UNIVERSITÉ PARIS SUD

MASTER THESIS

Use of Healthcare Services in the Residence and Workplace Neighbourhood:

The Effect of Spatial Accessibility to Healthcare Services

Student:
Ruben BRONDEEL

Supervisor:
Dr. Basile CHAIX

A thesis submitted in fulfilment of the requirements for the degree of Master

in the master program

Master 2 de Santé Publique - Recherche
Parcours Epidémiologie

June 2013

Acknowledgements

I would like to thank my supervisor, Dr. Basile Chaix, for giving me the opportunity of an internship in the RECORD Team and for the advice while completing the master thesis. I am very grateful to the other members of the RECORD Team for their advice and especially for making the internship a very enjoyable working experience. The corrections and suggestions to the thesis by prof. Nicole Vettenburg are also very much appreciated.

Special thanks go to my family for their constant support during my education.

Abstract

Use of Healthcare Services in the Residence and Workplace Neighbourhood: The Effect of Spatial Accessibility to Healthcare Services

by Ruben BRONDEEL

Objective. Previous literature on the effects of the spatial accessibility to healthcare services on service use has exclusively focused on the residential environment of people. This study investigates the effect of spatial accessibility to healthcare services in residence and workplace neighbourhoods on the use of healthcare services.

Method. Data from the first wave of the RECORD Cohord Study were used, from questionnaires or provided by public institutions (SNIIR-AM, CNAV). The dataset contained geographical information on the participants and healthcare services, the use of the services (through the linkage of administrative data) and demographical characteristics. To process the geographic information, Geographic information system (GIS) methods were applied. A novel method was developed to examine whether and for which participants there was clustering of the visits to healthcare services around the workplace. We examined the associations between spatial accessibility indicators and the use of four healthcare services: general practitioners, gynaecologists, cardiologists and psychiatrists. Negative binomial mixed regression models were used to test the associations between spatial accessibility to services from the residence and the workplace and the use of the four healthcare services.

Main Findings. A clustering of the use of healthcare services around the residence was found for most people. For only a small proportion of the participants (11%), we found also a clustering around the workplace. A logistic regression indicated that this use of services around the workplace was associated with commuting from the suburb to Paris, a high distance of commuting, and with a high occupational class and a high family income. No associations were found between the spatial accessibility to healthcare services and the use of healthcare services, neither in the residence neighbourhood nor in the workplace neighbourhood.

Conclusions. The use of healthcare services clustered around the workplace only for a small proportion of the participants. Spatial accessibility does not seem to have an influence on the use of healthcare services in a well-served area as Ile-de-France. Future research could benefit from focusing on how an individual overcomes spatial access barriers.

Contents

List of Figures

List of Tables

1 Introduction

1.1 Societal and scientific background

Accessibility to healthcare services has been a concern for researchers and policy makers for several decades [1–3]. A lack of accessibility to healthcare services is one of the major reasons for underusing or misusing healthcare services [4], which has implications for the patients health and efficiency in public health [5, 6]. A better accessibility to healthcare services is thus believed to have an effect on health outcomes through the use of healthcare services. Not only the size of the relation between health outcomes and accessibility makes accessibility an interesting topic for public health research. It is also a cause of health outcomes modifiable by and under direct responsibility of public health policy makers.

Throughout the literature, there are many definitions of accessibility proposed, due to the many dimensions of accessibility [7–9]. A useful definition is given by Penchansky and Thomas [1], defining accessibility to healthcare services in five sub-dimensions: accessibility, availability, acceptability, accommodation and affordability. These dimensions represent the spatial, social and economical context of accessibility. As Fortney, Rost, and Warren [10], we will focus on the two spatial dimensions defined by Penchansky and Thomas [1], namely accessibility and availability. Note that the term accessibility refers here to the distance to the nearest healthcare service. To overcome the ambiguity with other uses of accessibility, we will refer to this as proximity in the rest of this paper. The second spatial dimension is spatial availability[1], define by Penchansky and Thomas [1]

[1]As for the term 'accessibility', we acknowledge that the term 'availability' can be used in different ways. We prefer to follow Penchansky and Thomas [1], although some people prefer to divide the number physicians by the number of inhabitants in the neighbourhood. This would adjust the measure for the amount of people one has to share the physicians in her/his neighbourhood.

as the amount of healthcare services available in a predefined area. The concepts proximity and spatial availability indicate both the pure spatial dimension of accessibility and differentiates itself from dimensions such as individual mobility.

The interest in accessibility to healthcare services in the literature - and especially in spatial accessibility - has been revitalized recently through several interrelated trends within public health and epidemiology [11]. From a theoretical point of view, there is a common understanding that individual-based explanations of health outcomes are not sufficient and should be accompanied with group-based explanations. Group-based explanations can include the influence of neighbourhoods, work environment, network of peers, ... The (renewed) interest in these group based disparities within health research is strongly related to the methodological improvements. Multilevel modelling as a statistical development and more recently the very popular Geographic Information Systems (GIS) as a measurement development, have given a big impulse to revise and refine earlier findings as well as introducing completely new hypotheses [11].

Supported by these methodological developments, residence neighbourhood characteristics have been found to have an effect on health related variables [12]. More specifically, in research traditions as health geography, spatial accessibility is found to be an important determinant in treatment seeking behaviour [10]. Since people tend to limit the use of healthcare services to a relatively small area around their residence, neighbourhood differences in spatial access barriers are found to have an effect on healthcare use. Chaix et al. [5] found this for an elderly population with specific mobility issues, whereas Carr-Hill, Rice, and Roland [13] presents the same result but for a general population. The influence on healthcare use of spatial accessibility is not only found in rural but also in urban and well-served areas [14]. However, focusing on specialists only, this is not confirmed [15] or only partly confirmed (for men only) [16]. The exposure to neighbourhoods outside the residence is also found to be relevant to health outcomes

[12]. However, to our knowledge, no research has been done on the spatial accessibility to healthcare services in non-residential neighbourhoods and its effect on the use of healthcare services.

Social and contextual epidemiology pointed out the social importance of neighbourhood differences in healthcare use. The neighbourhood differences become important social differences, when considering that more socially and economically vulnerable people have worse health outcomes [12, 14] and therefore a higher need of healthcare services. At the same time, these people tend to live in neighbourhoods with less spatial accessibility to healthcare services [14]. So, the groups with the highest need of healthcare services are those with the lowest (spatial) accessibility.

1.2 Objectives of the study

In epidemiological studies on neighbourhood and health related behaviour, it has been found that people seek resources in the residence neighbourhood [17–19] and workplace neighbourhood [20]. Here, we hypothesize that people will turn to healthcare services in both the residence and the workplace neighbourhoods. Our first hypothesis is thus twofold:

Hypothesis 1.A. There is a clustering of the use of healthcare services in the residence neighbourhood.

Hypothesis 1.B. There is a clustering of the use of healthcare services in the workplace neighbourhood.

In earlier research, associations have been found between spatial accessibility and the use of healthcare services. However, most of the previous research has been done either in western societies comparing rural to urban areas [6, 21, 22], or in deserted areas in third world countries [23–26]. Within urban and peri-urban areas, there is also a variation in spatial accessibility. Supported by findings of Chandola [14], we examine if this variation within an urban and peri-urban area has an influence on the use of healthcare services. This leads to our second hypothesis:

Hypothesis 2. Spatial accessibility to healthcare services in the residence neighbourhood is associated with a person's use of healthcare services, even in the relatively well-served Ile-de-France region.

If it also appears that people use healthcare services in the vicinity of the workplace (Hypothesis 1.B), this quite naturally leads us to the third and last hypothesis.

Hypothesis 3. Spatial accessibility to healthcare services in the workplace neighbourhood is associated with a person's use of healthcare services.

Finally, we do not make the assumption that spatial accessibility to healthcare services has the same effect on the use of all types of healthcare services. It can be assumed that low spatial accessibility is more easily overcome for health conditions that need more specific intervention [3]. However, in Chaix et al. [16] and Saag et al. [22] there are associations found between the use of specialised healthcare services and spatial accessibility. Therefore, Hypotheses 2 and 3 are tested for four types of healthcare services separately; two services frequently used, viz. general practitioners and gynaecologists and two services for more specific treatments, viz. cardiologists and psychiatrists.

2 Method

2.1 Participants

In this study, we used the first wave of the RECORD Cohort Study (Residential Environment and CORonary heart Disease)[2]. The RECORD Cohort Study was established to investigate environmental determinants of territorial disparities in health [27]. It is a collaboration between the 'Groupe RECORD - UMR-S 707', the 'Centre d'Investigations Préventives et Cliniques' and the 'Centre de Recherche du Centre Hospitalier de l'Université de Montreal'[3].

Data from the first wave of the RECORD Cohort Study [27] were used in cross-sectional analyses. Overall, 7,290 participants, between 30 and 79 years old, were recruited between March 2007 and February 2008 during free preventive medical examinations [27–29]. The medical examinations are offered every five years by the French National Health Insurance System for Salaried Workers to all working and retired employees and their families. The participants to the RECORD Study were recruited during the medical examination in one of the four centres of the Centre d'Investigations Préventives et Cliniques (IPC) located in the Paris metropolitan area (Paris, Argenteuil, Trappes and Mantes-la-Jolie). People not insured by the National Health Insurance System for salaried workers could not be recruited for the RECORD Study: self-employed occupations (lawyers, architects, etc.), shopkeepers, craftsmen, farmers and salaried farm

[2] See www.record-study.org.

[3] The RECORD Study is funded by the Institute for Public Health Research (IReSP, Institut de Recherche en Santé Publique); the National Institute for Prevention and Health Education (INPES, Institut National de Prévention et d'Education pour la Santé) (Prevention Program 2007; 2010–2011 financial support; 2011–2013 financial support; 2012-2014 financial support); the National Institute of Public Health Surveillance (InVS, Institut de Veille Sanitaire) (Territory and Health Program); the French Ministries of Research and Health (Epidemiologic Cohorts Grant 2008); the National Health Insurance Office for Salaried Workers (CNAM-TS, Caisse Nationale d'Assurance Maladie des Travailleurs Salariés); the Ile-de-France Regional Health Agency (ARS, Agence Régionale de Santé); the Ile-de-France Regional Council (Conseil Régional d'Île-de-France, DIM SEnT and CODDIM); the National Research Agency (ANR, Agence Nationale de la Recherche) (Health–Environment Program 2005); the City of Paris (Ville de Paris); and the Ile-de-France Youth, Sports, and Social Cohesion Regional Direction (DRJSCS, Direction Régionale de la Jeunesse, des Sports et de la Cohésion Sociale).

workers. However, in the Ile-de-France region, working and retired employees and their families represent almost 95% of the population.

