Predicting future spatial patterns in COVID-19 booster vaccine uptake.

A.J. Wood[1], A.M. MacKintosh[3], M. Stead[3], and R.R. Kao*[1,2]

[1]*Roslin Institute, University of Edinburgh*
[2]*School of Physics and Astronomy, University of Edinburgh*
[3]*Institute for Social Marketing and Health, University of Stirling*

September 3, 2022

Abstract

Vaccination is a critical tool for controlling infectious diseases, with its use to protect against COVID-19 being a prime example. Where a disease is highly transmissible, even a small proportion of unvaccinated individuals can have substantial implications for disease burdens and compromise efforts for control. As socio-demographic factors such as deprivation and ethnicity have been shown to influence uptake rates, identifying how vaccine uptake varies with socio-demographic indicators is a critical step for reducing vaccine hesitancy and issues of access, and identifying plausible future uptake patterns.

Here, we analyse the numbers of COVID-19 vaccinations subdivided by age, gender, date, dose and geographical location. We use publicly available socio-demographic data, and use random forest models to capture patterns of uptake at high spatial resolution, with systematic variation restricted to fine spatial scale ($\sim$1km in urban areas). We show that uptake of first vaccine booster doses in Scotland can be used to predict with high precision the distribution of second booster doses across deprivation deciles, age and gender despite the substantially lower uptake of second boosters compared to first.

This analysis shows that while age and gender have the greatest impact on the model fit, there is a substantial influence of several deprivation factors and the proportion of BAME residents. The high correlation amongst these factors also suggests that, should vaccine uptake decrease, the impact of deprivation is likely to increase, furthering the disproportionate impact of COVID-19 on individuals living in highly deprived areas. As our analysis is based solely on publicly available socio-demographic data and readily recorded vaccination uptake figures, it would be easily adaptable to analysing vaccination uptake data from countries where data recording is similar, and for aiding vaccination campaigns against other infectious diseases.

1 Introduction

Vaccine hesitancy is a critical problem that severely impacts our ability to control important infectious diseases such as measles and seasonal influenza, and has been the subject of

*Corresponding author: rowland.kao@ed.ac.uk

much scrutiny during the COVID-19 pandemic. In a voluntary campaign, uptake will depend on individual decision-making and this is known to be influenced by sociological and demographic factors. Quantifying these factors can be challenging but is invaluable for understanding the context of individual decisions and developing strategies to improve uptake rates.

Scotland's COVID-19 vaccination programme began in December 2020 and as of August 2022 has delivered over 12.8 million doses to a population of 5.5 million, with 95% of the population aged over 20 receiving at least one dose. A primary course of two doses was followed by two subsequent rounds of booster vaccinations in Autumn/Winter 2021 (first aimed primarily at those aged over 50 [1], but rapidly expanded to all adults in response to a wave of the *Omicron* VOC in November 2021 [2]), and Spring 2022 (mainly to those aged 75+). While the initial acute threat presented by COVID-19 on healthcare systems in Scotland has subsided, waning immunity [3] and the continued circulation of infection amongst the population opens the question of how incentives aimed at increasing uptake in possible future rounds of vaccination should be targeted.

Given uncertainty in the future trajectory of the epidemic as well as vaccine accessibility, there are considerable challenges to estimating how uptake may evolve in future rollouts. Furthermore, an individual's attitude towards accepting follow-up booster doses may differ from their attitude towards the initial course of vaccination or first booster, especially as the perceived threat of COVID-19 changes over time. For example, surveys on hesitancy have thus far highlighted reasons for *accepting* a first dose that are specific to the context of the pandemic at the time of the initial rollout. These include, for example, a desire for life to get "back to normal", a feeling of moral duty, and a worry of a potential requirement of vaccination for travel [4, 5, 6, 7, 8]. Few or no legally binding government restrictions have been in place in Scotland since August 2021, and with this different context it is uncertain how uptake will settle in the longer term. Hesitancy has varied over the course of the epidemic [9, 10, 11] and in Scotland uptake has fallen with each successive dose. One may then reasonably expect uptake to fall further in the future, barring a radical change in virus virulence or transmissibility. The timing of this work coincides with a further round of booster doses in Scotland beginning in Autumn 2022, with all aged 50 and above eligible, as well as those otherwise considered vulnerable to severe disease [12, 13]. These eligibility criteria may become standard for booster doses in the longer term, being broadly consistent with those for the annual influenza vaccination in the UK [14].

Our aims are therefore twofold. First, to describe for previous booster campaigns, the spatial variation in uptake using known risk factors for vaccine hesitancy, refusal and access, thereby quantifying the importance of these risk factors in a geographically explicit context. Second, to use this analysis to *predict* [15] further spatial variation, should uptake continue to decrease as it has with each successive dose thus far.

2 Data

Vaccination data are provided by Public Health Scotland's electronic Data Research and Innovation Service (eDRIS). For each dose administered, the data specify a date of administration, a pseudonymised patient ID, dose number, patient sex, age range (in a five-year window, from 0–4 up to 75+), and residing *datazone* (DZ). DZs are non-overlapping census areas of order 500–1,000 individuals, each with an area as low as 0.15–0.4km^2 for densely populated areas. There are 6,976 DZs in total, covering the full area and population of

Scotland.

Population denominators used to evaluate overall uptake are taken from the 2021 *Small-Area Population Estimates* (SAPE) [16]. This is a DZ-level estimate of population in mid 2021, incorporating 2011 census data, subsequent births and deaths, and net immigration. As Scotland's most recently completed census was in 2011, there is likely to be considerable uncertainty in the population at a local level. Data on population breakdown by ethnicity are also obtained from census data.

Measures of deprivation are taken from the *Scottish Index of Multiple Deprivation* (SIMD) dataset [17]. This contains measures of different indicators of deprivation at DZ level (e.g., the percentage of residents living in overcrowded housing). The SIMD also ranks DZs by deprivation in each of *access*, *income*, *employment*, *education*, *health*, *crime* and *housing*. These ranks are derived from a weighted average of individual deprivation measures. A DZ with rank 1 would be considered to have the highest level of deprivation, and a DZ with rank 6,976 (out of 6,976) the lowest. An overall deprivation rank and decile are also given, from a weighted average of all measures.

COVID-19 vaccination in Scotland begins with administration of a first primary dose, and a second primary dose from eight weeks after. Three months from this initial course, adults then become eligible for a booster dose, and follow-up booster doses from then onwards, if available. In our analysis, we distinguish between two different characterisations of booster uptake:

- *Overall* uptake is the proportion of individuals to have received a booster vaccination. The denominator here is the population estimate. As this is often prone to substantial error, it is possible for the number of doses to exceed the population estimate.

- *Returning* uptake is the proportion of individuals *that have received at least one dose*, to have returned for a booster dose. The denominator is the number of individuals to have received at least one dose, derived from the eDRIS dataset.

The product of returning booster uptake and overall first dose uptake is then the overall booster uptake. Our model is fit to returning uptake (removing population uncertainty from the outcome variable), but we report in terms of overall uptake where appropriate.

