard night-time rates. Finally, Table 5.14 contains the attributes included in the best SVM model, based on age-adjusted night-time incidence rates. The degree of urbanisation and the average property value are positively related, whereas negative associations with the target attribute were found for the proportion of low income receivers, the proportion with other non-Western backgrounds, and recipients of WWB and labour disability benefits.

Time of day	Incidence rate	Accuracy	Attribute	OR (95% CI)	p-value	p < 0.05?
Full	*Standard*	71.1%	Inhabitants of at least 65 years	1.38 (1.15, 1.66)	<0.001	Yes
			Average household size	1.26 (0.96, 1.65)	0.097	
			Urbanisation	0.73 (0.34, 1.57)	0.419	
			Male inhabitants	1.09 (0.72, 1.64)	0.677	
Full	*Adjusted*	66.7%	Labour disability acts	1.42 (1.08, 1.87)	0.011	Yes
			Average property value	1.17 (1.03, 1.33)	0.017	Yes
			Married	0.90 (0.79, 1.02)	0.089	
			Other non-Western backgrounds	0.91 (0.72, 1.17)	0.470	
Night	*Standard*	65.6%	Married	0.80 (0.67, 0.96)	0.018	Yes
			Male inhabitants	0.64 (0.42, 1.00)	0.048	Yes
			Average household size	1.28 (0.97, 1.69)	0.080	
			Inhabitants of at least 65 years	1.13 (0.95, 1.35)	0.159	
			Turkish background	0.86 (0.70, 1.06)	0.164	
			Other non-Western backgrounds	0.86 (0.66, 1.13)	0.276	

Table 5.13. *Odds ratios for each configuration in Twente in which LR reported the highest accuracy.*

Time of day	Incidence rate	Accuracy	Attribute	Sign	Weight
Night	*Adjusted*	63.3%	Urbanisation	Positive	1365.65
			Low income receivers	Negative	-116.49
			Average property value	Positive	921.81
			Social protection (WWB)	Negative	-1241.83
			Social protection (labour disability)	Negative	-1165.74
			Other non-Western backgrounds	Negative	-1027.25

Table 5.14. *Results for the configuration in Twente in which SVM reported the highest accuracy.*

5.5 Resuscitation efforts

Table 5.15 shows the fraction of urban and rural districts found in North Holland and Twente. Table 5.16 contains the results of the LR analyses of bystander-initiated cardiopulmonary resuscitation (BCPR), defibrillator usage and AED deployment. The spatial analysis on AED deployment comprises of 163 out of 246 districts in North Holland and all 90 districts in Twente.

	NH (n = 246)	TW (n = 90)
Urban districts, %	55	37
Rural districts, %	45	63

Abbreviations: n = number of districts, NH = North Holland, TW = Twente.

Table 5.15. *Percentage of urban and rural districts in both study regions.*

In North Holland, we find a significantly higher proportion of applied AEDs (OR 1.02; 95% CI 1.00, 1.04) across districts that belong to the 50% with high OHCA incidence rates in the region, compared to the remaining 50% of districts with low rates. These disparities are attenuated when we separately examine urban districts, but remain significant across rural districts (OR 1.03; 95% CI 1.01, 1.05). No significant differences are found in terms of BPCR rates or AED deployment.

In Twente, we find significantly higher values for BCPR rates (OR 1.09; 95% CI 1.01, 1.18), AED usage (OR 1.15; 95% CI 1.04, 1.27), and the number of allocated AEDs per 10,000 population (OR 1.15; 95% CI 1.02, 1.30) across urban districts with relatively high OHCA incidence rates, compared to those with low rates. This relation is not observed when for rural districts. Although we do find disparities in terms of AED usage in both regions, no differences are found in the-proportion of ambulance defibrillator usage between districts with high and low incidence rates.

Region	*Attribute*	*Urbanisation*	*OR (95% CI)*	*p-value*	*p < 0.05?*
North Holland	Bystander CPR rate (%)	*All*[a]	1.01 (0.99, 1.03)	0.332	
		Urban	1.01 (0.98, 1.04)	0.582	
		Rural	1.01 (0.98, 1.04)	0.490	
	AED usage (%)	*All*	1.02 (1.00, 1.04)	0.023	Yes
		Urban	1.01 (0.98, 1.03)	0.638	
		Rural	1.03 (1.01, 1.05)	0.012	Yes
	mDFB usage (%)	*All*	0.98 (0.95, 1.01)	0.248	
		Urban	1.00 (0.95, 1.05)	0.934	
		Rural	0.96 (0.92, 1.01)	0.119	
	Registered AEDs[b,c] (n)	*All*	1.01 (0.99, 1.04)	0.191	
		Urban	1.09 (1.00, 1.19)	0.051	
		Rural	1.01 (0.99, 1.04)	0.348	
Twente	Bystander CPR rate (%)	*All*	1.02 (0.99, 1.04)	0.148	
		Urban	1.09 (1.01, 1.18)	0.026	Yes
		Rural	1.00 (0.97, 1.03)	0.940	
	AED usage (%)	*All*	1.01 (0.99, 1.02)	0.558	
		Urban	1.15 (1.04, 1.27)	0.005	Yes
		Rural	0.99 (0.97, 1.01)	0.377	
	mDFB usage (%)	*All*	0.99 (0.97, 1.02)	0.704	
		Urban	0.99 (0.90, 1.09)	0.863	
		Rural	1.00 (0.97, 1.03)	0.834	
	Registered AEDs[b,c] (n)	*All*	1.02 (1.00, 1.04)	0.079	
		Urban	1.15 (1.02, 1.30)	0.018	Yes
		Rural	1.01 (0.99, 1.04)	0.208	

CPR = cardiopulmonary resuscitation, AED = automated external defibrillator, mDFB = ambulance defibrillator.
[a]*All urbanisation degrees are considered.* [b]*Analysis is limited to 163 districts.* [c]*Per 10,000 population.*

Table 5.16. *Odds ratios of resuscitation efforts in North Holland (246 districts, 2006-16) and Twente (90 districts, 2010-16).*

Figures 5.8-5.13 show the values for BCPR rates, AED usage, and the number of AEDs per 10,000 population across the considered districts. Results are roughly similar between our regions. BCPR rates vary between 47 and 100% in North Holland, and from 40 to 100% in Twente. We observe considerable differences in the number of deployed AEDs, ranging from zero to 74.8 AEDs per 10,0000 population in North Holland, and from 4.5 to 130.4 AEDs per 10,0000 inhabitants in Twente. In North Holland, we observe seven 'high-risk' districts without any registered public access AED. By contrast, an AED is deployed in each district in Twente. The proportion of applied AEDs however also differs greatly. Nevertheless, an AED has been used at least once in 121 out of 123 high-risk areas in North Holland, and in 44 out of 45 high-risk areas in Twente.

Finally, Appendix P states the ten districts with the lowest observed values for BCPR, AED usage, and the number of deployed AEDs per capita, out of the 50% of districts in the region with relatively high incidence rates. Moreover, five out of ten districts in both study regions that correspond with the lowest BCPR rates, also correspond with the lowest proportions of AED usage. Moreover, no registered public access AEDs are located in two North Holland districts for which very high OHCA incidence rates are observed. In Twente, two districts are even listed as having the lowest values for all three considered resuscitation efforts. Finally, the three areas in Twente with fewest allocated AEDs per capita, are all located within the same municipality.

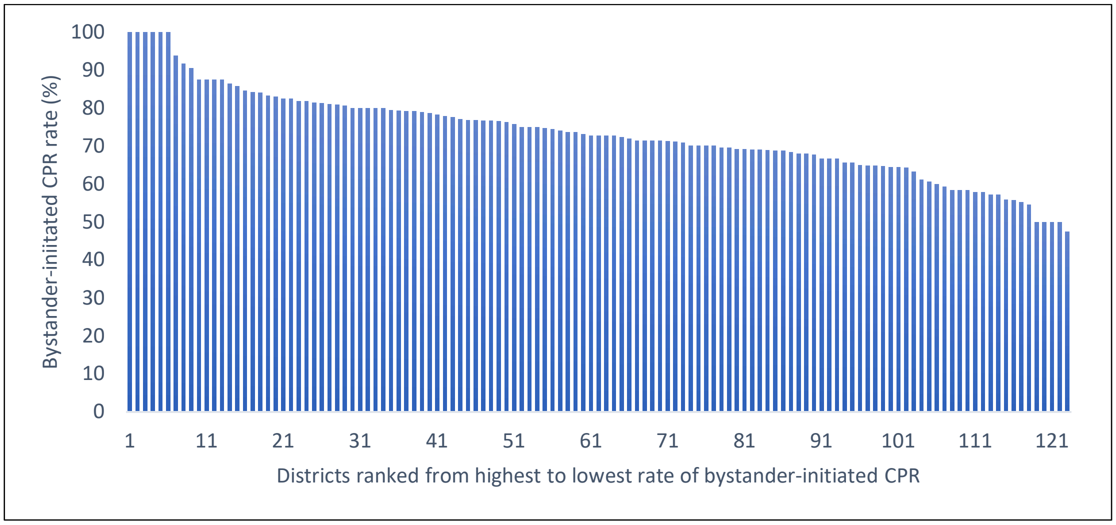


Figure 5.8. *Bystander-initiated CPR rates across the 50% of districts in North Holland (123 districts, 2006-16) with relatively high OHCA incidence rates (entire day, average annual standard OHCA incidence rate > 38.7).*

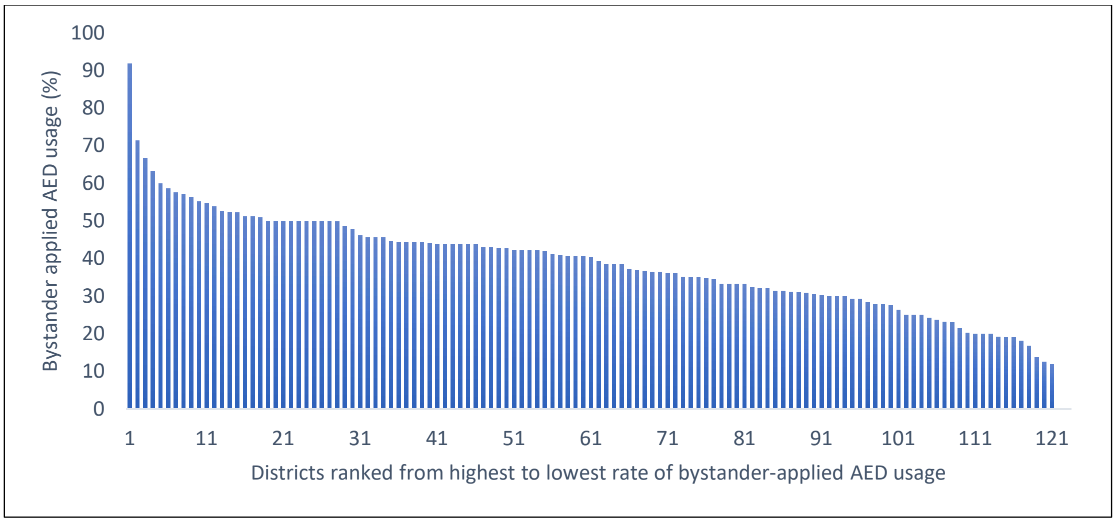


Figure 5.9. *Bystander-applied AED usage across the 50% of districts in North Holland (123 districts, 2006-16) with relatively high OHCA incidence rates (entire day, average annual standard OHCA incidence rate > 38.7).*

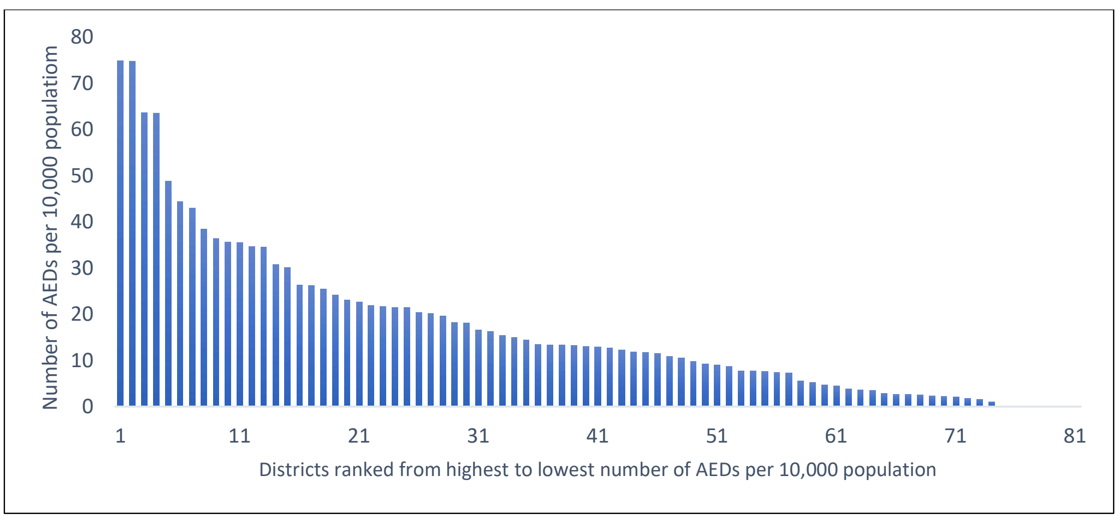


Figure 5.10. *Bystander-applied AED usage across the 50% of districts in North Holland (81 districts, 2006-16) with relatively high OHCA incidence rates (entire day, average annual standard OHCA incidence rate > 38.7).*

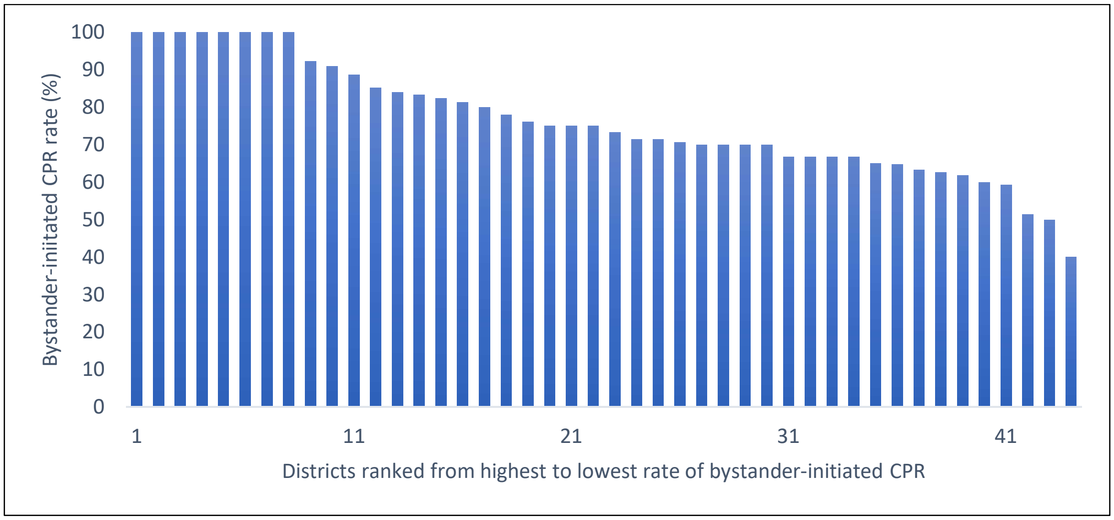


Figure 5.11. *Bystander-initiated CPR rates across the 50% of districts in Twente (44 districts, 2010-16) with relatively high OHCA incidence rates (entire day, average annual standard OHCA incidence rate > 40.6).*

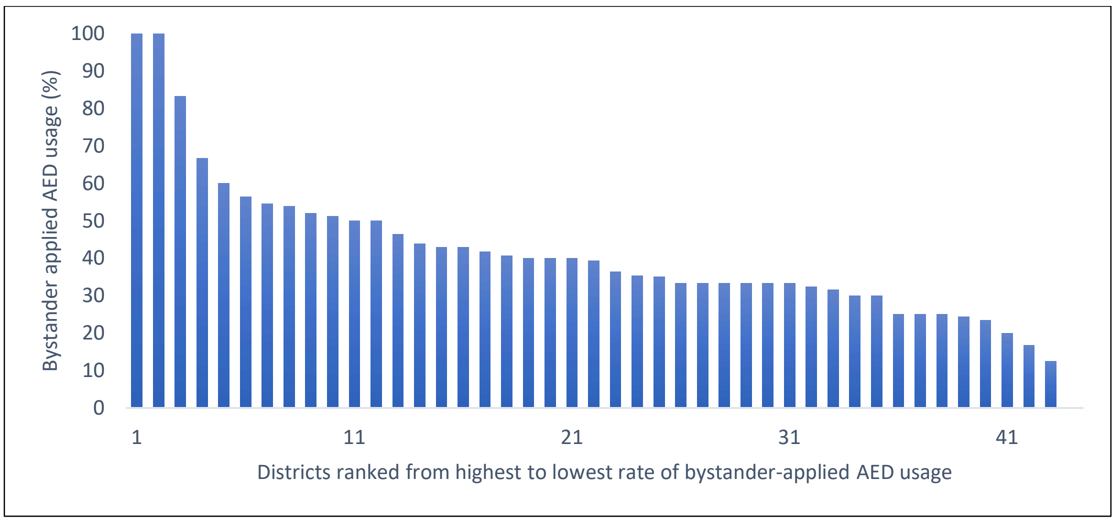


Figure 5.12. *Bystander-applied AED usage across the 50% of districts in Twente (44 districts, 2010-16) with relatively high OHCA incidence rates (entire day, average annual standard OHCA incidence rate > 40.6).*

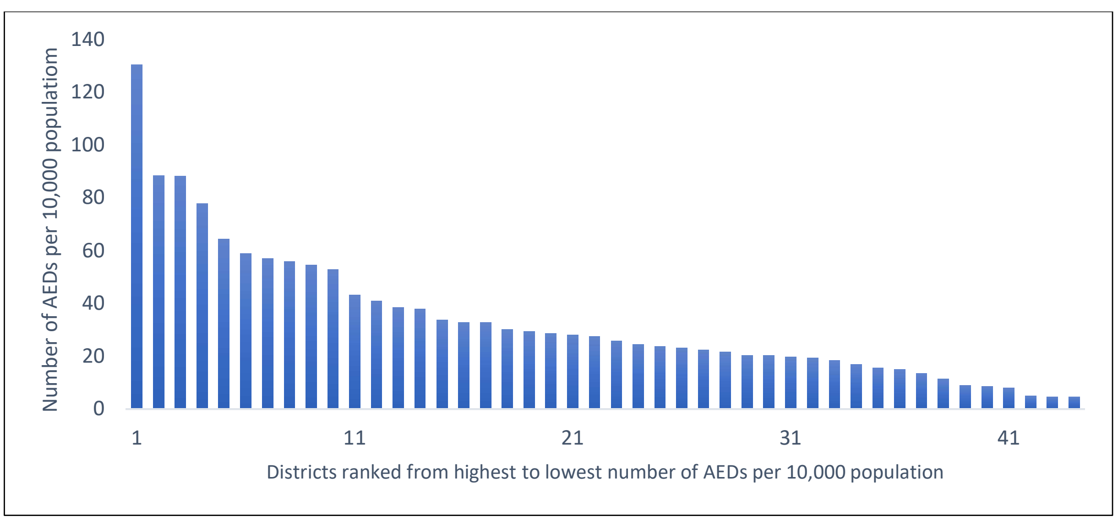


Figure 5.13. *Bystander-applied AED usage across the 50% of districts in Twente (44 districts, 2010-16) with relatively high OHCA incidence rates (entire day, average annual standard OHCA incidence rate > 40.6).*

CHAPTER 6 – DISCUSSION AND CONCLUSION

6.1 Discussion

Although aspects of the approach in this study have been examined separately, i.e. geographic risk distribution (e.g. Lerner et al., 2005; Sasson, Cudnik, et al., 2012), spatio-temporal variation (Demirtas et al., 2015; Sasson et al., 2010), the relation between OHCA incidence and 1) socio-economic factors (e.g. Reinier et al., 2011), 2) ethnicity (Straney et al., 2016), and 3) resuscitation efforts: disparities in CPR provision (e.g. Fosbøl et al., 2014; Raun et al., 2013), AED usage, and AED deployment (Moon et al., 2015), this study is the first to combine all aspects into one integrated approach. The intent of this approach is to gain a better understanding of disparities in OHCA incidence rates in terms of spatial variation, spatio-temporal variation, and underlying associations of spatial variation, which may in turn serve as input to improve the utility of public health efforts. This is done systematically by 1) identifying the high- and low-risk areas for OHCA, 2) determining whether these areas remain high- or low-risk over time, 3) finding underlying associations that may explain this phenomenon, and 4) exploring the current state of resuscitation efforts in these areas to identify the most beneficial opportunities for short-term improvement, by focusing on communities at greatest risk. The emphasis of this analysis is on determining the underlying associations between OHCA incidence and demographic, ethnicity and socio-economic factors, through the use of machine learning (ML) algorithms.

In both study regions, ARREST gathers data on all occurred EMS-treated OHCAs. Only 492 (≈4,0%) out of 12,455 eligible OHCAs were excluded after geocoding. We therefore believe that the excluded OHCAs does not significantly affect the results found in this study. Nevertheless, we recommend ARREST to regularly geocode future OHCA cases to increase the probability that each case can be used in future geographic studies. Next to that, we analysed two specific regions in the Netherlands, and results may not be generalizable to communities in other Dutch regions, or those in other countries. In addition, results for the region of North Holland are more likely to yield statistically significant results. This is due to the fact that the region comprises a larger set of districts, which in turn could result in more narrow 95% confidence intervals.

Geographic distributions of OHCA

We successfully identified areas with high and low OHCA incidence rates in both study regions. This is done for OHCAs that occurred during the entire day, or those that took place during night-time only. In both regions, over three out of four districts received the same classification (i.e., high or low OHCA incidence) for both rates. This suggests that a vast majority of high-risk areas during night-time can also be considered as having a high risk during the entire day. Even smaller differences in classifications are found between both standard and age-adjusted rates. We chose districts (in Dutch: 'wijken') as the most suitable level, as we considered the neighbourhood ('buurt') level too detailed to model relatively 'rare' OHCA events. In fact, only 34 percent of neighbourhoods[2] in North Holland had, on average, at least one cardiac arrest patient, compared to 74 percent of districts[3]. No cases were reported in about 10% of neighbourhoods in Twente[4],

---

[2] 293 out of 861 neighbourhoods, after applying structural reforms and filters as described in Section 4.7.