A priori, 10 (out of 20) administrative districts of Paris and 111 municipalities in the metropolitan area were selected for the study. The selection was based on a weighted sample that favours districts and municipalities of which was known that relatively many inhabitants would visit one of the four centres during the recruitment period. Favouring these districts and municipalities led to a sample with enough inhabitants from the same region to estimate neighbourhood effects. The selection also ensured inclusion of areas with different socio-economic backgrounds and from urban and peri-urban areas. No a priori sampling was performed on the level of the participants. Of the persons contacted for participation during their visit at the IPC medical centres, 83.6% agreed to participate and completed the data collection protocol. The French Data Protection Authority approved the study protocol.

From these 7290 participants, a selection had to be made for the analyses in this study. To answer our hypotheses, people were excluded when they were not working (n = 2787) or living closer than 2 km away from their work following the street network (see below: GIS) ($n = 440$). To clearly distinguish between the spatial access barriers in the residence neighbourhood and those in the workplace neighbourhood, there could not be overlap between the two neighbourhoods. For people living closer than 2 km away from their workplace, healthcare services could be closer than 1 km[4] away from both workplace and residence. Therefore, it is impossible to define the spatial access barriers to that healthcare service as a workplace characteristic or as a residence characteristic. People working outside of Ile-de-France ($n = 124$) had to be excluded since we had only information for the healthcare services of Ile-de-France. Other participants were excluded from the analyses if the workplace could not be located ($n = 48$), if the use of the healthcare services was not known ($n = 64$) or if there were missing values on one or more self-reported variables used in the analyses ($n = 113$). All analyses were performed

[4]The choice for a 1 km road distance will be explained later in this chapter.

on the cases for which full information was available in order to have a constant sample size and a stable sample throughout all our analyses.

2.2 Data Sources

For all the participants in the RECORD Study, multiple data sources were available. Firstly, the interviewer of the RECORD Study administered a survey to the participants. People were invited to report personal (e.g. education and income) and neighbourhood related information (e.g. neighbourhood satisfaction). Next to this general survey, the IPC medical centers administred a medical survey. The data from this survey as well as the results of the medical tests performed by the IPC centers, were made accessible for the RECORD Team.

Furthermore, several institutes provided the RECORD Study with additional data essential for this project. The 'Système National d'Informations Inter Régions d'Assurance Maladie' (SNIIR-AM) provided data on the use of healthcare services reimbursed by SNIIR-AM from 2006 to 2011 for all the participants. The data contains information on the health professionals the participants consulted and the date of the consultation. Thanks to the 'Institut d'amenagement et d'urbanisme' (IAU), all healthcare services in Ile-de-France could be geographically located. Linking the SNIIR-AM and the IAU data, it was possible to know for each participant which healthcare services was used, how many times and where these healthcare services were geographically located. The 'Caisse Nationale d'Assurance Vieillesse' (CNAV) provided us with the business identification code for the companies where the participants worked, as well as their salaries. The file received from the CNAV gave us yearly information on our participants. It indicated the amount of salary and the name of one or more different companies, but not the exact date of change. Based on the CNAV files, it was impossible to be certain about the work situation on the exact date of the recruitment for the study. In case of multiple employers within the same year, the workplace was determined as the employer of whom a participant had received the most wage. The 'Institut national de la statistique et des études économiques' (INSEE) gave us accessibility to the business identification codes for all companies in Ile-de-France together with the geographical coordinates. Thus, linking the data of the CNAV with those of the INSEE allowed us to geographically locate[5] most of the companies where the participants worked. About 6% of the workplaces had to be geolocated using Google Maps.

2.3 Variables

2.3.1 Use of Healthcare Services

We defined the use of healthcare services as the number of visits of a person to healthcare services. We did not distinguish between a participant that visits 10 times the same professional and a participant that visits 10 different professionals once. The variable contains the information on the number of visits in the 18 months following the date of recruitment, as reported in the data provided by the SNIIR-AM.

To describe the clustering of the use of healthcare services (Hypotheses 1.A and 1.B), the use of four services was included: general practitioners, dentists, pharmacies and medical laboratories. These are the four main healthcare services one would go to without a prior visit to a general practitioner. The number of visits for each type of service was determined for the 1 km road network buffers around residence and workplace and the 5 km combined buffer[6]. The use of these four services was summed up, resulting in one variable for each buffer.

To test the effect of spatial accessibility on the use of healthcare service, four services were considered: general practitioners, gynaecologists, cardiologists and psychiatrists. The

[5]Geographically locating or geolocating means that for each entity x and y coordinates are determined based on a pre-defined projection method. The geographical projection used in all analyses with ArcGIS is known as 'NTF Lambert II'.

[6]see section 2.4.1 for more information on the road network buffers.

four resulting variables reflected the total number of visits to each of these healthcare services within the Ile-de-France region. Each variable was analysed separately.

2.3.2 Spatial Accessibility: a GIS-approach

Spatial accessibility was measured by two variables: proximity and spatial availability. Proximity is defined as the road distance (in km) to the closest healthcare service. Spatial availability is defined as the amount of healthcare services within a given road distance. Determining the best road distance is a fairly arbitrary process. In this study, 1 km road distance was selected, based on several arguments. Firstly, other relevant research in densely populated areas used the same distance [7, 30]. Secondly, we have done some descriptive sensitivity analyses. For the four healthcare services considered in Hypotheses 1.A and 1.B, the Pearson correlation between the spatial availability in a buffer of 1 km around the residence and the spatial availability in buffers of 500 m and 2 km was respectively 0.78 and 0.89. For the spatial availability around the workplace, these correlations were respectively 0.90 and 0.84. Therefore, the results of analyses can be expected to be robust for the choice of road distance. Finally, a maximum walking distance of 1 km is reasonable. For longer distances than 1 km, people would use transportation modes for which the notion of neighbourhood is less important.

For both proximity and spatial availability, one could argue that the travel time instead of the road distance should be considered. Travel time, or the 'cost of space', is one of the costs associated with receiving treatment [10]. Without arguing that the travel time would not be a better measure, we give here a few arguments for using the road distance. Firstly, road distance is easier to measure than travel time. For the road distance, only the two addresses are needed. To measure travel time, the type of transport and/or other self-reported information is needed for every visit to a healthcare service. Considering all the visits a person can do in 18 months, this can lead to missing or biased data. Secondly, the spatial accessibility variables in this study apply to very short distances.

Spatial availability of healthcare services is measured for a 1 km network distance and the proximity rarely exceeds 1 km[7]. For these short walking distances, the transportation mode will not add much information in terms of spatial accessibility. Even in a less well-served area in Arkansa, Fortney et al. [10] found that virtually all of the variation in travel times could be explained by the road distances.

A Geographic Information System (GIS) approach

To measure the spatial access barriers, we processed the geolocated data in ArcGIS[8]. ArcGIS allows to position x and y coordinates on a map and to use the coordinates for spatial analyses. As an example, the data for one participant is shown in Figure 1. The participant's workplace and residence addresses are shown as respectively a square and a rectangle. The location of the practices of general practitioners are shown in Figure 1 by the dots. Finally, two street network buffers are shown around the workplace and the residence. These two buffers comprise all the locations within a 1 km road distance. In other words, they comprise all locations one can walk to, following the road network for 1 km or less[9]. After geographically locating all these places and buffers on a map, we calculated in ArcGIS for each participant the distance (in km) of the nearest healthcare service to the workplace and the residence, i.e. workplace proximity and residence proximity; and how many healthcare services were located in the 1 km buffers, i.e. workplace spatial availability and residence spatial availability. To correspond with the utilization variables, the spatial accessibility variables were constructed for the four healthcare services separately.

2.3.3 Other Variables

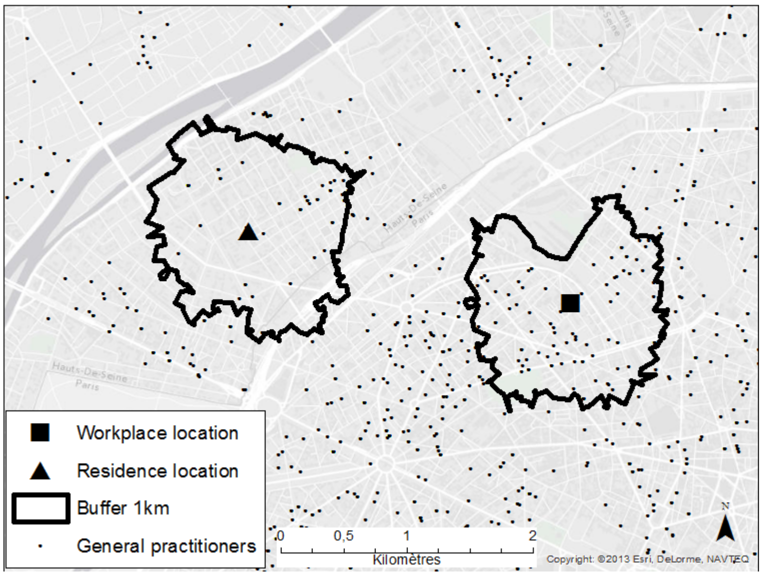
Individual sociodemographic variables used in the analyses were age, gender, education level, attachment to neighbourhood, occupation, household income per consumption

[7] see Results for more descriptive analyses on spatial accessibility variables.

[8] Processes in ArcGIS are automated with Python. More information on www.arcgis.com.

[9] The road network data was provided by the 'Institut National de l'Information Géographique et Forestière' (IGN). The analyses are performed in the module 'Network Analyst' of ArcGIS.

FIGURE 1: Data Handling and Data Analyses in ArcGIS - An Example by a Participant's Situation



unit, perceived financial strain and human development index of country of birth. One neighbourhood sociodemographic variable was included in the analyses: neighbourhood educational level. Finally, three spatial variables were included: commuting distance (distance work - home in km), living in Paris or in the rest of Ile-de-France and working in Paris or not.

