



OPEN

## Modelling and predicting the spatio-temporal spread of COVID-19, associated deaths and impact of key risk factors in England

B. Sartorius<sup>1,2,3</sup>✉, A. B. Lawson<sup>4</sup> & R. L. Pullan<sup>1</sup>

COVID-19 caseloads in England have passed through a first peak, and at the time of this analysis appeared to be gradually increasing, potentially signalling the emergence of a second wave. To ensure continued response to the epidemic is most effective, it is imperative to better understand both retrospectively and prospectively the geographical evolution of COVID-19 caseloads and deaths at small-area resolution, identify localised areas in space–time at significantly higher risk, quantify the impact of changes in localised population mobility (or movement) on caseloads, identify localised risk factors for increased mortality and project the likely course of the epidemic at high spatial resolution in coming weeks. We applied a Bayesian hierarchical space–time SEIR model to assess the spatiotemporal variability of COVID-19 caseloads (transmission) and deaths at small-area scale in England [Middle Layer Super Output Area (MSOA), 6791 units] and by week (using observed data from week 5 to 34 of 2020), including key determinants, the modelled transmission dynamics and spatial–temporal random effects. We also estimate the number of cases and deaths at small-area resolution with uncertainty projected forward in time by MSOA (up to week 51 of 2020), the impact mobility reductions (and subsequent easing) have had on COVID-19 caseloads and quantify the impact of key socio-demographic risk factors on COVID-19 related mortality risk by MSOA. Reductions in population mobility during the course of the first lockdown had a significant impact on the reduction of COVID-19 caseloads across England, however local authorities have had a varied rate of reduction in population movement which our model suggest has substantially impacted the geographic heterogeneity in caseloads at small-area scale. The steady gain in population mobility, observed from late April, appears to have contributed to a slowdown in caseload reductions towards late June and subsequent start of the second wave. MSOA with higher proportions of elderly (70+ years of age) and elderly living in deprivation, both with very distinct geographic distributions, have a significantly elevated COVID-19 mortality rates. While non-pharmaceutical interventions (that is, reductions in population mobility and social distancing) had a profound impact on the trajectory of the first wave of the COVID-19 outbreak in England, increased population mobility appears to have significantly contributed to the second wave. A number of contiguous small-areas appear to be at a significant elevated risk of high COVID-19 transmission, many of which are also at increased risk for higher mortality rates. A geographically staggered re-introduction of intensified social distancing measures is advised and limited cross MSOA movement if the magnitude and geographic extent of the second wave is to be reduced.

The first cases of COVID-19 in the United Kingdom were confirmed on 31 January 2020<sup>1</sup>.

<sup>1</sup>Department of Disease Control, Faculty of Tropical and Infectious Diseases, London School of Hygiene and Tropical Medicine, Keppel Street, Bloomsbury, London WC1E 7HT, UK. <sup>2</sup>Centre for Tropical Medicine and Global Health, Nuffield Department of Medicine, University of Oxford, Oxford, UK. <sup>3</sup>Department of Health Metrics Sciences, School of Medicine, University of Washington, Seattle, WA, USA. <sup>4</sup>Department of Public Health Sciences, Medical University of South Carolina, Charleston, SC 29425, USA. ✉email: benn.sartorius@ndm.ox.ac.uk

Despite implementation of non-pharmaceutical interventions, including closure of non-essential services and subsequent stay at home orders and school closures on 20 and 23 March 2020 respectively, this was not sufficient to contain the first wave of the UK outbreak, although these measures did serve to slow early transmission<sup>2</sup>. The full extent of the effect of the mobility restrictions on COVID-19 transmission has not been fully elucidated nor explicitly quantified<sup>3,4</sup>. In addition, the attributability of hypothesised socio-demographic risk factors for COVID-19 related mortality such as deprivation, crowding and ethnicity have also not been fully quantified nor unpacked at small-area resolution.

As COVID-19 caseloads in England appeared to be rising towards a second peak (wave) at the time of this analysis, to ensure continued response to the epidemic is most effective, it is imperative that we better understand, both retrospectively and prospectively, the geographical evolution of COVID-19 and localised areas in space–time at higher risk of severe disease burden and mortality. Assessing the impact of mobility and population-density (crowding) on differences in caseloads at small area scale will not only allow assessment of the impact that social distance measures have had on the magnitude and timing of the waves, but also allow counterfactual assessment of what this magnitude may have been in the absence of these non-pharmaceutical interventions. More importantly, with the emergence of a subsequent wave(s) and coupled increased likelihood of small-area localised outbreaks, this work could help identify small-areas at elevated risk of transmission and mortality and subsequently inform when it might be safe to start lifting social distancing measures at small-area scale in a geographic staggered approach. Furthermore, a better understanding of the spatiotemporal dynamics of this COVID-19 outbreak, and accurate characterisation of the likely spread and magnitude, will be critical for the design of timely and cost-effective control strategies to minimise the spread of future such pandemics, and to help establish better early warning systems<sup>5–7</sup>.

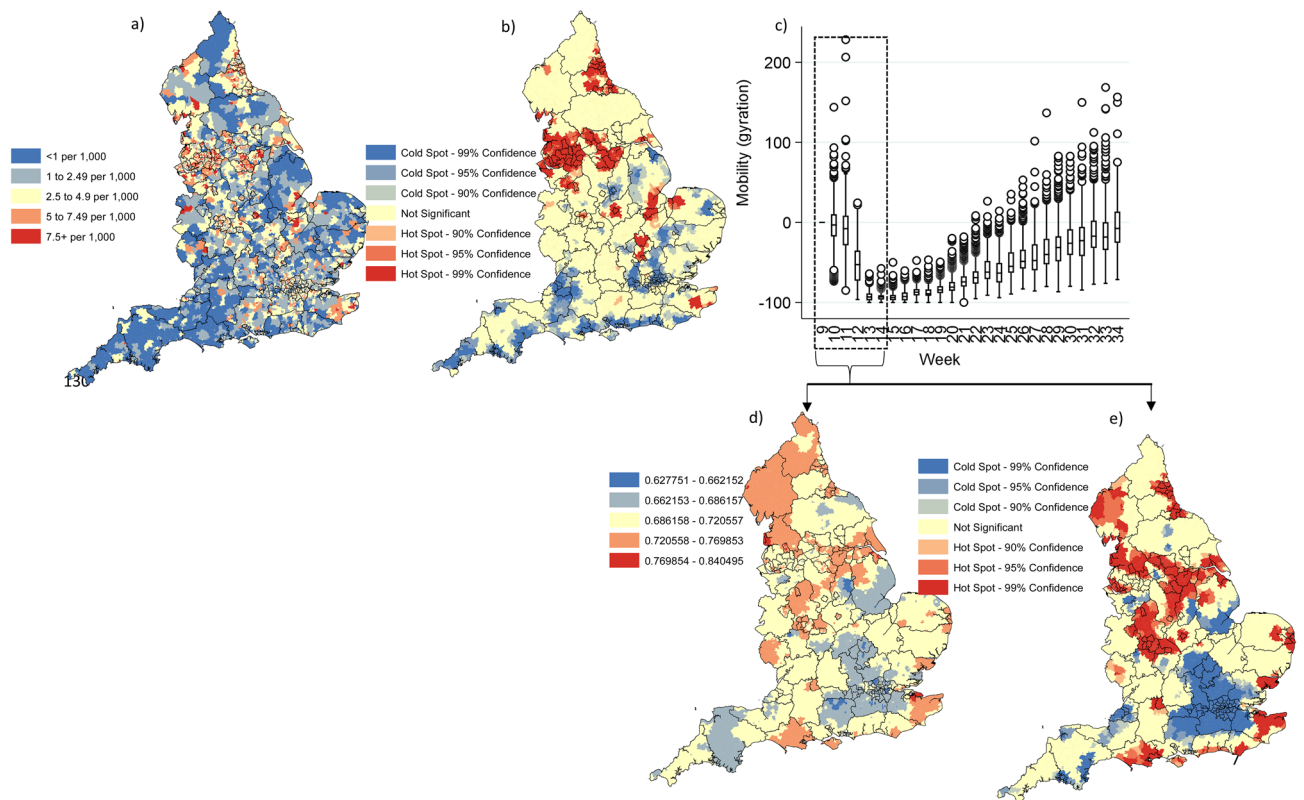
We apply a Bayesian hierarchical space–time Susceptible–Exposed–Infected–Removed (SEIR) model, previously applied to modelling of the spatial–temporal dynamics of influenza season outbreaks<sup>8</sup>, to publicly available COVID-19 confirmed case and death data in England at small area scale and by week to map the likely trajectory, peak and duration of the outbreak at MSOA level by week, quantifying the impact of mobility. This framework is an extension of the basic model<sup>9</sup> by combining the SEIR implementation with a spatial conditional autoregressive (CAR) model, and also extend this implementation by considering neighbourhood infection effects for the infection process<sup>8,10</sup>. We utilise weekly Clinical Commission Group (CCG) level population mobility data against observed confirmed COVID-19 case data to assess the impact mobility reduction at small-area scale has had on case transmission, and counterfactually what the magnitude may have been under the scenario of no mobility loss. Furthermore, the inclusion of the death compartment of the model using observed weekly COVID-19 related deaths at MSOA level and linkage to key socio-demographic risk factors allows this framework to identify not only small-areas with significantly elevated mortality risk but also what localised contextual factors may be significantly contributing to this. The space–time dynamics of COVID-19 related mortality in England and associated risk factors/determinants at MSOA level are also assessed to identify vulnerable MSOA.

