



Relative importance of perceived physical and social neighborhood characteristics for depression: a machine learning approach

Marco Helbich¹ · Julian Hagenauer¹ · Hannah Roberts¹

Received: 30 May 2019 / Accepted: 9 November 2019 / Published online: 14 November 2019
© The Author(s) 2019

Abstract

Purpose The physical and social neighborhood environments are increasingly recognized as determinants for depression. There is little evidence on combined effects of multiple neighborhood characteristics and their importance. Our aim was (1) to examine associations between depression severity and multiple perceived neighborhood environments; and (2) to assess their relative importance.

Methods Cross-sectional data were drawn from a population-representative sample ($N=9435$) from the Netherlands. Depression severity was screened with the Patient Health Questionnaire (PHQ-9) and neighborhood perceptions were surveyed. Supervised machine learning models were employed to assess depression severity-perceived neighborhood environment associations.

Results We found indications that neighborhood social cohesion, pleasantness, and safety inversely correlate with PHQ-9 scores, while increasing perceived distance to green space and traffic were correlated positively. Perceived distance to blue space and urbanicity seemed uncorrelated. Young adults, low-income earners, low-educated, unemployed, and divorced persons were more likely to have higher PHQ-9 scores. Neighborhood characteristics appeared to be less important than personal attributes (e.g., age, marital and employment status). Results were robust across different ML models.

Conclusions This study suggested that the perceived social environment plays, independent of socio-demographics, a role in depression severity. Contrasted with person-level and social neighborhood characteristics, the prominence of the physical neighborhood environment should not be overstated.

Keywords Depression · Social neighborhood environment · Physical neighborhood environment · Netherlands

Abbreviations

ALPHA	Instruments for assessing levels of physical activity and fitness
CV	cross-validation
NEEDS	Dynamic urban environmental exposures on depression and suicide
GLM	Generalized linear model
GIS	Geographic information system
GBM	Gradient boosting machine
MAE	Mean absolute error

ML	Machine learning
NNET	Neural network
PHQ	Patient Health Questionnaire
RF	Random forest
RMSE	Root-mean-square error

Introduction

It is gradually established that people's mental health is shaped, in addition to person-level attributes [1], by the neighborhood environment, which can be broadly categorized into physical and social characteristics [2]. Since mental illness contributes 13% of disability adjusted life-years lost to the global burden of mental disorders [3], it is necessary to understand how and to what extent the physical and social neighborhoods affect depression.

Recent reviews, mainly including cross-sectional and limited longitudinal evidence, suggest that socio-spatial

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00127-019-01808-5>) contains supplementary material, which is available to authorized users.

✉ Marco Helbich
m.helbich@uu.nl

¹ Utrecht University, Princetonlaan 8a, 3584 CB Utrecht, The Netherlands

aspects of people's living environment can contribute to or be protective against depression [4–6]. It was found that, for example, traffic-related air pollution [7], noise [8], safety concerns [9], and urbanicity [10] were harmful for mental health because they are usually experienced as undesirable and stressful for residents which may, in turn, promote depressive mood [6, 11].

In contrast, it is theorized that green and blue space [12], and social capital [13] were beneficial because such factors may be stress-reducing and buffer against negative thoughts [14], while neighborhood safety and social cohesion could act as coping mechanisms to safeguard from psychological distress [4]. Associations such as these are, however, not universally confirmed, and the mechanisms are yet to be fully understood.

Our present knowledge mainly originates from studies incorporating a single neighborhood characteristic [15–17], which may have resulted in misestimated neighborhood effects. In fact, as put forward by the socio-ecological model of health [18], multiple physical and social neighborhood characteristics may be involved at the same time, implying a complex interplay. Therefore, when assessing correlations, either directly or in interaction with person-level attributes, it is rational to assume that multiple neighborhood characteristics may re-inforce or level-out each other. Supportive empirical evidence is, however, scarce [19–21], as are studies that assess the relative importance of such characteristics in such a constellation [22].

Furthermore, it is suggested that neighborhood characteristics interact and are potentially non-linearly associated with depression [12]. Current insights from conventional (multilevel) regressions may be limited in that they assume correlations are linear [19–21]. This cannot be substantiated by theoretical considerations [18], and potentially results in overly simplistic models which may contribute to contradictory findings.

These issues might be overcome through machine learning (ML) [23]. ML includes a broad set of inductive models that learn to approximate unknown target functions from training data without being explicitly designed for a specific task. Echoing recent calls for methodological advances [12], many of the models allow for non-linear correlations, routinely assess variable importance, and explore interactions between person-level and neighborhood characteristics.

Given the inconsistencies between and the methodological limitations of studies conducted to date, this large-scale explorative study in the Netherlands aimed (1) to examine the associations of physical and social neighborhood characteristics on people's depression severity; and (2) to assess the relative importance of people's perception of physical and social neighborhood characteristics on depression severity through ML approaches.

Materials and methods

Study setting and participants

This cross-sectional study reports on a nationally representative population sample in the Netherlands. In the course of the NEEDS project [24], an online survey was carried out with Statistics Netherlands between September and December 2018.

Participants needed to fulfill the following eligibility criteria: to be aged between 18 and 65 years, and living in a private household. Through systematic sampling with probabilities proportional to the target population size, sub-municipalities were first selected from each COROP region (i.e., a regional Dutch division). Next, from those regions individuals registered in the Dutch National Personal Records Database were randomly sampled. Incentives were offered to increase the response. Of those 45,000 invited people, 11,524 completed the questionnaire resulting in an overall response rate of 25.6%. We conducted a complete case analysis excluding those with any missing information. After exclusions due to incomplete variables ($N=2089$), the final sample size was 9435. A full description of the study protocol [24] and the survey is available [25].

Ethical approval of the study design was obtained from the Ethics Review Board of the Faculty of Social and Behavioral Sciences of Utrecht University (FETC17–060). Informed consent was implied by conducting the questionnaire.

Data

Our survey in the Dutch language comprised various modules including sociodemographics, mental health, and perception of the residential neighborhood. Other questions were asked but were not included here. The survey was further enriched with selected register data, namely urbanicity and income, available through Statistics Netherlands. If not mentioned otherwise, register data refer to 1st July 2018.

