Socioeconomic and temporal heterogeneity in SARS-CoV-2 exposure and disease in England from May 2020 to February 2023

Authors: Christian Morgenstern[1, *], Thomas Rawson[1], Wes Hinsley[1], Pablo N. Perez Guzman[1], Samir Bhatt[1,2, †], Neil M. Ferguson[1, †]

Affiliations:

[1]MRC Centre for Global Infectious Disease Analysis, Jameel Institute, School of Public Health, Imperial College London, UK

[2]University of Copenhagen, Denmark

[†]Contributed equally

[*]Corresponding author: c.morgenstern@imperial.ac.uk

Abstract

The impact of COVID-19 varied significantly by deprivation, ethnicity, and policy measures. We analysed individual-level data on 12,310,485 first SARS-CoV-2 Pillar-2-PCR-confirmed infections, 439,083 hospitalisations, 107,823 deaths, and vaccination records in England from May 2020 to February 2022. Poisson regression models adjusted for demographic and temporal factors showed higher incidence rate ratios (IRRs) for severe outcomes in the most deprived areas compared to the least. We note higher IRRs for severe outcomes for all non-White relative to White ethnicities. The magnitude of IRRs for both deprivation and ethnicities declined from the wild-type to the omicron periods for severe outcomes. For infections, we observed IRRs above one for non-White ethnicities during the wild-type and alpha periods. Vaccination significantly reduced risks across all groups. For severe outcomes, pre-existing health inequalities led to large and persistent disparities. For infections, measures must be structured with ethnicity and deprivation in mind early in a pandemic.

Keywords: SARS-CoV-2, risk factors, deprivation, ethnicity, vaccine effectiveness.

Introduction

By December 2023, the COVID-19 pandemic had caused over 20.5 million confirmed cases and over 175,000 deaths in England **(*1, 2*)**. The pandemic did not impact individuals equally; individual-based studies have explored the links between deprivation, ethnicity, and other factors on health outcomes to varying levels **(*3–8*)**.

Past studies have explored heterogeneity in risk in England **(*9, 10*)** and Scotland **(*11*)** using the Index of Multiple Deprivation (IMD) **(*12*)** at the Lower-Tier Local Authority (LTLA) level. LTLAs are areas where local government provides services, and the UK Office of National Statistics (ONS) provides an extensive set of data, including population age distribution, ethnicity, and measures of deprivation. These measures vary substantially across LTLAs, as did the number of PCR-confirmed infections, hospitalisations, and deaths reported across England over the course of the pandemic (Figure 1). In this study, we aim to characterise this heterogeneity in COVID-19 outcomes and how it evolved during the pandemic **(*1, 13, 14*)**. Similar studies on the impact of socioeconomic heterogeneities on SARS-CoV-2 infections have been conducted in France **(*15*)** and Germany **(*16*)** and on outcomes in the USA **(*17*)**.

SARS-CoV-2 testing in England **(*18*)** had four "Pillars": Polymerase chain reaction (PCR) testing for health and care workers and individuals with clinical needs (Pillar 1), freely available PCR (and later antigen) testing for the general population (Pillar 2), serology (Pillar 3) and targeted surveillance (Pillar 4). Pillar 2 testing increased significantly from 10th May 2020, and most testing was halted on 27th February 2022 with the introduction of the 'Living with COVID' strategy **(*19*)** (see Methods and SI C.4).

We additionally investigate vaccine effectiveness (VE) in the post-vaccination period and the impact of vaccination on risk heterogeneity. The SARS-CoV-2 immunisation programme in England was one of the most rapid globally, with 89.9% of the adult population aged 20 or over receiving at least one dose and 86.6% receiving at least two doses by 27th February 2022 **(*20*)**. Test-negative case-control studies (TNCCs) were the standard epidemiological tool for evaluating SARS-CoV-2 VE during the pandemic **(*21–23*)**, though TNCCs can potentially suffer from biases **(*24*)**. The main alternative to TNCCs used for VE estimation during the COVID-19 pandemic is large population observational cohort studies that use individual-level healthcare and surveillance data linked to vaccination status **(*25–28*)**.

Last, we investigate how risk heterogeneity was modified by the public health measures ("restrictions") in force during the pandemic (see Methods and SI B.5) **(*1, 14*)**.

[Figure 1]

Results

We present the breakdown of the England population (56,427,863 people in England [Census 2021], 56,417,353 are in our synthetic population, and 56,344,410 are included in the analysis) in Table 1, including the number of first recorded PCR-confirmed SARS-CoV-2 infections (12,310,485) and associated hospitalisations (439,083) or deaths (107,823) linked to that first recorded infection from the week beginning 10th May 2020 (isoweek 18) to 27th February 2022 (isoweek 8) confirmed by a Pillar 2 PCR positive test. These 95 weeks cover the full period for which population-wide testing was available.

[Table 1]

We examined a total of 360 model variants, which varied by the covariates and interactions included (see SM A.2 for covariate definitions considered): age, definition of deprivation, regional disaggregation, ethnicity disaggregation, and which interaction terms to include in the model. SM fig. S2-S4 lists the top 88 selected models for each outcome of interest. Restriction levels are based on individual policy measures, such as school closure or gathering restrictions, at each time point (SM section A, C.3 and fig. S7-S9). We found that less granular age categories had better 10-fold cross-validation performance (under 40s, 10-year age bands up to 70, single group for 80+). Similarly, coarse spatial disaggregation (by NHS region of England) was favoured, and deprivation was quantified by IMD quintiles. We disaggregated ethnicity into six groups as the results were very similar for the range of ethnicity categorisations examined. The preferred model included covariates for sex, age, NHS region, vaccine status, IMD quintile, ethnicity and restriction level.

We compared the estimates from the preferred model against the data and found a good qualitative fit examining marginal incidence rate ratios (IRRs) (SM fig. S1).

We investigated the impact of overdispersion for each outcome of interest using negative binomial regression models. For Pillar 2 PCR confirmed cases, we found a moderate level of overdispersion (k=0.33) during the wild-type period. We found that either very high values of k (indicating marginal overdispersion) or the Poisson model were preferred for all other periods. For hospitalisation and death outcomes, the Poisson model was always preferred (SM section A and table S1).

Figure 2 presents the modelled IRRs for deprivation and ethnicity categories from the preferred model for each endpoint and by period defined by variant dominance. IRRs for sex, age and region are presented in SM fig. S11.

[Figure 2]

When we consider the entire period (May 2020-February 2022), the highest IMD quintile, corresponding to the most deprived LTLAs, has an IRR (relative to the least deprived IMD quintile) of 1.64 (95% CI: 1.60-1.67) for death as the outcome, 1.53 (95% CI: 1.51-1.55) for hospitalisation, 1.02

(95% CI: 1.02-1.02) for a positive Pillar 2 PCR test. For the Pillar 2 positive test outcome, we observed significant differences between periods defined by variant dominance. The biggest differences were observed for the Alpha variant with an IRR (for the highest quintile relative to the lowest) for that endpoint of 1.45 (95% CI: 1.45-1.46).

For deaths, over the entire study period, South Asian ethnicity had an IRR 1.63 (95% CI: 1.59-1.68), Asian (other) ethnicity an IRR 1.24 (95% CI: 1.17-1.32), and Black ethnicity an IRR 1.49 (95% CI: 1.43-1.54), with White ethnicity as the reference. For hospitalisation, over the entire study period, South Asian ethnicity had an IRR 1.42 (95% CI: 1.40-1.44), Asian (other) ethnicity an IRR 1.17 (95% CI: 1.15-1.20), and Black ethnicity an IRR 1.56 (95% CI: 1.53-1.58). Using Pillar 2 PCR positive tests as the endpoint, over the entire study period, South Asian ethnicity had an IRR 0.89 (95% CI: 0.88-0.89), Asian (other) ethnicity an IRR 0.78 (95% CI: 0.78-0.79), and Black ethnicity an IRR 0.89 (95% CI: 0.89-0.90). The IRRs for Mixed/Other ethnicity are reported in SM table S4.

Ethnicity differences declined for the Pillar 2 PCR positive test endpoint throughout the pandemic. IRR for Pillar 2 PCR positive cases declines for all non-White ethnicities (relative to White) over time, with $IRR_{WT} > IRR_{Alpha} > IRR_{Delta}$. During the Omicron period, non-White IRRs were broadly similar to the Delta period but typically slightly above IRR_{Delta}. For death, we observed a declining IRR over time for South Asian ethnicity, with IRR_{WT}, IRR_{Alpha}, and IRR_{Delta} exhibiting overlapping confidence intervals but $IRR_{Omicron}$ statistically significantly lower. For hospitalisation, we observed a declining IRR over time for South Asian ethnicity, with IRR_{WT}, IRR_{Alpha}, IRR_{Delta} and $IRR_{Omicron}$ statistically significantly lower than in the previous period. Other Asian ethnicities had high IRRs (overlapping confidence intervals) for the WT and Alpha periods and significantly lower IRRs for Delta and Omicron. Black ethnicities exhibited high IRRs up to Omicron in the range of 1.38-1.45 for death and for the entire period for hospitalisation 1.16-1.87.