Age was considered a continuous variable and was measured in years. Individual education was measured in three categories: (1) no education, primary or lower secondary education; (2) higher secondary and lower tertiary education; and (3) upper tertiary education. Attachment to the neighbourhood was measured by four questions; indicating the importance of the neighbourhood to the participant, the wish to keep the residence in the neighbourhood, whether or not having a good feeling in the neighbourhood and the degree to which the participant is proud about the neighbourhood. Occupation was measured in five categories[10]: (1) high white-collar workers; (2) intermediate occupations; (3) low white-collar workers; (4) blue-collar workers; and (5) a diverse category. Household income adjusted for household size was measured in Euro per consumption unit. Adults are considered one consumption unit and children half a unit. Perceived

[10]In accordance with the French National Institute of Statistics and Economic Studies.

financial strain was a dichotomous variable indicating presence of regular personal financial difficulties[11]. The Human Development Index was developed by the World Bank and has a continuous scale[31]. The index was based on the live expectancy, education and income indices of a country. Each participant was attributed the score of the index according to the self-reported country of birth[28]. The variable reflects the immigration status of a person, as well as aspects of the social context of the country of origin. Neighbourhood education level was the mean education level of the people living within the 1 km road network buffer around the participant's residence, based on all participants in the RECORD Study.

2.4 Statistical Analyses

To test the three hypotheses, we used two different methods. In order to describe the clustering of the use of healthcare services in the workplace neighbourhood (Hypothesis 1.A and 1.B), we performed for each individual a mid-P-value exact tests based on the binomial distribution. To test the effects of spatial accessibility on the number of visits of healthcare services (Hypotheses 2 and 3), we used multilevel negative binomial regression models.

2.4.1 Individual Mid-P-value Exact Tests

The use of healthcare services is considered to be clustered around the workplace when a person uses more healthcare services in direct proximity around the workplace than expected by chance. The direct proximity around the workplace is defined as a buffer based on 1 km road distance. The reference area is a combined buffer: the combination of the buffer based on 5 km road distance around the workplace and the buffer based on 5 km road distance around the residence. For people living closer than 10 km away from their workplace, there is an overlap between the 5 km residence buffer and the 5

[11]Question asked: During a regular month, are there moments when you encounter real financial difficulties to pay for basic needs (food, rent, electricity, ...). Answer possibilities: Yes or No.

km workplace buffer. To avoid counting the same healthcare services twice, we used the combined buffer. For the people with overlap between the two buffers, the combined buffer is smaller than for other people. However, we argue that the effort to go to a healthcare service within the combined buffer is equal for all people: a journey of maximum 5 km from either the workplace or the residence.

Within epidemiology, clustering given a known centre is called centre focused clustering [32]. These tests, are typically used to evaluate spatial patterns of exposure and disease. These tests are not applicable to our study however. From a practical point of view, there is no method yet to automize these tests over all 3777 persons. More importantly, the assumption made for these tests is that the spatial pattern for disease follows the spatial pattern of exposure. In other words, there are high exposure levels for a known centre of contamination and lower levels the further away from this centre. The decline in exposure levels can be modelled linear or non-linear. The patterns of disease are then believed to follow this pattern. In this study, there is no centre of exposure. For most cases, the exposure to healthcare services is quite evenly spread within a small region around residence or workplace. Finally, there is also a statistical problem to use the focused clustering tests. Most, if not all, of these tests are based on maximum likelihood (ML) estimation. It is well known that ML estimation is not very well equipped for estimation close to its range boundaries. With a considerable number of cases never using healthcare services or sporadically using them, the assumption of normally distributed estimates are no longer justified.

Therefore, an indicator was created for the present study based on the mid-P-values of the Clopper-Pearson exact test. The clustering problem translates to a simple inference problem based on the binomial distribution. For every individual, $p(k) = \binom{n}{k}\pi^k(1 - \pi)^{n-k}$, where k is the number of visits to healthcare services in the 1 km buffer around the workplace; $p(k)$ is the probability of observing k; n is the number of visits to healthcare services in the combined buffer of 5 km around the workplace and 5 km around the residence; and π is the probability under the null hypothesis H_0. This probability π is the ratio of the number of available healthcare services in the 1 km workplace buffer to the number of available healthcare services in the 5 km combined buffer. Under H_0, the probability of a visit to a healthcare service in the 1 km workplace buffer $(p(k))$ given the total of visits (n), is equal to or smaller than the proportion of healthcare services in the 1 km buffer given the total of healthcare services in the 5 km combined buffer. The alternative H_a is thus one-sided: the proportion of visits is higher in the 1 km buffer than expected by the proportion of healthcare services in the 1 km buffer. The use of healthcare services is considered to be clustered around the workplace if H_0 is rejected.

The ordinary Wald test is too optimistic for binomial distributions with probabilities close to 0 [33]. With small observed probabilities, the discrete nature of a count variable can not be ignored. The continuous approximation of the Wald test makes no longer sense. Therefore we used the Clopper-Pearson test, with mid-P-values. The Clopper-Pearson test is an exact test. The principle is to calculate the probability of every possible value k and then add the probabilities of the observed k and all values higher than k (for a one-sided test with a higher alternative). So, the Clopper-Pearson P-value $= P(\Pi \geq \pi_0)$. The Clopper-Pearson test is too conservative by nature. Therefore, a mid-P-value correction is suggested [33]. The mid-P-value $= \frac{1}{2}P(\Pi = \pi_0) + P(\Pi > \pi_0)$. This is an ad hoc correction, but simulation studies [33] have shown that this correction gives trustworthy p-values.

Testing the spatial clustering for each person results in a mid-P-value for each individual. These mid-P-values can be described and used to compare participants for different background variables.

2.4.2 Multilevel Negative Binomial Regression Models

To test the Hypotheses 2 and 3, a model is needed that allows two extensions to the well known generalized linear model for Poisson distributed data. In this traditional model for count data, the assumption is made that the mean equals the variance.Since several decades, it is clear that this assumption is too restrictive for most count data [33, 34]. In most cases, the variance of a count variable is larger than its mean. This is a phenomenon called overdispersion. One of the most popular extensions to the Poisson model to overcome this problem, is the use of the negative binomial model. In a negative binomial model, a dispersion parameter is introduced to allow for a variance bigger than the mean. In fact, one could consider Poisson regression model a special case of the negative binomial regression model where the dispersion parameter is equal to zero [33]. This model has also been used by Carr-Hill et al. [13] when analysing general practice consultation rates.

A second assumption in regular Poisson regression models violated in this data, is the assumption of independence between individuals. We consider the participants to have a spatial dependence, represented by a spatial cluster variable (TRIRIS) based on the census tracts created by the French National Institute of Statistic and Economical studies[12]. In case the algoritme did not converge with the TRIRIS variable, a regrouping was done, resulting in a variable indicating the departments and municipalities. To account for these dependencies, mixed models (or random effect models) can be used [33, 35]. Recently, mixed models have been used often in epidemiological and other research. However, the combination of mixed models for hierarchical data modelling and negative binomial models are less well known [34, 36].

The model can be written as following. Given that Y_{ij} denotes the outcome of the jth individual measured for cluster i

$$P(Y_{ij} = y_{ij}|\boldsymbol{b_i}) = \binom{\alpha_j + y_{ij} - 1}{\alpha_j - 1} \left(\frac{\beta_j}{1 + \kappa_{ij}\beta_j}\right)^{y_{ij}} \left(\frac{1}{1 + \kappa_{ij}\beta_j}\right)^{\alpha_j} \kappa_{ij}^{y_{ij}}, \tag{1}$$

where $\kappa_{ij} = \exp(x'_{ij}\boldsymbol{\beta} + z'_{ij}\boldsymbol{b_i})$ [34].

[12]The IRIS variable (Ilots Regroupés pour l'Information Statistique) is created by the French National Institute of Statistic and Economical studies (INSEE). More information on this variable can be found on www.insee.fr. The TRIRIS variable regroups three census tracts of the original IRIS variable.

The fixed effects $\boldsymbol{\beta}$ can be interpreted in the same way as in an ordinary Poisson regression for rates. The link function used in this model is the logarithm. Therefore, the natural exponent of the estimations can be interpreted as rate ratios.

Other models have been proposed [33] and applied [37, 38] to overcome the overdispersion problem in the context of the dependent data. Zero-inflated Poisson (ZIP) mixed models and Zero-inflated Negative Binomial (ZINB) mixed models are created for the specific case where the overdispersion is caused by an excessive amount of zeros. Zero-inflation models consider two subpopulations within the general target population of the study. The 'always zero' population includes those cases that - given there profile on the independent variables - are estimated to always have a zero value. The 'not always zero' population includes those cases that might have any positive value, including zero. To illustrate this idea, consider the use of cardiologists. We can imagine there is a subpopulation for which there is no need to visit a cardiologist. On the other hand, there is a subpopulation that does have a need to visit a cardiologist, with a variation in the number of visits including zero. In other words, there is a differentiation between an observed zero due to a lack of need and a observed zero due to other reasons. More practically, the zero-inflated models are mixtures models, fitting 2 models simultaneously. A logistic model distinguishes an 'always zero' population and a 'not always zero' population. And a count model (with Poisson or negative binomial distribution) that links the variation in observed counts within the 'not always zero' population.

Considering the complexity of these models, it is advised to use more parsimonious models when possible. Selecting the best model (Poisson, negative binomial, ZIP or ZINB) for a certain analysis, the fit of the models can be compared with the Bayesian Information Criterium (BIC)[13]. As far as we know, there are no exact tests available (yet) to compare these four count models in the context of dependent data. All models are fitted and compared with the glmmADMB package in R [39, 40].

[13] Also known as Schwarz's Bayesian criterion (SBC or SBIC).

All models on the use of healthcare services were tested for outliers. Outliers are defined by standardized residuals outside the interval [-3,3] based on regular negative binomial models. We have to rely on these simpler models, since no such tests for outliers are available for negative binomial mixed models. The models were also tested for nonlinear associations between the use of healthcare service and the spatial accessibility indicators. Interaction effects were tested between residence and workplace spatial accessibility indicators; as well as the interaction effects between gender and the spatial accessibility indicators. Where relevant, these analyses will be discussed.

