Geospatial Health
Geospatial Health 2026; 21:1477
doi:10.4081/gh.2026.1477

A Bayesian spatial ecological analysis of social conditions, infection burden and COVID-19 mortality across mainland Scottish councils 2020-2021

Zhiyuan Luo, Yancheng Liu, Wenqi Xu

School of Sociology and Anthropology, Sun Yat-sen University, Guangzhou, China

Abstract

The COVID-19 pandemic has generated substantial spatial and social inequalities in mortality, yet within-country council-level variations remain incompletely understood. Scotland offers a useful case, with a severe epidemic but rich administrative data and marked socio-economic gradients. We examined how infection burden, hospital-related indicators and selected area-level social indicators were associated with COVID-19 mortality across mainland Scotland 2020-2021. We assembled a two-year panel for 29 mainland councils, using deaths involving COVID-19, population denominators, COVID-19 testing and hospital activity, council-level indicators of marriage, unemployment, smoking cessation and migration. A Bayesian Poisson log-linear model with conditional autoregressive random effects was used to estimate area-level relative risks and covariate associations. Given the short panel, the analysis is interpreted primarily as a spatial ecological analysis conducted over two successive pandemic years. Crude mortality showed consistently higher COVID-19 mortality in the central belt than in many northern and rural councils. After adjustment, spatial differences remained but were modest: no council in 2021 had a posterior probability above 0.10 of exceeding the national mean mortality risk. Positive test burden was positively associated with mortality and is interpreted primarily as a proximal epidemiological indicator of infection burden. Social determinants were also important. Higher unemployment was associated with increased risk, whereas higher marriage counts were linked to lower mortality. Smoking quit rates showed a positive association with deaths, likely reflecting residual confounding by underlying deprivation and historical smoking prevalence. Hospital utilisation and migration indicators showed weaker and more uncertain effects. Together, these findings indicate that crude spatial disparities in COVID-19 mortality across mainland Scottish councils became more moderate after adjustment. Given the ecological design and the two-year panel, the findings should be interpreted as area-level associations rather than as evidence of specific local social mechanisms.

Key words: COVID-19 mortality; Bayesian disease mapping; council-level analysis; ecological analysis, Scotland.

Correspondence: Yancheng Liu, School of Sociology and Anthropology, Sun Yat-sen University, Guangzhou, China. E-mail: lyc001120@163.com

Introduction

The COVID-19 pandemic has generated an unprecedented global health and mortality shock, exposing vulnerabilities in health systems and societies worldwide. In 2023, the World Health Organization (WHO) estimated that at least six million deaths have been directly attributed to COVID-19, with total excess mortality likely to be substantially higher (WHO, 2023). Beyond acute mortality, emerging cohort evidence from England documents persistent multi-system sequelae among large populations, underlining that the health burden is both immediate and long-term (Atchison *et al.*, 2023). Within Europe, the United Kingdom has been a major COVID-19 epicentre: age-standardised mortality in the UK has remained elevated relative to pre-pandemic trends, even after mass vaccination (O'Dowd, 2023). Public debate and much empirical work have focused on England, while the Scotland situation has often been treated as secondary. Yet official statistics show substantial variation in COVID-19 mortality within Scotland itself (National Records of Scotland, NRS, 2023; National Health Service, NHS, Scotland, 2023), suggesting that a more granular spatial perspective is needed. COVID-19 did not strike into a social vacuum but interacted with pre-existing health inequalities. A large body of work now characterises the pandemic as an "unequal pandemic", with mortality and morbidity strongly stratified by socio-economic position, ethnicity and occupational exposure (Patel *et al.*, 2020; Bambra, Lynch and Smith, 2021). In the UK, more deprived populations, people in insecure or public-facing work, and communities already bearing a high burden of chronic disease experienced disproportionate high risks. These social gradients are not only individual-level characteristics; they are also spatially organised. Studies show that COVID-19 mortality is higher in more deprived neighbourhoods and that these disadvantages cluster in space, producing persistent spatial inequalities in risk (Sun *et al.*, 2021; Griffith *et al.*, 2021). Such findings reinforce a place-based understanding of the pandemic, in which geographical context – including local deprivation, labour-market conditions and healthcare capacity – shapes both exposure to infection and vulnerability to severe outcomes. Classic compartmental models, such as the Susceptible-Infectious-Recovered (SIR) framework, treat populations as homogeneous within compartments and largely abstract from social and spatial heterogeneity. Recent methodological developments have begun to challenge this abstraction. Nikolaou (2022) shows that extensions of the standard SIR model which incorporate heterogeneity and feedbacks

can substantially improve the prediction of epidemic peaks and timing. At a more conceptual level, Hébert-Dufresne *et al.* (2020) demonstrate that interacting contagions and socially reinforced behaviours can generate macroscopic patterns that mimic simple contagion processes, even though the underlying dynamics are mediated by complex social networks. Together, these contributions highlight the limitations of purely biomedical or homogeneous-mixing perspectives and point towards the need to embed infectious disease modelling within the social structure and spatial organisation of populations.

Spatial epidemiology provides a natural bridge between social determinants of health and infectious disease dynamics. Tobler's first law of geography – stating that "everything is related to everything else, but near things are more related than distant things" (Tobler, 1970) – has long guided geospatial analysis. In the context of health, this principle implies that neighbouring areas are likely to share exposures, health-care resources and policy environments, and to be linked through commuting and social interaction networks. The disease-mapping literature has operationalised this insight by explicitly modelling spatial autocorrelation in area-level health outcomes (Lawson, 2013; Moraga, 2019). Conditional Autoregressive (CAR) priors and related structures, introduced by Besag, York and Mollié (BYM) (1991), allow random effects for adjacent areas to be smoothed towards each other, borrowing strength and providing coherent estimates of risk in small populations. Applications of Bayesian hierarchical models with CAR components to COVID-19 have documented marked spatial clustering of incidence and mortality, and have shown how socio-economic factors help explain these patterns (Kianfar *et al.*, 2023; Sun *et al.*, 2021). Despite this growing body of spatial epidemiological work, Scotland remains relatively under-studied. National-level comparisons across the four UK nations emphasise that Scotland has experienced a somewhat lower cumulative death rate than England, but still a severe local burden (Scobie, 2022; O'Dowd, 2023). These aggregate contrasts, however, conceal substantial variation between Scottish local areas in demographic composition, deprivation, smoking prevalence, migration flows and unemployment. Scottish administrative and health data already document sharp geographical gradients in mortality and morbidity (National Records of Scotland, 2023; NHS Scotland, 2023), yet there has been limited effort to examine COVID-19 mortality across mainland Scottish councils using modern Bayesian disease-mapping methods. We know little about whether COVID-19 mortality in Scotland is spatially autocorrelated beyond what would be expected from demographic structure alone, and to what extent social and behavioural determinants can account for any observed clustering. Bayesian hierarchical spatial models provide an appropriate framework to address these questions. Building on the tradition of BYM (1991) and subsequent developments in Bayesian disease mapping (Lawson, 2013; Moraga, 2019), we employed models that combine fixed effects for measured covariates with spatially structured random effects for local administrative areas. The spatial component is specified using CAR-type priors that encode adjacency between local areas, allowing neighbouring units to share information while retaining the possibility of sharp local discontinuities (BYM, 1991; Leroux *et al.*, 2000). Implementation is carried out using the CAR Bayes and CAR Bayes 'ST' packages in R (Lee, 2013; Lee *et al.*, 2018), which are specifically designed for areal spatial and spatio-temporal data. This class of models has been widely used to estimate disease risk and quantify residual spatial dependence after adjusting for observable determinants (Lawson, 2013; Moraga, 2019; Kianfar *et al.*, 2023), making it well suited to the Scottish COVID-19 context. Against this background, our study had three aims. First, we described the crude spatial pattern of COVID-19 mortality across mainland Scottish councils in 2020 and 2021. Second, we examined whether area-level variation in mortality was associated with infection burden and selected social, behavioural, and hospital-related indicators. Third, we assessed the extent to which residual spatial variation remained after these observed factors were taken into account. Because the analysis is based on a two-year ecological panel, the study is intended to identify area-level associations rather than fine-scale local mechanisms. The overall study design and modelling strategy are summarized in Figure 1.

