Modelling the COVID-19 vaccine uptake rates in a geographical and socioeconomic context: a case study of England

Huanfa Chen*[1], Yanjia Cao†[2], Lingru Feng‡[1] and Qunshan Zhao§[3]

[1] Centre for Advanced Spatial Analysis, University College London
[2] School of Public Health and Human Longevity Science, University of California, San Diego
[3]Urban Big Data Centre, University of Glasgow

January 17, 2022

Summary

The global Covid-19 pandemic has posed unprecedented social and economic challenges to many countries, including the United Kingdom. One of the key strategies to contain the pandemic is mass vaccination. The Covid-19 vaccine uptake rate of a population group depends on a range of geographical and socio-economic factors, including accessibility to vaccination, ethnic composition, deprivation levels, etc. However, limited research has been conducted to obtain a quantitative understanding of how these factors are associated with the Covid-19 vaccine uptake rates. This study fills this gap by proposing a beta regression model for the small-area Covid-19 vaccine uptake rates in England. The findings have important implications for the practice and policymaking of advocating vaccination programmes and other healthcare services.

KEYWORDS: Covid-19 vaccine uptake, socioeconomic factors, beta regression, spatial accessibility, public health

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 since early 2020 (WHO 2020). As of November 2021, in the United Kingdom (UK), there are over 9 million reported cases and over 140,000 deaths. Mass vaccination has been utilised as one of the primary control strategies for minimising SARS-COV-2 infection cases and transmissions. In December 2020, the UK government issued emergency-use authorisation for COVID-19 vaccines and administered mass vaccination to all residents.

The acceptance of Covid-19 vaccination is heterogenous across space or population groups. It is acknowledged that a wide range of factors are likely associated with the vaccine uptake, which include accessibility to vaccination services, accessibility to transport, culture, and deprivation. However, limited research has been conducted that leads to a comprehensive and quantified understanding of Covid-19 vaccine uptake rates. This study fills this gap by proposing a beta regression model for the small-area Covid-19 vaccine uptake rates. We used the Middle Layer Super Output Areas (MSOA) from the 2011 UK census as the spatial unit.

This paper is organised as follows. In Section 2, we introduce the key methods in this study. In Section 3, we describe the key datasets for this research. In Section 4, we present the beta regression model and discuss the implications. Section 5 concludes this paper by discussing future research directions.

* huanfa.chen@ucl.ac.uk
† yacao@health.ucsd.edu
‡ lingru.feng.18@alumni.ucl.ac.uk
§ qunshan.zhao@glasgow.ac.uk

2. Data

We used England as the study area for this research but excluded Greater London. England was selected as the study area due to its various publicly available data of vaccination sites and vaccination usage. The main reason for excluding Greater London is that we used driving distance to represent travel impedance to vaccine sites and this distance does not reflect the travel cost in London.

2.1. Demographic data

The most recent MSOA-level demographic information for England is the resident population data from 2019 mid-year estimates (ONS 2020). The over-18 population estimate is used as the demand amount for vaccines. In addition, the MSOA ethnic composition is represented by the population proportion of five ethnic groups (i.e. 'White', 'Black', 'Asian', 'Mixed', and 'Other') (ONS 2011a), which is derived from the 2011 census data.

2.2. 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.

2.3. Multiple deprivation

We used the 2019 English Index of Multiple Deprivation (IMD) (MHCLG 2019) as a measure of relative small-area deprivation. The IMD is represented by nine binary variables, corresponding to Decile 2 to 10 of IMD with Decile 1 (the most deprived) being the reference level.

2.4. 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. The original site data contained address and postcode and were geocoded into latitude-longitude coordinates using the Geocoding API from Google Maps Platform.

2.5. Vaccine uptake rates

We retrieved the MSOA-level weekly number of vaccination uptake by age groups in England from the NHS England (NHS 2021), and used the accumulative second-dose vaccination uptake rates of the over-18 population up to 18 November 2021, as this date was close to when the vaccination site data were retrieved.

3. Methods

3.1. Beta regression model

Beta regression was proposed to perform a regression analysis for continuous dependent variable (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). We used the 'betareg' package (Cribari-Neto and Zeileis 2010) in R language (R Core Team 2015) to implement and interpret the beta regression models.