Results for deprivation were consistent for severe outcomes (hospitalisation and death) throughout the pandemic, with IRRs monotonically increasing as deprivation increased, but with IRRs during the Omicron period lower than other periods. This result also applied to a more granular representation of deprivation, using deciles (SM fig. S13) and for the different subcomponents of the IMD measure (SM fig. S20-S26). For Pillar 2 PCR positive cases, we observed an increasing IRR with increasing deprivation for the WT and Alpha periods but not for the Delta and Omicron periods. This resulted in the IRR for the whole period being close to 1. This result was consistent with IMD deciles and the subcomponents of the IMD index.

[Figure 3]

Figure 3 presents our VE estimates (SM section B.3 and table S3) for our preferred model. We obtain a VE (defined as 1-IRR) estimate for mRNA vaccines ranging from 86.8% (over 18 weeks after the second dose, 95% CI: 86.2-87.3) to 97.1% (less than two weeks after the booster dose, 95% CI: 96.6-

97.6) for protection against death for individuals with at least two doses, 84.8% (over 18 weeks after second dose, 95% CI: 84.5-85.0) to 93.3% (2-10 weeks after booster dose, 95% CI: 93.1-93.5) against hospitalisation and 30.4% (over 18 weeks after booster dose, 95% CI: 29.3-31.5) to 84.5% (2-10 weeks after second dose, 95% CI: 84.4-84.6) against Pillar 2 PCR positive confirmed infections for the full study period.

For adenovirus-based vaccines, we estimated VE for individuals with at least two doses, ranging from 77.2% (over 18 weeks after second dose, 95% CI: 76.2-78) to 95.9% (less than two weeks after booster dose, 95% CI: 95.3-96.5) for protection against death, 78.5% (over 18 weeks after second dose, 95% CI: 78.2-78.8) to 94.1% (2-10 weeks after booster dose, 95% CI: 94.0-94.3) against hospitalisation, and 29.6% (over 18 weeks after booster dose, 95% CI: 27.6-31.5) to 68.5% (2-10 weeks after second dose, 95% CI: 68.3-68.6) against Pillar 2 PCR-positive confirmed infections for the full study period.

We found that VE was consistently lower for the Omicron period. There was little to no difference in VE estimates for severe outcomes for the Alpha and Delta time periods.

We present the estimated VE results for all variants, vaccine status, and vaccine types in SM table S3 with their associated 95% CIs. The confidence intervals are narrow due to the large number of events observed (Table 1B); confidence intervals are wider for the variant period-specific estimates. A comparison with previous vaccine effectiveness estimates is included in SM section D.2.

[Figure 4]

Figure 4 shows the estimated VE results obtained from an extension of the preferred model, which included an additional interaction term between vaccine status and IMD quintile. Panel (A) presents results for the period when Delta was the dominant variant, while panel (B) covers Omicron. For deaths and hospitalisations, we did not observe significant differences in VE between IMD quintiles, except for the period over 18 weeks post-second dose during the Delta phase for deaths, where there was a tendency for the least deprived areas (IMD 1 and 2) to have lower VE than the more deprived areas (IMD 4 and 5).

For PCR-confirmed infections, we found different results for the Delta and Omicron periods. During Delta, we typically observed the ordering of IMD quintiles was higher VE for less deprived areas than more deprived areas, particularly less than 18 weeks since the second dose. VE estimates cluster together for over 18 weeks since the second and booster doses. For Omicron, we observed little variability in VE by IMD for second and booster doses, but VE estimates were higher for less deprived areas. The exception is for over 18 weeks since the booster dose, for which we observed higher VE for more deprived areas. Results for IRRs of all covariates for the model with an interaction term between vaccine status and IMD quintile are presented in SM fig. S17 and for vaccine status and ethnicity in SM fig. S18.

[Figure 5]

Figure 5 shows the estimated IRRs for the restriction level covariate for our preferred model. The IRRs for restriction levels were robust to regional resolution, up to and including the LTLA level (SM fig. S15 and S16). We estimated the IRR for restriction level over the full study period of 1.14 (95% CI: 1.12-1.15) for deaths, 1.15 (95% CI: 1.14-1.16) for hospitalisation, and 1.17 (95% CI: 1.17-1.17) for Pillar 2 PCR positive cases. These estimates indicate that periods of high levels of public health restrictions were associated with higher levels of risk than periods where restriction levels were lower. The estimated IRR for restriction level for the WT and Alpha periods was lower for deaths and hospitalisation than for the full-time period, which includes longer time intervals with no restrictions. However, for Pillar 2 PCR-confirmed infections, we found that the IRR for restriction level was higher during the WT period and lower for the Alpha period than for the full-time period (Figure 5). The IRRs for the Delta and Omicron periods are one, as no substantial restrictions were in place over those periods.

Discussion

We considered the impact of sex, age, region, deprivation, level of restrictions, and ethnicity on the relative risk of a SARS-CoV-2 Pillar 2 PCR-positive first recorded infection, hospitalisation, or death. Our proposed model was selected to provide a good fit across all three outcomes of interest. Our results extend previous work on the impact of deprivation **(*5*)** and evidence risk heterogeneity by ethnicity in England **(*4, 6, 8*)** and globally **(*7, 17*)**.

We find that modelling overdispersion for Pillar 2 PCR positive cases was necessary for the wild-type period, consistent with Morrissey et al. **(*13*)**, and the value for the overdispersion we find aligns with other estimates outside China **(*29, 30*)**. We did not find overdispersion to be present for severe outcomes, and the Poisson model provides the best fit to the data. Even though Pillar 2 PCR tests were available widely, this indicates that some groups, defined by higher levels of deprivation or non-White ethnicity, may have accessed testing less, leading to the observed overdispersion. This was less likely to be the case for more severe outcomes as testing would be more likely conducted due to the more severe nature of the infection.

The differences in IRRs we observe by outcome (PCR-confirmed infection, hospitalisation, or death) indicate a difference between severe outcomes and infections. For severe outcomes, we observed a significant risk heterogeneity by age, with older age groups at higher risks than younger age groups at an exponential rate (SM fig. S11(B)) consistent with the variation of the infection fatality ratio by age in the pre-vaccination era **(*31*)**. Increased risk was also positively associated with higher levels of deprivation (Figure 3A), and the difference between the most deprived quintile of LTLAs and the least deprived quintile was significant for all periods considered. Similarly, we found that non-White

ethnicities were at increased risk.

For Pillar 2 PCR-positive cases, the picture is more nuanced. We observed that the differences in risk between age groups were of a similar magnitude as the large differences we observed for severe outcomes. These differences are consistent with past studies and may reflect age-specific self-protective behaviours **(*32*)** and/or variation in contact rates (SM fig. S6). For deprivation and ethnicity, risk heterogeneity was similar in trends as for more severe outcomes, but the magnitude of differences was less. The results for Pillar 2 PCR-positive cases were consistent with other studies **(*32, 33*)**.

For severe outcomes, we observed less variation in deprivation or ethnicity related risk differences across the course of the pandemic, pointing to existing health inequalities driving the strong association between deprivation, ethnicity, and risk of severe outcomes. This held consistently for all sub-components of the index of multiple deprivation, except for the Living Environment component, and particularly strongly for the Income, Employment, Health, and Education sub-components (SM fig. S20-S26). This observation is consistent with known socio-economic and ethnic disparities in the UK for type-2 diabetes (*34*), which is a known risk factor for severe SARS-CoV-2 infection outcomes. Interestingly, we observed a high IRR for South Asian ethnicity, for both severe outcomes and infections, which reduced over time (Figure 3.B), perhaps suggesting the metabolic component impacting outcomes being reduced for omicron and delta as individuals of South Asian ethnicity are at higher risk of metabolic disease (*35*).