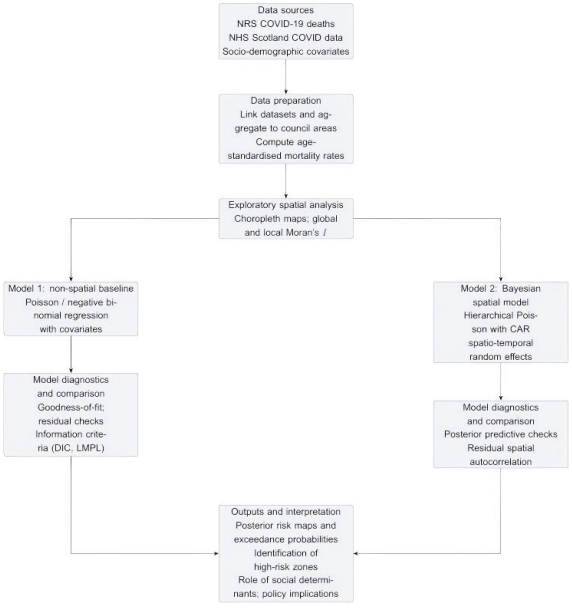

Figure 1. Overview of the study design and modelling strategy for analysing spatial inequalities in COVID-19 mortality in Scotland.

Materials and Methods

Study area

The study area comprises the 29 mainland council areas (k) of Scotland. The three island councils Orkney Islands, Shetland Islands and Na h-Eileanan Siar were excluded because they do not form part of the contiguous mainland and therefore have no first-order neighbours in the spatial adjacency structure, which would complicate the construction of a neighbourhood matrix. Polygon boundary data for Scottish council areas were obtained from the Scottish Spatial Data Infrastructure and processed in QGIS to create a consistent areal layer for analysis (Scottish Government, 2019). After removing the three island councils, the final spatial dataset consists of k= 29 non-overlapping areal units indexed by k = 1, ... ,29. The geographical overview is presented in Figure 2.

Spatial dependence was encoded through a binary adjacency matrix W of dimension 29 × 29, where w_{ij} = 1 if council areas i and j share a common boundary and w_{ij} = 0 otherwise (Besag *et al.*, 1991). Because the 29 areas in Figure 2 form a single landmass, each has at least one neighbour, so W defines a well-connected spatial graph (Scottish Government, 2019). This matrix underpins the CAR prior used in the hierarchical model.

Data sources and variables

We assembled a short two-year panel for the first two pandemic years 2020 and 2021 (T = 2). For each council area k and year t, the outcome Y_{kt} is the observed number of deaths involving COVID-19, obtained from NRS (2023). These data provide, for each council and year, the registered number of deaths where COVID-19 is mentioned on the death certificate. Mid-year population estimates by council and year from NRS were used as denominators to construct expected deaths. To adjust for differences in population size and age structure, we computed expected deaths E_{kt} under a reference Scottish mortality pattern. Age-specific COVID-19 mortality rates for Scotland were applied to the age distribution of each council area, and the resulting expected counts were summed to obtain E_{kt} for area k and year t. The Standardised Mortality Ratio (SMR) is then

$$SMR_{kt} = (Y_{kt} / E_{kt}) \quad \text{Eq.1}$$

where Y_{kt} is observed deaths; E_{kt} is the number of expected deaths in council k relative to the national average in year t (Lawson, 2013; Moraga, 2019). For example, if Glasgow in 2020 has Y_{kt} = 600 observed deaths (which were interpreted as the number of COVID-19 deaths); and E_{kt} = 400 expected deaths. Then SMR_{kt} = 1.5, meaning a 50% higher mortality rate than the Scottish average. These SMRs were used descriptively (Figure 3) to motivate the formal modelling, but all inferential statements were based on model-based relative risks because the crude SMRs can be unstable in areas with small populations. These descriptive measures are used to summarise broad spatial differences in mortality rather than to support finer-scale mechanistic claims. Information on testing and health-care activity related to COVID-19 was sourced from NHS Scotland open data (NHS Scotland, 2023). For each council area and year, we extracted the total number of positive tests, the number of PCR-only cases, the number of patients, and the average length of hospital stay. Total positive tests are treated primarily as a proximal epidemiological indicator of infection burden, whereas hospital-related indicators are interpreted more cautiously as markers of hospital utilisation, severity or service pressure rather than as direct measures of health-care capacity. To characterise broader area-level conditions we compiled indicators from NRS including marriages, households, smoking quit rates, quit attempts, unemployment rates and migration counts (NRS, 2023). Marriage counts and smoking quit rates are treated as exploratory ecological proxies rather than as direct measures of specific social

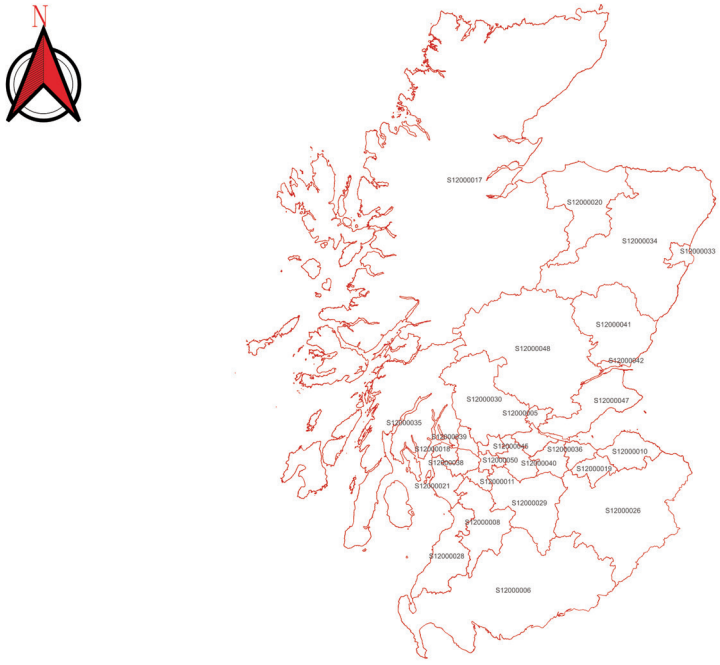

Figure 2. Council area boundaries for mainland Scotland. The three island councils are excluded to ensure a contiguous spatial graph for disease mapping.

mechanisms. Unemployment is treated as an area-level socio-economic indicator, while migration is interpreted more cautiously as a contextual demographic indicator. All data were linked using the council feature code, producing a panel with one record per council–year combination (29 areas × 2 years = 58 observations). Skewed count and rate variables, including total positive tests, patients per 100,000, average length of stay, and marriage counts, were log-transformed after adding a small constant (0.5) to avoid taking the logarithm of zero. Continuous rates such as unemployment, migration, and smoking quit rates were standardised to mean zero and unit variance. These transformations were applied to improve comparability across council-year observations and to reduce the influence of skewed distributions. The resulting p-dimensional covariate vector for area k and year t is denoted as:

$$\mathrm{x}kt = (xkt1,...,xktp)^{\top} \quad \text{Eq. 2}$$

and includes epidemiological indicators, hospital-related indicators, and selected area-level social and behavioural proxies. To make these distinctions fully transparent, we also added *Supplementary Table S1*, which summarises variable definitions, transformations, and analytical roles.

