Tilburg University

Healthcare expenditure prediction with neighbourhood variables

Mohnen, S. M.; Rotteveel, A. H.; Doornbos, G.; Polder, J. J.

[Link to publication in Tilburg University Research Portal](#)

Sigrid M. Mohnen, Adriënne H. Rotteveel*, Gerda Doornbos
and Johan J. Polder

Healthcare Expenditure Prediction with Neighbourhood Variables – A Random Forest Model

Abstract: We investigated the additional predictive value of an individual's neighbourhood (quality and location), and of changes therein on his/her healthcare costs. To this end, we combined several Dutch nationwide data sources from 2003 to 2014, and selected inhabitants who moved in 2010. We used random forest models to predict the area under the curve of the regular healthcare costs of individuals in the years 2011–2014. In our analyses, the quality of the neighbourhood before the move appeared to be quite important in predicting healthcare costs (i.e. importance rank 11 out of 126 socio-demographic and neighbourhood variables; rank 73 out of 261 in the full model with prior expenditure and medication). The predictive performance of the models was evaluated in terms of R^2 (or proportion of explained variance) and MAE (mean absolute (prediction) error). The model containing only socio-demographic information improved marginally when neighbourhood was added (R^2 +0.8%, MAE −€5). The full model remained the same for the study population (R^2 = 48.8%, MAE of €1556) and for subpopulations. These results indicate that only in prediction models in which prior expenditure and utilization cannot or ought not to be used neighbourhood might be an interesting source of information to improve predictive performance.

*Corresponding author: Adriënne H. Rotteveel, MSc, National Institute for Public Health and the Environment (RIVM), Centre for Nutrition, Prevention, and Health Services, PO Box 1, 3720 BA, Bilthoven, the Netherlands; Erasmus University Rotterdam, Erasmus School of Health Policy and Management, Rotterdam, the Netherlands; and University Medical Center Utrecht, Julius Center for Health Sciences and Primary Care, Utrecht, the Netherlands, E-mail: adrienne.rotteveel@rivm.nl. https://orcid.org/0000-0002-3395-0147
Dr. Sigrid M. Mohnen and Gerda Doornbos, National Institute for Public Health and the Environment (RIVM), Centre for Nutrition, Prevention, and Health Services, Bilthoven, the Netherlands, e-mail: sigrid.mohnen@rivm.nl (S.M. Mohnen), gerda.doornbos@rivm.nl (G. Doornbos). https://orcid.org/0000-0002-1537-8706 (S.M. Mohnen)
Johan J. Polder, National Institute for Public Health and the Environment (RIVM), Centre for Health and Society, Bilthoven, the Netherlands; and Tilburg University, Department Tranzo, Tilburg, the Netherlands, e-mail: johan.polder@rivm.nl

1 Introduction

This paper aims to improve the prediction of healthcare costs by introducing a new level: the neighbourhood. Our research interest in the predictive value of neighbourhood is based on the rich literature on neighbourhood health effects, i.e. characteristics of small geographical areas are associated with health status of inhabitants. Although causality is not guaranteed, we may assume that the neighbourhood affects health (Diez Roux and Mair 2010; Ellen and Turner 2003; Sampson, Morenoff, and Gannon-Rowley 2002) and since health is a major determinant of healthcare demand, we hypothesize that the exposure to the neighbourhood translates from healthcare demand to healthcare utilization and finally to healthcare costs.

In this study, we were interested in: 1) the importance of the neighbourhood quality and location compared to other variables in the prediction of healthcare costs, and 2) the improvement in the predictive performance when adding neighbourhood to an elaborate prediction model. Improving the prediction of healthcare expenditure is relevant for risk adjustment. An application of risk adjustment is risk-adjusted subsidies, which are used in competitive health insurance markets to prospectively compensate insurers for differences in case mix (i.e. in the people they insure). This compensation is necessary to reduce incentives for risk-rating, i.e. asking higher premiums for more expensive insured, and risk selection, i.e. using different means, such as advertisement, to contract the most cheap insured (Van de Ven 2011; Van de Ven et al. 2013). Another application of risk adjustment is capitation payment. Capitation payments are used to pay healthcare providers, and consist of a periodical lump sum per patient. Risk adjustment is used to differentiate capitation payments based on patient characteristics. This is necessary to prevent risk selection by providers, and to prevent under compensation of specialized providers treating mainly complex, expensive patients (Jegers et al. 2002; Shin, Schumacher, and Feess 2017). Risk-adjusted subsidies and differentiated capitation payments are calculated using risk adjustment models containing predictors of healthcare expenditure, such as demographic variables, regional variables, health status indicators, and prior healthcare expenditure and utilization (Shin, Schumacher, and Feess 2017; Van de Ven et al. 2007; Van Veen et al. 2015;). Despite the large set of variables included in risk adjustment models, these models still undercompensate insurers/healthcare providers for certain types of insured/patients (Buchner, Wasem, and Schillo 2017; Eijkenaar, van Vliet,

and van Kleef 2018; Sibley and Glazier 2012; Van Veen et al. 2017). For this reason, it is important to find new variables for risk adjustment models that improve the compensation for expensive insured/patients. This study gives insight in whether neighbourhood variables may be of additional value for risk adjustment models.

Furthermore, we like to study the predictive value of the neighbourhood because it could improve matching in observational studies. Our study might improve the accuracy of propensity scores which are most often used for matching to reduce imbalance in the distribution of the pre-treatment characteristics of the intervention and the control group (Stuart 2010).

First, this section proceeds with a subsection on the theoretical background in which we explain in short why the neighbourhood might matter for the prediction of healthcare expenditure. Subsequently, Section 2 describes the methods of our study, Section 3 describes the results and Section 4 discusses the implication of the results.

1.1 Theoretical Background

In this paper we like to emphasize that human beings are social beings, living their lives in a certain context, not in a laboratory (Barker 1968). Moreover, environmental inputs that are relevant to health, such as pollution control, greater public safety, expanded opportunities to improve physical fitness, or improved social housing, are beyond the control of a single individual (Leibowitz 2004). In this 'ecological approach', (Macintyre and Ellaway 2000; Sallis and Owen 2015; Sallis et al. 2006) applied in Public Health research, people and their health are studied within a physical and social environment, the neighbourhood. Hence, next to the direct effect of the healthcare system and individual characteristics, also the environment surrounding the individual is likely to predict need and utilization. Firstly, neighbourhoods might differ in distance, reachability, accessibility, opening hours as well as quantitative and qualitative characteristics of healthcare facilities (Ellen, Mijanovich, and Dillman 2001). Secondly, neighbourhoods might physiologically, thus directly affect an individual's health with a dose response relationship. Thirdly, neighbourhoods might also affect health indirectly via a psychological pathway or via health behavioural (Berkman et al. 2000). An example for the psychological pathway is the short-term restorative effect of contact with nature (e.g. green space) (Hartig et al. 2014) and its association with good perceived mental health (Van den Berg et al. 2015). An example of the behavioural pathway is, that walkable, social, or safe neighbourhoods provide more opportunities for physical activity which supports good health (Haskell et al. 2007) and health-related quality of life (Bize, Johnson, and Plotnikoff 2007). Lastly,

neighbourhood might affect healthcare utilization independent of need, i.e. neighbourhoods might differ in their level of neighbourhood social capital (and this might differently motivate people to demand and finally use preventive healthcare, e.g. screening for colorectal cancer (Leader and Michael 2013), preventive dental visits (Iida and Rozier 2013), and number of contacts with doctors (Nguyen, Ho, and Williams 2011)).

Pathways help to understand why neighbourhoods have the ability to harm and benefit health with consequences for the demand of healthcare (Mohnen and Schneider 2019). For example, it should be good for one's healthcare demand to live in a green neighbourhood as green space is associated with lower medical care use in Korea (Lee, Lee, and Kwon 2014) and less visits to mental health specialists and intake of mental health medication in Spain (Lee, Lee, and Kwon 2014). Furthermore, it should be bad for one's healthcare demand to live in neighbourhoods with air pollution as high levels of nitrogen dioxide are associated with premature birth (WHO 2013) and hospital admission for respiratory and cardiovascular symptoms (Dijkema et al. 2016). Another example of the negative influence of the neighbourhood on healthcare demand is the association between self-perceived neighbourhood disorder[1] and total health services usage (Martin-Storey et al. 2012). In reality, neighbourhood characteristics interact which makes it difficult to study the effect of a single neighbourhood characteristic. For example, playgrounds were only associated with a higher level of physical activity in adolescents in combination with a high level of neighbourhood social capital (Prins et al. 2012). Because it is difficult to study the effect of a single neighbourhood characteristic, we used an aggregate measure, the livability index of 2008, to differentiate between good and bad neighbourhoods. In this index, 49 items of social and physical neighbourhood characteristics were used to measure the quality of Dutch neighbourhoods (Leidelmeijer et al. 2009).