[3] 182 out of 246 districts.

[4] 39 out of 374 neighbourhoods, after applying structural reforms and filters.

versus only one of 90 districts. Although it is obvious that OHCA risk does not follow any predefined geographic borders, we believe that aggregating at least seven years of cardiac arrest data within those borders yields reliable OHCA hot spots. District boundaries may even give more meaningful and comprehensive results for public health programs, compared to outcomes of cluster methods that do not include the use of geographic structures. Furthermore, districts with low OHCA incidence may already have implemented successful community-based awareness programs. This may especially be the case for districts with low OHCA incidence, high rates of CPR provision, as well as a relatively high number of deployed AEDs. Further research is needed to examine the effect of these programs and the characteristics of those districts in which implementation has proven to be successful.

Spatio-temporal variation

We examined the overall temporal variation among the set of districts throughout the entire study period. Although we did not find significantly higher rates for any of the included years, we did observe significant variance for each residual random effect. This suggests that significant, unexplained variation within districts exist. Thus, as the number of OHCAs can differ, high-risk areas do not have to correspond with similar incidence rates from year to year. Further research is needed to examine underlying reasons for this. However, we may intuitively be able to partially explain this phenomenon as inherent to the geographic level of detail at which we examine relatively rare events. Hence, at a larger level (e.g. at a municipality level), we expect to find less temporal variation. However, a larger geographic level in turn implicates targeted public health efforts. Existing variation in OHCA incidence within districts over time also has implications for public health efforts, as solely considering spatial disparities seems to be insufficient. At the same time, this justifies our approach to include a multi-year study period when analysing underlying associations at this geographic level of detail.

The GLMM approach allows us to directly measure OHCA counts for spatial and temporal variation, while still including the annual population reported in each district. In addition, this approach does not violate any normality assumptions. Sample sizes in the considered study regions are relatively small, as we focus on the districts with an annual average of least two cardiac arrests. Therefore, estimates found by the model may differ when more districts are included. At the same time, we find this a feasible approach, as we consider the measurement of spatio-temporal variation to be especially important in those districts with the highest overall occurrence of cardiac arrests. Next to that, we correct the information criterion to the small sample size to still reliably measure the goodness of fit of our models. Compared to other agreement measures, mathematics behind GLMMs are relatively complex. As opposed to ICCs, the discriminating power of GLMMs does not depend on the variability in measurement results between subjects. The presence of between-subject variance and absence of within-subject variance would result in perfect fit of the model. In this case, the model classifies the random effects as redundant terms. Hence, performance of GLMMs does not depend on the homogeneity of the study population.. Furthermore, results of this study cannot be directly compared with those that include a different number of raters. More research is needed to explain the presence of unexplained variation found in this study by means of the significant residual random effects. One approach is suggested by Straney et al. (2016). The authors examine disparities to explain spatial variation in Victoria, Australia by including socio-demographic and cardiovascular risk factors into a mixed-effects Poisson regression model. The importance of including cardiovascular risk factors to explain spatial disparities is emphasized by a report of Graham et al. (2015).

In case all districts in the study region are included in the mixed model, a zero-inflated GLMM may provide a better fit. The same applies when assessing variation among annual measurements of night-time OHCAs with a considerable amount of zero values. The application of a non-zero inflated GLMM in this case is likely to produce biased parameter estimates and standard errors (Cameron & Trivedi, 2013; Zuur et al., 2009). However, the number of software packages that supports this more specialized version is more limited in comparison with the standard version of the model. The evaluation of more sophisticated models to assess spatio-temporal variation is recommended for future research.

Socio-economic factors

Among our considered characteristics, we considered several socio-economic attributes. The main advantage of including a wide range of spatial characteristics into a machine learning approach allows us to efficiently identify the most important community-level measures that are associated with OHCA incidence rates. Among the most important included measures of SES, income and occupation, we find income and not occupation to be significantly associated with increased odds of districts in North Holland having high OHCA incidence. Interestingly, the proportion of low income receivers seems to give comparable odds ratios for both standard (OR 1.16; 95% CI 1.06, 1.27) and age-adjusted rates (OR 1.12; 95% CI 1.02, 1.22). Thus, for each additional percentage point of receivers of low income that live in the area, at the cost of receivers of moderate to high income, the odds of the district having high OHCA incidence rates are expected to increase by 1.16 or 1.12 times (depending on the setting). The findings are in line with the results from Mirowsky et al. (2017) and Reinier et al. (2006). In Twente, we initially found only basic demographics to be significantly associated with OHCA incidence. These demographics correspond with the proportion of elderly, men, and married inhabitants. No significant socio-economic predictors were observed without adjusting for age structure. Through the use of age-adjusted rates in separate analyses, we observed the proportion of recipients of labour disability acts and the average property value to be significantly associated with OHCA incidence.

Several non-significant predictors were included in each of the models discussed in Section 5.4.2. It is likely that this phenomenon demonstrates the complexity of associating OHCA incidence rates exclusively with community-level measures. Results of this study suggest that, despite of several significant and obvious relations found, the problem of explaining disparities in incidence rates is in fact multifactorial. Additionally, differences in results between both regions may be explained by the possibility that validity of SES indicators differs between settings (Clark et al., 2009).

Ideally, we would have included education levels, median income and median property values. Data on these attributes was unfortunately not available. Average property value (WOZ) may not be a good substitute of the latter, due to significant bias in average income levels. WOZ may be biased towards a small proportion of properties that correspond with a considerably high property value. This presence of this possible bias was avoided for (average) income levels, as we were able to include the receivers of low income instead. Moreover, further research is needed to assess whether disparities in OHCA are associated with the proportion of receivers of low income in other regions.

As discussed in our literature review, the effect of low SES on health is mediated by material, behavioural and psychosocial factors. Hence, tackling inequalities in SES among districts is a multifactorial, culturally sensitive and costly task. Moreover, its effectiveness can only be measured on the long-term. Alternatively, reducing disparities in health without considering one's socio-economic position will seldom be adequate. Individuals with low SES may be well aware of behaviour-related risk factors for health. Hence, behavioural change may only be successful when other mediators such as poor psychosocial health, monetary problems as well as social norms within communities are addressed (Mackenbach, 2006). Due to the complexity of tackling these disparities, differences can be made on the short-term by optimizing public health efforts.

We used publicly available, regional data. Unfortunately, education level data is not publicly available at a district level. Nevertheless, we indirectly accounted for education level, as being an important determinant of both income and occupation (Van Oyen et al., 2011), which were both considered in our study. Since several prior studies show that education is related to OHCA incidence (Reinier et al., 2006; Straney et al., 2016), it is worthwhile to include education level data in future studies. As opposed to publicly available information, education data may be provided by Statistics Netherlands' Microdata Services for research purposes only ("Microdata: Conducting your own research," n.d.). In addition, regional incidence rates may be adjusted by using the average age of the population, rather than using a predefined age structure. Nevertheless, we believe that this alternative method to adjust for age differences between districts still gives reliable estimates. Two prior studies, facing a similar problem, also corrected OHCA incidence rate by making use of proportions of inhabitants of at least 65 years of age (Reinier et al., 2011; Straney et al., 2016).

We emphasize that the relation between OHCA incidence and socio-economic factors is explored at a community level, rather than at an individual level. Therefore, the observational nature of examining these factors implies that any observed relation is not causal, but associations only. The same accounts for the relations found between OHCA incidence and resuscitation efforts. This is due to the fact that correlation implies association, but not causation (Altman & Krzywinski, 2015). Moreover, we did not include SES scores as defined by the Netherlands Institute for Social Research (SCP). A main intent of this study was to go beyond the scope of a single predefined SES measure and identify which specific socio-economic factors contribute to disparities in OHCA incidence. The legitimacy of using community-level indicators to represent information at an individual level in the process is well established in literature. In addition, this approach permits use of a more sophisticated contextual analysis, compared to the use of individual-level indicators (Krieger, 1992). As a consequence, our study is expected to include victims of OHCAs that were not present in the same district as their residence at the time of the cardiac arrest. However, Reinier et al. (2006) show that 47% out of 693 patients in Portland, Oregon that suffered OHCA outside their residence, was still situated in the same census tract as their residence at the time of the arrest. Together with victims that had a cardiac arrest at home, a total of 84% of patients was still situated in the same census tract as their residence. Although we were unable to compare these findings with our own study, we assume that a vast majority of cardiac arrest occurred within the same district as the patient’s residence.

Ethnicity

This study is the first to examine disparities in OHCA incidence by including both western and non-western migrant groups. We include the four main non-western migrant groups in the Netherlands. Results of this study show that groups of each considered ethnic background were only marginally included in our models. Attributes that were included, were not found to be significantly associated with the odds of districts having high incidence. Based on that, we do not find the presence of certain ethnic backgrounds within communities to be associated with incidence of OHCA. These results are in line with the most similar study by Bos et al. (2005), taking into account substantial differences between the considered target attributes. The authors did not find large inequalities in cardiovascular mortality rates between Dutch areas with high and low SES among inhabitants with a Turkish, Moroccan, Antillean/Aruban, and Surinamese background. Regardless of the study results, we believe that public education campaigns should still be culturally sensitive and designed specifically for the particular community to be as effective as possible, as was emphasized by Sasson et al. (2013).

Since Dutch citizens are more likely to live in high SES districts, inhabitants of non-western backgrounds may make less use of the health care system. As mentioned earlier, inequalities among health care use may be caused by fees, language barriers, unfamiliarity with the health care system, or by obstacles in the administrative process (Kraler & Reichel, 2010; Padilla & Miguel, 2009; Rafnsson & Bhopal, 2008). A study by Groeneveld et al. (2003) examine Medicare beneficiaries throughout the United States, and find that among the elderly, whites are associated with improved OHCA outcomes. However, whites were also more likely to undergo coronary angioplasty or receive an implantable cardioverter defibrillator (ICD), possibly explaining the identified association among elderly survivors. Nevertheless, we conclude that possible (similar) disparities did not significantly influence our results, as we do not observe any significant associations between OHCA incidence rates and ethnicity backgrounds.

Statistics Netherlands’ definition of having a ‘western’ background is broad and includes backgrounds from countries all over the world (e.g. Japan, Indonesia and the United States). Combining these populations prior to analysis most likely hides considerable within-group heterogeneity (Bhopal, 2004). This categorisation diminishes our ability to examine the association of ethnic minorities with the probability of living in high-risk areas, as the underlying cultural differences are unclear. In fact, inequalities in health status have been reported among European countries (Mackenbach et al., 2008). A more detailed use of ethnic minorities with western backgrounds was preferred but not possible due to a lack of data. Nevertheless, these groups are unlikely to contribute to poor cardiovascular health, given the observed positive correlation between district-level SES and the proportion of western minorities (Knol, 2012). Fortunately, we were able to break down and separately analyse five ethnic groups with non-western backgrounds. These include the four main non-western minorities living in the Netherlands, namely those with backgrounds from the Netherlands Antilles/Aruba, Morocco, Suriname, and Turkey. As discussed in this study, cardiovascular health between these groups is reported to differ considerably (Kunst et al., 2011). Despite these reported differences, our results do not support disparities in incidence among inhabitants with different ethnicity backgrounds.

Grouping inhabitants based on his or her background, does not include the generation he or she belongs to. Significant differences in SES are likely to be observed between the first and second generation of migrants. In 2010, approximately half of the population with a non-western background was born in the Netherlands. This second generation performs better in education, and a higher individual-level SES is observed (Statistics Netherlands, 2010). Therefore, the relation between OHCA incidence and ethnic background may be stronger among first generation of migrants in comparison with subsequent generations with a similar cultural background.

Finally, the ethnic background of inhabitants may have increasing importance in medical studies in the near future. This is due to a population that tends to become more diverse in the next years, a trend observed throughout all of Europe. As Rechel et al. (2013) point out, an expected shortage of about one million jobs in the health sector by 2020 results in an increasing demand of low-qualified jobs, such as the provision of basic domestic care for the society's elderly people. Migrants that would fill those jobs are likely to live in communities with low SES. We recommend future studies to cardiovascular disease (CVD) incidence and SES to include both ethnic background as well as socio-economic factors. We recommend the ARREST registry to obtain the ethnic background and generation of OHCA victims, to be used in future studies.

Predictive modelling

In this study, we used predictive modelling to find relations between OHCA incidence and socio-economic factors as well as ethnicity backgrounds. We subdivide OHCA incidence rates into two distinct classes (i.e., high and low risk). We compared the predictive accuracy of four methods that have shown promising performance in health and cardiovascular risk prediction studies. These methods are Logistic Regression (LR), Artificial Neural Network (ANN), Support Vector Machine (SVM) with a radial basis function (RBF) and Adaptive Boosting (AdaBoost). Of these methods, LR is commonly used as the binary classification technique. The structurally higher predictive accuracy shown by LR in this study, compared to ANN, SVM and AdaBoost implies that a linear separation of districts that belong to different classes may be caused by underlying linear relations between the target and the considered predictor attributes. Moreover, LR does not need extensive data to train parameters or to avoid overfitting. Especially for ANNs, the small number of subjects may have led to improper estimates of parameters. We minimized the number of parameters by constructing a network with only one layer and a maximum of six neurons. Nevertheless, the amount of data may still have been insufficient to optimize the network's parameters, leading to overfitting and perhaps even to a high misclassification rate of training instances. In turn, low performance of the SVM may be caused by using the radial basis function, which is commonly used to explore the existence of complex classification patterns. Nevertheless, the ability of the SVM to reduce overfitting by establishing a maximum margin makes the algorithm powerful even when the number of subjects is limited.

LR also shows superior predictive performance in a similar study by Austin et al. (2013), and highest performance (after ANNs) in a comparable study by Weng et al. (2017). In line with the former, the reader should not conclude that the application of LR will result in superior predictive accuracy in other settings, solely based on the results in this study.

We chose our four algorithms based on their distinct classification capabilities and promising performance in prior health- and cardiovascular risk studies. It is possible that the choice of other strong classification algorithms would have resulted in higher classification accuracy. For example, the Random Forest algorithm shows promising performance in numerous studies. However, applying this algorithm would have resulted in losing interpretability of the constructed models. The same applies to the use of recent automated machine learning platforms such as Auto-WEKA (Thornton et al., 2012). The latter also comes at the cost of high computation complexity. The use of other methods that do not come with considerably losing interpretability is recommended to be used in future studies. This includes regularized (penalized) LR and polynomial SVM.

An increase in attributes could also improve results by correctly classifying districts of which risk levels are currently predicted wrong by either a majority or even by all considered ML techniques. We recommend future studies to include regional data on cardiovascular risk factors (e.g. Straney et al., 2016) as well as environmental factors (e.g. Rosenthal et al., 2008).

We did not examine whether the Genetic Algorithm (GA) outperforms greedy procedures such as forward selection or backward elimination, as this was out of the scope of this thesis. We instead carefully selected and extensively applied different attribute subsets found by folds of the GA, or by combining the attributes that are present in at least half of the folds. Moreover, the predictive accuracy of a learning algorithm is associated with the predictive power of the GA that uses the same algorithm as the base learner. This phenomenon can intuitively be explained, as the GA searches for the optimal set based on the base learner. Although not seen in similar studies, we chose to take into account this bias by implementing an additional GA with a different classifier. By combining the generally high predictive performance of GAs with two algorithms that are associated with high predictive ability in comparable studies, we believe that we correctly and extensively explored the attribute input space within our automated dimensionality reduction. Based on preliminary results, we limited the possible hyperparameter settings of the GA-SVM to reduce computational complexity. Results may differ when more hyperparameter values are evaluated. Ideally all parameters are assessed, but this generally is not a feasible approach. Results of the automated dimensionality reduction phase imply that combining subsets of CV-folds structurally result in a loss of information. Hence, we recommend studies with similar study designs to re-evaluate attributes found in each of the applied CV-folds.

Regardless of the predictive power of any learning algorithm, an essential prerequisite of applying predictive modelling is the use of high quality data. We consider the quality of our data as one of the main strengths of our study. Regional used in this study was initially gathered by Statistics Netherlands, a governmental statistical institution, and a department of the Dutch Ministry of Economic Affairs. Data was made available for all districts that met Statistics Netherlands' eligibility criteria. Although excluding 3.9% and 7.2% of districts respectively, on average we exclude 0.0008% of the total population in North Holland and 0.002% in Twente.

We dichotomized incidence rates to simplify the problem. As a result, we did not consider relatively small differences among the approximate 50% of districts with high rates and the 50%

with low rates. Based on the results of this study, we recommend future studies in the ARREST region to apply supervised regression techniques. These techniques can then be used to explore whether relations found in this study will hold when exact incidence rates are considered. In addition, new relations may be found compared to limiting the problem to two classes.

Future studies are needed to obtain a better understanding of the interaction between the prevalence of cardiac arrests and both socio-economic factors as well as ethnicity backgrounds. We recommend ARREST and other OHCA registries to conduct a cohort study concerning cardiac arrest prediction for individual cases with appropriate ML algorithms, similar to the successful prediction of myocardial infarction by Weng et al. (2017). This may give new insights in the importance of risk factors complementary to this study, to ensure that every patient receives the care he or she needs, regardless of age, gender, socio-economic status or ethnicity. A better understanding may especially be obtained after combining socio-economic factors and ethnic backgrounds with cardiovascular risk factors. To perform these studies at an individual level in the future, we recommend OHCA registries to expand current data collection with information of the patient's socio-economic status and ethnicity, within the national legal framework. Finally, although outside the scope of this study, we recommend ARREST to track recent developments in the area of automatic classification of ECG-based cardiac rhythms with ML (e.g. Ali et al., 2016).

Bystander CPR

We found significant disparities in terms of bystander-initiated cardiopulmonary resuscitation (BCPR) rates between urban districts in Twente. This implies that the probability of receiving CPR is higher in districts with higher incidence rates. The absence of a similar relation in North Holland may be partially explained due to overall BCPR rates that are on average already relatively high compared to other studies, as discussed in Section 2.4.

Disparities in CPR provision may also exist between areas of high and low SES. Prior studies have shown that cardiac arrest victims in neighbourhoods with high SES are more likely to receive BCPR than neighbourhoods with low SES (Sasson, Magid, et al., 2012; Vaillancourt et al., 2008). Sasson, Magid, et al. (2012) find that communities with lower income, lower education level, and higher percentage of African-American inhabitants are associated with lower prevalence of BCPR. Future studies are needed to examine whether this relation exists within districts in the Netherlands. In addition, high SES districts may already have implemented community-based awareness programs that increased overall knowledge of CPR. As CPR training programmes may not always be affordable for the most vulnerable populations (Dahan et al., 2017), disparities in knowledge of CPR may have affected results. Finally, another important factor which we did not take account, is the performance and willingness to attempt bystander CPR. The former depends on the presence of a person who is prepared to perform CPR (Sasson, Magid, et al., 2012), whereas the latter may include one's perception of the area's safety and quality of living, the fear or hurting the patient in the process, the physical inability to attempt CPR, and the perceived risk of transmitting infectious diseases (Bradley & Rea, 2011; Kanstad et al., 2011; Swor et al., 2006; Vaillancourt et al., 2008). The effect of district may thus be very important. Future research is required to investigate differences in performance and barriers in CPR provision among the considered districts in this study. More complex approaches to examine disparities in CPR provision could also be considered. For example, Straney et al. (2015) compute shrunken estimates to ensure the use of reliable estimates in less populated areas.

AED deployment

We find a significant positive relation between OHCA incidence rates and the number of deployed AEDs per 10,000 population among urban districts in Twente. Results are found non-significant when separately examining rural districts. No significant relation is observed in North Holland.

It is likely that the AED registry did not include all defibrillators. At the same time, only registered AEDs can be integrated into an EMS system. In addition, no upper bound exists for the number of AEDs to deploy in any area, as long as the fraction of area covered by these defibrillators is increased. It is clear that in practice, achieving 100% coverage is generally not possible. A trade-off therefore needs to be made where to locate the AEDs. In each study region, we list the districts with fewest AEDs per capita as a starting point for improvement of current AED allocation.

In agreement with Moon et al. (2015), we find this level of spatial analysis of both OHCA incidence rates and the number of AEDs per capita to be the first step in targeted AED deployment. A study by Chan et al. (2016), performed in Toronto, Canada, use mathematical modelling to optimize AED locations at a more granular level. Although outside the scope of this thesis, we see great potential to mathematically optimize AED deployment in our study regions as well. Finally, the number of deployed AEDs does not give us information on actual coverage, availability or the state of defibrillators. This is in line with the study design, as being a first step in the investigation of resuscitation efforts. Thus, further research is recommended to explore these relations.

Quality of life

The enhancement of survival after OHCA is not the only criterion to optimize public health campaigns that aim to improve resuscitation. Although outside the scope of our thesis, we briefly discuss the quality of life of survivors. A recent study by Boyce-van der Wal et al. (2015) shows that survivors with problematic cognitive outcomes after OHCA are associated with decreased quality of life as well as reduced levels of autonomy and participation in society. Fortunately, prior literature in North Holland conclude that most OHCA survivors generally have an acceptable quality of life after the event. Normal functioning and cognition among a majority of survivors is observed 6-12 months after the OHCA (Beesems et al., 2014). In addition, most survivors of older age show favourable neurologic outcome (Beesems et al., 2015). This emphasizes the importance of early, high-quality resuscitation to enhance outcome of OHCA. Indeed, Van Alem et al. (2004) find that early access, early CPR as well as early defibrillation improve cognitive, physical and psychosocial functioning in the same study region. Next to the ARREST regions, similar results of quality of life among survivors were found in a study by Smith et al. (2014).

