RESEARCH ARTICLE

SCoVMod – a spatially explicit mobility and deprivation adjusted model of first wave COVID-19 transmission dynamics [version 1; peer review: 2 approved]

Christopher J. Banks[1], Ewan Colman[1], Thomas Doherty[2], Oliver Tearne[3], Mark Arnold[3], Katherine E. Atkins[4,5], Daniel Balaz[1], Gaël Beaunée[6], Paul R. Bessell[1], Jessica Enright[7], Adam Kleczkowski[2], Gianluigi Rossi[1], Anne-Sophie Ruget[1], Rowland R. Kao[1,8]

[1]Roslin Institute, University of Edinburgh, Edinburgh, EH25 9RG, UK
[2]Mathematics & Statistics, University of Strathclyde, Glasgow, G1 1XH, UK
[3]The Animal and Plant Health Agency, Weybridge, Surrey, KT15 3NB, UK
[4]Faculty of Epidemiology and Population Health, on School of Hygiene and Tropical Medicine, WC1E 7HT, UK
[5]Usher Institute, University of Edinburgh, Edinburgh, EH8 9AG, UK
[6]INRAE, Nantes, 44300, France
[7]School of Computing Science, University of Glasgow, Glasgow, G12 8RZ, UK
[8]Royal (Dick) School of Veterinary Studies, University of Edinburgh, Edinburgh, EH25 9RG, UK

v1 **First published:** 26 May 2022, **7**:161
https://doi.org/10.12688/wellcomeopenres.17716.1
Latest published: 26 May 2022, **7**:161
https://doi.org/10.12688/wellcomeopenres.17716.1

Abstract

Background: Mobility restrictions prevent the spread of infections to disease-free areas, and early in the coronavirus disease 2019 (COVID-19) pandemic, most countries imposed severe restrictions on mobility as soon as it was clear that containment of local outbreaks was insufficient to control spread. These restrictions have adverse impacts on the economy and other aspects of human health, and it is important to quantify their impact for evaluating their future value.
Methods: Here we develop Scotland Coronavirus transmission Model (SCoVMod), a model for COVID-19 in Scotland, which presents unusual challenges because of its diverse geography and population conditions. Our fitted model captures spatio-temporal patterns of mortality in the first phase of the epidemic to a fine geographical scale.
Results: We find that lockdown restrictions reduced transmission rates down to an estimated 12\% of its pre-lockdown rate. We show that, while the timing of COVID-19 restrictions influences the role of the transmission rate on the number of COVID-related deaths, early reduction in long distance movements does not. However, poor health associated with deprivation has a considerable association with mortality; the Council Area (CA) with the greatest health-related deprivation was found to have a mortality rate 2.45 times greater than

Open Peer Review

Approval Status ✓ ✓

	1	2
version 1 26 May 2022	✓ view	✓ view

1. **Christopher Davis**, University of Warwick, Coventry, UK

2. **Cameron Zachreson**, The University of Melbourne, Melbourne, Australia

Any reports and responses or comments on the article can be found at the end of the article.

the CA with the lowest health-related deprivation considering all deaths occurring outside of carehomes.
Conclusions: We find that in even an early epidemic with poor case ascertainment, a useful spatially explicit model can be fit with meaningful parameters based on the spatio-temporal distribution of death counts. Our simple approach is useful to strategically examine trade-offs between travel related restrictions and physical distancing, and the effect of deprivation-related factors on outcomes.

Keywords
COVID-19, epidemic, spatio-temporal, model, deprivation, mobility, health

Corresponding author: Rowland R. Kao (Rowland.Kao@ed.ac.uk)

Author roles: Banks CJ: Data Curation, Formal Analysis, Investigation, Methodology, Resources, Software, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Colman E**: Data Curation, Formal Analysis, Investigation, Methodology, Software, Writing – Original Draft Preparation, Writing – Review & Editing; **Doherty T**: Methodology, Software, Writing – Original Draft Preparation, Writing – Review & Editing; **Tearne O**: Methodology, Software, Writing – Review & Editing; **Arnold M**: Supervision, Writing – Review & Editing; **Atkins KE**: Methodology, Writing – Review & Editing; **Balaz D**: Validation, Writing – Review & Editing; **Beaunée G**: Methodology, Writing – Review & Editing; **Bessell PR**: Formal Analysis, Investigation, Methodology, Writing – Review & Editing; **Enright J**: Formal Analysis, Methodology, Writing – Review & Editing; **Kleczkowski A**: Methodology, Writing – Review & Editing; **Rossi G**: Investigation, Methodology, Writing – Review & Editing; **Ruget AS**: Formal Analysis, Methodology, Visualization, Writing – Review & Editing; **Kao RR**: Conceptualization, Funding Acquisition, Investigation, Methodology, Project Administration, Supervision, Writing – Original Draft Preparation, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: This work was supported by Wellcome [209818; to RRK]; the Roslin ISP2 (theme 3) [BBS/E/D/20002174]; the BBSRC [BB/P010598/1]; the Strategic Blue Cloud Fund to Fight COVID-19 grant [003]. The simulation model was developed from a code base developed under Defra projects [SE3290 and SE3292].
The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

How to cite this article: Banks CJ, Colman E, Doherty T *et al.* **SCoVMod – a spatially explicit mobility and deprivation adjusted model of first wave COVID-19 transmission dynamics [version 1; peer review: 2 approved]** Wellcome Open Research 2022, **7**:161 https://doi.org/10.12688/wellcomeopenres.17716.1

First published: 26 May 2022, **7**:161 https://doi.org/10.12688/wellcomeopenres.17716.1

Introduction

The coronavirus disease 2019 (COVID-19) pandemic resulted in the introduction of severe non-pharmaceutical interventions (NPIs) to control its spread in most countries around the world. COVID-19 was introduced into Scotland no later than February 2020. Following a series of lesser restrictions and recommendations, on 23rd March a *lockdown* with widespread NPIs was put in place across the UK, which resulted in cessation of all non-essential activities, many businesses closed and most individuals except *key workers* were restricted to only short distance trips from their homes, initially with a maximum frequency of only once per day. These measures reduced the transmission of COVID-19 substantially, so that by the first week of April, the average infected individual in Scotland was estimated to be infecting fewer than one other individual each (i.e. causing the reproduction number R to fall below one).

While we therefore know that the combination of measures was effective, no study as yet has directly estimated the impact of different aspects of these control policies. One factor that makes this difficult for the evaluation of any such national level policy is the impact of variation in space of the disease transmission itself. For COVID-19, one critical issue is the known substantial variation in COVID-19 infection and mortality risk associated with deprivation[1]. The impact of deprivation is likely due to a combination of factors influencing both exposure (e.g. more crowded housing and working conditions) and mortality once exposed (e.g. due to already poorer health) and this should be accounted for in any assessment of the impact of restrictions.

Here, we develop an explicitly spatial agent-based simulation model that accounts for recorded movements-to-work (i.e. "commuter" patterns), modulated by recorded time-varying mobility statistics, and geographically explicit population age structures. Whilst this does not capture all human movement, it is expected to capture a large proportion of the long-range mobility especially affected by lockdown. We also modify the modelled epidemic using deprivation metrics. In these early epidemic stages case data were unevenly recorded, therefore we use this model to estimate transmission characteristics by fitting it to the observed number of COVID-19 related deaths using Approximate Bayesian inference. Our aim is to estimate the impact of travel restrictions and transmission reduction on the spread of COVID-19, and to assess the impact of measures of deprivation in order to estimate its impact on COVID-19 related mortality.

