Measuring and validating spatial accessibility to COVID-19 vaccination sites: a case study in England

Huanfa Chen¹, Yanjia Cao², Lingru Feng¹, Qunshan Zhao³

¹Centre for Advanced Spatial Analysis, University College London

²School of Public Health and Human Longevity Science, University of California, San Diego

³Urban Big Data Centre, School of Social & Political Sciences, University of Glasgow

Huanfa Chen (corresponding author) (huanfa.chen@ucl.ac.uk)(<https://orcid.org/0000-0002-4518-7601>)

Yanjia Cao (yanjia.cao@gmail.com)(<https://orcid.org/0000-0002-8294-8924>)

Lingru Feng (lingru.feng.18@alumni.ucl.ac.uk)(<https://orcid.org/0000-0001-8343-2383>)

Qunshan Zhao (qunshan.zhao@glasgow.ac.uk)(<https://orcid.org/0000-0002-5549-9457>)

Funding details

Dr. Qunshan Zhao has received UK ESRC's on-going support for the Urban Big Data Centre (UBDC) [ES/L011921/1 and ES/S007105/1]. We want to thank all the insightful comments and suggestions on earlier versions of this manuscript from anonymous reviewers.

Disclosure statement

The authors report there are no competing interests to declare.

Measuring and validating spatial accessibility to Covid-19 vaccination sites: a case study in England¹

Abstract

The global Covid-19 pandemic has caused numerous deaths and illnesses and posed unprecedented social and economic challenges to many countries. One of the key strategies to contain the pandemic is mass vaccination. While it is essential to ensure safe and easy accessibility to Covid-19 vaccines for all communities, limited research has been carried out to understand and validate the spatial accessibility of these vaccines. This study addresses this gap by measuring and validating the spatial accessibility to Covid-19 vaccines with a particular focus on England, United Kingdom. More specifically, we compare three floating catchment area (FCA) methods with differing parameters for measuring the small-scale spatial accessibility to vaccination services. Then, we calibrate these accessibility measurements using a beta regression model and the reported vaccination uptake rates. The results show that the three-step FCA method with a distance parameter of 30 miles is the optimal model for measuring the spatial accessibility to Covid-19 vaccines. The findings provide an improved understanding of the spatial inequality of vaccine services. Further, the framework of calibrating spatial accessibility to vaccine services is generalisable to other types of healthcare services.

Keywords: spatial accessibility, Covid-19 vaccination services, England, beta regression, validation

1. Introduction

The Coronavirus disease 2019 (COVID-19), caused by the severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2), has been a global pandemic that poses unprecedented health, social, and economic challenges (WHO 2020). As of November 2021, the world has confirmed approximately 0.2 billion COVID-19 cases; and in UK alone, over 9 million people have been contracted, with 140,000 deaths. In order to minimise SARS-COV-2 infection cases and

¹ This manuscript has been posted as a preprint: "Measuring and Validating Spatial Accessibility to Covid-19 Vaccination Sites: A Case Study in England." SocArXiv. osf.io/preprints/socarxiv/xvnps.

transmissions, public health officials have adopted social distancing as the primary control strategy, until mass vaccination becomes available (Chu *et al.* 2020, Hsiang *et al.* 2020, Lewnard and Lo 2020).

In December 2020, UK regulators issued emergency-use authorisation for COVID-19 vaccines from Pfizer and BioNTech, and later AstraZeneca (Ledford *et al.* 2020). In the COVID-19 vaccine delivery plan (UK Department of Health and Social Care 2021), it is stated that ‘We have always known that vaccines would be our best way out of this pandemic and towards a more normal way of life.’ To maximise the success of the vaccination plan, it is essential to “ensure safe easy access for the whole population’ (UK Department of Health and Social Care 2021) to the vaccine, which includes spatial accessibility (i.e. the ease of travelling to vaccine sites).

In general, fair and equitable accessibility to health-related resources is critical to maximising the benefits of healthcare. Population groups with poor accessibility to healthcare are less likely to receive routine or acute care, which results in irreversible health consequences. Accessibility to healthcare is impacted by a wide range of spatial and aspatial factors and processes, including service supply, the capability of paying for services, accessibility of information (e.g. language barrier, using digital systems), travel time, access of transport. The various definitions of accessibility to healthcare can be grouped according to two dichotomous criteria and into four types: revealed or potential, spatial or aspatial. Specifically, revealed accessibility concerns actual use of services, whereas potential accessibility focuses on the probable use of healthcare services but does not ensure the automatic utilisation of these services. Spatial accessibility emphasises how spatial/distance variables impact access, whilst aspatial accessibility focuses on the non-spatial barriers.

Accessibility to healthcare varies across the region, as the determinants of access include both local health resources (i.e., supply) and resident groups (i.e., demand), and neither of them is evenly distributed in space. Many methods have been developed to measure spatial accessibility to healthcare, including the regional availability models (Khan 1992), kernel density models (Guagliardo 2004), and floating catchment area (FCA) models (Luo and Wang 2003, Luo and Qi 2009, Wan *et al.* 2012). The regional availability models, which compares the total healthcare capacity and population demand in an area, have some limitations in its assumptions: first, the population within a given area have equal accessibility to healthcare within the area; second, people are restricted in an area and do not seek healthcare outside the area. On the other hand, the kernel density model creates a continuous surface showing the distribution of healthcare services and then calculates the accessibility at each demand point. However, this model is not useful in identifying low-access areas as it presents a disruptive pattern of access scores. The FCA models assume that a population group’s accessibility to medical sites decreases with its distance to nearby medical sites, conforming to the gravity law. Compared with the two aforementioned models, the FCA models are more flexible to implement and the result interpretation is more straightforward. Variants of FCA include Two-step FCA (2SFCA), Enhanced Two-step FCA (E2SFCA), and Three-step FCA (3SFCA). These variants differ in considerations of distance effect and competition effects between medical sites. A major drawback of the existing studies on spatial accessibility to healthcare is a lack of model-level calibration (Carley 2017), which is concerned with linking the estimated accessibility with real-world data and revealed accessibility. The calibration of healthcare accessibility serves two purposes. First, it identifies and parameterises the best-fitted model out of numerous candidates. Second, calibration is essential in establishing the accuracy of the theory of the models and moving the accessibility models from the theoretical to the practical realm (e.g. planning of healthcare services). A few studies have calibrated the distance decay function in the FCA models using taxi trip data (Wang *et al.* 2020, Gong *et al.* 2021) or willingness-to-travel survey data (Rader *et al.* 2021).

However, no existing study has proposed a general model-level calibration approach of healthcare accessibility models. The primary purpose of this study is to establish such a calibration framework for spatial accessibility to healthcare services, which adopts the COVID-19 vaccine as the case study.