Next to the neighbourhood location, we used the quality of the neighbourhood (i.e. livability) as prediction variable because we assumed that the quality of the neighbourhood – possibly more than the location - matters for the need for healthcare and thus for healthcare utilization and expenditure. To understand the relevance of using the quality of the neighbourhood as a prediction variable, we compared the importance of this variable with other, often-used prediction variables, e.g. age, gender, income and occupation (Shin, Schumacher, and Feess 2017; Van de Ven et al. 2007).

---

1 Neighborhood disorder are physical and social features of neighborhoods that may signal the breakdown of order and social control, for example, prostitution, drug dealing, fighting in the streets, abandoned cars, vandalized buildings, or litter in the streets (Marco et al. 2015).

The added value of a variable for a prediction model depends on the other prediction variables in the model. Therefore, we tested whether, next to socio-demographic characteristics, neighbourhood quality and location improved the prediction of regular healthcare costs, and if so, whether this added value vanished when prior expenditure and prior medication utilization were added to the model. For all these analyses, we conducted sensitivity analyses with *outcomes* that are expected to be more sensitive to neighbourhood effects (i.e. General practitioner (GP) consultation costs and medication utilization) and in chronically ill *subgroups* that are expected to be more sensitive to neighbourhood effects (i.e. diabetes type II, mental health and obstructive airway disease).

2 Material and Methods

2.1 Study Design

To test whether the neighbourhood in which an individual lives can predict individuals' healthcare expenditure, we followed individuals who moved (=movers). If the neighbourhood matters for healthcare expenditure we should find that the neighbourhood someone was exposed to for several years is an important prediction variable. Furthermore, and this is why we chose to work exclusively with movers, if a change in the quality of the neighbourhood (e.g. moving to a better quality neighbourhood) is of value for the prediction this would give a stronger indication that neighbourhood matters for prediction. In our study design, we aimed to minimize the effects of the supply side by following movers that changed neighbourhood but not healthcare supplier, by only including movers within a hospital catchment area (see Appendix A for information on Dutch hospital catchment areas).

2.2 Data

We combined several nationwide data sources. Below, we describe the data sources and Appendix B gives a complete overview of all prediction variables, with their data source and value labels. Via Statistics Netherlands (CBS), we had access to non-public microdata. This data was linked at the individual and neighbourhood level and encompasses the entire Dutch population. Anonymised data were analysed in a secure remote-access environment of CBS. Neighbourhood was operationalised using the neighbourhood code of CBS, a smaller and more precise operationalisation of the neighbourhood than 4-digit postal codes. In 2010, on average, 1418 (SD: 2000) people were living in each CBS neighbourhood.

2.2.1 Socioeconomic Variables (2003–2015): We used municipal register data including home address, relocation date, and socio- demographic characteristics, e.g. country of origin, marital status. CBS microdata also includes size of household and socio-economic status (occupation type and household income before tax).

2.2.2 Annual Health Insurance Claims Expenditure (2008–2014): In the Netherlands, a basic health insurance is obligatory by law, therefore almost all (99%) Dutch citizens have a basic health insurance (NZa 2016). The healthcare information centre Vektis collects and manages health claims of all Dutch health insurance companies on all healthcare procedures covered by the Health Insurance Act, including the costs of compulsory co-payments and deductible, excluding other out-of-pocket payments (de Boo 2011). The Vektis database covers 99% of all insured people. Vektis aggregated expenditures of claims per person, year and care category. Categories were the curative healthcare expenditures of primary and secondary care, prescribed medication, medical aids, patient transportation, and mental healthcare.

2.2.3 ATC Medication Codes (2005–2014): Based on claims data, the National Health Care Institute makes a yearly overview of prescribed medication per inhabitant, based on Anatomical Therapeutic Chemical (ATC) classification. CBS microdata included these ATC codes on 4-digit level, with a single code per person per year. Volume and actual intake of medication was not available. We were not able to differentiate between someone with missing values and someone with no prescribed medication.

2.2.4 'Livability Index of the Neighbourhood' (2008): Under the authority of the former Ministry of Housing, Spatial Planning, and Environment (VROM) the 'Livablity index of the neighbourhood' was developed based on scientific literature and empirical data. The index consists of 49 items from six disciplines: (1) housing, (2) public space, (3) public facilities, (4) composition of inhabitants (SES and ethnicity), (5) composition of inhabitants in terms of age, household size, and residential stability, and (6) Public safety (Leidelmeijer et al. 2009). All data were measured at 1.1.2008, except for environmental noise, measured in 2006, and a part of the dimension 'public space'. Uninhabited or very sparsely populated industrial and rural areas were not part of the index, and had a 'missing value' in this study. Content validity of the index was determined by a check by local policy makers of the scores of the neighbourhoods in the municipality they were responsible for (Leidelmeijer et al. 2008). The livability variables measure the quality of the neighbourhood a person lived in before (livability_pre) and after (livability_post) their move in 2010. The improvement_in_move variable measures whether the quality of the neighbourhood improved after the move compared to before the move.

2.3 Study Population

We had access to all registered citizens living between 2005 and 2015 in the Netherlands (n = 21,559,510). Since the health claims data are only available from 2008 onwards, we decided to analyse the event of moving in the year 2010. Each registered inhabitant lives in an object and each of these objects has a unique object number. We interpret a change in object number as a move to another address. We included people in our analysis who have been living stable (= on the same address) between 2005 and 2009, were born before 1 January 2005 and were not deceased in 2005–2010 (n = 14,981,058). We only included people who moved once in 2010 (n = 478,462) within a hospital catchment area (n = 310,653). People with incomplete municipality registration data between 2005 until the end of 2010 were not part of the analyses. Reasons for the gaps in the registration were death, moving abroad, losing permanent home, or registration errors. When these gaps occurred in the follow-up period (2010–2014), we used the information of the movers until the gap and ignored the information after the start of the gap. Finally, we selected only one

person per household (n = 207,614). Sensitivity analyses were conducted for the chronically ill subgroups diabetes type II (n = 9496), mental health disease (n = 20,337) and obstructive airway disease (n = 20,124). See Appendix C for subgroups definitions.

2.4 Dependent Variables

The dependent variable was the average over a fixed period of the annual individual's regular healthcare costs, which included all costs that were covered by the basic health insurance in the Netherlands. It included the deductible costs but excluded both intramural mental healthcare costs and out-of-pocket payments. The average healthcare cost was defined as the area under the polygonal curve of an individual's healthcare expenditures during the years 2011–2014, computed with the trapezoidal rule, divided by the length (i.e. number of years) of the period of observation. For individuals with missing data it was computed over a smaller number of points and/or shorter period. This definition of individual average costs allows us to include people with different lengths of follow-up and even deceased people. Next to being generally convenient from a statistical point of view, including more data has the important advantage of creating a more diverse population, that is more similar to the general population rather than a distinct - probably healthier - subsample.

The dependent variables for the sensitivity analyses were 1) the annual individual's GP consultation costs and 2) the annual individual's sum of ATC codes. Both variables were calculated in a similar way to regular healthcare costs from the area under the curve during the years 2011–2014. GP consultation costs included all costs for GP visits that are covered by the basic health insurance. The sum of ATC codes were the number of different level-4 ATC groups per individual in a year. All costs are reported in Euros (1 Euro = 1.1045 US dollar – exchange rate of 11 September 2019).