The impact of NPIs is of considerable interest and has been the subject of several analyses, however previous results have focused on statistical comparisons of interventions across multiple countries[2–4]. Such approaches are extremely useful for identifying the impact of events, but do not directly allow for disentangling the impact of simultaneous events. Long distance movement and local restrictions are typically applied contemporaneously, and our approach that includes processes at the two scales (local and regional) allows us to estimate the impact of each.

Methods

Data sources

We use a combination of publicly available Scottish census data from National Records for Scotland (NRS)[1] and data on COVID-19 held by Public Health Scotland (PHS) and made available via the PHS Electronic Data Research and Innovation Service (eDRIS)[2]. We used data zone (DZ) level resolution where DZs are population census units of approximately 500 to 1,000 residents. The data for assignment of individuals to work locations is drawn from the NRS Census Flows data[3], Table WU01UK, which provides origin/destination workplace data for the population from the 2011 census. We adjust these with respect to the 2018 population estimates.

Age demographics and movement to work patterns are available at the level of Census Output Areas (OA), each of which contains approximately 20 households or 50 people[4].

Census data on the Scottish Index of Multiple Deprivation (SIMD)[5] considers multiple relative deprivation measures and combines them into a single value. Deprivation data are publicly available at the DZ level.

We also used publicly available data from Google to estimate mobility levels over time, with respect to commuting patterns[6].

Model overview

The model we present here breaks down into a number of distinct parts, each of which we describe over the following sections. The code is available from GitHub and is archived with Zenodo. First, to account for the variation in mortality risk due to deprivation factors we develop a statistical model that provides a mortality parameter adjustment to the main model. The core of the simulation model then breaks down into the following parts:

- Local transmission—a homogeneous mixing compartmental model for each unit area of the country;
- National transmission—a network-based simulation of the movement of individuals between unit areas;
- Lockdown simulation—the reduction of both local and national transmission to simulate non-pharmaceutical interventions;
- Parameter inference—a Bayesian estimation of the parameters for local transmission.

[1]https://statistics.gov.scot/

[2]https://www.isdscotland.org/Products-and-Services/EDRIS/COVID-19/

[3]https://wicid.ukdataservice.ac.uk

[4]https://www.scotlandscensus.gov.uk/about/2011-census/2011-census-geographies/, https://www.nrscotland.gov.uk/files/geography/2011-census/geography-bckground-info-comparison-of-thresholds.pdf

[5]https://www.gov.scot/collections/scottish-index-of-multipledeprivation-2020/

[6]https://www.google.com/covid19/mobility/

The simulation framework (Scotland Coronavirus transmission Model, or SCoVMod) considers key aspects of COVID-19 epidemiology including phases for latent infection, infectious and mildly infected (showing few or no clinical signs) and severely infected (with substantial clinical signs) individuals, hospitalised, recovered and died, similar to other investigations[5,6]. These epidemiological processes are captured as individual disease states (Figure 1). Individuals are also stratified into three age groups: young (0–15), adult (16–64) and elderly (65+). Within-OA transmission is assumed to be homogeneously mixed, while between-OA transmission is determined by the empirical age-specific patterns of home and work contact (creating day/night patterns of contact). We do not consider overnight shifts in location or introductions from outside Scotland beyond the impact on the initial seeding. Death rates for all age classes are assumed to be the same, but with differing recovery rates, resulting in age-dependent differences in the infection-related mortality, as is consistent with the data at the time (see Table 1). Deprivation is also known to influence COVID-19 mortality[7]; we therefore adjust mortality in the model with the average health index in the local council area (see below).

We make a number of simplifying assumptions regarding transmission pathways. First, infections in care homes are not modelled at this stage of the epidemic, as they are assumed to result in few additional infections outside of these locations. Second, hospital-acquired infections are considered 'dead ends', following evidence that they do not result in substantial community outbreaks[8]. Third, population mobility patterns are determined by the patterns of movements to work, recorded in Scottish Census data. We assume that only adults contribute to commuter movement, in the daytime; the remaining proportion of adults and all young and elderly individuals are assumed to move primarily within their local OAs, which also account for non-work activities. Finally, commuting is restricted to healthy and exposed or mildly symptomatic individuals; severely infected and hospitalised individuals do not commute. The day/night patterns also result in two transmission rates.

Health index and mortality

Using the SIMD we constructed two initial statistical models, with a view to guiding the parameters of the core simulation model. These consisted of:

1. A population level mortality model
2. An excess deaths model

The models were generalised linear models with binomial error structures at the local Council Area (CA) level (N=32). For the population level model the outcome variable was c(Covid deaths, population – Covid deaths). For the excess mortality model c(Covid deaths, all deaths – Covid deaths). To correct for overdispersion we fitted the model with an individual level random effect for each data point using the glmer function in the lme4 package for R (R package: lme4, RRID:SCR_015654), in multivariable models variance inflation factors were checked using the car package (R package: car, RRID:SCR_022137) and overdispersion using the DHARMa package (R package: DHARMa, RRID:SCR_022136). We also checked the results against a model fitted using the quasi-binomial family. In both models we tested the following covariates in univariable analysis: Population (from SIMD 2020), Population density, SIMD 2016 score, score without the access component, income indicator, employment indicator, health indicator, education indicator, housing indicator, access indicator, and crime indicator.

Local transmission (within-OA)

Within each OA (i) the infection process is governed by a compartmental model for which the frequency dependent force of infection $\Lambda_i(t)$ defined in Figure 2. In the compartmental model are infection classes S (susceptible), E (exposed), I^M

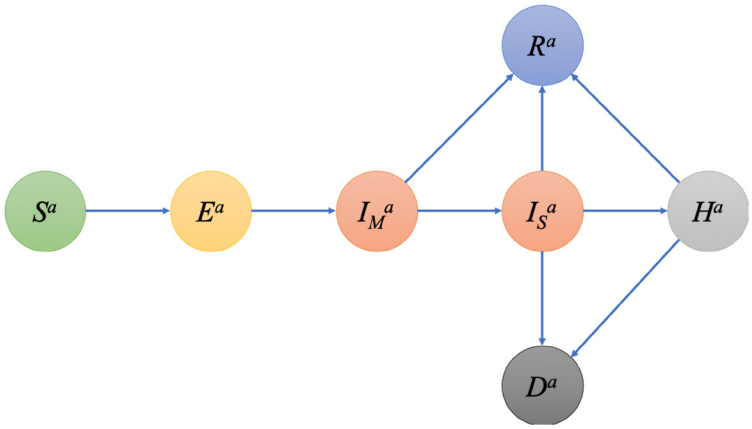

Figure 1. Schematic of infection stages in SCoVMod. Individuals pass through stages post infection as described by arrows. Not all stages are obligatory for all infected individuals (e.g. some individuals recover without going to hospital). SCoVMod, Scotland Coronavirus transmission Model.

Table 1. Epidemiological parameters in SCoVMod, with priors and fixed values as appropriate. Where age is not indicated, parameters are assumed to be age independent. All times are measured in days. SCoVMod, Scotland Coronavirus transmission Model; eDRIS, Electronic Data Research and Innovation Service.