Our analysis considers uptake amongst those aged 20 or over. While present eligibility is broadly the same for all aged between 16 and 50 (a two-dose course and a single booster), the 15–19 bracket (per eDRIS data) includes those aged 15 who are not generally eligible for a booster, thus the range is excluded. Finally, a small fraction of individuals with severely weakened immune systems are eligible for a third (and since, fourth) primary dose of COVID-19 vaccine, on top of a booster [18]. To simplify matters, we define booster uptake for all individuals as uptake of the third available dose, of any type.

3 Distribution of booster uptake, model fit

Booster uptake is summarised in Fig. 1, with respect to age, sex, and deprivation decile. In ages 20+, nationwide overall booster uptake was 78.2% (returning uptake 82.7%). Uptake falls in younger subpopulations in more deprived DZs, as well as greater decline in men compared to women. The oldest groups sustain high uptake at all deprivation indices, with a smaller decline with increasing deprivation. Fig. 1 also highlights the underlying uncertainty in the population estimate, where the trends in overall uptake (that rely on an estimated population denominator) are less consistent compared to returning uptake

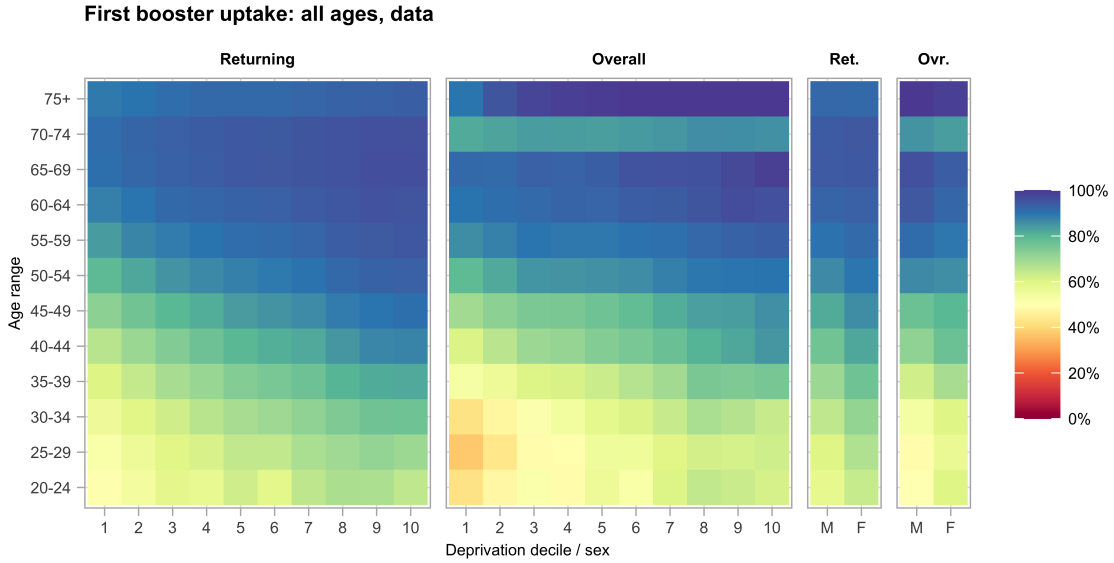

Figure 1: Booster vaccination uptake with respect to age range, sex and DZ deprivation decile. Decile 1 contains the most deprived DZs, and decile 10 the least deprived. Values of "overall" uptake have greater variability, owing to uncertainty in the population denominator.

(where the denominator of first doses administered is known). There is an anomalous fall in overall uptake across the 70–74 range, compared to 65–69 and 75+. This is not seen in returning uptake, thus this drop in overall uptake is likely an anomaly stemming from an inaccurate population denominator.

The 6,976 DZs, 12 age ranges, and 2 sexes divide the population into 167,424 subpopulations. We term these *cohorts*, with each containing of order 0–50 individuals. A random forest regression is then fit to cohort-level returning booster uptake, informed by:

- age range (5 year window, up to 75+);
- sex;
- ethnicity (% population belonging to a black or minority ethnicity), and;
- DZ-level deprivation ranks, by *access*, *income*, *employment*, *education*, *health*, *crime* and *housing*.

To keep the model simple, as well as suitably defined for generating scenarios with lower uptake (later in Section 4), we use DZ-level ranks rather than individual measures of deprivation. Details of model hyperparameters and performance are provided in Appendix A. Summarising, the fit explains 72.8% of cohort-level variation in returning uptake (fit: 74.9%, test: 64.5%), and 90.8% of DZ-level variation (fit: 92.4%, test: 83.6%) (Fig. 5B–C). There is evidence of clustering in residuals that falls on distance scales of order 5km (Fig. 6), suggesting variation in uptake from local factors, beyond the variables used to inform the model. The fit accurately captures observed trends with respect to age and deprivation (Fig. 5D).

Per the census data, fewer than 5% of Scotland's DZs have a black and minority ethnic population over 25%. The eDRIS vaccination data do not specify individual-level ethnicity so we can not infer a direct relation here, however other studies show uptake to be notably

lower in black and minority ethnic communities [19, 20, 21, 22]. Nonetheless, the inclusion of ethnicity (measured here as the percentage of a cohort belonging to a black or minority ethnicity) improves the model performance (Fig. 5A), with a high node purity indicative of a stronger influence on cohorts where the 20+ BAME population significantly deviates from the DZ mean of 7.3%.

4 Prediction of uptake distributions in a future rollout

Our regression model takes input data on population structure and deprivation, and is able to accurately produce the spatial distribution of booster dose uptake. With this, we are free to modify variables in the input data set (from which the model generates an uptake distribution), to generate *counterfactual* scenarios. Taking these modified inputs, the model may output a different distribution of uptake, a different overall nationwide uptake, or both.

Anticipating that uptake may fall with future doses, our interest is in designing input scenarios that result in *lower* booster uptake overall, and probing the resultant uptake distributions. Under a hypothesis that the relationship between deprivation and uptake will persist in future rounds of boosters, we design counterfactual scenarios where the population structure is unchanged, but all DZs have a higher degree of deprivation than in reality. We propose a simple method of adjusting the input data in order to do this, by changing each deprivation rank by some "shift" Δ, where larger negative values of Δ correspond to a higher degree of deprivation. We emphasise that Δ is an abstract parameter without a physical analogue, however it allows for a univariate deprivation adjustment to the input data, to counterfactual scenarios where all DZs have higher deprivation than they do in reality, whilst retaining the rank order.

Our methodology is detailed in Appendix B. Briefly, for each cohort, we obtain modelled uptake values for a range of values of Δ. We then fit a sigmoid curve to these modelled values, as a function of Δ. By taking different values of Δ, we then build uptake distributions (Fig. 7), in counterfactual scenarios where deprivation ranks fall below the "floor" rank of 1. This is equivalent to a hypothesis that cohorts particularly sensitive to small changes in deprivation will be at greater risk of putative further falls in uptake. Thus these cohorts are then suffer a greater projected fall in uptake compared to other cohorts, as nationwide uptake decreases.