2.5 Method: Random Forest Models Statistics and Variable Importance

In this study, random forest was used to predict healthcare utilization and costs of individuals. Random forest (Breiman 2001; Hastie, Tibshirani, and Friedman 2009) is a machine learning or statistical prediction algorithm that generates and in some sense averages the predictions of a large number of 'decision trees'. Random forest is well established as a useful statistical tool and it is increasingly applied in prediction problems because of its flexibility and prediction accuracy. In particular, random forest can cope with many predictor variables (covariates) of various kinds (numerical, ordinal or categorical), collinearity of predictor variables or unusual distributional forms (e.g. asymmetry or lack of normality), and tends to show up among the most accurate prediction methods in comparative prediction studies (Shrestha et al. 2018).

2.5.1 Error Statistics: We used the package 'ranger' (Wright and Ziegler 2017) of the open source statistical software R (R 3.5.1) to produce output such as the mean and median prediction errors, MAE (mean/median *absolute* error), or the average/median absolute difference between the actual and the predicted values of the outcome of interest, R^2 or PEV (proportion of explained variance), which is defined by 1-MSE[2]/Var (outcome) and normally assumes values between 0 and 1, higher values indicating a greater usefulness of the predictor variables; as measures of prediction accuracy.

---

2 MSE is the mean squared error; or the average squared difference between actual and predicted values.

2.5.2 Variable Importance: In addition, random forest produces a ranking of the predictor variables in terms of the 'importance' they have for producing predictions. Roughly speaking, the importance of a variable is proportional to the worsening – namely the relative increase in MSE[2] – of the prediction error that results from permuting the values of that variable randomly in the data set. If a variable is irrelevant for predicting, replacing the value of that variable for an individual by an arbitrary value will hardly affect the prediction for that individual; if on the contrary the variable really matters for prediction then 'confusing' the variable will tend to worsen the predictions substantially.

Variable importance was used in this study to assess whether and to what extent neighbourhood variables play a role in the prediction of healthcare costs. To understand the role of neighbourhood in the prediction of regular healthcare costs several models, summarized in Table 1, were used for comparison (see Appendix B for a list of all variables). In each model, 1000 trees were built.

3 Results

3.1 Descriptive Information

The study population was compared to the Dutch population on pre-move annual healthcare expenditures (Table 2) and socio-demographic variables (Appendix D). The average regular healthcare costs in the study population were €2156 and the average GP consultation costs were €58, which was slightly higher than in the Dutch population (regular costs: €1763, GP consultation costs: €46). The study population used medications from on average 3.3 different ATC groups, which was slightly lower than the Dutch population (3.9). The regular healthcare cost of the study population remained quite stable between 2011 and 2014 (i.e. the years used

Table 1: Overview of prediction models.

Model	Prediction variables	Number of prediction variables
Model 1	All socio-demographic information available	116
Model 2	Model 1 + neighbourhood	126
Model 3	Model 1 + prior expenditure and medication use	251
Model 4 = full model	Model 3 + neighbourhood	261[a]

[a]Model 4 on GP consultation costs has one extra variable: area under the curve of GP consultation cost before the move.

for the dependent variable) with averages of €2217 in 2011, €2165 in 2012, €2234 in 2013 and €2228 in 2014; which is in line with the average regular costs of the Dutch

Table 2: Healthcare costs (€) and sum of ATC codes of study population (and sub populations) vs. Dutch population in 2009.

	Regular costs[a]					GP consultation costs					Sum of ATC codes				
				Quartiles					Quartiles					Quartiles	
	Mean	St.dev.	Median	1st	3rd	Mean	St. dev.	Median	1st	3rd	Mean	St. dev.	Median	1st	3rd
NL population n = 16,570,523	1763	5424	397	135	1326	46	62	27	9	61	3.9	3.4	3.0	1.0	5.0
Study population n = 208,399	2156	6354	463	146	1659	58	81	32	9	77	3.3	3.7	2.0	1.0	5.0
Subgroup: Diabetes type 2[b] n = 9496	6377	9800	3170	1641	7158	166	124	136	81	218	9.2	4.6	9.0	6.0	12.0
Subgroup: Mental health disease[c] n = 20,337	4894	10186	1922	677	5115	139	125	105	55	185	7.2	4.8	6.0	3.0	10.0
Subgroup: Obstructive airway disease[d] n = 20,124	4639	9350	1726	652	4673	121	116	86	45	162	7.5	4.9	6.0	4.0	10.0

[a] Includes all costs that are covered by the basic benefit package in the Netherlands (including the deductible, excluding other out-of-pocket payments), but no intramural healthcare costs for mental health.
[b] Users of Blood glucose lowering drugs.
[c] Users of calming and arousing mental health drugs.
[d] Users of medication for obstructive airway disease.

population (2011: €1866, 2012: €1861, 2013: €1943, 2014: not available). Average GP consultation costs remained quite stable as well (€45 in 2011, €42 in 2012, €41 in 2013 and €43 in 2014). Average sum of ATC decreased slightly (3.3 in 2011, 3.1 in 2012, 3.0 in 2013 and 3.0 in 2014). The average age of the study population was higher than in the Dutch population (43.1 vs. 39.9 years) and the percentage of males was slightly lower (48.3 vs. 49.5%). The mortality in 2011–2015 was higher than in the Dutch population (2.5 vs. 0.8%). Furthermore, less people were married or had a registered partner (25 vs. 41%) and the household income was higher (€39,493 compared to €23,300) than in the Dutch population.

Unsurprisingly, the chronically ill subpopulations had clearly higher regular healthcare costs (diabetes: €6,377; mental health: €4,894; obstructive airway: €4,639) and GP consultation costs (diabetes: €166; mental health; €139, obstructive airway; €121) than the whole study population. The amount of ATC groups used was also clearly higher in the chronically ill (diabetes: 9.2; mental health: 7.2; obstructive airway: 7.5). People with diabetes type 2 were older (average 72.7 years) than people in the other chronically ill subgroups (average 56.3 and 52.3, respectively) and the whole study population (average 43.1). Furthermore, they were more often married or widowed, were more often pensioners, and had lower household incomes than the other subpopulations and the study population. The subpopulation with mental health problems was more often a recipient of some kind of welfare benefits compared to the other subpopulations and the study population. The subpopulation with obstructive airway diseases was quite comparable to the study population on all socio-demographic variables reported in Appendix D.

In 2008 on average 165,735 (SD: 66,293; Range: 25,285–388,945) people were living in the 103 hospital catchment areas in the Netherlands and each catchment area consisted on average of 127.7 (SD: 64.5; Range: 14–353) neighbourhoods (Appendix A).

3.2 Random Forest Model Results

3.2.1 Quality vs. Location of the Neighbourhood

In the model with socio-demographic and neighbourhood variables, the quality of the neighbourhood mattered more for the prediction of regular costs than the location with an importance value of 62.1 and 61.2 for livability pre and post and an importance value between 40 and 50 for the location variables (Figure 1). Neighbourhood quality and location were equally important for the prediction of GP consultation costs and sum of ATC codes (Appendix E and F). In the Full model, the quality of the neighbourhood was equally important as the location in the

prediction of regular costs, GP consultation costs and sum of ATC codes (Figure 2, Appendix G and H). A change in the exposure to neighbourhood quality (i.e. improvement in move) was of some importance in model 2 (with importance ranks of 26–31 out of 126) but low ranked in the Full model (102–132 out of 261, Figures 1 and 2, and Appendix E–H).