Severity of depression

Depression severity was operationalized through the depression module of the Patient Health Questionnaire (PHQ-9) [26]. This instrument was recognized as having good diagnostic performance, good sensitivity, and good specificity in a meta-analysis [27]. The 9-item long screener assesses people's experience within the last 2 weeks. The statements address whether a respondent felt down or depressed, had pleasure in doing things, had thoughts of suicide, etc. Each

item is on a 4-point Likert scale ranging from “not at all” to “nearly every day”. We summed the individual item scores per question. The total score was our outcome measure, assumed to be continuous. A PHQ-9 total score of 1 refers to no evidence of depressive symptoms, while 27 represents highest depression severity. The internal consistency of the PHQ-9 in our sample had a Cronbach’s alpha of 0.887.

Physical neighborhood environment

Residential exposure to natural environments was measured twofold: first, we asked for the perceived distance to the nearest green space (defined as parks, play areas, sports fields or forests), and second, perceived distance to the nearest blue space (rivers, lakes, beaches). Distances were categorized into < 300 m, 300–1 km, > 1–5 km, and > 5 km. These were in line with others [28, 29].

To capture the perceived density of traffic, respondents were asked to evaluate traffic in their neighborhood based on their experiences in the past 6 months. The variable was on a 4-point Likert scale from “very busy/congestion” to “very quiet”, with a greater higher score indicating less perceived density of traffic.

Pleasantness was operationalized using four questions from the ALPHA questionnaire [30]. Respondents were asked to what extent they agree that the environment is pleasant for walking and cycling, the amount of incivilities (e.g., litter, graffiti) present, the number of trees in the street and the maintenance of buildings. Each question was answered on a 4-point Likert scale, with inverse scoring on negatively stated items. A greater score overall indicated a more pleasant neighborhood. The Cronbach’s alpha was with 0.620 low. It is argued elsewhere [30] that for such environmental constructs a reduced Cronbach’s alpha is acceptable because the involved indicators are often not intercorrelated.

Data on urbanicity refer to the urbanicity of the neighborhood (‘buurt’) of the person’s home according to Statistics Netherlands. The variable was grouped into quintiles ranging from “not urban” (< 500 addresses/km²) to “very strongly urban” (> 2500 addresses/km²). Within this range, class breaks were set every additional 500 addresses/km².

Social neighborhood environment

To operationalize social cohesion, participants were asked to rate their agreement on a 5-point scale ranging from 1 (totally agree) to 5 (totally disagree), with the following statements [31]: ‘People around here are willing to help their neighbors’, ‘I live in a cozy neighborhood, people in this neighborhood can be trusted’, ‘people in this neighborhood generally cannot get along so well’, and ‘people in this neighborhood do not share the same values’. Negatively stated items were inversely scored, with a higher score

overall indicating greater social cohesion. The internal consistency was with a Cronbach’s alpha of 0.829.

Questions on perceived safety were drawn from the “neighborhood safety” module within the ALPHA questionnaire [30]. Participants rated their level of agreement with a set of five statements on a 4-point Likert scale from 1 (strongly disagree) to 4 (strongly agree). Statements included: ‘It is dangerous to leave a bicycle locked in my neighborhood’ and ‘it is dangerous in my neighborhood during the day because of the level of crime’. Responses were reverse coded and then summed. The Cronbach’s alpha was 0.822.

Covariates

The following routinely considered covariates were included [1, 19, 32]: age (grouped into 5-year categories), gender (men, woman), ethnicity (Dutch, Western background, non-Western background), marital status (married, divorced, widowed, unmarried), employment status (employed, unemployed), and education (re-coded into low (up to lower secondary education), medium (up to upper secondary education), and high (university education and further)). Household income was obtained via Statistics Netherlands; the most recent data available are from 1st January 2016. The data were classified into quintiles (1 = lowest, 5 = highest).

Statistical analyses

Machine learning models

We undertook a supervised machine learning (ML) [33] approach to assess associations between depression severity and neighborhood characteristics while adjusting for person-level attributes. Generally spoken, supervised ML models seek for patterns (i.e., complex relationships and interactions) in training data and use this information to conduct inference or predictions for unseen data without relying on strict model assumptions.

Since the repertoire of available regression-based ML algorithms is large and the performance may depend on the data set at hand [23], we selected well-established regression-based models. We fitted a generalized linear model (GLM) [34] as base model, and the following three ML models: an artificial neural network (NNET) [35], a random forest (RF) [36], and a gradient boosting machine (GBM) [37]. For brief model descriptions, see the supplementary materials. Our model pre-selection was also guided by benchmark studies [23]. Each model was fitted with depression severity as outcome variable and full covariate adjustment. All input variables were scaled to have zero mean and unit variance to make them comparable.

The goodness-of-fit of each model depends on the chosen hyper-parameters. We tested different settings from and

evaluated the models' root-mean-square error (RMSE), the mean absolute error (MAE), and R^2 using 10 times repeated 10-fold cross-validation (CV). CV randomly partitions the data into 10 disjoint subsets. Subsets are used one at a time for model testing while the remaining ones are used for model building. Table A1 in the supplementary materials lists the final parameter settings. All analyses were carried

out in the R programming environment [38] and the caret package [39] facilitated parameter tuning and model validation, while providing a unified interface for each algorithm.