Parameter	Transition	Symbol	Age	Value	Prior	References
Latency period	$E \to I^M$	$1/\gamma$	All	fitted	U(1.67,28)	5,9–12
Days from mild infectiousness to recovery	$I^M \to R$	$1/\rho_M$	All	fitted	U(0.67,28)	6,12
Symptom onset time after infectiousness	$I^M \to I^S$	$1/\gamma_M$	All	fitted	U(2,28)	5,9–16
Transmission rate for severe infectors (baseline, daytime)	$S \to E$	β_d	All	fitted	U(0,2.8)	
Transmission rate for severe infectors (baseline, nightime)	$S \to E$	β_n	All	fitted	U(0,2.6)	
Transmission rate multiplier for mild infectors	$S \to E$	y	All	fitted	U(0,2.6)	
Severe symptom onset to hospitalization	$I^S \to H$	$1/\eta$	All	4		15–22
Severe symptom onset to recovery for non-hospitalised	$I^S \to R$	$1/\rho_S$	Young Adults Elderly	19 20.7 21.6		15
Days hospitalisation to death	$H \to D$	$1/\mu_H$	Young Adults Elderly	 6.97 6.62		eDRIS data eDRIS data eDRIS data
Proportion of hospitalised who recover	$H \to R$	$\rho_H/(\rho_H+\mu_H)$	Young Adults Elderly	1 0.96 0.84		eDRIS data eDRIS data eDRIS data
Symptoms onset to death	$I^S \to D$	$1/\mu_S$	Adults	16		16,18,21–24
Mortality rate multiplier (relative to average health index)		μ_{mod}	All	fitted	U(0,0.08)	
Number of seed infections		N_s	N/A	fitted	U(0, 2000)	

Within each OA i the infection process is governed by the frequency dependent force of infection at time t:

$$\Lambda_i(t) = -\Bigg[(\beta_N + \beta_D) \sum_{a \in Y,E} \left(y I_{ia}^M + I_{ia}^S \right) + \beta_N \left(y I_{ia}^M + I_{ia}^S \right) + \beta_D I_{ia}^S + \beta_D \left\{ \left\{ \sum_j \left(1 - \sum_j x_{ij} \right) \left(y I_{ia}^M \right) \right\} + \sum_j \left\{ x_{ij} \left(\sum_k (1 - x_{jk}) \left(y I_{ja}^M \right) + I_{ja}^S + \sum_{a \in Y,E} \left(y I_{ja}^M + I_{ja}^S \right) \right) \right\} \right\} \Bigg] / N_i$$

where:

- β_N and β_D are the nighttime and daytime transmission rates,
- $a \in \{Y, A, E\}$ is the age class (Young, Adult, Elderly),
- y is the transmission rate modifier for mildly infected individuals,
- I_{ia}^M is the number of mildly infected at location i in age class a,
- I_{ia}^S is the number of severely infected at location i in age class a,
- x_{ij} is the proportion of individuals commuting between locations i and j,
- N_i is the population of location i.

Figure 2. Equation: Force of infection for location *i* at time *t*.

(mildly infected), I^S (severely infected), H (hospitalised). Model equations for individuals residing in one OA i and for age class a are therefore:

$$\frac{dS_{ia}}{dt} = -\Lambda_i(t)S_{ia}$$
$$\frac{dE_{ia}}{dt} = \Lambda_i(t)S_{ia} - \gamma E_{ia}$$
$$\frac{dI^M_{ia}}{dt} = \gamma E_{ia} - (\gamma_M + \rho_M)I^M_{ia}$$
$$\frac{dI^S_{ia}}{dt} = \gamma_M I^M_{ia} - (\rho_{Sa} + \mu_{Sia} + \eta)I^S_{ia}$$
$$\frac{dH_{ia}}{dt} = \eta I^S_{ia} - (\rho_{Ha} + \mu_{Hia})I^S_{ia}$$
$$\frac{dR_{ia}}{dt} = \rho_M I^M_{ia} + \rho_{Sa} I^S_{ia} + \rho_{Ha} H_{ia}$$

The number of deaths is determined by:

$$D_{ia} = N_{ia} - (E_{ia} + I^M_{ia} + I^S_{ia} + H_{ia} + R_{ia})$$

The rates of infection are detailed in Figure 2 and other state transition rates are given by:

- γ for $E \rightarrow I^M$
- γ_M for $I^M \rightarrow I^S$
- η for $I^S \rightarrow H$
- ρ_M for $I^M \rightarrow R$
- ρ_{Sa} for $I^S \rightarrow R$ for age class a
- ρ_{Ha} for $H \rightarrow R$ for age class a
- μ_{Sia} for $I^S \rightarrow D$ for age class a and location i
- μ_{Hia} for $H \rightarrow D$ for age class a and location i

Mortality rates are adjusted by location, according to Health Index (see results of Health Index and Mortality):

$$\mu_{CA} = \mu_{av}\mu_{mod}\left(1 + \left(\frac{k_{CA} - k_{av}}{k_{av}}\right)\right)$$

where μ_{CA} is the COVID-19 related mortality rate for a given CA, μ_{av} is the average across CAs, k_{CA} is the CA mean health index value (from the SIMD), and μ_{mod} is a fitted parameter given a prior range of [0, 0.08] in order to preclude negative values for low values of k_{CA}.

The values for all parameters are either established from the literature (Table 1) or fit (see below).

National transmission (between-OA movements)

Between OAs individuals move daily across a network of locations defined by Scottish Census data adjusted by Google mobility data.

From the current population estimates we draw the number of individuals whose primary residence is mapped onto an OA, with their age group. The total population of Scotland from this estimate is 5,438,054 (Young: 919,580; Adult: 3,492,421; Elderly: 1,026,053). Of the adults 1,960,712 commute to work (reduced to 647,034 under lockdown (see details below).

An individual's workplace is assigned by distributing a proportion of the population of each location to each work location, weighted by the proportion of individuals from each home location in the census flows data who work in another location. For each origin o and destination d we assign a weight w_{od} from the census flow data:

$$w_{od} = \frac{n_{od}}{t_o}$$

where n_{od} is the total number of people who move from o to d to work, and t_o is the total number who move from origin o to any location for work. We take the individuals of each home location if they are eligible to work (total n_o); in this case we assume all individuals of adult age 16–65. Each destination is assigned to $n_o \times w_{od}$ of these individuals. The individuals who remain have no assigned workplace—either they do not work, or they work within their home location.

For each day of the simulation we consider two time steps: a day step where individuals can move to their place of work, and a night step where those individuals move back to their home location. In each day step, we take each destination location d. Let λ_d be the number of eligible workers who may move to the destination location. For each day the sampled number who move s is drawn from a Poisson distribution: $s \sim Poisson(\lambda_d)$. The sampled number of moves s is then scaled according to the percent change in mobility m (see below) for the given day: $s_m = \lfloor s(1 + \frac{m}{100}) \rfloor$.

In order to improve the computational efficiency of the simulation, movements of commuters between OAs were batched into groups of five, with movements between OAs of fewer than five individuals per day retained at a proportionate rate by drawing from a binomial distribution: $s_{mt} \sim B(s_m, \frac{1}{5})$. If the sampled number of workers s_{mt} is less than or equal to the number of workers who may normally move to destination d, then those who move are sampled randomly from those who may normally move. However, if s_{mt} is greater than the number of workers who may normally move to d, then the additional workers are drawn randomly from workers who have no assigned destination location. While this reduces the overall network link density, the effect on transmission dynamics is expected to be small. We note that this means that interpretation of the combined β_D and β_N must be made with caution.

