Time varying association between deprivation, ethnicity and SARS-CoV-2 infections in England: a space-time study

Tullia Padellini, Radka Jersakova, Peter J. Diggle, Chris Holmes, Ruairidh E. King, Brieuc C. L. Lehmann, Ann-Marie Mallon, George Nicholson, Sylvia Richardson, Marta Blangiardo

Supplementary Material
1 Model specification

Let \( i = 1, \ldots, n = 312 \) index the spatial unit (LTLA), \( t = 1, \ldots, 68 \) index the week and \( j = 5 - 12, \ldots, 65 + \) index age groups.

1.1 Test Positivity

Let \( B_{itj} \) be the number of Pillar 2 positive tests in LTLA \( i \) during week \( t \) for age-group \( j \) and \( n_{itj} \) the total number of processed Pillar 2 tests. We assume \( B_{itj} \) follows a Binomial distribution

\[
B_{itj} | p_{itj} \sim \text{Bin}(p_{itj}, n_{itj}) \quad \forall i, j, t.
\]

We use the following logistic regression to assess the effect of the variables of interest on the test positivity \( p_{itj} \):

\[
\text{logit}(p_{itj}) = \beta_0 + \alpha_j + \beta_{1, \text{IMD}_i} + \beta_{2, \text{BAME}_i} + \gamma_1 \text{Urban}_i + \gamma_2 \text{Vax}_i + \delta_{1, m, \text{IMD}_i} + \delta_{2, m, \text{BAME}_i} + \lambda_i + \omega_t + \xi_{it} \quad \forall i, j, t.
\]

When considering the distinct BAME subgroups, the model can be reformulated as:

\[
\text{logit}(p_{itj}) = \beta_0 + \alpha_j + \beta_{1, \text{IMD}_i} + \beta_{2, \text{Black}_i} + \beta_{3, \text{SouthAsian}_i} + \beta_{4, \text{Other}_i} + \gamma_1 \text{Urbanicity}_i + \gamma_2 \text{Vax}_i + \delta_{1, m, \text{IMD}_i} + \delta_{2, m, \text{Black}_i} + \delta_{3, m, \text{SouthAsian}_i} + \delta_{4, m, \text{Other}_i} + \lambda_i + \omega_t + \xi_{it} \quad \forall i, j, t.
\]

The model specification includes a global intercept \((\beta_0)\), an age specific intercept \((\alpha_j)\), the overall effects of the covariates of interest (captured at the LTLA level), as well as month-specific covariate effects \(\delta_{k, m(t)}\), where \(m(t)\) represents which month week \(t\) belongs to. We impose a Random Walk 1 structure with sum to zero constraints on \(\delta_{k, m(t)}\), i.e.

\[
\delta_{k, m(t)} | \delta_{k, m(t) - 1} \sim \text{Norm} \left( \delta_{k, m(t) - 1}, \sigma_{\delta_k}^2 \right) \quad k = 1, \ldots, 4 \quad \forall t.
\]

The sum-to-zero constraints ensure that each \(\delta_{k, m(t)}\) measures the difference of the effect of month \(m(t)\) with respect to the global average \(\beta_k\).

Following [5, 4] we model \(\lambda = (\lambda_1, \ldots, \lambda_n)\), the random effect accounting for the spatial autocorrelation across LTLAs, as

\[
\lambda = \frac{1}{\tau} \left( \sqrt{1 - \rho} v + \sqrt{\rho} u \right).
\]

\(u = (u_1, \ldots, u_n)\) is a spatially structured random effect with prior distribution

\[
u \sim \text{Norm}(0, Q^-)
\]

where \(Q^-\), is the inverse of the precision matrix of a Besag model, scaled in the sense of [6]. \(v = (v_1, \ldots, v_n)\) is an i.i.d. Gaussian random effect [1], that is

\[
v \sim \text{Norm}(0, I)
\]

where \(I\) is the identity matrix. To account for the time dependence, \(\omega_t\) can be modelled through a random walk of order 2, that, given \(\Delta^2 \omega_t = \omega_t - 2\omega_{t+1} + \omega_{t+2}\), can be formalized as

\[
\Delta^2 \omega_t | \sigma_\omega^2 \sim \text{Norm} \left( 0, \sigma_\omega^2 \right).
\]
Both $\lambda$ and $\omega$ imply some smoothing, hence they help highlight the persistent patterns in the data, but they cannot account for transient anomalies. We thus include a spatio-temporal interaction term $\xi_{it}$ to account for the residual local variability from the general, structured spatio-temporal trend. We assume that

$$
\xi_{it} \sim \text{Norm}(0, \sigma_\xi^2)
$$

so that $\xi_{it}$ is unstructured in both space and time (or Type I interaction in the sense of [2]). This is critical for highlighting anomalies as the lack of temporal and spatial structure prevents the interaction term to smooth away unexpected but transient behavior.