Table 1 Sample characteristics

Variable	Category	%	Min.	Max.	Mean	SD
PHQ-9 scores			0.000	27.000	4.797	4.861
Physical neighborhood						
Green space	< 300 m	74.118				
	300 m to 1 km	21.293				
	> 1 km to 5 km	4.240				
	> 5 km	0.350				
Blue space	< 300 m	21.791				
	300 m to 1 km	29.889				
	> 1 km to 5 km	33.333				
	> 5 km	14.987				
Traffic	Very low	22.946				
	Low	43.932				
	High	29.995				
	Very high	3.127				
Pleasantness			4.000	16.000	6.487	2.009
Urbanicity	Very low	9.274				
	Low	24.123				
	Middle	16.789				
	High	30.069				
	Very high	19.746				
Social neighborhood						
Social cohesion			5.000	25.000	19.388	3.250
Safety			6.000	24.000	10.088	2.947
Covariates						
Age			18.000	65.000	44.904	13.815
Education	Low	19.194				
	Middle	35.601				
	High	45.204				
Employment	Unemployed	24.759				
Ethnicity	Dutch	87.790				
	Western	7.451				
	Non-western	4.759				
Gender	Male	47.843				
Income	Very low	9.126				
	Low	10.429				
	Middle	18.654				
	High	26.762				
	Very high	35.029				
Marital status	Married	53.672				
	Divorced	8.511				
	Widowed	1.526				
	Unmarried	36.290				

Model interpretability

Different strategies were conducted for an in-depth model understanding. First, we assessed the variable importance relative to each other by permuting one variable at a time and measuring the change in performance [36]. To explore commonalities in variable importance, a heat map was generated. Second, to investigate the directions of the relationships and possible non-linearities, partial dependence plots were used. These plots show the change in the average predicted value as one or more variables vary over their marginal distribution [40]. Third, through the H statistic we quantified either the total interaction of one variable with all others or the interaction of two variables [41]. A value of 0 means no interaction; 1 means that the entire variance is explained by the partial dependence functions.

Results

Sample description

Of 11,524 participants, 9435 (81.8%) had complete data. The Mann–Whitney–Wilcoxon test confirmed that omitting survey respondents with incomplete information resulted in no significant differences ($p=0.333$) in depression severity between the full and retained sample. Our sample had a mean PHQ-9 score of 4.857 with a standard deviation (SD) of 4.913. Table 1 summarizes the socio-demographic characteristics of our sample.

Model fits

Figure 1 shows the cross-validated model fits. The median magnitudes of model performances were, independent of the fit measure, rather similar. More specifically, the lowest median MAE and RMSE were achieved by GBM, while GLM had the highest errors. GBM also achieved the highest R^2 , while GLM had the lowest. The R^2 s were modestly high. Wilcoxon tests showed that the median performance of GBM was always significantly better than the one of GLM ($p < 0.050$). Generally, no significant difference in median

performance were found for RF and NNET. The Table A2 in the supplementary materials lists the detailed test results.

Variable importance

The clustered heat map in Fig. 2 shows two groups of variables with different levels of importance. The three most important variables for predicting depression severity were

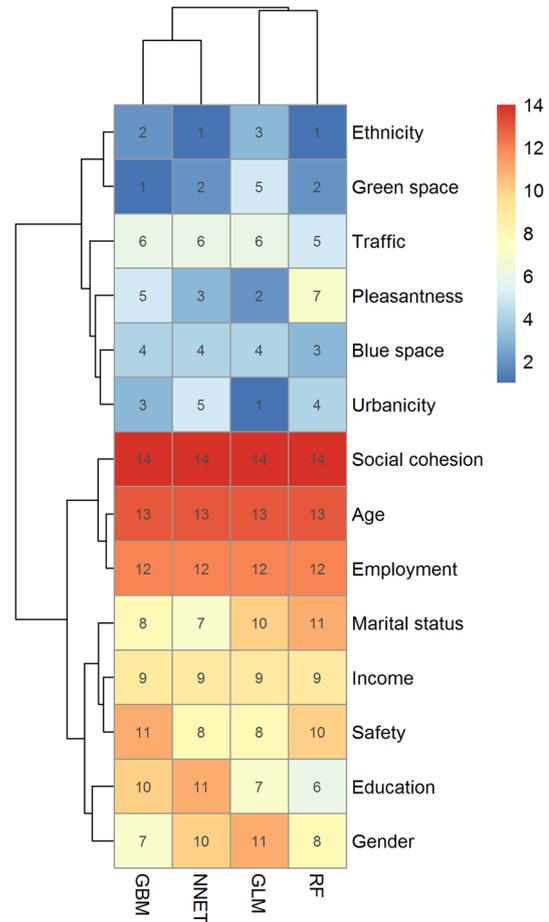
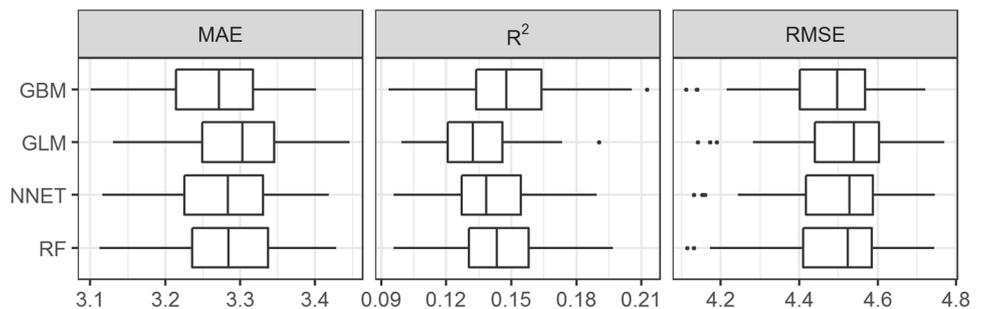


Fig. 2 Clustered heat map of the variable importance. The number per cell refers to the variable rank. Higher ranks indicate more important variables

Fig. 1 Summary of the cross-validated model fits. *RMSE* root-mean-square error, *MAE* mean absolute error