3 Results

3.1 Descriptive data

The 3777 participants in our sample were predominantly men (72.1%) and had a mean age of 46 years (se = 9.4). Overall, 1075 (28.5%) participants had obtained no, primary or lower secondary education; 1109 (29.4%) obtained higher secondary or lower tertiary education and 1593 (42.2%) obtained higher tertiary education. Their mean household income per consumption unit was 1700€. In our sample, 1180 (31.2%) participants were born outside of France. The participants lived in 641 different TRIRIS census tracts.

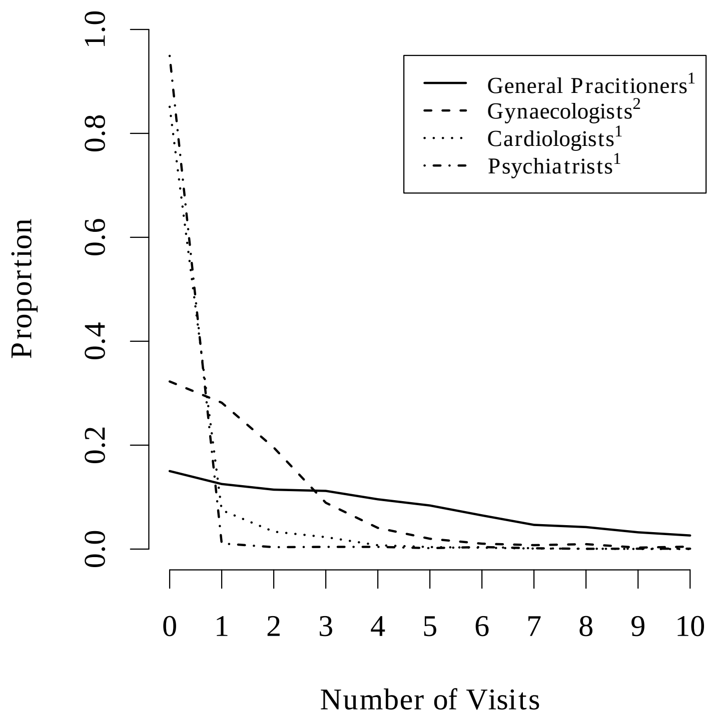
In Appendix A.1, descriptive data is given for the observed use of healthcare services. However, since the distributions are very skewed, a plot is more informative. In Figure 2, the proportions are shown for the number of visits to respectively general practitioners, gynaecologists, cardiologists and psychiatrists. In order to keep the graph readable, number of visits higher than 10 to a certain healthcare service are not shown even though the highest go up to 102 (for vistis to psychiatrists).

Figure 2 indicates a very high skewness for all four variables, especially for cardiologists and psychiatrists. Eighty five percent of the participants did not go to a cardiologist in the 18 months following the recruitment for the study; whereas 95% did not got to a psychiatrist. For exactly one visit, the proportions drop to respectively 8.5% and 1.1%. For higher numbers of visits, Figure 2 indicates proportions of nearly 0.

3.2 Clustering of Healthcare Service Use

Hypotheses 1.A and 1.B state that there is a clustering of use of healthcare services in respectively the residence and the workplace neighbourhood. We described the mid-P-values and linked them to variables of interest. From the total of 3777, two participants had to be excluded from these analyses since they had a workplace buffer crossing

FIGURE 2: Number of Visits to 4 HC Services



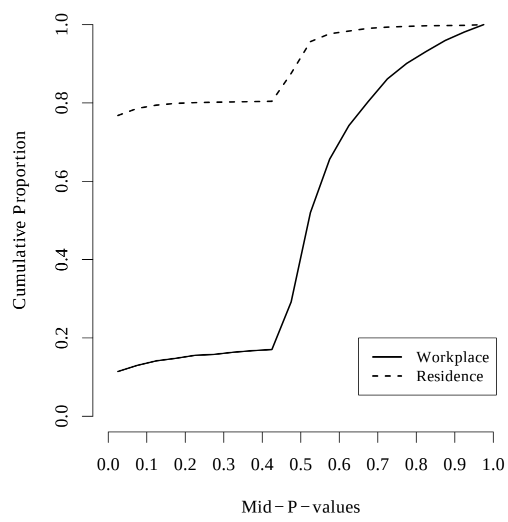
Note: ^{1}n = 3777; ^{2}n = 1054 (only women)

the border of Ile-de-France. Therefore there was missing data on the use and spatial availability of healthcare services within the evaluated buffers.

Figure 3 presents the cumulative proportions of the mid-P-values for clustering around the residence and around the workplace. The clustering was higher when the mid-P-value approaches 0. The full line indicates the cumulative proportions of the mid-P-values for the clustering test in the workplace neighbourhood. About 11% had a mid-P-value lower than 0.05. As expected, the clustering was higher around the residence than around the workplace. About 77% of the participants had a mid-P-value lower than 0.05.

Appendix A.4 presents logistic models on the probability of mid-P-values smaller than 0.05. There was a higher probability of clustered use around the residence for older people ($\beta = 0.022$, s.e. $= 0.005$, $p < 0.001$), women ($\beta = -0.520$, s.e. $= 0.102$, $p < 0.001$) and people with less commuting distance ($\beta = -0.017$, s.e. $= 0.005$, p

FIGURE 3: Mid-P-values of Clustering Tests



Note: n = 3775

< 0.001). Slightly higher probabilities were found for people with less family income ($\beta = -0.094$, s.e. = 0.042, p = 0.027).

The regression on the cluster indicator around the workplace (Appendix A.4, column 2), showed a higher probability for women ($\beta = -0.705$, s.e. = 0.118, p < 0.001), people with a lower occupation (gradual decline in numbers of visits between different occupation classes, LRT-Chi-square = 30.692, df = 4, p < 0.001), people with a higher income ($\beta = 0.262$, s.e. = 0.047, p < 0.001) and people living closer to the workplace ($\beta = -0.062$, s.e. = 0.010, p < 0.001). There was also an interesting interaction effect between the workplace location and the residence location ($\beta = -0.631$, s.e. = 0.237, p = 0.008). This interaction effect indicated that people working in Paris have a higher probability of clustered use around the workplace, except if they also live in Paris. In other words, the group of people coming to Paris to work, also used the healthcare services around the workplace. Finally, there was a slightly higher probability of use around the workplace for people living in a neighbourhood with a higher educational level ($\beta = -1.151$, s.e. = 0.527, p = 0.029).

3.3 Use of Healthcare Services and Spatial Accessibility

Spatial accessibility was measured by two indicators: proximity (distance to nearest healthcare service) and spatial availability (number of services in a 1 km road network buffer). Previously, these measures were used as indicators of two different phenomena [1, 10]. However, when the services were perfectly evenly distributed within an area, the measures were perfectly correlated (except for some noise). Therefore, in a densely populated and well-served area like Ile-de-France, it is more sensible to use the two variables as two indicators of the same phenomena: spatial accessibility. The Pearson correlations between the two indicators ranged between -0.47 for the spatial accessibility to general practitioners in the workplace neighbourhood, and -0.39 for the spatial accessibility to general practitioners in the residence neighbourhood[14]. To avoid over-controlling, all analyses were done for the two indicators separately.

The highly skewed distribution for the visits to healthcare services (see Figure 2) suggests that regular Poisson mixed models will not be appropriate to analyse these variables. Four models were compared by the BICs in Table 1: Poisson mixed models with the Negative binomial mixed models and the zero-inflated models (ZIP and ZINB). This comparison has been done for each type of healthcare service. The BICs were extracted from models with all independent variables included (see Appendices A.7 and A.8 for a full list of independent variables). For the zero-inflated models, there were only constant terms included for the logistic regression[15]. The best fitting model for the number of visits to gynaecologists, cardiologists and psychiatrists was the regular negative binomial mixed model. For the general practitioners, the ZINB had a better BIC than the negative binomial model. We decided to use the negative binomial model for all following regressions. This allowed us to use the same model for all dependent variables, and the negative binomial model is more parsimonious than the ZINB model. The lack of

[14]Correlations are calculated for general practitioners, gynaecologists, cardiologists and psychiatrists in both the residence and workplace neighbourhood.

[15]To date, the glmmADMB package does not allow for other terms to be included in the logistic part of zero-inflated models.

convergence for several models based on the Poisson distribution, confirmed that these models were not appropriate.

Table 1: Comparing Fit of Count Models by BIC

	General.	Gynaec.	Cardio.	Psychiat.
Proximity				
Poisson	26,313.3	4,321.3	*nc*	*nc*
Negative binomial	19,990.3	3,839.8	4,853.1	3,026.5
ZIP	*nc*	4,095.6	4,955.4	*nc*
ZINB	19,977.4	3,846.8	4,861.3	3,034.7
Spatial Availability				
Poisson	26,300.7	4,328.0	6,240.5	*nc*
Negative binomial	19,993.8	3,845.9	4,856.6	3,035.5
ZIP	*nc*	4,102.8	4,966.6	3,641.4
ZINB	19,980.4	3,852.9	4,864.8	3,043.7

Note: nc = No Convergence obtained

3.3.1 Effect of Residence Neighbourhood Spatial Accessibility

Hypothesis 2 states that spatial accessibility to healthcare services in the residence neighbourhood are associated with the use of healthcare services, even in the relatively well-served Ile-de-France region. Table 2 presents the results of four analyses on each of the four dependent variables: the number of visits to respectively general practitioners, gynaecologists, cardiologists and psychiatrists. Only the coefficients for the spatial accessibility variables in the residence neighbourhood are presented. The complete regression outputs can be found in Appendices A.5 and A.6. The first and the second line of Table 2 presents the association of proximity with the number of visits based on respectively unadjusted and adjusted regressions. In the same manner, the third and forth line present the association of spatial availability with the number of visits. Adjustment was done for the distance between home and work, living in Paris (yes/no), working in Paris (yes/no), the perceived attachment to the residence neighbourhood, the age, the gender, the educational level, the occupation, the household income per member, the perceived financial strain and the Human Developmental Index of the country of origin.