## Results

**Observed spatial patterns for COVID-19 caseloads at MSOA.** A spatial assessment of COVID-19 caseloads (scaled by population totals) for the observed data input period (week 9 to 34 of 2020), suggests a non-random distribution to high incidence MSOA across England with higher risk areas concentrated in and around metropolitan areas (Fig. 1a) for example Leicester, Birmingham, Liverpool, and Manchester. The analysis of localised spatial clustering identified significant hot spots (at MSOA level) in Bedford/Bedfordshire, Leicester, Peterborough, large contiguous band in Cheshire and Merseyside, another contiguous band spanning Leeds–Wakefield–Barnsley–Sheffield and in the north Newcastle–Sunderland (Fig. 1b). Additional pockets of significant excess COVID-19 case rates in less metropolitan type areas were identified in pockets throughout England, for example Ashford, Kings Lynn and West Norfolk and Burrow-in-Furness.

**Spatial temporal patterns in population mobility at MSOA.** Despite an overall significant reduction in population mobility from early March (week 9) to mid-April (week 14) of 2020 (Fig. 1c), mobility rates began to increase from week 16 and median mobility levels across MSOA had been almost attained pre-lock March/April down levels by week 34 of 2020. The rate of cumulative mobility loss by mid-April varied across MSOA (Fig. 1d), with lower levels of mobility loss generally in part of West and East Midlands, Northwest, York and the Humber land as well as North East England. A large band of significant hotspots for lower cumulative mobility loss at week 14 were identified in a large contiguous aggregation of MSOA spanning from Birmingham through to Leicester, Nottingham, Sheffield, Leeds, Manchester and Liverpool (Fig. 1e). Furthermore, significant clustering of lower cumulative mobility loss was also observed in Newcastle and surrounds, Norwich and coastline areas immediately to the east of Norwich, Tendering/Colchester, Canterbury/Dover and west of Southampton for example.

**Spatial–temporal patterns for COVID-19 mortality.** The highest cumulative death rates for COVID-19 at MSOA resolution from week 9 to 26 of 2020 were observed in parts of North London, most of Birmingham and immediate surrounding MSOA, a large area spanning Cheshire–Merseyside, and lastly in and surrounding Newcastle (Fig. 2a). Significant hotspots for COVID-19 related mortality were highly clustered and were identified in contiguous MSOA located in North London (Ealing–Brent–Hillingdon–Harrow–Barnet–Enfield–Three Rivers–Hertsmere), West Midlands (Birmingham–Sandwell–Walsall–Dudley–Wolverhampton), North West (Liverpool–Manchester–Cheshire), Sheffield, and Newcastle (County Durham–Sunderland–Gateshead) (Fig. 2b). The two principle risk factors associated with COVID-19 mortality risk and leveraged in this study, namely proportion of MSOA population aged 70 years or older and elderly population living in deprivation have two