For each night of the simulation, the workers who moved in the day step are moved back to their origin location.

Modelling lockdown

To model reductions in activity that are the aim of lockdown, we consider two factors. First, we thin movements in the simulation (mobility reduction) in proportion to the observed changes in mobility[7] (as above), also checking consistency with social contact surveys[25]. This was corroborated for remote areas, where mobility data are few, with an independent dataset (see Figure 3). Second, physical distancing is incorporated via a reduction in contacts applied to both daytime and nighttime transmission rates (transmission reduction). For the post-lockdown period, we contingently fit a reduction in transmission only, assuming that posterior distributions for all other parameters estimated based on the pre-lockdown fit remain relevant, and considering a range of values consistent with independent estimates of the reproduction number R_t[8].

Model inference and computation

Simulated epidemics are compared to the spatio-temporal pattern of COVID-19 spread in Scotland. Non-observable parameters were estimated using the number of deaths where COVID-19 is recorded, considering all weeks beginning 9th March and ending on the 12th April 2020. Recorded deaths are from the weekly PHS records identified by the DZ of residence[9]. These are the most complete and unbiased indicators of infection available during the early epidemic. However, they differ from other official sources as they record the date of registration of death. We assume that for each reported case, death occurs in the week prior to registration, which is allowed up to 8 days postmortem.

Estimation was performed using a Sequential Monte Carlo implementation of Approximate Bayesian Computation (ABC-SMC)[26,27]. We calibrated the model output to the cumulative weekly number of deaths due to COVID-19 aggregated at the level of CAs, using this spatial variation in deaths across Scotland to provide the necessary signature to properly calibrate the role of human mobility. Preliminary attempts to fit the data using weekly incidence had poor results (results not shown). We hypothesise that this is a result of a combination of very few deaths per CA to constrain the fits in the early stages of the epidemic, and also the complications from the multiple changes in the control efforts especially through March. Thus we chose to emphasise the later stages using as our observation the cumulative number of all COVID-19 related deaths per CA.

Simulated and observed summary statistics are compared via a score equal to a sum of squared errors, recorded weekly:

$$score = \sum_{week}\sum_{CA}\left((D_{sim} - D_{obs})(max(D^Z_{obs}))\right)^2 + \sum_{week}\sum_{CA}\left((D^Z_{sim} - D^Z_{obs})(max(D_{obs}))\right)^2$$

where D_{sim} is the cumulative number of deaths per CA simulated and D_{obs} its observed value, D^Z_{sim} is the cumulative number of data zones with deaths per CA simulated and D^Z_{obs} its observed value. $max(x)$ is the maximum number of x over all CAs, used to balance the relative contributions of the two elements

[7]Google Community Mobility Reports: https://www.google.com/covid19/mobility/

[8]COVID-19: Framework for Decision Making, Further Information (published 23rd April, 2020), accessed 11th May, 2020: https://www.gov.scot/publications/coronavirus-COVID-19-framework-decision-making-further-information/

[9]Data provided by eDRIS is subject to data sharing agreement due to potential for individual disclosure.

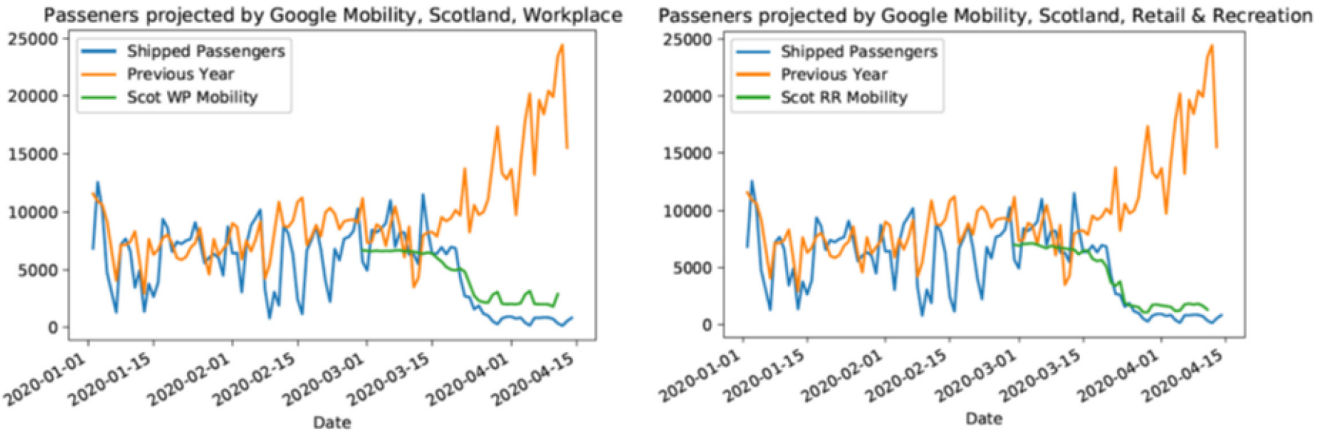

Figure 3. Comparison of Google mobility data for Scotland to CalMac Ferry records. Comparison to workplace mobility (left) Comparison to Recreation mobility (right). The comparison is relative to the mean value prior to lockdown on March 23rd, 2020. In order to provide confidence that inclusion of mobility reductions across regions is appropriate, we assume that urban areas such as Glasgow and Edinburgh are likely to be well represented, but that rural areas may be less so. To check this, we compare an independent dataset on independent sailings and passenger numbers for ferry services run by Caledonian MacBrayne, who operate all ferry services in the west of Scotland. A comparison of data from 2019 to 2020 and to Google Mobility data, shows a strong fidelity between the two datasets, as well as a substantial reduction in activity at point of lockdown. The similarity prior to lockdown between 2019 and 2020 also suggests that patterns of increased summer activity are unlikely to have had strong influences on our assumptions regarding commuter movements, at least in this area.

of the metrics to the inference. The total number of infected individuals at the start of the simulation (the seeds) are fitted as part of the inference. The seeds are randomly assigned a disease state from E, I^M, and I^S. Seed locations are stochastically assigned according to the cumulative proportion of deaths registered per Intermediate Zone, as recorded up to the week of March 23rd, 2020. Intermediate zones are aggregates of approximately five DZs; this scale is chosen to account for clustering of infections in areas near to identified deaths.

Uniform prior distributions constrain all parameter values to plausible ranges based on the available literature relevant to the early, pre-lockdown period. Infection dynamics are simulated via a τ-leap algorithm using halfday timesteps[28]. All parameters are listed in Table 1. The inference framework is run on a distributed application framework (Akka)[10] running on a cloud computing infrastructure (Amazon AWS2)[11]. The model code has been written using industry grade software engineering practices including agile development for project task planning, test driven development, pair programming and code reviews to produce unit tested, robust, and reusable software components. The majority of the code has been reviewed by a second software developer.