TABLE 2: Use of Healthcare Services and Residence Spatial Access Barriers - Simple and Adjusted Multilevel Analyses

	Number of Visits			
	General.	Gynaec.	Cardio.	Psychiat.
Distance to NHS				
Simple analysis	0.207***	−0.005	−0.007	−0.215*
	(0.054)	(0.060)	(0.083)	(0.100)
Adjusted	0.134*	−0.005	0.002	0.014
	(0.053)	(0.067)	(0.089)	(0.198)
Spatial Availability				
Simple analysis	−0.003***	0.006	−0.007	0.030*
	(0.001)	(0.006)	(0.013)	(0.012)
Adjusted	−0.001	0.096	−0.068	−0.171
	(0.001)	(0.008)	(0.016)	(0.018)
Observations	3,777	1,054	3,777	3,777

Note: *p<0.05; **p<0.01; ***p<0.001; n=3777;
NHS = Nearest Healthcare Service

In the simple regressions, we found associations for proximity and spatial availability for both general practitioners and psychiatrists. But after adjustment, only the proximity to the residence of general practitioners had a small positive association on the number of visits. The further someone lived from the closest general practitioner, the higher the number of visits to general practitioners. The other associations found in the simple analyses are explained by one or more of the confounding variables.

3.3.2 Effect of Workplace Neighbourhood Spatial Accessibility

The third hypothesis states that spatial accessibility to healthcare services in the workplace neighbourhood are associated with the use of healthcare services. The first three rows in Table 3 present the associations of the four dependent variables and the respective measures of proximity. The first row presents the coefficients of unadjusted regression analyses; the second presents the coefficients adjusted for the residence proximity; and the third row presents the coefficients adjusted for residence proximity and the control variables. The control variables were the same as in the tests of Hypothesis 2. Complete regression outputs can be found in Appendices A.7 and A.8.

Table 3: Use of Healthcare Services and Workplace Spatial Access Barriers - Simple and Adjusted Multilevel Analyses

	Number of Visits			
	General.	Gynaec.	Cardio.	Psychiat.
Distance to NHS				
Simple analysis	0.004	-0.107^{*}	-0.052	-0.108
	(0.032)	(0.046)	(0.049)	(0.115)
Adjusted on Residence	0.001	-0.108^{*}	-0.052	-0.098
	(0.032)	(0.047)	(0.049)	(0.118)
Full Adjustment	-0.014	-0.081	0.015	-0.332^{*}
	(0.033)	(0.051)	(0.055)	(0.158)
Spatial Availability				
Simple analysis	-0.001	0.0004	0.027^{**}	0.007
	(0.0004)	0.003)	(0.009)	(0.006)
Adjusted on Residence	-0.001	-0.00003	0.027^{**}	0.005
	(0.0004)	(0.003)	(0.009)	(0.006)
Full Adjustment	-0.001^{*}	-0.004	0.021^{*}	0.010
	(0.001)	(0.004)	(0.010)	(0.012)
Observations	3,777	1,054	3,777	3,777

Note: *p<0.05; **p<0.01; ***p<0.001; n=3777;
NHS = Nearest Healthcare Service

A negative association was found on the use of general practitioners after full adjustment. The simple regression and the regression adjusting for the residence spatial availability, did not not show an association. A negative association means that people working in a neighbourhood with less general practitioners, visit more often general practitioners. For the use of gynaecologists, an negative association for the distance to the nearest gynaecologist in the workplace environment was found in the simple analysis. This was not confirmed in the adjusted regression. There was no association found for the spatial availability of gynaecologists in the workplace environment.

A positive association was found between the spatial availability in the workplace environment and the use of cardiologists (β = 0.021, s.e. = 0.010, p = 0.033). After adjustment, the size of the coefficient was smaller than the unadjusted estimation, but remained statistically significant. As described above, outlier analyses were done for all fully adjusted regression analyses. We found one participant meeting the outlier criterion (standardized residual = 4.45). Looking at the observed values for this participant, we saw that he had no less than 72 visits to a cardiologist in the 18 months under consideration. After removing this outlier and refitting the model, we found a non-significant coefficient (β = 0.009, s.e. = 0.010, p = 0.383). There was small negative association found between the distance to the nearest psychiatrist and the use of psychiatrists ($\beta = -0.332$, s.e. = 0.158, p = 0.036). This association became even weaker after deleting an outlier (102 visits in 18 months) ($\beta = -0.294$, s.e. = 0.160, p = 0.067). The association was not found in the simpler analyses. There is not really an explanation for the association to appear after adjustment. Because of this instability in the coefficients and the instability due to a single outlier, we can not exclude the null hypothesis of no association between the proximity and the use of psychiatrist.

4 Discussion

Main findings

In this study, we have found a clustering of the use of healthcare services around the residence for most people. For only a small proportion of people, we also found a clustering around the workplace. No associations were documented between the spatial accessibility to healthcare services and the use of healthcare services, neither in the residence neighbourhood nor in the workplace neighbourhood.

Strengths and limitations

Thanks to the RECORD Study and its partners, this study could use advanced measurement and analysis methods on a large and extensive data set. For each participant, there was information on the amount of visits and the exact location of every visit. We also had full data on the healthcare services located in the residence and workplace neighbourhood.

A source of measurement bias of spatial accessibility, is the definition of neighbourhood. The concept of activity space might give better measures than the road network buffers centred on a single location such as the residence or the workplace [41]. The activity space is an area defined by the locations frequently visited by a person [42]. Conceptually, activity space is therefore a closer approximation of the person's real patterns of movement over space[41]. On the other hand, given a fairly evenly spread of services, the spatial accessibility in the activity space would be highly correlated to the road network buffers around the residence. This is also the case if people regularly visit places with a socioeconomic status comparable to their own neighbourhood. Moreover, measuring the activity space is based on data reported by the participants, most likely leading to missing data. So it remains to be tested if the theoretically more sophisticated concept of activity space would really enhance the measurements. A second bias in the measurement of the spatial accessibility is caused by participants moving or changing workplace in the 18 months following the recruitment to the study. Residence and workplace addresses were administered on the date of recruitment. On the other hand, the use of healthcare services is measured in the 18 months following the recruitment. Although this makes the causal argument stronger, this biases the measurement of spatial accessibility for people changing work or residence address during these 18 months. A related bias is the uncertainty of the workplace at the date of recruitment (see Method). It was known for which employer(s) a person worked during the year of recruitment. But in case of several employers, it was unclear which one was the employer at the exact date of recruitment. There were similar problems for people who had been unemployed or retired during a part of the year.

This study made use of advanced methodologies. Combining the information on the use of healthcare services to the home and work addresses, very precise measures of the spatial accessibility could be calculated with GIS techniques. Using the road network information, it was possible to define street network buffers instead of simpler circular buffers. Another strength in the methodology of this study is the use of negative binomial mixed model. In most previous research, more basic models were applied such as logistic or count models based on a Poisson distribution. Logistic regression models not only answer a different research question, they also ignore much of the variability in a count variable, losing precision in the estimates. Disregarding overdispersion by using a regular Poisson regression, overestimates precision. Disregarding dependencies between observations can both under- and overestimate precision. Overall, these models will generally provide less trustworthy results than the more advanced analysis techniques applied here.

No technique incorporated in an existing software allowed us to address our research question on clustering of healthcare use around the workplace. Therefore, we developed an indicator based on the mid-P-values of a Clopper-Pearson exact test. This indicator was found to be a useful and relatively simple instrument to describe spatial clustering of events. Unfortunately, the scale of the indicator is not meaningful, making it hard to interpret a single value. However, the indicator does lend itself to comparing different groups. Further investigation on this indicator is necessary, especially on the problems with very small amount of events. Even though the mid-P-value method corrects for the problem of discreteness in small samples, the correction might still not be sufficient for the very small numbers of events. Future research will have to show whether this causes bias or just noise in this indicator of clustering.

Social and scientific relevance

To our knowledge, this is the first study to investigate the clustering of the use of healthcare services around the residence and workplace in Ile-de-France. Previous research had focused mainly on rural or deserted areas and on the residential neighbourhood. Here, we assess whether spatial accessibility can add to the understanding of healthcare seeking behaviour in well-served areas. Furthermore, we studied the importance of the workplace environment; a non-residential environment where other health related behaviour has been observed in previous studies.

We have found a high clustering of healthcare use in the residence neighbourhood. Contradicting the findings in more remote areas [3], we have found only a limited amount of clustering in the workplace neighbourhood. The association found in remote areas, might be caused by a coincidence of workplace location and the location of healthcare services; and not as much by healthcare seeking behaviour in the workplace neighbourhood. In better served areas as Ile-de-France, nearly everyone can find healthcare services around the residence which makes it unnecessary to seek for services elsewhere. A relatively small group of people does use healthcare services in the workplace neighbourhood. The logistic regression showed that this use around the workplace is linked with commuting to Paris, the distance of commuting, a high level of occupation and a high family income. This indicates that a subgroup of people with a high time investment in their work seek for healthcare services around the workplace, especially when commuting from a less well-served area to a better served area.

Contradicting Chandola [14] but confirming others [15, 22], no association was found between the use of healthcare services and the spatial accessibility in the residence neighbourhood. Chaix et al. [16] had found an association between spatial accessibility and healthcare use in the French context. However in their study, there was a focus on elder people. The association was primarily found for disabled elderly, indicating that spatial accessibility in a well-served area is an issue for people with a low mobility. In this study, we focused on pure spatial accessibility, disregarding individual ability to overcome spatial access barriers. Also Ensor and Cooper [25] argued that the demand side (or individual) access barriers are likely to be as important as supply side (or structural) barriers. The findings in this study and those of Chaix et al. [16] indicate that in the context of well-served areas, the demand side access barrier 'mobility' could be important on its own as well as an important moderator for spatial accessibility.

Saag et al. [22] noticed that not only the personal mobility influences the way people handle spatial access barriers. In his study, people with arthritis had relatively low mobility but were not hindered by distance to visit general practitioners or other healthcare services. This suggest that the reason for the visit might also be a moderator for the association between spatial accessibility and the use of healthcare services, as well as a moderator for the association between mobility and healthcare use. Unfortunately, the reason for a visit is difficult to collect, and is unknown in large administrative databases as ours. It is confidential information; and there might be several reasons for one visit or different reasons for different visits over time. Health outcomes could be considered as a proxy of the reason to visit. However, in this study we considered health outcomes as potential consequences and not as causes of the use of healthcare services. Including consequences of the outcome in a model could cause biased and less precise estimators.