We assessed vaccine effectiveness, and our results broadly align with previously published VE estimates for England by UKHSA(*21*) for similar periods and VE definitions. We find vaccine effectiveness for the mRNA vaccines is consistent with UKHSA estimates of 88.0% (95% CI: 85.3-90.1) for two doses of BNT162b2 for infections for persons with the Delta variant. Similarly, we find vaccine effectiveness for the adenovirus vaccines is in line with UKHSA estimates of 67.0% (95% CI: 61.3-71.8) for persons with two doses of the ChAdOx1 nCoV-19 vaccine. For the Omicron variant, estimates vary more widely(*36*) with VE against hospitalisation, ranging from 50.5% to 89.2%, broadly consistent with 67.8% to 92.4% from the literature. Similarly, VE against Pillar 2 PCR confirmed infection ranged from 17% to 69.5% compared to 11.7% to 75.3% in the literature. We note that we only consider the first recorded infections for the Omicron period, which may bias our VE estimates to be higher, which is offset by our longer study period.

We considered VE stratified by deprivation (Figure 4). For Pillar 2 infections, observed differences in VE may be explained by differences in the propensity to seek testing which was associated with both deprivation level and vaccination status. Later in the pandemic, the propensity to seek testing was substantially lower in the most deprived areas than the least deprived (*37*). Higher VE for the least deprived areas might also be due to higher levels of infection-induced immunity in more deprived areas; if mild infections were less likely to be detected in more deprived areas, the unvaccinated group would

have higher immunity levels, resulting in lower VE estimates for the most deprived IMD quintile.

The policy implications of our study are that both protective and support measures need to be structured with ethnicity and deprivation in mind, particularly in the early stages of a pandemic. Deprivation had a particularly significant impact during the Alpha variant era for PCR-confirmed infections. For ethnicity, we observed a reduction in IRRs as the pandemic progressed (Figure 3). This suggests that more deprived areas were less able to isolate effectively, potentially due to work, during heightened public health restrictions. Ethnic minorities were more likely to work in occupations where workplace attendance continued even during lockdowns and also had less access to savings (*38*), potentially exposing them to a higher risk of community transmission. This is particularly true of, Bangladeshi and Pakistani men working in the hospitality, leisure and transport sectors. South Asian and African-Black men are also overrepresented in health and social care roles in England, potentially leading to higher levels of infection risk during the pandemic (*38*).

Considering the impact of the public health restriction level, we found IRRs>1 for all outcomes in the wild-type and Alpha periods. Restrictions were highest during periods of high hospitalisation and death incidence (SM fig. S8). The higher IRR for deaths during Alpha, relative to Wildtype, was consistent with the hospitalisation fatality ratio (HFR) being higher for Alpha than Wildtype (*39*). The IRR for deaths over the entire study period was comparable to the IRR for deaths during Alpha, suggesting that vaccination during Delta and Omicron, both periods of low to no restrictions, contributed to lower risk relative to Wildtype and Alpha. We find an IRR greater than one for hospitalisations but higher IRRs during Wildtype than Alpha. Similarly to deaths, we find the highest IRRs for the entire study period, pointing to the impact of vaccination in the second half of the study period. For Pillar 2 PCR cases, we observe an IRR greater than one, with IRRs highest during Wildtype. The IRRs for the full study period are lower than for Wildtype, as we observed significant infections in both the unvaccinated and vaccinated groups during Omicron, a period with no restrictions. We note that this IRR for the entire study period would be lower if we had taken Lateral Flow Tests (LFTs) into account.

This highlights the impact of the different timing of restrictions between regions, as we already account for week and region effects in the model. We note that other results were not sensitive to whether restrictions were included as a predictor in the model, and the IRRs for restrictions were robust to different resolutions of the regional covariate.

A limitation is that we only considered first recorded infections in this study and did attempt to investigate reinfections. However, we note that up to November 2021, with the arrival of the Omicron variant of SARS-CoV-2, 93% of all recorded infections were first infections. In addition, we only quantify incidence among those seeking tests, not underlying infection incidence. This may impact our results if the propensity to seek a test varied by age, deprivation or ethnicity, especially if underlying true infection incidence varied substantially by those variables. Hence for the Pillar-2 test outcome, risk

and VE differences between population groups/strata may reflect true exposure differences, differences in propensity to seek a test, or both. Such considerations do not affect the hospitalisation and death endpoints, as testing was universal for anyone entering hospital for the large majority of the study period. For Omicron, we only consider PCR test results and do not include Lateral Flow Tests (LFTs).

Future research priorities include a multi-outcome survival model to incorporate infections and severe outcomes in the same model. Integrating serology and infection prevalence data is essential in addressing questions such as the impact of pre-existing comorbidities on the risk heterogeneity by ethnicity or socio-economic group.

Material & Methods

Data sources

The UK Health Security Agency (UKHSA) maintains a single unified database of all SARS-CoV-2 positive test results in the country, categorised by Pillar. The resulting database contains a hashed version of a unique identifier (NHS number), the age, sex, ethnic group, home LTLA, symptom status, testing Pillar, and date of specimen for each positive case.

Second, linked to the case database by (hashed) National Health Service (NHS) number, UKHSA maintains a database of all deaths within 28 days of a positive PCR test for SARS-CoV-2. This database records the date of death, the date of hospital admission, and how the death was identified (e.g., via hospital reporting, death registration, or both).

Third, two NHS data sources on hospitalisation were also linked by NHS number to the case database: the Secondary Uses Service (SUS) data on hospital episodes and the Emergency Care Data Set (ECDS) on attendance at Accident and Emergency departments. An individual in these datasets was classified as hospitalised with COVID-19 if they had a positive PCR test between 14 days before admission and the day before discharge, and they were classified as an inpatient.

Lastly, England maintains a national vaccination register, the National Immunisation Management System (NIMS) (*20*). Every SARS-CoV-2 vaccine dose given in England is recorded at the time of vaccine administration. The anonymised version of the dataset used here contained a unique identifier (hashed NHS number), age, sex, ethnic group, LTLA and administration date.

We extracted data from the versions of these databases with data up to 25th July 2023. Cases with a valid NHS number were then linked to the NIMS vaccination dataset. Where a match was found, the case was classified as having received a vaccine (either before or after testing positive). Cases with no matching immunisation record were classified as having not been vaccinated. In our analysis, we used the vaccination status of an individual at the time of testing. The vaccination status includes the type of vaccine, number of doses received and time since the last vaccination.

PCR testing (Pillars 1 and 2) was initially focused on key workers in the NHS, social care and other sectors due to limited testing capacity. From May 2020, testing was widened to individuals with symptoms in the general public and regular asymptomatic testing of individuals working in high-risk settings, such as care homes, was introduced **(*40*)**. Individuals with a positive SARS-CoV-2 test were initially advised to self-isolate for 14 days, with subsequent rules being more complex and dependent on testing (fig. S9). Limited support was available to individuals who had to self-isolate, provided they met several criteria, such as not being able to work at home and being in receipt of at least one state benefit **(*41*)**.

Population denominator data was used to calculate the number of people who have neither received a vaccine nor been diagnosed with SARS-CoV-2 infection. We used the 2021 Census population estimates generated by the ONS (*42*), stratified by age, sex and LTLA (*43*). We made the simplifying assumption that in the absence of COVID-19, the size of the English population and its age distribution would have been at a steady state for our analysis period consistent with the Census 2021 estimates. Individuals with missing information on sex, age, or residence location were not included in the analysis.

We use the Index of Multiple Deprivation (IMD) dataset released in 2019 (*12*) by the UK Ministry of Housing, Local Communities, and Local Government, aggregated at the LTLA level and its subcomponents (SM section B.4).

The public health and social restrictions level was computed from ONS LTLA data up to December 2020 (*44*) and the Oxford COVID-19 Government Response Tracker (*1*) (SM section A and C.3). Several non-pharmaceutical interventions (NPIs) were used across the course of the pandemic, including containment, school closure and stay at home orders. The “roadmap out of lockdown” policy in spring 2021 had three intermediate steps between national lockdown and no restrictions. We map the restrictions over time onto a 5-level scale based on the individual policy measures in place at each time point (SM section C.3).

Statistical Analysis

Survival models were used to estimate Incidence Rate Ratios (IRR), an estimate of the increased risk of a particular outcome for a specific population group relative to a reference group. We used Poisson regression using an offset to adjust for person-days at risk to generate IRRs for the different strata of interest, proving a good approximation to continuous time survival models (SM section A)(*45*). Poisson regressions are more general than Cox proportional hazard survival models, allowing for non-proportional hazards. We investigated using negative binomial regressions to account for overdispersion. We parametrically captured variation by systematically examining evidence for interactions between covariates. We conducted model selection to identify the model which performed

best across all three outcomes of interest (infection, hospitalisation, and deaths) by considering the average Akaike's information criterion (AIC) and coefficient of determination (R^2) (*46*, *47*) across the three respective models.