In the context of COVID-19, limited research has been conducted to measure, calibrate, and understand the spatial accessibility to vaccines or other healthcare services, with two exceptions. Kang *et al.* (2020) used an E2SFCA method to rapidly estimate the spatial accessibility to COVID-19 healthcare resources and to identify geographic areas with low spatial accessibility through a parallel computing strategy. While this study improves the computing efficiency of E2SFCA, it is limiting in the external validation of the spatial accessibility, as the estimated spatial accessibility is not linked to the revealed accessibility to healthcare resources. On the other hand, Rader *et al.* (2021) employed the E2SFCA method to identify “vaccine deserts” in US cities, areas with low spatial accessibility to COVID-19 vaccine supply. The methodological contribution is that they calibrated the E2SFCA’s travel-time decay function using people’s time-willing-to-travel for vaccination from a large-scale survey. However, the estimated spatial accessibility to vaccination lacks rigorous calibration. Our study builds upon existing research, and proposes to calibrate the spatial accessibility to COVID-19 vaccination via linking it to the revealed accessibility and use of vaccination while controlling for socio-economic confounders.

This study makes contributions in the following ways: first, for the first time in literature, it presents a framework for calibrating spatial healthcare accessibility models using external data. Specifically, the calibration is conducted using regression models between estimated spatial accessibility and usage of healthcare (i.e., vaccine uptake rate). We tested this framework in a case study of COVID-19 vaccines and argue that it can be generalised to a wide range of healthcare services. Second, it contributes to measuring and understanding the spatial accessibility of COVID-19 vaccines. Unlike prior research that used a predefined accessibility method, this study systematically compares and calibrates a range of spatial accessibility methods. Third, it proposes a generalised linear model (specifically, beta regression model) for small-area vaccination uptake that accounts for various geographical and socio-economic variables. The model can inform policymakers and public health practitioners to promote vaccination and to optimise the allocation of COVID-19 vaccination resources.

The rest of this paper is organised as follows. Section 2 describes the key methods in the study. Section 3 introduces relevant datasets and the study area of this study. Section 4 evaluates the results of measuring and validating spatial accessibility to COVID-19 vaccines. Section 5 concludes this paper with a discussion about future research directions.

2. Methodology

This section firstly introduces the overall framework for measuring and calibrating spatial accessibility to healthcare services. After that, it describes the calibration of spatial accessibility models and then the FCA methods that measure spatial accessibility to healthcare services, followed by a brief introduction of the beta regression model.

2.1. Overall Framework

The framework of measuring and calibrating the spatial accessibility to vaccines is illustrated in Figure 1. It consists of two parts. Part one measures the spatial accessibility to vaccines, which utilises three FCA methods (associated with varying parameter values) and generates 15 models of

spatial accessibility. These models account for various factors, including impedance to vaccination sites and congestion of vaccination sites. Part two calibrates spatial accessibility models from Part one by firstly building regression models of small-area vaccine uptake, which account for demographic and socio-economic covariates. Then, the calibration consists of two steps. First, the spatial accessibility models with significant and positive coefficients with vaccine uptake are retained. Second, among the retained accessibility models, the one that leads to the best model fitness is selected as the final model. The rationale and details of the calibration are described in Section 2.2.

Figure 1. Methodology framework of measuring and validating accessibility to vaccination services

2.2. Calibration of spatial accessibility models

In social science research, the validation of computational models refers to testing the model validity, the extent to which the model results measure the real concept under investigation (Bryman 2012). The validation can be classified into six types: conceptual, internal, external, cross-model, data, and security (Carley 2017). Among them, external validation refers to the adequacy and accuracy of the model in matching real-world data, which is based on the linkage between the simulated (or predicted) and the real. More specifically, techniques for external validation are classified into four types, namely grounding, calibrating, verifying, and harmonising (Carley 2017). Among them, calibration is the process of tuning a model to fit granular real data, which is often used for proving that it is feasible for the model to generate results that are aligned with real-world data. Herein, we focus on the calibration of FCA models to measure spatial accessibility to healthcare services (or COVID-19 vaccination services in the case study).

The calibration of spatial accessibility to healthcare services occurs at two levels. The component-level calibration concerns determining the parameters of a component of a predefined model using empirical data. In one study, when measuring the accessibility to hospitals, the distance decay function of the E2SFCA model is calibrated using the empirical relationship between taxi trip frequency and trip duration (Wang *et al.* 2020). In another study (Rader *et al.* 2021), when building an E2SFCA model for the spatial accessibility to COVID-19 vaccine services, the travel-time decay function was calibrated by the survey result regarding people's willingness to travel to the COVID-19 vaccine. On the other hand, the model-level calibration concerns comparing the estimates of spatial accessibility score against the real-world data (e.g. use of medical services and vaccine uptake). To

our knowledge, no existing study has addressed model-level calibration. There are two potential reasons. First, the model-level calibration is hindered by a lack of real data on how people access and use healthcare services due to data privacy concerns. Second, as the revealed accessibility is simultaneously associated with both spatial and non-spatial factors, the model-level calibration requires consideration and control of multiple covariates, which is methodologically more complicated than fitting the distance decay function.

In the case study, we calibrated the spatial accessibility models by using the real-world COVID-19 vaccine uptake data, while controlling relevant covariates, as shown in Figure 1. This calibration is built on the assumption that the valid spatial accessibility to vaccine services should have a positive association with vaccination uptake rates when covariates are controlled. In other words, when the other relevant variables are fixed, the better spatial accessibility of a population group to vaccine services lead to a higher vaccination rate.

2.3. FCA methods

2.3.1. 2SFCA

The 2SFCA method (Luo and Wang 2003) has been introduced to account for supply-to-demand ratios and the number of population in the service area of a supply. It is an innovative extension of the earlier version of FCA, which only considered supply locations in computing the catchment area. This method consists of two steps:

The first step assigns a supply-to-demand ratio to each supply point as a measure of supply availability. The numerator is the supply at location j , and the denominator is the sum of the demand in the catchment of location j . The equation is:

$$R_j = \frac{S_j}{\sum_i D_i I(d_{ij} < d_t)} \quad (1)$$

where R_j and S_j are the supply-to-demand ratio and the capacity at supply point j , respectively, D_i is the size of demand at point i , d_{ij} is the distance between supply j and demand i , d_t is the threshold of service distance, and $I(\cdot)$ is the indicator function.