Finally we assign a Gaussian prior with sum to zero constraint to the parameter $\alpha_j$ for all $j$, we set a minimally informative $\text{Gamma}(1, 5 \times 10^{-5})$ on the inverse of $\sigma_\delta^2, \sigma_\omega^2, \sigma_\xi^2$ and a minimally informative $\text{Norm}(0, 1000)$ prior for the fixed effect coefficients $\beta_0, \beta_k, \gamma_s$, with $k = 1, \ldots, 4$ and $s = 1, 2$.

### 1.2 Debiasd Prevalence

We build on the framework of [3], which allows us to estimate a distribution for the number of cases. Let $I_{itj}$ be the posterior median of such distribution for LTLA $i$, week $t$ and age-group $j$. We assume $I_{itj}$ follows a Binomial distribution

$$
I_{itj} \mid p_{itj} \sim \text{Bin}(p_{itj}, n_{ij}) \quad \forall i, j, t
$$

where $n_{ij}$ the population of LTLA $i$ for age-group $j$ as retrieved by the ONS mid-year population estimates for 2020. Similar to the model for test positivity, we model the debiased prevalence $p_{itj}$ as

$$
\logit(p_{itj}) = \beta_0 + \alpha_j + \\
\beta_1 \text{IMD}_i + \beta_2 \text{Black}_i + \beta_3 \text{SouthAsian}_i + \beta_4 \text{Other}_i + \\
\gamma_1 \text{Urbanicity}_i + \gamma_2 \text{Vax}_i + \\
\delta_1,m \text{IMD}_i + \delta_2,m \text{Black}_i + \delta_3,m \text{SouthAsian}_i + \delta_4,m \text{Other}_i + \\
\lambda_i + \omega_t + \xi_{it} \quad \forall i, j, t.
$$

with all model components the same as specified in the previous section.
|                         | Outcome                           | Exposures                                      | Confounders                          |
|-------------------------|-----------------------------------|-----------------------------------------------|--------------------------------------|
| **Main Analysis 1**     | PCR Test Positivity              | IMD score                                     | Age                                  |
|                         |                                   | % BAME population                             | Urbanicity                           |
|                         |                                   |                                               | Vaccination Uptake                   |
| **Main Analysis 2**     | Prevalence Estimate              | IMD score                                     | Age                                  |
|                         |                                   | % BAME population                             | Urbanicity                           |
|                         |                                   |                                               | Vaccination Uptake                   |
| **Main Analysis 3**     | PCR Test Positivity              | IMD score                                     | Age                                  |
|                         |                                   | % Black population                            | Urbanicity                           |
|                         |                                   | % South-Asian population                       | Vaccination Uptake                   |
|                         |                                   | % Other BAME population                        |                                      |
| **Main Analysis 4**     | Prevalence Estimate              | IMD score                                     | Age                                  |
|                         |                                   | % Black population                            | Urbanicity                           |
|                         |                                   | % South-Asian population                       | Vaccination Uptake                   |
|                         |                                   | % Other BAME population                        |                                      |
| **Sensitivity Analysis 1** | PCR and LF Test Positivity   | IMD score                                     | Age                                  |
|                         |                                   | % Black population                            | Urbanicity                           |
|                         |                                   | % South-Asian population                       | Vaccination Uptake                   |
|                         |                                   | % Other BAME population                        |                                      |
| **Sensitivity Analysis 2** | PCR Test Positivity               | IMD score                                     | Age                                  |
|                         |                                   | % Black population                            | Urbanicity                           |
|                         |                                   | % South-Asian population                       | Vaccination Uptake                   |
|                         |                                   | % Other BAME population                        | Lockdown                              |
| **Sensitivity Analysis 3** | PCR Test Positivity               | IMD score                                     | Age                                  |
|                         |                                   | % Black population                            | Urbanicity                           |
|                         |                                   | % South-Asian population                       |                                      |
|                         |                                   | % Other BAME population                        |                                      |

Table 1: Summary of outcomes and covariates included in the all the Analyses (Main and Sensitivity). Random effect specification remains the same as specified in Section 1.
2 Additional figures

Figure 1: Spatial distribution of the two main exposures, IMD score and proportion of resident BAME population, by LTLA.