We included terms for sex, ethnic group, vaccination status and IMD in all models. We used a categorical IMD variable, ranking LTLAs by average IMD score and then binning them into quintiles (1 being least deprived, five being most). The four other covariates included in all models were geographic region (either upper tier local authority or region), age (in 10-year bands with the last band being 85+ or 10-year bands with all individuals below 40 grouped and a max-age of 80+), restriction levels (used as a parametric covariate as restriction levels changed several times, creating non-continuous time intervals) and epidemiological week. We examined all pairwise interactions between these three variables. The combined restriction covariate was also included in the preferred model.

All survival analyses used the test specimen date as the outcome event date. Three SARS-CoV-2-related events were defined: death within 28 days of any positive PCR test, hospitalisation where the individual hospitalised tested PCR-positive between 14 days before admission and the day before discharge and was classed as an inpatient, and any PCR-confirmed infection within Pillar 2. We only consider the first recorded SARS-CoV-2 infection for all three events and associated hospitalisations and deaths.

We fitted the model for all three outcomes of interest. Individuals were treated as censored after a positive SARS-CoV-2 PCR test to exclude post-primary infections and avoid bias since the UK vaccination programme excluded people from being vaccinated within 28 days of a positive test.

Since the UKHSA/NHS and ONS data used different ethnicity categories, we aggregated ethnicities into six groups that could be identified consistently across all datasets: White, South Asian, Asian (other), Black, and Mixed/Other to be consistent across two datasets (SM section B.1). Details are provided in Supplementary Materials (SM section B.1). The computation of age for each individual is outlined in the SM section B.2.

We estimated VE accounting for the type of vaccine (mRNA or adenovirus-based), time since vaccination, and outcome of interest (death, hospitalisation, Pillar 2 PCR positive case). Details are provided in SM section B.3. By comparing currently vaccinated individuals against people who would be vaccinated in the future, we controlled for potential differences in behaviour and exposure risk between those who are never vaccinated and those who eventually are. Censoring after the first recorded positive test partially mitigated estimates from being affected by the accumulation of naturally acquired immunity in the population.

To investigate how estimates varied over time, for example, in periods dominated by particular variants, we restricted model fitting to data from the specific period of interest (censoring individuals with positive tests before that period).

We used negative binomial regressions as a sensitivity analysis to explore over-dispersion in the count data. In all figures, we display only results with a p-value of 0.05 or less.

Interactions between VE and deprivation were examined by running the preferred model with an additional interaction term for vaccine status and IMD. Similarly, we investigated associations between VE and any ethnicity by running the preferred model with an interaction term for vaccine status and ethnicity.

All analyses were undertaken in R version 4.3.1 using the H2O.ai machine learning package version 3.42.0.2 (*48*), which offers high-performance parallelised algorithms for fitting general linear models to large datasets. A 32-core Intel-based server with 128GB RAM was used to conduct the analyses.

Data access

While all data used in this analysis were anonymised, the individual-level nature of the data used risks individuals being identified or being able to self-identify if it is released publicly. Requests for access to the underlying source data should be directed to UKHSA. The access to protected data request is available at https://www.gov.uk/government/publications/accessing-ukhsa-protected-data/accessing-ukhsa-protected-data. All other data needed to evaluate the conclusions of the paper are present in the paper and/or the Supplementary Materials.

Ethical approval

Surveillance of COVID-19 testing and vaccination is undertaken under Regulation 3 of The Health Service (Control of Patient Information) Regulations 2002 to collect confidential patient information (www.legislation.gov.uk/uksi/2002/1438/regulation/3/made) under Sections 3(i) (a) to (c), 3(i)(d) (i) and (ii) and 3(3). Data were shared with the investigators as part of the UK's emergency response to the COVID-19 pandemic via the SPI-M subcommittee of the UK Scientific Advisory Group for Emergencies (SAGE). Ethics permission was sought for the study via Imperial College London's standard ethical review processes, and the study was approved by the College's Research Governance and Integrity Team (ICREC reference: 21IC6945).

Role of the funding source

The funders of the study had no role in the study design, data collection, data analysis, data interpretation, or writing of the report.

References from SI (need to be at the end of the main text references; delete these lines once unlinked for resubmission): (*1*, *12*, *14*, *21*, *42*, *45*, *48*, *49*)

References

1. T. Hale, N. Angrist, R. Goldszmidt, B. Kira, A. Petherick, T. Phillips, S. Webster, E. Cameron-Blake, L. Hallas, S. Majumdar, H. Tatlow, A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker). *Nat. Hum. Behav.* **5**, 529–538 (2021).

2. T. Phillips, A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker); https://github.com/OxCGRT/covid-policy-dataset.

3. E. Kontopantelis, M. A. Mamas, R. T. Webb, A. Castro, M. K. Rutter, C. P. Gale, D. M. Ashcroft, M. Pierce, K. M. Abel, G. Price, C. Faivre-Finn, H. G. C. Van Spall, M. M. Graham, M. Morciano, G. P. Martin, T. Doran, Excess deaths from COVID-19 and other causes by region, neighbourhood deprivation level and place of death during the first 30 weeks of the pandemic in England and Wales: A retrospective registry study. *Lancet Reg. Health - Eur.* **7**, 100144 (2021).

4. H. Ward, C. Atchison, M. Whitaker, K. E. C. Ainslie, J. Elliott, L. Okell, R. Redd, D. Ashby, C. A. Donnelly, W. Barclay, A. Darzi, G. Cooke, S. Riley, P. Elliott, SARS-CoV-2 antibody prevalence in England following the first peak of the pandemic. *Nat. Commun.* **12**, 905 (2021).

5. V. J. McGowan, C. Bambra, COVID-19 mortality and deprivation: pandemic, syndemic, and endemic health inequalities. *Lancet Public Health* **7**, e966–e975 (2022).

6. C.-H. Lo, L. H. Nguyen, D. A. Drew, E. T. Warner, A. D. Joshi, M. S. Graham, A. Anyane-Yeboa, F. M. Shebl, C. M. Astley, J. C. Figueiredo, C.-G. Guo, W. Ma, R. S. Mehta, S. Kwon, M. Song, R. Davies, J. Capdevila, C. H. Sudre, J. Wolf, Y. C. Cozier, L. Rosenberg, L. R. Wilkens, C. A. Haiman, L. L. Marchand, J. R. Palmer, T. D. Spector, S. Ourselin, C. J. Steves, A. T. Chan, Race, ethnicity, community-level socioeconomic factors, and risk of COVID-19 in the United States and the United Kingdom. *eClinicalMedicine* **38**, 101029 (2021).

7. D. Pan, S. Sze, J. S. Minhas, M. N. Bangash, N. Pareek, P. Divall, C. Ml. Williams, M. R. Oggioni, I. B. Squire, L. B. Nellums, W. Hanif, K. Khunti, M. Pareek, The impact of ethnicity on clinical outcomes in COVID-19: A systematic review. *EClinicalMedicine* **23**, 100404 (2020).

8. R. Mathur, C. T. Rentsch, C. E. Morton, W. J. Hulme, A. Schultze, B. MacKenna, R. M. Eggo, K. Bhaskaran, A. Y. S. Wong, E. J. Williamson, H. Forbes, K. Wing, H. I. McDonald, C. Bates, S. Bacon, A. J. Walker, D. Evans, P. Inglesby, A. Mehrkar, H. J. Curtis, N. J. DeVito, R. Croker, H. Drysdale, J. Cockburn, J. Parry, F. Hester, S. Harper, I. J. Douglas, L. Tomlinson, S. J. W. Evans, R. Grieve, D. Harrison, K. Rowan, K. Khunti, N. Chaturvedi, L. Smeeth, B. Goldacre, Ethnic differences in SARS-CoV-2 infection and COVID-19-related hospitalisation, intensive care unit admission, and death in 17 million adults in England: an observational cohort study using the OpenSAFELY platform. *The Lancet* **397**, 1711–1724 (2021).

9. T. Padellini, R. Jersakova, P. J. Diggle, C. Holmes, R. E. King, B. C. L. Lehmann, A.-M. Mallon, G. Nicholson, S. Richardson, M. Blangiardo, Time varying association between deprivation, ethnicity and SARS-CoV-2 infections in England: A population-based ecological study. *Lancet Reg. Health - Eur.* **15**, 100322 (2022).

10. C. Razieh, S. Shabnam, H. Dambha-Miller, E. J. A. Morris, T. Yates, Y. Chudasama, F. Zaccardi, C. Gillies, A. Banerjee, M. Pareek, B. Lacey, M. White, K. Khunti, N. Islam, Socioeconomic inequalities in risk of infection with SARS-CoV-2 delta and omicron variants in the UK, 2020-22: analysis of the longitudinal COVID-19 Infection Survey. *BMJ Med.* **3**, e000624 (2024).