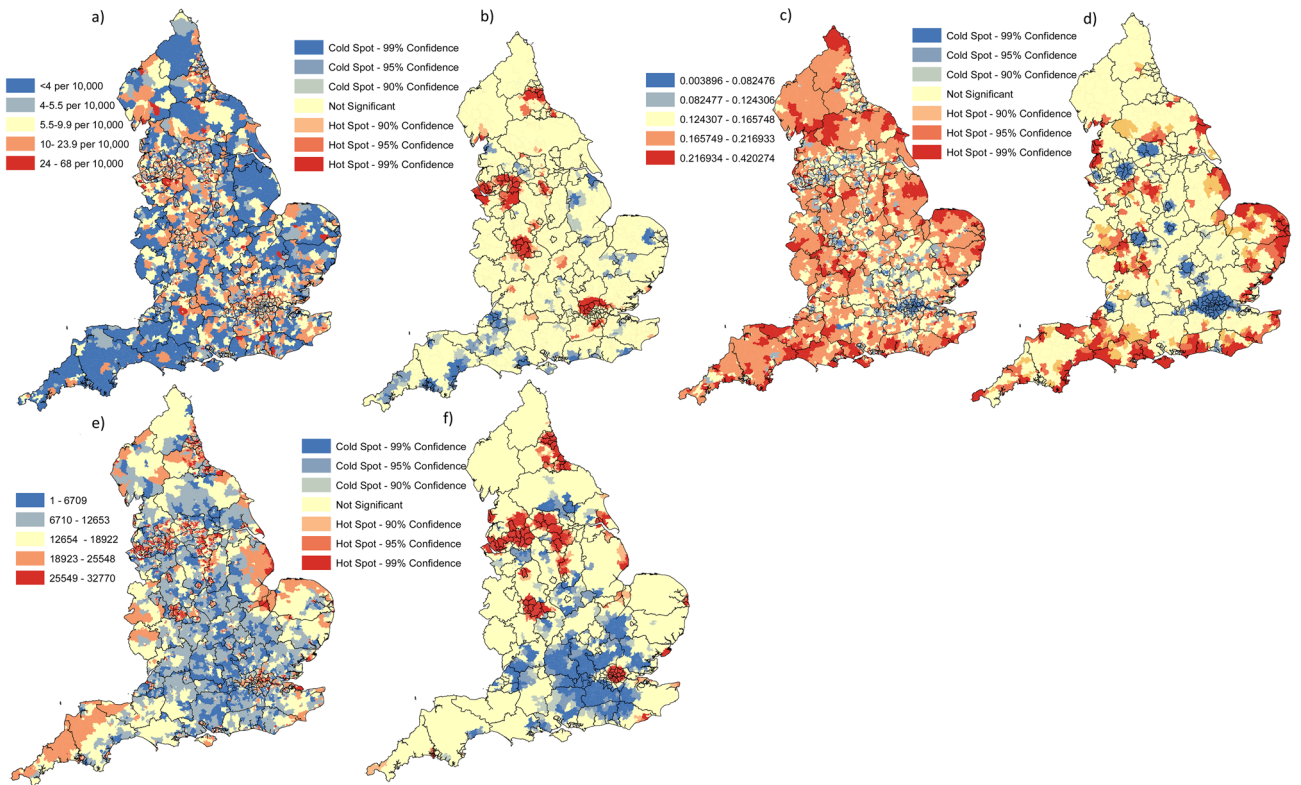


**Figure 1.** (a) cumulative case rates at week 34 at English MSAO level ( $N=6791$ ); (b) statistically significant localised hot and cold spots for higher and lower cumulative COVID-19 cases at MSAO resolution at week 34 of 2020 using the Getis-Ord  $G_i^*$  localised clustering statistic; (c) Box plot of mobility by MSAO and week; (d) cumulative mobility loss (darker red indicates lower mobility loss) by MSAO in England ( $n=6791$ ) from week 9 to 14 of 2020; (e) statistically significant localised hot and cold spots for higher and lower cumulative mobility loss at MSAO resolution from week 9 to 14 of 2020 using the Getis-Ord  $G_i^*$  statistic. Maps constructed using ArcGIS 10.5 (<https://desktop.arcgis.com/en/>).

very distinct spatial distributions across England (Fig. 2c,e respectively). The MSAO with higher proportions of population aged 70 years or older appear significantly clustered along the East of England coastline and the coastline in Southwest and South East regions of England (Fig. 2d). Additional significant hot spots for elderly population clustering were identified outside key metropolitan centres, for example: west of Birmingham, north of Liverpool, south of Manchester, north of Nottingham and north of Bradford-Leeds (Fig. 2d).

**Spatial patterns for deprivation.** The most deprived (worst 10 percentile) neighbourhoods (680 MSAO) are not evenly distributed across England (Fig. 2e,f), with 50% of these areas located within just 20 local authority districts, notably—Birmingham with 62 such neighbourhoods; Liverpool, 35; Leeds and Manchester, 29 each, and Bradford, 20. Significant spatial clustering of the most deprived neighbourhoods is concentrated in large areas of contiguous MSAO located in London, West Midlands, Northwest, Yorkshire and Northeast England, with pockets of significant deprivation surrounded by less/least deprived neighbourhoods (Fig. 2f). Furthermore, concentrations of deprivation among these 680 MSAO are also disproportionately concentrated in large urban conurbations (142 Industrial and Multi-ethnic, 125 Larger Towns and Cities; 80 Ethnically Diverse Metropolitan Living; 33 urban living), as well as in areas that are (or historically were) classified as heavy industry manufacturing and/or mining sectors (74 manufacturing legacy, 42 service economy; 39 mining legacy). A statistical assessment of spatial clustering for elderly population living in deprivation confirms the significant concentration of hot spots in the metropolitans, namely: London, Birmingham, Liverpool, Manchester, Nottingham, Sheffield-Leeds and Newcastle (Fig. 2f).

**Space–time SEIR dynamics.** The full multivariate Bayesian space–time model estimated an asymptomatic fraction of 0.2 (95% UI: 0.06–0.39) i.e. for every 5 confirmed cases there is likely to be 1 additional asymptomatic case (Table 2). A decrease in mobility at MSAO scale has had a significant reduction on caseloads i.e. each unit increase in cumulative mobility (rescaled to be between 0 and 1) appears to be associated with a +3.6-unit increase in cases on the log scale (95%UI: 3.57–3.61) (Table 1). The change in cumulative mobility also appears to significantly interact with the natural logarithm of population density within a MSAO i.e. the impact of higher mobility is amplified by increasing population density ( $\beta = +0.05$ ). Increasingly proportion of a MSAO population aged 70 years or older was strongly and significantly associated with increased COVID-19 related deaths ( $\beta = +4.86$ ,



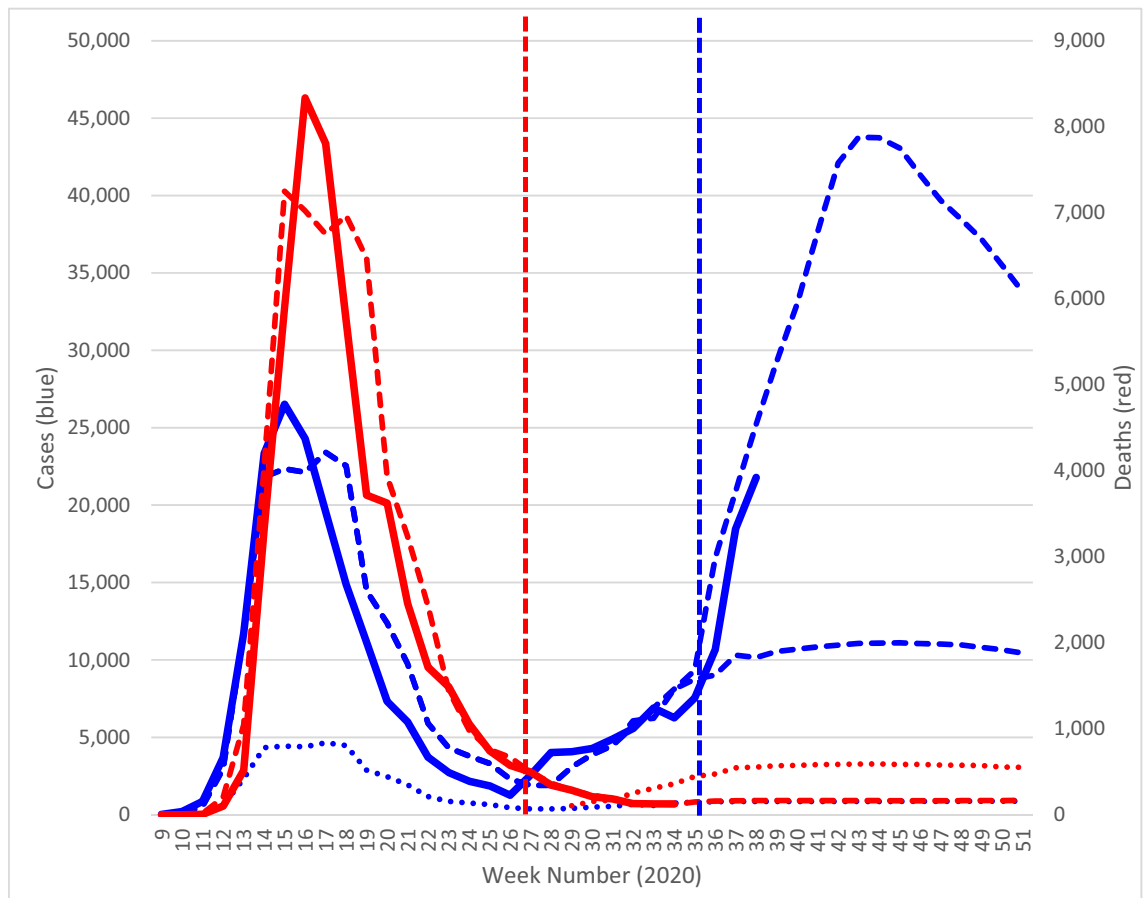
**Figure 2.** (a) Cumulative case rates at week 34 at English MSOA level (N = 6791); (b) statistically significant localised hot and cold spots for higher and lower cumulative COVID-19 cases at MSOA resolution at week 34 of 2020 using the Getis-Ord  $G_i^*$  statistic; (c) proportion of population aged 70+ at MSOA resolution; (d) statistically significant localised hot and cold spots for higher and lower proportion of population aged 70+ at MSOA resolution using the Getis-Ord  $G_i^*$  statistic; (e) Index of multiple deprivation at MSOA resolution; (f) statistically significant localised hot and cold spots for higher and lower proportion of IMD at MSOA resolution using the Getis-Ord  $G_i^*$  statistic. Maps constructed using ArcGIS 10.5 (<https://desktop.arcgis.com/en/>).