The second step computes an accessibility measure at demand point i (A_i) by summing up the supply-to-demand ratio of supply points within the catchment area of i . The distance decay effect of supply ratios is not considered. The equation is:

$$A_i = \sum_j R_j I(d_{ij} < d_t) \quad (2)$$

2.3.2. E2SFCA

The E2SFCA was introduced by Luo and Qi (2009) in order to consider the distance decay effects in 2SFCA, which indicates greater accessibility at demand points that are closer to supply. It includes a distance decay weight $w(d_{ij})$ in the computation of supply-to-demand ratio at supply points (Step 1) and the accessibility measure at demand points (Step 2). Specifically, in Step 1, the equation is:

$$R_j = \frac{S_j}{\sum_i D_i w(d_{ij})} \quad (3)$$

In Step 2, the equation is:

$$A_i = \sum_j R_j w(d_{ij}) \quad (4)$$

There are several considerations in E2SFCA. First, the distance weight function may take different forms, such as a step function, a gravity model, or a Gaussian function. Herein, this function is defined as a Gaussian function that has a bandwidth parameter d_b . The Gaussian function declines slowly first and then declines sharply as the distance increases. Second, the E2SFCA here does not include a distance threshold d_t . This is different from the original E2SFCA (Luo and Qi 2009) that incorporates d_t that defines the catchment area and represents the cutoff effect. The assumption underlying d_t is that healthcare sites only provide services for population within the catchment area, which is not realistic, especially for rural areas (Wang *et al.* 2020, Gong *et al.* 2021). Therefore, d_t is not adopted in our E2SFCA and 3SFCA (see below).

2.3.3. 3SFCA

The 3SFCA (Wan *et al.* 2012) was introduced to account for the congestion effect at supply points by adding new weight to the computation of the weighted demand and supply-to-demand ratio. The congestion effect indicates that congestion in nearby supplies might lead demand population to seek service from more distant supplies. In other words, demand is allocated to accessible supplies, depending on the distance and congestion at supplies.

The three steps are as follows:

Step 1:

$$g_{ij} = \frac{w(d_{ij})}{\sum_j w(d_{ij})} \quad (5)$$

where g_{ij} is the additional weight between demand i and supply j .

Step 2:

$$R_j = \frac{S_j}{\sum_i D_i g_{ij} w(d_{ij})} \quad (6)$$

Step 3:

$$A_i = \sum_j R_j g_{ij} w(d_{ij}) \quad (7)$$

The considerations in 3SFCA are similar to E2SFCA. First, $w(d_{ij})$ takes the form of a Gaussian function and has a bandwidth parameter d_b . Second, the distance threshold d_t is not adopted here.

The specifications of FCA methods in this study are shown in Table 1. More details of the data will be introduced in the later sections.

Table 1. Specifications of FCA methods in this study

Variable/parameter	Description
demand (indexed by i)	MSOA population, with population size denoted by D_i
supply (indexed by j)	Covid-19 vaccine sites in England, with unit supply capacity
d_{ij}	Driving distance between demand and supply
$w(d_{ij})$	The Gaussian distance decay function, with a bandwidth parameter d_b (used in E2SFCA and 3SFCA)

2.4. Beta regression model

The implementation and computation of beta regression models were conducted via the ‘betareg’ package (Cribari-Neto and Zeileis 2010) in R language (R Core Team 2015). Beta regression was proposed to perform a regression analysis for continuous dependent variable (called y) that have values in the open unit interval (0,1), e.g. proportions and concentration (Ferrari and Cribari-Neto 2004, Cribari-Neto and Zeileis 2010). A key assumption of this method is that y follows a beta distribution. This method has several advantages: first, the model easily accommodates heteroskedasticity and asymmetries in y . These y variables (e.g. proportions) are typically heteroskedastic as they display more variation around the mean and less in the lower and upper limits, and they are usually asymmetrically distributed. Second, the resulting parameters are interpretable in terms of the mean of y . Third, the model is able to accommodate a variety of y distributions (e.g., flat, left- or right-skewed) due to the flexible combinations of parameter values.

The beta regression method is based on an alternative formulation of the beta density. Originally, the beta density is expressed as:

$$f(y; p, q) = \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} y^{p-1}(1-y)^{q-1}, 0 < y < 1 \quad (8)$$

where $p, q > 0$ and $\Gamma(\cdot)$ is the gamma function. A different parameterisation that sets $\mu = p/(p+q)$ and $\phi = p+q$ (with $0 < \mu < 1$ and $\phi > 0$) is as follows:

$$f(y; \mu, \phi) = \frac{\Gamma(\phi)}{\Gamma(\mu\phi)\Gamma((1-\mu)\phi)} y^{\mu\phi-1}(1-y)^{(1-\mu)\phi-1}, 0 < y < 1 \quad (9)$$

In beta regression, a common notation is $y \sim B(\mu, \phi)$, and $E(y) = \mu$ and $VAR(y) = \mu(1-\mu)/(1+\phi)$.

Herein, we applied a fixed-variation beta regression model where the mean parameter μ_i is not a constant and varies with independent variables, whilst the precision parameter ϕ is constant for all instances. Let $y_1, y_2, \dots, y_n$ be a random sample and $y \sim B(\mu_i, \phi)$. In beta regression, the link function that connects μ_i and independent variables (called x) is defined as:

$$g(\mu_i) = x_i^T \beta = \eta_i \quad (10)$$

where x_i is the vector of k variables, β is a $k \times 1$ vector of regression parameters. The link function $g(\cdot)$ is strictly increasing and twice differentiable, and can assume different types. The common link functions include logit, probit, log-log, and complementary log-log. Here, the logit link function is used, as this link function is either too sensitive or insensitive to the change of the linear predictor, and thus is good for most applications (Smithson and Verkuilen 2006). It is defined as:

$$g(\mu) = \log \frac{\mu}{1-\mu} \quad (11)$$

In this model, the parameter β can be estimated using maximum likelihood. Note that it is also possible to adopt a variable dispersion beta regression model, where the precision parameter ϕ is modelled in a similar way as μ_i . More details of this model can be found in Cribari-Neto and Zeileis (2010).

The metrics of Akaike Information Criterion (AIC) and Pseudo R-squared are used to evaluate the goodness-of-fit of the estimated beta regression and to choose the best-fitted model (Cribari-Neto and Zeileis 2010). The AIC statistic accounts for both the model performance (i.e. how well the model fits the data) and the model complexity (i.e. the number of parameters of the model). For a fitted model M , it is defined as:

$$AIC = -2\ell(M) + 2k \quad (12)$$

where $\ell(M)$ represents the log-likelihood of the model, and k is the number of parameters in the model. Given a set of candidate models, the preferable model is the one with the minimum AIC value.