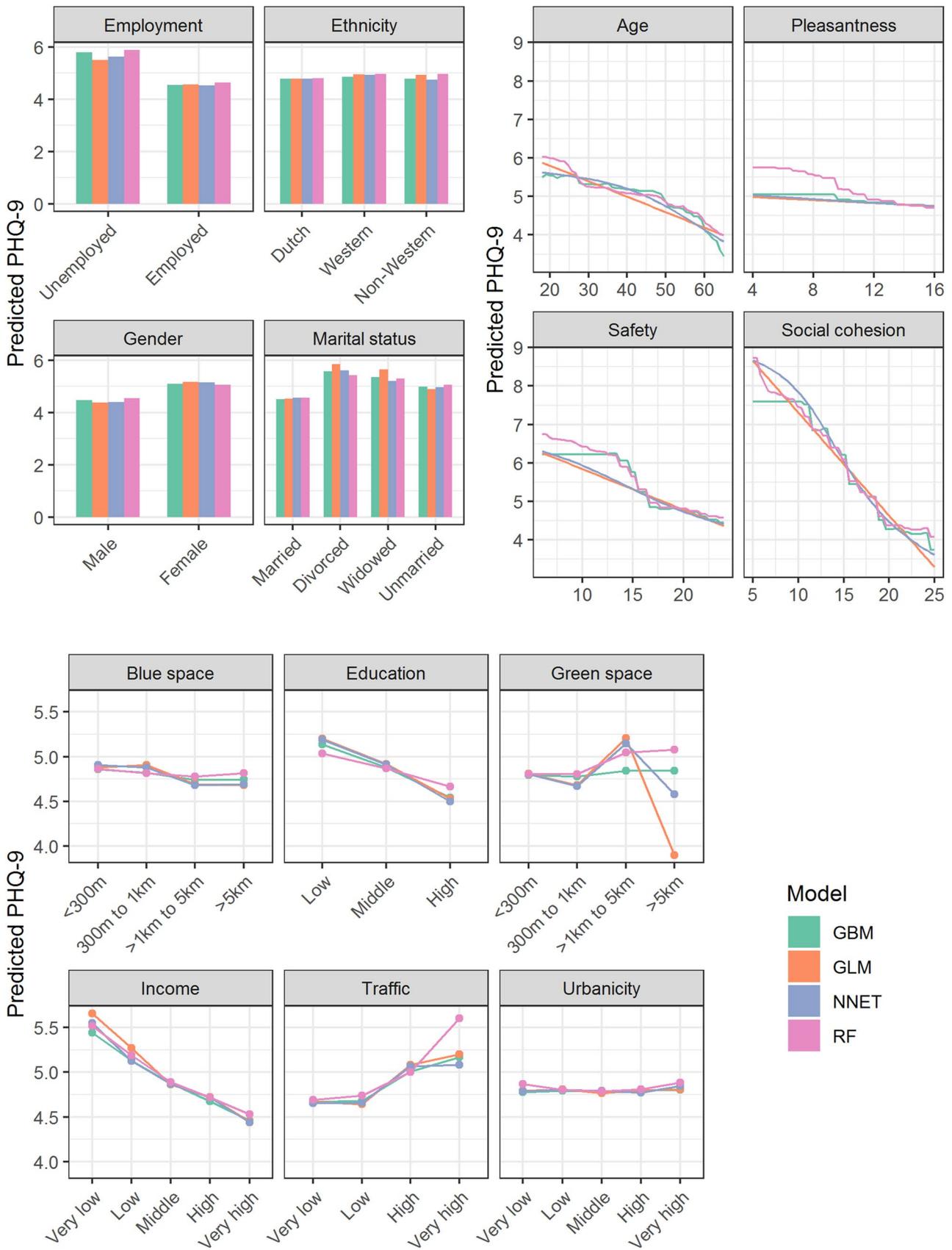


Fig. 3 Partial dependence plots relating each predictor to the PHQ-9 scores. Models are based on full covariate adjustment

social cohesion, age, and employment status. Of minor importance were urbanicity, ethnicity, and perceived distance to green and blue spaces. Only minor changes appeared across the models.

Correlation assessments

Correlations between PHQ-9 and the person-level and neighborhood characteristics are displayed in Fig. 3. As before, model differences were mostly small. Perceived safety and pleasantness were both negatively, and roughly linearly, associated with PHQ-9 scores. Social cohesion was inversely correlated with PHQ-9 scores. Perceived traffic was positively correlated with PHQ-9 scores. Both RF and GBM suggested that perceived distance to green space was positively correlated with PHQ-9 scores. Yet, the results for the distances > 5 km are inconclusive and diverge across the models. However, the category perceived distance to green space > 5 km was sparsely populated (Table 1). Blue space seemed to be uncorrelated with PHQ-9; as was urbanicity.

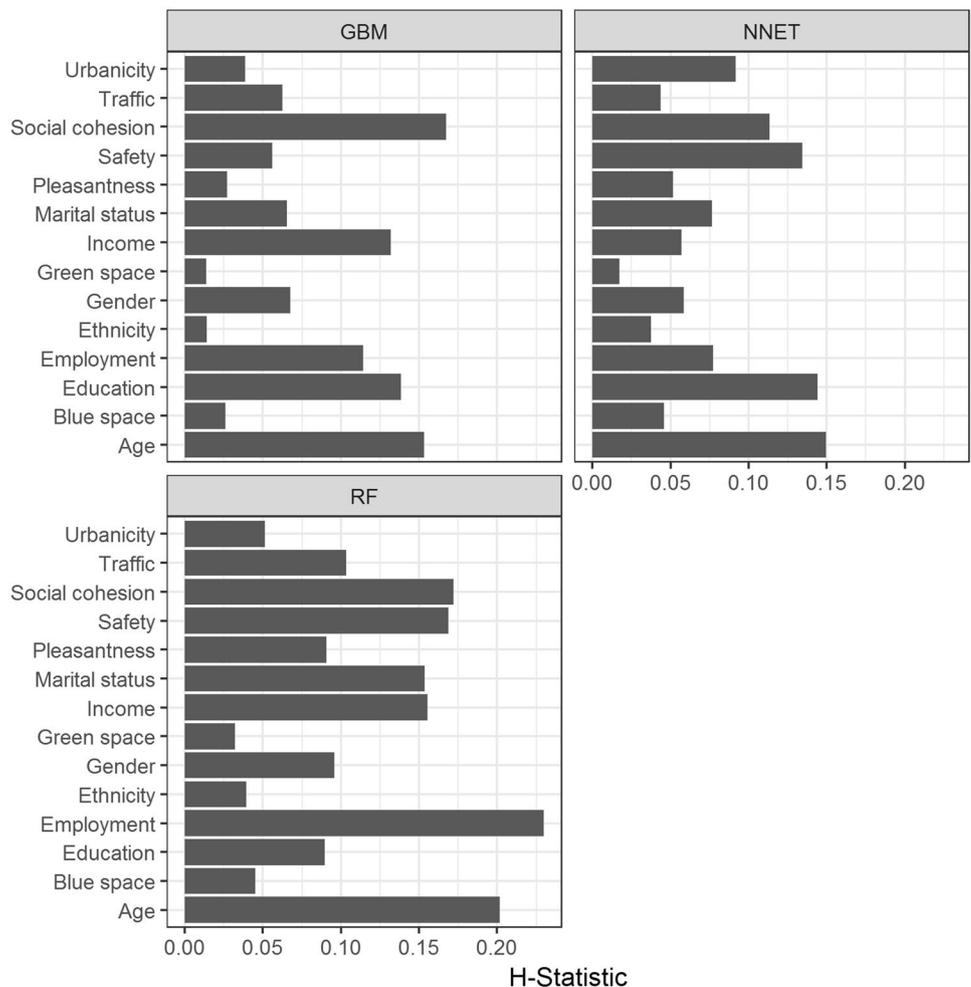
Unemployed, female, and divorced people showed higher PHQ-9 scores. PHQ-9 scores were substantially higher also for low earners and lower educated people. No differences were observed across ethnicities.