Compartment	Node	Coef. ( $\beta$ )	2.50%	97.50%
Cases	Asymptomatic fraction	+ 0.20	+ 0.06	+ 0.39
	Cumulative mobility loss	+ 3.59	+ 3.57	+ 3.61
	Cumulative mobility loss X ln (Population per km <sup>2</sup> )	+ 0.05	+ 0.05	+ 0.05
	$\tau (v_{1i})$ —spatially structured random effect	2.35	2.05	2.70
	$\tau (U_{1i})$ —unstructured random effect	14.53	12.06	17.47
Deaths	Proportion aged 70 years and older	+ 4.86	+ 4.59	+ 5.15
	Elderly population living in material deprivation index <sup>1</sup>	- 0.25	- 0.34	- 0.10
	$\tau (v_{2i})$ —spatially structured random effect	0.81	0.73	0.88
	$\tau (U_{2i})$ —unstructured random effect	1.44	1.34	1.55

**Table 1.** Posterior statistics from the Bayesian multivariable space–time SEIR model using WinBUGS. i: increasing score indicate lower level of material deprivation.

95% CI: 4.59–5.15) while a decreasing proportion of the elderly population in a given MSOA living in deprivation appears to have a strong and significant negative association ( $\beta = -0.25$ , 95% CI:  $-0.34, -0.10$ ) (Table 1).

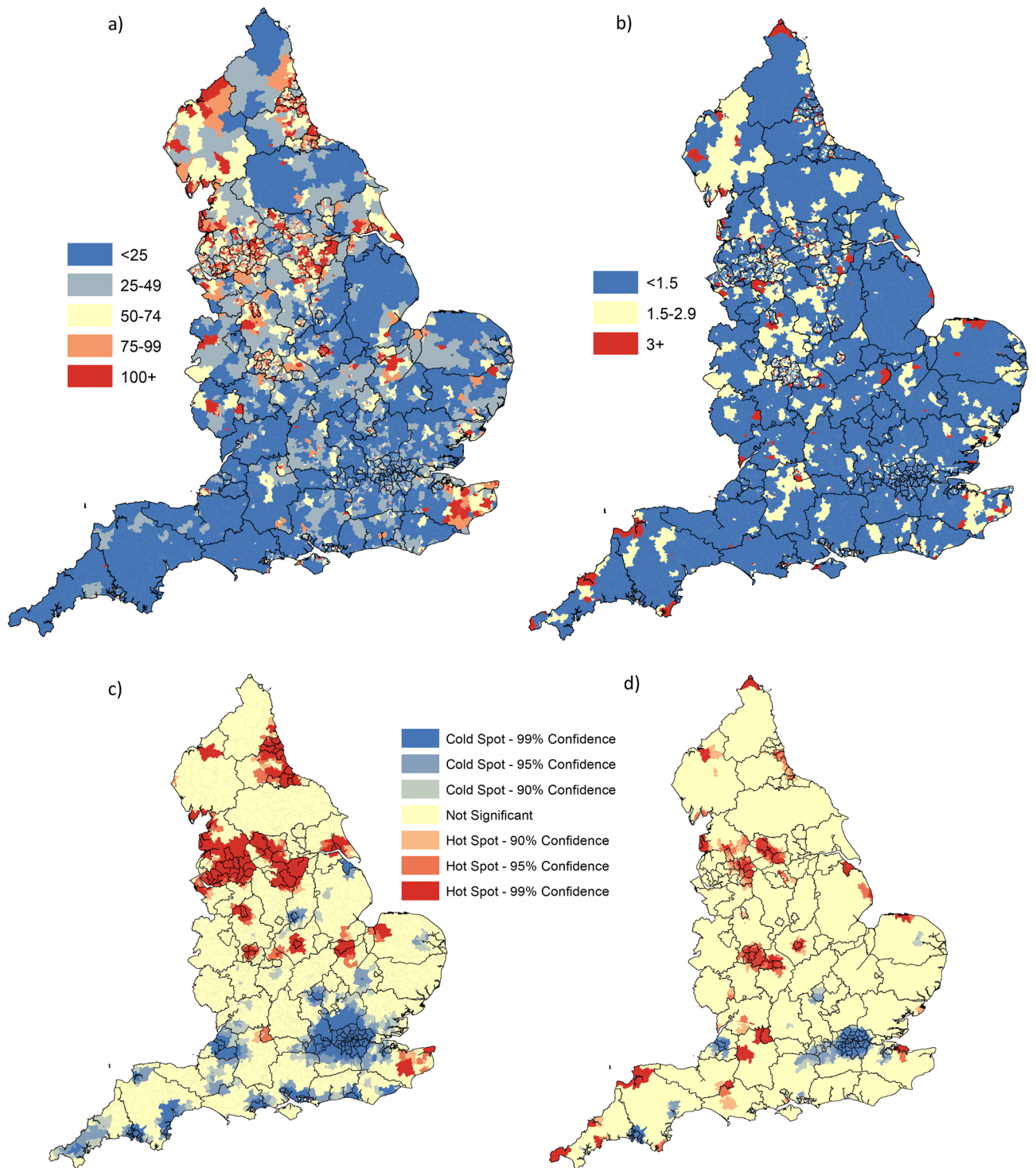
**Posterior space–time projections from ST-SEIR model.** The posterior projections from the full space–time SEIR model closely fit with the observed temporal course of case and deaths for COVID-19 over the observed data input period (spanning the first wave), namely week 9 to 26 for deaths and week 9 to 34 for cases (Fig. 3). From week 27 the model fairly closely tracks the increase in caseloads which appears to signal the start of the second wave. Post week 34 i.e. the model out of prediction phase, we note an exponential rise in observed cases beginning around week 35. The model prediction based on averted cases from lockdown in the first phase combined with the impact of steadily increasing population mobility appears to fairly accurately track the exponential rise from week 35 and suggests a likely second wave peak of almost 45,000 weekly cases



**Figure 3.** Observed and model fitted number of confirmed cases and deaths in England from week 9 to 34 of 2020, with projections to week 51 of 2020. Vertical coloured line represents the truncation points at the time of this analysis for observed input data for cases (week 34) and deaths (week 26) respectively. Blue solid line = observed number of weekly COVID-19 cases; blue dashed line = model fitted cases; dotted blue line = asymptomatic cases; red solid line = observed number of weekly COVID-19 deaths, red dashed line = model fitted deaths; red dotted line = COVID-19 deaths based on mortality rate from first wave.

by week 43/44 (i.e. latter half of October) of 2020 (Fig. 3). Secondly, the model prediction based on the impact of observed mobility change alone, without accounting for the averted (or counterfactual) caseloads in the first wave, appears to significantly underestimate the trajectory of the outbreak in the current phase post week 34 (Fig. 3). Interestingly while the death compartment of the model closely fits the observed death rates in the first wave, the counterfactual projection post week 26 in the death compartment which assumes similar death rates to that observed in the first wave appears to significantly overestimate the death rate observed at the time of this analysis for the initial part of the second wave (Fig. 3).