On the other hand, the Pseudo R-squared, R_p^2 , is defined as the square of the sample correlation between the linear predictor $\hat{\eta}$ and link-transformed response $g(y)$. Note that $0 \leq R_p^2 \leq 1$, and $R_p^2 = 1$ when there is a perfect agreement between $\hat{\eta}$ and $g(y)$, or $\hat{\mu}$ and y . Given a set of candidate models, the one with the maximum R_p^2 value is preferred and selected.

In this study, we propose a beta regression of vaccine uptake rate as follows:

$$\log \frac{VUR}{1-VUR} = \beta_0 + \beta_1 AS + \beta_2 \text{IMD_2} + \beta_3 \text{IMD_3} + \beta_4 \text{IMD_4} + \beta_5 \text{IMD_5} + \beta_6 \text{IMD_6} + \beta_7 \text{IMD_7} + \beta_8 \text{IMD_8} + \beta_9 \text{IMD_9} + \beta_{10} \text{CO} + \beta_{11} \text{E_Asian} + \beta_{12} \text{E_Black} + \beta_{13} \text{E_Mixed} + \beta_{14} \text{E_Other} \quad (13)$$

where VUR denotes vaccine uptake rate, AS denotes accessibility score, IMD_i denotes the i-th decile of IMD, CO denotes car ownership, E_Asiian/Black/Mixed/Other denote the ethnicity composition of Asian/Black/Mixed/Other. These variables have been tested with no significant multi-collinearity using variance inflation factor approach (VIF). The coefficients will be learned using maximum likelihood estimation.

3. Data and study area

We used England, one of the four nations of the United Kingdom, as the study area for this research. Figure 2 shows the map of England, with the boundary of nine regions and local authority districts. In 2020, England has a population of 56.3 million, comprising 84% of the UK population (ONS 2020). England was selected as the study area due to its various publicly available data of vaccination sites and vaccination usage.

Note that Greater London was excluded from this analysis. The main reason is that we used driving distance to represent travel impedance to vaccine sites and this distance does not reflect the travel cost in London. According to statistical data in 2018/19 (UK Department For Transport 2020), private transport accounts for 34.8% of trip-based mode in London, whereas this proportion is over 63% in the other regions of England.

Figure 2. Map of England with the boundary of nine regions and local authority districts

3.1. Demographic Data

We used the Middle Layer Super Output Areas (MSOA) from the 2011 UK census as the spatial unit, which is a geographic area designed to improve the reporting of small area statistics in England and Wales. By definition, each MSOA has an average population of 7,500 residents or 4,000 households. England is divided into 6,791 MSOAs, in which 5,808 MSOAs are outside London.

The most recent MSOA-level demographic information for England is the resident population data from 2019 mid-year estimates (ONS 2020), which includes the estimated population size by age group and sex. This dataset is used as the demand amount for vaccines. MSOAs are represented by population-weighted centroids based on 2011 census data, which is in line with the demand unit representation used by Luo and Wang (2003). c

3.2. Car ownership

The MSOA-level car ownership (i.e. the proportion of households having at least one car or vans) was used as a measure of local travel accessibility (ONS 2011b). The reason for including this variable is that car ownership and local travel accessibility might exhibit a considerable association with the vaccine uptake rates, especially in rural areas where public transport is not widely available.

3.3. Multiple deprivation

We used the 2019 English Index of Multiple Deprivation (IMD) (MHCLG 2019) as a measure of relative small-area deprivation. The IMD was originally calculated for Lower-layer Super Output Areas in England and then aggregated to the MSOA level. ONS recommended using the ranks and deciles of IMD rather than the scores as the scores were less easy to interpret. As such, the IMD is represented by nine binary variables, corresponding to Decile 2 to 10 of IMD with Decile 1 (the most deprived) being the reference level.

3.4. Vaccination sites

The addresses of publicly accessible SARS-CoV-2 vaccination sites ($N=2,868$) in England were retrieved from the NHS England (NHS England 2021) up to 17 November 2021. As the vaccine supply in each site is not available, we assume an equal vaccine supply across sites. These sites fall into three types, including vaccination centres ($N=108$), hospital hubs ($N=230$), and local vaccination services ($N=2530$). The local vaccination services include pharmacies and GP-led vaccination services. Although another type of vaccination site, mobile units, is in operation, the locations of these units are not publicly available and therefore they were excluded herein.

The original vaccination site data contain address and postcode (e.g. 'Airedale Hospital NHS FT, Skipton Road, Keighley, West Yorkshire, BD20 6TD'), which were geocoded into WGS-84 coordinates of (longitude, latitude) using the Geocoding API from Google Maps Platform (Google 2021).

3.5. Vaccination uptake rates

There are different ways of calculating the vaccine uptake rate, either across the whole population or for an age group. We collected the MSOA-level weekly number of vaccination uptake by age groups in England from the NHS England (NHS 2021), which started on 8 December 2020. According to the vaccine delivery plan (UK Department of Health and Social Care 2021), NHS has been prioritising high-risk residents (including the elderly) and offering vaccines to different age groups in stages. In this study, we used the accumulative second-dose vaccination uptake rates of the over-18 population up to 18 November 2021, as this date was close to the date when the vaccination site data were retrieved.

3.6. Travel Distance

The shortest driving distance between MSOA centroids and vaccination sites was computed by the Open Source Routing Machine (OSRM) (OSRM Project 2021). OSRM is a high-performance and open-source routing engine, which is designed to retrieve and apply OpenStreetMap data via HTTP API and other interfaces. OSRM provides varying services including snapping coordinates to a street network, identifying the fastest route between coordinates, computing the pairwise shortest travel times/distances between supplied coordinates.

We consider the driving distance as the measurement of spatial barriers, as driving is the main commute and travel mode in most areas of England (UK Department For Transport 2020). In addition, using a single mode of driving simplifies the computing of distance and accessibility. As Greater London has significantly lower car ownership compared to other England regions and driving is not the main travel mode there, Greater London was excluded in the following analysis.

4. Results and discussion

4.1. Measuring accessibility to vaccination

We computed the accessibility score using three methods (i.e. 2SFCA, E2SFCA, and 3SFCA) at varying distance parameter values (i.e., 10, 15, 20, 25, 30 miles) by using the Python module 'access' (Saxon *et al.* 2021). The descriptive statistics of the 15 accessibility scores are presented in Table 2. Using 2SFCA, with the increase in the distance threshold, both the average and variance of accessibility score gradually increase. In contrast, a different trend is observed in the E2SFCA and 3SFCA methods: the mean accessibility score slightly increases with the bandwidth, as is similar to 2SFCA; however, the variance of accessibility score decreases when the bandwidth increases from 10 to 30 miles. This can be explained by that a larger bandwidth leads to a greater smoothing effect of the kernel function and therefore the accessibility score.