Statistical model

We model COVID-19 mortality using a Bayesian hierarchical Poisson log– linear model for areal count data (Besag *et al.*, 1991; Lawson, 2013; Moraga, 2019). For each council $k = 1, \ldots, 29$ and year $t = 1,2$ we assume:

$$Y_{kt} \mid \theta_{kt} \sim \text{Poisson}(E_{kt}\,\theta_{kt}) \quad \text{Eq. 3}$$

where θ_{kt} is the relative risk in area k and year t, and E_{kt} is the expected number of COVID-19 deaths given the local age structure." If $\theta_{kt} = 1$, council k has exactly the expected mortality in year t; if $\theta_{kt} = 1.5$, mortality is 50% higher than expected. We relate the log-relative risk to the covariates and random effects through:

$$\log \theta_{kt} = \alpha + \mathrm{x}^{\top}_{kt}\beta + \psi_{kt} \quad \text{Eq. 4}$$

where α is an intercept; $\beta = (\beta_1,...,\beta_p)^{\top}$ a vector of regression coefficients linking the covariates to mortality; and ψ_{kt} a spatio-temporal random effect for area k and year t.

Substituting Eq. 4 into Eq. 3, the linear predictor for the Poisson mean becomes:

$$\eta_{kt} = \log(E_{kt}) + \alpha + \mathrm{x}^{\top}_{kt}\beta + \psi_{kt} \quad \text{Eq. 5}$$

so that $\log(E_{kt})$ acts as a fixed offset. Under this formulation, $\exp(\beta_j)$ represents the multiplicative change in COVID-19 mortality associated with a one-standard-deviation increase in covariate x_{ktj}, holding other factors constant. The regression coefficients are given a weakly informative multivariate Gaussian prior:

$$\beta \sim \mathrm{N}(0,\sigma_{\beta}^{2}\mathrm{I}_{p}) \quad \text{Eq. 6}$$

with a large variance σ_{β}^{2} to reflect limited prior knowledge while providing some regularisation (Lawson, 2013; Gelman *et al.*, 2013). Residual spatial and temporal structure in COVID-19 mortality is captured by the random effects ψ_{kt}. We adopted the spatio-temporal autoregressive CAR model implemented in the 'ST.CARar' function of the 'CARBayesST' package in R (Rushworth *et al.*, 2017; Lee *et al.*, 2018; Lee, 2020). If you let $\varphi_t = (\varphi_{1t},...,\varphi_{Kt})^{\top}$ denote the vector of random effects at time t and set $\psi_{kt} = \varphi_{kt}$, the temporal evolution of φ_t is specified as:

$$\varphi_1 \sim \mathrm{N}(0,\tau^2 Q(\mathrm{W},\rho_S)^{-1}) \quad \text{Eq. 7}$$

$$\varphi_t \mid \varphi_{t-1} \sim \mathrm{N}\left(\rho_T\varphi_{t-1},\tau^2 Q(\mathrm{W},\rho_S)^{-1}\right),\ t = 2,...,T \quad \text{Eq. 8}$$

where ρ_S is a spatial dependence parameter; ρ_T a temporal autoregressive parameter; and τ^2 controls the overall variance. The precision matrix:

$$Q(\mathrm{W},\rho_S) = \rho_S\{\mathrm{diag}(\mathrm{W}1) - \mathrm{W}\} + (1-\rho_S)\mathrm{I}_K \quad \text{Eq. 9}$$

is the Leroux CAR structure, which smoothly interpolates between independent random effects ($\rho_S \approx 0$) and a strong spatial-

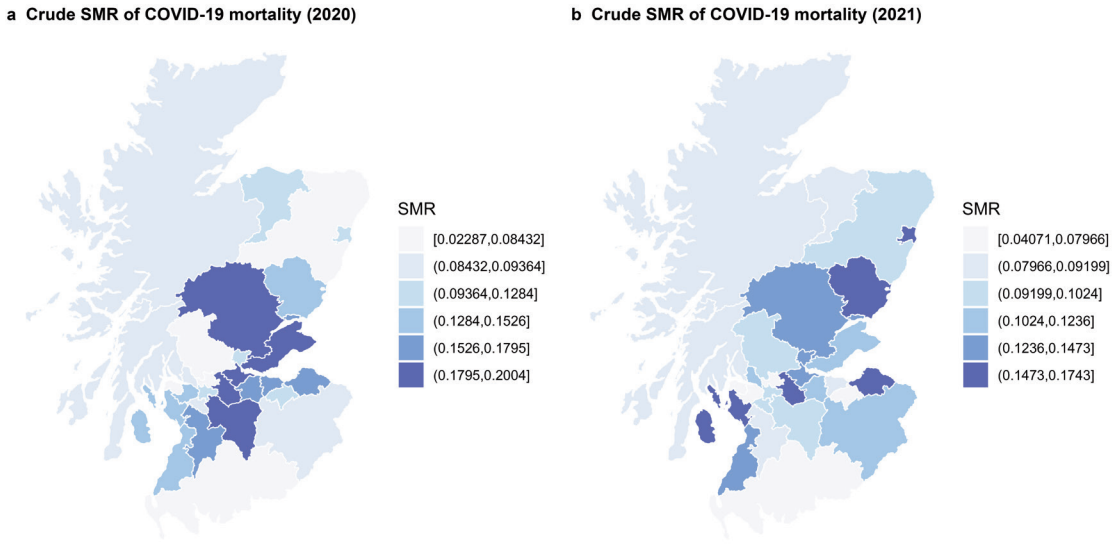

Figure 3. Crude standardized mortality ratios (SMRs) for deaths involving COVID-19 in the 29 mainland council areas of Scotland. (**a**) in 2020; (**b**) in 2021.

ly smoothed surface (ρ_S close to 1) (Leroux *et al*., 2000).

In the Scottish data, this structure allows councils with shared boundaries (for example, Glasgow and its neighbouring authorities) to borrow strength from each other when estimating COVID-19 mortality risk. We also examined the dispersion of the council–year death counts using non-spatial negative-binomial benchmarks and posterior predictive checks (see Model performance and sensitivity analysis). These diagnostics indicated that, once spatio-temporal random effects are included, the Poisson log–linear specification with an offset provides an adequate description of the data, and we therefore retained it as our primary model. The hyper-parameters are assigned weakly informative priors. Following previous work on spatio-temporal disease mapping (Rushworth *et al.*, 2017; Lawson, 2013), we used independent Uniform (0,1) priors for ρ_S and ρ_T and an inverse-gamma prior for the variance:

$$\tau^2 \sim \text{Inverse-Gamma}(a,b) \qquad \text{Eq. 10}$$

with $(a,b) = (1,0.01)$, which yields a diffuse but proper prior.