6.2 Conclusion

Spatial disparities in standard and age-adjusted OHCA incidence are observed in the Dutch regions of North Holland and Twente. A vast majority of areas with high incidence given the entire day also have high incidence during night-time. The proportion of income receivers that belong to the national 40% with the lowest income, the proportion of at least 65 years of age, and the proportion of male inhabitants are significantly associated with increased odds of a district in

North Holland having high OHCA incidence. Average household size is significantly associated with similar, decreased odds. In addition, the proportion of receivers of low income and average household size are significant determinants of age-adjusted incidence rates. The average household size and the proportion of married inhabitants are significantly associated with the odds of a district in North Holland having high OHCA incidence during night-time. In Twente, we find the proportion of elderly, and male as well as married inhabitants to be significantly associated with the odds of a district having high OHCA incidence. After adjusting for age structure, we observe the proportion of recipients of labour disability acts and the average property value to be significantly and positively associated with OHCA incidence. In both regions, ethnicity was not found to be significantly associated with high OHCA incidence. LR structurally provides higher predictive accuracy than ANNs, SVMs and AdaBoost in predicting areas with high and low incidence, based on socio-economic attributes and ethnicity backgrounds.

Public health campaigns should target districts with a high proportion of receivers of low income, recipients of labour disability acts, elderly population, male inhabitants, a low proportion of married inhabitants as well as those with small household sizes. Next to that, public health campaigns that aim to improve resuscitation in areas with known high OHCA incidence are advised to prioritize on those resuscitation efforts for which unfavoured disparities are found: the number of AEDs per capita, AED usage, and, to a smaller extent, BCPR rates. These campaigns should take into account the presence of significant temporal variation of OHCA incidence rates within districts.

FUNDING

There was no financial support from any organisation for the submitted work.

CONFLICT OF INTEREST STATEMENT

The author declares no conflict of interest.

BIBLIOGRAPHY

Afbakening generaties met migratieachtergrond. (n.d.). Retrieved October 30, 2017, from https://www.cbs.nl/nl-nl/achtergrond/2016/47/afbakening-generaties-met-migratieachtergrond

Agudelo-Suárez, A. A., Ronda-Pérez, E., & Benavides, F. G. (2011). Occupational health. In B. Rechel, P. Mladovsky, W. Devillé, B. Rijks, R. Petrova-Benedict, & M. McKee (Eds.), *Migration and health in the European Union* (pp. 155–168). Maidenhead: Open University Press.

Ahn, K. O., Shin, S. Do, Hwang, S. S., Oh, J., Kawachi, I., Kim, Y. T., … Hong, S. O. (2011). Association between deprivation status at community level and outcomes from out-of-hospital cardiac arrest: A nationwide observational study. *Resuscitation*, *82*(3), 270–276.

Ahonen, E. Q., Benavides, F. G., & Benach, J. (2007). Immigrant populations, work and health—a systematic literature review. *Scandinavian Journal of Work, Environment & Health*, *33*(2), 96–104.

Altman, D. G., & Bland, J. M. (1983). Measurement in Medicine: The Analysis of Method Comparison Studies. *The Statistician*, *32*(3), 307–317.

Altman, N., & Krzywinski, M. (2015). Points of Significance: Association, correlation and causation. *Nature Methods*, *12*(10), 899–900.

Anderson, R. T., Sorlie, P., Backlund, E., Johnson, N., & Kaplan, G. A. (1997). Mortality Effects of Community Socioeconomic Status. *Epidemiology*, *8*(1), 42–47.

Austin, P. C., Tu, J. V, Ho, J. E., Levy, D., & Lee, D. S. (2013). Using methods from the data mining and machine learning literature for disease classification and prediction: A case study examining classification of heart failure sub-types. *Journal of Clinical Epidemiology*, *66*(4), 398–407.

Baker, E. H. (2014). Socioeconomic Status, Definition. In *The Wiley Blackwell Encyclopedia of Health, Illness, Behavior, and Society* (pp. 2210–2214).

Banerjee, A. (2017). Predicting the future of cardiovascular risk prediction. *Heart*, *103*(12), 891.2-892.

Bardai, A., Berdowski, J., Van Der Werf, C., Blom, M. T., Ceelen, M., Van Langen, I. M., … Tan, H. L. (2011). Incidence, Causes, and Outcomes of Ot-of-Hospital Cardiac Arrest in Children: A Comprehensive, Prospective, Population-Based Study in the Netherlands. *Journal of the American College of Cardiology*, *57*(18), 1822–1828.

Beck, B., Bray, J. E., Smith, K., Walker, T., Grantham, H., Hein, C., … Finn, J. (2016). Description of the ambulance services participating in the Aus-ROC Australian and New Zealand out-of-hospital cardiac arrest Epistry. *EMA - Emergency Medicine Australasia*, *28*(6), 673–683.

Becker, L. B., Han, B. H., Meyer, P. M., Wright, F. A., Rhodes, K. V., Smith, D. W., & Barrett, J. (1993). Racial differences in the incidence of cardiac arrest and subsequent survival. *New England Journal of Medicine*, *329*(9), 600–606.

Beesems, S. G., Blom, M. T., Van der Pas, M. H. A., Hulleman, M., Van de Glind, E. M. M., van Munster, B. C., … Koster, R. W. (2015). Comorbidity and favorable neurologic outcome after out-of-hospital cardiac arrest in patients of 70 years and older. *Resuscitation*, *94*, 33–39.

Beesems, S. G., Wittebrood, K. M., de Haan, R. J., & Koster, R. W. (2014). Cognitive function and quality of life after successful resuscitation from cardiac arrest. *Resuscitation*, *85*(9), 1269–1274.

Beesems, S. G., Zijlstra, J. A., Stieglis, R., & Koster, R. W. (2012). Reanimatie buiten het ziekenhuis in Noord-Holland en Twente: resultaten ARREST-onderzoek over 2006-2011. *Hart- En Vaatziekten in Nederland*, 101–112.

Berchtold, A. (2016). Test–retest: Agreement or reliability? *Methodological Innovations*, *9*, 205979911667287.

Berdowski, J., Beekhuis, F., Zwinderman, A. H., Tijssen, J. G. P., & Koster, R. W. (2009). Importance of the first link: Description and recognition of an out-of-hospital cardiac arrest in an emergency call. *Circulation*, *119*(15), 2096–2102.

Berdowski, J., Blom, M. T., Bardai, A., Tan, H. L., Tijssen, J. G. P., & Koster, R. W. (2011). Impact of onsite or dispatched automated external defibrillator use on survival after out-of-hospital cardiac arrest. *Circulation*, *124*(20), 2225–2232.

Berdowski, J., Kuiper, M. J., Dijkgraaf, M. G. W., Tijssen, J. G. P., & Koster, R. W. (2010). Survival and health care costs until hospital discharge of patients treated with onsite, dispatched or without automated external defibrillator. *Resuscitation*, *81*(8), 962–967.

Bezzina, C. R., Pazoki, R., Bardai, A., Marsman, R. F., de Jong, J. S. S. G., Blom, M. T., … Wilde, A. A. M. (2010). Genome-wide association study identifies a susceptibility locus at 21q21 for ventricular fibrillation in acute myocardial infarction. *Nature Genetics*, *42*(8), 688–691.

Bhopal, R. (2004). Glossary of terms relating to ethnicity and race: for reflection and debate. *Journal of Epidemiology & Community Health*, *58*(6), 441–445.

Bijl, R., Boelhouwer, J., Cloïn, M., & Pommer, E. (2011). *De sociale staat van Nederland 2011*. The Hague: The Netherlands Institute for Social Research.

Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*.

Bland, J. M., & Altman, D. G. (1986). Statistical Methods for Assessing Agreement Between Two Methods of Clinical Measurement. *The Lancet*, *327*(8476), 307–310.

Bland, J. M., & Altman, D. G. (1999). Measuring agreement in method comparison studies. *Statistical Methods in Medical Research*, *8*, 135–160.

Blom, M. T., Beesems, S. G., Homma, P. C. M., Zijlstra, J. A., Hulleman, M., Van Hoeijen, D. A., … Koster, R. W. (2014). Improved survival after out-of-hospital cardiac arrest and use of automated external defibrillators. *Circulation*, *130*(21), 1868–1875.

Blom, M. T., van Hoeijen, D. A., Bardai, A., Berdowski, J., Souverein, P. C., De Bruin, M. L., …

Tan, H. L. (2014). Genetic, clinical and pharmacological determinants of out-of-hospital cardiac arrest: rationale and outline of the AmsteRdam Resuscitation Studies (ARREST) registry. *Open Heart*, *1*(1), e000112.

Blom, M. T., Warnier, M. J., Bardai, A., Berdowski, J., Koster, R. W., Souverein, P. C., … Tan, H. L. (2013). Reduced in-hospital survival rates of out-of-hospital cardiac arrest victims with obstructive pulmonary disease. *Resuscitation*, *84*(5), 569–574.

Blommesteijn, M., & Mallee, L. (2009). *Minimum income scheme: Work and Social Assistance Act*.

Bolker, B. M., Brooks, M. E., Clark, C. J., Geange, S. W., Poulsen, J. R., Stevens, M. H. H., & White, J. S. S. (2009). Generalized linear mixed models: a practical guide for ecology and evolution. *Trends in Ecology and Evolution*, *24*(3), 127–135.

Bos, V., Kunst, A. E., Garssen, J., & Mackenbach, J. P. (2005). Socioeconomic inequalities in mortality within ethnic groups in the Netherlands, 1995–2000. *Journal of Epidemiology & Community Health*, *59*, 329–335.

Boyce-van der Wal, L. W., Volker, W. G., Vliet Vlieland, T. P. M., Van den Heuvel, D. M. J., Van Exel, H., & Goossens, P. H. (2015). Cognitive problems in patients in a cardiac rehabilitation program after an out-of-hospital cardiac arrest. *Resuscitation*, *93*, 63–68.

Bradley, S. M., & Rea, T. D. (2011). Improving bystander cardiopulmonary resuscitation. *Current Opinion in Critical Care*, *17*(3), 219–224.

Braveman, P., Cubbin, C., Egerter, S., Chideya, S., Marchi, K. S., Metzler, M., & Posner, S. (2005). Socioeconomic Status in Health Research: One Size Does Not Fit All. *JAMA*, *294*(22), 2879–2888.

Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification and Regression Trees*. CRC press.

Brookoff, D., Kellermann, A. L., Hackman, B. B., Somes, G., & Dobyns, P. (1994). Do Blacks Get Bystander Cardiopulmonary Resuscitation as Often as Whites? *Annals of Emergency Medicine*, *24*(6), 1147–1150.

Brooks, S. C., Schmicker, R. H., Rea, T. D., Aufderheide, T. P., Davis, D. P., Morrison, L. J., … Dorian, P. (2010). Out-of-hospital cardiac arrest frequency and survival: Evidence for temporal variability. *Resuscitation*, *81*(2), 175–181.

Buick, J. E., Ray, J. G., Kiss, A., & Morrison, L. J. (2016). The association between neighborhood effects and out-of-hospital cardiac arrest outcomes. *Resuscitation*, *103*, 14–19.

Cameron, A. C., & Trivedi, P. K. (2013). *Regression Analysis of Count Data*. *Cambridge University Press* (Vol. 53). Cambridge, United Kingdom.

Caruana, R., & Niculescu-Mizil, A. (2006). An empirical comparison of supervised learning algorithms. *ICML'06 Proceedings of the 23rd International Conference on Machine Learning*, 161–168.

Chan, T. C. Y., Demirtas, D., & Kwon, R. H. (2016). Optimizing the Deployment of Public

Access Defibrillators. *Management Science*, *62*(12), 3617–3635.

Chicote, B., Irusta, U., Alcaraz, R., Rieta, J. J., Aramendi, E., Isasi, I., … Ibarguren, K. (2016). Application of Entropy-Based Features to Predict Defibrillation Outcome in Cardiac Arrest. *Entropy*, *18*(9), 1–17.

Chu, K., Swor, R., Jackson, R., Domeier, R., Sadler, E., Basse, E., … Gitlin, J. (1998). Race and Survival After Out-of-Hospital Cardiac Arrest in a Suburban Community. *Annals of Emergency Medicine*, *31*(4), 478–482.

Chugh, S. S., Reinier, K., Teodorescu, C., Evanado, A., Kehr, E., Al Samara, M., … Jui, J. (2008). Epidemiology of Sudden Cardiac Death: Clinical and Research Implications. *Progress in Cardiovascular Diseases*, *51*(3), 213–228.

Clark, A. M., DesMeules, M., Luo, W., Duncan, A. S., & Wielgosz, A. (2009). Socioeconomic status and cardiovascular disease: risks and implications for care. *Nature Reviews Cardiology*, *6*(11), 712–722.

Clarke, S. O., Schellenbaum, G. D., & Rea, T. D. (2005). Socioeconomic status and survival from out-of-hospital cardiac arrest. *Academic Emergency Medicine*, *12*(10), 941–947.

Costa-Santos, C., Antunes, L., Souto, A., & Bernardes, J. (2010). Assessment of Disagreement: A New Information-Based Approach. *Annals of Epidemiology*, *20*(7), 554–561.

Costa-Santos, C., Costa Pereira, A., & Bernardes, J. (2005). Agreement studies in obstetrics and gynaecology: Inappropriateness, controversies and consequences. *BJOG: An International Journal of Obstetrics and Gynaecology*, *112*(5), 667–669.

Cowie, M. R., Fahrenbruch, C. E., Cobb, L. A., & Hallstrom, A. P. (1993). Out-of-Hospital Cardiac Arrest: Racial Differences in Outcome in Seattle. *American Journal of Public Health*, *83*(7), 955–959.

Cox, D. R. (1958). The Regression Analysis of Binary Sequences, *20*(2), 215–242.

Cristianini, N., & Shawe-Taylor, J. (2000). *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*. Cambridge: Cambridge University Press.

Cummins, R. O., Ornato, J. P., Thies, W. H., Pepe, P. E., Billi, J. E., Seidel, J., … Membrino, G. E. (1991). Improving survival from sudden cardiac arrest: The chain of survival'' concept. A statement for health professionals from the advanced cardiac life support subcommittee and the emergency cardiac care committee, American Heart Association. *Circulation*.

Dahan, B., Jabre, P., Karam, N., Misslin, R., Tafflet, M., Bougouin, W., … Jouven, X. (2017). Impact of neighbourhood socio-economic status on bystander cardiopulmonary resuscitation in Paris. *Resuscitation*, *110*, 107–113.

Das, R., Turkoglu, I., & Sengur, A. (2009). Effective diagnosis of heart disease through neural networks ensembles. *Expert Systems with Applications*, *36*(4), 7675–7680.

Davies, A., Basten, A., & Frattini, C. (2009). *Migration: a social determinant of the health of* *migrants, background paper developed within the framework of the IOM project "Assisting migrants and communities (AMAC): analysis of social determinants of health and health inequalities." Eurohealth* (Vol. 16).

De Jong, K. (1988). Learning with Genetic Algorithms: An Overview. *Machine Learning*, *3*(2), 121–138.

De Vet, H. C. W. (1998). Observer reliability and agreement. In *Armitage P, Colton T, editors, Encyclopedia biostatistica, Vol 4. John, Chichester: Wiley & Sons, Ltd.* (pp. 3123–8).

De Vet, H. C. W., Terwee, C. B., Knol, D. L., & Bouter, L. M. (2006). When to use agreement versus reliability measures. *Journal of Clinical Epidemiology*, *59*(10), 1033–1039.

De Vreede-Swagemakers, J. J. M., Gorgels, A. P. M., Dubois-Arbouw, W. I., Dalstra, J., Daemen, M. J. A. P., Van Ree, J. W., … Wellens, H. J. (1998). Circumstances and causes of out-of-hospital cardiac arrest in sudden death survivors. *Heart*, *79*, 356–361.

DeCarlo Santiago, C., Wadsworth, M. E., & Stump, J. (2011). Socioeconomic status, neighborhood disadvantage, and poverty-related stress: Prospective effects on psychological syndromes among diverse low-income families. *Journal of Economic Psychology*, *32*(2), 218–230.

Definitions. (n.d.). Retrieved October 15, 2017, from https://www.cbs.nl/en-gb/our-services/methods/definitions

Demirtas, D., Brooks, S. C., Morrison, L. J., & Chan, T. C. (2015). Abstract 15003: Spatiotemporal Stability of Public Cardiac Arrests. *Circulation*, *132*(Suppl 3), A15003 LP-A15003.

Diez-Roux, A. V, Nieto, F. J., Muntaner, C., Tyroler, H. A., Comstock, G. W., Shahar, E., … Szklo, M. (1997). Neighbourhood Environments and Coronary Heart Disease: A Multilevel Analysis. *American Journal of Epidemiology*, *146*(1), 48–63.

Diez Roux, A. V, Nieto, F. J., Caulfield, L., Tyroler, H. A., Watson, R. L., & Szklo, M. (1999). Neighbourhood differences in diet: the Atherosclerosis Risk in Communities (ARIC) Study. *Journal of Epidemiology & Community Health*, *53*, 55–63.

Disability schemes. (n.d.). Retrieved October 15, 2017, from https://www.cbs.nl/en-gb/artikelen/nieuws/2008/30/social-benefits-reduced-by-over-300-thousand/disability-schemes

Dominic, V., Gupta, D., & Khare, S. (2015). An Effective Performance Analysis of Machine Learning Techniques for Cardiovascular Disease. *Appl Med Inform*, *36*(1), 23–32.

Duda, R. O., Hart, P. E., & Stork, D. G. (2012). *Pattern classification*. John Wiley & Sons.

Eisenberg, M., Bergner, L., & Hallstrom, A. (1979). Paramedic programs and out-of-hospital cardiac arrest: I. Factors associated with successful resuscitation. *American Journal of Public Health*, *69*(1), 30–38.

Feero, S., Hedges, J. R., & Stevens, P. (1995). Demographics of Cardiac Arrest: Association with Residence in a Low-income Area. *Academic Emergency Medicine*, *2*(1), 11–16.

Fisher, R. A. (1921). On the “probable error” of a coefficient of correlation deduced from a small sample. *Metron 1*.

Fleiss, J. L. (1981). *Statistical Methods for Rates and Proportions*. (J. W. & Sons, Ed.). New York.

Forrest, S. (1993). Genetic Algorithms: Principles of Natural Selection Applied to Computation. *Source: Science, New Series*, *261*(13), 872–878.

Fosbøl, E. L., Dupre, M. E., Strauss, B., Swanson, D. R., Myers, B., McNally, B. F., … Granger, C. B. (2014). Association of neighborhood characteristics with incidence of out-of-hospital cardiac arrest and rates of bystander-initiated CPR: Implications for community-based education intervention. *Resuscitation*, *85*(11), 1512–1517.

Freund, Y., & Schapire, R. (1996). Experiments with a new boosting algorithm. *Machine Learning: Proceedings of the Thirteenth International Conference: San Francisco, California: Morgan Kauffman*.

Fröhlich, H., Chapelle, O., & Schölkopf, B. (2004). Feature Selection for Support Vector Machines Using Genetic Algorithms. *International Journal on Artificial Intelligence Tools*, *13*(4), 791–800.

Galea, S., Blaney, S., Nandi, A., Silverman, R., Vlahov, D., Foltin, G., … Richmond, N. (2007). Explaining Racial Disparities in Incidence of and Survival from Out-of-Hospital Cardiac Arrest. *American Journal of Epidemiology*, *166*(5), 534–543.

Gallo, L. C., & Matthews, K. A. (2003). Understanding the Association Between Socioeconomic Status and Physical Health: Do Negative Emotions Play a Role? *Psychological Bulletin*, *129*(1), 10–51.

Galobardes, B., Shaw, M., Lawlor, D. A., Lynch, J. W., & Smith, G. D. (2006). Indicators of socioeconomic position (part 1). *Journal of Epidemiology & Community Health*, *60*(1), 7–12.

Ghobrial, J., Heckbert, S. R., Bartz, T. M., Lovasi, G., Wallace, E., Lemaitre, R. N., … Sotoodehnia, N. (2016). Ethnic differences in sudden cardiac arrest resuscitation. *Heart*, *102*(17), 1363–1370.

Gillum, R. F. (1997). Sudden Cardiac Death in Hispanic Americans and African Americans. *American Journal of Public Health*, *87*(9), 1461–1466.

Goldberg, D. E., & Deb, K. (1991). A Comparative Analysis of Selection Schemes Used in Genetic Algorithms. *Foundations of Genetic Algorithms*, *1*, 69–93.

Goldstein, B. A., Navar, A. M., & Carter, R. E. (2016). Moving beyond regression techniques in cardiovascular risk prediction: applying machine learning to address analytic challenges. *European Heart Journal*, ehw302.

Gow, R. M., Barrowman, N. J., Lai, L., & Moher, D. (2008). A review of five cardiology journals found that observer variability of measured variables was infrequently reported. *Journal of Clinical Epidemiology*, *61*(4), 394–401.

Graham, R., McCoy, M. A., & Schultz, A. M. (Eds. . (2015). *Strategies to Improve Cardiac Arrest Survival: A Time to Act.* National Academies Press.

Gräsner, J. T., Herlitz, J., Koster, R. W., Rosell-Ortiz, F., Stamatakis, L., & Bossaert, L. (2011). Quality management in resuscitation - Towards a European Cardiac Arrest Registry (EuReCa). *Resuscitation*, *82*(8), 989–994.

Gräsner, J. T., Lefering, R., Koster, R. W., Masterson, S., Böttiger, B. W., Herlitz, J., … Bossaert, L. L. (2016). EuReCa ONE—27 Nations, ONE Europe, ONE Registry: A prospective one month analysis of out-of-hospital cardiac arrest outcomes in 27 countries in Europe. *Resuscitation*, *105*, 188–195.

Groeneveld, P. W., Heidenreich, P. A., & Garber, A. M. (2003). Racial disparity in cardiac procedures and mortality among long-term survivors of cardiac arrest. *Circulation*, *108*(3), 286–291.

Gu, C., & Ma, P. (2005). Generalized Nonparametric Mixed-Effect Models: Computation and Smoothing Parameter Selection. *Journal of Computational and Graphical Statistics*, *14*(2), 485–504.

Gushulak, B., Pace, P., & Weekers, J. (2010). Migration and health of migrants. In T. Koller (Ed.), *Poverty and social exclusion in the WHO European Region: health systems respond* (pp. 257–81). Copenhagen: WHO Regional Office for Europe.

Guyatt, G., Walter, S., & Norman, G. (1987). Measuring change over time: Assessing the usefulness of evaluative instruments. *Journal of Chronic Diseases*, *40*(2), 171–178.