3.2.2 Quality of the Neighbourhood in Perspective

In the prediction model with socio-demographic and neighbourhood variables, the quality of the neighbourhood (livability_pre) appeared to be an important predictor with ranks of 14–17 out of 126 for regular costs, GP consultation costs and sum of ATC codes (Figure 1 and Appendix E and F). In these models, the quality of the neighbourhood was equally (or more) important as age in predicting all three dependent variables (Figure 1 and Appendix E and F). This was not the case in the Full model. In the Full model on regular costs the importance rank of neighbourhood quality dropped to 73 out of 261 (Figure 2). In this model, age was twice as important as the quality of the neighbourhood (Figure 2). In the models on GP consultation costs (Appendix G) and sum of ATC codes (Appendix H), age was the most important variable and was 2–3 times as important as livability.
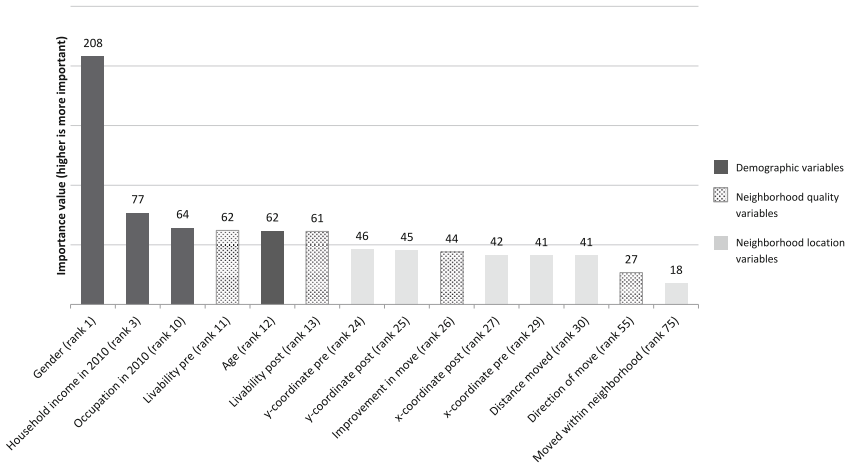

Figure 1: Random forest importance values (& ranks) of all neighbourhood variables and a selection of socio-demographic variables in Model 2 on regular costs (in total 126 variables).
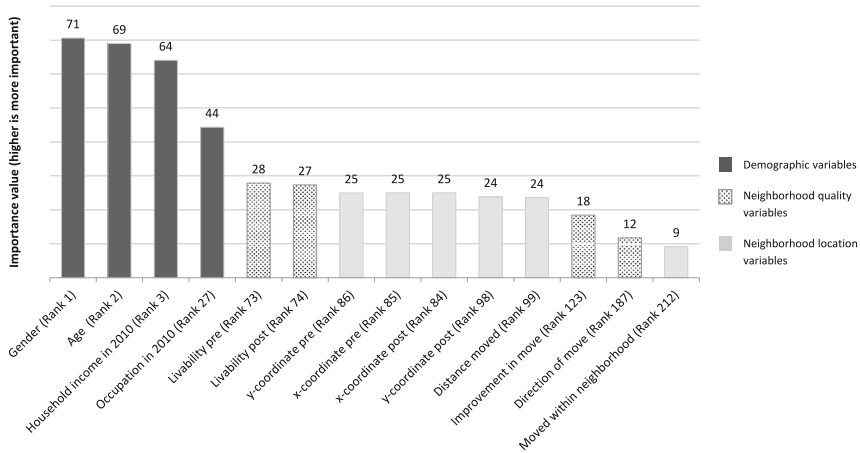
Figure 2: Random forest importance values (& ranks) of neighbourhood variables and four selected socio-demographic variables of model 4 on regular costs (in total 261 variables).

3.2.3 Neighbourhood Less Important for Chronically Ill than for the Study Population.

In the prediction model on regular healthcare costs, the importance ranks of the livability_pre variable in the chronically ill subpopulations were above 142 and higher (Appendix I), indicating a lower importance of these variables for chronically ill than for the whole study population (rank 73). A similar pattern was found for GP consultation costs (Rank 95 and higher vs. 81) and sum of ATC codes (Rank 159 and higher vs. 70).

3.2.4 Small Additional Value of Neighbourhood Variables Next to Socio-demographic Variables

When the neighbourhood variables were added to a prediction model with a rich set of socio-demographic information (comparing Model 1 and 2, Table 3), the R^2 of the prediction model on regular costs increased with 0.8%. Furthermore, mean and median absolute prediction error improved (i.e. error decreased) with €5 and €4, respectively. Prediction error showed contradicting results, with a deterioration of mean prediction error of €12 (error increased) and an improvement in median prediction error of €1 (error decreased). The dependent variables that were chosen because they might be more sensitive to neighbourhood effects (i.e. GP

Table 3: Predictive performance of adding neighbourhood to the prediction models for the three outcome variables; applied to the study population and subpopulations.

Model	Prediction variables	R^2 (%)[a]	Prediction error (€/N)[b]				Absolute prediction error (€/N)[§]	
			Mean	Median	Percentiles		Mean	Median
					0.05	0.95		
Regular costs (main outcome)								
Study population								
Model 1	All socio-demographic information	26.9	1326	−4	−1349	7732	1864	500
Model 2	Model 1 + neighbourhood	27.7	1338	−3	−1292	7724	1859	496
Model 3	Model 1 + prior expenditure and medication use	48.8	1048	−27	−1175	6225	1556	405
Model 4	Model 3 + neighbourhood (full model)	48.8	1051	−28	−1167	6231	1556	404
Subgroup: Diabetes type 2								
Model 3	Model 1 + prior expenditure and medication use	34.4	2543	301	−2902	14485	3855	1699
Model 4	Model 3 + neighbourhood (full model)	34.6	2549	314	−2929	14354	3856	1681
Subgroup: Mental health disease								
Model 3	Model 1 + prior expenditure and medication use	42.4	1880	125	−1992	10936	2723	943
Model 4	Model 3 + neighbourhood (full model)	42.4	1889	131	−1976	10942	2724	947
Subgroup: Obstructive airway disease								
Model 3		49.8	1866	43	−2386	11283	2853	946

Table 3: (continued)

Model	Prediction variables	R^2 (%)[a]	Prediction error (€/N)[b]				Absolute prediction error (€/N)[§]	
			Mean	Median	Percentiles		Mean	Median
					0.05	0.95		
	Model 1 + prior expenditure and medication use							
Model 4	Model 3 + neighbourhood (full model)	49.8	1871	40	−2428	11421	2859	949
GP consultation costs (secondary outcome)								
Study population								
Model 1	All socio-demographic information	23.2	19	4	−25	112	28	14
Model 2	Model 1 + neighbourhood	24.9	19	4	−24	112	29	14
Model 3	Model 1 + prior expenditure and medication use	48.4	15	3	−22	91	24	11
Model 4	Model 3 + neighbourhood (full model)	48.8	15	3	−22	91	24	11
Subgroup: Diabetes type 2								
Model 3	Model 1 + prior expenditure and medication use	43.1	38	134	−43	194	55	29
Model 4	Model 3 + neighbourhood (full model)	43.2	38	134	−43	192	55	29
Subgroup: Mental health disease								
Model 3	Model 1 + prior expenditure and medication use	44.3	31	9	−36	168	45	22
Model 4	Model 3 + neighbourhood (full model)	44.5	31	9	−36	168	45	22

Table 3: (continued)

Model	Prediction variables	R^2 (%)[a]	Prediction error (€/N)[b]				Absolute prediction error (€/N)[§]	
			Mean	Median	Percentiles		Mean	Median
					0.05	0.95		
	Model 3 + neighbourhood (full model)							
Subgroup: Obstructive airway disease								
Model 3	Model 1 + prior expenditure and medication use	43.5	27	9	−38	158	42	20
Model 4	Model 3 + neighbourhood (full model)	43.6	27	9	−37	156	42	20
Sum ATC codes (secondary outcome)								
Study population								
Model 1	All socio-demographic information	36.4	0.8	<−0.1	−2.8	7.2	2.0	1.1
Model 2	Model 1 + neighbourhood	37.5	0.8	<−0.1	−2.7	7.2	2.0	1.1
Model 3	Model 1 + prior expenditure and medication use	68.2	0.5	<0.1	−1.7	4.4	1.3	0.7
Model 4	Model 3 + neighbourhood (full model)	68.2	0.5	<0.1	−1.7	4.4	1.3	0.7
Subgroup: Diabetes type 2								
Model 3	Model 1 + prior expenditure and medication use	52.1	1.4	1.1	−4.2	7.7	2.9	2.1
Model 4	Model 3 + neighbourhood (full model)	52.2	1.5	1.1	−4.1	7.8	2.9	2.1
Subgroup: Mental health disease								

Table 3: (continued)

Model	Prediction variables	R^2 (%)[a]	Prediction error (€/N)[b]				Absolute prediction error (€/N)[§]	
			Mean	Median	Percentiles		Mean	Median
					0.05	0.95		
Model 3	Model 1 + prior expenditure and medication use	57.7	1.1	0.5	−3.0	7.1	2.4	1.5
Model 4	Model 3 + neighbourhood (full model)	57.8	1.1	0.5	−3.0	7.1	2.4	1.5
Subgroup: Obstructive airway disease								
Model 3	Model 1 + prior expenditure and medication use	62.6	1.0	0.5	−3.0	6.8	2.3	1.4
Model 4	Model 3 + neighbourhood (full model)	62.6	1.0	0.5	−3.0	6.8	2.3	1.4

[a]R^2 = 1 − (sum of squares of residual expenses/sum of squares of total expenses).
[b]Prediction Error = mean/median/percentiles of the prediction error (i.e. actual costs − predicted cost); in Euros for regular costs and GP consultation costs; in numbers for sum of ATC codes.
[§]Absolute prediction Error = mean/median of the absolute prediction error (i.e. |actual costs − predicted cost|); in Euros for regular costs and GP consultation costs; in numbers for sum of ATC codes.

consultation costs and sum of ATC codes), did not substantially benefit more from adding the neighbourhood variables (Table 3: GP consultation costs R^2 +1.7%; sum ATC codes R^2 +1.1%). This indicates that the additional value of the neighbourhood variables next to socio-demographic information in predicting regular costs, GP consultation costs and sum of ATC codes was small.