Estimation and sensitivity analysis

All models were fitted in R using the 'CARBayesST' package (version 4.0) (Lee *et al.*, 2018; Lee, 2020). We used the 'ST.CARar' function with a Poisson likelihood, a log link and the expected deaths E_{kt} as an offset, following standard practice in Bayesian disease mapping (Lawson, 2013; Moraga, 2019). Two parallel Markov chain Monte Carlo (MCMC) chains were run, each with a burn-in of 100,000 iterations followed by 800,000 sampling iterations thinned every 2,000 steps. Initial exploratory runs indicated relatively slow mixing for some hyper-parameters, so we adopted longer chains and thinning to ensure adequate effective sample sizes for posterior summaries. These settings match the complexity of the model relative to the 58 council–year observations and provided a large effective sample size for posterior summaries.

Convergence was assessed using trace plots and Gelman–Rubin diagnostics for all regression coefficients and key hyperparameters (ρ_S, ρ_T and τ^2) (Gelman *et al.*, 2013). Posterior medians and 95% credible intervals were computed for each coefficient and for the area–year-specific relative risks θ_{kt}. These estimates were then mapped to produce smoothed risk surfaces and to identify high-risk council areas. As a simple sensitivity analysis, we also fitted a reduced 'ST.CARar' model that included only the COVID-19 testing and hospital indicators as covariates, omitting the social variables. The Deviance Information Criterion (DIC) and Log Marginal Predictive Likelihood (LMPL) from 'CARBayesST' were used to compare model fit and predictive performance between the full and reduced specifications (Lee *et al*., 2018; Lee, 2020). The full model, which combines epidemiological, hospital-related, and selected area-level social indicators, provided the basis for the main results. The sensitivity analyses are treated as internal checks across related model forms rather than as definitive resolution of broader conceptual limitations.

Results

Crude spatial pattern of COVID-19 mortality

COVID-19 mortality in Scotland displayed clear spatial variation over the first two pandemic years. Figure 3 maps the crude SMRs for deaths involving COVID-19 in 2020 and 2021 based on counts and mid-year population denominators (NRS, 2023). In both years the central belt – including Glasgow City, North and South Lanarkshire and neighbouring councils – had higher crude SMRs than many rural and northern areas, whereas several mainland councils in the north recorded comparatively low mortality. Between 2020 and 2021 the overall national burden evolved in line with successive waves of infection and changing public health measures (Scobie, 2022; WHO, 2023), but the broad geography of risk remained relatively stable. Some councils, particularly in the North-east, moved from medium to higher SMR, while others showed reductions, reflecting local differences in timing of waves, vaccine roll-out and health-system strain (O'Dowd, 2023). These descriptive patterns motivate the use of formal spatial models to

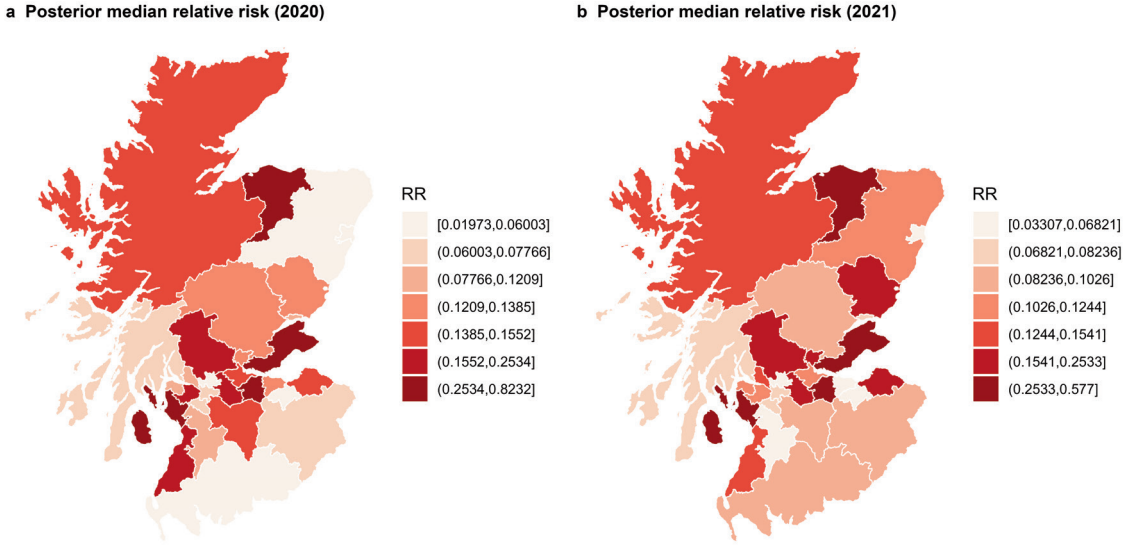

Figure 4. Posterior median relative risk (RR) of COVID-19 mortality from the spatio-temporal ST.CARar model (**a**) in 2020 and (**b**) in 2021.

separate persistent geographic risk from random variation in council-level counts.

SMR is defined as the ratio of observed to expected deaths (based on Scotland-wide age-specific COVID-19 mortality rates); values above 1 indicate higher-than-expected mortality relative to the national pattern. Darker shading in Figure 3 indicates higher SMR. Council boundaries follow the 2019 Scottish Spatial Data Infrastructure layer, and mortality counts and denominators are drawn from NRS. Class breaks are based on quantiles of the SMR distribution within each year.

Model-based relative risks and spatio-temporal structure

The Bayesian 'ST.CARar' model yielded smoothed estimates of council-specific Relative Risks (RR) after accounting for expected deaths and covariates. Figure 4 presents the posterior median RR for 2020 and 2021. Compared with the crude SMR, the modelled surfaces were smoother, with extreme values in sparsely populated councils shrunk towards the national mean. Elevated risks nonetheless persisted in parts of the central belt in both years, while many northern councils remain below unity, indicating that the spatial pattern is unlikely to be attributable solely to chance, even after adjustment for testing, hospital utilisation and social context. Posterior summaries for the variance and autocorrelation parameters (Table 1) indicate moderate residual spatio-temporal structure. The variance parameter τ^2 has a posterior mean of 0.133 (95% CrI 0.073-0.243), implying appreciable unexplained heterogeneity in log-risk between-councils and years. The spatial dependence parameter ρ_S had a mean of 0.134 (95% CrI 0.004-0.721) suggesting positive but uncertain spatial autocorrelation consistent with CAR smoothing (Leroux *et al*., 2000; Rushworth *et al*., 2017). The temporal autoregressive parameter ρ_T has a mean of 0.343 (95% CrI 0.024-0.803) indicating that council-specific risks in 2021 tended to move in the same direction as in 2020, but with substantial year-to-year variation. Together, these results support the use of the spatio-temporal model while also indicating that dependence is not overwhelming in either dimension. However, as the temporal dimension comprised only two annual observations, it should be interpreted cautiously, with the model primarily used to support a spatial ecological analysis conducted over two successive pandemic years.

The Relative Risks (RRs) are expressed relative to the Scotland-wide mean in each year and are derived from a Poisson log-linear model with expected deaths as an offset and CAR random effects. Darker shading in Figure 4 indicates higher model-based RR. Both panels are shown with the same colour scale and class breaks to facilitate visual comparison across years.