[10]https://akka.io

[11]https://aws.amazon.com

Results

Health index and mortality

The results of univariable analyses of the SIMD variables for the population level is presented in Table 2. As access has the lowest Akaike information criterion (AIC) and Bayesian information criterion (BIC) and is orthogonal to the other variables we tested it against the remaining variables in turn to check for a multivariable model. The model with access and health had the lowest AIC (276.3), BIC (282.2) and low variance inflation factor (1.291) and was not overdispersed (Table 3). We analysed death records from Public Health Scotland (PHS) data on non-care home deaths due to COVID-19. We examined the impact of different deprivation factors relevant to this period; whilst deprivation overall is significantly associated with increased COVID-19 mortality, this could be further disaggregated:

1. Population level risk of COVID-19 mortality is associated with the SIMD indicator that describes (good) accessibility and orthogonally with the SIMD indicator that

Table 2. Results of the univariable model with Covid deaths / population as the outcome. SIMD, Scottish Index of Multiple Deprivation; AIC, Akaike information criterion; BIC, Bayesian information criterion.

Predictor	Estimate	St. error	p-value	AIC	BIC
Population / 100000	0.0137	0.067	1.063	303.2	307.6
SIMD score	0.046	0.012	<0.001	291.5	295.9
SIMD education indicator	0.036	0.013	0.005	297	301.4
SIMD health indicator	0.042	0.008	<0.001	283.1	287.5
SIMD housing indicator	0.025	0.009	0.006	297.4	301.8
SIMD access indicator	-0.04	0.009	<0.001	283.8	288.2
SIMD crime indicator	0.049	0.013	<0.001	290.7	295.1
SIMD income indicator	0.04	0.009	<0.001	289.8	294.2
SIMD employment indicator	0.041	0.009	<0.001	287.2	291.6
SIMD score (exc access)	0.047	0.011	<0.001	288.2	292.6
$\sqrt{\text{population density}}$	0.017	0.006	0.003	295.9	300.3

Table 3. Results of the multivariable model with Covid deaths / population as the outcome. SIMD, Scottish Index of Multiple Deprivation.

Variable	Estimate	St. error	p-value
Intercept	-7.91	0.320	<0.001
SIMD access indicator	-0.024	0.008	0.004
SIMD health indicator	0.029	0.008	<0.001

describes (poor) health. Indicating that areas with poorest health and good access experienced higher COVID-19 mortality.

2. Risk of excess COVID-19 mortality (COVID-19 deaths as a fraction of all deaths) is most closely associated with the access indicator component of SIMD. This indicates that the areas that have good local connectivity and transport will have higher rates of COVID-19 transmission.

3. We hypothesise that access is a proxy for earlier introduction of infection and the mobility patterns in our data. Thus the observed differences in mortality due to access is already accounted for in our model (via the initial seeding plus transmission dynamics). We therefore fit a modifier to COVID-19 mortality considering the average health index (only) in each CA.

Movement and network patterns

The pattern of movement generated at the OA level shows substantial differences in both the average distance travelled, and the connectedness between OAs across the country. Individuals in remote areas move the farthest to work on average, and individuals in the densely populated "Central Belt" are the most connected (Figure 4).

Parameter posteriors and model fit

In the fitted model, parameter posteriors are strongly unimodal (Figure 5) with only weak pair-wise correlations. The most likely (mean) estimate for the mortality modifier is $\mu_{mod} = 0.03$, resulting in COVID-19 mortality rates that are 2.45× higher in the CA with the worst health index, compared to the best. Post-lockdown, the best fit value (lowest score) occurs when transmission rates are reduced to 0.12× the pre-lockdown value, starting from 28th March. The number of observed deaths per

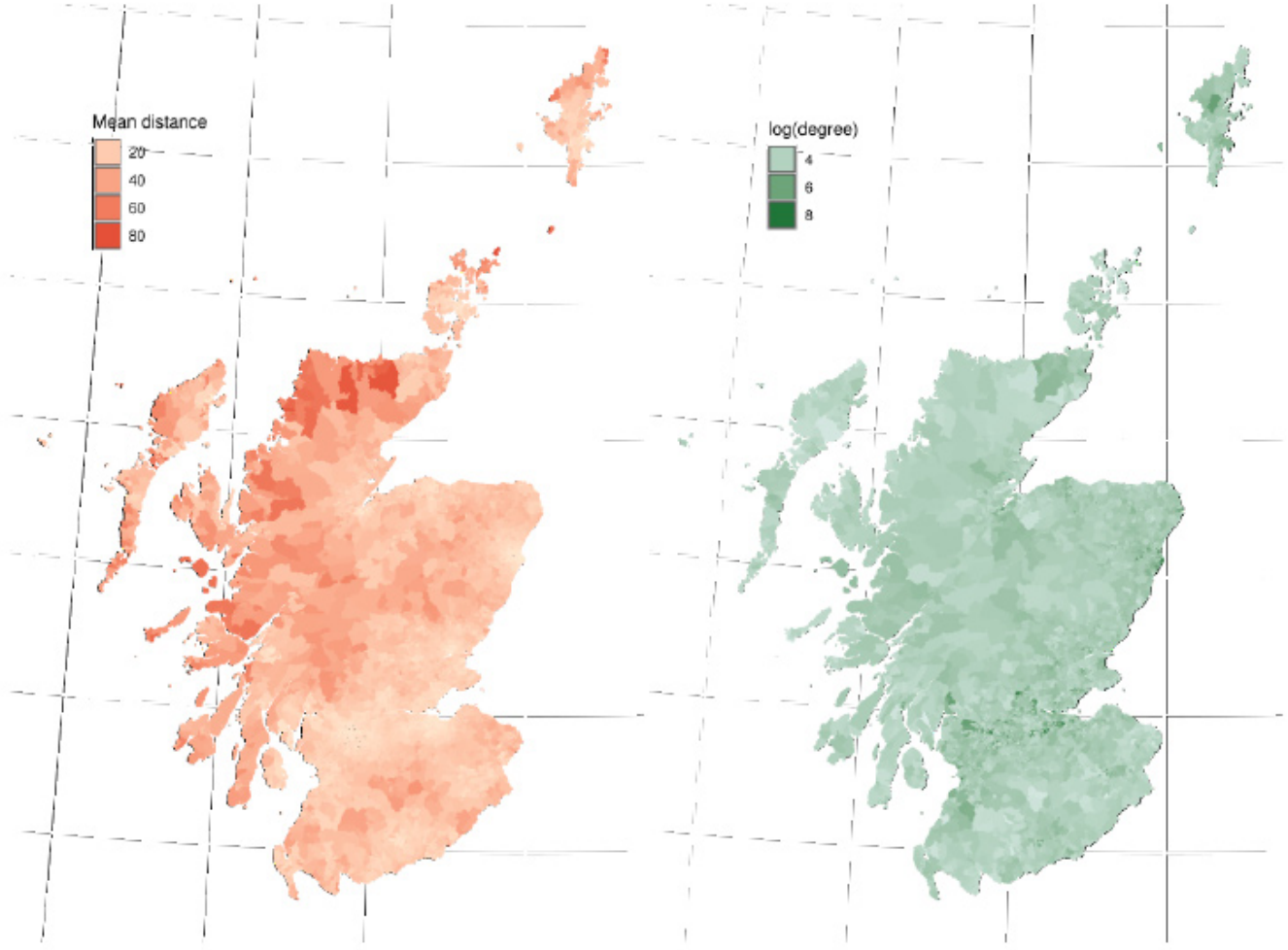

Figure 4. Commuter movement patterns, with commuters aggregated by Output Area (OA) (Census areas with typically 50–500 individuals, maximum of 2081). Data according to the 2011 census. Mean distance travelled from OA in km (left) and mean number of OA's to which each OA is connected to (right). The greatest distances are travelled on average by individuals in remote locations. The greatest network degree is found in highly urbanised areas.