Conclusion

Spatial accessibility does not seem to have an influence on the use of healthcare services in a well-served area such as the Ile-de-France region, neither from the residence nor the workplace neighbourhood. This has been found for four different types of healthcare services: general practitioners, gynaecologists, cardiologists and psychiatrists. To better understand the healthcare seeking behaviour around the residence and around the workplace, future research might benefit of measuring in detail the time spend at the workplace and the time spend commuting. Future research interested in the association between spatial accessibility and the use of healthcare services in well-served areas, could focus on how an individual overcomes spatial access barriers. Specifically, we propose to include the ability (e.g. individual mobility) and individual motivation (e.g. reason to visit healthcare service) to overcome spatial access barriers. For health policy makers, these findings imply that promoting healthcare accessibility should not only be focused on the further spatial component of the accessibility to healthcare services, but also on lowering the access barriers for vulnerable subpopulations to the already existing services.

List of Appendices

Table A.1: Descriptive Statistics for Utilization Variables

	Descriptive statistics						
	Min	q_1	$\widetilde{x}$	q_3	Max	$\bar{x}$	s
Number of Visits							
General Practitioners	0.0	1.0	3.0	7.0	73.0	4.8	5.3
Gynaecologists	0.0	0.0	1.0	2.0	28.0	1.7	2.4
Cardiologists	0.0	0.0	0.0	0.0	72.0	0.3	1.5
Psychiatrists	0.0	0.0	0.0	0.0	102.0	0.7	5.3

Note: Min = Minimum; q_1 = 1^e quartile; $\widetilde{x}$ = Median; q_3 = 3^e quartile; Max = Maximum; $\bar{x}$ = Mean; s = Standard Deviation; n-total=3777

Table A.2: Descriptive Statistics for Continuous Control Variables

	Descriptive statistics						
	Min	q_1	$\widetilde{x}$	q_3	Max	$\bar{x}$	s
Workplace Access Barriers							
Distance to NHS (km)							
General Practitioners	0.0	0.1	0.2	0.5	6.2	0.4	0.5
Gynaecologists	0.0	0.2	0.4	0.7	7.5	0.7	0.9
Cardiologists	0.0	0.3	0.6	1.1	21.1	0.9	1.1
Psychiatrists	0.0	0.2	0.5	1.1	19.7	0.8	1.2
Spatial Availability							
General Practitioners	0.0	9.0	27.0	66.0	263.0	39.9	38.5
Gynaecologists	0.0	1.2	7.0	16.0	59.0	11.0	12.1
Cardiologists	0.0	0.0	3.0	7.0	31.0	4.9	6.3
Psychiatrists	0.0	0.0	4.0	32.0	121.0	17.3	24.0
Residence Access Barriers							
Distance to NHS (km)							
General Practitioners	0.0	0.1	0.2	0.4	8.1	0.3	0.3
Gynaecologists	0.0	0.3	0.6	0.9	8.8	0.7	0.7
Cardiologists	0.0	0.4	0.6	1.1	9.5	0.8	0.7
Psychiatrists	0.0	0.3	0.5	1.0	10.1	0.7	0.7
Spatial Availability							
General Practitioners	0.0	9.0	20.0	51.0	259.0	32.3	31.2
Gynaecologists	0.0	1.0	3.0	8.0	56.0	5.5	7.0
Cardiologists	0.0	0.0	2.0	4.0	29.0	3.3	4.4
Psychiatrists	0.0	0.0	2.0	10.0	95.0	8.8	14.8
Commuting Distance (km)	2.0	5.5	9.1	15.0	78.4	11.7	9.1
Age (years)	30.0	39.0	46.0	54.0	77.0	46.3	9.4
Family Income (/1000€)	0.0	0.8	1.4	2.2	8.5	1.7	1.1
HDI 2004	0.3	0.9	0.9	0.9	1.0	0.9	0.1
Attachment to Neighbourhood	3.0	9.0	10.0	12.0	13.0	10.2	2.2

Note: Min = Minimum; q_1 = 1^e quartile; $\widetilde{x}$ = Median; q_3 = 3^e quartile; Max = Maximum; $\bar{x}$ = Mean; s = Standard Deviation; n-total=3777
NHS = Nearest Healthcare Service

Table A.3: Descriptive Statistics for Categorical Control Variables

	Descriptive statistics		
	n	%	$\sum$%
Gender			
Female	1054	27.9	27.9
Male	2723	72.1	100.0
Educational level			
No - primary - low secondary	1075	28.5	28.5
High secondary - low tertiary	1109	29.4	57.8
High tertiary	1593	42.2	100.0
Occupation			
High white-collar	1735	45.9	45.9
Intermediate	267	7.1	53.0
Low white-collar	1199	31.7	84.8
Blue-collar	511	13.5	98.3
Divers	65	1.7	100.0
Financial strain			
No	3146	83.3	83.3
Yes	631	16.7	100.0
Work location			
Outside of Paris	2204	58.4	58.4
Paris	1573	41.6	100.0
Home Location			
Outside of Paris	2687	71.1	71.1
Paris	1090	28.9	100.0

Note: n = number of participants; % = percentage; $\sum$% = Cumulative percentage; n-total = 3777

Table A.4: Clustering of Use of Healthcare Services and Background Variables

	$P(MPV < 0.05)$	
	Residence	Workplace
Work Location [a]		
Paris	0.165	0.585^{***}
	(0.103)	(0.134)
Work Location [a]		
Paris	0.088	0.291
	(0.154)	(0.184)
Work by Home Location (int)	−0.092	-0.631^{**}
	(0.193)	(0.237)
Commuting Distance (km)	-0.017^{***}	-0.062^{***}
	(0.005)	(0.010)
Attachment Neighbourhood	0.027	−0.034
	(0.020)	(0.026)
Age (years)	0.022^{***}	−0.005
	(0.005)	(0.006)
Gender [b]		
Male	-0.520^{***}	-0.705^{***}
	(0.102)	(0.118)
Education [c]		
High sec - low ter	−0.033	0.132
	(0.118)	(0.172)
High tertiary	−0.014	0.086
	(0.132)	(0.185)
Occupation [d]		
Intermediate	0.038	−0.265
	(0.179)	(0.225)
Low white-collar	−0.071	-0.696^{***}
	(0.117)	(0.154)
Blue collar	0.053	-0.789^{**}
	(0.163)	(0.264)
Divers	0.112	-2.365^{*}
	(0.332)	(1.037)
Family Income (/1000€)	-0.094^{*}	0.262^{***}
	(0.042)	(0.047)
Financial Strain	−0.090	−0.163
	(0.119)	(0.187)
HDI 2004	0.769	-1.151^{*}
	(0.427)	(0.527)
Education Neighbourhood	0.530	1.330
	(0.388)	(0.883)
(intercept)	−0.073	−1.614
	(0.493)	(0.923)
Observations	3,775	3,775

Note: $P(MPV < 0.05)$ = the risk of a mid-P-value smaller than 0.05
HDI = Human Development Index;
*p<0.05; **p<0.01; ***p<0.001
[a] : ref = 'Outside Paris'; [b] : ref = 'Female';
[c] : ref = 'No, primary or low secondary';
[d] : ref = 'High white-collar'; [e] : ref = 'No'; n=3775

Table A.5: Multiple Negative Binomial Regressions - Use of Healthcare Services and Residential Proximity

	Number of Visits			
	General.	Gynaec.	Cardio.	Psychiat.
Distance to NHS (km)	0.134*	0.102	−0.066	−0.134
	(0.053)	(0.067)	(0.089)	(0.198)
Commuting Distance (km)	−0.0002	0.0003	−0.001	−0.011
	(0.002)	(0.005)	(0.006)	(0.021)
Work Location [a]				
Paris	−0.050	0.117	0.213	−0.260
	(0.034)	(0.077)	(0.111)	(0.339)
Home Location [a]				
Paris	−0.063	0.029	−0.406**	0.439
	(0.044)	(0.099)	(0.144)	(0.385)
Attachment Neighbourhood	−0.010	−0.042*	−0.021	−0.228**
	(0.008)	(0.017)	(0.026)	(0.078)
Age (years)	0.011***	−0.025***	0.066***	−0.031
	(0.002)	(0.004)	(0.006)	(0.018)
Gender [b]				
Male	−0.491***		0.140	−0.557
	(0.037)		(0.125)	(0.344)
Education [c]				
High sec - low ter	−0.053	0.039	−0.237	0.891
	(0.046)	(0.099)	(0.148)	(0.491)
High tertiary	−0.066	0.014	−0.297	1.247*
	(0.051)	(0.116)	(0.167)	(0.515)
Occupation [d]				
Intermediate	0.143*	−0.213	0.118	−0.919
	(0.069)	(0.158)	(0.230)	(0.631)
Low white-collar	0.091*	−0.208*	0.237	0.822
	(0.047)	(0.097)	(0.153)	(0.439)
Blue collar	0.132*	−0.220	−0.064	0.458
	(0.065)	(0.193)	(0.214)	(0.688)
Divers	0.188	−0.140	0.247	2.212
	(0.128)	(0.227)	(0.434)	(1.142)
Family Income (/1000€)	−0.003	0.012	0.021	−0.033
	(0.017)	(0.041)	(0.053)	(0.130)
Financial Strain [e]				
Yes	0.108*	−0.147	0.084	−0.229
	(0.047)	(0.100)	(0.154)	(0.425)
HDI 2004	0.312	0.107	0.764	4.639*
	(0.166)	(0.362)	(0.609)	(2.180)
Neighbourhood Education	−0.154	0.850*	0.133	2.087
	(0.159)	(0.367)	(0.506)	(1.704)
(intercept)	1.213***	1.564***	−4.973***	−2.448
	(0.199)	(0.437)	(0.717)	(1.950)
Observations	3,777	1,054	3,777	3,777

Note: *p<0.05; **p<0.01; ***p<0.001
General. = general practitioners; Gynaec. = gynaecologists; Cardio. = cardiologists;
Psychiat. = psychiatrists; NHS = Nearest Healthcare Service;
HDI = Human Development Index;
[a] : ref = 'Outside Paris'; [b] : ref = 'Female'; [c] : ref = 'No, primary or low secondary';
[d] : ref = 'High white-collar'; [e] : ref = 'No'; n=3777

Table A.6: Multiple Negative Binomial Regressions - Use of Healthcare Services and Residential Spatial Availability