Covariate effects on COVID-19 mortality

Several covariates were associated with area-level COVID-19 mortality (Table 2). After adjustment for all predictors and spatio-temporal random effects, the log-transformed total number of positive tests showed a positive association with mortality (posterior mean log-RR 0.211; 95% CrI 0.051-0.358; RR 1.24, 95% CrI 1.05-1.43). Councils with higher apparent infection burden, as captured by positive tests, tended to experience higher death rates, consistent with evidence linking community transmission to severe outcomes (Nikolaou, 2022; Atchison *et al*., 2023). Social indicators also contributed to explaining variation. Higher marriage counts were associated with lower mortality (mean log-RR −0.326; 95% CrI −0.494–−0.180; RR 0.72, 95% CrI 0.61-0.84). In this study, marriage counts were interpreted as an exploratory area-level proxy and should not be read as direct evidence of a specific social mechanism. In contrast, both the standardised unemployment rate and the standardised quit-smoking rate showed positive associations, although the latter should be interpreted cautiously as an ecological association that may reflect residual confounding. A one-standard-deviation increase in unemployment was associated with a 16% increase in mortality risk (RR 1.16, 95% CrI 1.01-1.30) and higher smoking quit rates were associated with an 18% increase (RR 1.18, 95% CrI 1.05-1.32). Migration rates and hospital-related measures had more uncertain associations, with credible intervals including the null. Overall, these results indicate that, beyond direct epidemiological indicators, selected social and behavioural indicators contribute to the area-level distribution of

Table 1. Posterior summaries for the variance and autocorrelation parameters in the ST.CARar model.

Parameter	Mean	2.5%	97.5%
τ^2	0.133	0.073	0.243
ρ_S	0.134	0.004	0.721
ρ_T	0.343	0.024	0.803

Table 2. Posterior covariate effects from the ST.CARar model.

Covariate	Mean	2.5%	97.5%	RR	RR-2.5%	RR-97.5%
Total positive tests	0.211	0.051	0.358	1.235	1.053	1.431
Hospital patients	-0.862	-2.066	0.333	0.422	0.127	1.395
Mean length of hospital stay	0.223	-0.221	0.627	1.249	0.802	1.873
Number of marriages	-0.326	-0.494	-0.180	0.722	0.610	0.835
Unemployment rate	0.145	0.006	0.263	1.156	1.006	1.301
Net migration rate	0.064	-0.026	0.157	1.066	0.975	1.170
Smoking quit rate	0.162	0.048	0.278	1.176	1.049	1.321

RR, relative risk; coefficients are on the log-relative risk scale; RR columns show exponential values.

COVID-19 mortality in Scotland, echoing broader evidence on socio-economic disadvantage and pandemic outcomes (Rigby, 2021; Scobie, 2022).

National-mean comparison and uncertainty in 2021

To focus on the most recent year, Figure 5 compares council-specific risks in 2021 with the Scotland-wide mean. Panel a shows the posterior median RR relative to the national mean. Several councils in the centre and the Southwest had RRs above 1.2, whereas many northern and rural councils were below the mean. When uncertainty was considered (panel b), only a small number of councils had 95% CrIs entirely below the national mean, with none credibly above it. Exceedance probabilities $P(RR > 1)$ were below 0.10 for all councils (panel c), indicating limited posterior support for clearly defined hotspots of excess mortality in 2021 once covariates and spatio-temporal smoothing were incorporated. These results suggest that, by the second year of the pandemic, council inequalities in COVID-19 mortality within mainland Scotland were modest after adjustment for demographic, testing and social characteristics, in contrast to the pronounced regional disparities observed across all four UK nations.

Model performance and sensitivity analysis

Model performance and robustness were examined through a set of comparisons. A spatial-only CAR model and a non-spatial Poisson log–linear model were fitted using the same covariates and offset. For key predictors such as the log-transformed count of positive COVID-19 tests, the estimated effects of the standardised unemployment rate and the log-transformed number of marriages were similar across specifications (Table 3). For example, the log-transformed total positive tests coefficient was 0.21 (95% CrI 0.05-0.36) in the ST.CARar model, 0.23 (0.07-0.39) in the spatial-only CAR model and 0.26 (0.12-0.40) in the non-spatial Poisson model. The direction and magnitude of the unemployment and marriage effects were likewise consistent, suggesting that the substantive conclusions were not sensitive to the precise form of spatial and temporal dependence. Differences in the estimated RR surfaces between the spatial-only and spatio-temporal models were small (Figure 6). Across all councils and both years, absolute changes in posterior median RR were typically below 0.1, with slightly larger shifts for a few central councils. Red shading (positive ΔRR) corresponds to higher estimated risk under the spatial-only model, while blue shading indicates higher risk under the spatio-temporal

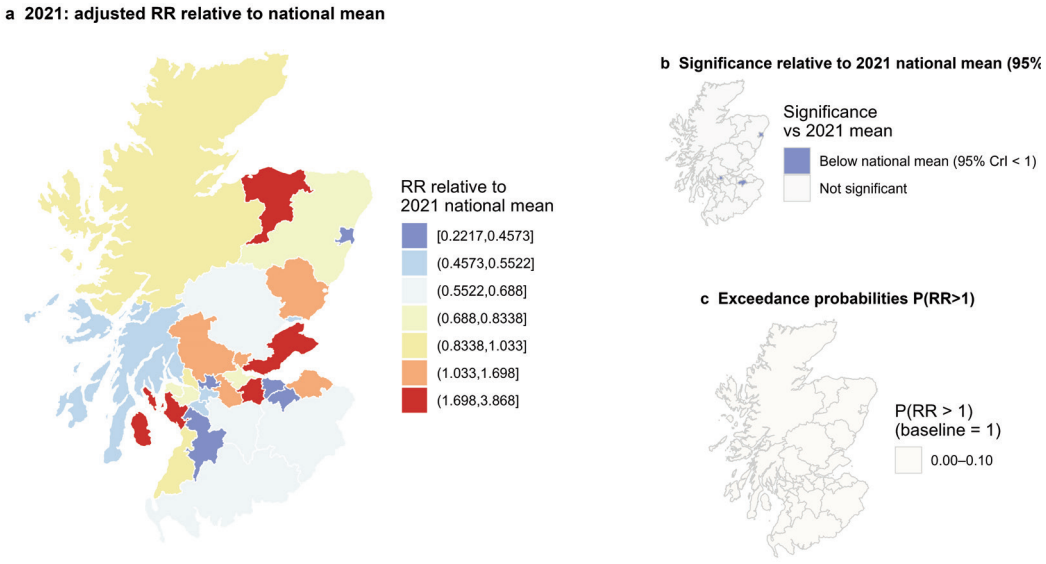

Figure 5. Modelled COVID-19 mortality risk in 2021 relative to the Scotland-wide mean from the ST.CARar model. **a**) Posterior median RR for each council area; **b**) Councils with 95% credible intervals for RR entirely below the national mean (blue) versus councils whose intervals include the mean (grey); no councils have 95% credible intervals entirely above the mean; **c**) Posterior exceedance probabilities $P(RR>1)$ for each council, with a reference line at 0.50. Together, the panels summarise the magnitude and uncertainty of council-level deviations from the national mean mortality risk in 2021.

Table 3. Comparison of key covariate effects across models.