4.1 Comparison of model in 75+ to Spring 2022 rollout

This hypothesis underpins our modelled distributions. To test it, we first compare to *observed* uptake of the *second* booster dose [23] rolled out in Spring 2022, which mainly targeted the over 75 age category. These data are limited to a single age group, but provide an initial opportunity to compare outputs from our model against observed trends.

Per data up to 18 August 2022, 75+ second booster uptake had plateaued, with returning uptake at 81.6%. This is down from 92.0% returning in the first booster round. Fig. 2 shows returning uptake in the least deprived decile fell from 93.7% to 86.4% (-7.3%), and in the most deprived decile from 89.1% to 74.2% (-14.9%). Uptake therefore started at a lower baseline value in more deprived deciles, but also fell more sharply in the second round, exacerbating the inequality in uptake with respect to deprivation. Differences in the predicted and actual uptake are increased at the lowest decile where the requirement to extrapolate individual deprivation rankings is most common. However, these differences

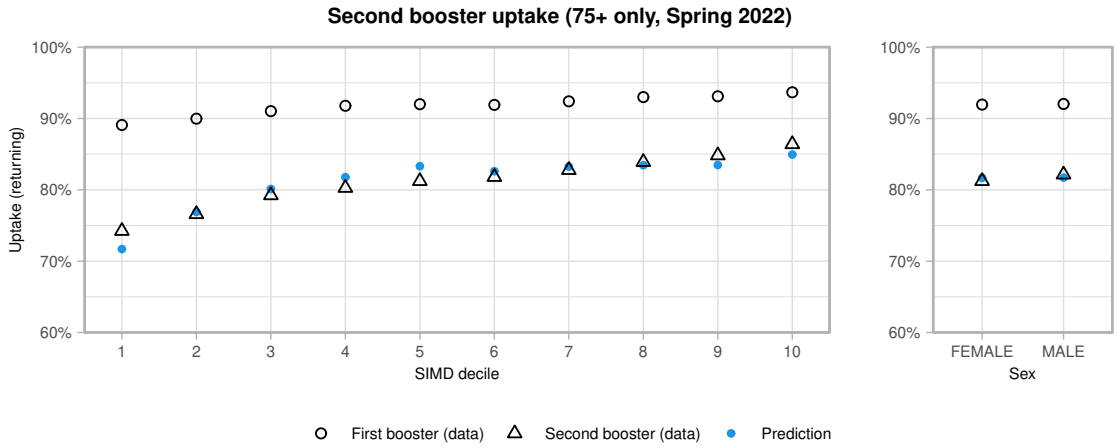

Figure 2: Returning uptake in the second booster rollout, as compared to first booster doses. Filled, blue points are model predictions from the first booster rollout, at the uptake achieved in the second booster round. Note that populations in each decile are not necessarily equal, where less deprived DZs generally skew older.

remain small, and the model captures the observed trend of uptake becoming increasingly skewed by factors relating to deprivation.

4.2 Predicted uptake distributions amongst ages 50+

In anticipation of a nationwide programme of boosters in Scotland beginning in Autumn 2022, we move on to estimate distributions of uptake across the full 50+ population. Fig. 3 shows projected distributions by age range, sex and deprivation, for overall nationwide uptake amongst ages 50+ falling from the baseline of 92.0% from the first booster rollout, to as low as 50%. At these lower uptakes, the estimated distributions echo the observed trend in Fig. 2 but now across all age groups. As nationwide uptake falls, uptake declines fastest in younger, more deprived groups, whereas it is sustained at higher levels in the oldest, least deprived groups. Fig. 4 shows that spatial clustering of communities with respect to deprivation (A) results in clusters of low overall booster uptake (B), which continues in the projected distributions (C–F) as nationwide uptake decreases.

5 Discussion

We have used high-resolution data to describe patterns in COVID-19 vaccination uptake across communities in Scotland. To understand fine-scale differences in booster uptake (of those returning after a first dose), we have fit a random forest regression model, informed by local population structure and deprivation. This explains substantial spatial variation in returning booster uptake (Fig. 5C). For a given DZ, booster uptake can then be estimated with high accuracy from its population structure and relative level of deprivation alone, without any information on location or neighbouring communities.

Data on first booster vaccinations reveal falls in uptake as a cohort becomes younger, and the level of deprivation increases. These are accurately captured by our model (Fig. 5D). There is a risk of uptake falling further in the longer term as the context of the pandemic evolves, and we have used our model to estimate longer-term distributions in uptake un-