	Number of Visits			
	General.	Gynaec.	Cardio.	Psychiat.
Spatial Availability	−0.001	−0.005	0.002	0.014
	(0.001)	(0.008)	(0.016)	(0.018)
Commuting Distance (km)	0.001	0.002	−0.001	−0.015
	(0.002)	(0.005)	(0.006)	(0.021)
Work Location [a]				
Paris	−0.054	0.109	0.216	−0.205
	(0.034)	(0.077)	(0.111)	(0.340)
Home Location [a]				
Paris	−0.011	0.040	−0.391**	0.139
	(0.062)	(0.115)	(0.151)	(0.549)
Attachment Neighbourhood	−0.010	−0.042*	−0.022	−0.228**
	(0.008)	(0.017)	(0.026)	(0.078)
Age (years)	0.011***	−0.025***	0.066***	−0.035
	(0.002)	(0.004)	(0.006)	(0.018)
Gender [b]				
Male	−0.494***		0.142	−0.506
	(0.037)		(0.125)	(0.343)
Education [c]				
High sec - low ter	−0.042	0.047	−0.241	0.858
	(0.046)	(0.099)	(0.148)	(0.487)
High tertiary	−0.057	0.018	−0.295	1.227*
	(0.051)	(0.116)	(0.167)	(0.512)
Occupation [d]				
Intermediate	0.145*	−0.205	0.120	−0.972
	(0.069)	(0.158)	(0.231)	(0.642)
Low white-collar	0.097*	−0.201*	0.239	0.876*
	(0.046)	(0.097)	(0.153)	(0.440)
Blue collar	0.145*	−0.202	−0.061	0.546
	(0.064)	(0.193)	(0.214)	(0.687)
Divers	0.198	−0.143	0.243	2.276*
	(0.128)	(0.227)	(0.435)	(1.145)
Family Income (/1000€)	−0.002	0.014	0.020	−0.030
	(0.017)	(0.041)	(0.053)	(0.128)
Financial Strain [e]				
Yes	0.107*	−0.140	0.083	−0.141
	(0.047)	(0.100)	(0.154)	(0.433)
HDI 2004	0.315	0.114	0.765	4.973*
	(0.165)	(0.363)	(0.609)	(2.173)
Neighbourhood Education	−0.126	0.784*	0.184	2.005
	(0.163)	(0.377)	(0.521)	(1.708)
(intercept)	1.259***	1.653***	−5.042***	−2.721
	(0.197)	(0.434)	(0.710)	(1.928)
Observations	3,777	1,054	3,777	3,777

Note: *p<0.05; **p<0.01; ***p<0.001
General. = general practitioners; Gynaec. = gynaecologists; Cardio. = cardiologists;
Psychiat. = psychiatrists;
HDI = Human Development Index;
[a] : ref = 'Outside Paris'; [b] : ref = 'Female'; [c] : ref = 'No, primary or low secondary';
[d] : ref = 'High white-collar'; [e] : ref = 'No'; n=3777

Table A.7: Multiple Negative Binomial Regressions - Use of Healthcare Services and Workplace Proximity

	Number of Visits			
	General.	Gynaec.	Cardio.	Psychiat.
Work Distance to NHS (km)	−0.014	−0.081	0.015	−0.332*
	(0.033)	(0.051)	(0.055)	(0.158)
Residence Distance to NHS (km)	0.130*	0.097	−0.066	−0.169
	(0.052)	(0.067)	(0.089)	(0.209)
Commuting Distance (km)	0.0001	0.002	−0.001	−0.0004
	(0.002)	(0.005)	(0.007)	(0.021)
Work Location [a]				
Paris	−0.057	0.030	0.225	−0.529
	(0.036)	(0.099)	(0.120)	(0.367)
Home Location [a]				
Paris	−0.059	−0.035	−0.408**	0.348
	(0.047)	(0.142)	(0.144)	(0.388)
Attachment Neighbourhood	−0.011	−0.040*	−0.021	−0.247**
	(0.008)	(0.017)	(0.026)	(0.078)
Age (years)	0.011***	−0.024***	0.066***	−0.031
	(0.002)	(0.004)	(0.006)	(0.018)
Gender [b]				
Male	−0.492***		0.140	−0.597
	(0.037)		(0.125)	(0.344)
Education [c]				
High sec - low ter	−0.048	0.036	−0.234	1.007*
	(0.046)	(0.099)	(0.148)	(0.492)
High tertiary	−0.064	0.013	−0.295	1.419**
	(0.051)	(0.116)	(0.167)	(0.523)
Occupation [d]				
Intermediate	0.147*	−0.212	0.119	−0.976
	(0.069)	(0.158)	(0.230)	(0.632)
Low white-collar	0.093*	−0.207*	0.237	0.929*
	(0.046)	(0.097)	(0.153)	(0.442)
Blue collar	0.143*	−0.208	−0.066	0.687
	(0.064)	(0.193)	(0.214)	(0.700)
Divers	0.192	−0.125	0.243	2.466*
	(0.128)	(0.227)	(0.434)	(1.156)
Family Income (/1000€)	−0.003	0.015	0.020	−0.022
	(0.017)	(0.041)	(0.053)	(0.130)
Financial Strain [e]				
Yes	0.109*	−0.144	0.084	−0.058
	(0.047)	(0.100)	(0.154)	(0.434)
HDI 2004	0.318	0.071	0.767	4.837*
	(0.165)	(0.363)	(0.609)	(2.136)
Neighbourhood Education	−0.156	0.831*	0.132	2.497
	(0.158)	(0.367)	(0.506)	(1.708)
(intercept)	1.224***	1.630***	−4.996***	−2.538
	(0.198)	(0.439)	(0.722)	(1.944)
Observations	3,777	1,054	3,777	3,777

Note: *p<0.05; **p<0.01; ***p<0.001
General. = general practitioners; Gynaec. = gynaecologists; Cardio. = cardiologists;
Psychiat. = psychiatrists; NHS = Nearest Healthcare Service;
HDI = Human Development Index;
[a] : ref = 'Outside Paris'; [b] : ref = 'Female'; [c] : ref = 'No, primary or low secondary';
[d] : ref = 'High white-collar'; [e] : ref = 'No'; n=3777

Table A.8: Multiple Negative Binomial Regressions -
Use of Healthcare Services and Workplace Spatial Availability

	Number of Visits			
	General.	Gynaec.	Cardio.	Psychiat.
Work Spatial Availability	-0.001^{*}	-0.004	0.021^{*}	0.010
	(0.001)	(0.004)	(0.010)	(0.012)
Residence Spatial Availability	-0.001	-0.005	0.004	0.014
	(0.001)	(0.008)	(0.015)	(0.018)
Commuting Distance (km)	0.0002	0.002	-0.0002	-0.013
	(0.002)	(0.005)	(0.006)	(0.021)
Work Location ^{a}				
Paris	0.031	0.168	0.059	-0.542
	(0.050)	(0.103)	(0.134)	(0.524)
Home Location ^{a}				
Paris	-0.022	0.036	-0.361^{*}	0.184
	(0.061)	(0.115)	(0.147)	(0.551)
Attachment Neighbourhood	-0.009	-0.041^{*}	-0.022	-0.222^{**}
	(0.008)	(0.017)	(0.025)	(0.078)
Age (years)	0.011^{***}	-0.025^{***}	0.066^{***}	-0.036^{*}
	(0.002)	(0.004)	(0.006)	(0.018)
Gender ^{b}				
Male	-0.496^{***}		0.167	-0.508
	(0.037)		(0.123)	(0.341)
Education ^{c}				
High sec - low ter	-0.048	0.046	-0.246	0.797
	(0.046)	(0.099)	(0.147)	(0.490)
High tertiary	-0.056	0.018	-0.310	1.210^{*}
	(0.051)	(0.116)	(0.165)	(0.508)
Occupation ^{d}				
Intermediate	0.141^{*}	-0.199	0.115	-0.940
	(0.069)	(0.158)	(0.229)	(0.642)
Low white-collar	0.101^{*}	-0.192^{*}	0.250	0.832
	(0.047)	(0.098)	(0.152)	(0.437)
Blue collar	0.133^{*}	-0.201	-0.032	0.561
	(0.065)	(0.193)	(0.213)	(0.686)
Divers	0.194	-0.138	0.253	2.306^{*}
	(0.128)	(0.227)	(0.435)	(1.139)
Family Income (/1000€)	-0.001	0.015	0.023	-0.040
	(0.017)	(0.041)	(0.052)	(0.128)
Financial Strain ^{e}				
Yes	0.107^{*}	-0.140	0.053	-0.131
	(0.047)	(0.100)	(0.152)	(0.432)
HDI 2004	0.306	0.115	0.775	5.027^{*}
	(0.166)	(0.363)	(0.606)	(2.160)
Neighbourhood Education	-0.133	0.806^{*}	0.178	1.963
	(0.165)	(0.377)	(0.507)	(1.704)
(intercept)	1.279^{***}	1.654^{***}	-5.082^{***}	-2.780
	(0.198)	(0.433)	(0.707)	(1.926)
Observations	3,777	1,054	3,777	3,777

Note: *p<0.05; **p<0.01; ***p<0.001
General. = general practitioners; Gynaec. = gynaecologists; Cardio. = cardiologists;
Psychiat. = psychiatrists;
HDI = Human Development Index;
^{a} : ref = 'Outside Paris'; ^{b} : ref = 'Female'; ^{c} : ref = 'No, primary or low secondary';
^{d} : ref = 'High white-collar'; ^{e} : ref = 'No'; n=3777

Bibliography

[1] Penchansky R. and Thomas J. W. The concept of access: definition and relationship to consumer satisfaction. *Medical care*, 19(2):127–40, 1981. URL http://www.ncbi.nlm.nih.gov/pubmed/7206846.

[2] Millman M. *Access to Health Care in America.* National Academies Press, Washington, DC, 1993.

[3] Gesler W. M. and Meade M. S. Locational and population factors in health care-seeking behavior in Savannah, Georgia. *Health services research*, 23(3):443–62, 1988. URL http://www.ncbi.nlm.nih.gov/pubmed/3403277.

[4] Parchman M. L. and Culler S. D. Preventable hospitalizations in primary care shortage areas. An analysis of vulnerable Medicare beneficiaries. *Archives of family medicine*, 8(6):487–91, 1999. URL http://www.ncbi.nlm.nih.gov/pubmed/10575386.

[5] Chaix B., Veugelers P. J., Boelle P. Y., and Chauvin P. Access to general practitioner services: the disabled elderly lag behind in underserved areas. *European journal of public health*, 15(3):282–7, 2005. URL http://www.ncbi.nlm.nih.gov/pubmed/15941749.