3.2.5 No Additional Value of Neighbourhood Variables in the Full Model

Neighbourhood variables used in this study had no additional value in predicting healthcare expenditures next to a rich set of socio-demographic variables and prior healthcare expenditures and medication (Table 3). The dependent variables that were chosen because they might be more sensitive to neighbourhood effects (i.e. GP consultation costs and sum of ATC codes), did not benefit from adding the neighbourhood variables to the model as well. The subpopulations that were chosen because they might be more sensitive to neighbourhood effect (chronically ill of diseases known for its link with the neighbourhood) showed similar results. Furthermore, sensitivity analyses within three different age groups and within females and males also showed no additional predictive value of neighbourhood (Appendix J). Besides, in Appendix K, we calculated differences in prediction error for different groups of people. Categories were ethnic background, household income, occupation, having one of three chronic diseases, patients with multiple diseases, health care utilization (specialist care and mental healthcare) and people with healthcare expenditures in the top 25% in the past 2 years. These results showed no improvement in prediction error for any of these groups.

3.2.6 Accuracy of Prediction

In the full model on the study population, Random Forest models showed an R^2 of 48.8%, a mean absolute prediction error of €1556, and a median absolute prediction error of €404 for predicting regular costs (Table 3).

The predictive performance of the full model on regular costs was lower in the subpopulation with diabetes type 2 (R^2: 34.6, mean & median absolute prediction error: €3855 & €1699) and in the subpopulation with mental health disease (R^2: 42.4, mean & median absolute prediction error: €2724 & €947,) compared to the study population (Table 3). In the subpopulation obstructive airway disease, the R^2 was higher (49.6) while the mean and median absolute prediction errors were higher (€2859, €949) than in the study population.

The R^2 for predicting GP consultation costs (48.8) was similar to the R^2 for predicting regular costs for the full model in the study population. The R^2 for predicting sum of ATC codes was higher (68.2) than the R^2 for predicting regular

costs. The mean and median absolute prediction error for GP consultation costs were €24 and €11, respectively, and for sum of ATC codes 2.0 and 1.1, respectively.

4 Discussion

The aim of this study was to explore the additional predictive value of using neighbourhood variables next to other commonly used variables to predict healthcare costs. As we followed movers in time, we could not only study the quality and the location of the neighbourhood but also whether someone moved to a 'better' neighbourhood and whether this information helps to predict healthcare costs in the three years following a move to a new address within an hospital catchment area.

In this study, we found that the quality of the neighbourhood was in general more important in predicting healthcare costs than the location of the neighbourhood. To put the importance of the quality of the neighbourhood into perspective, we showed that it is equally important as age in the prediction of healthcare costs with a prediction model containing socio-demographic and neighbourhood variables. However, in a prediction model to which prior expenditure and medication were added, the importance rank of the quality of the neighbourhood dropped, while the importance rank of age increased, making age much more important than neighbourhood in this model. Besides, our study showed that a change to a 'better' neighbourhood is not important for the prediction of healthcare utilization and costs.

Furthermore, in this study we found that, only when adding neighbourhood to the prediction model with socio-demographic information the predictive performance slightly improved. No improvement in predictive performance was observed when adding neighbourhood to the prediction model with socio-demographic information, and prior expenditure and medication use. Sensitivity analyses showed same results for different outcome variables and subpopulations. Hence, the neighbourhood is only of additional value for prediction models in contexts in which data on prior healthcare utilization and expenditure cannot or ought not to be used.

Finally, this study demonstrated that random forest is an important tool for variable screening for healthcare expenditure prediction while producing a high R^2. The high accuracy of prediction suggests (1) that we have used interesting variables for the prediction and (2) that the random forest method was able to discover underlying interactions which traditional methods (e.g. OLS) are not able to find. The latter is in line with Shrestha et al. (2018) who showed that Random

forest models can outperform more traditional OLS regressions in healthcare prediction (Shrestha et al. 2018).

4.1 Strengths and Limitations

Since the decision to move and where to move was in the hands of the movers themselves, a limitation of our study is that we did not study the effect of a 'natural experiment' (Craig et al. 2012). In a real natural experiment, movers would have to move randomly. An example of a real natural experiment is the 'Moving to Opportunity' (MTO) study, where people moved randomly from one neighbourhood to another (Katz, Kling, and Liebman 2001). No experiment in this kind exists in the Netherlands. Hence, because of selection biases causality cannot be proven. However, by using a prediction model we were able to study the value of the neighbourhood in the prediction of healthcare utilization and expenditure.

Following movers in time enables studying neighbourhood effects because people were exposed to different neighbourhoods. However, a move might also go along with a change in healthcare supplier (we were not able to study this with our data). Therefore, others have used populations of (far distance) movers to disentangle the supply effect on healthcare expenditures from the demand effect (Finkelstein, Gentzkow, and Williams 2016; Moura et al. 2019). The aim of our study, however, was to object the demand side effect, not the supply side effect. We hypothesized that the demand side is affected by the neighbourhood and that a change to the neighbourhood quality is associated with a change in healthcare utilization. In order to study the importance of the neighbourhood in the prediction of healthcare expenditure, we restricted our study population to people moving within hospital catchment areas (because we assumed that these people keep going to the same hospital). However, it may be that people moving within hospital catchment areas changed GP (in the Netherlands, almost every neighbourhood has one or more GP practices). As GP's might differ in the frequencies of consultation with the patients, in referral behaviour and in prescribing medication (Grytten and Sørensen 2003; Sinnige et al. 2016; Van Dijk et al. 2013), a possible change in GP may have confounded an effect of neighbourhood in our study. We believe, however, that the number of people changing a GP is rather small in our study - and thus the impact of this limitation can be neglected - because of the relative short distance of moves and because of a study among Dutch elderly showing that these elderly consider continuity of GP care (i.e. having the same GP) more important than distance to GP care (Berkelmans et al. 2010).

We may have found only limited additional value of the neighbourhood in our prediction model because neighbourhood might affect healthcare utilization only

on the long run. Hence, it may be that the timeframe of this study was too short to pick up the effect of neighbourhood on healthcare expenditure. Moreover, although livability varied within hospital catchment areas, the variation in neighbourhood exposure to, for example, blue space ('Blue space' showed to be associated with health (Wheeler et al. 2012)) would have been larger if our study population would also consist people who moved from, for example, the middle of the Netherlands to the West at the coast. Besides, this study only showed that neighbourhood location and quality (measured with the livability index) were not able to improve prediction models. However, a single neighbourhood characteristic might do a better job. Next, due to data restrictions (liveability was measured radically different in 2012 compared to 2008 and therefore longitudinal use of the liveability score was not possible) neighbourhood quality change was limited to livability data from 2008. This limitation might have affected the predictive value of neighbourhood change in the prediction model.

Finally, our study population may not be representative for the entire Dutch population because people who moved might have a different need of healthcare and subsequently different healthcare costs. Moreover, our study design may be overshadowed by the global financial crisis, which also affected the housing market in the Netherlands in 2010. Therefore, people moving in 2010 may be even more different from the Dutch population than movers in general.