**Model predicted hotspots for weeks 35–40.** The forward projection of the model for weeks 35 to 40 at MSOA level suggests high and significant case rates forecast for large contiguous area spanning North West and Yorkshire regions (Liverpool-Manchester complex, contiguous band spanning Bradford-Calderdale-Kirklees-Leeds-Wakefield-Barnsley-Doncaster-Rotherham-Sheffield, contiguous band of MSOA in and around Kingston upon Hull), large band of connected MSOA in the north east spanning from Newcastle through to Stockton-On-Tees, high risk pockets through the midlands (Stoke-on-Trent, Wolverhampton-Walsall-Dudley-Sandwell, Leicester, Peterborough) and two additional pockets in southern England (Swindon, Ashford-Canterbury-Thanel) (Fig. 4a/c). The forecast pockets for significantly high COVID-19 related mortality are more focused/smaller in geographic extent compared to the caseloads with some overlap between areas with project high caseloads and high death rates: Blackpool, Manchester, band spanning Bradford-Leeds-Kirklees-Wakefield, South Tyneside-Sunderland, large band spanning Wolverhampton-Birmingham-Coventry, Leicester and Swindon (Fig. 4b/d). Additional pockets of high mortality rate risk but with low-moderate projected caseload rates were identified, with multiple pockets in the south of England (Wiltshire, North Devon, Cornwall, South Somerset-Dorset and Swale).



**Figure 4.** Model projected COVID case (a) and death (b) rates per 100,000 at MSOA level for week 34 to 40 of 2020; statistically significant localised hot and cold spots for COVID-19 case (c) and death (d) rates at MSOA level for week 34 to 40, 2020. Maps constructed using ArcGIS 10.5 (<https://desktop.arcgis.com/en/>).

## Discussion

To both support and respond effectively to surveillance activities such as NHS Test and Trace or universal weekly testing<sup>11</sup>, it is essential to unpack the spatial–temporal dynamics of the COVID-19 outbreak and related mortality at high spatial resolution (e.g. MSOA) over time and assess the contribution of key risk factors to differences in risk across space and time. Our high resolution spatial–temporal (MSOA by week) model suggest that social distancing measures, as estimated by the impact on population mobility patterns, have had a significant impact on the reduction of COVID-19 caseloads across England, in line with previous studies, for example in China and

Italy<sup>12–14</sup>. However, local authorities have seen varied rates of reduction of population movement, significantly impacting the cumulative caseloads at small-area scale. A sensitivity analysis whereby mobility restrictions are eased at various intensities at MSOA scale suggests that a steady gain in mobility, as observed from late April, may have contributed to a slowdown in caseload reductions towards late June and subsequent steady rise thereafter until late August (Fig. 3). Subsequent to that the model prediction and observed caseloads data suggest an exponential rise from week 35 (early September). As such, it would appear that a critical tipping point had been reached at that stage, and our model suggested particular groupings of MSOA at significant increased risk of future caseloads and deaths. Further targeted implementation of local lockdown is essential at this stage. A geographically staggered approach combined with enhanced community surveillance<sup>11</sup> will become increasingly important if we are to reduce the magnitude of the second wave in England and plan for subsequent waves/cycles in the absence of an effective vaccine.

While previous work<sup>12</sup> has confirmed that digital sources of mobility data correlate well with COVID-19 incidence and contact mixing<sup>15</sup> at country level, the extent of the effect of mobility restrictions on COVID-19 transmission in time at small-area resolution has not been fully elucidated nor explicitly quantified. Furthermore, the impact of demographic and socio-economic factors such as elderly population living in deprivation on COVID-19 related mortality risk at high spatial resolution has not been adequately assessed and identification of these areas at excess risk can help guide local policy to better protect these highly vulnerable sub-groups. Lastly, as the Covid-19 pandemic has progressed and evolved, a variety of modelling approaches have been proposed and/or implemented, many of which are time series-based, and do not explicitly incorporate spatial structure in the modelling process. However, it is clear that spatial position (i.e. small area) and influence of neighbouring small-areas is important and so ignoring limits the ability of the models to account for this phenomenon. Additionally, the use of unstructured random effects to allow for unexplained variability and confounding is often ignored. In this article we demonstrate the use of a Bayesian spatial–temporal susceptible–exposed–infected–removed (SEIR) model with spatial geo-referencing at small area (MSOA) resolution in England by week—this combined implementation approach has not been used extensively and presents some distinct advantages (as briefly summarised above) over either formulation in isolation.

Crucially, our results highlight differences in epidemic dynamics across small areas in England, emphasising the importance of monitoring at fine sub-national scale. Our findings suggest that a significant proportion of caseload and death variance is spatially structured, and that caseloads in neighbouring local authorities of a given MSOA in the preceding time point are significantly associated with caseloads in current time point. This—coupled with the large significant regional clustering of high caseloads observed—has important implications for informing limitations on movement between areas in the subsequent phases of lockdown easing. Small-area lockdowns could help reduce COVID-19 transmission while minimizing the size of the population (and economy) that needs to be disrupted to continue to release reduced transmission. Furthermore, deaths due to COVID-19 appear to be significantly clustered in space, with some overlap between areas in and around metropolitan centres with higher levels of elderly population living in deprivation (prominent examples: Wolverhampton–Birmingham complex, Blackpool, Manchester, Bradford–Leeds–Wakefield complex and south of Newcastle). However additional pockets of excess mortality risk were also identified in areas in less deprived coastal areas, however from our analyses many of these localities have much higher proportions of elderly population which may in part explain these clusters.

Human mobility has been demonstrated to play an important role in the early stages of the COVID-19 epidemic<sup>12</sup> and in particular reductions observed in digital data sources of mobility appear to correlate well with incidence<sup>13</sup> as well as contact patterns<sup>15</sup>. As many countries continue to bring their COVID-19 outbreaks under control, evidence on the effectiveness of current lockdowns continues to emerge<sup>16–19</sup>. Social distancing measures have had a significant impact on the reduction of COVID-19 caseloads across England, however local authorities have had a varied rate of reduction of population movement, which has substantially impacted the cumulative caseloads at small-area scale. While non-pharmaceutical interventions such as social distancing have had a profound impact on the trajectory of the COVID-19 outbreak in England, country wide relaxing of these measures from June 2020 and gradual observed increases in population mobility which started even prior to this, risks a rebound in a number of local authorities which still have caseloads above the threshold required to realise MSOA cessation.