11. A. J. Wood, A. R. Sanchez, P. R. Bessell, R. Wightman, R. R. Kao, Assessing the importance of demographic risk factors across two waves of SARS-CoV-2 using fine-scale case data. *PLOS Comput. Biol.* **19**, e1011611 (2023).

12. UK Ministry of Housing, Communities and Local Government, The English Indices of Deprivation 2019 (2019). https://assets.publishing.service.gov.uk/media/5d8e26f6ed915d5570c6cc55/IoD2019_Statistical_Release.pdf.

13. K. Morrissey, F. Spooner, J. Salter, G. Shaddick, Area level deprivation and monthly COVID-19 cases: The impact of government policy in England. *Soc. Sci. Med.* **289**, 114413 (2021).

14. T. Phillips, Oxford Covid-19 Government Response Tracker (OxCGRT), Blavatnik School of Government, University of Oxford, GitHub; https://github.com/OxCGRT/covid-policy-tracker.

15. S. Vandentorren, S. Smaïli, E. Chatignoux, M. Maurel, C. Alleaume, L. Neufcourt, M. Kelly-Irving, C. Delpierre, The effect of social deprivation on the dynamic of SARS-CoV-2 infection in France: a population-based analysis. *Lancet Public Health* **7**, e240–e249 (2022).

16. S. Rohleder, Dr. D. Costa, P. K. Bozorgmehr, Area-level socioeconomic deprivation, non-national residency, and Covid-19 incidence: A longitudinal spatiotemporal analysis in Germany. *eClinicalMedicine* **49**, 101485 (2022).

17. S. Magesh, D. John, W. T. Li, Y. Li, A. Mattingly-app, S. Jain, E. Y. Chang, W. M. Ongkeko, Disparities in COVID-19 Outcomes by Race, Ethnicity, and Socioeconomic Status: A Systematic Review and Meta-analysis. *JAMA Netw. Open* **4**, e2134147 (2021).

18. Department of Health and Social Care, "Coronavirus (COVID-19): scaling up testing programmes" (2020); https://assets.publishing.service.gov.uk/media/5e888f05e90e0707799498b3/coronavirus-covid-19-testing-strategy.pdf.

19. UK Health Security Agency, Testing Times: striving for quality in the COVID-19 national testing service, April 2020 to March 2022. https://www.gov.uk/government/publications/coronavirus-covid-19-scaling-up-testing-programmes.

20. NHS England, National vaccination programmes. https://www.england.nhs.uk/contact-us/privacy-notice/national-flu-vaccination-programme/.

21. J. Lopez Bernal, N. Andrews, C. Gower, E. Gallagher, R. Simmons, S. Thelwall, J. Stowe, E. Tessier, N. Groves, G. Dabrera, R. Myers, C. N. J. Campbell, G. Amirthalingam, M. Edmunds, M. Zambon, K. E. Brown, S. Hopkins, M. Chand, M. Ramsay, Effectiveness of Covid-19 Vaccines against the B.1.617.2 (Delta) Variant. *N. Engl. J. Med.* **385**, 585–594 (2021).

22. J. Lopez Bernal, N. Andrews, C. Gower, C. Robertson, J. Stowe, E. Tessier, R. Simmons, S. Cottrell, R. Roberts, M. O'Doherty, K. Brown, C. Cameron, D. Stockton, J. McMenamin, M. Ramsay, Effectiveness of the Pfizer-BioNTech and Oxford-AstraZeneca vaccines on covid-19 related symptoms, hospital admissions, and mortality in older adults in England: test negative case-control study. *BMJ*, n1088 (2021).

23. H. Chua, S. Feng, J. A. Lewnard, S. G. Sullivan, C. C. Blyth, M. Lipsitch, B. J. Cowling, The Use of Test-negative Controls to Monitor Vaccine Effectiveness: A Systematic Review of Methodology. *Epidemiology* **31**, 43–64 (2020).

24. WHO, "Evaluation of COVID-19 vaccine effectiveness: interim guidance, 17 March 2021" (2021); https://iris.who.int/handle/10665/340301.

25. Y. M. Bar-On, Y. Goldberg, M. Mandel, O. Bodenheimer, L. Freedman, N. Kalkstein, B. Mizrahi, S. Alroy-Preis, N. Ash, R. Milo, A. Huppert, Protection of BNT162b2 Vaccine Booster against

Covid-19 in Israel. *N. Engl. J. Med.* **385**, 1393–1400 (2021).

26. H. Chemaitelly, H. M. Yassine, F. M. Benslimane, H. A. Al Khatib, P. Tang, M. R. Hasan, J. A. Malek, P. Coyle, H. H. Ayoub, Z. Al Kanaani, E. Al Kuwari, A. Jeremijenko, A. H. Kaleeckal, A. N. Latif, R. M. Shaik, H. F. Abdul Rahim, G. K. Nasrallah, M. G. Al Kuwari, H. E. Al Romaihi, M. H. Al-Thani, A. Al Khal, A. A. Butt, R. Bertollini, L. J. Abu-Raddad, mRNA-1273 COVID-19 vaccine effectiveness against the B.1.1.7 and B.1.351 variants and severe COVID-19 disease in Qatar. *Nat. Med.* **27**, 1614–1621 (2021).

27. E. Vasileiou, C. R. Simpson, T. Shi, S. Kerr, U. Agrawal, A. Akbari, S. Bedston, J. Beggs, D. Bradley, A. Chuter, S. De Lusignan, A. B. Docherty, D. Ford, F. R. Hobbs, M. Joy, S. V. Katikireddi, J. Marple, C. McCowan, D. McGagh, J. McMenamin, E. Moore, J. L. Murray, J. Pan, L. Ritchie, S. A. Shah, S. Stock, F. Torabi, R. S. Tsang, R. Wood, M. Woolhouse, C. Robertson, A. Sheikh, Interim findings from first-dose mass COVID-19 vaccination roll-out and COVID-19 hospital admissions in Scotland: a national prospective cohort study. *The Lancet* **397**, 1646–1657 (2021).

28. V. Nafilyan, C. R. Bermingham, I. L. Ward, J. Morgan, F. Zaccardi, K. Khunti, J. Stanborough, A. Banerjee, J. C. Doidge, Risk of death following COVID-19 vaccination or positive SARS-CoV-2 test in young people in England. *Nat. Commun.* **14**, 1541 (2023).

29. A. Endo, Centre for the Mathematical Modelling of Infectious Diseases COVID-19 Working Group, S. Abbott, A. J. Kucharski, S. Funk, Estimating the overdispersion in COVID-19 transmission using outbreak sizes outside China. *Wellcome Open Res.* **5**, 67 (2020).

30. E. B. Hodcroft, M. S. Wohlfender, R. A. Neher, J. Riou, C. L. Althaus, Estimating R_e and overdispersion in secondary cases from the size of identical sequence clusters of SARS-CoV-2. [Preprint] (2024). https://doi.org/10.1101/2024.05.26.24307940.

31. COVID-19 Forecasting Team, Variation in the COVID-19 infection–fatality ratio by age, time, and geography during the pre-vaccine era: a systematic analysis. *The Lancet* **399**, 1469–1488 (2022).

32. N. W. Papageorge, M. V. Zahn, M. Belot, E. Van Den Broek-Altenburg, S. Choi, J. C. Jamison, E. Tripodi, Socio-demographic factors associated with self-protecting behavior during the Covid-19 pandemic. *J. Popul. Econ.* **34**, 691–738 (2021).

33. T. Rawson, W. Hinsley, R. Sonabend, E. Semenova, A. Cori, N. M. Ferguson, The impact of health inequity on regional variation of COVID-19 transmission in England. [Preprint] (2024). https://doi.org/10.1101/2024.04.20.24306121.

34. S. D. Nagar, A. M. Nápoles, I. K. Jordan, L. Mariño-Ramírez, Socioeconomic deprivation and genetic ancestry interact to modify type 2 diabetes ethnic disparities in the United Kingdom. *eClinicalMedicine* **37**, 100960 (2021).

35. A. Ramachandran, Trends in prevalence of diabetes in Asian countries. *World J. Diabetes* **3**, 110 (2012).

36. J. Stowe, N. Andrews, F. Kirsebom, M. Ramsay, J. L. Bernal, Effectiveness of COVID-19 vaccines against Omicron and Delta hospitalisation, a test negative case-control study. *Nat. Commun.* **13**, 5736 (2022).