The results of this study may be valuable to improve risk adjustment models because our study predicts healthcare costs (regular costs) in a similar way to the Dutch 'curative' risk adjustment models (i.e. excluding mental healthcare costs) (Van Veen et al. 2017). However, as we did not have access to the original Dutch risk equalization model, we could not directly test the added value of the neighbourhood for this model. Instead, we chose all variables relevant to healthcare utilization and available at CBS. Hence, our model included more socio-demographic and expenditure information than the Dutch risk equalization model, which may have underestimated the additional effect of neighbourhood for risk adjustment models. Besides, as many other countries do not have access to as many prediction variables as in the Netherlands, the additional effect of neighbourhood in risk adjustment models may be even further underestimated in these countries. Finally, as the influence of the neighbourhood on utilization may be modest, it may be a limitation of this study that we were not able to measure the amount of a medication that was used but only the number of ATC4 codes, a rather rough outcome.

A strength of this study is the use of a large set of linked information – up to 261 predictive variables. On the contrary to many other studies using claims data of only one or a few health insurers, our study used claims data of all Dutch health insurers covering almost the entire Dutch population. Hence, we were able to

select all people living in the Netherlands who applied to our inclusion criteria and repeated our analyses (with same findings) on different random selections of this pool of people. Furthermore, a rich set of high quality socio-demographic information gathered by CBS was used in this study. We believe that the amount and quality of the data provided in these datasets and the representativeness of the study population improved the reliability of our results.

In this study, next to predicting regular healthcare costs, we also predicted costs/utilization that are expected to be more sensitive to the neighbourhood and less effected by the supply side. Furthermore, we not only tested the effect of neighbourhood in the regular population, but also studied the effect in populations that are expected to be more sensitive to a change in neighbourhood. Because of this effort, this paper is able to more confidently show that the added value of the neighbourhood variables in the prediction of healthcare utilization and expenditure is very limited, at least for the neighbourhood variables used in this study.

4.2 Comparison of our Findings with Previous Studies

The MTO study, mentioned earlier, showed that personal health (Ludwig et al. 2012) and wealth (Chetty, Hendren, and Katz 2016) improved when children below the age of 13 moved from public housing to a low-poverty area. A consequence of this finding could be that moving to a better neighbourhood decreases the need for healthcare and that the neighbourhood is of importance for prediction of healthcare. Our study, however, found neighbourhood to be of limited to no additional value in the prediction of healthcare costs. Three recent studies have also tested the association between neighbourhood and healthcare costs/utilization. One study, measuring neighbourhood environment by looking at crime, safety and neighbourhood physical and social disorder, did not find any association, as well, between neighbourhood and the probability of having high healthcare costs (Sterling et al. 2018). However, the two other studies, either measuring neighbourhood with the Ontario marginalization index or by looking at neighbourhood social-economic status (SES), showed an association between neighbourhood and healthcare costs/utilization (Filc et al. 2014; Thavorn et al. 2017). The study of Thavorn et al. did not include prior utilization and expenditure in their model. Besides, the model included a less elaborate set of socio-demographic variables than the model used in our study. This is in line with our findings that the quality of the neighbourhood was of importance in model 2 but not in model 4. In the study by Filc et al. the neighbourhood SES is used as a proxy for individual SES, meaning that individual SES on itself was not included in the model. In our study, individual

SES is measured with annual household income, occupation, value of the house, non-mortgage debt and household asset percentile (see Appendix B for more information on these variables). In our analyses, these variables have high importance ranks/values (Appendix L). Hence, it may be that, in the study by Filc. et al. neighbourhood SES did only have an effect because of an underlying not measured effect of individual SES.

The study by Ash et al. (Ash et al. 2017) also tested the predictive value of neighbourhood in a risk adjustment model. Ash et al. measured neighbourhood using the neighbourhood stress score (NSS), which indicates the neighbourhood economic stress based on the percentage of household incomes below federal poverty level, unemployment, public assistance, having no car, single parents, and adults with no high school degree. They found that including social determinants, such as mental illness, unstable housing and NSS, in the model, improves prediction compared to a model only including medical information, age and gender. However, the NSS only had a minor contribution to the improvement in the predictive value of the model (Ash et al. 2017). Therefore, the findings of Ash et al. confirm our finding that neighbourhood is only of limited additional value in the prediction of healthcare costs. Several other risk adjustment studies have included a more broader region variable than neighbourhood. Region variables used are urbanization, county, province and region (not further specified) (Newhouse et al. 1989; Van Barneveld et al. 1998; Van Kleef, Van Vliet, and Van de Ven 2013; Van Veen et al. 2015, 2017). Two Dutch studies have tested the additional predictive effect of these region variables. The first study found that adding province to a risk adjustment model containing age, gender and supplementary insurance increased the R^2 from 2.3 to 2.4% (Van Vliet and Van de Ven 1992). The second study found that adding region to a model with only age and gender increased the R^2 from 5.97 to 6.01% (Van Kleef, Van Vliet, and Van de Ven 2013). Hence, although these studies added the region variable to a less elaborate prediction model, they found an even smaller improvement in R^2 than we did in our study. This may be because the region variables included in these models cover larger areas in the Netherlands than the region variable in our study, i.e. these variables contain less detail of the environment people live in.

In addition to studying the value of including region in risk adjustment models, studies have also explored the predictive value of including interactions between predictive factors in the model. These studies used one or several regression trees to identify valuable interactions (which were in some studies later on included in the traditional OLS regression model). The studies showed that including these interactions in risk adjustment models could marginally improve the predictive performance of the model (Buchner, Wasem, and Schillo 2017; Robinson 2008; Van Veen et al. 2017). In our study, the method random forest build

several regression trees that also include relevant interactions. However, as the number of regression trees that are build is large (i.e. 1000), and these trees include different interactions, it is difficult to determine what the additional predictive value of these interactions was in our study. Recent prediction models in the risk adjustment literature have reported R^2 values of 25–36% (Buchner, Wasem, and Schillo 2017; Van Veen et al. 2015, 2017). The models in these studies have been estimated using ordinary least squares regression, weighted least squares regression or regression trees. Our study used the random forest method to estimate a prediction model and obtained a much higher R^2 of 49% for regular costs and GP consultation costs and of 68% for sum of ATC codes. This large improvement in R^2 may be partly explained by the rich set of variables and mainly by the method used. As is well known, random forests provide an important improvement upon trees (a forest being made up of many trees, in our case 1000) and other prediction methods, which may explain the large R^2 found in the current study. For this reason, machine learning methods such as random forest may be promising in improving risk adjustment. Traditional risk adjustment models, such as ordinary least squares (OLS) regression, have been shown to be ill-equipped to deal with skewness, complex non-linear associations, and interactions, resulting in under- or overcompensation of certain types of insured (Eijkenaar and van Vliet 2017; Irvin et al. 2020). Machine learning methods are able to include non-linearity, skewedness and a large number of complex interactions. For this reason, a recent study using US insurance data found that the machine learning method 'gradient boosted trees' outperforms OLS in predicting healthcare expenditure, showing a 0.06% higher R^2 based on the same predictor variables (Irvin et al. 2020). Despite of the advantages of machine learning, as far as we are aware these methods have not been adopted in risk adjustment schemes so far, probably being due to unfamiliarity with the methods and the complexity of the models and their results (Irvin et al. 2020; Kan et al. 2019). To pursue this direction, the first question to answer is which machine learning method performs best in prediction healthcare expenditure of individuals; as done for example by Morid et al. (2017). Next, and a more difficult task is the implementation of the machine learning method in current risk adjustment schemes.

4.3 Conclusions

This study shows that neighbourhood has a small additional predictive value when added to a model with only socio-demographic information. No improvement in predictive performance was observed when adding neighbourhood to the prediction model with socio-demographic information, prior expenditure, and medication use.

Hence, only in prediction models in contexts with poor access to prior expenditure and utilization or a wish to minor the use of these variables, the quality of the neighbourhood should be considered as a possible prediction variable.

Furthermore, future research might also investigate 1) the value of other neighbourhood characteristics in the prediction of healthcare expenditures, 2) the long-term effect of neighbourhood on healthcare expenditures, 3) and how to integrate the 'random forest' method into risk adjustment.