Variable interactions

Figure 4 shows pronounced overall variable interactions for social cohesion, age, employment, and education; neighborhood characteristics showed only little interaction.

Further, Figs. A1 and A2 detail with which variables social cohesion and age interact most (e.g., education, employment), which is the basis for Fig. 5 showing the bivariate interactions of these variables. The effect of employment status on the PHQ-9 scores varied over age, with unemployed persons always having a higher risk. Between 30 and 50 years of age, the differences were most pronounced; this gap in risk shrinks from 50 years onwards. Independent of the model, respondents that were less well educated and younger than 30 years old had pronounced

Fig. 4 Overall variable interaction (note that the *H* statistic is not available for GLMs)



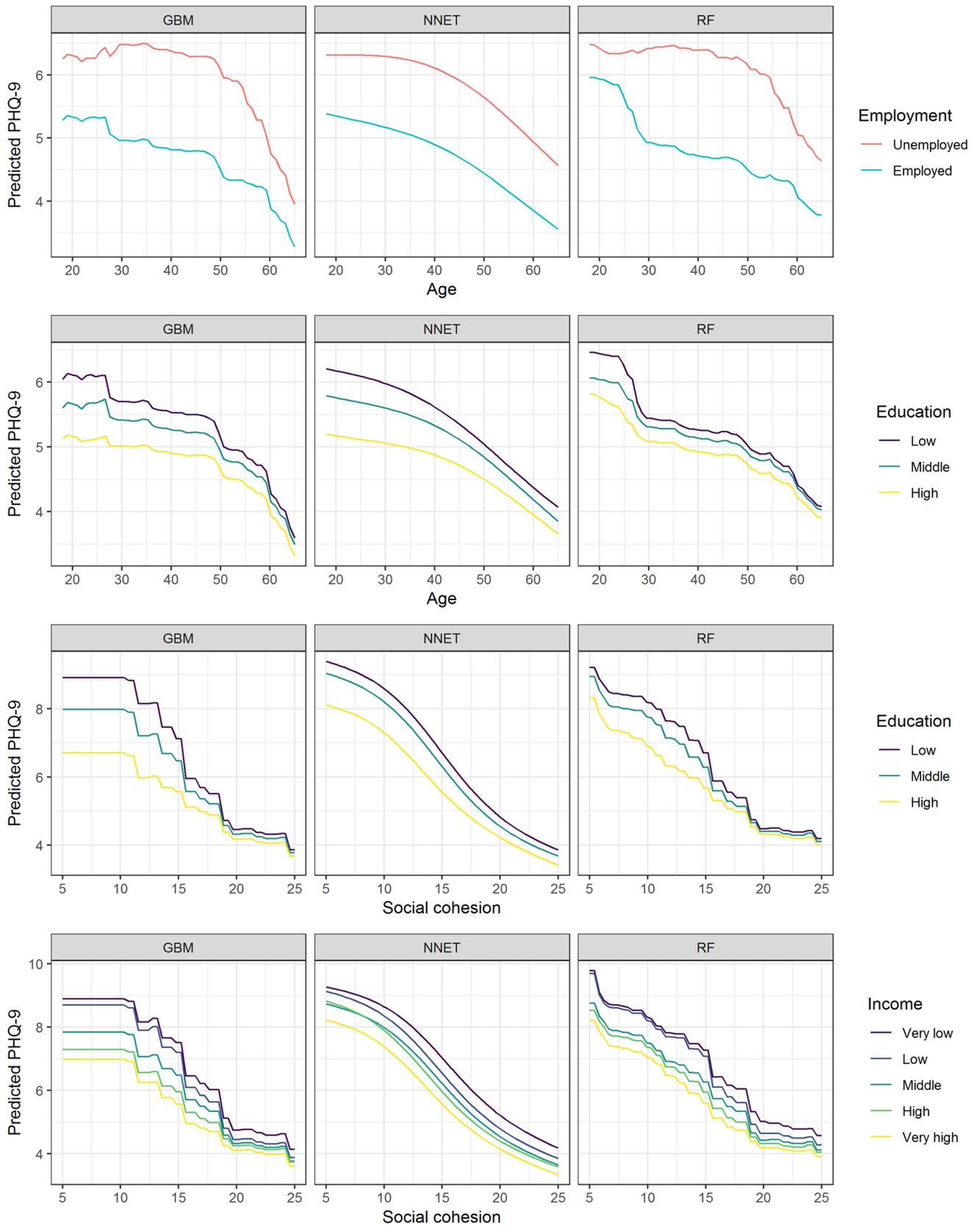


Fig. 5 Two-dimensional partial dependence plots relating predictors to the outcome PHQ-9 scores across the ML models. Only variables with a pronounced interaction are shown. Models are based on full covariate adjustment

PHQ-9 scores. With increasing age, the differences between lower and higher educated groups aligned each other.