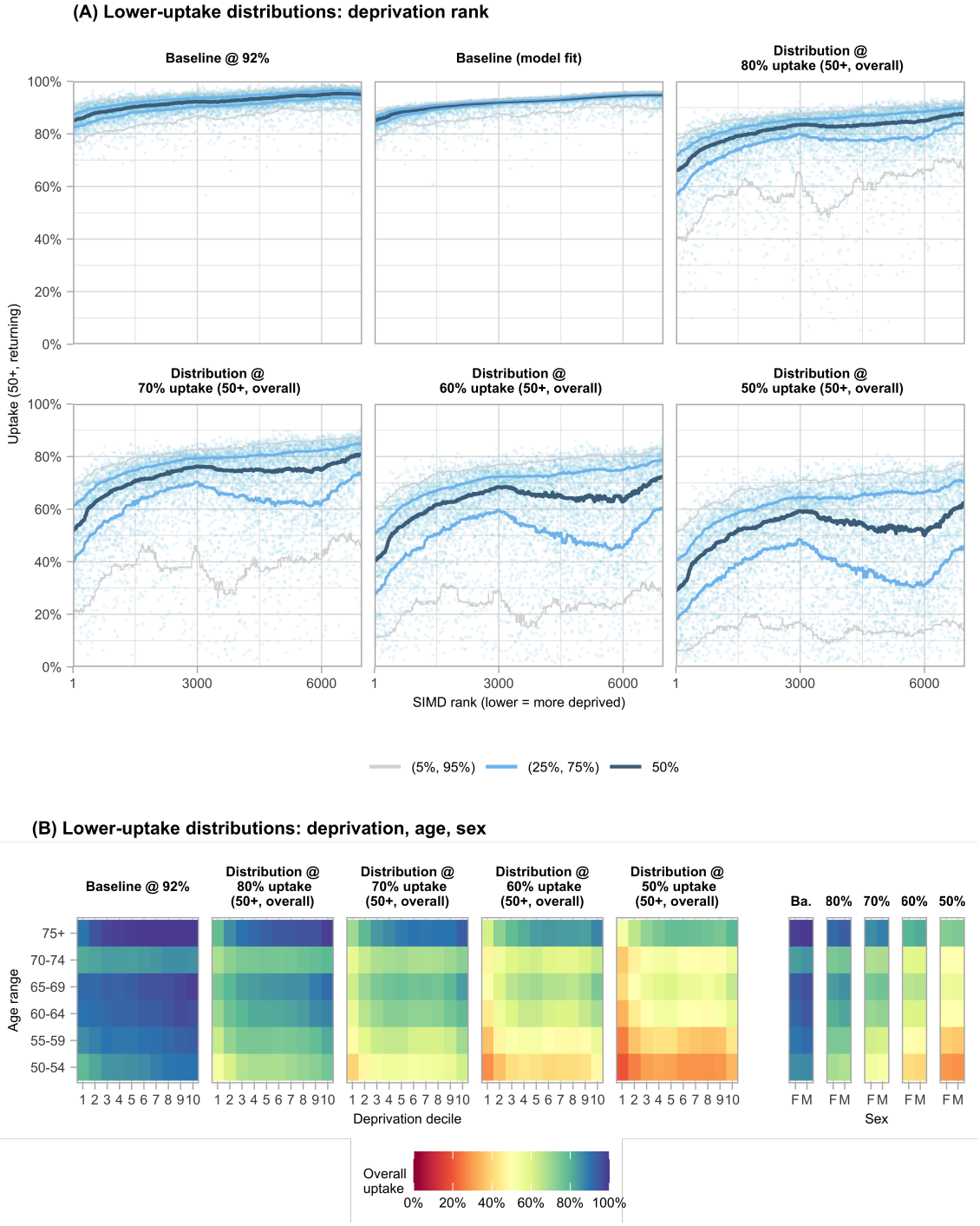

Figure 3: Model distribution of uptake in ages 50+ with declining nationwide uptake, across (A) DZs ordered by overall deprivation rank, tracking associated quantiles, and (B) with respect to deprivation decile, age and sex. These are in comparison to the "baseline" of the first booster rollout starting in Autumn 2021, where overall uptake in all aged 50+ is 92.0%.

der the assumption that nationwide uptake will continue to decline, and the trends across deprivation will persist. We fed into our model counterfactual data where each cohort has a higher degree of deprivation than in reality, and fit a relation between estimated

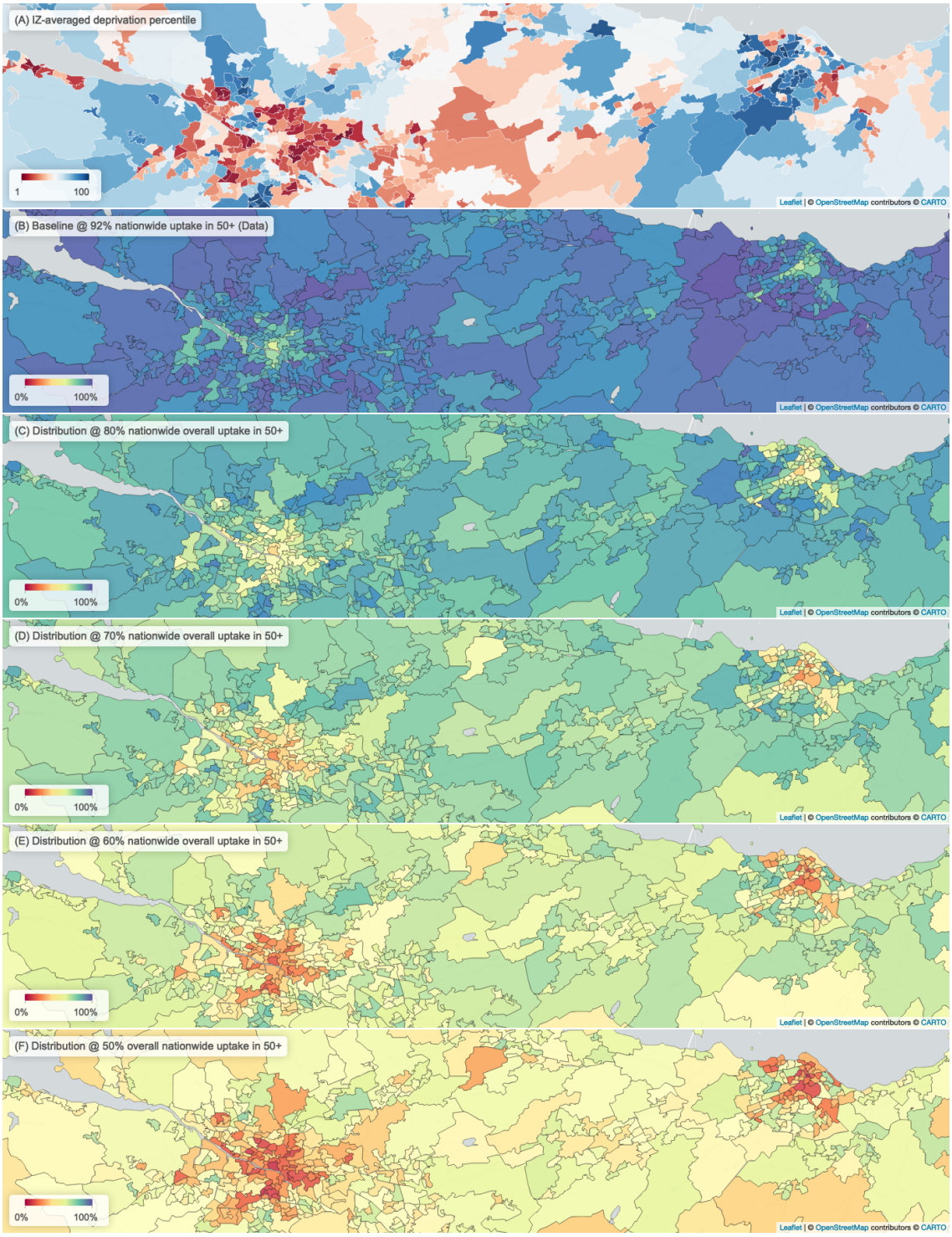

Figure 4: Map view of deprivation, overall 50+ first booster uptake, and projected 50+ uptake across the *central belt* of Scotland, containing the cities of Glasgow (left cluster) and Edinburgh (right cluster). These are presented at *intermediate zone* level; geographical areas typically containing of 4–6 DZs. Clusters of higher deprivation (A) coincide with clusters of low overall first booster uptake (B). In the projected scenarios (C–F) from 80% down to 50% nationwide uptake, the clustering of areas with low and high uptake persists.

uptake and changing deprivation. Cohorts more sensitive to small changes in deprivation are interpreted as at risk of steeper falls in uptake, relative to others. These predictions (Fig. 3, 4C–F) are initialised at the baseline of observed first booster uptake, and may serve as plausible distributions for future uptake, if nationwide uptake were to fall further. Model predictions for those aged 75 and over align with *second* booster uptake in the Spring 2022 rollout (Fig. 2, that the model was not informed by), with dose-to-dose uptake falling at the highest absolute rates in the most deprived communities where uptake was already the lowest. We use here *prediction* in the sense of prediction forecasting, i.e. to provide estimates of what might actually occur in a particular time and place, under the assumption that the underlying conditions of the prediction do not change [15]. With further rounds of booster doses planned, future analyses could refine the function used to extrapolate the impact of deprivation at the lower deciles. While trends may have shifted between the rollouts of the first and second booster campaigns, our analysis suggests that any such changes were insufficient to change the overall pattern of uptake and that drivers for low uptake likely remained very similar, at least in the age groups to which the second boosters applied. Such trends may not necessarily hold in the future or in other age groups, especially as the urgency of COVID-19 control subsides. Should they change, similar analyses can provide useful counterfactuals, invaluable in improving our understanding of why uptake trends may change in the future.