References

Ash, A. S., E. O. Mick, R. P. Ellis, C. I. Kiefe, J. J. Allison, and M. A. Clark. 2017. "Social Determinants of Health in Managed Care Payment Formulas." *JAMA Internal Medicine* 177 (10): 1424–30.

Barker, R. G. 1968. *Ecological Psychology: Concepts and Methods for Studying the Environment of Human Behavior*: Stanford University Press.

Berkelmans, P. G., A. J. Berendsen, P. F. Verhaak, and K. van der Meer. 2010. "Characteristics of General Practice Care: what Do Senior Citizens Value?. A Qualitative Study." *BMC Geriatrics* 10: 80.

Berkman, L. F., T. Glass, I. Brissette, and T. E. Seeman. 2000. "From Social Integration to Health: Durkheim in the New Millennium." *Social Science & Medicine* 51 (6): 843–57.

Bize, R., J. A. Johnson, and R. C. Plotnikoff. 2007. "Physical Activity Level and Health-Related Quality of Life in the General Adult Population: a Systematic Review." *Preventive Medicine* 45 (6): 401–15.

Breiman, L. 2001. "Random Forests." *Machine Learning* 45: 5–32.

Buchner, F., J. Wasem, and S. Schillo. 2017. "Regression Trees Identify Relevant Interactions: Can This Improve the Predictive Performance of Risk Adjustment?." *Health Economics* 26 (1): 74–85.

Chetty, R., Hendren, N., and Katz, L. F. (2016), "The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment." *The American Economic Review* 106 (4): 855–902.

Craig, P., C. Cooper, D. Gunnell, S. Haw, K. Lawson, S. Macintyre, D. Ogilvie, M. Petticrew, B. Reeves, M. Sutton, and S. Thompson. 2012. "Using Natural Experiments to Evaluate Population Health Interventions: New Medical Research Council Guidance." *Journal of Epidemiology & Community Health* 66 (12): 1182–6.

de Boo, A. 2011. "The Health Care Informaton Centre Vektis [Vektis' Informatiecentrum voor de zorg', Dutch]." *Tijdschrift voor gezondheidswetenschappen* 89 (7): 358–9.

Diez Roux, A. V. and C. Mair. 2010. "Neighborhoods and Health." *Annals of the New York Academy of Sciences* 1186: 125–45.

Dijkema, M. B. A., R. T. van Strien, S. C. van der Zee, S. F. Mallant, P. Fischer, G. Hoek, B. Brunekreef, and U. Gehring. 2016. "Spatial Variation in Nitrogen Dioxide Concentrations and Cardiopulmonary Hospital Admissions." *Environmental Research* 151: 721–7.

Eijkenaar, F. and R. C. J. A. van Vliet. 2017. "Improving Risk Equalization for Individuals with Persistently High Costs: Experiences from the Netherlands." *Health Policy* 121 (11): 1169–76.

Eijkenaar, F., R. C. J. A. van Vliet, and R. C. van Kleef. 2018. "Diagnosis-based Cost Groups in the Dutch Risk-Equalization Model: Effects of Clustering Diagnoses and of Allowing Patients to Be Classified into Multiple Risk-Classes." *Medical Care* 56 (1): 91–6.

Ellen, I. G., T. Mijanovich, and K. N. Dillman. 2001. "Neighborhood Effects on Health: Exploring the Links and Assessing the Evidence." *Journal of Urban Affairs* 23 (3-4): 391–408.

Ellen, I. G. and M. A. Turner. 2003. "Do neighborhoods Matter and Why?." In *Choosing a Better Life? Evaluating the Moving to Opportunity Social Experiment*, edited by J. M. Goering and J. D. Feins, 313–8. Washington, D.C.: Urban Institute Press.

Filc, D., N. Davidovich, L. Novack, and R. D. Balicer. 2014. "Is Socioeconomic Status Associated with Utilization of Health Care Services in a Single-Payer Universal Health Care System?." *International Journal for Equity in Health* 13: 115.

Finkelstein, A., M. Gentzkow, and H. Williams. 2016. "Sources of Geographic Variation in Health Care: Evidence from Patient Migration." *Quarterly Journal of Economics* 131 (4): 1681–726.

Grytten, J. and R. Sørensen. 2003. "Practice Variation and Physician-specific Effects." *Journal of Health Economics* 22 (3): 403–18.

Hartig, T., R. Mitchell, S. de Vries, and H. Frumkin. 2014. "Nature and Health." *Annual Review of Public Health* 35: 207–28.

Haskell, W. L., I.-M. Lee, R. R. Pate, K. E. Powell, S. N. Blair, B. A. Franklin, C. A. Macera, G. W. Heath, P. D. Thompson, and A. Bauman. 2007. "Physical Activity and Public Health: Updated Recommendation for Adults from the American College of Sports Medicine and the American Heart Association." *Circulation* 116 (9): 1081.

Hastie, T., R. Tibshirani, and J. Friedman. 2009. *The Elements of Statistical Learning - Data Mining, Inference, and Prediction*, Vol. 2. New York: Springer.

Iida, H. and R. G. Rozier. 2013. "Mother-perceived Social Capital and Children's Oral Health and Use of Dental Care in the United States." *American Journal of Public Health* 103 (3): 480–7.

Irvin, J. A., A. A. Kondrich, M. Ko, P. Rajpurkar, B. Haghgoo, B. E. Landon, R. L. Phillips, S. Petterson, A. Y. Ng, and S. Basu. 2020. "Incorporating Machine Learning and Social Determinants of Health Indicators into Prospective Risk Adjustment for Health Plan Payments." *BMC Public Health* 20 (1): 608.

Jegers, M., K. Kesteloot, D. De Graeve, and W. Gilles. 2002. "A Typology for Provider Payment Systems in Health Care." *Health Policy* 60 (3): 255–73.

Kan, H. J., H. Kharrazi, H.-Y. Chang, D. Bodycombe, K. Lemke, and J. P. Weiner. 2019. "Exploring the Use of Machine Learning for Risk Adjustment: A Comparison of Standard and Penalized

Linear Regression Models in Predicting Health Care Costs in Older Adults." *PloS One* 14 (3): e0213258.

Katz, L. F., J. R. Kling, and J. B. Liebman. 2001. "Moving to Opportunity in Boston: Early Results of a Randomized Mobility Experiment." *Quarterly Journal of Economics* 116 (2): 607–54.

Leader, A. E. and Y. L. Michael. 2013. "The Association between Neighborhood Social Capital and Cancer Screening." *American Journal of Health Behavior* 37 (5): 683–92.

Lee, K.-S., J.-S. Lee, and J.-H. Kwon. 2014. "The Effects of Urban Forests on the Medical Care Use for Respiratory Disease in Korea: A Structural Equation Model Approach." *International Journal of Public Policy* 10 (4-5): 195–208.

Leibowitz, A. A. 2004. "The Demand for Health and Health Concerns after 30 Years." *Journal of Health Economics* 23 (4): 663–71.

Leidelmeijer, K., G. Marlet, J. van Iersel, C. van Woerkens, and H. van der Reijden. 2008. *De Leefbarometer: Leefbaarheid in Nederlandse wijken en buurten gemeten en vergeleken; rapportage instrumentontwikkeling*: RIGO Research en Advies BV and Atlas voor gemeenten.

Leidelmeijer, K., G. Marlet, C. Van Woerkens, N. Van den Berg, M. Bosker, H. Van der Reijden, R. Schulenberg, E. Cozijnsen, and J. Van Iersel. 2009. *Leefbarometer meting 2008 - Eerste uitkomsten en methodische verantwoording*: R I G O Research en Advies B V & Stichting Atlas voor gemeenten.

Ludwig, J., G. J. Duncan, L. A. Gennetian, L. F. Katz, R. C. Kessler, J. R. Kling, and L. Sanbonmatsu. 2012. "Neighborhood Effects on the Long-Term Well-Being of Low-Income Adults." *Science* 337 (6101): 1505–10.

Macintyre, S. and A. Ellaway. 2000. *Ecological Approaches: Rediscovering the Role of the Physical and Social Environment*, Social Epidemiology, edited by L. F. Berkman and I. Kawachi, 332–48. New York, NY: Oxford University Press.

Marco, M., E. Gracia, J. M. Tomás, and A. López-Quílez. 2015. "Assessing Neighborhood Disorder: Validation of a Three-Factor Observational Scale." *The European Journal of Psychology Applied to Legal Context* 7 (2): 81–9.