[6] Briggs L. W., Rohrer J. E., Ludke R. L., Hilsenrath P. E., and Phillips K. T. Geographic variation in primary care visits in Iowa. *Health services research*, 30(5):657–71, 1995. URL http://www.ncbi.nlm.nih.gov/pubmed/8537225.

[7] Charreire H., Casey R., Salze P., Simon C., Chaix B., Banos A., Badariotti D., Weber C., and Oppert J. M. Measuring the food environment using geographical information systems: a methodological review. *Public health nutrition*, 13(11):1773–85, 2010. URL http://www.ncbi.nlm.nih.gov/pubmed/20409354.

[8] Gould P. R. *Spatial Diffusion, Resource Paper No. 4.* Association of American Geographers, 1710 Sixteenth Street, N.W., Washington, D.C. 20009, 1969. URL http://www.eric.ed.gov/ERICWebPortal/detail?accno=ED120029.

[9] Garrett C. R., Gask L. L., Hays R., Cherrington A., Bundy C., Dickens C., Waheed W., and Coventry P. A. Accessing primary health care: a meta-ethnography of the experiences of British South Asian patients with diabetes, coronary heart disease or a mental health problem. *Chronic Illness*, 8(2):135–155, 2012. URL http://chi.sagepub.com/content/8/2/135.abstract.

[10] Fortney D. J., Rost D. K., and Warren J. Comparing Alternative Methods of Measuring Geographic Access to Health Services. *Health Services and Outcomes Research Methodology*, 1(2):173–184, 2000. URL http://www.ingentaconnect.com/content/klu/hsor/2000/00000001/00000002/00265228.

[11] Diez Roux A. V. and Mair C. Neighborhoods and health. *Annals of the New York Academy of Sciences*, 1186:125–45, 2010. URL http://www.ncbi.nlm.nih.gov/pubmed/20201871.

[12] Inagami S., Cohen D. A., and Finch B. K. Non-residential neighborhood exposures suppress neighborhood effects on self-rated health. *Social science & medicine*, 65 (8):1779–91, 2007.

[13] Carr-Hill R. A., Rice N., and Roland M. Socioeconomic determinants of rates of consultation in general practice based on fourth national morbidity survey of general practices. *BMJ*, 312(7037):1008–12, 1996. URL http://www.ncbi.nlm.nih.gov/pubmed/8616346.

[14] Chandola T. Spatial and social determinants of urban health in low-, middle- and high-income countries. *Public health*, 126(3):259–261, 2012. URL http://www.sciencedirect.com/science/article/pii/S0033350611003970.

[15] Earle C. C., Neumann P. J., Gelber R. D., Weinstein M. C., and Weeks J. C. Impact of referral patterns on the use of chemotherapy for lung cancer. *Journal of clinical oncology*, 20(7):1786–92, 2002. URL http://www.ncbi.nlm.nih.gov/pubmed/11919235.

[16] Chaix B., Boelle P. Y., Guilbert P., and Chauvin P. Area-level determinants of specialty care utilization in France: a multilevel analysis. *Public health*, 119(2):97–104, 2005. URL http://www.ncbi.nlm.nih.gov/pubmed/15694956.

[17] Casey R., Chaix B., Weber C., Schweitzer B., Charreire H., Salze P., Badariotti D., Banos A., Oppert J. M., and Simon C. Spatial accessibility to physical activity facilities and to food outlets and overweight in French youth. *International journal of obesity*, 36(7):914–9, 2012. URL http://www.ncbi.nlm.nih.gov/pubmed/22310474.

[18] Karusisi N., Bean K., Oppert J. M., Pannier B., and Chaix B. Multiple dimensions of residential environments, neighborhood experiences, and jogging behavior in the RECORD Study. *Preventive medicine*, 55(1):50–5, 2012. URL http://www.ncbi.nlm.nih.gov/pubmed/22564774.

[19] Stafford M., Cummins S., Macintyre S., Ellaway A., and Marmot M. Gender differences in the associations between health and neighbourhood environment. *Social science & medicine (1982)*, 60(8):1681–1692, 2005. URL http://europepmc.org/abstract/MED/15686801.

[20] Christian W. J. Using geospatial technologies to explore activity-based retail food environments. *Spatial and spatio-temporal epidemiology*, 3(4):287–95, 2012. URL http://www.ncbi.nlm.nih.gov/pubmed/23149325.

[21] Rosenthal T. C. and Fox C. Access to health care for the rural elderly. *JAMA : the journal of the American Medical Association*, 284(16):2034–6, 2000. URL http://www.ncbi.nlm.nih.gov/pubmed/11042733.

[22] Saag K. G., Doebbeling B. N., Rohrer J. E., Kolluri S., Mitchell T. A., and Wallace R. B. Arthritis health service utilization among the elderly: the role of urban-rural residence and other utilization factors. *Arthritis care and research : the official journal of the Arthritis Health Professions Association*, 11(3):177–85, 1998. URL http://www.ncbi.nlm.nih.gov/pubmed/9782809.

[23] Gething P. W., Johnson F. A., Frempong-Ainguah F., Nyarko P., Baschieri A., Aboagye P., Falkingham J., Matthews Z., and Atkinson P. M. Geographical access to care at birth in Ghana: a barrier to safe motherhood. *BMC public health*, 12:991, 2012. URL http://www.ncbi.nlm.nih.gov/pubmed/23158554.

[24] Huerta Munoz U. and Kallestal C. Geographical accessibility and spatial coverage modeling of the primary health care network in the Western Province of Rwanda. *International journal of health geographics*, 11:40, 2012. URL http://www.ncbi.nlm.nih.gov/pubmed/22984920.

[25] Ensor T. and Cooper S. Overcoming barriers to health service access: influencing the demand side. *Health Policy Plan*, 19(2):69–79, 2004.

[26] Cheng Y., Wang J., and Rosenberg M. W. Spatial access to residential care resources in Beijing, China. *International journal of health geographics*, 11:32, 2012. URL http://www.ncbi.nlm.nih.gov/pubmed/22877360.

[27] Chaix B., Kestens Y., Bean K., Leal C., Karusisi N., Meghiref K., Burban J., Fon Sing M., Perchoux C., Thomas F., Merlo J., and Pannier B. Cohort profile: residential and non-residential environments, individual activity spaces and cardiovascular risk factors and diseases–the RECORD Cohort Study. *International journal of epidemiology*, 41(5):1283–92, 2012. URL http://www.ncbi.nlm.nih.gov/pubmed/21737405.

[28] Chaix B., Bean K., Leal C., Thomas F., Havard S., Evans D., Jego B., and Pannier B. Individual/neighborhood social factors and blood pressure in the RECORD Cohort Study: which risk factors explain the associations? *Hypertension*, 55(3): 769–75, 2010. URL http://www.ncbi.nlm.nih.gov/pubmed/20100998.

[29] Leal C., Bean K., Thomas F., and Chaix B. Are associations between neighborhood socioeconomic characteristics and body mass index or waist circumference based on model extrapolations? *Epidemiology*, 22(5):694–703, 2011. URL http://www.ncbi.nlm.nih.gov/pubmed/21709560.

[30] Frank L., Kerr J., Chapman J., and Sallis J. Urban Form Relationships With Walk Trip Frequency and Distance Among Youth. *American Journal of Health Promotion*, 21(4s):305–311, 2007. URL http://dx.doi.org/10.4278/0890-1171-21.4s.305.

[31] United Nation Development Program. UNDP Annual Report 2004: Mobilizing Global Partnerships, May 2004. URL http://www.undp.org/.

[32] Puett R. C., Lawson A. B., Clark A. B., Hebert J. R., and Kulldorff M. Power evaluation of focused cluster tests. *Environmental and Ecological Statistics*, 17(3): 303–316, 2010. URL http://dx.doi.org/10.1007/s10651-009-0108-1.

[33] Agresti A. *Categorical Data Analysis*. Wiley, 2002.

[34] Molenberghs G., Verbeke G., and Demetrio C. G. An extended random-effects approach to modeling repeated, overdispersed count data. *Lifetime data analysis*, 13(4):513–31, 2007. URL http://www.ncbi.nlm.nih.gov/pubmed/17999182.

[35] Fitzmaurice G. M., Laird N. M., and Ware J. H. *Applied longitudinal analysis*. Wiley-Interscience, Hoboken, N.J. ; [Great Britain], 2004.

[36] Ye F., Yue C., and Yang Y. Modeling time-dependent overdispersion in longitudinal count data. *Computational Statistics & Data Analysis*, 58:257–264, 2013. URL http://ac.els-cdn.com/S016794731200312X/1-s2.0-S016794731200312X-main.pdf?_tid=63a1f666-a8e8-11e2-9991-00000aacb362&acdnat=1366372863_aca7d16a5eb803965717ef1b30538a42.

[37] Atkins D. C. and Gallop R. J. Rethinking how family researchers model infrequent outcomes: a tutorial on count regression and zero-inflated models. *Journal of family psychology*, 21(4):726–35, 2007. URL `http://www.ncbi.nlm.nih.gov/pubmed/18179344`.

[38] De Smet O., Buysse A., and Brondeel R. Effect of the breakup context on unwanted pursuit behavior perpetration between former partners. *Journal of forensic sciences*, 56(4):934–41, 2011. URL `http://www.ncbi.nlm.nih.gov/pubmed/21470223`.

[39] Skaug H., Fournier D., Nielsen A., Magnusson A., and (2012) Bolker B. Generalized Linear Mixed Models using AD Model Builder. R package version 0.7.3.

[40] Fournier D. A., Skaug H. J., Ancheta J., Ianelli J., Magnusson A., Maunder M. N., Nielsen A., and Sibert J. AD Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. *Optimization Methods and Software*, 27(2):233–249, 2012. URL `http://dx.doi.org/10.1080/10556788.2011.597854`.

[41] Sherman J. E., Spencer J., Preisser J. S., Gesler W. M., and Arcury T. A. A suite of methods for representing activity space in a healthcare accessibility study. *International journal of health geographics*, 4:24, 2005. URL `http://www.ncbi.nlm.nih.gov/pubmed/16236174`.

[42] Golledge R. G. and Stimson R. R. J. *Spatial behavior: A geographic perspective*. The Guilford Press, 1997.