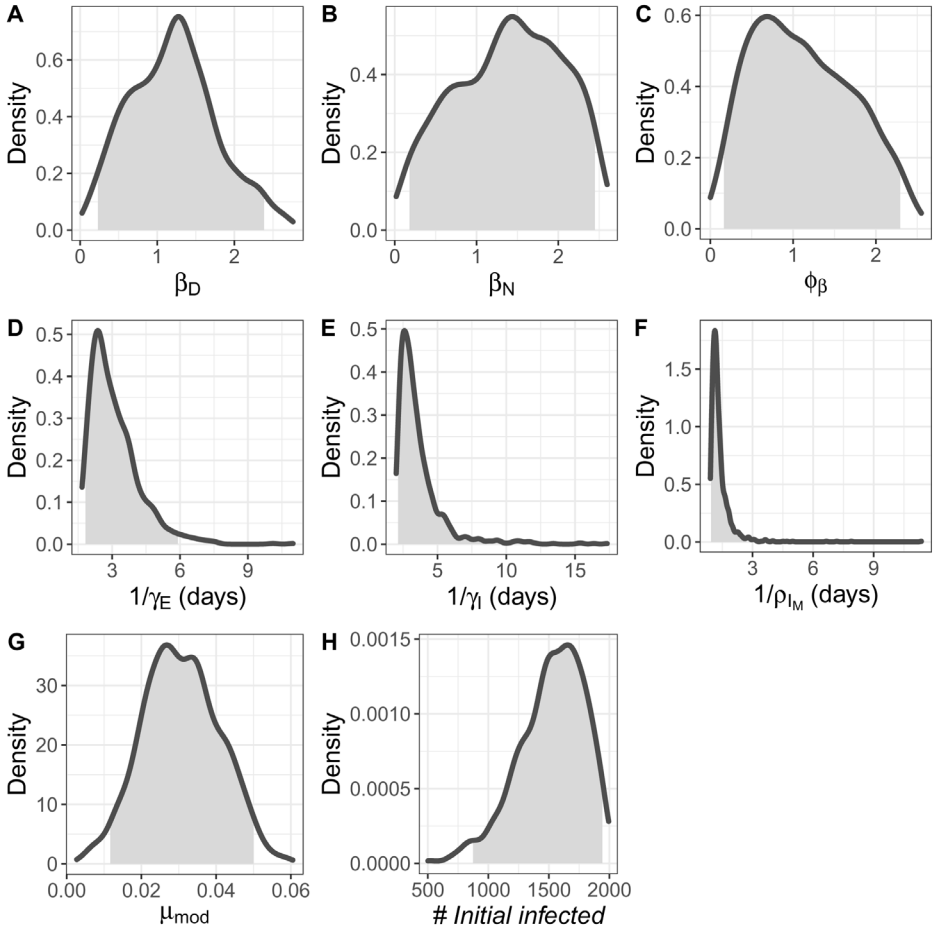

Figure 5. Posterior distributions of fitted parameters. From top left to lower right: (**A**) frequency dependent transmission rate for severely infectious individuals in daytime locations (per five severely infectious individuals, per half day), (**B**) frequency dependent transmission rate for severely infectious individuals in nighttime locations (per severely infectious, per half day), (**C**) multiplier for mildly infectious individuals. (**D**) duration of the exposed stage of infection (in days), (**E**) duration of the mildly infectious period in the absence of recovery (in days), modifier (**F**) time to recovery for mildly infectious individuals in the absence of progression to severely infected (days) (**G**) mortality rate modifier, and (**H**) number of seed infections. Shaded areas show 95% credible intervals for all parameters.

CA per week mostly lie well within the range of 95% of the fitted simulations (parameters drawn in sets from the posterior parameter distributions; see Figure 6). The number of DZs with deaths is more likely to exceed these limits (Figure 7) though with low numbers and therefore greater stochastic variability in the data.

Impact of distance reduction and transmission reduction on COVID-19 spread

Lockdown restrictions impact both the local spread of infection via physical distancing measures that change transmission rates and geographical spread (e.g. reductions in travel-to-work). In only a few cases over 25% of infection occurs outside the OA of residence (Figure 8).

These are typically urban areas, most likely with considerable inward commuting traffic (Figure 9). However, disease spread is predominantly local, with the most likely outcome being more than 90% of infection occurring within OA.

We consider counterfactuals where some lockdown restrictions are imposed on 9th March, two weeks prior to the actual date, and just before the first death due to COVID-19 in Scotland was reported. The reduction in transmission rates is attributable to physical distancing, or reduction in mobility, or both. With earlier lockdown, we predict a median 581 deaths (95% of simulations within 377 to 1,010 deaths) by 26th April 2020, compared to 2,722 (95% of simulations within 1,294 to 4,050) in the baseline scenario (observed number is 2,795, assuming that all deaths occur in the week prior to the week the death is registered in). Most of this difference is due to the reduction in transmission rates (Figure 10). The reduction in the total number of deaths also results in a reduction in geographical spread with many fewer DZs affected by COVID-19 mortality in the early lockdown scenario (Figure 11).

Discussion

Exploiting our estimates of the relative contribution of longer distance movements to work and local transmission that include