Martin-Storey, A., C. E. Temcheff, P. L. Ruttle, L. A. Serbin, D. M. Stack, A. E. Schwartzman, and J. E. Ledingham. 2012. "Perception of Neighborhood Disorder and Health Service Usage in a Canadian Sample." *Annals of Behavioral Medicine* 43 (2): 162–72.

Mohnen, S. M., and S. Schneider. 2019. "Neighborhood Characteristics as Determinants of Healthcare Utilization – A Theoretical Model." *Health Economics Review* 9 (7). https://doi.org/10.1186/s13561-019-0226-x.

Morid, M. A., K. Kawamoto, T. Ault, J. Dorius, and S. Abdelrahman. 2017. "Supervised Learning Methods for Predicting Healthcare Costs: Systematic Literature Review and Empirical Evaluation." *AMIA Annu Symp Proc*: 1312–21. 2017.

Moura, A., M. Salm, R. Douven, and M. Remmerswaal. 2019. "Causes of Regional Variation in Dutch Healthcare Expenditures: Evidence from Movers." *Health Economics (United Kingdom)* 28 (9): 1088–98.

Newhouse, J. P., W. G. Manning, E. B. Keeler, and E. M. Sloss. 1989. "Adjusting Capitation Rates Using Objective Health Measures and Prior Utilization." *Health Care Financing Review* 10 (3): 41–54.

Nguyen, D. D., K. H. Ho, and J. H. Williams. 2011. "Social Determinants and Health Service Use Among Racial and Ethnic Minorities: Findings from a Community Sample." *Social Work in Health Care* 50 (5): 390–405.

NZa. 2016. *Marktscan Zorgverzekeringsmarkt 2016*: Nederlandse Zorgautoriteit. Dutch Healthcare Authority.

Prins, R. G., S. M. Mohnen, F. J. van Lenthe, J. Brug, and A. Oenema. 2012. "Are Neighbourhood Social Capital and Availability of Sports Facilities Related to Sports Participation Among Dutch Adolescents?." *The International Journal of Behavioral Nutrition and Physical Activity* 9: 90.

Robinson, J. W. 2008. "Regression Tree Boosting to Adjust Health Care Cost Predictions for Diagnostic Mix." *Health Services Research* 43 (2): 755–72.

Sallis, J. F., R. B. Cervero, W. Ascher, K. A. Henderson, M. K. Kraft, and J. Kerr. 2006. "An Ecological Approach to Creating Active Living Communities." *Annual Review of Public Health* 27: 297–322.

Sallis, J. F. and N. Owen. 2015. "Ecological Models of Health Behavior." In *Health Behavior - Theory, Research, and Practice*, edited by K. Glanz, B. K. Rimer and K. Viswanath, 43–64. San Francisco, CA: Jossey-Bass.

Sampson, R. J., J. D. Morenoff, and T. Gannon-Rowley. 2002. "Assessing Neighborhood Effects": Social Processes and New Directions in Research." *Annual Review of Sociology* 28: 443–78.

Shin, S., C. Schumacher, and E. Feess. 2017. "Do Capitation-based Reimbursement Systems Underfund Tertiary Healthcare Providers? Evidence from New Zealand." *Health Economics (United Kingdom)* 26 (12): e81–102.

Shrestha, A., S. Bergquist, E. Montz, and S. Rose. 2018. "Mental Health Risk Adjustment with Clinical Categories and Machine Learning." *Health Services Research* 53 (Suppl. 1): 3189–206.

Sibley, L. M. and R. H. Glazier. 2012. "Evaluation of the Equity of Age–Sex Adjusted Primary Care Capitation Payments in Ontario, Canada." *Health Policy* 104 (2): 186–92.

Sinnige, J., Braspenning, J. C., Schellevis, F. G., Hek, K., Stirbu, I., Westert, G. P. and Korevaar, J. C. 2016. "Inter-practice Variation in Polypharmacy Prevalence Amongst Older Patients in Primary Care." *Pharmacoepidemiology and Drug Safety*, 25 (9), 1033–41.

Sterling, S., F. Chi, C. Weisner, R. Grant, A. Pruzansky, S. Bui, P. Madvig, and R. Pearl. 2018. "Association of Behavioral Health Factors and Social Determinants of Health with High and Persistently High Healthcare Costs." *Prev Med Rep* 11: 154–9.

Stuart, E. A. 2010. "Matching Methods for Causal Inference: A Review and a Look Forward." *Statistical Science* 25 (1): 1–21.

Thavorn, K., C. J. Maxwell, A. Gruneir, S. E. Bronskill, Y. Bai, A. J. Kone Pefoyo, Y. Petrosyan, and W. P. Wodchis. 2017. "Effect of Socio-Demographic Factors on the Association between Multimorbidity and Healthcare Costs: a Population-Based, Retrospective Cohort Study." *BMJ Open* 7 (10): e017264.

Van Barneveld, E. M., L. M. Lamers, R. C. van Vliet, and W. P. van de Ven. 1998. "Mandatory Pooling as a Supplement to Risk-Adjusted Capitation Payments in a Competitive Health Insurance Market." *Social Science & Medicine* 47 (2): 223–32.

Van de Ven, W. P. 2011. "Risk Adjustment and Risk Equalization: what Needs to Be Done?." *Health Economics, Policy and Law* 6 (1): 147–56.

Van de Ven, W. P., K. Beck, F. Buchner, E. Schokkaert, F. T. Schut, A. Shmueli, and J. Wasem. 2013. "Preconditions for Efficiency and Affordability in Competitive Healthcare Markets: Are They Fulfilled in Belgium, Germany, Israel, the Netherlands and Switzerland?." *Health Policy* 109 (3): 226–45.

Van de Ven, W. P., K. Beck, C. Van de Voorde, J. Wasem, and I. Zmora. 2007. "Risk Adjustment and Risk Selection in Europe: 6 Years Later." *Health Policy* 83 (2): 162–79.

Van den Berg, M., W. Wendel-Vos, M. Van Poppel, H. Kemper, W. Van Mechelen, and J. Maas. 2015. "Health Benefits of Green Spaces in the Living Environment: A Systematic Review of Epidemiological Studies." *Urban Forestry and Urban Greening* 14 (4): 806–16.

Van Dijk, C. E., J. C. Korevaar, J. D. De Jong, B. Koopmans, M. Van Dijk, and D. H. De Bakker. 2013. *Kennisvraag: Ruimte Voor Substitutie? Verschuivingen Van Tweedelijns- Naar Eerstelijnszorg*: NIVEL.

Van Kleef, R. C., R. C. Van Vliet, and W. P. Van de Ven. 2013. "Risk Equalization in The Netherlands: an Empirical Evaluation." *Expert Review of Pharmacoeconomics & Outcomes Research* 13 (6): 829–39.

Van Veen, S. H., R. C. Van Kleef, W. P. Van de Ven, and R. C. Van Vliet. 2015. "Improving the Prediction Model Used in Risk Equalization: Cost and Diagnostic Information from Multiple Prior Years." *The European Journal of Health Economics* 16 (2): 201–18.

Van Veen, S. H., R. C. Van Kleef, W. P. Van de Ven, and R. C. J. A. Van Vliet. 2017. "Exploring the Predictive Power of Interaction Terms in a Sophisticated Risk Equalization Model Using Regression Trees." *Health Economics* 27: e1–e12.

Van Vliet, R. C. and W. P. Van de Ven. 1992. "Towards a Capitation Formula for Competing Health Insurers. An Empirical Analysis." *Social Science & Medicine* 34 (9): 1035–48.

Wheeler, B. W., M. White, W. Stahl-Timmins, and M. H. Depledge. 2012. "Does Living by the Coast Improve Health and Wellbeing?." *Health & Place* 18 (5): 1198–201.

WHO. 2013. *Review of Evidence on Health Aspects of Air Pollution – REVIHAAP Project Technical Report*: The WHO European Centre for Environment and Health.

Wright, M. N. and A. Ziegler. 2017. "Ranger: A Fast Implementation of Random Forests for High Dimensional Data in C++ and R. Journal of Statistical Software." *Journal of Statistical Software* 77 (1): 1–17.