Table 2. Descriptive statistics of 15 accessibility scores (10^{-5})

Panel 1: 2SFCA					
bandwidth (miles)	10	15	20	25	30
Mean	6.468	6.485	6.482	6.484	6.480
Std. Dev.	4.086	3.215	2.882	2.735	2.681
Min	0.000	0.000	0.000	0.000	0.000
Max	64.300	21.400	15.700	16.900	16.800
Percentiles					
5%	0.000	0.000	0.000	0.000	0.000
25%	4.100	4.700	5.100	5.200	5.200
50%	6.200	6.500	6.500	6.600	6.700

75%	8.425	8.300	8.400	8.200	8.200
95%	13.000	11.800	10.900	10.500	10.300
Panel 2: E2SFCA					
bandwidth (miles)	10	15	20	25	30
Mean	6.405	6.440	6.458	6.467	6.471
Std. Dev.	4.910	3.907	3.389	3.089	2.906
Min	0.000	0.000	0.000	0.000	0.000
Max	54.900	30.400	23.000	18.000	14.900
Percentiles					
	0				
5%	0.000	0.000	0.000	0.000	0.000
25%	3.300	4.000	4.400	4.700	5.000
50%	5.850	6.100	6.300	6.500	6.500
75%	8.700	8.500	8.300	8.300	8.200
95%	14.400	13.100	12.300	11.600	11.200
Panel 3: 3SFCA					
bandwidth (miles)	10	15	20	25	30
Mean	6.444	6.462	6.472	6.479	6.481
Std. Dev.	5.288	4.124	3.510	3.154	2.928
Min	0.000	0.000	0.000	0.000	0.000
Max	74.600	56.900	36.400	25.100	19.800
Percentiles					
5%	0.000	0.000	0.000	0.000	0.000
25%	3.500	4.300	4.700	5.000	5.100
50%	5.900	6.100	6.300	6.400	6.500
75%	8.500	8.300	8.200	8.100	8.000
95%	14.200	12.900	12.100	11.500	11.165

Figure A. 1–A. 3 (see Appendix A) show the maps of these accessibility scores. It is observed that areas close to London, the North West, and North East region had higher accessibility scores, whereas the South West and South East regions had relatively low accessibility to vaccination. In addition, the maps reveal that a larger d_t or d_b (distance threshold or bandwidth) led to a stronger spatial smoothing effect.

4.2. Selecting the optimal accessibility model

Table 3 shows the results of beta regression of vaccine uptake rates, which uses different accessibility models while controlling for a range of covariates, including ethnic composition, IMD decile, and car ownership. The selection of the optimal accessibility model follows two steps: first, the accessibility models that exhibit a positive and significant coefficient (p -value less than 0.05) in the beta regression are retained, while the other models are excluded. Second, among the retained models, the one with the minimum AIC value is selected as the optimum. As shown in Table 3, the 3SFCA accessibility score with $d_b=30$ miles is the optimum, with an AIC value of -16858.2. Although the 2SFCA accessibility score with $d_t=10$ miles has a slightly higher Pseudo R^2 than 3SFCA accessibility with $d_b=30$ miles, we consider AIC value as a more important model performance metric than Pseudo R^2 as AIC accounts for both the model fitness and model complexity.

The spatial distribution of accessibility to vaccination calculated by 3SFCA with $d_b = 30$ miles is shown in Figure 3. An overall pattern is that areas in the North East and North Western regions have higher accessibility scores, whereas the South West, West Midlands, and East Midlands region have relatively low accessibility. Interestingly, in South East region, some highest-accessibility-level MSOAs are adjacent to MSOAs associated with the lowest accessibility level, reflecting a sharp change of accessibility.

Figure 3. Maps of 3SFCA accessibility score with $d_b=30$ miles

Table 3. Beta regression results of vaccine uptake rates using 15 accessibility scores

Method	bandwidth (Miles)	Coefficient of accessibility score		Model performance	
		Estimate	Significance	AIC	Pseudo R^2
2SFCA	10	1292.2	***	-16858.0	0.3113
	15	1202.0	***	-16838.4	0.3101
	20	483.4	*	-16819.8	0.3089
	25	-331.1		-16818.2	0.3086
	30	-639.7	**	-16821.2	0.3088
E2SFCA	10	570.4	***	-16827.0	0.3090

	15	1096.0	***	-16841.8	0.3097
	20	1485.7	***	-16852.4	0.3104
	25	1533.8	***	-16848.7	0.3103
	30	1329.0	***	-16838.2	0.3098
3SFCA	10	613.4	***	-16829.8	0.3090
	15	888.9	***	-16834.8	0.3093
	20	1162.0	***	-16840.6	0.3097
	25	1487.7	***	-16849.2	0.3103
	30	1799.7	***	-16858.2	0.3109

Significance Code: ***:<0.001, **:<0.01, *:<0.05

According to the sensitivity analysis in Table 3, the accessibility bandwidth d_b plays an important role in the validity and goodness-of-fit of the regression models of vaccine rates. Specifically, given a method of measuring accessibility score, a larger bandwidth leads to a less realistic accessibility score and vaccine rate model. For example, when the threshold d_t equals 25 or 30 miles for 2SFCA, the beta regression coefficient of the accessibility score becomes negative, which is counter-intuitive and invalid. This might result from the considerable spatial smoothing effect due to a large bandwidth parameter.

4.3. Further analysis of the optimal vaccine uptake model

We performed a detailed analysis of the optimal regression model of vaccine uptake rate (3SFCA with $d_b=30$ miles). To measure the multicollinearity of this model, we computed the variance inflation factor (VIF) of each independent variable. Results show that all VIF values are below five, which indicates insignificant multicollinearity between variables. Table 4 is a summary of this regression model. As shown in Panel 1, the Pseudo R-squared is 0.311, meaning that 31.1% of the variation of the logit of vaccine uptake can be explained by the predictors. This R_p^2 is comparable to a previous study (Bauer *et al.* 2020), in which the R^2 values between FCA accessibility scores and actual hospital visits due to six medical diagnoses fall into a range between 0.132 and 0.473. The low explained variance can be explained by the omittance of important variables, such as the supply capacity of vaccine sites.