Furthermore, at the time of this analysis and the period covered by the input data, vaccination was not yet available for SARS-CoV-2. An application of this model post vaccine rollout would need to explicitly account for increasing cumulative vaccination of the population in a given MSOA with time. This could be achieved using a modified version of the SEIR model that accounts for vaccination, for example<sup>20</sup>.

SARS-CoV-2 co-infection with other infectious agents at admission have been documented<sup>21</sup> and appear to significantly increase mortality risk<sup>22</sup>. Future applications should consider additional potentially useful and important predictors for COVID-19 mortality such as co-infection rates at admission.

This study is subject to limitations. First, confirmed COVID-19 cases were linked to the MSOA where the test was carried out, rather than patient residence. This is of greater concern for metropolitan areas, where a substantial proportion of patients may have been admitted or transferred to neighbouring trusts before being tested. Furthermore, we model the impact of mobility within a given MSOA and week on caseloads and thus do not include a measure of between MSOA mobility by week which would also be an important additional driver of wider COVID-19 related spread and subsequent caseloads. Secondly, the data are also vulnerable to bias introduced by variable testing rates between Trusts, which may have contributed to greater between MSOA variability. Thirdly, we are assessing the association between aggregated cases/deaths at MSOA level with aggregated risk factor variables at the same level. Thus we cannot discount the possible impact of ecological fallacy in the observed association. Fourthly, both the case fatality risk and the fraction of asymptomatic cases of COVID-19 are known to vary widely by age<sup>23,24</sup>. Thus our model does not fully consider the age-structured transmission and

Model component	Variable	Source (open access unless indicated*)	Spatial resolution	Temporal resolution	Compartment
Outcomes	Confirmed cases	UK COVID-19 Dashboard: <a href="https://coronavirus.data.gov.uk/">https://coronavirus.data.gov.uk/</a> Number of people with a lab-confirmed positive COVID-19 PCR test. Data include only pillar 1 cases until 2 July, from when pillar 2 cases are also included. Cases are allocated to the person's area of residence	MSOA	Weekly	Cases
	Deaths within 28 days of positive test result for COVID-19	UK COVID-19 Dashboard: <a href="https://coronavirus.data.gov.uk/">https://coronavirus.data.gov.uk/</a> Total number of people who had a positive test result for COVID-19 and died within 28 days of the first positive test, reported on or up to the date of death or reporting date (depending on availability). Deaths are allocated to the deceased's usual area of residence	MSOA	Weekly	Deaths
Covariate (daily varying)	Mobility (population movement index)	Oxford COVID-19 impact monitor (Cuebiq) ( <a href="https://oxford-covid-19.com/">https://oxford-covid-19.com/</a> )	Clinical Commission Group (CCG)	Daily	Cases
Covariates (annual estimates, weekly fixed)	Age, ethnic structure	ONS ( <a href="https://www.ons.gov.uk/">https://www.ons.gov.uk/</a> )	MSOA	2020	Deaths
	Population density	ONS ( <a href="https://www.ons.gov.uk/">https://www.ons.gov.uk/</a> )	MSOA	2020	Cases
	Household occupancy (overcrowding)	ONS ( <a href="https://www.ons.gov.uk/">https://www.ons.gov.uk/</a> )	MSOA	2019	Cases/deaths
	Index of Multiple Deprivation (IMD)	Multiple deprivation experienced by Lower Layer Super Output Area (LSOA) in England ( <a href="https://www.gov.uk/government/statistics/english-indices-of-deprivation-2019">https://www.gov.uk/government/statistics/english-indices-of-deprivation-2019</a> )	MSOA	2019	Deaths
	Elderly population proportion living in deprivation	Public Health England – Local Health: <a href="https://www.localhealth.org.uk/">https://www.localhealth.org.uk/</a>	MSOA	2019	Deaths
	Emergency hospital admissions for chronic disease	ONS ( <a href="https://www.ons.gov.uk/">https://www.ons.gov.uk/</a> )	MSOA	2017/18	Deaths

**Table 2.** List of model components, relevant variables and data sources.

mortality dynamics and could this lead to biased estimates due to possible variations in the age-distribution of each MSA. A final limitation relates to the mobility data and lack of detail regarding movement between local authorities. While we have demonstrated the importance of mobility loss within a given MSA on its subsequent caseloads, the impact of movement between local authorities (and density thereof) would provide additional important information regarding cross locality mixing and further targeted restriction of movement between particular higher risk areas.

## Conclusions

This study highlights the importance and usefulness of space–time framework to unpack the dynamics of the COVID-19 outbreak at high resolution (MSOA) and in time, identify particular small-areas at elevated risk of transmission as well as COVID-19 related mortality. Furthermore, at small area scale we demonstrated the utility of longitudinal mobility data for real-time surveillance of the impact of social distancing interventions, and thereby can also allow national authorities to assess the impact on outbreak dynamics at small-area. It is advisable to pre-emptively assess what impact changes in mobility post lockdown easing will have on outbreak rebound potential at localised geographic scale, and to include monitoring of changes in daily/weekly mobility patterns. Lastly geographic areas with high proportions of elderly population and in particularly elderly population living in deprivation are at significantly greater risk of COVID-19 related mortality, and therefore enhanced surveillance/case follow-up should be ensured in these more vulnerable areas.

## Methods

**Overview.** This analysis adheres to the Guidelines for Accurate and Transparent Health Estimates Reporting standards (GATHER)<sup>25</sup> (Supplementary Material Section 1). The data used are all in the public domain. Data processing and analyses were performed using Stata 16, and WINBUGS. The statistical code for implementing the Bayesian model in WINBUGS is provided in Supplementary Material (Supplementary Material Section 2). We constructed all maps using ArcGIS 10.5.

**Data.** A summary of the input data and sources is presented in Table 2. The weekly number of confirmed COVID-19 cases (Pillar 1—Government's mass testing programme) and deaths for each MSA Middle Layer Super Output Area (MSOA) (6791 areas) in England were obtained from the UK COVID-19 dashboard (<https://coronavirus.data.gov.uk/>) for the period of February 24, 2020 (week 5) to August 23, 2020 (week 34). The MSA shape file, used to create the adjacency matrix for the model below and for mapping of the model posteriors, was



obtained from the United Kingdom government geoportal platform: <https://geoportal.statistics.gov.uk/datasets/middle-layer-super-output-areas-december-2011-boundaries-ew-bfc-1>. Weekly COVID-19 deaths at MSOA resolution were also extracted from the NHS PHE Dashboard (<https://coronavirus.data.gov.uk/>).

Data on population size, population density and the proportion of the population aged 70 years and older as well as proportion of population that is of Black or mixed ethnicity by MSOA were obtained from the Office for National Statistics (ONS) (<https://www.ons.gov.uk/>). Furthermore, we also extracted the Index of Multiple Deprivation (IMD) score and proportion of the elderly population living in deprivation for each MSOA in England for 2019 (Public Health England: Local Health. [Accessed: 24/08/2020] <https://www.localhealth.org.uk/#c=home>).

Daily population movement data by Clinical Commission Group (CCG) was extracted from the COVID-19 Impact monitor (<https://www.oxford-covid-19.com/>). Weekly mobility in a given MSOA were assumed to be the same as the weekly mobility in the higher level CCG containing most of that MSOA. Furthermore, rail passenger numbers and crowding statistics (2018) were extracted for use a commuter density index by MSOA prior to the local down for use a proxy for likely daily contacts with other individuals and mixing density (<http://maps.dft.gov.uk/rail-passengers-and-crowding/interactive-dashboard/index.html>).