In Scotland, overall first dose uptake amongst those aged 20+ has plateaued at 95%, with second (primary) dose uptake at 90%, and first booster dose uptake at 78%. Regular future rounds of vaccination are likely for those most vulnerable to severe disease, with early studies suggesting that protection wanes on timescales of order six months [24, 25]. With these falls in uptake, there is concern about future immunity. These analyses confirm an observed inequality of increased vaccine refusal in younger groups living in more deprived communities [26, 27, 19]. Our analyses illustrate how these inequalities in uptake manifest as larger clusters of low uptake consistent with clustering of communities with a higher degree of deprivation, that may become more polarised as nationwide uptake falls.

Throughout the pandemic (including before vaccine availability), age and deprivation have been significant risk factors for poor COVID-19 outcomes [28, 29, 30, 31]. Given these risk factors, we highlight a potential *twofold* risk in the longer term, even with a modest drop in nationwide uptake. A fall in uptake will clearly reduce the population with vaccine-induced protection against severe COVID-19 disease. Compounding this, however, our analyses suggest that any shrinking pool of protection may become increasingly with those living in less deprived neighbourhoods, already at lower prior risk of developing severe disease. Our model predicts that a fall in 50+ nationwide uptake from 92% to 80% would see a mean 9% fall in uptake (96% to 87%) in the least deprived decile of DZs. In the most deprived decile, the mean fall is 19% (87% to 68%). This larger fall in uptake from a lower baseline would see clusters of communities that are most vulnerable disproportionately more exposed, and at higher risk of infection spread, hospitalisation and mortality.

Vaccination uptake data, while routinely collected, is not often made available at such a fine spatial resolution to researchers. Our access to these data was exceptional, owing to the need for rapid, policy relevant analysis during the COVID-19 pandemic [32]. Our model otherwise has been entirely informed using public data, using methods widely applicable to diseases beyond COVID-19, and to other indicators for vaccine hesitancy and accessibility beyond deprivation. With the wealth of public demographic data available, our results provide a strong argument for the value of access to such vaccine uptake data for COVID-19 as well as other infectious diseases where there is a strong public health interest.

6 Author contributions

R.R.K. conceived the project. A.J.W. wrote the model code, performed the model analysis, and with R.R.K. wrote the manuscript. A.M.M. and M.S. motivated conceptual work on evolving motives for vaccination, and risk factors for declining uptake. R.R.K. and A.J.W. conceived the technique for generating lower-uptake scenarios. All authors commented on and approved the manuscript.

7 Competing interests

The authors declare no competing interests.

8 Funding Statement

This work has been funded by the ESRC grant ES/W001489/1: *Real-time monitoring and predictive modelling of the impact of human behaviour and vaccine characteristics on COVID-19 vaccination in Scotland.*

References

[1] Cabinet Secretary for Health and Social Care. Scotland's autumn/ winter vaccination strategy 2021. `https://www.gov.scot/binaries/content/documents/govscot/publications/strategy-plan/2021/09/scotlands-autumn-winter-vaccination-strategy-20212/documents/scotlands-autumn-winter-vaccination-strategy/scotlands-autumn-winter-vaccination-strategy/govscot%3Adocument/scotlands-autumn-winter-vaccination-strategy.pdf`.

[2] Department of Health and Social Care. Jcvi advice on the uk vaccine response to the omicron variant. `https://www.gov.uk/government/publications/uk-vaccine-response-to-the-omicron-variant-jcvi-advice/jcvi-advice-on-the-uk-vaccine-response-to-the-omicron-variant/`.

[3] Nick Andrews, Elise Tessier, Julia Stowe, Charlotte Gower, Freja Kirsebom, Ruth Simmons, Eileen Gallagher, Simon Thelwall, Natalie Groves, Gavin Dabrera, et al. Duration of protection against mild and severe disease by covid-19 vaccines. *New England Journal of Medicine*, 386(4):340–350, 2022.

[4] Martine Stead, Anne Marie MacKintosh, Allison Ford, Douglas Eadie, Kathryn Angus, Michael Ussher, Kate Hunt, Curtis Jessop, Andy MacGregor, Hanna Biggs, Claire Elliott, and Helen Bedford. Latest insights from optimum: Young adults and vaccine acceptance, why people change their minds about vaccination, and what this means for future vaccination. *In prep.*

[5] Rhiannon Phillips, David Gillespie, Britt Hallingberg, Jennifer Evans, Khadijeh Taiyari, Anna Torrens-Burton, Rebecca Cannings-John, Denitza Williams, Elizabeth Sheils, Pauline Ashfield-Watt, et al. Perceived threat of covid-19, attitudes towards vaccination, and vaccine hesitancy: A prospective longitudinal study in the uk. *British Journal of Health Psychology*, 2022.

[6] Hannah Brindle, Nina T Rogers, Rosalind M Eggo, Luisa Enria, Shelley Lees, et al. Vaccine confidence and hesitancy at the start of covid-19 vaccine deployment in the uk: An embedded mixed-methods study. *Frontiers in public health*, 9, 2021.

[7] Bridget Lockyer, Rachael H Moss, Charlotte Endacott, Shahid Islam, Laura Sheard, et al. Compliant citizens, defiant rebels or neither? exploring changing covid-19 vaccine attitudes and decisions in bradford, uk: Findings from a follow-up qualitative study. *medRxiv*, 2022.

[8] Martine Stead, Allison Ford, Douglas Eadie, Hannah Biggs, Claire Elliott, Michael Ussher, Helen Bedford, Kathryn Angus, Kate Hunt, Anne Marie MacKintosh, et al. A "step too far" or "perfect sense"? a qualitative study of british adults' views on mandating covid-19 vaccination and vaccine passports. *Vaccine*, 2022.

[9] Philip Hyland, Frédérique Vallières, Mark Shevlin, Richard P Bentall, Ryan McKay, Todd K Hartman, Orla McBride, and Jamie Murphy. Resistance to covid-19 vaccination has increased in ireland and the united kingdom during the pandemic. *Public Health*, 195:54–56, 2021.

[10] Philip Hyland, Frédérique Vallières, Todd K Hartman, Ryan McKay, Sarah Butter, Richard P Bentall, Orla McBride, Mark Shevlin, Kate Bennett, Liam Mason, et al. Detecting and describing stability and change in covid-19 vaccine receptibility in the united kingdom and ireland. *Plos one*, 16(11):e0258871, 2021.

[11] Elise Paul and Daisy Fancourt. Predictors of uncertainty and unwillingness to receive the covid-19 booster vaccine: An observational study of 22,139 fully vaccinated adults in the uk. *The Lancet Regional Health-Europe*, 14:100317, 2022.