Table 4. Summary of the optimal regression model of vaccine rates

Panel 1 (model summary)

Degree of freedom	17
Log-likelihood	8,450.00
AIC	- 16,858.20
Pseudo R-squared	0.311

Panel 2 (model coefficient)

Variable	Coefficient estimate	Std Error	z value	Pr(> z)
precision_parameter	21.352	0.403	52.929	0.000***
Intercept	-0.519	0.078	-6.662	0.000***

Access_score	1800	288	6.259	0.000***
IMD_Decile_1 (reference level)	-	-	-	-
IMD_Decile_2	-0.127	0.031	-4.094	0.000***
IMD_Decile_3	-0.144	0.035	-4.093	0.000***
IMD_Decile_4	-0.171	0.038	-4.518	0.000***
IMD_Decile_5	-0.181	0.040	-4.480	0.000***
IMD_Decile_6	-0.144	0.043	-3.377	0.001***
IMD_Decile_7	0.052	0.046	1.129	0.259
IMD_Decile_8	0.131	0.048	2.709	0.007***
IMD_Decile_9	0.335	0.050	6.628	0.000***
IMD_Decile_10	0.561	0.054	10.468	0.000***
car_ownership	3.305	0.116	28.524	0.000***
White (reference level)	-	-	-	-
Mixed	-5.910	0.870	-6.796	0.000***
Asian	-0.379	0.080	-4.720	0.000***
Black	-0.027	0.344	-0.080	0.937
Other	-4.520	0.942	-4.798	0.000***

Significance Code: ***:<0.001, **:<0.01, *:<0.05

Table 4 reveals the level of association between each explanatory variable and vaccine uptake. To illustrate this association, we focus on two MSOAs, m_1 with the lowest fitted vaccine uptake and m_2 with the lowest accessibility score (see Table 5). The accessibility score demonstrates a statistically significant association, with $p<0.001$. Assuming all other variables are fixed, if the accessibility score increases by 10^{-5} , the estimated odds of vaccine uptake rate would increase by a factor of 1.018, which is calculated by $e^{1800*10^{-5}} = 1.018$. In other words, if the accessibility score of m_1 increases to $7.7000*10^{-5}$, its estimated vaccine uptake would increase from 0.2707 to 0.2742. Likewise, the estimated vaccine uptake of m_2 would become 0.7731 if its accessibility score becomes $1*10^{-5}$.

Table 5. Attributes and predicted vaccine uptake rates of two MSOAs

MSOA Code	Vaccine update rate		IMD decile	Car ownership	Ethnicity proportion
-----------	---------------------	--	------------	---------------	----------------------

		Actual	Predicted*	Accessibility to vaccination (10 ⁻⁵)			Mixed	Asian	Black	Other
m_1	E02001385	0.5823	0.2707	6.7000	1	0.3148	0.3548	0.1040	0.0918	0.2332
m_2	E02003245	0.7450	0.7699	0.0000	1	0.6175	0.7895	0.0358	0.1294	0.0337

*Note: the predicted vaccine uptake rate is calculated using attributes in this table and coefficient in Table 4, based on Equation (13).

In the model, the multiple deprivation is represented by nine dummy variables, corresponding to Decile 2 to 10 of IMD with Decile 1 being the reference level. All Deciles demonstrate a significant association with vaccine uptake except for Decile 7. Moreover, the coefficient estimates reveal that given all other variables being fixed, Decile 10 leads to the highest vaccine uptake, followed by Decile 9, 8, 7, 1, 2, 3, 6, 4, and 5. To specify the impact of IMD deciles on vaccine uptake, we examine the relationship between decile and vaccine uptake per decile pair. An observation is that among 32 pairs of 45 possible combinations, the high decile (or low deprivation) of an MSOA corresponds to a high estimated vaccine uptake rate. For example, if m_1 is promoted to Decile 10, its vaccine uptake would increase to 0.3941, which is a considerable improvement.

Notably, car ownership shows a significant association with the vaccine uptake rate. Quantitatively, assuming the other variables are fixed, a one percent increase in car ownership would increase the odds of vaccine uptake by a factor of 1.034. More specifically, if the car ownership of m_1 and m_2 increases by 0.01, the estimated vaccine uptake would become 0.2773 and 0.7757, respectively. This finding is consistent with the report by the UK Parliament (Mutebi 2021), which points out that accessibility of transport to and from vaccine facilities is one factor of the barrier to accessibility of vaccines.

The ethnic composition exhibits a statistically significant association with the vaccine uptake, except for the 'Black' group. The 'White' group is associated with the highest vaccine uptake, followed by 'Black', 'Asian', 'Other', and 'Mixed' groups. This result is largely consistent with the reported vaccination rates (from 8 December 2020 to 15 May 2021) by socio-demographic groups among people over 40 years living in England (Office for National Statistics 2021). Specifically, it is reported that the 'White British' has the highest vaccination rate, followed by 'Bangladeshi', 'Black African', 'Black Caribbean', 'Chinese', 'Indian', 'Mixed', 'Other', 'Pakistani', and 'White Other'. Although disparities exist in the two rankings and the two datasets do not completely match in timescale and age groups, the trend in common is that 'White' group outperforms the other ethnicities in vaccination rates and 'Black' outperforms 'Asian' and 'Mixed'.

Next, we conducted residual analysis for this model, with the standardised residual plot given in Figure 4. The standardised residual of the beta regression model is defined as follows (Ferrari and Cribari-Neto 2004):

$$r_t = \frac{y_t - \hat{y}_t}{\sqrt{\widehat{\text{var}}(y_t)}} = \frac{y_t - \hat{y}_t}{\hat{y}_t(1 - \hat{y}_t)/(1 + \phi)} \quad (14)$$

where y_t and $\hat{y}_t$ denote the actual and predicted vaccine uptake rate of the t-th MSOA, respectively; $\sqrt{\widehat{\text{var}}(y_t)}$ denotes the estimated standard deviation of the vaccine uptake rate of the t-th MSOA. If

the beta regression model is appropriate, a plot of the standardised residual against the observation index should show no detectable patterns and no significant outliers. In practice, the observations that are far from the horizontal axis are likely outliers.

An inspection of this residual plot reveals that the five instances with the absolute standardised residual greater than 6 are likely to be outliers, as highlighted by the rectangle in Figure 4. Here, the threshold value of 6 is selected based on the visual interpretation of the residual pattern. We re-fitted a beta regression model without these 5 outlier instances and noted a slight improvement in the AIC value from -16,858.2 to -16,997.3 (see Table 6). However, the coefficient estimates (point estimate +/- standard error) do not considerably differ from the original model. Therefore, we stick to the original beta regression model.

Figure 4. The standardised residual plot for the beta regression model (using 3SFCA, $d_i=30$ mile). The x-axis denotes the observation number of MOSAs. Outliers with standardised residual greater than 6 are highlighted in rectangles.