Assuming that the vaccine uptake rates follow a beta distribution, we propose a beta regression of MSOA-level vaccine uptake rate as follows:

$$\log\frac{VUR}{1-VUR} = \beta_0 + \beta_1 AS + \beta_2 \mathrm{IMD\_2} + \beta_3 \mathrm{IMD\_3} + \beta_4 \mathrm{IMD\_4} + \beta_5 \mathrm{IMD\_5} + \beta_6 \mathrm{IMD\_6} + \beta_7 \mathrm{IMD\_7} + \beta_8 \mathrm{IMD\_8} + \beta_9 \mathrm{IMD\_9} + \beta_{10} \mathrm{CO} + \beta_{11} \mathrm{E\_Asian} + \beta_{12} \mathrm{E\_Black} + \beta_{13} \mathrm{E\_Mixed} + \beta_{14} \mathrm{E\_Other} \quad (1)$$

where VUR denotes vaccine uptake rate; AS denotes accessibility score that is calculated by the Three-step Floating Catchment Area (3SFCA) with a bandwidth parameter of 30 miles; IMD_i denotes the i-th decile of IMD; CO denotes car ownership; E_Asian/Black/Mixed/Other denote the ethnicity composition of Asian/Black/Mixed/Other. These variables have no significant multi-collinearity using variance inflation factor approach (VIF). The coefficients will be learned using maximum likelihood estimation.

4. Results and discussion

The beta regression model of vaccine uptake rates is presented in Table 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-square value is comparable to a previous study (Bauer et al. 2020), in which the R^2 values between spatial 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.

All independent variables exhibited a statistically significant relationship with the vaccine uptake rates, except for IMD Decile 7 and the Black ethnicity composition. Specifically, a larger accessibility score or a higher car ownership will likely lead to a higher vaccine uptake rate. In terms of IMD, a general trend is that a lower level of deprivation will lead to a higher vaccine uptake rate. The ethnicity composition has a considerable association with the vaccine uptake rates. More specifically, 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). This report states 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'.

Table 1. Summary of the beta regression model of vaccine uptake rates

Panel 1 (model summary)				
Degree of freedom	17			
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***
Accessibility 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***

5. Conclusions

In this study, we proposed a beta regression model for the small-area vaccine uptake rates in England. This model provides insights into how the geographical and socioeconomic factors are associated with the vaccine uptake rates and have important implications for the vaccination programmes. Further study will seek to incorporate other influential variables or spatial autocorrelation of the vaccine uptake rates into this model in order to improve model fitting performance.

6. Acknowledgements

We acknowledge the Office for National Statistics and NHS England for their data support.

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. https://doi.org/10.1186/s12942-020-00223-3

Cribari-Neto, F. and Zeileis, A. (2010) Beta regression in R. *Journal of Statistical Software*, 34 (2),1–24. https://doi.org/10.18637/jss.v034.i02

Ferrari, S. and Cribari-Neto, F. (2004) Beta Regression for Modelling Rates and Proportions. *http://dx.doi.org/101080/0266476042000214501*, 31 (7),799–815. https://doi.org/10.1080/0266476042000214501

MHCLG (2019) National Statistics English Indices of Deprivation 2019

NHS (2021) COVID-19 Vaccinations Archive. https://www.england.nhs.uk/statistics/statistical-work-areas/covid-19-vaccinations/covid-19-vaccinations-archive/

Office for National Statistics (2021) COVID-19 vaccination rates and odds ratios by socio-demographic group. https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/healthinequalities/datasets/covid19vaccinationratesandoddsratiosbysociodemographicgroup. [Accessed 20 Aug 2021]

ONS (2020) Population estimates for the UK, England and Wales, Scotland and Northern Ireland. https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/bulletins/annualmidyearpopulationestimates/latest#population-growth-in-england-wales-scotland-and-northern-ireland. [Accessed 13 Aug 2021]

ONS (2011a) LC2101EW - Ethnic group by sex by age. https://www.nomisweb.co.uk/query/construct/summary.asp?mode=construct&version=0&dataset=801

ONS (2011b) Car or van availability. In: Nomis. https://www.nomisweb.co.uk/census/2011/qs416ew

R Core Team (2015) R: A language and environment for statistical computing. https://www.r-project.org/. [Accessed 1 Sep 2017]

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. https://doi.org/10.1080/13658816.2011.624987

WHO (2020) WHO announces outbreak of COVID-19 as a pandemic. https://www.euro.who.int/en/health-topics/health-emergencies/coronavirus-covid-19/news/news/2020/3/who-announces-covid-19-outbreak-a-pandemic

Biographies

Huanfa Chen is a Lecturer in Spatial Data Science and Quantitative Methods in the Centre for Advanced Spatial Analysis (CASA) at University College London. His research interests span spatial optimisation, agent-based modelling, and applied machine learning, with applications in transport, crime, and public health.
Yanjia Cao is a postdoctoral scholar at school of public health in University of California San Diego. She holds PhD in Geography from University of Maryland. Her research interest focuses on health geography, spatial data science and geographic information retrieval.
Lingru Feng obtained MSc in Spatial Data Science and Visualisation in 2021 from Centre for Advanced Spatial Analysis, University College London. Her research interests span geographic information science and spatial analysis, with applications in public health and urban planning.
Qunshan Zhao is a Lecturer in Urban Analytics in the Urban Big Data Centre at University of Glasgow. His research interests span urban analytics, spatial optimization/statistics, remote sensing, and geographic information science, with applications in transportation, housing, public health, and environmental monitoring.