[12] Department of Health and Social Care. Joint committee on vaccination and immunisation (jcvi) updated statement on the covid-19 vaccination programme for autumn 2022. `https://www.gov.uk/government/publications/jcvi-updated-statement-on-the-covid-19-vaccination-programme-for-autumn-2022/joint-committee-on-vaccination-and-immunisation-jcvi-updated-statement-on-the-covid-1`

[13] Scottish Government. Autumn/winter covid-19 booster programme. `https://www.gov.scot/news/autumnwinter-covid-19-booster-programme/`.

[14] Richard G Pebody, Fiona Warburton, Nick Andrews, Mary Sinnathamby, Ivelina Yonova, Arlene Reynolds, Chris Robertson, Simon Cottrell, Muhammad Sartaj, Rory Gunson, et al. Uptake and effectiveness of influenza vaccine in those aged 65 years and older in the united kingdom, influenza seasons 2010/11 to 2016/17. *Eurosurveillance*, 23(39):1800092, 2018.

[15] Eric Winsberg and Stephanie Harvard. Purposes and duties in scientific modelling. *J Epidemiol Community Health*, 76(5):512–517, 2022.

[16] National Records of Scotland. Mid-2021 small area population estimates for 2011 data zones. `https://www.nrscotland.gov.uk/statistics-and-data/statistics/statistics-by-theme/population/population-estimates/small-area-population-estimates-2011-data-zone-based/mid-2021/`.

[17] Scottish Government. Scottish index of multiple deprivation 2020. https://www.gov.scot/publications/scottish-index-of-multiple-deprivation-2020v2-indicator-data/.

[18] Joint Committee on Vaccination and Immunisation. Joint committee on vaccination and immunisation (jcvi) advice on third primary dose vaccination. https://www.gov.uk/government/publications/third-primary-covid-19-vaccine-dose-for-people-who-are-immunosuppressed-jcvi-advice/joint-committee-on-vaccination-and-immunisation-jcvi-advice-on-third-primary-dose-vac

[19] Office for National Statistics. Coronavirus and vaccination rates in people aged 18 years and over by socio-demographic characteristic and occupation, england: 8 december 2020 to 31 december 2021. https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/healthinequalities/bulletins/coronavirusandvaccinationratesinpeopleaged18yearsandoverbysociodemographiccharacterist 8december2020to31december2021.

[20] Mohammad S Razai, Tasnime Osama, Douglas GJ McKechnie, and Azeem Majeed. Covid-19 vaccine hesitancy among ethnic minority groups, 2021.

[21] Long H Nguyen, Amit D Joshi, David A Drew, Jordi Merino, Wenjie Ma, Chun-Han Lo, Sohee Kwon, Kai Wang, Mark S Graham, Lorenzo Polidori, et al. Self-reported covid-19 vaccine hesitancy and uptake among participants from different racial and ethnic groups in the united states and united kingdom. *Nature communications*, 13(1):1–9, 2022.

[22] Ruth Elizabeth Watkinson, Richard Williams, Stephanie Gillibrand, Caroline Sanders, and Matt Sutton. Ethnic inequalities in covid-19 vaccine uptake and comparison to seasonal influenza vaccine uptake in greater manchester, uk: A cohort study. *PLoS medicine*, 19(3):e1003932, 2022.

[23] Scottish Government. Spring rollout of boosters begins. https://www.gov.scot/news/spring-rollout-of-boosters-begins/.

[24] Nick Andrews, Elise Tessier, Julia Stowe, Charlotte Gower, Freja Kirsebom, Ruth Simmons, Eileen Gallagher, Meera Chand, Kevin Brown, Shamez Ladhani, et al. Vaccine effectiveness and duration of protection of comirnaty, vaxzevria and spikevax against mild and severe covid-19 in the uk. *medrxiv*, 2021.

[25] Einav G Levin, Yaniv Lustig, Carmit Cohen, Ronen Fluss, Victoria Indenbaum, Sharon Amit, Ram Doolman, Keren Asraf, Ella Mendelson, Arnona Ziv, et al. Waning immune humoral response to bnt162b2 covid-19 vaccine over 6 months. *New England Journal of Medicine*, 385(24):e84, 2021.

[26] Martine Stead, Curtis Jessop, Kathryn Angus, Helen Bedford, Michael Ussher, Allison Ford, Douglas Eadie, Andy MacGregor, Kate Hunt, and Anne Marie MacKintosh. National survey of attitudes towards and intentions to vaccinate against covid-19: implications for communications. *BMJ open*, 11(10):e055085, 2021.

[27] Ted Dolby, Katie Finning, Allan Baker, Leigh Fowler-Dowd, Kamlesh Khunti, Cameron Razieh, Thomas Yates, and Vahe Nafilyan. Monitoring sociodemographic inequality in covid-19 vaccination uptake in england: a national linked data study. *J Epidemiol Community Health*, 76(7):646–652, 2022.

[28] Anthony J Wood, Aeron R Sanchez, Paul R Bessell, and Rowland R Kao. Spatio-temporal characteristics of the sars-cov-2 omicron variant spread at fine geographical scales, and comparison to earlier variants. *medRxiv*, 2022.

[29] Nazir I Lone, Joanne McPeake, Neil I Stewart, Michael C Blayney, Robert Chan Seem, Lorraine Donaldson, Elaine Glass, Catriona Haddow, Ros Hall, Caroline Martin, et al. Influence of socioeconomic deprivation on interventions and outcomes for patients admitted with covid-19 to critical care units in scotland: a national cohort study. *The Lancet Regional Health-Europe*, 1:100005, 2021.

[30] Office for National Statistics. Deaths due to covid-19, registered in england and wales: 2021. `https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/articles/deathsregisteredduetocovid19/2021`.

[31] Aniruddh P Patel, Manish D Paranjpe, Nina P Kathiresan, Manuel A Rivas, and Amit V Khera. Race, socioeconomic deprivation, and hospitalization for covid-19 in english participants of a national biobank. *International Journal for Equity in Health*, 19(1):1–4, 2020.

[32] Scottish Government. Coronavirus (covid-19): modelling the epidemic (collection). `https://www.gov.scot/collections/coronavirus-covid-19-modelling-the-epidemic/`.

[33] Patrick AP Moran. Notes on continuous stochastic phenomena. *Biometrika*, 37(1/2):17–23, 1950.

A Model details

The random forest model was fit using the *RandomForest* package (version 4.6–14) in *R* (version 4.1.0). Model script is available at `https://git.ecdf.ed.ac.uk/awood310/covid-19-vaccination-analysis-and-projection`.

The model was fit with 800 trees, each with a maximum node size of 5,000 and 3 variables tried per split. These parameters served to maximise the cohort-level variation explained, in the data that the model were not informed with. Cohorts from 80% of randomly selected DZs are used to fit the model, with the remaining 20% reserved to test the model against data not trained on.

Fig. 5A shows explanatory variable importance outputs (node purity and accuracy loss), with age emerging as the dominant variable.