Table 6. Summary of the refitted regression model of vaccine uptake

Panel 1 (model summary)				
Degree of freedom	17			
Log-likelihood	8520			
AIC	-16997.3			
Pseudo R-squared	0.311			

Panel 2 (model coefficient)				
Variable	Coefficient estimate	Std Error	z value	Pr(> z)
Precision parameter	21.960	0.415	52.919	0.000***
Intercept	-0.472	0.077	-6.105	0.000***
Access score	1858	285	6.516	0.000***
IMD_Decile_1 (reference level)	-	-	-	-
IMD_Decile_2	-0.119	0.031	-3.858	0.000***
IMD_Decile_3	-0.129	0.035	-3.729	0.000***
IMD_Decile_4	-0.151	0.037	-4.047	0.000***
IMD_Decile_5	-0.158	0.040	-3.950	0.000***
IMD_Decile_6	-0.103	0.042	-2.425	0.015*
IMD_Decile_7	0.083	0.046	1.809	0.070
IMD_Decile_8	0.172	0.048	3.589	0.000***
IMD_Decile_9	0.376	0.050	7.492	0.000***
IMD_Decile_10	0.604	0.053	11.346	0.000***
Car ownership	3.214	0.115	27.921	0.000***
White (reference level)	-	-	-	-
Mixed	-5.744	0.860	-6.678	0.000***
Asian	-0.370	0.079	-4.664	0.000***
Black	-0.109	0.340	-0.320	0.749
Other	-4.739	0.931	-5.089	0.000***

Significance Code: ***:<0.001, **:<0.01, *:<0.05

5. Conclusions

This study has proposed a model-level calibration framework for spatial accessibility to healthcare services. We adopted the COVID-19 mass vaccination programme in England as a case study. Specifically, we measured and calibrated the MSOA-level spatial accessibility to vaccination services, using a variety of floating catchment area models. The model calibration was conducted using beta regression analysis to test the association between the vaccination uptake rates and spatial accessibility to vaccine sites, controlled for socio-economic and demographic covariates. Results

show that the best-fitting model of spatial accessibility to vaccines is the three-step FCA model with a kernel bandwidth of 30 miles, which exhibits a statistically significant and positive effect on vaccination uptake. More specifically, if the spatial accessibility score increases by 10^{-5} , the estimated odds of vaccine uptake would increase by a factor of 1.018. The covariates including multiple deprivation deciles, car ownership percentage, and ethnic composition also have a statistically significant association with the vaccination uptake. This case study proves the feasibility of using revealed accessibility to validate and calibrate the spatial accessibility to healthcare services at a model level.

The contributions of this study are three-fold. First, for the first time in literature, it presents a framework for calibrating spatial healthcare accessibility models using revealed accessibility data. This framework is tested in the case study of COVID-19 vaccination services in England, UK. Second, it contributes to the understanding of the spatial accessibility of COVID-19 vaccines. Unlike prior research that uses a predefined accessibility method (Rader *et al.* 2021), this study systematically compares and calibrates a range of accessibility models, leading to a more reliable estimate of spatial accessibility to vaccination. Third, it proposes a generalised linear model for small-area vaccination uptake rates, which incorporates various geographical and socio-economic variables. The model can inform policymakers and public health practitioners about high-risk regions with low accessibility to the COVID-19 vaccine and help them address resource allocation in the country to battle COVID-19.

This study has several limitations, which opens avenues for future research. First, it assumes that residents travel by car to get the vaccine and use driving distance in computing spatial accessibility. However, this is likely not true in city centres where driving is much less popular than public transport. We excluded London for this reason, where COVID-19 severely impacted residents. Future work needs to integrate the various travel modes in the accessibility to vaccination. Second, we do not account for the spatial autocorrelation of vaccine uptake in the regression models. In future work, it would be promising to build spatial regression models for vaccine uptake (e.g. using geographically weighted regression), which would reveal the spatial heterogeneity and local effect of how vaccine uptake rate is related to spatial and socio-economic variables.

Disclosure statement

The authors report there are no competing interests to declare.

500

501 Figure A. 1. Maps of 2SFCA with $d_t=10,15,20,25,30$ miles. From left to right and top to bottom are (a)

502 $d_t=10$, (b) $d_t=15$, (c) $d_t=20$, (d) $d_t=25$, and (e) $d_t=30$, unit in miles

503

504

Figure A. 2. Maps of E2SFCA with $d_b=10,15,20,25,30$ miles. From left to right and top to bottom are (a) $d_b=10$, (b) $d_b=15$, (c) $d_b=20$, (d) $d_b=25$, and (e) $d_b=30$, unit in miles

Figure A. 3. Maps of 3SFCA with $d_b=10,15,20,25,30$ miles. From left to right and top to bottom are (a) $d_b=10$, (b) $d_b=15$, (c) $d_b=20$, (d) $d_b=25$, and (e) $d_b=30$, unit in miles

References

- Bauer, J., Klingelhöfer, D., Maier, W., Schwettmann, L., and Groneberg, D.A., 2020. Prediction of hospital visits for the general inpatient care using floating catchment area methods: A reconceptualization of spatial accessibility. *International Journal of Health Geographics*, 19 (1), 1–11.
- Bryman, A., 2012. *Social Research Methods*. 4th ed. Oxford, UK: Oxford University Press.
- Carley, K.M., 2017. *Validating Computational Models*.
- Chu, D.K., Akl, E.A., Duda, S., Solo, K., Yaacoub, S., Schünemann, H.J., El-harakeh, A., Bognanni, A., Lotfi, T., Loeb, M., Hajizadeh, A., Bak, A., Izcovich, A., Cuello-Garcia, C.A., Chen, C., Harris, D.J., Borowiack, E., Chamseddine, F., Schünemann, F., Morgano, G.P., Muti Schünemann, G.E.U., Chen, G., Zhao, H., Neumann, I., Chan, J., Khabisa, J., Hneiny, L., Harrison, L., Smith, M., Rizk, N., Giorgi Rossi, P., AbiHanna, P., El-khoury, R., Stalteri, R., Baldeh, T., Piggott, T., Zhang, Y., Saad, Z., Khamis, A., and Reinap, M., 2020. Physical distancing, face masks, and eye protection to prevent person-to-person transmission of SARS-CoV-2 and COVID-19: a systematic review and meta-analysis. *The Lancet*, 395 (10242), 1973–1987.
- Cribari-Neto, F. and Zeileis, A., 2010. Beta regression in R. *Journal of Statistical Software*, 34 (2), 1–24.

529 Ferrari, S. and Cribari-Neto, F., 2004. Beta Regression for Modelling Rates and Proportions.
530 <http://dx.doi.org/10.1080/0266476042000214501>, 31 (7), 799–815.

531 Gong, S., Gao, Y., Zhang, F., Mu, L., Kang, C., and Liu, Y., 2021. Evaluating healthcare resource
532 inequality in Beijing, China based on an improved spatial accessibility measurement.
533 *Transactions in GIS*, 25 (3), 1504–1521.