Covariate	ST.CARar β (95% CrI)	Spatial CAR β (95% CrI)	Poisson GLM β (95% CI)
Total positive COVID-19 tests (log-transformed count)	0.21 (0.05, 0.36)	0.23 (0.07, 0.39)	0.26 (0.12, 0.40)
Unemployment rate (standardised)	0.15 (0.01, 0.26)	0.16 (0.02, 0.29)	0.18 (0.05, 0.31)
Marriage registrations (log-transformed count)	-0.33 (-0.49, -0.18)	-0.35 (-0.51, -0.20)	-0.32 (-0.48, -0.17)

β denotes the regression coefficient on the log-relative-risk scale. For the Bayesian models, values are posterior means with 95% credible intervals (CrI); for the Poisson generalised linear model (GLM), values are maximum-likelihood estimates with 95% confidence intervals (CI). Positive values indicate higher mortality risk, whereas negative values indicate lower mortality risk.

model. The modest magnitude of these differences implies that incorporating temporal autocorrelation improves fitness without materially altering the spatial pattern of risk. MCMC diagnostics support model adequacy. Effective sample sizes for fixed effects exceeded 500, and Geweke statistics lie within the conventional interval (−2,2) for all coefficients. Trace plots of regression parameters and hyperparameters showed good mixing and no evidence of lack of convergence, in line with standard guidance for Bayesian hierarchical models (Gelman *et al*., 2013). The narrow range of differences shown in Figure 6 suggests that allowing for temporal autocorrelation improves model fit with only modest impact on the estimated spatial pattern of COVID-19 mortality.

Next, to assess overdispersion and the appropriateness of the Poisson assumption, we first compared a non-spatial Poisson Generalised Linear Model (GLM) with a non-spatial negative-binomial GLM using the same covariates and offset. The Poisson model showed clear evidence of overdispersion (Pearson dispersion 14.9), and the negative-binomial GLM provided a substantially better fit (AIC 635.6 *vs.* 1163.6) (Table 4). This supports the presence of extra-Poisson variation in the council–year death counts when spatial and temporal structure are ignored.

We then examined whether the spatio-temporal ST.CARar model adequately absorbed this extra-Poisson variability using a paired posterior predictive check based on a Pearson-type dispersion discrepancy. For each MCMC iteration, we computed the discrepancy $D_{obs}^{(s)}$ for the observed counts and $D_{rep}^{(s)}$ for a Poisson replicate generated from the corresponding fitted means. Here, s indexes the MCMC iteration, $D_{obs}^{(s)}$ denotes the dispersion discrepancy for the observed data, and $D_{rep}^{(s)}$ denotes the corresponding discrepancy for the replicated data The mean of $D_{obs}^{(s)}$ was 1.21 (SD 0.21), very close to the mean of $D_{rep}^{(s)}$ at 1.17 (SD 0.22), yielding a paired posterior predictive *p*-value of 0.44 (Table 4, Figure 7). Taken together, these diagnostics indicate that while a simple Poisson GLM is strongly overdispersed, the hierarchical ST.CARar specification provides an adequate description of the dispersion in the data. Besides, a posterior predictive histogram of the dispersion discrepancy further illustrates the agreement between the observed and replicated dispersion measures, D_{obs}(s) and D_{rep}(s), under the ST.CARar model (Figure 7).

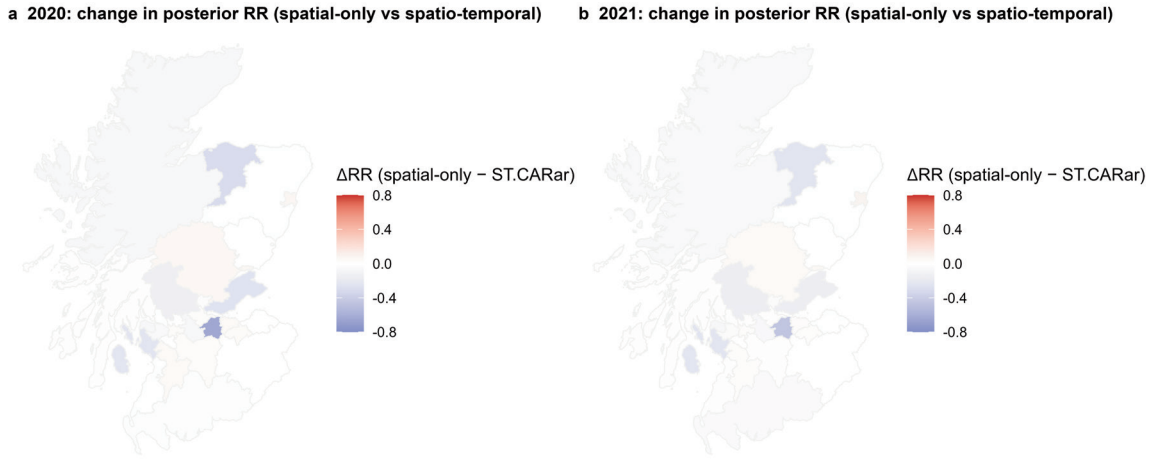

Figure 6. Sensitivity of modelled relative risk to inclusion of temporal autocorrelation. The maps show the difference in posterior median RR between a spatial-only CAR model and the full spatio-temporal ST.CARar model for (**a**) 2020 and (**b**) 2021, computed as $RR_{spatial\text{-}only} - RR_{spatial\text{-}temporal}$; red shading indicates higher estimated risk under the spatial-only model (positive differences); blue shading indicates higher estimated risk under the spatio-temporal model (negative differences).

Table 4. Overdispersion diagnostics and posterior predictive checks for non-spatial and spatio-temporal models.

Statistic model	Non-spatial Poisson GLM	Non-spatial negative- binomial GLM	ST.CARar (Poisson)
AIC	1163.6	635.6	–
Pearson dispersion	14.9	–	–
Mean D_{obs} (SD)	–	–	1.21 (0.21)
Mean D_{rep} (SD)	–	–	1.17 (0.22)
Predictive p_{paired}	–		0.44

AIC, Akaike information criterion; GLM, generalized linear model; ST.CARar, a function in the R statistical package CARBayesST used to fit Bayesian spatiotemporal GLM models to areal unit data; SD, standard deviation; D_{obs} denotes the Pearson-type dispersion discrepancy for the observed counts; D_{rep} denotes the Pearson-type dispersion discrepancy for the replicated counts.

Discussion

Summary in relation to the study aims

The study had three aims: to assess whether COVID-19 mortality in mainland Scotland was spatially structured, to examine the contribution of selected council-level covariates, and to determine the extent of residual clustering once these factors were considered. Descriptive maps showed a concentration of mortality in parts of the central belt, but the Bayesian spatio-temporal CAR model suggested that adjusted differences across councils were more moderate than the crude pattern. More generally, the results point to spatial gradients in mortality rather than a sharply polarised geography of risk. In line with the ecological design and the short two-year panel, these findings are best understood as evidence of council-level patterning rather than as a basis for strong claims about finely identified local mechanisms or temporal dynamics. The regression results identified a limited number of adjusted associations that were interpretable at the council level. Higher positive test counts were associated with higher mortality, unemployment was positively related to risk, and marriage counts were inversely related to mortality, while migration and hospital indicators were more weakly associated. These findings are more appropriately interpreted as council-level ecological associations than as evidence of clearly identified social mechanisms.