Haan, M., Kaplan, G. A., & Camacho, T. (1987). Poverty and health: prospective evidence from the Alameda County Study. *American Journal of Epidemiology*, *125*(6), 989–998.

Hall, M. A. (1999). *Correlation-based Feature Selection for Machine Learning*.

Hallstrom, A. P., Boutin, P., Cobb, L., & Johnson, E. (1993). Socioeconomic Status and Prediction of Ventricular Fibrillation Survival. *American Journal of Public Health*, *83*(2), 245–248.

Hallstrom, A. P., Ornato, J. P., Weisfeldt, M., Travers, A., Christenson, J., McBurnie, M. A., … Proschan, M. (2004). Public-Access Defibrillation and Survival after Out-of-Hospital Cardiac Arrest. *New England Journal of Medicine*, *351*(7), 637–646.

Hansen, C. M., Wissenberg, M., Weeke, P., Ruwald, M. H., Lamberts, M., Lippert, F. K., … Folke, F. (2013). Automated External Defibrillators Inaccessible to More Than Half of Nearby Cardiac Arrests in Public Locations During Evening, Nighttime, and Weekends. *Circulation*, *128*, 2224–2231.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (2nd Editio). Springer.

Heck, Ronald H.; Thomas, Scott; Tabata, L. (2013). *Multilevel Modeling of Categorical Outcomes Using IBM SPSS*.

Hendrix, A., Vaartjes, I., Mosterd, A., Reitsma, J. B., Doevendans, P. A., Grobbee, D. E., & Bots, M. L. (2010). Regional differences in incidence of sudden cardiac death in the young. *Neth J Med*, *68*(6), 274–279.

Henriques, T., Antunes, L., Bernardes, J., Matias, M., Sato, D., & Costa-Santos, C. (2013). Information-based measure of disagreement for more than two observers: a useful tool to compare the degree of observer disagreement. *BMC Medical Research Methodology*, *13*(1), 47.

Herd, P., Goesling, B., & House, J. S. (2007). Socioeconomic Position and Health: The Differential Effects of Education versus Income on the Onset versus Progression of Health Problems, *48*, 223–238.

Holland, J. H. (1975). Adaptation in Natural and Artificial Systems. *Ann Arbor MI University of Michigan Press*, *Ann Arbor*, 183.

Howe, A., Escalona, O. J., Di Maio, R., Massot, B., Cromie, N. A., Darragh, K. M., … McEneaney, D. J. (2014). A support vector machine for predicting defibrillation outcomes from waveform metrics. *Resuscitation*, *85*(3), 343–349.

I-Kuei Lin, L. (1989). A Concordance Correlation Coefficient to Evaluate Reproducibility. *Biometrics*, *45*, 255–268.

Ingleby, D. (2009). *European Research on migration and health, background paper developed within the framework of the IOM project "Assisting migrants and communities (AMAC): analysis of social determinants of health and health inequalities."*

Jaglal, S. B., & Goel, V. (1994). Social inequity in risk of coronary artery disease in Ontario. *Canadian Journal of Cardiology*, *10*(4), 439–443.

Jones, M., Dobson, A., & O'Brian, S. (2011). A graphical method for assessing agreement with the mean between multiple observers using continuous measures. *International Journal of Epidemiology*, *40*(5), 1308–1313.

Kadi, I., Idri, A., & Fernandez-Aleman, J. L. (2017). Knowledge discovery in cardiology: A systematic literature review. *International Journal of Medical Informatics*, *97*, 12–32.

Kanstad, B. K., Nilsen, S. A., & Fredriksen, K. (2011). CPR knowledge and attitude to performing bystander CPR among secondary school students in Norway. *Resuscitation*, *82*(8), 1053–1059.

Kao, J.-H., Chan, T.-C., Lai, F., Lin, B.-C., Sun, W.-Z., Chang, K.-W., … Lin, J.-W. (2017). Spatial analysis and data mining techniques for identifying risk factors of Out-of-Hospital Cardiac Arrest. *International Journal of Information Management*, *37*(1), 1528–1538.

Kavakiotis, I., Tsave, O., Salifoglou, A., Maglaveras, N., Vlahavas, I., & Chouvarda, I. (2017). Machine Learning and Data Mining Methods in Diabetes Research. *Computational and Structural Biotechnology Journal*, *15*, 104–116.

Keogh, E., & Mueen, A. (2011). *Encyclopedia of Machine Learning*. Springer.

Keuning, H. (2012). De invloed van gemeentelijke herindelingen op de regionale statistieken van het CBS. *Geo-Info*, *9*(1), 14–17.

Kincaid, C. (2005). *Guidelines for Selecting the Covariance Structure in Mixed Model Analysis*.

Proceedings of the Thirtieth Annual SAS Users Group International Conference (Vol. 198–30).

Kitamura, T., Iwami, T., Kawamura, T., Nagao, K., Tanaka, H., & Hiraide, A. (2010). Nationwide public-access defibrillation in Japan. *The New England Journal of Medicine*, *362*(11), 994–1004.

Knol, F. (2012). Statusontwikkeling van wijken in Nederland 1998-2010.

Kohavi, R. (1995). A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection. *IJCAI*, *14*, 1137–1143.

Kohavi, R., & John, G. H. (1997). Wrappers for feature subset selection. *Artificial Intelligence*, *97*(1–2), 273–324.

Kottner, J., Audige, L., Brorson, S., Donner, A., Gajewski, B. J., Hróbjartsson, A., … Streiner, D. L. (2011). Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed. *International Journal of Nursing Studies*, *48*(6), 661–671.

Kourou, K., Exarchos, T. P., Exarchos, K. P., Karamouzis, M. V., & Fotiadis, D. I. (2015). Machine learning applications in cancer prognosis and prediction. *Computational and Structural Biotechnology Journal*, *13*, 8–17.

Kraler, A., & Reichel, D. (2010). Statistics on migration, integration and discrimination in Europe. PROMINSTAT final report. Brussels: European Commission.

Krieger, N. (1992). Overcoming the absence of socioecnomic data in medical records: validation and application of census-based methodology. *American Journal of Public Health*, *82*(5), 703–710.

Krishnan, S., Cozier, Y. C., Rosenberg, L., & Palmer, J. R. (2010). Socioeconomic Status and Incidence of Type 2 Diabetes: Results From the Black Women's Health Study. *American Journal of Epidemiology*, *171*(5), 564–570.

Krizmaric, M., Verlic, M., Stiglic, G., Grmec, S., & Kokol, P. (2009). Intelligent analysis in predicting outcome of out-of-hospital cardiac arrest. *Computer Methods and Programs in Biomedicine*, *95*(2 SUPPL.), 22–32.

Kucharska-Newton, A. M., Harald, K., Rosamond, W. D., Rose, K. M., Rea, T. D., & Salomaa, V. (2011). Socioeconomic Indicators and the Risk of Acute Coronary Heart Disease Events: Comparison of Population-Based Data from the United States and Finland. *Annals of Epidemiology*, *21*, 572–579.

Kunst, A. E. (2010). Een overzicht van sociaal-economische verschillen in gezondheid in Europa. In *De gezonde levensloop: Een geschenk van vele generaties* (pp. 39–56). Amsterdam: Amsterdam University Press.

Kunst, A. E., Stronks, K., & Agyemang, C. (2011). Non-communicable diseases. In B. Rechel, P. Mladovsky, W. Devillé, B. Rijks, R. Petrova-Benedict, & M. McKee (Eds.), *Migration and Health in the European Union* (pp. 101–120). Maidenhead: Open University Press.

Lantz, P. M., Golberstein, E., House, J. S., & Morenoff, J. (2010). Socioeconomic and behavioral risk factors for mortality in a national 19-year prospective study of U.S.

adults. *Social Science & Medicine*, *70*(10), 1558–1566.

Larsen, M. P., Eisenberg, M. S., Cummins, R. O., & Hallstrom, A. P. (1993). Predicting survival from out-of-hospital cardiac arrest: A graphic model. *Annals of Emergency Medicine*, *22*(11), 1652–1658.

Leray, P., & Gallinari, P. (1999). Feature selection with Neural Networks. *Behaviormetrika*, *26*(1), 145–166.

Lerner, E. B., Fairbanks, R. J., & Shah, M. N. (2005). Identification of out-of-hospital cardiac arrest clusters using a geographic information system. *Academic Emergency Medicine*, *12*(1), 81–84.

Libbrecht, M. W., & Noble, W. S. (2015). Machine learning applications in genetics and genomics. *Nature Reviews Genetics*, *16*(6), 321–332.

Linden, A., Krensky, P., Hare, J., Idoine, C. J., Sicular, S., & Vashisth, S. (2017). *Magic Quadrant for Data Science Platforms*.

Liu, T., Lin, Z., Ong, M. E. H., Koh, Z. X., Pek, P. P., Yeo, Y. K., … Liu, N. (2015). Manifold ranking based scoring system with its application to cardiac arrest prediction: A retrospective study in emergency department patients. *Computers in Biology and Medicine*, *67*, 74–82.

Lorant, V., Deliège, D., Eaton, W., Robert, A., Philippot, P., & Ansseau, M. (2003). Socioeconomic Inequalities in Depression: A Meta-Analysis. *American Journal of Epidemiology*, *157*(2), 98–112.

Ma, M. H. M., Chiang, W. C., Ko, P. C. I., Huang, J. C. C., Lin, C. H., Wang, H. C., … Lin, F. Y. (2007). Outcomes from out-of-hospital cardiac arrest in Metropolitan Taipei: Does an advanced life support service make a difference? *Resuscitation*, *74*(3), 461–469.

Macintyre, S. (1997). The Black Report and beyond: what are the issues? *Social Science and Medicine*, *44*(6), 723–745.

Mackenbach, J. P. (2006). *Health Inequalities: Europe in Profile*. European Commission.

Mackenbach, J. P., Stirbu, I., Roskam, A.-J. R., Schaap, M. M., Menvielle, G., Leinsalu, M., & Kunst, A. E. (2008). Socioeconomic Inequalities in Health in 22 European Countries. *New England Journal of Medicine*, *358*(23), 2468–2481.

Mackenbach, J. P., Van de Mheen, H., & Stronks, K. (1994). A prospective cohort study investigating the explanation of socio-economic inequalities in health in the Netherlands. *Social Science and Medicine*, *38*(2), 299–308.

Majer, I. M., Nusselder, W. J., Mackenbach, J. P., & Kunst, A. E. (2011). Socioeconomic inequalities in life and health expectancies around official retirement age in 10 Western-European countries. *Journal of Epidemiology & Community Health*, *65*(11), 972–979.

McNally, B., Stokes, A., Crouch, A., & Kellermann, A. L. (2009). CARES: Cardiac Arrest Registry to Enhance Survival. *Annals of Emergency Medicine*, *54*(5).

Mensah, G. A., Mokdad, A. H., Ford, E. S., Greenlund, K. J., & Croft, J. B. (2005). State of Disparities in Cardiovascular Health in the United States. *Circulation*, *111*, 1233–1241.

Microdata: Conducting your own research. (n.d.). Retrieved November 6, 2017, from https://www.cbs.nl/en-gb/our-services/customised-services-microdata/microdata-conducting-your-own-research

Microsoft. (n.d.). Microsoft Excel 2016.

Miller, B. L., & Goldberg, D. E. (1995). Genetic Algorithms, Tournament Selection, and the Effects of Noise. *Complex Systems*, *9*(3), 193–212.

Mirowsky, J. E., Devlin, R. B., Diaz-Sanchez, D., Cascio, W., Grabich, S. C., Haynes, C., … Neas, L. (2017). A novel approach for measuring residential socioeconomic factors associated with cardiovascular and metabolic health. *Journal of Exposure Science and Environmental Epidemiology*, *27*, 281–289.

Moon, S., Bobrow, B. J., Vadeboncoeur, T. F., Kortuem, W., Kisakye, M., Sasson, C., … Spaite, D. W. (2014). Disparities in bystander CPR provision and survival from out-of-hospital cardiac arrest according to neighborhood ethnicity. *American Journal of Emergency Medicine*, *32*(9), 1041–1045.

Moon, S., Vadeboncoeur, T. F., Kortuem, W., Kisakye, M., Karamooz, M., White, B., … Bobrow, B. J. (2015). Analysis of out-of-hospital cardiac arrest location and public access defibrillator placement in Metropolitan Phoenix, Arizona. *Resuscitation*, *89*(C), 43–49.

Moore, L. V., Diez Roux, A. V., Evenson, K. R., McGinn, A. P., & Brines, S. J. (2008). Availability of Recreational Resources in Minority and Low Socioeconomic Status Areas. *American Journal of Preventive Medicine*, *34*(1), 16–22.

Morland, K., Wing, S., & Diez Roux, A. (2002). The Contextual Effect of the Local Food Environment on Residents' Diets: The Atherosclerosis Risk in Communities Study. *American Journal of Public Health*, *92*(11), 1761–1767.

Morrison, L. J., Nichol, G., Rea, T. D., Christenson, J., Callaway, C. W., Stephens, S., … Newgard, C. (2008). Rationale, development and implementation of the Resuscitation Outcomes Consortium Epistry-Cardiac Arrest. *Resuscitation*, *78*(2), 161–169.

Mozaffarian, D., Benjamin, E. J., Go, A. S., Arnett, D. K., Blaha, M. J., Cushman, M., … Turner, M. B. (2016). Heart Disease and Stroke Statistics—2016 Update: A Report From the American Heart Association. *Circulation*, *133*(4), e38–e360.

Müller, R., & Büttner, P. (1994). A critical discussion of intraclass correlation coefficients. *Statistics in Medicine*, *13*(23–24), 2465–2476.

Nassel, A. F., Root, E. D., Haukoos, J. S., McVaney, K., Colwell, C., Robinson, J., … Sasson, C. (2014). Multiple cluster analysis for the identification of high-risk census tracts for out-of-hospital cardiac arrest (OHCA) in Denver, Colorado, *85*(12), 1667–1673.

Neison, F. G. P. (1844). On a Method Recently Proposed for Conducting Inquiries into the Comparative Sanatory Condition of Various Districts, with Illustrations, Derived from

Numerous Places in Great Britain at the Period of the Last Census. *Journal of the Statistical Society of London*, *7*(1), 40–68.

Nichol, G., Thomas, E., Callaway, C. C. W., Hedges, J., Powell, J. L., Aufderheide, T. P., … Stiell, I. (2008). Regional variation in out-of-hospital cardiac arrest incidence and outcome. *Journal of the American Medical Association*, *300*(12), 1423–1431.

Old age pension. (n.d.). Retrieved October 15, 2017, from https://www.cbs.nl/en-gb/artikelen/nieuws/2009/13/sharp-increase-in-partial-aow-benefits/old-age-pension

Ong, M. E. H., Ng, C. H. L., Goh, K., Liu, N., Koh, Z. X., Shahidah, N., … Lin, Z. (2012). Prediction of cardiac arrest in critically ill patients presenting to the emergency department using a machine learning score incorporating heart rate variability compared with the modified early warning score. *Crit Care*, *16*(3), R108.

Ong, M. E. H., Shin, S. Do, Tanaka, H., Ma, M. H. M., Khruekarnchana, P., Hisamuddin, N., … Khan, M. N. (2011). Pan-Asian Resuscitation Outcomes Study (PAROS): Rationale, methodology, and implementation. *Academic Emergency Medicine*, *18*(8), 890–897.

Ong, M. E. H., Tan, E. H., Yan, X., Anushia, P., Lim, S. H., Leong, B. S. H., … Anantharaman, V. (2008). An observational study describing the geographic-time distribution of cardiac arrests in Singapore: What is the utility of geographic information systems for planning public access defibrillation? (PADS Phase I). *Resuscitation*, *76*(3), 388–396.

Padilla, B., & Miguel, J. P. (2009). *Health and migration in the EU: building a shared vision for action*.

Perkins, G. D., Handley, A. J., Koster, R. W., Castrén, M., Smyth, M. A., Olasveengen, T., … Greif, R. (2015). European Resuscitation Council Guidelines for Resuscitation 2015. Section 2. Adult basic life support and automated external defibrillation. *Resuscitation*, *95*, 81–99.

Perkins, G. D., Jacobs, I. G., Nadkarni, V. M., Berg, R. A., Bhanji, F., Biarent, D., … Zideman, D. A. (2015). Cardiac Arrest and Cardiopulmonary Resuscitation Outcome Reports: Update of the Utstein Resuscitation Registry Templates for Out-of-Hospital Cardiac Arrest. *Resuscitation*, *96*, 328–340.

Pijls, R. W. M., Nelemans, P. J., Rahel, B. M., & Gorgels, A. P. M. (2016). A text message alert system for trained volunteers improves out-of-hospital cardiac arrest survival. *Resuscitation*, *105*, 182–187.

Rafnsson, S., & Bhopal, R. (2008). Migrant and ethnic health research: report on the European Public Health Association Conference 2007. *European Journal of Public Health*, *122*, 532–34.

Rajan, S., Wissenberg, M., Folke, F., Hansen, C. M., Lippert, F. K., Weeke, P., … Torp-Pedersen, C. (2015). Out-of-hospital cardiac arrests in children and adolescents: Incidences, outcomes, and household socioeconomic status. *Resuscitation*, *88*, 12–19.

Raun, L. H., Jefferson, L. S., Persse, D., & Ensor, K. B. (2013). Geospatial analysis for targeting out-of-hospital cardiac arrest intervention. *American Journal of Preventive Medicine*, *45*(2), 137–142.

Rechel, B., Mladovsky, P., Devill, W., Rijks, B., Petrova-Benedict, R., & McKee, M. (2011). Migration and health in the European Union: an introduction. In B. Rechel, P. Mladovsky, W. Devillé, B. Rijks, R. Petrova-Benedict, & M. McKee (Eds.), *Migration and Health in the European Union* (pp. 3–13). Maidenhead: Open University Press.

Rechel, B., Mladovsky, P., Ingleby, D., Mackenbach, J. P., & McKee, M. (2013). Migration and health in an increasingly diverse Europe. *The Lancet*, *381*(9873), 1235–1245.

Reinier, K., Stecker, E. C., Vickers, C., Gunson, K., Jui, J., & Chugh, S. S. (2006). Incidence of sudden cardiac arrest is higher in areas of low socioeconomic status: A prospective two year study in a large United States community. *Resuscitation*, *70*(2), 186–192.

Reinier, K., Thomas, E., Andrusiek, D. L., Aufderheide, T. P., Brooks, S. C., Callaway, C. W., … Chugh, S. S. (2011). Socioeconomic status and incidence of cardiac arrest. *Canadian Medical Association Journal*, *183*(15), 1705–1712.

Ro, Y. S., Hwang, S. S., Shin, S. Do, Han, D., Kang, S., Song, K. J., & Cho, S. Il. (2015). Presumed regional incidence rate of out-of-hospital cardiac arrest in Korea. *Journal of Korean Medical Science*, *30*(10), 1396–1404.

Root, E. D., Gonzales, L., Persse, D. E., Hinchey, P. R., McNally, B., & Sasson, C. (2013). A tale of two cities: The role of neighborhood socioeconomic status in spatial clustering of bystander CPR in Austin and Houston. *Resuscitation*, *84*(6), 752–759.

Rosenblatt, F. (1962). Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms. *Spartan*.

Rosenthal, F. S., Carney, J. P., & Olinger, M. L. (2008). Increased Risk of Out-of-Hospital Cardiac Arrest is Associated with Short-Term PM2.5 Exposure. *Epidemiology*, *19*(6), S91.

Roth, R., Stewart, R. D., Rogers, K., & Cannon, G. M. (1984). Out-of-hospital cardiac arrest: Factors associated with survival. *Annals of Emergency Medicine*, *13*(4), 237–243.

Rozanski, A., Blumenthal, J. A., Davidson, K. W., Saab, P. G., & Kubzansky, L. (2005). The epidemiology, pathophysiology, and management of psychosocial risk factors in cardiac practice: The emerging field of behavioral cardiology. *Journal of the American College of Cardiology*.

Sasson, C., Cudnik, M. T., Nassel, A., Semple, H., Magid, D. J., Sayre, M., … Warden, C. R. (2012). Identifying high-risk geographic areas for cardiac arrest using three methods for cluster analysis. *Academic Emergency Medicine*, *19*(2), 139–146.

Sasson, C., Hendrika, M., Abella, B. S., Berg, R. A., Bobrow, B. J., Chan, P. S., … Rea, T. D. (2013). Increasing Cardiopulmonary Resuscitation Provision in Communities With Low Bystander Cardiopulmonary Resuscitation Rates. *Circulation*, *127*, 1342–1350.

Sasson, C., Keirns, C. C., Smith, D., Sayre, M., Macy, M., Meurer, W., … Iwashyna, T. J. (2010). Small Area Variations in Out-of-Hospital Cardiac Arrest: Does the Neighbourhood Matter? *Annals of Internal Medicine*, *153*(1), 19–22.

Sasson, C., Magid, D. J., Chan, P., Root, E. D., McNally, B. F., Kellermann, A. L., & Haukoos, J.

S. (2012). Association of Neighborhood Characteristics with Bystander-Initiated CPR. *New England Journal of Medicine*, *367*(17), 1607–1615.

Savastano, S., De Servi, S., Oltrona Visconti, L., & Raimondi, M. (2015). The Pavia Cardiac Arrest REgistry - Pavia CARE. It's time for Cardiologists to be engaged in cardiac arrest. *International Journal of Cardiology*, *185*, 93–94.

Saydah, S. H., Imperatore, G., & Beckles, G. L. (2013). Socioeconomic status and mortality: Contribution of health care access and psychological distress among U.S. adults with diagnosed diabetes. *Diabetes Care*, *36*(1), 49–55.

Sayegh, A. J., Swor, R., Chu, K. H., Jackson, R., Gitlin, J., Domeier, R. M., … Fales, W. (1999). Does race or socioeconomic status predict adverse outcome after out of hospital cardiac arrest: a multi-center study. *Resuscitation*, *40*(3), 141–146.

Schaffer, C. (1994). A conservation law for generalization performance. *Proceedings of the 11th International Conference on Machine Learning*, 259–265.

Schenker, M. B. (2008). Work-related injuries among immigrants: a growing global health disparity. *Occupational and Environmental Medicine*, *65*(11), 717–8.