534 Google, 2021. Geocoding API of Google Maps Platform [online]. Available from:
535 <https://developers.google.com/maps/documentation/geocoding/start> [Accessed 23 Jul 2021].

536 Guagliardo, M.F., 2004. Spatial accessibility of primary care: concepts, methods and challenges.
537 *International Journal of Health Geographics*, 3, 3.

538 Hsiang, S., Allen, D., Annan-Phan, S., Bell, K., Bolliger, I., Chong, T., Druckenmiller, H., Huang, L.Y.,
539 Hultgren, A., Krasovich, E., Lau, P., Lee, J., Rolf, E., Tseng, J., and Wu, T., 2020. The effect of
540 large-scale anti-contagion policies on the COVID-19 pandemic. *Nature*, 584 (7820), 262–267.

541 Kang, J.Y., Michels, A., Lyu, F., Wang, S., Agbodo, N., Freeman, V.L., and Wang, S., 2020. Rapidly
542 measuring spatial accessibility of COVID-19 healthcare resources: a case study of Illinois, USA.
543 *International Journal of Health Geographics*, 19 (1).

544 Khan, A.A., 1992. An integrated approach to measuring potential spatial access to health care
545 services. *Socio-Economic Planning Sciences*, 26 (4), 275–287.

546 Ledford, H., Cyranoski, D., and Van Noorden, R., 2020. The UK has approved a COVID vaccine - here's
547 what scientists now want to know. *Nature*, 588 (7837), 205–206.

548 Lewnard, J.A. and Lo, N.C., 2020. Scientific and ethical basis for social-distancing interventions
549 against COVID-19. *The Lancet Infectious Diseases*, 20 (6), 631–633.

550 Luo, W. and Qi, Y., 2009. An enhanced two-step floating catchment area (E2SFCA) method for
551 measuring spatial accessibility to primary care physicians. *Health and Place*, 15 (4), 1100–1107.

552 Luo, W. and Wang, F., 2003. Measures of spatial accessibility to health care in a GIS environment:
553 Synthesis and a case study in the Chicago region. *Environment and Planning B: Planning and*
554 *Design*, 30 (6), 865–884.

555 MHCLG, 2019. *National Statistics English Indices of Deprivation 2019*.

556 Mutebi, N., 2021. COVID-19 vaccine coverage and targeted interventions to improve vaccination
557 uptake - POST [online]. Available from: [https://post.parliament.uk/covid-19-vaccine-coverage-](https://post.parliament.uk/covid-19-vaccine-coverage-and-targeted-interventions-to-improve-vaccination-uptake/)
558 [and-targeted-interventions-to-improve-vaccination-uptake/](https://post.parliament.uk/covid-19-vaccine-coverage-and-targeted-interventions-to-improve-vaccination-uptake/) [Accessed 20 Aug 2021].

559 NHS, 2021. COVID-19 Vaccinations Archive [online]. Available from:
560 [https://www.england.nhs.uk/statistics/statistical-work-areas/covid-19-vaccinations/covid-19-](https://www.england.nhs.uk/statistics/statistical-work-areas/covid-19-vaccinations/covid-19-vaccinations-archive/)
561 [vaccinations-archive/](https://www.england.nhs.uk/statistics/statistical-work-areas/covid-19-vaccinations/covid-19-vaccinations-archive/).

562 NHS England, 2021. Coronavirus Vaccination Sites [online]. Available from:
563 <https://www.england.nhs.uk/coronavirus/publication/vaccination-sites/> [Accessed 23 Jul
564 2021].

565 Office for National Statistics, 2021. COVID-19 vaccination rates and odds ratios by socio-
566 demographic group [online]. Available from:
567 <https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/healthinequalit>

- ies/datasets/covid19vaccinationratesandoddsratiosbysociodemographicgroup [Accessed 20 Aug 2021].
- ONS, 2011a. LC2101EW - Ethnic group by sex by age [online]. Available from: <https://www.nomisweb.co.uk/query/construct/summary.asp?mode=construct&version=0&dataset=801>.
- ONS, 2011b. Car or van availability [online]. *Nomis*. Available from: <https://www.nomisweb.co.uk/census/2011/qs416ew>.
- ONS, 2020. Population estimates for the UK, England and Wales, Scotland and Northern Ireland [online]. Available from: <https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/bulletins/annualmidyearpopulationestimates/latest#population-growth-in-england-wales-scotland-and-northern-ireland> [Accessed 13 Aug 2021].
- OSRM Project, 2021. Open Source Routing Machine. Modern C++ Routing Engine for Shortest Paths in Road Networks [online]. Available from: <http://project-osrm.org/> [Accessed 23 Jul 2021].
- R Core Team, 2015. R: A language and environment for statistical computing [online]. Available from: <https://www.r-project.org/> [Accessed 1 Sep 2017].
- Rader, B., Astley, C.M., Sewalk, K., Paul, L., Cordiano, K., Wronski, L., Rivera, J.M., Hallberg, K., Pera, M.F., Cantor, J., Whaley, C.M., Bravata, D.M., and Brownstein, J.S., 2021. Spatial Accessibility Modeling of Vaccine Deserts as Barriers to Controlling SARS-CoV-2. *medRxiv*, 1–20.
- Saxon, J., Koschinsky, J., Acosta, K., Anguiano, V., Anselin, L., and Rey, S., 2021. An open software environment to make spatial access metrics more accessible. *Journal of Computational Social Science*.
- Smithson, M. and Verkuilen, J., 2006. A better lemon squeezer? Maximum-likelihood regression with beta-distributed dependent variables. *Psychological Methods*, 11 (1), 54–71.
- UK Department For Transport, 2020. Mode of travel - GOV.UK [online]. Available from: <https://www.gov.uk/government/statistical-data-sets/nts03-modal-comparisons#mode-by-region> [Accessed 13 Aug 2021].
- UK Department of Health and Social Care, 2021. UK COVID-19 vaccines delivery plan [online]. Available from: <https://www.gov.uk/government/publications/uk-covid-19-vaccines-delivery-plan>.
- Wan, N., Zou, B., and Sternberg, T., 2012. A three-step floating catchment area method for analyzing spatial access to health services. *International Journal of Geographical Information Science*, 26 (6), 1073–1089.
- Wang, J., Du, F., Huang, J., and Liu, Y., 2020. Access to hospitals: Potential vs. observed. *Cities*, 100, 102671.
- WHO, 2020. WHO announces outbreak of COVID-19 as a pandemic [online]. Available from: <https://www.euro.who.int/en/health-topics/health-emergencies/coronavirus-covid-19/news/news/2020/3/who-announces-covid-19-outbreak-a-pandemic>.