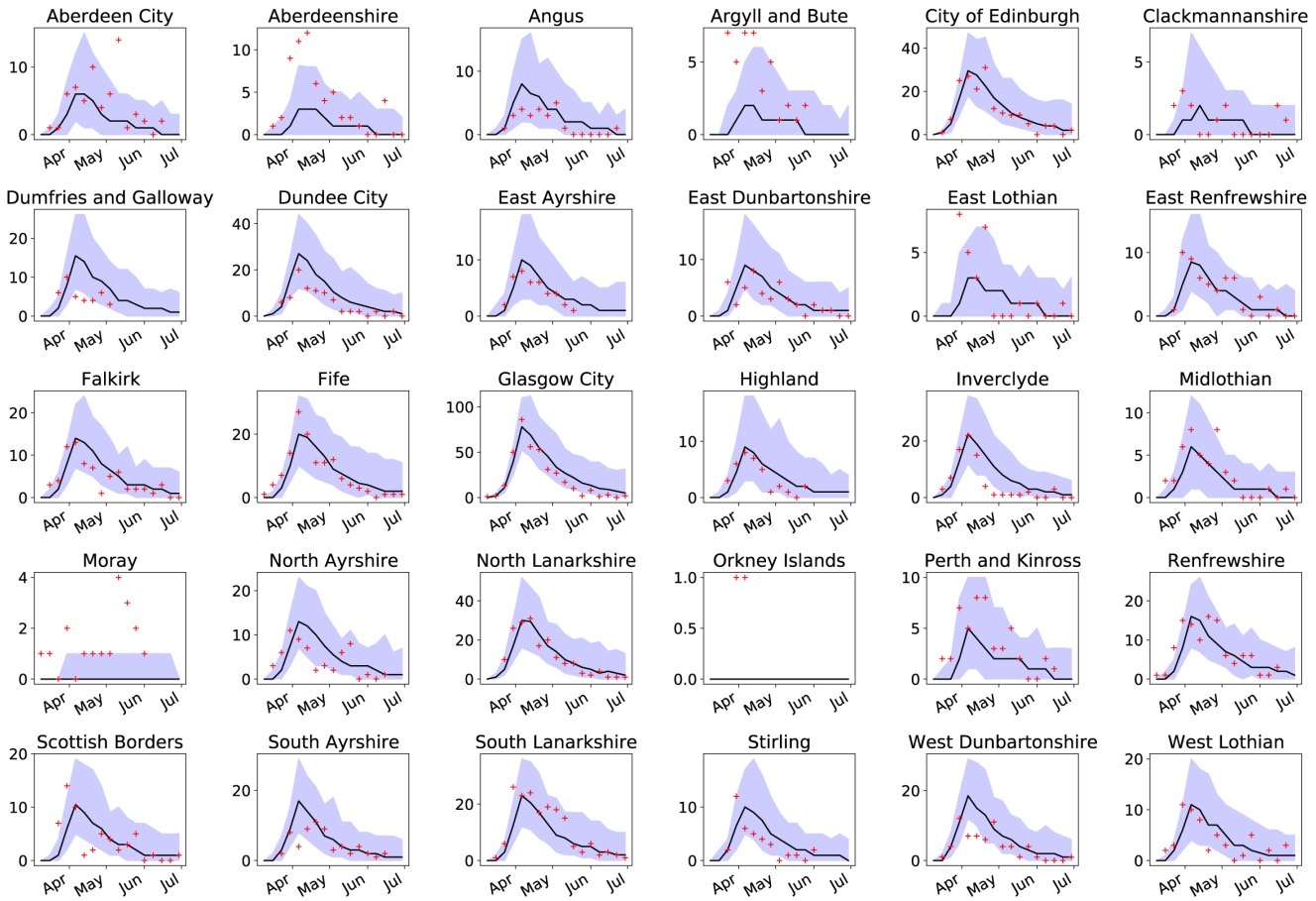

Figure 6. Number of deaths per week for all Council Areas in Scotland (bar Shetland). Incidence number (red crosses) of coronavirus disease 2019 related deaths compared to the median of 100 simulations (black line), and 95% confidence intervals (purple areas).

work and non-work interactions, we show that at this early stage physical distancing had a dominant impact on the death rate, with little evidence that longer range movement restrictions were important for reducing deaths.

Our model is fitted to the spatio-temporal pattern of deaths due to early stage COVID-19 under the assumption that it maps well onto the patterns of infection at this spatial scale and at this stage of the epidemic. While an inference done on case distributions would be a more direct approach to estimating transmission parameters, testing and hence case ascertainment in the early stages of the pandemic was poor. Our modelling approach shows that, even though the spatio-temporal pattern of deaths is subject to substantial stochastic variability, it can be used to generate meaningful opinions in our posterior parameter estimates and form the basis for prediction of counterfactuals.

Movement restrictions are of course critical for preventing the seeding of new areas and therefore a direct estimate of its role in preventing spread at these early stages of the epidemic can provide vital insights into how important rapidly applied measures are and what elements of those measures are most important. Here we show that a movement ban on its own, while it would restrict the important outcome of geographical spread would on its own have had minimal impact at this stage, on the scale of the epidemic. Instead, local physical distancing measures would have had the greatest impact.

Our analysis does not preclude increased transmission due to deprivation at the local level (e.g. due to greater use of public transport, for example) as these factors are not embedded in the model. Of course over the longer term, both elements of long distance travel restrictions and transmission reduction are needed; reductions in both the areas being affected (allowing for potentially more targeted control efforts) and total deaths (both directly saving lives, and reducing impact on the health care system). More localised or clustered infections also will ultimately result in a higher proportion of recovered individuals in hotspot areas, promoting localised herd immunity

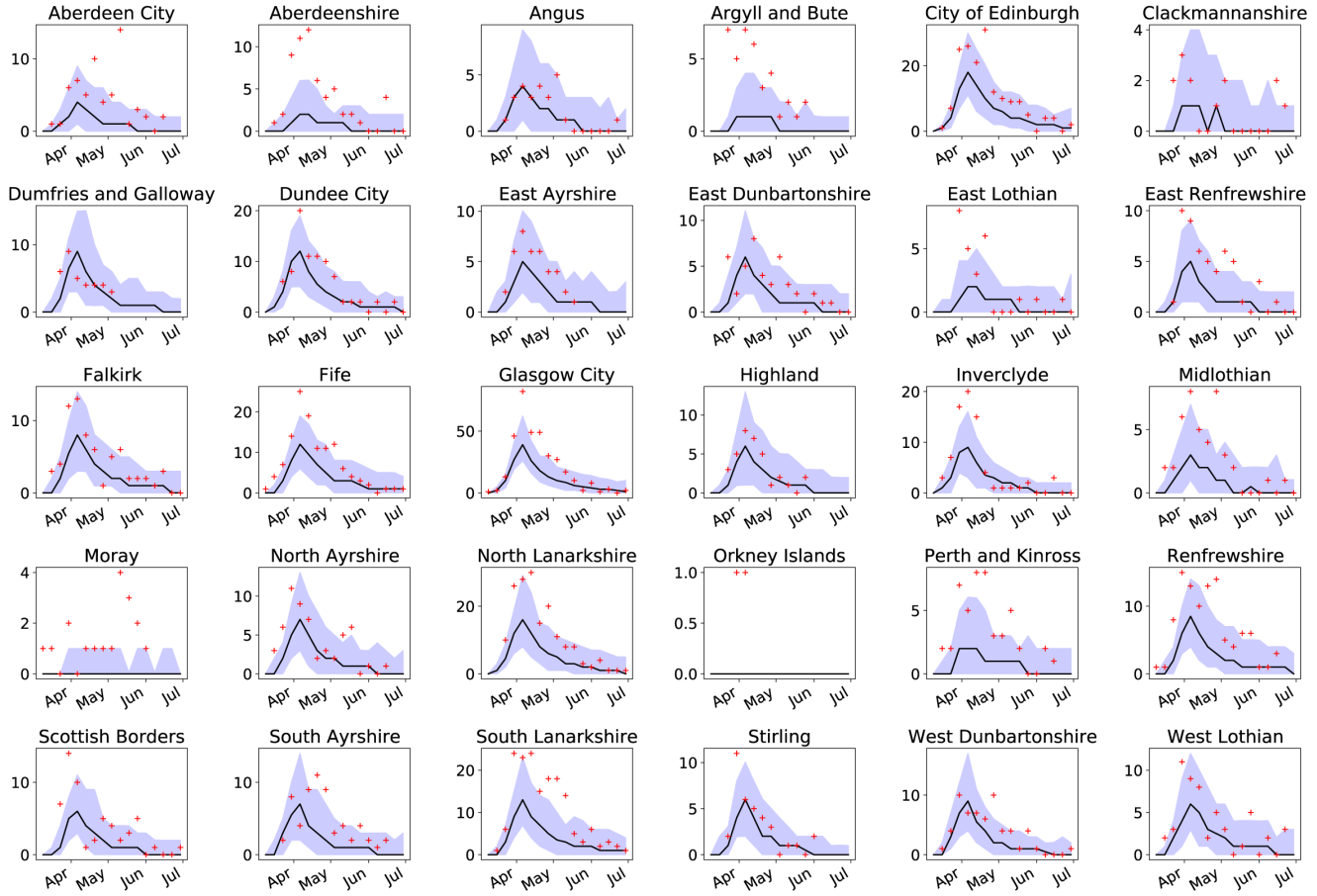

Figure 7. Number of data zones (DZ) with deaths per week for all Council Areas in Scotland (bar Shetland). Incidence number (red crosses) of DZs with coronavirus disease 2019 related deaths compared to the median of 100 simulations (black line), and 95% confidence intervals (purple areas).