With regards to residuals, a map visualisation of DZ-to-DZ residuals (e.g., Fig. 6B with focus on central Glasgow, the largest city in Scotland by population and with substantial variation in DZ-to-DZ deprivation) shows evidence of clustering of residuals of the same sign, with broader areas where uptake was different from that expected by its population structure and relative deprivation, suggesting a local effect outwith these factors. Fig. 6B–C quantifies this correlation using the Moran's I statistic [33] (*ape* package, version 5.5),

showing spatial autocorrelation fall substantially over distances of 5–10km, or within 1–5 nearest neighbours (with nearest neighbours being DZs that share a border).

B Details of modelling lower-uptake uptake distributions

We use uptake in the first booster vaccination rollout, in conjunction with our random forest model informed with SIMD deprivation ranks, to create estimates of longer term distributions, with *lower* nationwide uptake.

Per Section 4, we create counterfactual *input* data by making a univariate shift of all deprivation ranks in dataset by some amount Δ. For larger negative values of Δ (thus the ranks being lower, and representative of higher levels of deprivation than in reality), projected uptake is anticipated to fall. A limitation to this, however, is when a counterfactual rank runs outside the range 1–6,976 (e.g., a rank 55 at shift $\Delta = -500$ will be "rank -445"). Random forest models, being a form of stepwise regression, generally perform poorly with data outside the range they are fit against. To address this and project beyond the "floor" rank of each DZ, then, we fit a sigmoid function to each cohort i, for the fit returning uptake U_i as a function of Δ

$$U_i(\Delta) = \frac{1}{1 - \mathrm{e}^{-a_i(\Delta - b_i)}} \ .$$

a_i and b_i are parameters to be fit for each cohort, based on the model fit values for the range of counterfactual Δ values that do not exceed the range 1–6,976. Finally, we shift each fit curve along the y-axis by an amount ϵ_i, such that $U_i(0)$ is the exact observed first booster uptake for cohort i. To avoid scenarios where this shift introduces a modelled uptake greater than 100% or below 0%, we bound $U_i(\Delta) = 0$ if $\left[1 - \mathrm{e}^{-a_i(\Delta - b_i)}\right]^{-1} - \epsilon_i \leq 0$, and 1 if $\left[1 - \mathrm{e}^{-a_i(\Delta - b_i)}\right]^{-1} - \epsilon_i \geq 1$. The $\Delta = 0$ distribution then exactly reproduces that observed for first booster uptake at nationwide uptake amongst those 50+ of 92.0%. We then reduce Δ to generate projected distributions at lower uptakes (cohort-level examples of this are given in Figure 7).

For a small number of DZs, the range of valid values for Δ is too narrow to reasonably fit a curve (i.e., instances where a DZ has one very high deprivation rank, and another that is very low). For cohorts where the range is lower than 500 ranks, then, we instead use the average value of parameters a_i and b_i from similar cohorts; having the same age range, sex, DZ deprivation *decile*, and similar first booster uptake.

The underlying random forest model is fit to returning uptake rather than overall uptake. We then calculate the corresponding overall uptake by multiplying relative uptake by overall first dose uptake. By making lower-uptake distributions by decreasing returning uptake, then, we make an implicit assumption that all future doses will be given to individuals that have received at least one vaccination before.

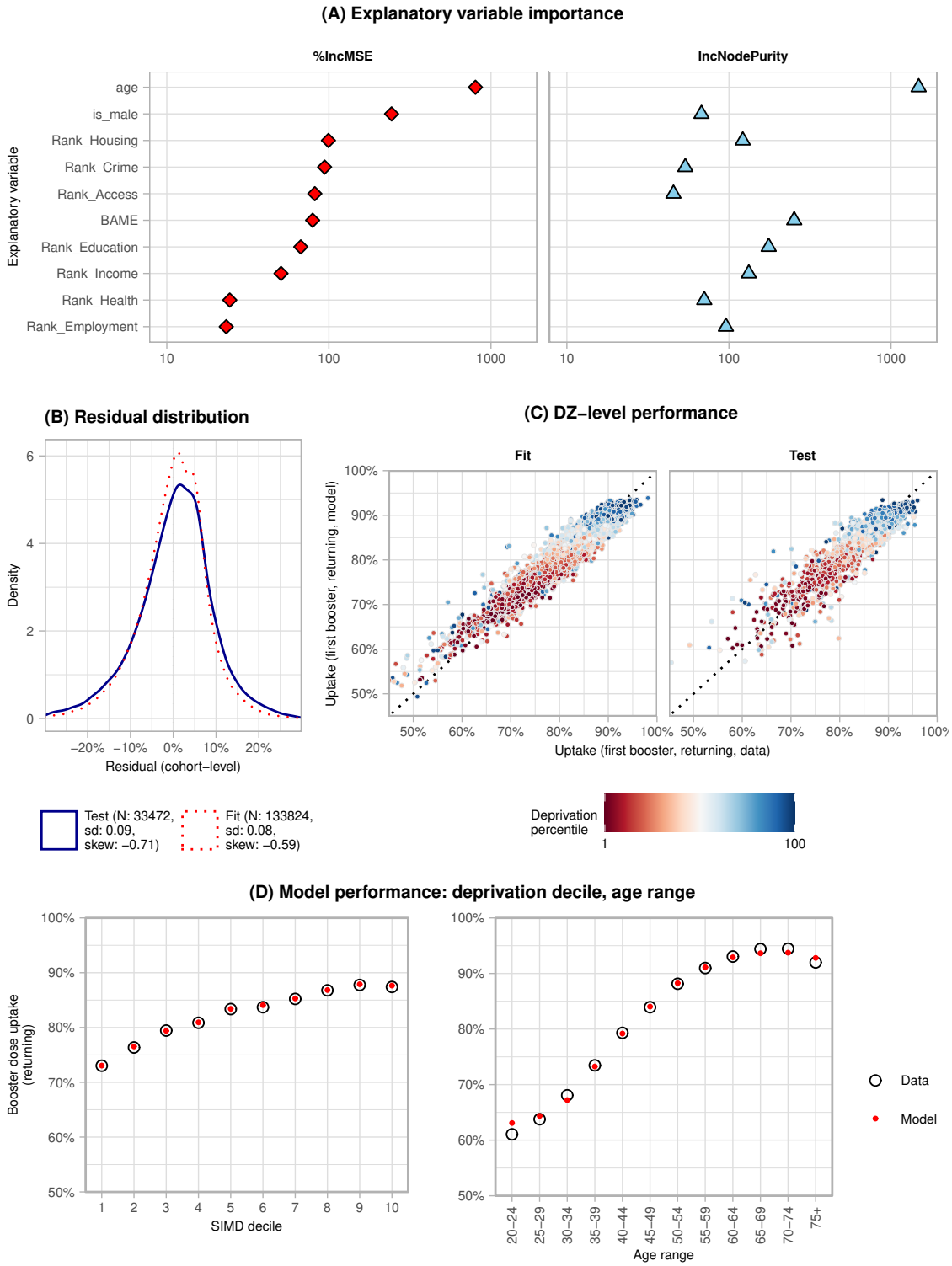

Figure 5: Model output. (A) Explanatory variable importance output from RF model (MSE loss and node purity). (B) Residual distributions of the fit and test data sets. (C) Performance comparing data and fit values for returning uptake at the DZ level, over individual DZs, with deprivation indicated. (D) Performance aggregating cohorts over deprivation decile, and age range.