Spatial patterns in a comparative perspective

Placed alongside work for England and other European settings, the Scottish pattern looks familiar in structure but less extreme in magnitude. Sun *et al.* (2021) show marked spatial inequalities in COVID-19 mortality across English local authorities, tightly linked to socio-economic and environmental disadvantage and displaying pronounced clustering in large urban regions. Reviews of spatial and GIS-based analyses of the pandemic similarly emphasise substantial regional heterogeneity in incidence and mortality across Europe and North America (Franch-Pardo *et al.*, 2020). Against that background, the relatively modest residual spatial autocorrelation observed here suggests that, within mainland Scotland, the geography of COVID-19 mortality was somewhat less polarised than in some comparator contexts, even though the same underlying mechanisms appear to be at work. This reading is consistent with arguments that COVID-19 should be understood as a "syndemic", in which infectious disease interacts with preexisting chronic conditions and socio-economic deprivation (Bambra and Smith, 2021; McGowan and Bambra, 2022). For example, McGowan and Bambra (2022) document strong deprivation gradients in COVID-19 mortality across the UK and situate them in a longer history of health inequalities. The present study contributes to this literature by showing that, once age structure, infection burden, and a limited set of area-level indicators are taken into account, geographic differences in COVID-19 mortality across Scottish councils persist but remain moderate in scale.

Interpretation of selected area-level associations

The positive association between infection burden and mortality provides an important internal check on the analysis: councils with higher test-positive counts, interpreted as higher community transmission, experienced higher death rates even after adjustment for social covariates. Similar patterns have been reported in other national settings where local incidence has been an important proximate driver of spatial differences in mortality (Franch-Pardo *et al.*, 2020). In this study, the result is most appropriately interpreted as evidence that infection burden functioned as a proximal epidemiological indicator of mortality risk. The positive association between unemployment and mortality is consistent with broader evidence linking labour-market disadvantage to poorer health outcomes (Roelfs *et al.*, 2011). At the same time, in the present study unemployment is interpreted as an area-level socio-economic indicator rather than as direct evidence of a specific causal pathway. Because the analysis is ecological, the observed association should be read as a contextual relationship rather than as an identified mechanism operating at the individual level. The inverse association between marriage counts and mortality is an interesting empirical finding, but it should be interpreted cautiously. At the ecological scale used here, marriage counts are best regarded as an exploratory area-level proxy that may capture a mixture of demographic structure, household formation patterns, and broader local conditions. The measure is too coarse to identify a specific mechanism, and the present analysis does not allow these possibilities to

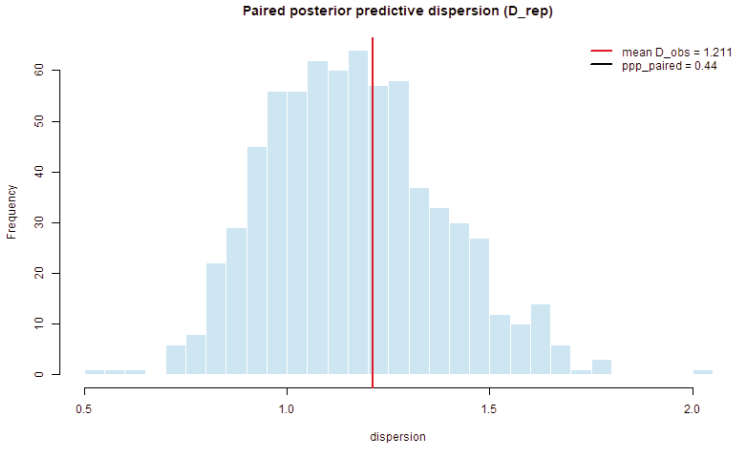

Figure 7. Paired posterior predictive check for dispersion. The histogram shows the distribution of the replicated Pearson-type dispersion discrepancy under the ST.CARar model. The vertical red line marks the mean observed discrepancy. The paired posterior predictive p-value was 0.44.

be disentangled. The positive association between quit-smoking rates and mortality is also difficult to interpret. One possibility is that cessation activity is greater in councils with historically higher smoking prevalence, poorer baseline health, or more intensive cessation targeting. More generally, this variable is best treated as an exploratory ecological proxy rather than as evidence of a clearly identified behavioural mechanism. Residual confounding therefore remains a plausible explanation.

Methodological considerations

Our model assessment highlights the importance of accounting for spatial and temporal structure when analysing COVID-19 mortality counts. In non-spatial sensitivity analyses, a negative-binomial GLM provided a much better fit than a simple Poisson GLM, with an AIC reduction of more than 500 units and a Pearson dispersion statistic of 14.9 for the Poisson model. This pattern is consistent with substantial extra-Poisson variation in the crude counts when area-level heterogeneity and temporal dependence are ignored, and is in line with previous work on overdispersed disease-mapping data (Lawson, 2013; Moraga, 2019). These findings justify the use of a hierarchical framework that explicitly models spatial and temporal random effects rather than relying on a non-spatial Poisson regression.

At the same time, the paired posterior predictive checks suggest that the spatio-temporal ST.CARar specification adequately captures this extra-Poisson variability. When dispersion was evaluated using a Pearson-type discrepancy for each posterior draw, the mean observed and replicated discrepancies were almost identical (1.21 and 1.17, respectively), and the paired posterior predictive *p*-value was 0.44. A *p*-value close to 0.5 indicates no systematic mismatch between the dispersion observed in the data and that expected under the fitted model, once spatial and temporal random effects are included (Gelman *et al.*, 2013). Taken together, the non-spatial negative-binomial comparison and the posterior predictive diagnostics support the Poisson ST.CARar model as an appropriate primary analysis for these data, while making explicit the presence of overdispersion in simpler non-spatial specifications.

Nevertheless, some limitations remain. First, we only considered a Poisson likelihood for the spatio-temporal CAR model; alternative formulations such as spatial negative-binomial or Poisson–lognormal models could in principle offer additional flexibility, especially in settings with more extreme overdispersion or zero inflation. Second, our overdispersion diagnostics are based on aggregate council–year counts and cannot capture within-area clustering at finer spatial or temporal scales. Finally, as with any ecological study, residual confounding by unmeasured area-level factors cannot be ruled out. Future work could explore more flexible hierarchical likelihoods and richer confounder structures, building on the diagnostics reported here. The ecological design means that council-level associations cannot be interpreted causally at the individual level, and residual confounding by unmeasured compositional and contextual factors is likely. In addition, the relatively small number of spatial units and time points limits the scope for more flexible modelling structures.

Implications for policy and future research

Several implications follow from these findings. First, the positive association between infection burden and mortality underlines the continuing importance of surveillance systems capable of detecting local surges in incidence. Spatial analyses in other settings have shown how case data linked to local areas can be used to identify emerging clusters and to guide the allocation of testing and control resources (Franch-Pardo *et al*., 2020). Maintaining, and when necessary scaling, such systems remains a core component of preparedness. Second, the unemployment gradient suggests that labour-market conditions form part of the wider risk context of the pandemic. Evidence from pre-pandemic cohorts indicates that unemployment and labour-market exit are associated with elevated mortality risk (Roelfs *et al*., 2011). Insofar as similar contextual mechanisms operated during the COVID-19 crisis, income protection, job-retention schemes, and related social protections are likely to have health as well as social benefits during future shocks. Finally, the absence of clearly delineated post-adjustment hot spots suggests that, at least in 2020–2021, Scotland did not develop the extreme internal geographic inequalities in COVID-19 mortality observed in some other settings. Further reductions in mortality are therefore likely to depend less on geographically narrow remediation of a few councils and more on broader efforts to reduce underlying social vulnerabilities while maintaining robust systems for monitoring changes in infection burden. Future research should extend the time series, use finer spatial units where possible, and incorporate richer measures of deprivation, occupation, housing, and vaccination.