Similar patterns were observed for social cohesion. High education and income were associated with lower PHQ-9 scores. However, while the differences in PHQ-9 scores between education groups decreased with social cohesion score, a decrease in the differences of the PHQ-9 scores between income groups was not always observable. For instance, there was always a notable gap in PHQ-9 scores between low and middle income groups, regardless of the social cohesion score. Moreover, when comparing the different models, it can be seen that for GBM, no decrease in PHQ-9 scores can be observed for a social cohesion score from 5 to 11.

Discussion

Main findings in the context of available evidence

This study assessed how multiple physical and social neighborhood characteristics together are correlated with depression severity after adjusting for individual socio-demographic factors, using a ML approach. The four models fitted on our large-scale data resulted in robust evidence that demonstrates which perceived neighborhood characteristics are cross-sectionally correlated with depression severity. All ML models showed a better fit than basic regression, however, the differences were more of a statistical nature (Fig. 1).

We went a step further than previous studies [19–21] in also assessing the relative importance of individuals' perceptions of the physical and social neighborhood with respect to depression severity. Our models consistently showed that perceived physical neighborhood environment only played a minor role in explaining depression severity (Fig. 2). In contrast, social cohesion and safety were found to be important overall. Our result that the neighborhood social environment is of greater importance than the physical one replicates a study from the UK [22].

In line with a systematic review [42], we observed a negative relation between depression and age. It seems that older aged people's susceptibility to depression declines which could result from diminishing emotional responsiveness or psychological immunization against stressful situations [43]. Moreover, age was found to interact strongly with other variables, primarily personal-level (e.g., employment status [44]) and to a minor extent with environmental ones (e.g., perceived green space) (Fig. A1), over the life span. Such a co-variation [42] is not surprising because, for instance, unemployment may pose a higher risk for a young adult than someone close to retirement.

Perceived physical neighborhood characteristics including green and blue space, pleasantness, and urbanicity were found to be less important. This may partly be due to the way we assessed neighborhood features; some variables (e.g., green space) also showed limited variance. To circumvent methodological issues we employed, as frequently done [22, 45], people's neighborhood perceptions instead of geographic information system (GIS)-based measures per administrative area or buffer. Both ways cause spatial [46] and temporal context uncertainties (i.e., temporally ill-aligned GIS and survey data) [47] potentially translating into biased outcomes. Work undertaken in metropolitan Chicago found that perceived but not objectively measured neighborhood deterioration was correlated with higher depressive symptoms, which further supports our reasoning [48].

Some neighborhood characteristics were identified as relevant, but not all turned out to be related to depression severity (Fig. 3). In what follows the neighborhood characteristics are discussed in accordance to their descending order of importance (Fig. 2). First, our study supports previous findings suggesting that pronounced neighborhood social cohesion seems to correlate with reduced depression severity [45, 49]. It is assumed that in socially cohesive neighborhoods it is more likely that people help, support, and trust each other, and that a tightly knit social network may facilitate the spread of information among neighbors [50]. Through such pathways living in a cohesive environment may promote mental health.

Second, neighborhood safety was confirmed in our study to be negatively associated with depression severity. Another Dutch cross-sectional study has concluded the same [19], but overall findings are inconclusive [51]. Among different conceivable mechanisms, we speculate that living in an unsafe neighborhood enhances experienced stress, which in turn is a depression risk factor [52]. Alternatively, it has been theorized that a lack of safety limits social cohesion due to mistrusting others in the neighborhood [50].

Third, perceived traffic appeared to be positively correlated with depression severity. While our data did not allow us to disentangle pollutants emitted from traffic, we believe that air pollution and noise are conceivable underlying pathways. This is underpinned by a meta-analysis on air pollution and risk of depression [7], but contradicts a European multi-cohort study [16]. Traffic noise is regarded as a psychosocial stressor causing annoyance and negative emotions [53], and in a German study was significantly related to depressive symptoms [8, 19].

Fourth, we found pleasantness was negatively correlated with depression severity. This is in line with previous research concerning neighborhood quality and depression. For example, walkable neighborhoods have previously been associated with reduced depressive symptoms [54]. It is suggested that this is due to increased opportunity for social

interaction, which in turn can improve depressive symptoms. Poor maintenance of buildings and incivilities in the street, or neighborhood social disorder, has been linked to increased risk of depression [2]. This may be the result of reduced neighborhood satisfaction [55], or via enhanced stress [2].

Fifth, we found no indication that depression severity differed between urban and rural areas. While contradicting an international meta-analysis on mood disorders and urbanization [56], our results confirm another Dutch study reporting an insignificant correlation [19]. Further, in a recent analysis of eight Dutch cohort studies, inconsistent results were found for the effect of urbanization on depression severity [57]. It is suggested this is due to the use of different research designs, measures of depression, and confounders.

Lastly, we could not confirm that blue space within people's living environment is correlated with depression severity. This finding aligns with a series of others reporting insignificant associations on the 5% level [19, 21]. However, our findings were suggestive for beneficial mental health effects of perceived closeness to green space, though no causality can be inferred. Similar results were reported elsewhere [12, 20, 29]. The assumed mechanisms may operate through stress recovery, attention restoration, physical activity, and social interaction [14].

Strengths and limitations

A number of key strengths of this study need to be emphasized. Our study is innovative in the way correlations were assessed. While earlier studies were limited to linear associations without examining variable interactions and non-linearities [15, 19, 22], we put these challenges central and fitted flexible ML models in a data-driven manner. Our study also used a large nationally representative data set for the Netherlands. This produced a large sample size where our results are deemed to be robust. However, whether and how our findings can be generalized for a wider European or other cultural contexts needs further, ideally longitudinal [58], exploration.

Despite these strengths, several limitations are recognized. The cross-sectional nature of the data has limited capability to establish causal links. We were unable to assess whether the social causation hypothesis or the social drift hypothesis applies [59]. While the former posits that adversity linked with low socio-economic status contribute to depression, the latter argues that depressed people experience a downward drift towards neighborhoods with lower socio-economic status [59, 60]. Our findings may also be biased because depressed people might be more likely to view their environment negatively [11].