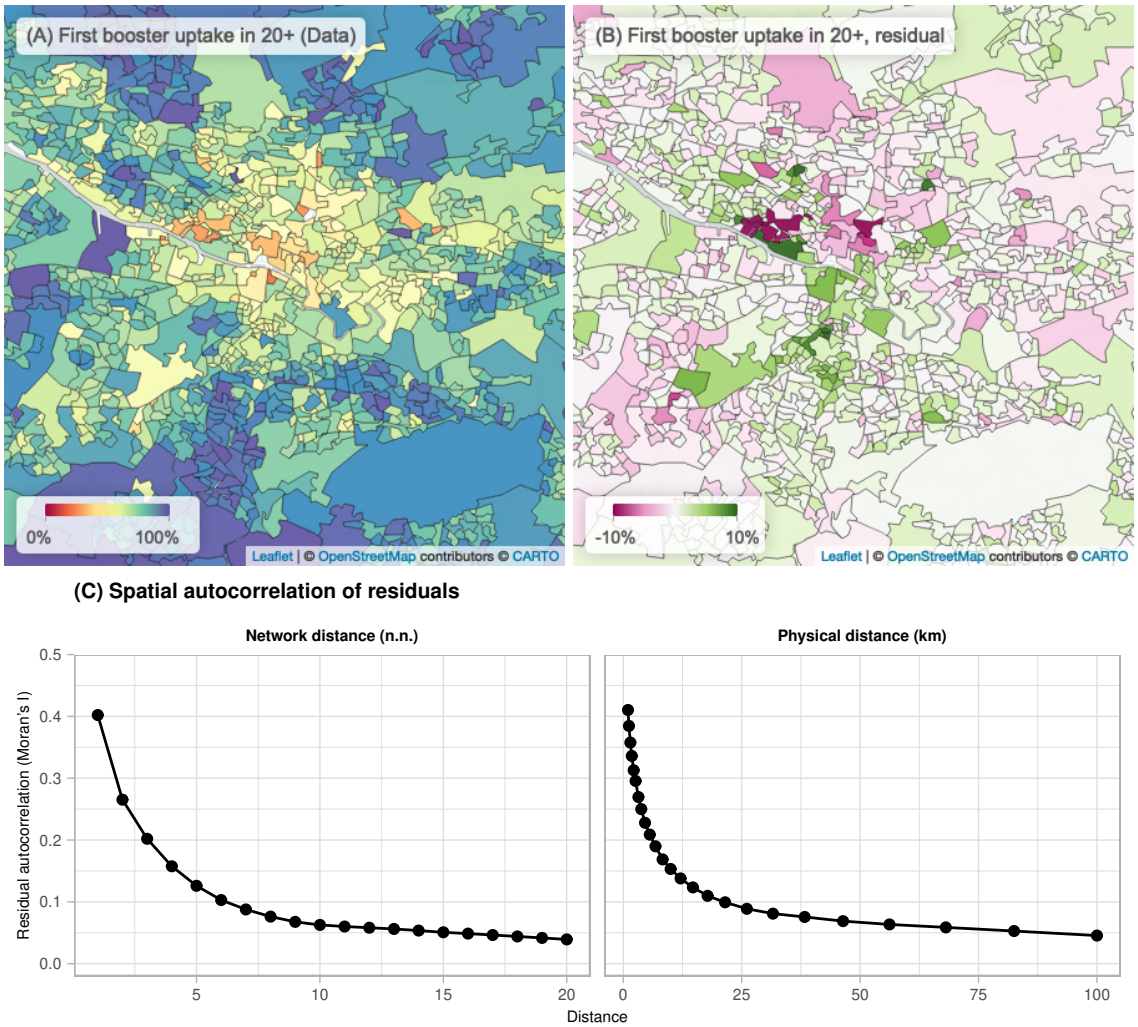

Figure 6: (A) overall booster uptake across an 18km × 18km area of central Glasgow, and (B) corresponding residuals (the difference between actual uptake, and the fit value). Green DZs are those where observed uptake is higher than the fit value, and pink DZs indicate where observed uptake was lower. (C) Residual autocorrelation as measured by the Moran's I statistic, comparing autocorrelation between residuals (y-axis) within a certain locus (x-axis).

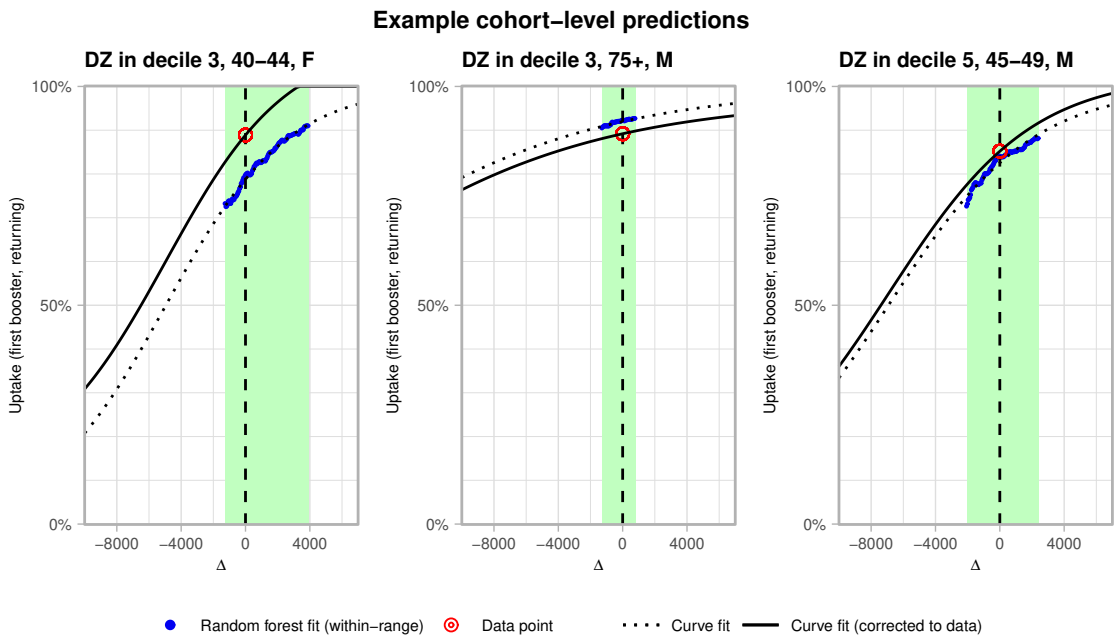

Figure 7: Graphical representation of prediction of uptake with respect to level of deprivation. For each cohort, the green box bounds the "floor" and "ceiling" values of deprivation rank shift Δ. Within this range, the projected uptake (blue points) falls for decreasing Δ (increasing level of deprivation). A sigmoid function (black, dotted) is fit to these fit values, which is then shifted to match the actual returning first booster uptake (red circle) at $\Delta = 0$ (vertical dashed line).