Schenker, M. B. (2010). A Global Perspective of Migration and Occupational Health. *American Journal of Industrial Medicine*, *53*(4), 329–337.

Schmidt, F. L., & Hunter, J. E. (1989). Interrater Reliability Coefficients Cannot Be Computed When Only One Stimulus Is Rated. *Journal of Applied Psychology*, *74*(2), 368–370.

Scholten, A. C., van Manen, J. G., van der Worp, W. E., IJzerman, M. J., & Doggen, C. J. M. (2011). Early cardiopulmonary resuscitation and use of Automated External Defibrillators by laypersons in out-of-hospital cardiac arrest using an SMS alert service. *Resuscitation*, *82*(10), 1273–1278.

Schootman, M., Andresen, E. M., Wolinsky, F. D., Malmstrom, T. K., Miller, J. P., Yan, Y., & Miller, D. K. (2007). The Effect of Adverse Housing and Neighborhood Conditions on the Development of Diabetes Mellitus among Middle-aged African Americans. *American Journal of Epidemiology*, *166*(4), 379–387.

Semple, H. M., Cudnik, M. T., Sayre, M., Keseg, D., Warden, C. R., & Sasson, C. (2013). Identification of high-risk communities for unattended out-of-hospital cardiac arrests using GIS. *Journal of Community Health*, *38*(2), 277–284.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, *86*(2), 420–428.

Smith, K., Andrew, E., Lijovic, M., Nehme, Z., & Bernard, S. (2014). Quality of life and functional outcomes 12 months after out-of-hospital cardiac arrest. *Circulation*, *131*(2), 174–181.

Søholm, H., Hassager, C., Lippert, F., Winther-Jensen, M., Hartvig-Thomsen, J., Friberg, H., … Kjaergaard, J. (2014). Factors Associated With Successful Resuscitation After Out-of-Hospital Cardiac Arrest and Temporal Trends in Survival and Comorbidity. *Annals of Emergency Medicine*, *65*(5), 1–9.

Soo, L., Huff, N., Gray, D., & Hampton, J. R. (2001). Geographical distribution of cardiac arrest in Nottinghamshire. *Resuscitation*, *48*(2), 137–147.

Statistics Netherlands. (2010). *Jaarrapport Integratie*. Den Haag/Heerlen.

Statistics Netherlands. (2016). *Toelichting Wijk- en Buurtkaart*.

Statistics Netherlands StatLine database. (n.d.). Retrieved June 18, 2017, from http://statline.cbs.nl/Statweb/

Straney, L. D., Bray, J. E., Beck, B., Bernard, S., Lijovic, M., & Smith, K. (2016). Are sociodemographic characteristics associated with spatial variation in the incidence of OHCA and bystander CPR rates? A population-based observational study in Victoria, Australia. *BMJ Open*, *6*, e012434.

Straney, L. D., Bray, J. E., Beck, B., Finn, J., Bernard, S., Dyson, K., … Smith, K. (2015). Regions of high out-of-hospital cardiac arrest incidence and low bystander CPR rates in Victoria, Australia. *PLoS ONE*, *10*(10), 1–14.

Stratford, P. W., & Goldsmith, C. H. (1997). Use of the standard error as a reliability index of interest: an applied example using elbow flexor strength data. *Physical Therapy*, *77*(7), 745–750.

Stringhini, S., Carmeli, C., Jokela, M., Avendaño, M., Muennig, P., Guida, F., … Kivimäki, M. (2017). Socioeconomic status and the 25 x 25 risk factors as determinants of premature mortality: a multicohort study and meta-analysis of 1.7 million men and women. *The Lancet*, 7–9.

Stringhini, S., Séverine, S., Shipley, M., Brunner, M., Nabi, H., Kivimaki, M., & Singh-Manoux. (2010). Association of Socioeconomic Position With Health Behaviors and Mortality. *JAMA*, *303*(12), 1159–1166.

Strömsöe, A., Svensson, L., Claesson, A., Lindkvist, J., Lundström, A., & Herlitz, J. (2011). Association between population density and reported incidence, characteristics and outcome after out-of-hospital cardiac arrest in Sweden. *Resuscitation*, *82*(10), 1307–1313.

Sun, C. L. F., Demirtas, D., Brooks, S. C., Morrison, L. J., & Chan, T. C. Y. (2016). Overcoming Spatial and Temporal Barriers to Public Access Defibrillators Via Optimization. *Journal of the American College of Cardiology*, *68*(8), 836–845.

Sundquist, K., Theobald, H., Yang, M., Li, X., Johansson, S. E., & Sundquist, J. (2006). Neighborhood violent crime and unemployment increase the risk of coronary heart disease: A multilevel study in an urban setting. *Social Science and Medicine*, *62*(8), 2061–2071.

Swor, R., Khan, I., Domeier, R., Honeycutt, L., Chu, K., & Compton, S. (2006). CPR Training and CPR Performance: Do CPR-trained Bystanders Perform CPR? *Academic Emergency Medicine*, *13*(6), 596–601.

Tableau Software. (n.d.). Tableau Desktop. Retrieved June 21, 2017, from https://www.tableau.com/products/desktop

Thornton, C., Hutter, F., Hoos, H. H., & Leyton-Brown, K. (2012). Auto-WEKA: Combined Selection and Hyperparameter Optimization of Classification Algorithms.

Tsai, Y.-S., Ko, P. C.-I., Huang, C.-Y., & Wen, T.-H. (2012). Optimizing locations for the installation of automated external defibrillators (AEDs) in urban public streets through the use of spatial and temporal weighting schemes. *Applied Geography*, *35*(1–2), 394–404.

Tu, J. V. (1996). Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. *Journal of Clinical Epidemiology*, *49*(11), 1225–31.

Turrell, G., Lynch, J. W., Leite, C., Raghunathan, T., & Kaplan, G. A. (2007). Socioeconomic disadvantage in childhood and across the life course and all-cause mortality and physical function in adulthood: evidence from the Alameda County Study. *Journal of Epidemiology & Community Health*, *61*(8), 723–730.

Ujcic-Voortman, J. K., Baan, C. A., Seidell, J. C., & Verhoeff, A. P. (2012). Obesity and cardiovascular disease risk among Turkish and Moroccan migrant groups in Europe: A systematic review. *Obesity Reviews*, *13*(1), 2–16.

Vafaie, H., & de Jong, K. (1992). Genetic algorithms as a tool for feature selection in machine learning. *Proceedings Fourth International Conference on Tools with Artificial Intelligence TAI '92*, *6*(July), 267–281.

Vaillancourt, C., Lui, A., De Maio, V. J., Wells, G. A., & Stiell, I. G. (2008). Socioeconomic status influences bystander CPR and survival rates for out-of-hospital cardiac arrest victims. *Resuscitation*, *79*(3), 417–423.

Vaillancourt, C., Stiell, I. G., & Wells, G. A. (2008). Understanding and improving low bystander CPR rates: A systematic review of the literature. *Canadian Journal of Emergency Medicine*.

van Alem, A. P., Waalewijn, R. A., Koster, R. W., & de Vos, R. (2004). Assessment of quality of life and cognitive function after out-of-hospital cardiac arrest with successful resuscitation. *The American Journal of Cardiology*, *93*(2), 131–135.

Van Oort, F. V. A., Van Lenthe, F. J., & Mackenbach, J. P. (2005). Material, psychosocial, and behavioural factors in the explanation of educational inequalities in mortality in the Netherlands. *Journal of Epidemiology & Community Health*, *59*, 214–220.

Van Oyen, H., Deboosere, P., & Lorant, V. (2011). *Sociale ongelijkheden in gezondheid in België*.

Vapnik, V. (1995). The Nature of Statistical Learning Theory.

Warden, C., Cudnik, M. T., Sasson, C., Schwartz, G., & Semple, H. (2012). Poisson Cluster Analysis of Cardiac Arrest Incidence in Columbus, Ohio. *Prehospital Emergency Care*, *16*(3), 338–346.

Wells, D. M., White, L. L. Y., Fahrenbruch, C. E., & Rea, T. D. (2016). Socioeconomic status and survival from ventricular fibrillation out-of-hospital cardiac arrest. *Annals of*

Epidemiology, *26*(6), 418–423.e1.

Weng, S. F., Reps, J., Kai, J., Garibaldi, J. M., & Qureshi, N. (2017). Can machine-learning improve cardiovascular risk prediction using routine clinical data? *PLoS ONE*, *12*(4), 1–14.

WHO. (2010). *How health systems can address health inequities linked to migration and ethnicity*.

Wilde, E. T., Robbins, L. S., & Pressley, J. C. (2012). Racial differences in out-of-hospital cardiac arrest survival and treatment. *Emergency Medicine Journal*, *29*(5), 415–419.

Witten, I. H., Frank, E., Hall, M. A., & Pal, C. J. (2016). *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann.

Yang, J., & Honavar, V. (1997). Feature subset selection using a genetic algorithm. *IEEE Intelligent Systems*, *13*(2), 44–49.

Yen, I. H., & Kaplan, G. A. (1998). Poverty Area Residence and Changes in Physical Activity Level: Evidence From the Alameda County Study. *American Journal of Public Health*, *88*(11), 1709–1712.

Zhang, J., & Yu, K. F. (1998). What's the relative risk? A method of correcting the odds ratio in cohort studies of common outcomes. *JAMA : The Journal of the American Medical Association*, *280*(19), 1690–1.

Zhong, J., Hu, X., Zhang, J., & Gu, M. (2005). Comparison of Performance between Different Selection Strategies on Simple Genetic Algorithms. *International Conference on Computational Intelligence for Modelling, Control and Automation and International Conference on Intelligent Agents, Web Technologies and Internet Commerce*, *2*, 1115–1121.

Zijlstra, J. A., Radstok, A., Pijls, R., Nas, J., Beesems, S. G., Hulleman, M., … Blom, M. T. (2016). Overleving na een reanimatie buiten het ziekenhuis: vergelijking van de resultaten van 6 verschillende Nederlandse regio's. *Reanimatie in Nederland*, 9–23.

Zijlstra, J. A., Stieglis, R., Riedijk, F., Smeekes, M., van der Worp, W. E., & Koster, R. W. (2014). Local lay rescuers with AEDs, alerted by text messages, contribute to early defibrillation in a Dutch out-of-hospital cardiac arrest dispatch system. *Resuscitation*, *85*(11), 1444–1449.

Zivin, J. A., & Bartko, J. J. (1976). Statistics for disinterested scientists. *Life Sciences*, *18*(1), 15–26.

Zuur, A. F., Ieno, E. N., & Elphick, C. S. (2010). A protocol for data exploration to avoid common statistical problems. *Methods in Ecology and Evolution*, *1*(1), 3–14.

Zuur, A. F., Ieno, E. N., Walker, N. J., Saveliev, A., & Smith, G. M. (2009). Zero-Truncated and Zero-Inflated Models for Count Data. In *Mixed effects models and extensions in ecology with R* (pp. 261–293).

APPENDICES

Appendix A. CONFIDENTIAL - Demographic and EMS characteristics of OHCA cases

Appendix B. Manual attribute selection from Statistics Netherlands

This Appendix describes which available indicators from Statistics Netherlands are included.

Basic demographics

Urbanisation is defined by the surrounding address density (OAD). OAD refers to the average number of addresses within a one kilometre radius. Statistics Netherlands uses five categories to classify the OAD, ranging from Category 1 ('strongly urbanised') to Category 5 ('not urbanised') (Statistics Netherlands, 2016). Population density is another available indicator and defined as the number of inhabitants per km^2 of land surface. The indicator naturally correlates considerably with the degree of urbanisation. We aim to avoid including attributes that measure similar information. Based on our literature review in Section 3.3.2, we may prefer to include the population density measure. However, the indicator has a wide numeric range. To avoid grouping this wide range subjectively, we instead use the degree of urbanisation as used by Statistics Netherlands.

In addition, data on five distinct age proportions is available: 0-14 years, 15-24 years, 25-44 years, 45-64 years, and at least 65 years of age. We consider only the latter, since this threshold lies close to the average age of OHCA victims in both study regions (≈ 66 years, see Chapter 2). Furthermore, data is accessible concerning gender and marital status (including married, unmarried, divorced and widowed). From the latter, we consider the proportion of married inhabitants. Finally, data on birth and mortality rates are available, but both are not considered in this research.

Housing

The average property value (WOZ) is attainable, as well as the average household size and proportions of several housing characteristics, namely: (1) single households, (2) households with children, (3) households without children, (4) properties before the year 2000, (5) properties built in or after the year 2000, (6) private properties, (7) rental properties, (8) single-family properties, and (9) multi-family properties[5].

We do not consider (1-3), as it measures similar information as the average household size. Although the proportion of old houses is related to low district SES (Knol, 2012), we do not believe that taking the year 2000 as a threshold in this context gives us relevant measures. Hence, we do not consider (4) and (5). Furthermore, (6) and (7) are mutually exclusive, and (6) is only accounted for in prior studies as an alternative measure of WOZ. Since the latter is available, we do not consider both indicators. Finally, we believe (8) and (9) are similar alternative measures of WOZ, and therefore not considered in this study.

Income

Statistics Netherlands measures average income levels per income receiver as well as per capita. We find it more relevant to measure income per income receiver rather than per capita, as we

[5] "Every dwelling that forms an entire property together with other dwellings and business premises. This includes apartment blocks, flats with gallery entrances or entrance halls, dwellings occupying a lower or an upper floor or floors, and dwellings above company premises when they have a separate entrance outside the company premises." ("Definitions," n.d.).

already include the fraction of unemployed workers out of the working population. However, as opposed to median income levels commonly used in prior studies, average income levels may be biased towards a small proportion of considerably high income levels. In addition, if both available, we prefer other income related measures to be given per income receiver as well (rather than per household). The following income level information remains and is given as proportions:

1. Income receivers that belong to the national 40% of people with the lowest income.
2. Income receivers that belong to the national 20% of people with the highest income.
3. Households with income lower than or around minimum wage[6].
4. Proportion of households with low purchasing power.

Indicators (1), (3) and (4) measure similar information and are likely to be highly correlated. At the same time, less information (i.e. less frequent measurements) is available for (3) and (4). In addition, we consider both (1) and (2) as highly intercorrelated, and as alternative measures of the average income per income receiver. We choose to include (1) as our single income measure, as we aim to investigate whether a high proportion of inhabitants that receives a low income, is related with high OHCA incidence rates.

Social protection

The Dutch social protection system consists of social assistance, social security and state pension provisions. Social assistance only applies to the working age population (15-65 years) that is able to work. It is primarily covered by the Work and Social Assistance Act (WWB). Next to that, social security can be divided into employee insurance and national insurance. Statistics Netherlands measures the proportion of individual recipients that are covered by (1) the Work and Social Assistance Act (WWB), (2) the Unemployment Benefits Act (WW), (3) long-term occupational disability acts, and (4) the General Old Age Pensions Act (AOW). Additional information on these measures of social protection is given in Appendix C.

The relation between occupation and both (1) and (2) is obvious. Additionally, in communities with low SES, physical disabilities are encountered more often and at a much younger age. This complicates the ability of the population in these communities to retain sufficient income up until the pension eligibility age is reached (Majer et al., 2011). Therefore, we also include (3) as a social protection measure.

Given the pension eligibility age of 65 years old, the proportion of recipients of AOW and the proportion of inhabitants of at least 65 years old measure similar information. We do not consider the former, since the number of recipients that received AOW is not known for each year.

Occupation

Finally, the proportion of unemployed inhabitants out of the working age population is available. More specifically, the proportion of income receivers aged between 15 and 65 years old with 52 weeks of income that claimed social benefits[7] is measured up until 2012. A slightly different

[6] The minimum poverty threshold as determined in political decision making (“Definitions”).

[7] Including AOW, ANW, ZW, WAO, WAZ, Wajong, WW, benefits for hold-ups due to frost, ABW, IOAW and IOAZ.

measure was introduced in 2013, measuring the proportion of inhabitants aged between 15 and 75 years old in households with as personal main revenue, income from labour or income from a private company. These indicators, together with the proportion of recipients of the WWB, measure similar information and are expected to have high intercorrelation. We only include the proportion of recipients of the WWB, since more years of data is available for this measure.

Next to that, we do not consider the working population per sector or ethnicity group[8], since data is only available prior to 2010. We are therefore unable to include this information for both regions.

Education

Statistics Netherlands does not gather information on education levels at a district level. We do not consider the number of students enrolled in higher education, the average distance to nearest school, and the average number of schools in a predefined area.

Ethnicity

Statistics Netherlands measures the proportion of inhabitants with a Western or non-Western background. An individual's background is considered to be Western if he or she, or at least one of his or her parents is born in Europe (excluding Turkey), North America, Oceania, Indonesia or Japan. Inhabitants with a Dutch background are not included in this group. Inhabitants with a non-Western background are subdivided into inhabitants from: (1) the Netherlands Antilles or Aruba, (2) Morocco, (3) Suriname, (4) Turkey, and (5) countries in Africa, Latin-America and Asia other than the previously mentioned ("Afbakening generaties met migratieachtergrond," n.d.).

Other

Measures that do not fall into one of the above categories concern electricity consumption (in kWh) and natural gas consumption (in m^3) in total as well as per property type, the proportion of properties connected to a district heating network of renewable energy, average distances to nearest facility types, the average number of facility types within a certain range, vehicle ownership, the proportion of companies per sector and proportions of terrain use. These attributes will not be considered in this research, but are shown in Table B.1 for the sake of completeness[9].

[8] All given in proportions: working population per sector (including agriculture, forestry, fishing, industry, energy, trade, hospitality, transport, information, communication, financial services, real estate, commercial services, non-commercial services, culture, recreation and other services) and working population per ethnicity group (Dutch inhabitants, inhabitants with a Western background, inhabitants with a non-Western background).

[9] Not shown: attributes corresponding to mutations, scaling factors, number of inhabitants, number of households, number of residences, most common zip code, covering percentage of most common zip code, attributes that only partially cover current definitions (e.g. prior measures of social protection).

Attribute	Value
Rental properties in possession of housing corporations	%
Rental properties in possession of owners other than housing corporations	%
Properties with unknown ownership	%
Stacked properties	%
Occupied properties	%
Empty properties	%
Average electricity consumption per type of property[1] (8 attributes)	kWh
Average natural gas consumption per type of property[1] (8 attributes)	m^3
Average distance to nearest type of facility[2,3] (17 attributes)	km
Average number of facility types[4] within a certain range[3] (10 attributes)	#
Passenger cars per household	#.#
Passenger cars per km^2 of land surface	#.#
Passenger cars per household that are no more than 6 years old	%
Passenger cars per household that is at least 6 years old	%
Passenger cars that run on petrol	#
Passenger cars that run on fuel other than petrol[5]	#
Motorcycles	#
Commercial motor vehicles	#
Average electricity consumption per household	kWh
Average natural gas consumption per household	m^3
District heating share	%
Companies per sector[6] (14 attributes)	#
Urban area[7] (2 attributes)	ha
Terrain use[8] (6 attributes)	ha
Surface[9] (3 attributes)	ha

[1] Private property, rental property, detached house, semi-detached house, terraced house, corner house, apartment and unknown type of property.
[2] GP practice, GP centre, hospital (excl. outpatient clinics), nursery, school (all types), school (primary education), school (secondary education), school (prevocational secondary education (vmbo)), school (senior general secondary education (havo) & pre-university education (vwo)), large retail supermarket, restaurant, cinema, library, swimming pool, public green space, sports complex, train station, slip road of provincial road or motorway.
[3] For some facility types, more than one range is given (e.g. 1, 3 and 5 km). One facility type can therefore be included by more than one unique attribute.
[4] GP practice, hospital (excl. outpatient clinics), nursery, school (all types), school (primary education), school (secondary education), school (prevocational secondary education (vmbo)), school (senior general secondary education (havo) & pre-university education (vwo)), large retail supermarket, restaurant, cinema.
[5] Including diesel, LPG, electricity (incl. hybrid), hydrogen, alcohol, liquefied natural gas and compressed natural gas.
[6] Including agriculture, forestry, fishing, industry, energy, trade, hospitality, transport, information, communication, financial services, real estate, commercial services, non-commercial services, culture, recreation and other services.
[7] Urban area, rural area.
[8] Traffic area, built-up area, semi-built-up area, recreational area, agricultural terrain, forest and open natural terrain.
[9] Total surface, water surface, land surface.

Table B.1. *Redundant attributes.*

Appendix C. Social protection measures in the Netherlands

This Appendix discusses social protection measures in the Netherlands of which data has been made publicly available at a regional level.

Work and Social assistance Act

The Work and Social assistance Act (WWB) came into force in 2004 and provides income support benefits to citizens that lack the means to adequately support themselves. It also includes subsidised work and reintegration facilities (Blommesteijn & Mallee, 2009). One can only apply if he or she does not possesses sufficient assets or lives with a partner or parents with sufficient income. Hence, the WWB is referred to as the social safety net. In 2015, the WWB was replaced by the Participation Act.

Unemployment

The Unemployment Benefits Act (WW) insures employees against the financial consequences of unemployment. The duration for which the employee receives the unemployment benefit depends on the employment record ("Definitions," n.d.).

Disablement legislation

Four different Dutch laws to provide for the financial consequences of long-term occupational disability exist: the Occupational Disability Insurance Act (WAO), the Disablement Act for Self-employed (WAZ), the Act on Work and Income according to Labour Capacity (WIA) and the Disablement Provision Act for Disabled from an Early Age (Wajong). These labour disability schemes are financed directly from public funds ("Disability schemes," n.d.).

State pension provisions

The General Old Age Pensions Act (AOW) is a social insurance that provides all people with a basic state pension after reaching the entitlement age. The size of the entitlement gradually increases with the number of years that the recipient is insured ("Old age pension," n.d.). Throughout the study period, the pension eligibility age was set at 65 years. However, one can choose to receive AOW five years before or after the entitlement age.

Appendix D. Hyperparameter settings of Genetic Algorithm

This Appendix describes the operators and contains the settings of the Genetic Algorithms (GAs).