37. S. Bajaj, S. Chen, R. Creswell, R. Naidoo, J. L.-H. Tsui, O. Kolade, G. Nicholson, B. Lehmann, J. A. Hay, M. U. G. Kraemer, R. Aguas, C. A. Donnelly, T. Fowler, S. Hopkins, L. Cantrell, P. Dahal, L. J. White, K. Stepniewska, M. Voysey, B. Lambert, R. Aguas, M. Amswych, B.

Andersen-Waine, S. Bajaj, K. Bimpong, A. Bodley, L. Cantrell, S. Chen, R. Creswell, P. Dahal, S. Dickinson, S. Dittrich, T. Evans, A. Ferguson-Lewis, C. Franco, B. Gao, R. Hounsell, M. Kasim, C. Keene, B. Lambert, U. Mahmood, M. Mills, A. Moldokmatova, S. Molyneux, R. Naidoo, R. Ngwafor Anye, J. Norman, W. Pan-Ngum, S. Pokharel, A. Polner, E. Rowe, S. Saralamba, R. Shretta, S. Silal, K. Stepniewska, J. L-H Tsui, M. Voysey, M. Wanat, L. J. White, COVID-19 testing and reporting behaviours in England across different sociodemographic groups: a population-based study using testing data and data from community prevalence surveillance surveys. *Lancet Digit. Health* **6**, e778–e790 (2024).

38. L. Platt, R. Warwick, COVID-19 and Ethnic Inequalities in England and Wales [*]. *Fisc. Stud.* **41**, 259–289 (2020).

39. P. N. Perez-Guzman, E. Knock, N. Imai, T. Rawson, Y. Elmaci, J. Alcada, L. K. Whittles, D. Thekke Kanapram, R. Sonabend, K. A. M. Gaythorpe, W. Hinsley, R. G. FitzJohn, E. Volz, R. Verity, N. M. Ferguson, A. Cori, M. Baguelin, Epidemiological drivers of transmissibility and severity of SARS-CoV-2 in England. *Nat. Commun.* **14**, 4279 (2023).

40. Independent report Chapter 6: testing (2023). https://www.gov.uk/government/publications/technical-report-on-the-covid-19-pandemic-in-the-uk/chapter-6-testing#introduction.

41. Stay at home: guidance for households with possible or confirmed coronavirus (COVID-19) infection (2022). https://www.gov.uk/government/publications/covid-19-stay-at-home-guidance/stay-at-home-guidance-for-households-with-possible-coronavirus-covid-19-infection#financial-or-other-practical-support.

42. ONS, “Census;” https://www.ons.gov.uk/census.

43. Office of National Statistics, Office of National Statistics Census 2021 Create a custom dataset; https://www.ons.gov.uk/datasets/create.

44. ONS, Full list of local COVID alert levels by area. https://www.gov.uk/guidance/full-list-of-local-covid-alert-levels-by-area.

45. N. Laird, D. Olivier, Covariance Analysis of Censored Survival Data Using Log-Linear Analysis Techniques. *J. Am. Stat. Assoc.* **76**, 231–240 (1981).

46. P. Stoica, Y. Selen, Model-order selection. *IEEE Signal Process. Mag.* **21**, 36–47 (2004).

47. T. O. Kvalseth, Cautionary Note about R 2. *Am. Stat.* **39**, 279 (1985).

48. H2O, H2O.ai. h2o: R Interface for H2O. R package version 3.42.0.2 2023; https://github.com/h2oai/h2o-3.

49. Imperial College London/YouGov, COVID-19 behaviour tracker, (2020); https://www.imperial.ac.uk/global-health-innovation/what-we-do/our-response-to-covid-19/covid-19-behaviour-tracker/.

Declarations

Funding All authors acknowledge funding from the Medical Research Council (MRC) Centre for Global Infectious Disease Analysis (MR/X020258/1) funded by the UK MRC and carried out in the frame of the Global Health EDCTP3 Joint Undertaking supported by the EU; the NIHR for support for the Health Research Protection Unit in Modelling and Health Economics, a partnership between the UK Health Security Agency (UKHSA), Imperial College London, and London School of Hygiene & Tropical Medicine (grant code NIHR200908); a philanthropic donation from Community Jameel supporting the work of the Jameel Institute S.B. acknowledges support from the Novo Nordisk Foundation via The Novo Nordisk Young Investigator Award (NNF20OC0059309). SB acknowledges the Danish National Research Foundation (DNRF160) through the chair grant. S.B. acknowledges support from Schmidt Sciences via the Schmidt Polymath Award (G-22-63345), which also supports C.M. The funders of the study had no role in the study design, data collection, data analysis, data interpretation, or writing of the report. For the purpose of open access, the author has applied a 'Creative Commons Attribution' (CC BY) licence to any Author Accepted Manuscript version arising from this submission.

Availability of data and materials While all data used in this analysis were anonymised, the individual-level nature of the data used risks individuals being identified or being able to self-identify if it is released publicly. Requests for access to the underlying source data should be directed to UKHSA. The access to protected data request is available at https://www.gov.uk/government/publications/accessing-ukhsa-protected-data/accessing-ukhsa-protected-data. All other data needed to evaluate the conclusions of the paper are present in the paper and/or the Supplementary Materials.

Code availability https://github.com/cm401/covid19_risk_heterogeneity_england (https://doi.org/10.5281/zenodo.15202350)

Competing interests. PPG has performed paid consulting for Pfizer and Munich Re in matters related to pandemic modelling; neither entity had any form of involvement in the present study. The authors declare no other competing interests.

Authors' Contributions

Conceptualisation: NMF, CM, SB

Data curation & validation: CM, WH, PPG, NMF

Methodology: CM, TR, WH, PPG, SB, NMF

Investigation: CM

Visualisation: CM

Supervision: NMF, SB

Writing – original draft: CM

Writing – review & editing: CM, TR, WH, PNPG, SB, NMF

All authors discussed, edited, and approved the final version of the manuscript. All authors had final responsibility for the decision to submit the manuscript for publication.