Figure 2: Spatial distribution of the proportion of resident BAME population in the total resident population for the three subgroups considered in the analyses: Black, South Asian and Other BAME.
Figure 3: Posterior median and 95% credible intervals (CI) for the temporal trend included in the spatio-temporal model on the logit scale for the main analysis with disaggregated BAME subgroups. Top panel refers to the model with test positivity as outcome, bottom panel to debiased prevalence.
Figure 4: Posterior median and 95% credible intervals (CI) for the temporal trend included in the spatio-temporal model on the logit scale for the main analysis with disaggregated BAME subgroups. Sensitivity analysis 1 (yellow - top panel) uses combined PCR and LF test positivity as outcome measure while sensitivity analyses 2 (green - central panel) and 3 (light blue - bottom panel) use only PCR test positivity. Sensitivity analysis 1 includes IMD, BAME%, age, urbanicity and vaccination as covariates; sensitivity analysis 2 includes lockdown in addition to these covariates; sensitivity analysis 3 removes vaccination from the covariates. When including LFT in the analysis the summer peak is lower than the two winter ones (top panel). This might potentially reflect the change in testing policy: LFT became more widespread as businesses were encouraged to sign up for a testing scheme for their staff, increasing the denominator of the positivity rate. The summer peak is also somewhat less pronounced when removing vaccine uptake from the confounders; in this case the temporal random effect implicitly accounts for the fact that higher vaccination rates lead to a reduction in infections (bottom panel). Finally, the time pattern does not change when we include a lockdown indicator, suggesting that the temporal component already captures much of this effect (central panel).
Figure 5: Posterior median (in black) of the fixed effects ($\beta$, $\gamma$) for the main analysis on test positivity and debiased prevalence and corresponding 95% credible intervals (CI). Results are reported on the Odds Ratio (OR) scale.
Figure 6: Test positivity for profiles of ethnicity (disaggregated by sub-group) and deprivation. Each tile represents the average weekly test positivity over the entire period of analysis, obtained as output of the model including ethnicity, IMD, confounders and the spatio-temporal correlation structure. In parenthesis we report the relative change in the outcome between each profile and the reference, defined as low deprivation and low Black, South Asian and Other BAME population.
Figure 7: Debiased prevalence for profiles of ethnicity (disaggregated by sub-group) and deprivation. Each tile represents the average weekly debiased prevalence over the entire period of analysis, obtained as output of the model including ethnicity, IMD, confounders and the spatio-temporal correlation structure. In parenthesis we report the relative change in the outcome between each profile and the reference, defined as low deprivation and low Black, South Asian and Other BAME population.
Figure 8: Time-varying test positivity (top) and debiased prevalence (bottom) for profiles of ethnicity disaggregated by BAME subgroups (Black, South Asian, Other). Each line represents the monthly median odds ratio for each profile relative to a population with a low percentage of all BAME subgroups. The mean odds ratio is the output from the model including ethnicity, IMD, confounders and the spatio-temporal correlation structure.
Figure 9: Posterior median (in black) of the fixed effects ($\beta, \gamma$) for the sensitivity analyses and corresponding 95% credible intervals (CI). Results are reported on the Odds Ratio (OR) scale. Sensitivity analysis 1 (yellow) uses combined PCR and LF test positivity as outcome measure while sensitivity analyses 2 (green) and 3 (light blue) use only PCR test positivity. Sensitivity analysis 1 includes IMD, BAME%, age, urbanicity and vaccination as covariates; sensitivity analysis 2 includes lockdown in addition to these covariates; sensitivity analysis 3 removes vaccination from the covariates. The effect of the disaggregated BAME subgroups is robust with respect to the testing strategy and does not change when we include Lateral Flow (LF) tests in the outcome while the effect of IMD is stronger when considering combined PCR and LF test positivity (sensitivity analysis 1).
References

[1] Julian Besag, Jeremy York, and Annie Mollié. Bayesian image restoration, with two applications in spatial statistics. *Annals of the institute of statistical mathematics*, 43(1):1–20, 1991.

[2] Leonhard Knorr-Held. Bayesian modelling of inseparable space-time variation in disease risk. *Statistics in medicine*, 19(17-18):2555–2567, 2000.

[3] George Nicholson, Brieuc Lehmann, Tullia Padellini, Koen B Pouwels, Radka Jersakova, James Lomax, Ruairidh E King, Ann-Marie Mallon, Peter J Diggle, Sylvia Richardson, Marta Blangiardo, and Chris Holmes. Local prevalence of transmissible sars-cov-2 infection: an integrative causal model for debiasing fine-scale targeted testing data. *medRxiv*, 2021.

[4] Andrea Riebler, Sigrunn H Sorbye, Daniel Simpson, and Håvard Rue. An intuitive bayesian spatial model for disease mapping that accounts for scaling. *Statistical methods in medical research*, 25(4):1145–1165, 2016.

[5] Daniel Simpson, Håvard Rue, Andrea Riebler, Thiago G Martins, Sigrunn H Sorbye, et al. Penalising model component complexity: A principled, practical approach to constructing priors. *Statistical science*, 32(1):1–28, 2017.

[6] Sigrunn Holbek Sorbye and Håvard Rue. Scaling intrinsic gaussian markov random field priors in spatial modelling. *Spatial Statistics*, 8:39–51, 2014.