Selection

Evolution of binary strings between two successive generations j and $j+1$ in our GA occurs differently than evolution of individuals in natural populations. We choose tournament selection as our type of selection scheme due to its simplicity, its ability to apply dynamic selection pressure, as well as its ability to run computations in parallel (Miller & Goldberg, 1995). The selection scheme's computation time equals $O(n)$ and is therefore faster than the majority of other suitable schemes (Goldberg & Deb, 1991). In particular, Zhong et al. (2005) show that tournament selection outperforms the main alternative scheme, roulette wheel selection, based on more efficient algorithmic convergence. Tournament selection randomly gathers a fraction k out of all individuals with replacement, and only selects the individual with the highest fitness. This step is repeated until the number of selected individuals equals the population size. In this way, the probability that an individual is selected to form a new subset increases with its fitness to incorporate survival of the fittest. We initially set $k = 0.25$. As the population evolves, we apply dynamic selection pressure by gradually increasing the tournament size k. As a result, the probability that a relatively weak individual is selected, decreases during the procedure. In addition, we always select the current generation's most fit individual.

Reproduction

The selected individuals then enter the reproduction stage, in which offspring is created by applying two genetic operators to all of these individuals, namely crossover and mutation. Crossover corresponds to combining two binary substrings (i.e. individuals) of generation j into two new binary substrings of generation $j+1$ (also referred to as offspring). The probability of performing crossover on any of the preselected individuals is generally set relatively large. We choose to use $p_{crossover} = 0.6$, as proposed by Yang & Honavar (1997). We use the uniform crossover type, as it can handle complex reproductions. Mutation refers to switching the value of any individual and prevents that all combinations throughout the procedure are based on subsets of the initial population. As this probability commonly is set very low, we set $p_{mutation} = 0.033$. Examples of applying these operators are shown in Figure D.1 and Figure D.2.

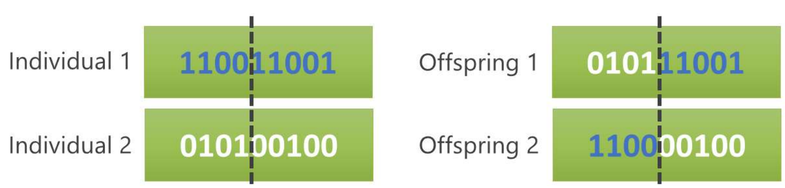


Figure D.1. *A simplified example of crossover between two individuals in generation j and the resulting offspring in generation $j+1$ within the GA.*

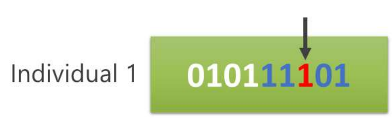


Figure D.2. *An example of mutated offspring in generation $j+1$ within the GA. The arrow refers to the particular mutation of the n^{th} gene (shown in red).*

Hyperparameter settings

Based on initial results, we limit the possible values of the SVM's regularization parameter C and scale parameter γ to reduce computational complexity.

Hyperparameter	Setting
Initialization	
Population size	25
Initial probability for an attribute to be switched on	0.5
Minimum number of attributes	1
Reproduction	
Fitness measure	Accuracy
Fitness function	SVM, LR
Kernel function*	Radial basis function
Regularization parameter C*	$[1,100]$
Scale parameter γ*	$[0.1]$
Selection scheme	Tournament
Tournament fraction size	0.25
Dynamic selection pressure	Yes
Keep best individual of each generation	Yes
Crossover probability	0.6
Crossover type	Uniform
Mutation probability	0.033
Maximal fitness	Infinity
Maximum number of generations	50
Use early stopping	Yes
Generations without improvement	15

**Only applicable to SVM.*

Table D.3. *Summary of the Genetic Algorithm's hyperparameter settings.*

Appendix E. CONFIDENTIAL - Data processing of cardiac arrest data

Appendix F. Flowcharts and geographic maps of structural reforms

This Appendix contains all flowcharts and geographic maps from the structural reforming process. As can be seen, the number of changes was considerably higher in North Holland, although a longer period was considered. In addition, no changes in geographic structures were reported in Twente between 2013 and 2016.

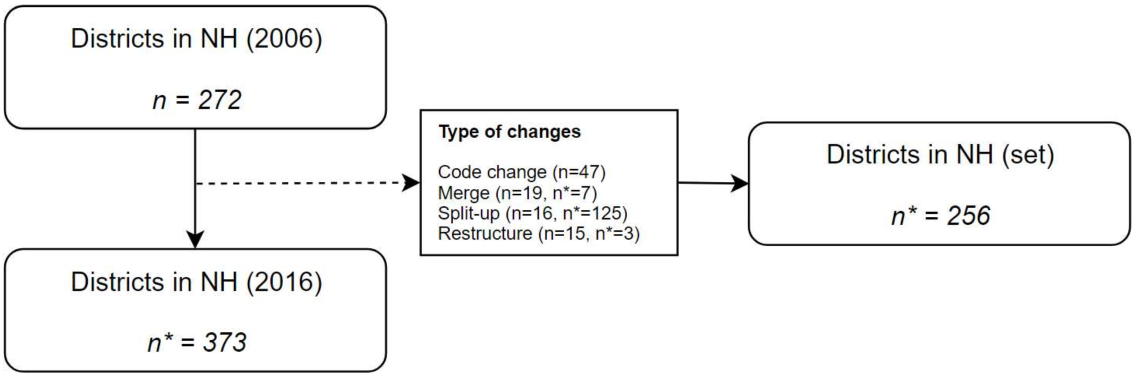


Figure F.1. *Flowchart of regional changes at a district level in North Holland between 2006 and 2016. The resulting number of districts is marked as n*.*

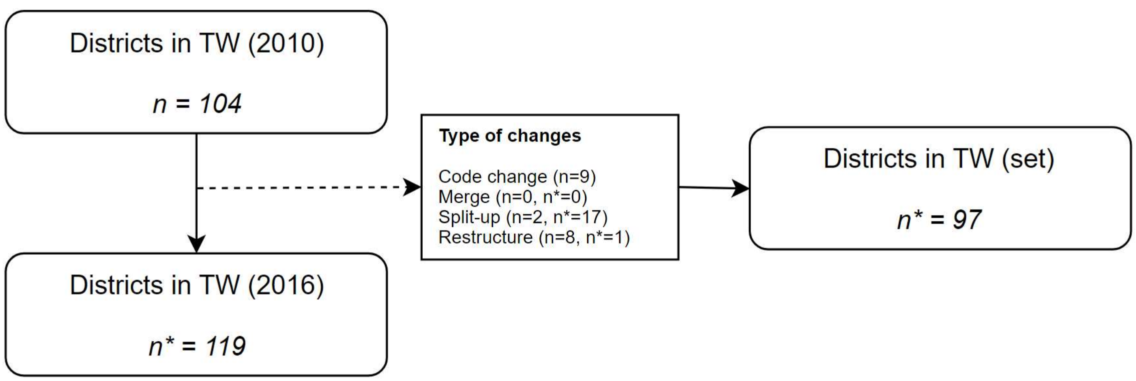


Figure F.2. *Flowchart of regional changes at a district level in Twente between 2010 and 2016. The resulting number of districts is marked as n*.*

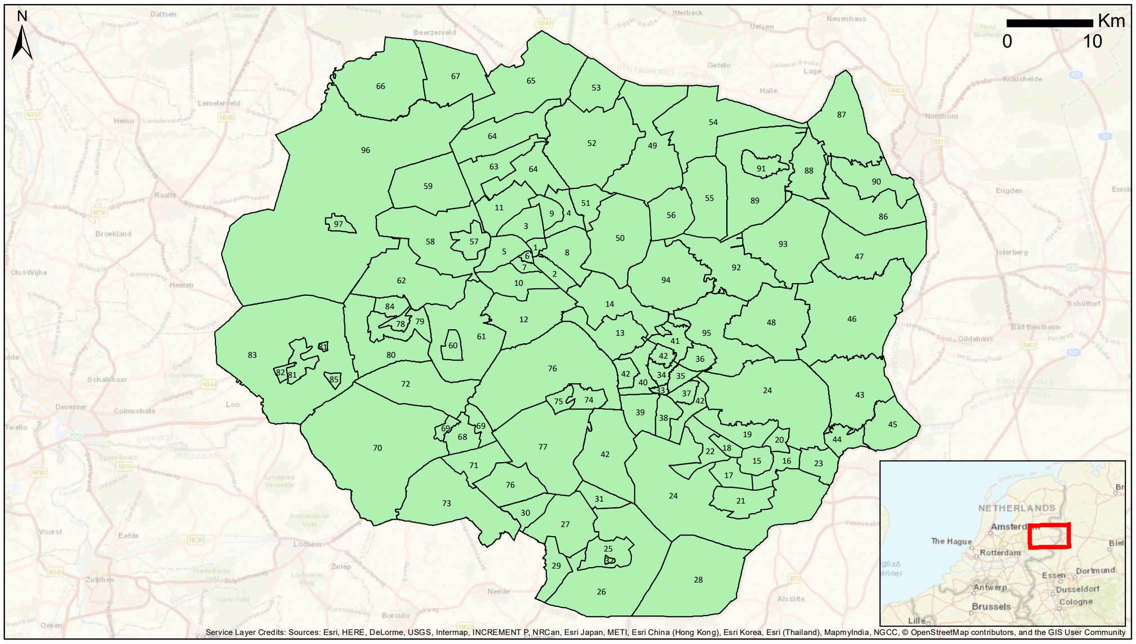


Figure F.3. *Constructed set of districts in Twente (97 districts).*

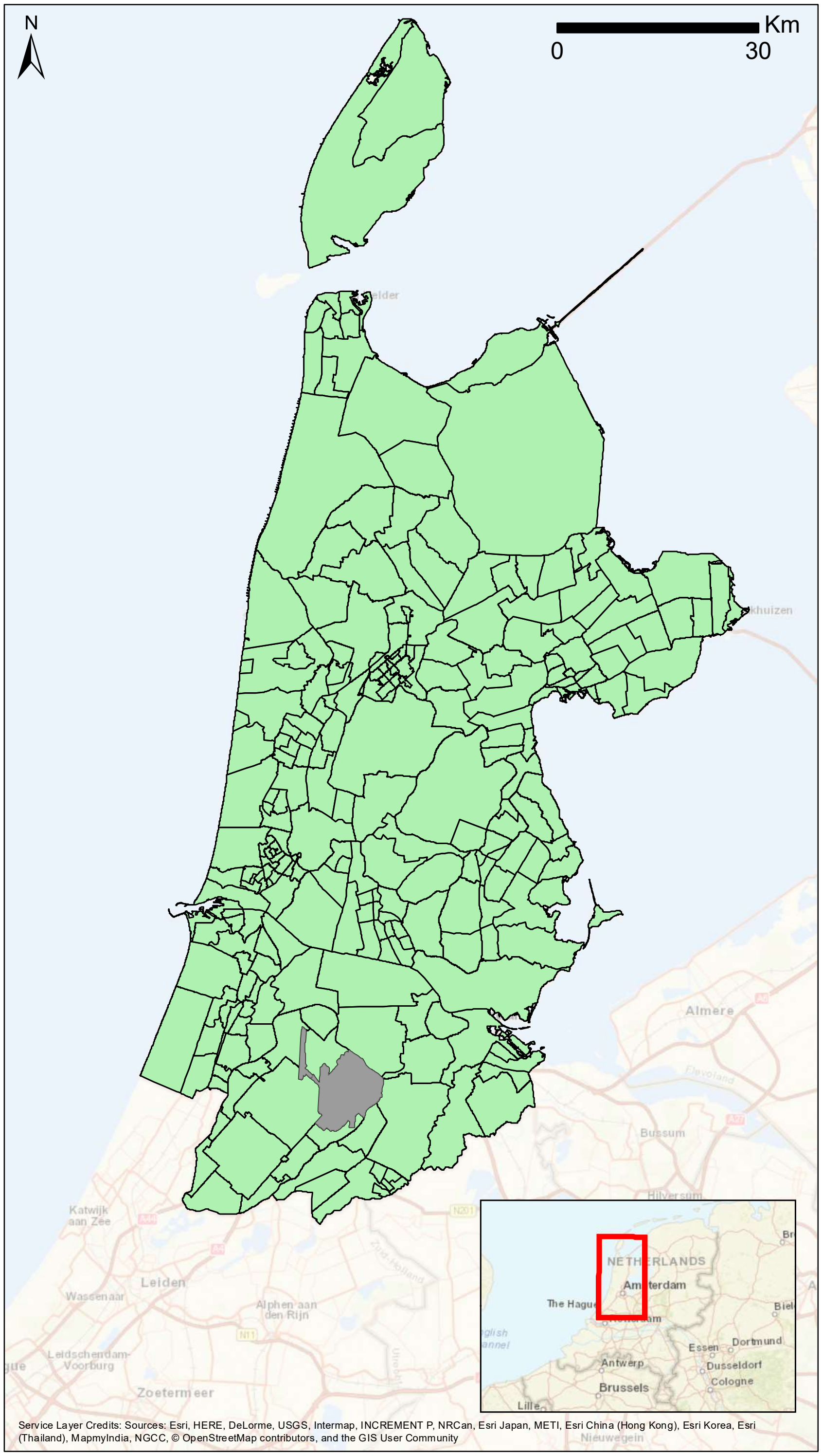


Figure F.4. *Constructed set of districts in North Holland (256 districts)*

Appendix G. Exclusion of districts

This Appendix contains an overview of the number of districts that do not meet our eligibility criteria.

Step	*Attribute*	*Threshold*	*North Holland (n = 256)*	*Twente (n = 97)*
1	Inhabitants	< 200	8	5
2	Households	< 50	0	0
3	Income receivers	< 100	0	0
4	Region name[1]	company-related[2]	0	2
5	Company/population	> 1	2	0
6	All	unknown[3]	0	0
		Excluded, n (%)	10 (3.9)	7 (7.2)

[1]Names are unknown for restructured regions.
[2]Either 'company', 'companies', 'industry', 'harbour'.
[3]For each year throughout the study period.

Table G.1. *Stepwise exclusion of districts in North Holland (2006-16) and Twente (2010-16).*

Appendix H. Missing data

This Appendix contains the percentage of missing data for each attribute in both study regions.

	North Holland (n = 246)	Twente (n = 90)
Socio-demographic factors		
Population	0.0	0.0
Degree of urbanisation	0.0	0.0
At least 65 years old	0.0	0.0
Male inhabitants	0.0	0.0
Married inhabitants	18.2	0.0
Average household size	0.0	0.0
Average property value	0.0	0.8
Receivers of low income	18.4	14.6
Recipients of the WWB	18.4	14.6
Recipients of labour disability acts	18.4	14.6
Recipients of the WW	27.4	14.6
Ethnicity background		
Western	0.0	0.0
The Netherlands Antilles or Aruba	2.7	15.7
Morocco	2.7	15.7
Suriname	2.7	15.7
Turkey	2.7	15.7
Remaining non-Western	2.7	15.7

Note: n = number of districts.

Table H.1. *Missing data (in percentage) per attribute in North Holland (2006-16) and Twente (2010-16).*

Appendix I. Geographic OHCA risk maps

In this Appendix, geographic risk distributions based on incidence of EMS-attended OHCA in both study regions are illustrated.

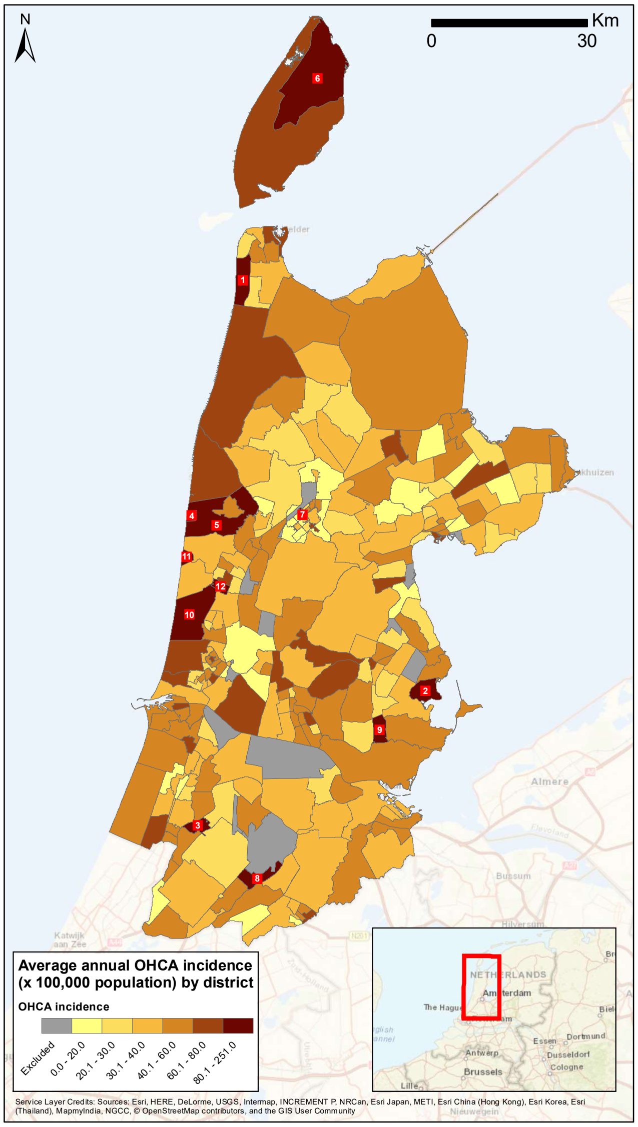


Figure I.1. *Average annual standard OHCA incidence rates (entire day) by district in North Holland (256 districts, 2006-16). The numbers indicate the rank (from highest to lowest incidence rate) of the particular district, across all considered districts with average annual standard OHCA incidence rates (entire day) above 80.*

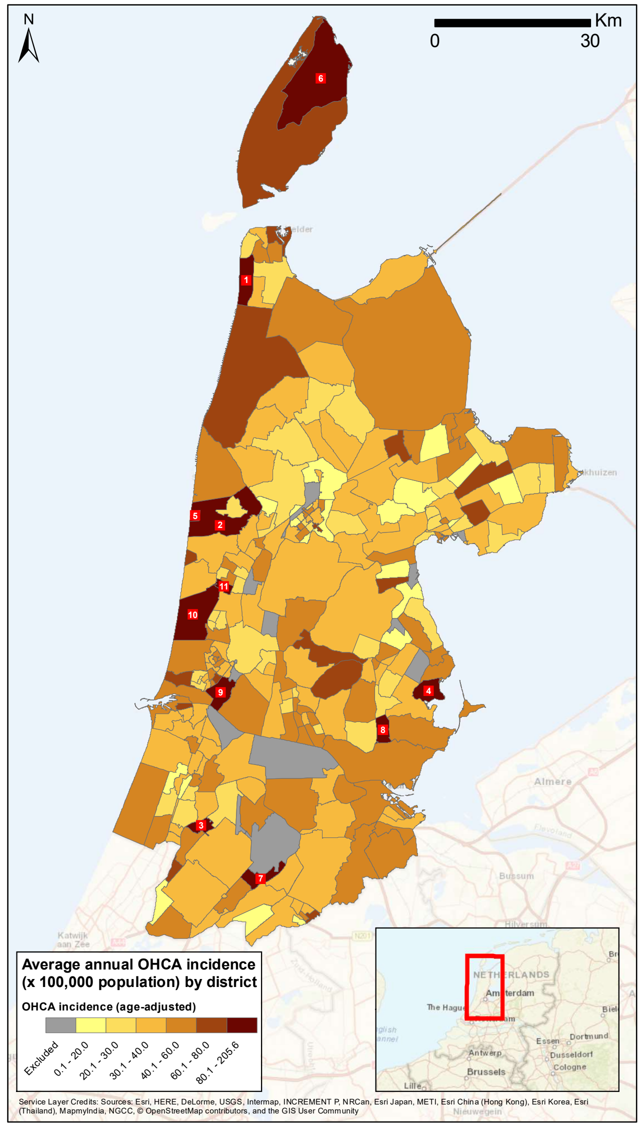


Figure I.2. *Average annual age-adjusted OHCA incidence rates (entire day) by district in North Holland (256 districts, 2006-16). The numbers indicate the rank (from highest to lowest incidence rate) of the particular district, across all considered districts with average annual age-adjusted OHCA incidence rates (entire day) above 80.*

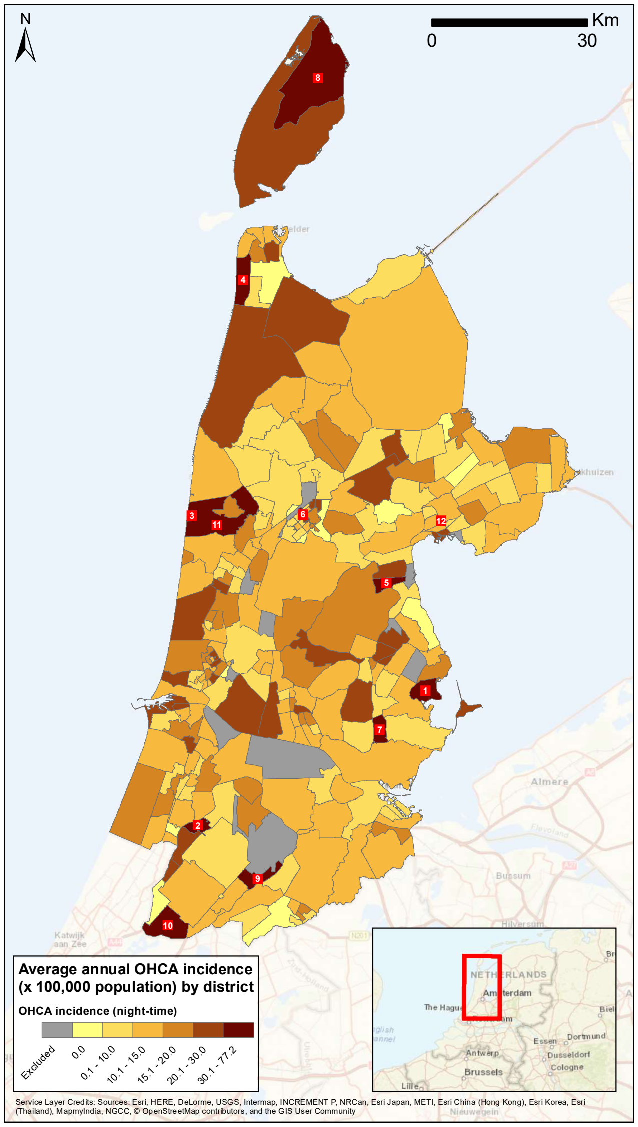


Figure I.3. *Average annual standard OHCA incidence rates (night-time only) by district in North Holland (256 districts, 2006-16). The numbers indicate the rank (from highest to lowest incidence rate) of the particular district, across all considered districts with average annual standard OHCA incidence rates (night-time only) above 30.*