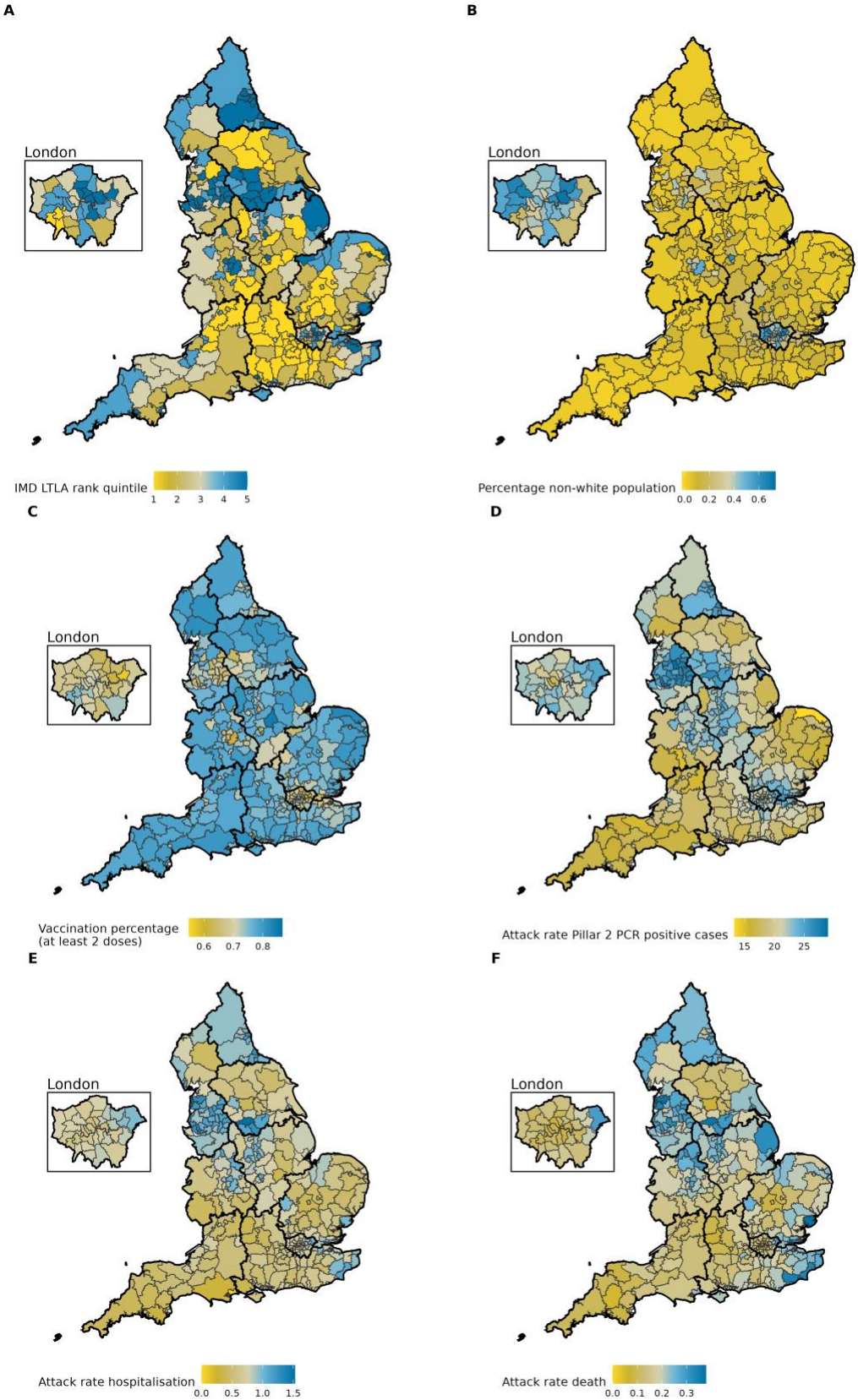

Figure 1: Maps of the population of England at the LTLA level between 10 May 2020 – 27 February 2022. (A) Index of Multiple Deprivation quintile by LTLA, (B) percentage of non-White population within the LTLA by LTLA, (C) percentage of individuals vaccinated with at least two doses for all ages, (D) percentage of positive Pillar 2 PCR confirmed first recorded cases by LTLA, (E) percentage of hospitalisation by LTLA, (F) percentage deaths by LTLA

A

Variable	Category	n	Person days at risk	Cases (Pillar 2 PCR)	Hospitalisations	Deaths
Sex	Female	28,685,081	17,345,400,702	6,517,293	222,651	49,389
	Male	27,659,329	16,807,116,305	5,793,192	216,432	58,434
Age	under_40	28,863,746	17,378,043,704	7,986,079	97,480	1,246
	40_49	7,126,246	4,298,749,448	1,832,839	40,577	2,211
	50_59	7,574,556	4,665,987,550	1,400,202	56,807	5,989
	60_69	5,734,391	3,547,069,498	635,915	61,340	12,753
	70_79	4,559,170	2,762,316,669	284,970	79,674	26,508
	above_80	2,486,301	1,500,350,138	170,480	103,205	59,116
Region	East Midlands	4,842,841	2,923,964,911	1,086,738	41,310	10,572
	East of England	6,278,538	3,821,826,715	1,296,183	43,192	12,411
	London	8,837,013	5,309,194,394	1,937,334	67,441	12,285
	North East	2,651,524	1,599,576,314	638,786	26,268	6,407
	North West	7,420,784	4,455,050,061	1,849,257	78,606	19,004
	South East	9,264,652	5,678,938,650	1,890,892	59,760	16,179
	South West	5,680,938	3,491,446,658	1,050,950	28,731	7,439
	West Midlands	5,894,953	3,564,918,434	1,296,108	48,408	12,840
	Yorkshire and The Humber	5,473,167	3,307,600,870	1,264,237	45,367	10,686
IMD Quintile	IMD1	7,674,923	4,699,769,599	1,569,158	44,322	11,696
	IMD2	9,445,584	5,762,949,346	1,921,063	63,003	16,629
	IMD3	10,876,412	6,615,134,330	2,302,369	78,105	20,659
	IMD4	11,801,363	7,137,336,955	2,586,990	95,823	22,278
	IMD5	16,546,128	9,937,326,777	3,930,905	157,830	36,561
Ethnicity	Asian (other)	1,548,640	913,540,586	322,769	9,381	1,209
	Black	2,307,809	1,356,279,568	601,101	23,642	3,152
	Mixed/Other	2,698,916	1,643,470,031	395,571	11,892	1,143
	South Asian	4,058,419	2,381,064,042	984,164	33,238	5,381
	White	45,730,626	27,858,162,780	10,006,880	360,930	96,938

B

Variable	Category	n	Person days at risk	Cases (Pillar 2 PCR)	Hospitalisations	Deaths
Vaccination status	Not Vaccinated	12,524,706	21,652,863,421	7,030,900	289,451	80,601
	First Dose >3w Adenovirus	343,341	1,168,792,975	97,771	6,782	2,419
	Second Dose over 18w Adenovirus	2,781,388	1,561,139,076	1,068,220	33,061	5,879
	Booster Dose <2w Adenovirus	7,434,225	772,914,618	428,088	10,506	966
	Booster Dose >2w Adenovirus	9,188,109	332,906,007	137,555	11,487	1,522
	Booster over 18w Adenovirus	508,392	12,907,790	5,238	661	148
	First Dose <21d mRNA	253,760	438,532,528	210,475	7,628	2,364
	First Dose >3w mRNA	2,462,228	1,187,931,423	542,990	8,482	2,388
	Second Dose <2w mRNA	331,407	234,175,466	47,748	627	66
	Second Dose >2w mRNA	1,206,396	829,053,400	113,626	1,686	132
	Second Dose 10-18w mRNA	967,000	919,526,650	373,924	4,419	398
	Second Dose over 18w mRNA	4,066,262	1,129,558,164	751,049	18,615	3,765
	Booster 10-18w mRNA	301,068	178,424,050	168,635	1,738	151
	Booster Dose <2w mRNA	6,054,894	586,137,083	294,073	7,356	770
	Booster Dose >2w mRNA	5,512,367	326,242,322	207,494	14,032	1,809
	Booster over 18w mRNA	1,954,193	43,244,152	16,974	2,059	385

Table 1: Population for England with numbers of first recorded Pillar 2 PCR confirmed SARS-CoV-2, hospitalisations and deaths in England between 10th May 2020 and 27th February 2022. (A) Population breakdown by sex, age, region, deprivation, and ethnicity. (B) Breakdown by vaccination status for categories with more than 200,000 doses at the end of the study period (n is the count of individuals in this category at the end of the study period, person days at risk and counts of Pillar 2 PCR positive cases, hospitalisations and deaths are over the entire study period). Vaccination status categories are by vaccination type (Adenovirus or mRNA), vaccine dose, and time since vaccination. The whole table for panel B is available in SI Figure D.9.

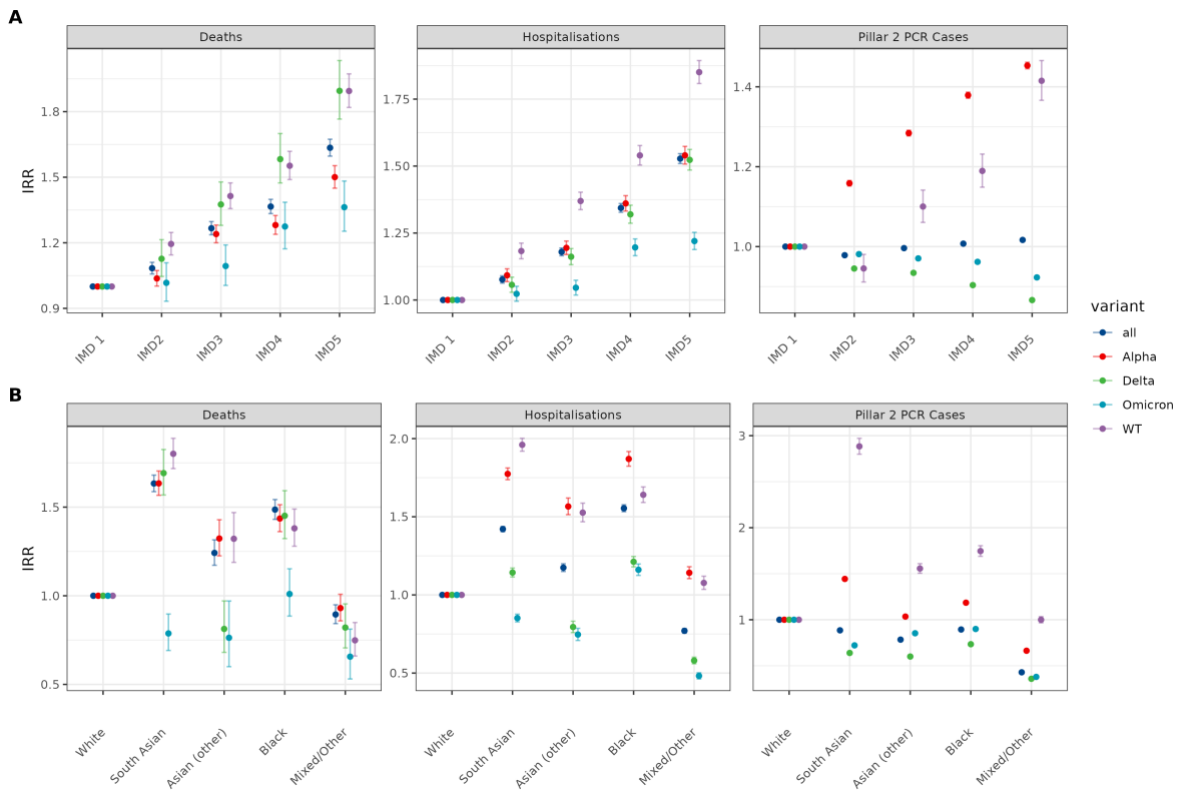

Figure 2: Estimated IRRs for (A) deprivation and (B) ethnicity covariates from preferred models. Deprivation was categorised by quintiles of the Index of Multiple Deprivation score defined at the LTLA level (results for sub-components are available in the SI), with the first (least deprived) quintile (IMD1) being the reference group (IRR=1). Results are shown for the whole pandemic ('all' - test dates between May 2020 and February 2022) and for the time intervals within that period where specific viral variants dominated (see SI for details). WT=Wild-Type (pre-December 2020).