As a proxy for underlying chronic circulatory and respiratory disease burden, we also utilised mortality rates from chronic obstructive pulmonary disease (COPD) from 2016–2018 by MSOA. (Public Health England: Local Health. [Accessed: 24/08/2020] <https://www.localhealth.org.uk/#c=home>).

**Data analysis.** For initial exploratory spatial analysis, we used the Getis–Ord  $G_i^*$  statistic ( $G_i^*$ )<sup>26</sup>, also known as hot-spot analysis, to identify significant higher and lower risk MSOA in terms of case and deaths rates for the observed input data period, namely week 9 to 34. The significance level was set as 0.05. A previously developed Bayesian space–time SEIR formulation<sup>8</sup> was then applied to assess the spatiotemporal variability of COVID-19 transmission at small area scale (MSOA) and by week in England, by accounting for the modelled transmission dynamics of the pathogen, inherent spatial–temporal correlation in the data, and important contextual risk factors for both COVID-19 cases and deaths. We also assessed the sum of cases in shared neighbours in the preceding time point as an additional parameter in our model to further assess the dependence of case-loads in adjoined areas<sup>27</sup>. Based on the available data provided we the starting study week was denoted as week 9 of 2020 (starting 24 February 2020), with available case data at the time of extraction for this analysis available up to week 34 (starting 17 August 2020) and death data available up to week 26 (starting 22 June 2020). We also limited the end date for the input data at week 34 to assess the predictive performance of the model into the future if we are to prove its predictive utility and thus use as a pre-emptive resource allocation/prioritization tool.

Let  $I_{ij}$  be weekly confirmed COVID-19 cases in MSOA region  $i$  and week  $j$ , ( $i = 1, \dots, 151$ ;  $j = 9, \dots, 51$ ). A discrete form SEIR model for the size of susceptible population at week  $(j + 1)$  and MSOA region  $i$  is given by  $S_{i,j+1} = S_{i,j} - I_{i,j} - R_{i,j}$ , where  $S$ ,  $I$  and  $R$  represent the susceptible, infectious and removed (deaths and recovered) populations, respectively. The number of the infected COVID-19 cases in region  $i$  and week  $j$  were assumed to follow a Poisson distribution, with expected number of cases is a function of the serial interval or generation time as a time measure of disease communicability and defined as the number of infected cases in one median incubation period back ( $j - 1$  week<sup>28</sup>) i.e. and the susceptible population size in week  $j$ . The number of asymptomatic cases by week and MSOA is assumed to be a function of the observed case load in a given MSOA/week multiplied by a parameter for the additional fraction that are asymptomatic.

Hence,  $I_{ij} \sim \text{Pois}(\mu_{ij})$ , and  $\mu_{ij}$  is given by<sup>8</sup>:

$$\mu_{ij} = \beta_{ij} \times S_{ij} \times I_{i,j-1}$$

$$S_{ij} = S_{i,j-1} - I_{i,j-1} - \text{Asym}_{i,j-1} - \text{deaths}_{i,j-1}$$

$$\log(\mu_{ij}) = \log(\beta_{1ij}) + \log(S_{ij}) + \log(I_{i,j-1})$$

$$\log(\beta_{1ij}) = b_1 0 + b_1 \times (\text{mobility})_{i,j} + b_2 \times \text{mobility}_{i,j} \times \ln(\text{population density}_i) + u_{1i} + v_{1i}$$

$$\text{Asym}_{ij} \sim \text{Pois}(U_{ij} * \text{asym})$$

$$\text{asym} \sim \text{Uniform}(0.05, 0.4)$$

And  $\text{deaths}_{ij} \sim \text{Pois}(dmu_{ij})$ , and  $dmu_{ij}$  is dependent on caseloads 2 weeks prior<sup>29</sup> and given by

$$\log(dm_{ij}) = \log(\beta_{2ij}) + \log(I_{i,j-2})$$

$$\log(\beta_{2ij}) = b_2 0 + b_3 \times (p_{70plus})_i + b_4 \times (e\text{-IMD})_i + u_{2i} + v_{2i}$$

Here  $b = (b_0, b_1, b_2)$  is the vector of regression coefficients for the intercept (representing the log-transformed baseline transmission rate across all locations), mobility represents the observed weekly mobility in a given MSOA based on the mobility in the CCG containing most of that MSOA; mobility  $\times$  population density represents an interaction term to capture the varying impact of mobility change as a function of changing population density at MSOA level;  $p_{70plus}$  plus represents the proportion of the population in a given MSOA that were 70 years or older in 2020 in a given MSOA;  $e\text{-IMD}$  represents the proportion of the elderly people living in deprivation in a given MSOA in 2019;  $u_i$  corresponds to structured (spatial) heterogeneity and represents spatial variation in transmission rate between regions that captures the effects of unobserved variables with an underlying spatial pattern;  $v_i$  correspond to geographically unstructured (i.e. random effect) heterogeneity in the transmission rate; and  $\mu_{ij}$  is the weekly transmission rate at MSOA  $i$  and week  $j$  after incorporating the spatiotemporal effects of local social-demographic factors, the transmission dynamics of COVID-19 outbreak (i.e. SEIR model) and random effects ( $u_i$  and  $v_i$ ). All covariate coefficients had diffuse normal priors, given by  $b \sim N(0.0, 1.0E6)$ . The variances

of the random effects had uniform priors,  $\sigma_u \sim U(0,5)$  and  $\sigma_v \sim U(0,5)$ . We assumed that the asymptomatic proportion parameter, *asym*, was constrained to between 5 to 40%. This is guided by two systematic reviews<sup>30,31</sup> which suggest that between 17% (95% CI 14–20%) and 20% (95% CI 17–25%) of infections respectively are asymptomatic. However, given that there is still substantial heterogeneity across these studies, we chose a more conservative (wider uncertainty) range for this fraction in our model, hence the distributional assumption of 5 to 40% in our model.

A convolution CAR model<sup>32</sup> was applied to decompose the log of area-level relative risks into the sum of two random effects, namely a structured spatial dependency effect and unstructured areal level heterogeneity effect through the random effects  $u_i$  and  $v_i$  respectively. The convolution approach has demonstrated good robust estimates across multiple studies and a range of disease clustering/modelling scenarios, for example<sup>33</sup>. This spatial structured effect models the effect of proximity using a first-order neighbourhood structure (i.e. MSOA adjacency matrix), whereby the random effect is assumed to have a normal distribution, with the conditional weighted mean given by the average of the neighbours.

We conducted various sensitivity analyses to estimate the impact of social distancing (as proxied by the daily population mobility covariate) on COVID-19 caseloads in England. The model coefficients for the mobility covariates from the full multivariable space–time model was applied to the baseline mobility prior to lockdown and propagated throughout the period assuming no change in mobility from baseline (that is, daily mobility patterns remained unchanged). Secondly, the observed rate of change of mobility at week 34 was propagated forward to week 51. Lastly, we assumed various fractions of mobility gain post week 34 and applied these to the difference between the counterfactual of no reduction in population movement and observed first wave outbreak dynamics under full lockdown to characterise the likely prospective trajectory in England in the coming weeks.