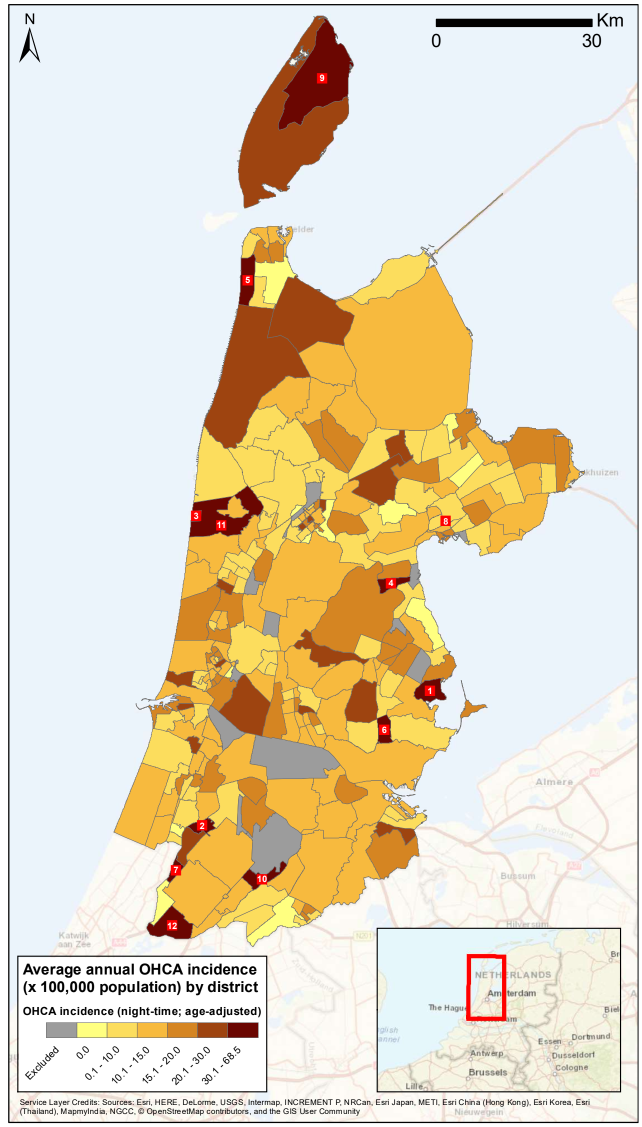


Figure I.4. *Average annual age-adjusted OHCA incidence rates (night-time only) by district in North Holland (256 districts, 2006-16). The numbers indicate the rank (from highest to lowest incidence rate) of the particular district, across all considered districts with average annual age-adjusted OHCA incidence rates (night-time only) above 30.*

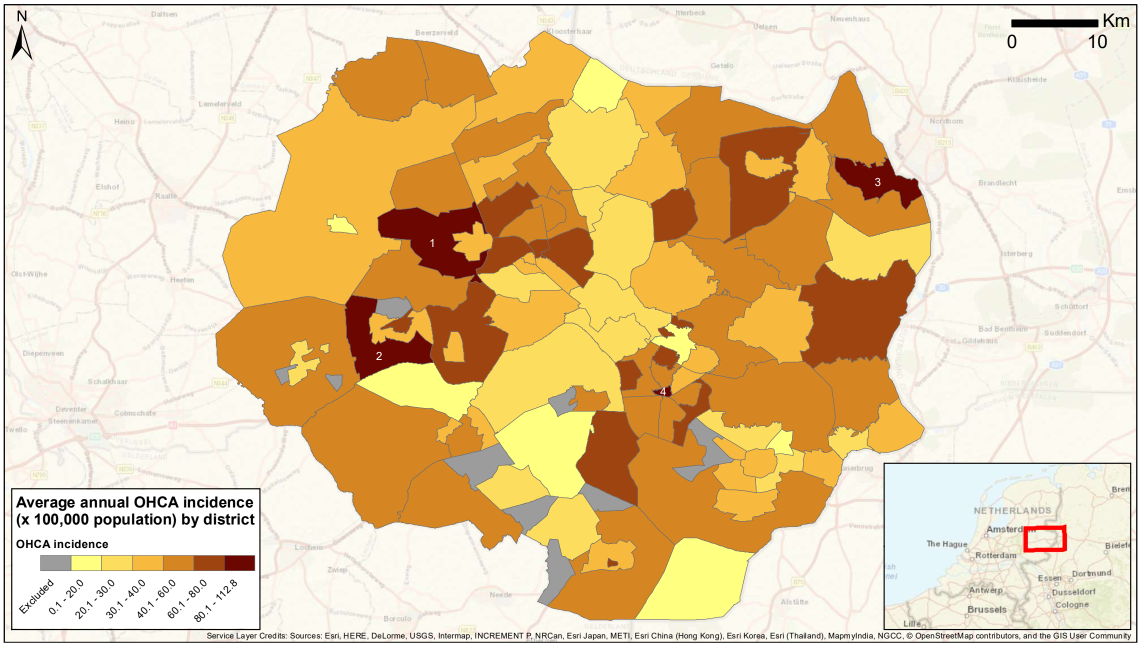


Figure I.5. *Average annual standard OHCA incidence rates (entire day) by district in Twente (97 districts, 2010-16). The numbers indicate the rank (from highest to lowest incidence rate) of the particular district, across all considered districts with average annual standard OHCA incidence rates (entire day) above 80.*

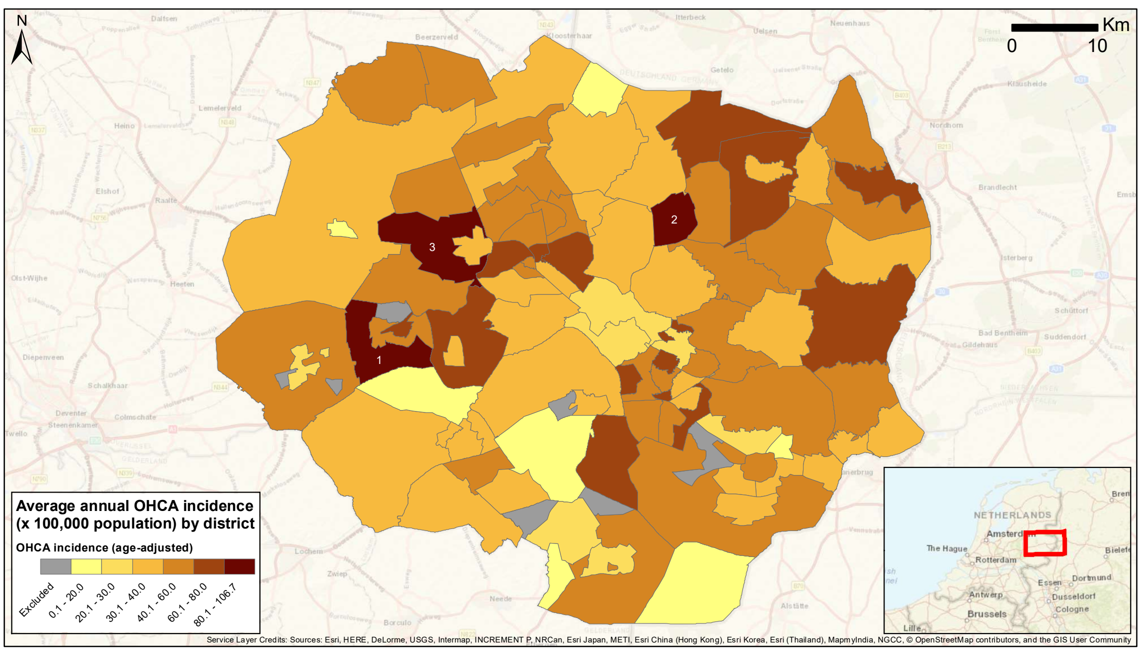


Figure I.6. *Average annual age-adjusted OHCA incidence rates (entire day) by district in Twente (97 districts, 2010-16). The numbers indicate the rank (from highest to lowest incidence rate) of the particular district, across all considered districts with average annual age-adjusted OHCA incidence rates (entire day) above 80.*

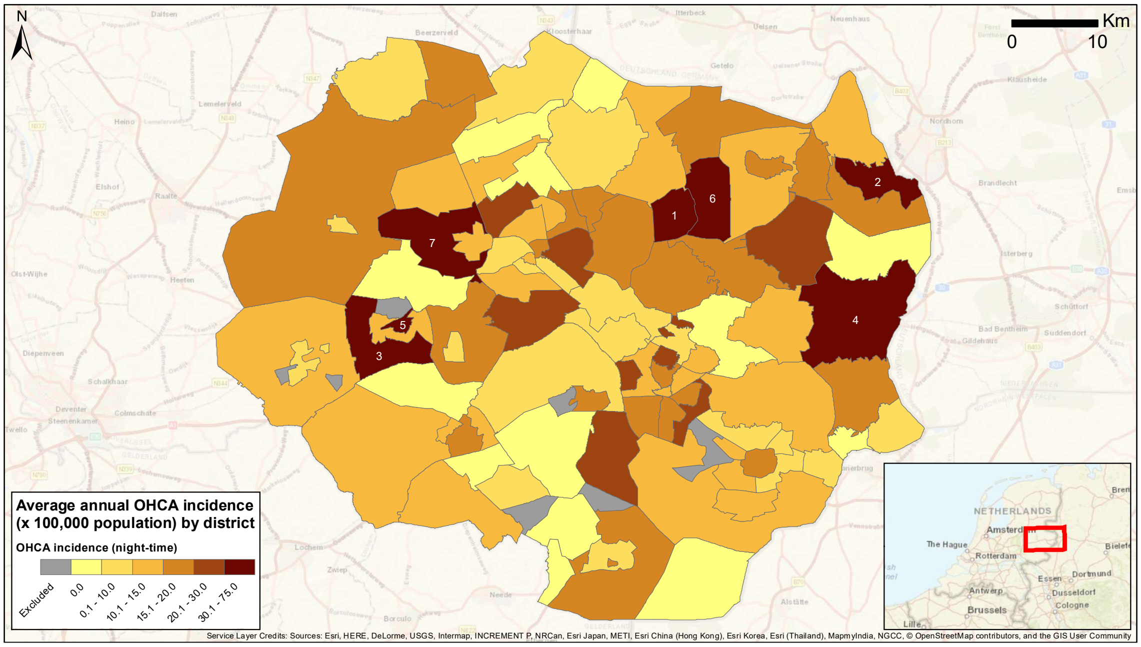


Figure I.7. *Average annual standard OHCA incidence rates (night-time only) by district in Twente (97 districts, 2010-16). The numbers indicate the rank (from highest to lowest incidence rate) of the particular district, across all considered districts with average annual standard OHCA incidence rates (night-time only) above 30.*

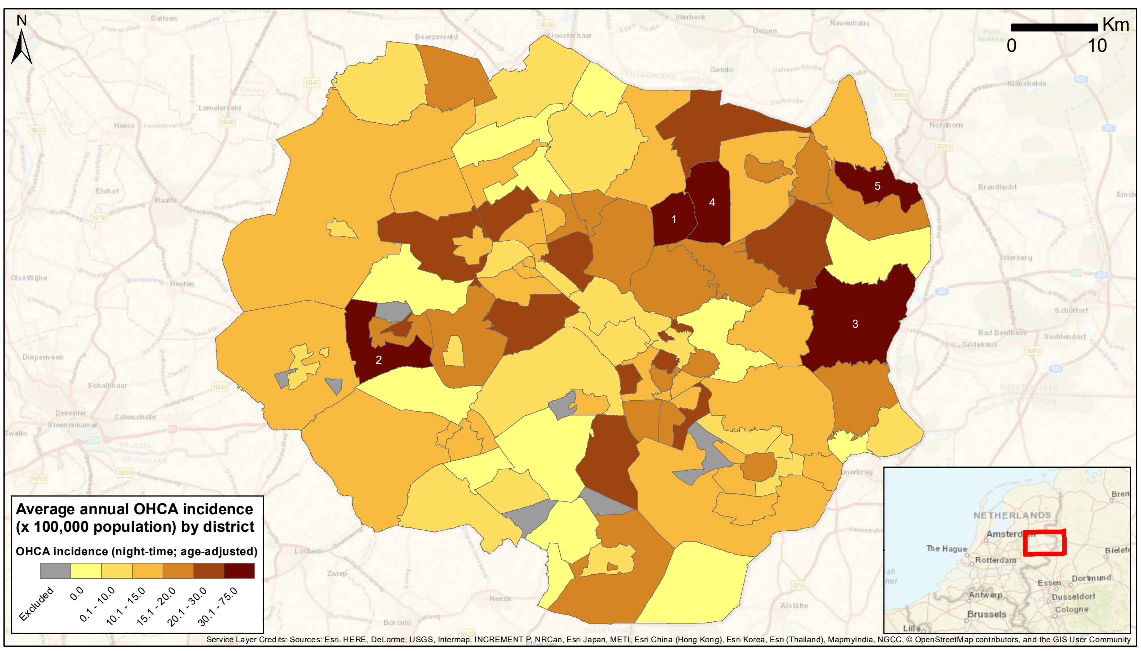


Figure I.8. *Average annual age-adjusted OHCA incidence rates (night-time only) by district in Twente (97 districts, (2010-16). The numbers indicate the rank (from highest to lowest incidence rate) of the particular district, across all considered districts with average annual age-adjusted OHCA incidence rates (night-time only) above 30.*

Appendix J. CONFIDENTIAL - Districts with highest incidence rates

Appendix K. Accuracy graphs of the automated dimensionality reduction phase

To demonstrate the development of the Genetic Algorithm over generations, this Appendix contains graphs of the respective average and highest accuracy achieved in North Holland, given standard incidence rates (entire day), and Logistic Regression as base learner.

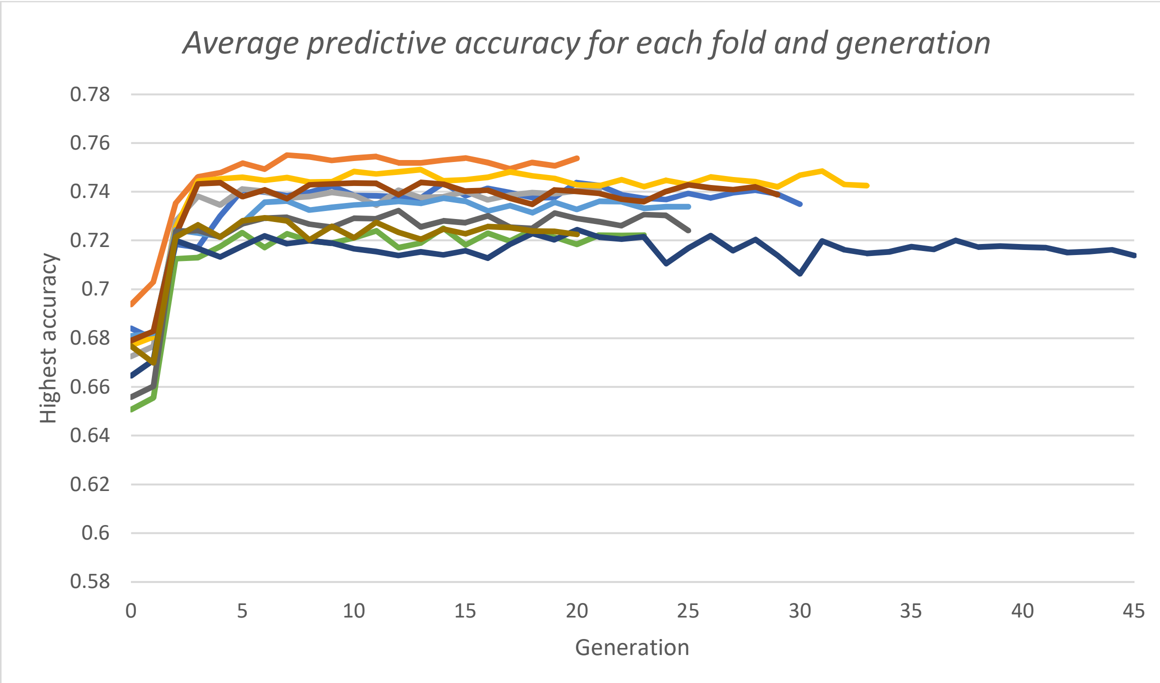


Figure K.1. *Average accuracy achieved in North Holland (246 districts, 2006-16), given standard incidence rates based on OHCAs that occurred at any time of day.*

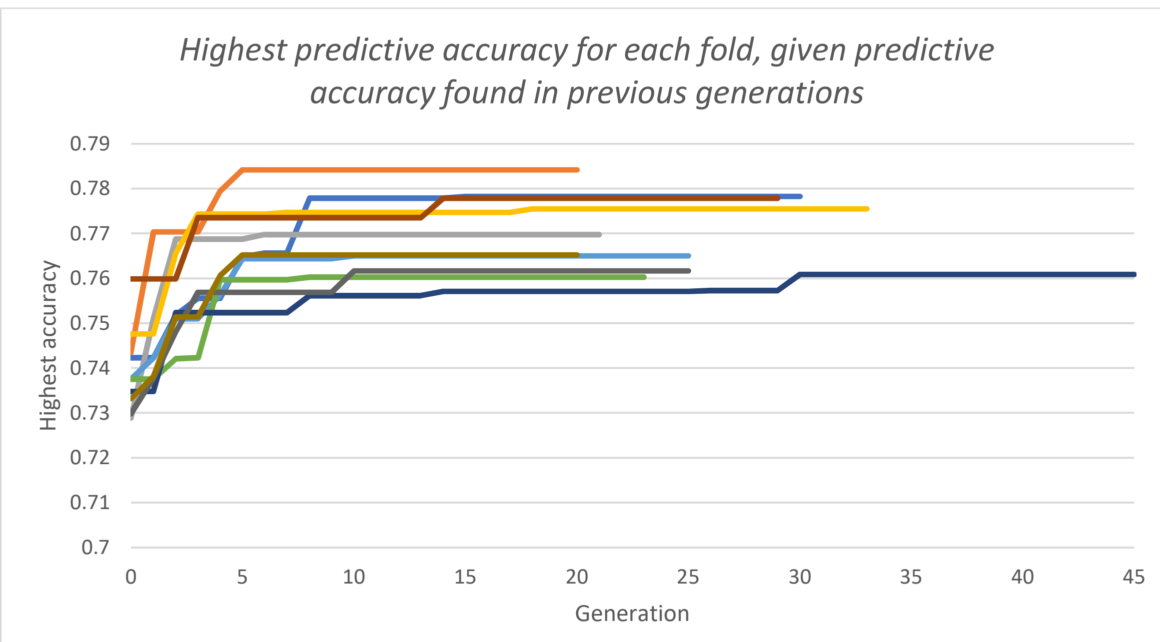


Figure K.2. *Highest generational accuracy achieved in North Holland (246 districts, 2006-16), given standard incidence rates based on OHCAs that occurred at any time of day.*

Appendix L. Results of the automated dimensionality reduction

This Appendix includes the number of folds the pre-selected attributes are found in each configuration as a result of the automated dimensionality reduction phase. Results are obtained by means of Genetic Algorithms (GAs).

	North Holland				**Twente**			
Incidence rate	*Standard*		*Adjusted*		*Standard*		*Adjusted*	
Time of day	*Full*	*Night*	*Full*	*Night*	*Full*	*Night*	*Full*	*Night*
Urbanisation	3	2	7	2	2	1	8	5
At least 65 years	9	6	-	-	10	6	-	-
Men	8	4	3	10	5	9	7	5
Married	7	7	3	2	2	10	5	6
Household size	10	7	10	8	5	4	3	7
Property value	0	2	5	2	5	5	10	2
Low income	10	8	7	6	2	3	2	3
Recipients of WWB	3	2	5	5	2	4	4	1
Recipients of AO	6	5	2	5	7	5	10	9
Recipients of WW	1	5	7	7	5	5	4	5
Western	1	5	5	1	3	4	5	7
Antilleans/Aruban	3	2	7	4	-	-	-	-
Moroccan	3	6	5	0	-	-	-	-
Surinamese	4	3	4	2	-	-	-	-
Turkish	2	4	5	2	4	3	5	7
Other non-Western	4	3	4	5	2	6	4	1

Table L.1. *Number of folds each attribute is found in after applying the GA-LR.*

	North Holland				**Twente**			
Incidence rate	*Standard*		*Adjusted*		*Standard*		*Adjusted*	
Time of day	*Full*	*Night*	*Full*	*Night*	*Full*	*Night*	*Full*	*Night*
Urbanisation	9	5	8	4	7	3	3	7
At least 65 years	7	6	-	-	10	10	-	-
Men	6	2	3	10	7	9	2	1
Married	7	7	5	7	2	3	7	2
Household size	9	8	7	8	6	6	8	7
Property value	4	3	5	2	3	5	4	9
Low income	7	5	6	6	7	3	5	9
Recipients of WWB	2	3	2	3	2	5	5	5
Recipients of AO	2	1	3	1	3	7	4	8
Recipients of WW	3	6	5	6	10	6	7	3
Western	6	4	7	5	7	0	4	4
Antilleans/Aruban	7	1	8	3	-	-	-	-
Moroccan	4	2	7	5	-	-	-	-
Surinamese	1	3	3	2	-	-	-	-
Turkish	5	3	6	5	4	5	5	3
Other non-Western	6	3	5	2	5	5	1	9

Table L.2. *Number of folds each attribute is found in after applying the GA-SVM.*

Appendix M. Machine learning classification results

Tables M.1 and M.2 shows the classification results for each method and configuration.