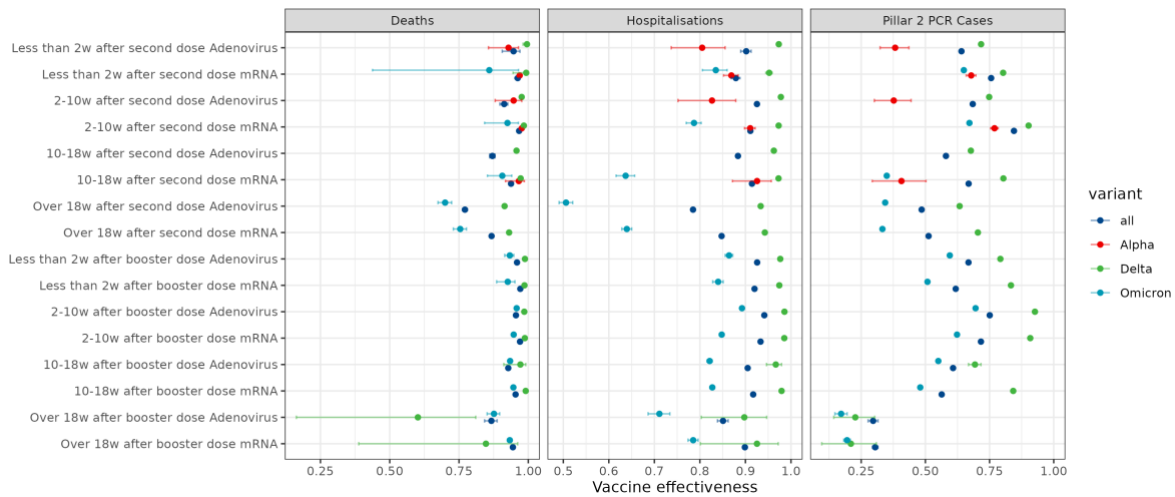

Figure 3: Vaccine effectiveness (VE=1-RR) for different vaccination statuses with not vaccinated as the reference group (IRR=1 / VE=0). Results are shown for the whole pandemic ('all' - test dates between May 2020 and February 2022) and for the time intervals within that period where specific viral variants dominated (see SI for details). Results for the Second Dose and Booster Dose categories are displayed for mRNA and Adenovirus vaccines (all other results are in SI Table D.3 and Figure D.11). WT=Wild-Type (pre-December 2020).

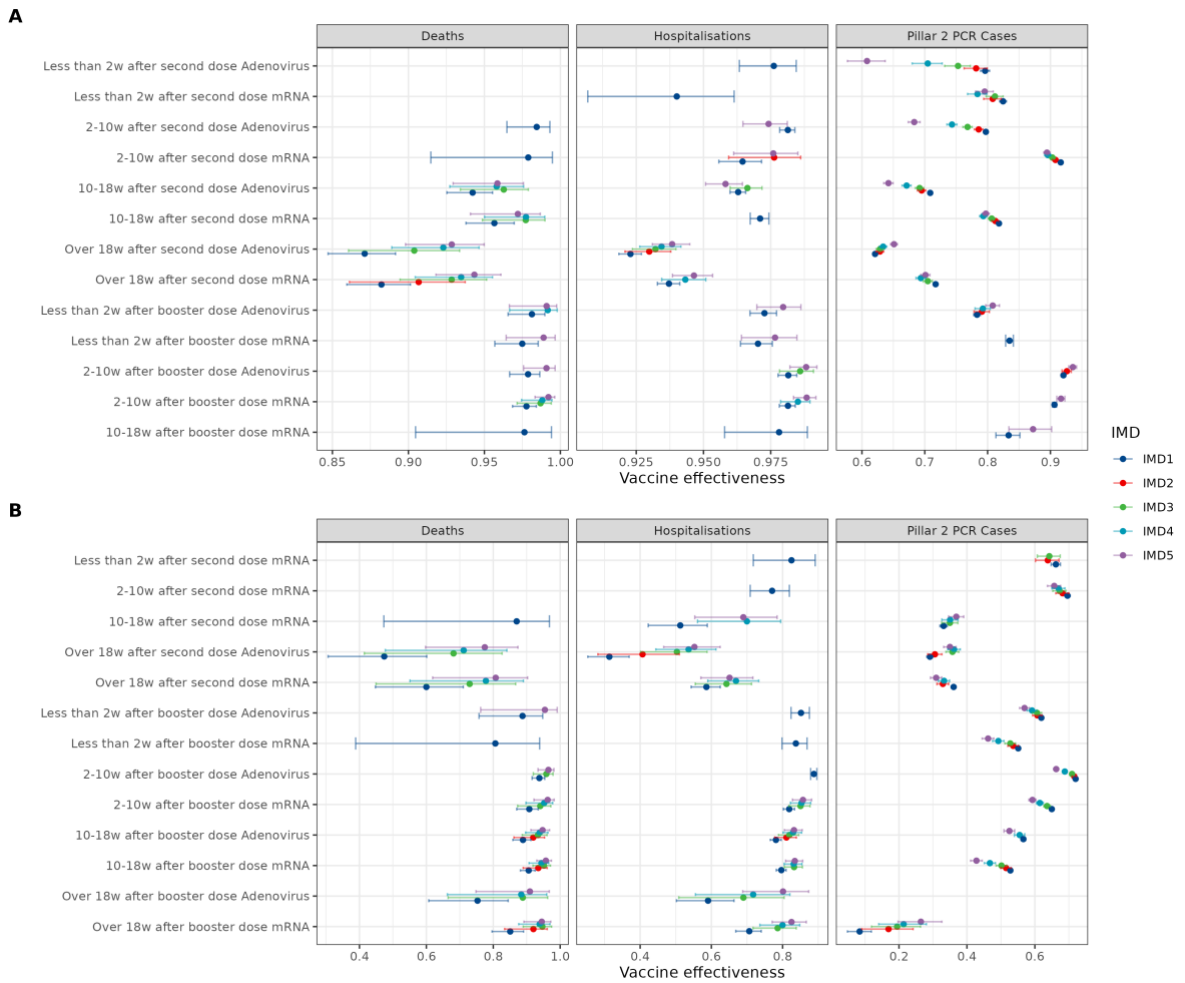

Figure 4: Vaccine effectiveness (VE=1-RR) for the preferred model with an additional interaction term for vaccine status and IMD, with not vaccinated as the reference group (IRR=1 / VE=0). Results are shown for the Delta period in panel (A) and Omicron period in panel (B), and results for the Second Dose and Booster Dose categories are displayed for mRNA and Adenovirus.

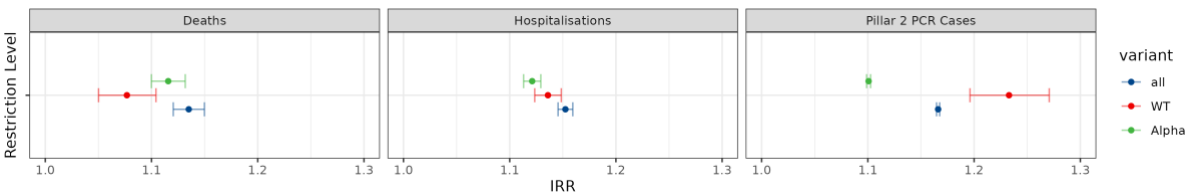

Figure 5: Estimated IRRs for restriction levels. Results are shown for the whole pandemic ('all' - test dates between May 2020 and February 2022) and for the time intervals within that period where specific viral variants dominated (see SI for details). Restriction level is a numeric covariate in the model.