Our survey benefited from the inclusion of well-tested questionnaires (e.g., PHQ-9), which facilitates comparability

with other studies, but they may be subject to self-reporting bias. We cannot eliminate that the perception of depressed people is impaired [61]. As some survey questions relate to people's living environment, ambiguities concerning the neighborhood size and the environmental perception may arise; which potentially have attenuated the relationships. Despite the fact we adjusted for several socio-economic characteristics, another final consideration is that we cannot rule out unmeasured and residual confounding. However, our findings were robust to adjustment for many potential confounding factors but some, for example people's physical activity levels [62], were not available to us on a personal level.

Conclusions

The results reported here are from a large nationally representative sample from the Netherlands and provide support for a relationship between perceived physical and social neighborhood characteristics and people's severity of depression. The importance of the physical neighborhood environment is, however, limited relative to the social environment and individual attributes. We observed specifically that neighborhood social cohesion, pleasantness and safety were inversely correlated with depression severity, while distance to green space and traffic were positively correlated. No association was found for urbanicity and blue space. While confirmation through longitudinal research is required, our study suggests that modification of physical and social neighborhood characteristics could represent an effective intervention to promote mental health.

Acknowledgements We thank the anonymous reviewers for their valuable comments on an earlier draft of this paper.

Funding The project was funded by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (Grant Agreement No. 714993). In the course of the NEEDS project, JH's contribution was based on a research stay at Utrecht University.

Compliance with ethical standards

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Malhi GS, Mann JJ (2018) Depression. *Lancet* 392:2299–2312
- Diez Roux AV, Mair C (2010) Neighborhoods and health. *Ann N Y Acad Sci* 1186:125–145
- Vigo D, Thornicroft G, Atun R (2016) Estimating the true global burden of mental illness. *Lancet Psychiatry* 3:171–178
- Ehsan AM, De Silva MJ (2015) Social capital and common mental disorder: a systematic review. *J Epidemiol Community Health* 69:1021–1028
- Richardson R, Westley T, Gariépy G et al (2015) Neighborhood socioeconomic conditions and depression: a systematic review and meta-analysis. *Soc Psychiatry Psychiatr Epidemiol* 50:1641–1656
- Rautio N, Filatova S, Lehtiniemi H, Miettunen J (2018) Living environment and its relationship to depressive mood: a systematic review. *Int J Soc Psychiatry* 64:92–103
- Gu X, Liu Q, Deng F et al (2019) Association between particulate matter air pollution and risk of depression and suicide: systematic review and meta-analysis. *Br J Psychiatry* 215(2):456–467
- Orban E, McDonald K, Sutcliffe R et al (2016) Residential road traffic noise and high depressive symptoms after five years of follow-up: results from the Heinz Nixdorf recall study. *Environ Health Perspect* 124:578–585
- Wilson-Genderson M, Pruchno R (2013) Effects of neighborhood violence and perceptions of neighborhood safety on depressive symptoms of older adults. *Soc Sci Med* 85:43–49
- Purtle J, Nelson KL, Yang Y et al (2019) Urban-rural differences in older adult depression: a systematic review and meta-analysis of comparative studies. *Am J Prev Med* 56(4):603–613
- Gong Y, Palmer S, Gallacher J et al (2016) A systematic review of the relationship between objective measurements of the urban environment and psychological distress. *Environ Int* 96:48–57
- Helbich M, Klein N, Roberts H et al (2018) More green space is related to less antidepressant prescription rates in the Netherlands: a Bayesian geospatial quantile regression approach. *Environ Res* 166:290–297
- Kawachi I, Berkman LF (2001) Social ties and mental health. *J Urban Health* 78:458–467
- Hartig T, Mitchell R, De Vries S, Frumkin H (2014) Nature and health. *Annu Rev Public Health* 35:207–228
- de Vries S, ten Have M, van Dorsselaer S et al (2016) Local availability of green and blue space and prevalence of common mental disorders in the Netherlands. *Br J Psychiatry Open* 2:366–372
- Zijlema WL, Wolf K, Emeny R et al (2016) The association of air pollution and depressed mood in 70,928 individuals from four European cohorts. *Int J Hyg Environ Health* 219:212–219
- Klijs B, Kibele EUB, Ellwardt L et al (2016) Neighborhood income and major depressive disorder in a large Dutch population: results from the LifeLines Cohort study. *BMC Public Health*. <https://doi.org/10.1186/s12889-016-3332-2>
- Stokols D (1992) Establishing and maintaining healthy environments: toward a social ecology of health promotion. *Am Psychol* 47:6–22
- Generaal E, Timmermans EJ, Dekkers JEC et al (2019) Not urbanization level but socioeconomic, physical and social neighbourhood characteristics are associated with presence and severity of depressive and anxiety disorders. *Psychol Med* 49:149–161
- Groenewegen PP, Zock JP, Spreeuwenberg P et al (2018) Neighbourhood social and physical environment and general practitioner assessed morbidity. *Health Place* 49:68–84
- Zock JP, Verheij R, Helbich M et al (2018) The impact of social capital, land use, air pollution and noise on individual morbidity in Dutch neighbourhoods. *Environ Int* 121:453–460
- Gidlow C, Cochrane T, Davey RC et al (2010) Relative importance of physical and social aspects of perceived neighbourhood environment for self-reported health. *Prev Med (Baltim)* 51:157–163
- Fernández-Delgado M, Cernadas E, Barro S et al (2014) Do we need hundreds of classifiers to solve real world classification problems? *J Mach Learn Res* 15:3133–3181
- Helbich M (2019) Dynamic urban environmental exposures on depression and suicide (NEEDS) in the Netherlands: a protocol for a cross-sectional smartphone tracking study and a longitudinal population register study. *BMJ Open* 9:e030075
- de Groot J, Roels J, Shah S, de Witt S (2018) Stemming and Leefomgeving. Onderzoeksdocumentatie Kwaliteitsanalyse (NEEDS), Den Haag
- Kroenke K, Spitzer RL, Williams JBW (2001) The PHQ-9. Validity of a brief depression severity measure. *J Gen Intern Med* 16:606–613
- Gilbody S, Richards D, Brealey S, Hewitt C (2007) Screening for depression in medical settings with the Patient Health Questionnaire (PHQ): a diagnostic meta-analysis. *J Gen Intern Med* 22:1596–1602
- Reklaitiene R, Grazuleviciene R, Dedele A et al (2014) The relationship of green space, depressive symptoms and perceived general health in urban population. *Scand J Public Health* 42:669–676
- Stigsdotter UK, Ekholm O, Schipperijn J et al (2010) Health promoting outdoor environments-Associations between green space, and health, health-related quality of life and stress based on a Danish national representative survey. *Scand J Public Health* 38:411–417
- Spittaels H, Verloigne M, Gidlow C et al (2010) Measuring physical activity-related environmental factors: reliability and predictive validity of the European environmental questionnaire ALPHA. *Int J Behav Nutr Phys Act* 7:48
- Sampson RJ, Raudenbush SW, Earls F (1997) Neighborhoods and violent crime: A multilevel study of collective efficacy. *Science (80-)* 277:918–924
- Lorant V, Deliège D, Eaton W et al (2003) Socioeconomic inequalities in depression: a meta-analysis. *Am J Epidemiol* 157:98–112
- Hastie T, Tibshirani R, Friedman J (2017) The elements of statistical learning. Data mining, inference and prediction, 2nd edn. Springer, New York
- McCullagh P, Nelder JA (1989) Generalized Linear Models, 2nd edn. Chapman and Hall/CRC, Boca Raton, FL. ISBN 0-412-31760-5.
- Haykin S (1994) Neural networks: a comprehensive foundation. Prentice Hall PTR
- Breiman L (2001) Random forests. *Mach Learn* 45:5–32
- Friedman JH (2001) Greedy function approximation: a gradient boosting machine. *Ann Stat* 29:1189–1232
- R Core Team (2018) R Development Core Team. *RA Lang Environ Stat Comput* 55:275–286
- Kuhn M (2008) Building predictive models in R using the caret package. *J Stat Softw* 28:1–26
- Goldstein A, Kapelner A, Bleich J, Pitkin E (2015) Peeking inside the black box: visualizing statistical learning with plots of individual conditional expectation. *J Comput Graph Stat* 24:44–65
- Friedman JH, Popescu BE (2008) Predictive learning via rule ensembles. *Ann Appl Stat* 2:916–954
- Jorm AF (2000) Does old age reduce the risk of anxiety and depression? A review of epidemiological studies across the adult life span. *Psychol Med* 30(1):11–22
- Gross JJ (1998) Antecedent- and response-focused emotion regulation: divergent consequences for experience, expression, and physiology. *J Pers Soc Psychol* 74:224–237
- Jefferis BJ, Nazareth I, Marston L et al (2011) Associations between unemployment and major depressive disorder: evidence