Settings					*Performance*			
Time	*IR*	*GA*	*Weights*	*Model*	*Attributes*	*Accuracy*	*Sensitivity*	*Specificity*
Full	Standard	LR	Overall	LR	6	0.763 ± 0.053	0.768 ± 0.053	0.757 ± 0.054
Full	Standard	LR	Overall	ANN	6	0.747 ± 0.054	0.705 ± 0.057	0.789 ± 0.051
Full	Standard	LR	Overall	SVM	6	0.739 ± 0.055	0.694 ± 0.058	0.783 ± 0.052
Full	Standard	LR	Overall	ADA	6	0.665 ± 0.059	0.589 ± 0.061	0.742 ± 0.055
Full	Standard	LR	Best fold	LR	7	**0.772 ± 0.052**	**0.781 ± 0.052**	0.765 ± 0.053
Full	Standard	LR	Best fold	ANN	6	0.752 ± 0.054	0.714 ± 0.056	0.788 ± 0.051
Full	Standard	LR	Best fold	SVM	7	0.752 ± 0.054	0.720 ± 0.056	0.788 ± 0.051
Full	Standard	LR	Best fold	ADA	10	0.732 ± 0.055	0.677 ± 0.058	0.788 ± 0.051
Full	Standard	SVM	Overall	LR	9	0.719 ± 0.056	0.712 ± 0.057	0.724 ± 0.056
Full	Standard	SVM	Overall	ANN	9	0.711 ± 0.057	0.655 ± 0.059	0.766 ± 0.053
Full	Standard	SVM	Overall	SVM	9	0.715 ± 0.056	0.674 ± 0.059	0.758 ± 0.053
Full	Standard	SVM	Overall	ADA	9	0.662 ± 0.059	0.647 ± 0.060	0.676 ± 0.059
Full	Standard	SVM	Best fold	LR	11	0.749 ± 0.054	0.755 ± 0.054	0.741 ± 0.055
Full	Standard	SVM	Best fold	ANN	10	0.723 ± 0.056	0.692 ± 0.058	0.756 ± 0.054
Full	Standard	SVM	Best fold	SVM	11	0.756 ± 0.054	0.689 ± 0.058	**0.820 ± 0.048**
Full	Standard	SVM	Best fold	ADA	10	0.720 ± 0.056	0.659 ± 0.059	0.779 ± 0.052
Full	Adjusted	LR	Overall	LR	5	0.662 ± 0.059	0.642 ± 0.060	0.685 ± 0.058
Full	Adjusted	LR	Overall	ANN	5	0.642 ± 0.060	0.585 ± 0.062	0.699 ± 0.057
Full	Adjusted	LR	Overall	SVM	5	0.633 ± 0.060	0.506 ± 0.062	0.765 ± 0.053
Full	Adjusted	LR	Overall	ADA	5	0.585 ± 0.062	0.469 ± 0.062	0.701 ± 0.057
Full	Adjusted	LR	Best fold	LR	6	**0.684 ± 0.058**	0.648 ± 0.060	0.724 ± 0.056
Full	Adjusted	LR	Best fold	ANN	6	0.658 ± 0.059	0.587 ± 0.062	0.731 ± 0.055
Full	Adjusted	LR	Best fold	SVM	10	0.614 ± 0.061	**0.651 ± 0.060**	0.582 ± 0.062
Full	Adjusted	LR	Best fold	ADA	10	0.642 ± 0.060	0.570 ± 0.062	0.716 ± 0.056
Full	Adjusted	SVM	Overall	LR	7	0.662 ± 0.059	0.642 ± 0.060	0.685 ± 0.058
Full	Adjusted	SVM	Overall	ANN	7	0.634 ± 0.060	0.580 ± 0.062	0.692 ± 0.058
Full	Adjusted	SVM	Overall	SVM	7	0.621 ± 0.061	0.502 ± 0.062	0.742 ± 0.055
Full	Adjusted	SVM	Overall	ADA	7	0.629 ± 0.060	0.516 ± 0.062	0.740 ± 0.055
Full	Adjusted	SVM	Best fold	LR	7	0.675 ± 0.059	0.623 ± 0.061	0.731 ± 0.055
Full	Adjusted	SVM	Best fold	ANN	6	0.662 ± 0.059	0.539 ± 0.062	**0.788 ± 0.051**
Full	Adjusted	SVM	Best fold	SVM	7	0.666 ± 0.059	0.560 ± 0.062	0.772 ± 0.052
Full	Adjusted	SVM	Best fold	ADA	11	0.659 ± 0.059	0.613 ± 0.061	0.706 ± 0.057
Night	Standard	LR	Overall	LR	5	0.635 ± 0.060	0.642 ± 0.060	0.633 ± 0.060
Night	Standard	LR	Overall	ANN	5	0.603 ± 0.061	0.652 ± 0.060	0.558 ± 0.062
Night	Standard	LR	Overall	SVM	5	0.558 ± 0.062	0.544 ± 0.062	0.575 ± 0.062
Night	Standard	LR	Overall	ADA	5	0.618 ± 0.061	0.637 ± 0.060	0.599 ± 0.061
Night	Standard	LR	Best fold	LR	4	**0.650 ± 0.060**	0.648 ± 0.060	0.654 ± 0.059
Night	Standard	LR	Best fold	ANN	8	0.625 ± 0.060	0.616 ± 0.061	0.640 ± 0.060
Night	Standard	LR	Best fold	SVM	10	0.614 ± 0.061	0.651 ± 0.060	0.582 ± 0.062
Night	Standard	LR	Best fold	ADA	10	0.621 ± 0.061	0.625 ± 0.060	0.622 ± 0.061
Night	Standard	SVM	Overall	LR	4	0.623 ± 0.061	0.621 ± 0.061	0.632 ± 0.060
Night	Standard	SVM	Overall	ANN	4	0.619 ± 0.061	**0.679 ± 0.058**	0.565 ± 0.062
Night	Standard	SVM	Overall	SVM	4	0.582 ± 0.062	0.578 ± 0.062	0.592 ± 0.061
Night	Standard	SVM	Overall	ADA	4	0.591 ± 0.061	0.643 ± 0.060	0.540 ± 0.062
Night	Standard	SVM	Best fold	LR	5	0.638 ± 0.060	0.676 ± 0.059	0.602 ± 0.061
Night	Standard	SVM	Best fold	ANN	5	0.626 ± 0.060	0.655 ± 0.059	0.596 ± 0.061
Night	Standard	SVM	Best fold	SVM	5	0.645 ± 0.060	0.587 ± 0.062	**0.704 ± 0.057**
Night	Standard	SVM	Best fold	ADA	5	0.606 ± 0.061	0.527 ± 0.062	0.687 ± 0.058
Night	Adjusted	LR	Overall	LR	4	0.605 ± 0.061	0.593 ± 0.061	0.622 ± 0.061
Night	Adjusted	LR	Overall	ANN	4	0.613 ± 0.061	0.526 ± 0.062	0.702 ± 0.057
Night	Adjusted	LR	Overall	SVM	4	0.589 ± 0.061	0.528 ± 0.062	0.654 ± 0.059
Night	Adjusted	LR	Overall	ADA	4	0.524 ± 0.062	0.488 ± 0.062	0.566 ± 0.062
Night	Adjusted	LR	Best fold	LR	4	0.658 ± 0.059	0.572 ± 0.062	0.744 ± 0.055

Night	Adjusted	LR	Best fold	ANN	4	0.638 ± 0.060	0.565 ± 0.062	0.712 ± 0.057
Night	Adjusted	LR	Best fold	SVM	4	0.638 ± 0.060	0.448 ± 0.062	**0.822 ± 0.048**
Night	Adjusted	LR	Best fold	ADA	5	0.598 ± 0.061	0.470 ± 0.062	0.716 ± 0.056
Night	Adjusted	SVM	Overall	LR	6	0.605 ± 0.061	0.568 ± 0.062	0.646 ± 0.060
Night	Adjusted	SVM	Overall	ANN	6	0.577 ± 0.062	0.528 ± 0.062	0.631 ± 0.060
Night	Adjusted	SVM	Overall	SVM	6	0.557 ± 0.062	0.409 ± 0.061	0.703 ± 0.057
Night	Adjusted	SVM	Overall	ADA	6	0.507 ± 0.062	0.516 ± 0.062	0.503 ± 0.062
Night	Adjusted	SVM	Best fold	LR	7	0.630 ± 0.060	0.591 ± 0.061	0.671 ± 0.059
Night	Adjusted	SVM	Best fold	ANN	9	0.598 ± 0.061	0.486 ± 0.062	0.712 ± 0.057
Night	Adjusted	SVM	Best fold	SVM	4	**0.666 ± 0.059**	**0.627 ± 0.060**	0.697 ± 0.057
Night	Adjusted	SVM	Best fold	ADA	7	0.557 ± 0.062	0.558 ± 0.062	0.565 ± 0.062

IR = Incidence rate, LR = Logistic Regression, ANN = Artificial Neural Network, SVM = Support Vector Machine, ADA = Adaptive Boosting, GA = Genetic Algorithm.

Table M.1. *Performance measures in North Holland (2006-16, 246 districts) as a result of the predictive modelling analysis, given the time of day, the type of incidence rate, the attribute selection method, the weights used as a result from the attribute selection method, as well as the learning algorithm.*

Settings					*Performance*			
Time	*IR*	*GA*	*Weights*	*Model*	*Attributes*	*Accuracy*	*Sensitivity*	*Specificity*
Full	Standard	LR	Overall	LR	2	0.622 ± 0.100	0.560 ± 0.103	0.685 ± 0.096
Full	Standard	LR	Overall	ANN	2	0.644 ± 0.099	0.540 ± 0.103	0.745 ± 0.090
Full	Standard	LR	Overall	SVM	2	0.644 ± 0.099	0.475 ± 0.103	**0.805 ± 0.082**
Full	Standard	LR	Overall	ADA	2	0.622 ± 0.100	0.590 ± 0.102	0.665 ± 0.098
Full	Standard	LR	Best fold	LR	4	**0.711 ± 0.094**	0.645 ± 0.099	0.795 ± 0.083
Full	Standard	LR	Best fold	ANN	4	0.656 ± 0.098	0.665 ± 0.098	0.655 ± 0.098
Full	Standard	LR	Best fold	SVM	4	0.633 ± 0.100	0.580 ± 0.102	0.695 ± 0.095
Full	Standard	LR	Best fold	ADA	4	0.667 ± 0.097	0.680 ± 0.096	0.665 ± 0.098
Full	Standard	SVM	Overall	LR	7	0.589 ± 0.102	0.570 ± 0.102	0.615 ± 0.101
Full	Standard	SVM	Overall	ANN	7	0.656 ± 0.098	**0.755 ± 0.089**	0.580 ± 0.102
Full	Standard	SVM	Overall	SVM	7	0.611 ± 0.101	0.595 ± 0.101	0.640 ± 0.099
Full	Standard	SVM	Overall	ADA	7	0.622 ± 0.100	0.550 ± 0.103	0.715 ± 0.093
Full	Standard	SVM	Best fold	LR	4	0.633 ± 0.100	0.600 ± 0.101	0.665 ± 0.098
Full	Standard	SVM	Best fold	ANN	6	0.678 ± 0.097	0.735 ± 0.091	0.630 ± 0.100
Full	Standard	SVM	Best fold	SVM	6	0.678 ± 0.076	0.670 ± 0.104	0.665 ± 0.091
Full	Standard	SVM	Best fold	ADA	7	0.656 ± 0.098	0.590 ± 0.102	0.725 ± 0.092
Full	Adjusted	LR	Overall	LR	4	0.489 ± 0.103	0.405 ± 0.101	0.580 ± 0.102
Full	Adjusted	LR	Overall	ANN	4	0.511 ± 0.103	0.340 ± 0.098	0.685 ± 0.096
Full	Adjusted	LR	Overall	SVM	4	0.478 ± 0.103	0.345 ± 0.098	0.610 ± 0.101
Full	Adjusted	LR	Overall	ADA	4	0.444 ± 0.103	0.345 ± 0.098	0.600 ± 0.101
Full	Adjusted	LR	Best fold	LR	4	**0.667 ± 0.097**	**0.590 ± 0.102**	0.740 ± 0.091
Full	Adjusted	LR	Best fold	ANN	6	0.600 ± 0.101	0.585 ± 0.102	0.620 ± 0.100
Full	Adjusted	LR	Best fold	SVM	8	0.633 ± 0.100	0.540 ± 0.103	0.700 ± 0.095
Full	Adjusted	LR	Best fold	ADA	3	0.611 ± 0.101	0.480 ± 0.103	0.720 ± 0.093
Full	Adjusted	SVM	Overall	LR	3	0.522 ± 0.103	0.500 ± 0.103	0.555 ± 0.103
Full	Adjusted	SVM	Overall	ANN	3	0.556 ± 0.103	0.470 ± 0.103	0.645 ± 0.099
Full	Adjusted	SVM	Overall	SVM	3	0.578 ± 0.102	0.505 ± 0.103	0.670 ± 0.097
Full	Adjusted	SVM	Overall	ADA	3	0.411 ± 0.102	0.310 ± 0.096	0.535 ± 0.103
Full	Adjusted	SVM	Best fold	LR	6	0.600 ± 0.101	0.545 ± 0.103	0.665 ± 0.098
Full	Adjusted	SVM	Best fold	ANN	7	0.622 ± 0.100	0.495 ± 0.103	**0.750 ± 0.089**
Full	Adjusted	SVM	Best fold	SVM	7	0.622 ± 0.082	0.550 ± 0.101	0.690 ± 0.126
Full	Adjusted	SVM	Best fold	ADA	4	0.567 ± 0.102	0.435 ± 0.102	0.695 ± 0.095
Night	Standard	LR	Overall	LR	4	0.600 ± 0.101	0.590 ± 0.102	0.630 ± 0.100
Night	Standard	LR	Overall	ANN	4	0.600 ± 0.101	0.585 ± 0.102	0.625 ± 0.100
Night	Standard	LR	Overall	SVM	4	0.533 ± 0.103	0.460 ± 0.103	0.630 ± 0.100
Night	Standard	LR	Overall	ADA	4	0.578 ± 0.102	0.490 ± 0.103	0.680 ± 0.096
Night	Standard	LR	Best fold	LR	6	**0.656 ± 0.098**	0.635 ± 0.099	0.670 ± 0.097
Night	Standard	LR	Best fold	ANN	8	0.625 ± 0.060	0.616 ± 0.061	0.640 ± 0.060
Night	Standard	LR	Best fold	SVM	7	0.589 ± 0.102	0.525 ± 0.103	0.655 ± 0.098
Night	Standard	LR	Best fold	ADA	10	0.621 ± 0.061	0.625 ± 0.060	0.622 ± 0.061
Night	Standard	SVM	Overall	LR	5	0.578 ± 0.102	0.545 ± 0.103	0.625 ± 0.100
Night	Standard	SVM	Overall	ANN	5	0.533 ± 0.103	0.565 ± 0.102	0.530 ± 0.103
Night	Standard	SVM	Overall	SVM	5	0.600 ± 0.101	0.500 ± 0.103	**0.720 ± 0.093**
Night	Standard	SVM	Overall	ADA	5	0.500 ± 0.103	0.530 ± 0.103	0.515 ± 0.103
Night	Standard	SVM	Best fold	LR	9	0.589 ± 0.102	0.560 ± 0.103	0.625 ± 0.100
Night	Standard	SVM	Best fold	ANN	5	0.633 ± 0.100	**0.640 ± 0.099**	0.625 ± 0.100
Night	Standard	SVM	Best fold	SVM	5	0.645 ± 0.060	0.587 ± 0.062	0.704 ± 0.057
Night	Standard	SVM	Best fold	ADA	5	0.606 ± 0.061	0.527 ± 0.062	0.687 ± 0.058
Night	Adjusted	LR	Overall	LR	5	0.467 ± 0.103	0.490 ± 0.103	0.455 ± 0.103
Night	Adjusted	LR	Overall	ANN	5	0.456 ± 0.103	0.505 ± 0.103	0.415 ± 0.102
Night	Adjusted	LR	Overall	SVM	5	0.433 ± 0.102	0.380 ± 0.100	0.485 ± 0.103
Night	Adjusted	LR	Overall	ADA	5	0.467 ± 0.103	0.410 ± 0.102	0.540 ± 0.103
Night	Adjusted	LR	Best fold	LR	5	0.600 ± 0.101	0.605 ± 0.101	0.600 ± 0.101
Night	Adjusted	LR	Best fold	ANN	4	0.556 ± 0.103	0.530 ± 0.103	0.585 ± 0.102
Night	Adjusted	LR	Best fold	SVM	3	0.611 ± 0.101	0.600 ± 0.101	0.635 ± 0.099
Night	Adjusted	LR	Best fold	ADA	5	0.598 ± 0.061	0.470 ± 0.062	**0.716 ± 0.056**
Night	Adjusted	SVM	Overall	LR	6	0.522 ± 0.103	0.480 ± 0.103	0.555 ± 0.103

Night	Adjusted	SVM	Overall	ANN	6	0.533 ± 0.103	0.520 ± 0.103	0.535 ± 0.103
Night	Adjusted	SVM	Overall	SVM	6	0.522 ± 0.103	0.465 ± 0.103	0.575 ± 0.102
Night	Adjusted	SVM	Overall	ADA	6	0.522 ± 0.103	0.355 ± 0.099	0.700 ± 0.095
Night	Adjusted	SVM	Best fold	LR	7	0.544 ± 0.103	0.530 ± 0.103	0.565 ± 0.102
Night	Adjusted	SVM	Best fold	ANN	9	0.567 ± 0.102	**0.665 ± 0.098**	0.450 ± 0.103
Night	Adjusted	SVM	Best fold	SVM	6	**0.633 ± 0.072**	0.595 ± 0.086	0.660 ± 0.101
Night	Adjusted	SVM	Best fold	ADA	7	0.544 ± 0.103	0.455 ± 0.103	0.665 ± 0.098

IR = Incidence rate, LR = Logistic Regression, ANN = Artificial Neural Network, SVM = Support Vector Machine, ADA = Adaptive Boosting, GA = Genetic Algorithm.

Table M.2. *Performance measures in Twente (2010-16, 90 districts) as a result of the predictive modelling analysis, given the time of day, the type of incidence rate, the attribute selection method, the weights used as a result from the attribute selection method, as well as the learning algorithm.*

Appendix N. Settings of machine learning algorithms

This Appendix contains the settings of the Artificial Neural Network (ANN), Support Vector Machine (SVM) and Adaptive Boosting (AdaBoost) applied in this study.

ANN settings

Parameter	Setting
Type of neural network	Multilayer feed-forward network
Network algorithm	Backpropagation
Activation function	Sigmoid
Number of hidden layers	1
Number of neurons in the hidden layer	[3, 4, 5, 6]
Training cycles	500
Learning rate	[0.3, 0.5, 0.7, 0.9]
Momentum	[0.3, 0.5, 0.7, 0.9]
Weight decay	Yes

Table N.1. *Summary of the Artificial Neural Network's hyperparameters.*

SVM settings

Hyperparameters are tuned on a logarithmic scale.

Parameter	Setting
Kernel function	Radial basis function
Regularization parameter C	$[10^{-3}, 10^{-2}, 10^{-1}, 1, 10^{1}, 10^{-2}, 10^{3}]$
Scale parameter γ	$[10^{-3}, 10^{-2}, 10^{-1}, 1, 10^{1}, 10^{-2}, 10^{3}]$

Table N.2. *Summary of the Support Vector Machine's hyperparameters.*

AdaBoost settings

We have more districts and therefore training instances in North Holland compared to Twente. Therefore, we choose slightly lower values for the minimal leaf size and the minimal size that allows for a split to be made.

Parameter	Setting
Iterations	10
Base learner	Decision Tree
Criterion	Information gain
Maximal depth	6
Pre-pruning	Yes
Minimal gain	0.01
Minimal leaf size	NH: 6; TW: 4
Minimal size for split	NH: 6; TW: 4
Pruning	Yes
Confidence level in pruning	0.1

Abbreviations: NH = North Holland, TW = Twente.

Table N.3. *Summary of the hyperparameter values used in the AdaBoost algorithm.*

Appendix O. Linear discriminant functions

Linear discriminant functions use hyperplanes to subdivide a D-dimensional input space into decision regions. These decision boundaries are modelled as linear functions of the input vector $\mathbf{x}$, which corresponds to the attribute values. Linear classifications functions are modelled as $(D - 1)$-dimensional hyperplanes within the defined input space. The formula $y(\mathbf{x})$ for a hyperplane is as follows (Bishop, 2006):

$$y(\mathbf{x}) = \sum_{j=0}^{D} w_j x_j = \mathbf{w}^{\mathrm{T}}\mathbf{x} = 0 \tag{O.1}$$

Here, $\mathbf{x}$ corresponds to the vector of attribute values (including x_0) and $\mathbf{w}$ to the attribute weight vector. Similar to the offset in linear regression, x_0 refers to the bias and always contains the value $x_0 = 1$. Its purpose is to ensure that $\mathbf{w}$ always starts in the origin. The slope of the hyperplane is then defined by the values of $\mathbf{w} = (w_1 \ldots w_A)$. An example of a hyperplane in a two-dimensional input space is given in Figure O.1.

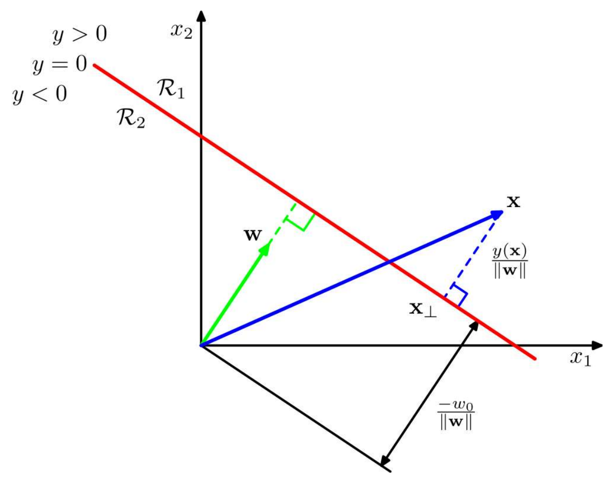


Figure O.1. *Geometry of a linear discriminant function. The hyperplane is illustrated in red, the weight vector* $\mathbf{w}$ *in green, and the vector to point* $\boldsymbol{x}$ *in the input space in blue. From: Bishop (2006).*

Appendix P. CONFIDENTIAL - Districts with lowest values for resuscitation efforts