A comparison of model fit was performed using the deviance information criterion (DIC). Posterior distributions for parameters of interest were obtained through Markov chain Monte Carlo (MCMC) sampling. Convergence was assessed by visual inspection of the trace/autocorrelation plots for the sample chains, using Gelman–Rubin statistical plots<sup>34</sup> and by confirming the MCMC error for all posteriors were less than 5% of the standard deviation for a given posterior. We ran 100,000 MCMC iterations and discarded the first 10,000 MCMC iterations as part of the model burn-in. We extracted the mean point estimate for a given posterior as well as the 95% uncertainty intervals (2.5 to 97.5 percentile). This analysis was performed using WinBUGS version 1.4.3.

We validated the performance of the model using out of sample prediction by comparing the predicted number of cases from the model post week 34 (observed case data truncation point) with the observed cases to prove its utility as an early warning (pre-emptive) tool for resource allocation and prioritization by MSOA resolution (Supplementary Material Section 3).

Received: 22 October 2020; Accepted: 8 February 2021

Published online: 08 March 2021

## References

- Lillie, P. J. *et al.* Novel coronavirus disease (Covid-19): the first two patients in the UK with person to person transmission. *J. Infect.* **80**, 578–606 (2020).
- Hellewell, J. *et al.* Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts. *Lancet Glob. Health* **8**(4), e488–96 (2020).
- Du, Z. *et al.* Risk for transportation of coronavirus disease from Wuhan to other cities in China. *Emerg. Infect. Dis.* **26**, 1049 (2020).
- Wu, J. T., Leung, K. & Leung, G. M. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. *The Lancet* **395**, 689–697 (2020).
- Bootsma, M. C. & Ferguson, N. M. The effect of public health measures on the 1918 influenza pandemic in US cities. *Proc. Natl. Acad. Sci.* **104**, 7588–7593 (2007).
- Eggo, R. M., Cauchemez, S. & Ferguson, N. M. Spatial dynamics of the 1918 influenza pandemic in England, Wales and the United States. *J. R. Soc. Interface* **8**, 233–243 (2011).
- Fuhrmann, C. The effects of weather and climate on the seasonality of influenza: what we know and what we need to know. *Geogr. Compass* **4**, 718–730 (2010).
- Lawson, A. B. & Song, H.-R. Bayesian hierarchical modeling of the dynamics of spatio-temporal influenza season outbreaks. *Spat. Spatio-temporal Epidemiol.* **1**, 187–195 (2010).
- Lawson, A. B. *Statistical Methods in Spatial Epidemiology* (Wiley, Hoboken, 2013).
- Lawson, A. B. & Kim, J. Space-Time Covid-19 Bayesian SIR modeling in South Carolina. arXiv preprint [arXiv:2010.15207](https://arxiv.org/abs/2010.15207) (2020).
- Peto, J. *et al.* Universal weekly testing as the UK COVID-19 lockdown exit strategy. *The Lancet* **395**, 1420–1421 (2020).
- Kraemer, M. U. *et al.* The effect of human mobility and control measures on the COVID-19 epidemic in China. *Science* **368**, 493–497 (2020).
- Chinazzi, M. *et al.* The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. *Science* **368**, 395–400 (2020).
- Vollmer, M. A. *et al.* Report 20: using mobility to estimate the transmission intensity of COVID-19 in Italy: a subnational analysis with future scenarios. (Technical Report May. Imperial College COVID-19 Response Team, 2020).
- Prem, K. *et al.* The effect of control strategies to reduce social mixing on outcomes of the COVID-19 epidemic in Wuhan, China: a modelling study. *Lancet Public Health* (2020).
- Lau, H. *et al.* The positive impact of lockdown in Wuhan on containing the COVID-19 outbreak in China. *J. Travel Med.* **27**, taaa037 (2020).
- Ji, T. *et al.* Lockdown contained the spread of 2019 novel coronavirus disease in Huangshi city, China: early epidemiological findings. *Clin. Infect. Dis.* **71**(6), 1454–60 (2020).
- Chintalapudi, N., Battineni, G. & Amenta, F. COVID-19 disease outbreak forecasting of registered and recovered cases after sixty day lockdown in Italy: A data driven model approach. *J. Microbiol. Immunol. Infect.* **53**(3), 396–403 (2020).
- Matrajt, L. & Leung, T. Evaluating the effectiveness of social distancing interventions to delay or flatten the epidemic curve of coronavirus disease. *Emerg. Infect. Dis.* **26**(8), 1740 (2020).
- Chauhan, S., Misra, O. P. & Dhar, J. Stability analysis of SIR model with vaccination. *Am. J. Comput. Appl. Math.* **4**, 17–23 (2014).

21. Ripa, M. *et al.* Secondary infections in patients hospitalized with COVID-19: incidence and predictive factors. *Clin. Microbiol. Infect.* <https://doi.org/10.1016/j.cmi.2020.10.021> (2020).
22. Chen, N. *et al.* Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *The Lancet* **395**, 507–513 (2020).
23. Levin, A. T. *et al.* Assessing the age specificity of infection fatality rates for COVID-19: systematic review, meta-analysis, and public policy implications. *Eur. J. Epidemiol.* **8**, 1–16 (2020).
24. Davies, N. G. *et al.* Age-dependent effects in the transmission and control of COVID-19 epidemics. *MedRxiv* (2020).
25. Stevens, G. A. *et al.* Guidelines for accurate and transparent health estimates reporting: the GATHER statement. *PLoS Med.* **13**, e1002056 (2016).
26. Ord, J. K. & Getis, A. Local spatial autocorrelation statistics: distributional issues and an application. *Geogr. Anal.* **27**, 286–306 (1995).
27. Lawson, A. B., Rotejanaprasert, C., Moraga, P. & Choi, J. A shared neighbor conditional autoregressive model for small area spatial data. *Environmetrics* **26**, 383–392 (2015).
28. Lauer, S. A. *et al.* The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: estimation and application. *Ann. Intern. Med.* **172**, 577–582 (2020).
29. Baud, D. *et al.* Real estimates of mortality following COVID-19 infection. *Lancet Infect. Dis.* **20**(7), 773 (2020).
30. Buitrago-Garcia, D. *et al.* Occurrence and transmission potential of asymptomatic and presymptomatic SARS-CoV-2 infections: a living systematic review and meta-analysis. *PLoS Med.* **17**, e1003346 (2020).
31. Byambasuren, O. *et al.* Estimating the extent of true asymptomatic COVID-19 and its potential for community transmission: systematic review and meta-analysis. Available at SSRN 3586675 (2020).
32. Besag, J., York, J. & Mollié, A. Bayesian image restoration, with two applications in spatial statistics. *Ann. Inst. Stat. Math.* **43**, 1–20 (1991).
33. Best, N., Richardson, S. & Thomson, A. A comparison of Bayesian spatial models for disease mapping. *Stat. Methods Med. Res.* **14**, 35–59 (2005).
34. Gelman, A. & Rubin, D. B. Inference from iterative simulation using multiple sequences. *Stat. Sci.* **7**, 457–472 (1992).

## Acknowledgements

We thank Professor Graham Medley and Professor Robin Bailey for their useful comments and insights.

## Author contributions

B.S. and R.P. conceived and designed the study. A.L. assisted with adaptation of the statistical framework. B.S. acquired and analysed the data. All authors interpreted the data. B.S. drafted the manuscript. All authors critically revised the manuscript and approved the final version for submission.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-83780-2>.

**Correspondence** and requests for materials should be addressed to B.S.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021, corrected publication 2021