- from an international, prospective study (the predict cohort). *Soc Sci Med* 73:1627–1634
45. Ruijsbroek A, Mohnen SM, Droomers M et al (2017) Neighbourhood green space, social environment and mental health: an examination in four European cities. *Int J Public Health* 62:657–667
 46. Kwan M-P (2012) The uncertain geographic context problem. *Ann Assoc Am Geogr* 102:958–968
 47. Helbich M (2019) Spatiotemporal contextual uncertainties in green space exposure measures: exploring a time series of the normalized difference vegetation indices. *Int J Environ Res Public Health* 16(5):852
 48. Wilbur J, Zenk S, Wang E et al (2009) Neighborhood characteristics, adherence to walking, and depressive symptoms in midlife African American women. *J Women's Health* 18:1201–1210
 49. Mair C, Diez Roux AV, Shen M et al (2009) Cross-sectional and longitudinal associations of neighborhood cohesion and stressors with depressive symptoms in the multiethnic study of atherosclerosis. *Ann Epidemiol* 19:49–57
 50. Kawachi I, Berkman L (2000) Social cohesion, social capital, and health. In: Berkman LF, Kawachi I (eds) *Social Epidemiology*. Oxford University Press, New York, pp 174–190
 51. Lorenc T, Clayton S, Neary D et al (2012) Crime, fear of crime, environment, and mental health and wellbeing: mapping review of theories and causal pathways. *Health Place* 18:757–765
 52. Chandola T (2001) The fear of crime and area differences in health. *Health Place* 7:105–116
 53. Rylander R (2004) Physiological aspects of noise-induced stress and annoyance. *J Sound Vib* 277:471–478
 54. Koohsari MJ, McCormack GR, Nakaya T et al (2019) Urban design and Japanese older adults' depressive symptoms. *Cities* 87:166–173
 55. Leslie E, Cerin E (2008) Are perceptions of the local environment related to neighbourhood satisfaction and mental health in adults? *Prev Med (Baltim)* 47:273–278
 56. Peen J, Schoevers RA, Beekman AT, Dekker J (2010) The current status of urban-rural differences in psychiatric disorders. *Acta Psychiatr Scand* 121:84–93
 57. Generaal E, Hoogendijk EO, Stam M et al (2019) Neighbourhood characteristics and prevalence and severity of depression: pooled analysis of eight Dutch cohort studies. *Br J Psychiatry* 215(2):468–475
 58. Murphy JM, Olivier DC, Monson RR et al (1991) Depression and anxiety in relation to social status: a prospective epidemiologic study. *Arch Gen Psychiatry* 48:223–229
 59. Lund C, Stansfeld S, De Silva M et al (2014) Social determinants of mental health. In: Patel V, Minas H, Cohen A, Prince M (eds) *Global mental health: principles and practice*. Oxford University Press, Oxford, pp 116–136
 60. Ritscher JEB, Warner V, Johnson JG, Dohrenwend BP (2001) Intergenerational longitudinal study of social class and depression: a test of social causation and social selection models. *Br J Psychiatry* 178:s84–s90
 61. Althubaiti A (2016) Information bias in health research: definition, pitfalls, and adjustment methods. *J Multidiscip Healthc* 9:211–217
 62. Schuch FB, Vancampfort D, Firth J et al (2018) Physical activity and incident depression: a meta-analysis of prospective cohort studies. *Am J Psychiatry* 175:631–648