Conclusions

Overall, the study contributes to the literature by showing that council-level spatial differences in COVID-19 mortality in Scotland became less pronounced after adjustment for infection burden and selected area-level indicators. This is an area-level ecological finding and should not be interpreted as direct evidence of specific local social mechanisms. Future work should extend the time series, use finer spatial units where possible, and incorporate richer measures of socio-economic conditions, occupation, housing, and vaccination. Such extensions would help to clarify which area-level associations observed here reflect broader contextual conditions and which may warrant closer investigation in more detailed designs.

References

Atchison CJ, Davies B, Cooper E, Lound A, Whitaker M, Hampshire A, Azor A, Donnelly CA, Chadeau-Hyam M, Cooke GS, Ward H, Elliott P, 2023. Long-term health impacts of COVID-19 among 242,712 adults in England, Nat Commun 14:6588.

Bambra C, Lynch J, Smith KE, 2021. The Unequal Pandemic: COVID-19 and Health Inequalities. Bristol: Policy Press.

Bambra C, Smith KE, 2021. The syndemic pandemic: COVID-19 and social inequality. In: Andrews GJ, Crooks VA, Pearce J et al., eds. COVID-19 and Similar Futures: Global Perspectives on Health Geography. Cham: Springer, pp. 147–154.

Besag J, York J, Mollié A, 1991. Bayesian image restoration, with two applications in spatial statistics, Ann Inst Stat Math 43:1–20.

Franch-Pardo I, Napoletano BM, Rosete-Verges F, Billa L, 2020. Spatial analysis and GIS in the study of COVID-19: A review, Sci Total Environ 739:140033.

Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB, 2013. Bayesian Data Analysis. 3rd ed. Boca Raton, FL: Chapman and Hall/CRC.

Griffith GJ, Davey Smith G, Manley D, Howe LD, Owen G, 2021. Interrogating structural inequalities in COVID-19 mortality in England and Wales, J Epidemiol Community Health 75:1165–71.

Hébert-Dufresne L, Scarpino SV, Young J-G, 2020. Macroscopic patterns of interacting contagions are indistinguishable from social reinforcement, Nat Phys 16:426–431.

Kianfar K, Zayeri F, Mahaki B, 2023. Spatial modeling of COVID-19 incidence and mortality: A Bayesian hierarchical approach using conditional autoregressive priors. Soc Sci Med 335:116298.

Lawson AB, 2013. Bayesian Disease Mapping: Hierarchical Modeling in Spatial Epidemiology. 2nd ed. Boca Raton, FL: CRC Press.

Lee D, 2013. CARBayes: An R package for Bayesian spatial modeling with conditional autoregressive priors, J Stat Softw 55:1–24.

Lee D, Rushworth A, Napier G, 2018. Spatio-temporal areal unit modeling in R with conditional autoregressive priors using the CARBayesST package, J Stat Softw 84:1–39.

Lee D, 2020. A tutorial on spatio-temporal disease risk modelling in R using Markov chain Monte Carlo simulation and the CARBayesST package, Spat Spatiotemporal Epidemiol 34:100353.

Leroux BG, Lei X, Breslow N, 2000. Estimation of disease rates in small areas: A new mixed model for spatial dependence. In: Halloran ME, Berry D, eds. Statistical Models in Epidemiology, the Environment and Clinical Trials. New York: Springer, pp. 179–191.

McGowan VJ, Bambra C, 2022. COVID-19 mortality and deprivation: pandemic, syndemic, and endemic health inequalities, Lancet Public Health 7:e966–75.

Moraga P, 2019. Geospatial Health Data: Modeling and Visualization with R-INLA and Shiny. Boca Raton, FL: Chapman and Hall/CRC.

National Records of Scotland, 2023. Deaths involving coronavirus (COVID-19). Edinburgh: National Records of Scotland. Available from: https://statistics.gov.scot/data/deaths-involving-coronavirus-covid-19.

NHS Scotland, 2023. COVID-19 Statistical Data in Scotland. Edinburgh: NHS Scotland. Available from: https://www.opendata.nhs.scot/dataset/covid-19-in-scotland.

Nikolaou M, 2022. Revisiting the standard for modelling the spread of infectious diseases, Sci Rep 12:7077.

O'Dowd A, 2023. Covid-19: UK death rate is still higher than before pandemic, BMJ 383:2371.

Patel JA, Nielsen FBH, Badiani AA, Assi S, Unadkat VA et al., 2020. Poverty, inequality and COVID-19: The forgotten vulnerable, Public Health 183:110–11.

Rigby E, 2021. The COVID-19 economy, unemployment insurance, and population health, JAMA Netw Open 4:e2035955.

Roelfs DJ, Shor E, Davidson KW, Schwartz JE, 2011. Losing life and livelihood: A systematic review and meta-analysis of unemployment and all-cause mortality, Soc Sci Med 72:840–54.

Rushworth A, Lee D, Sarran C, 2017. An adaptive spatio-temporal smoothing model for estimating trends and step changes in disease risk, J R Stat Soc Ser C Appl Stat 66:141–57.

Scobie S, 2022. Covid-19: How has the pandemic differed across the four UK nations? BMJ 377:o1482.

Scottish Government, 2019. Scottish Spatial Data Infrastructure. Edinburgh: Scottish Government. Available from: https://www.spatialdata.gov.scot/.

Sun Y, Hu X, Xie J, 2021. Spatial inequalities of COVID-19 mortality rate in relation to socioeconomic and environmental factors across England, Sci Total Environ 758:143595.

Tobler WR, 1970. A computer movie simulating urban growth in the Detroit region, Econ Geogr 46:234–40.

World Health Organization, 2023. Weekly epidemiological update on COVID-19, 1 September 2023. Geneva: World Health Organization. Available from: https://www.who.int/publications/m/item/weekly-epidemiological-update-on-covid-19-1-september-2023.

Online supplementary materials
Supplementary Table S1. Variables used in the analysis.

Received: 15 December 2025; Accepted: 12 April 2026.

Contributions: Zhiyuan Luo: conceptualization, software, data curation, formal analysis, methodology, writing-original draft, and writing-review and editing; Yancheng Liu: validation, critical review, language editing, and writing-review and editing; Wenqi Xu: validation, critical review, and writing-review and editing.

Conflict of interest: the authors declare no potential conflict of interest, and all authors confirm accuracy.

Availability of data and materials: the data used in this study were obtained from publicly available sources, including National Records of Scotland, NHS Scotland Open Data, and the Scottish Spatial Data Infrastructure. Details of the variables used in the analysis are provided in Supplementary Table S1. The processed analytic dataset is available from the corresponding author upon reasonable request.

Funding: the authors received no financial support for the research, authorship, and/or publication of this article.

AI use statement: ChatGPT was used only for limited language editing during manuscript preparation. All study design, data collection, data processing, statistical analysis, interpretation, and figure preparation were undertaken by the authors, who take full responsibility for the content of the manuscript.