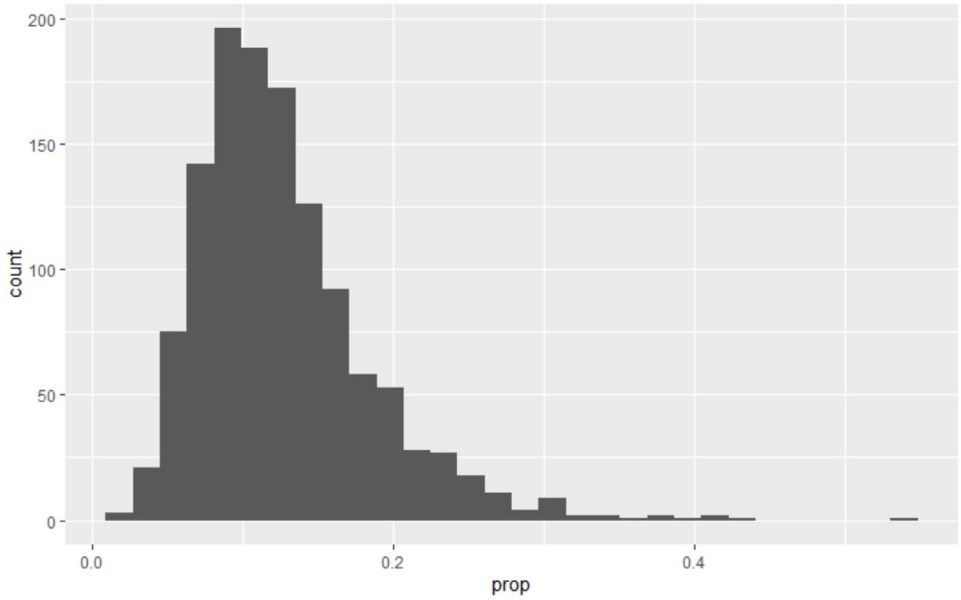

Figure 8. Proportion of transmission occurring outside of Output Area (OA) of residence, across all OAs in Scotland. For most OAs the majority of transmission is estimated to occur within the home location.

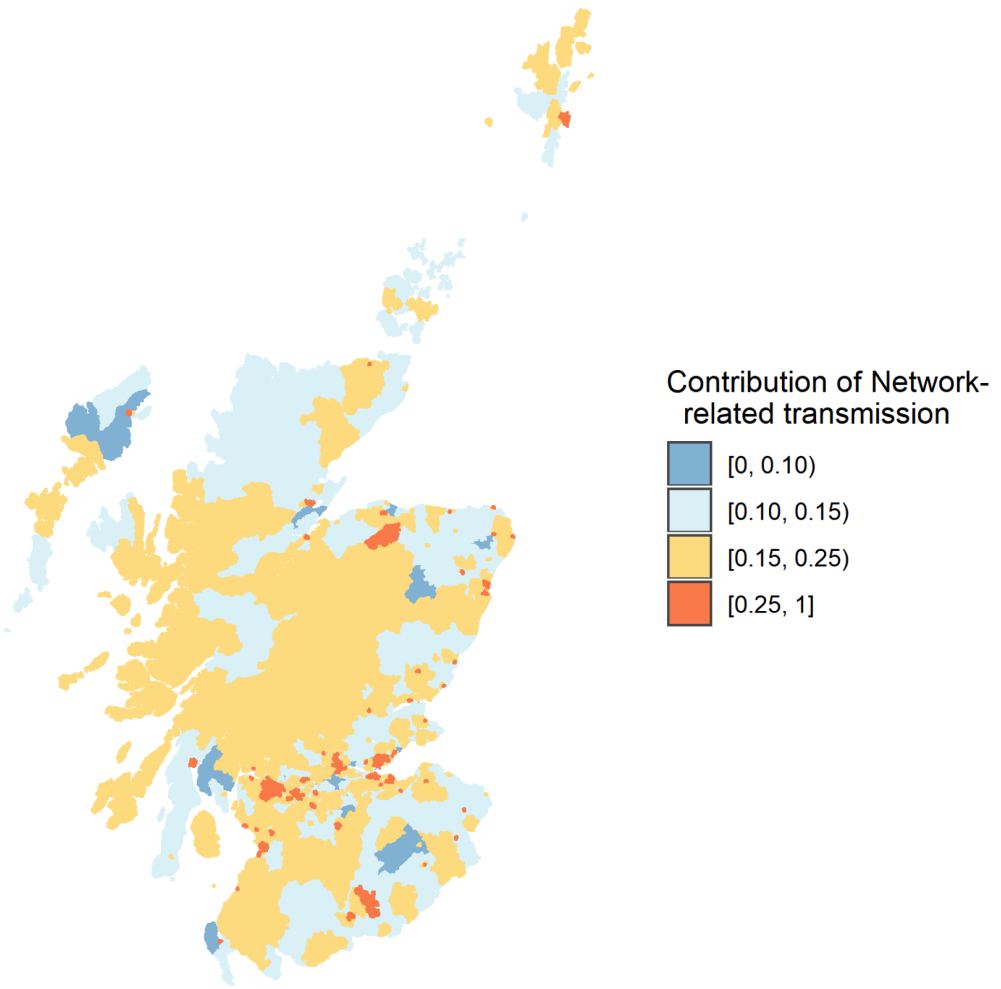

Figure 9. Estimated relative proportion of transmission events likely to occur within the Output Areas of residence of a single infected individual introduced into an otherwise susceptible population. Areas in yellow (high) and orange (extremely high) generally represent highly urban areas, with considerable inward commuting traffic.

effects. Further, as these hotspots were typically in more urban areas, this also has the important consideration of reducing infection rates in remote areas, where access to hospitals and intensive care units (ICUs) is typically poor. Of course, not all long distance movements occur for work purposes and should the patterns of non-work movements be substantially different to movements-to-work patterns, and be of sufficient volume, then it could have an influence on our results. However, the available mobility data suggest that national lockdown restrictions would have had a greater effect on at least the volume of non-work movements.

The best parameter estimates give a high probability that mildly infected individuals have higher transmission rates than the severely infected, though with shorter duration. This does not imply that they are necessarily producing more virus, as those who have severe symptoms may be self-isolating or be more strictly physically distancing. Our model also estimates the number of individuals with infection in early March, with the most likely case being on the order of 1500 infected individuals in Scotland at this time. While our approach is crude and does not take account of continuing importation of individuals over time, it is substantially greater than the estimate of 113 introductions of COVID-19 into Scotland based on viral sequence data[29], but is plausible if one takes into account the additional infections these introductions would have caused by the start of our simulations.

By explicitly modelling transmission dynamics, this allows us with relatively few data, to infer in our model that health-related deprivation results in a 2.45× difference in the death rate due to COVID-19 across CAs, rather than because these areas

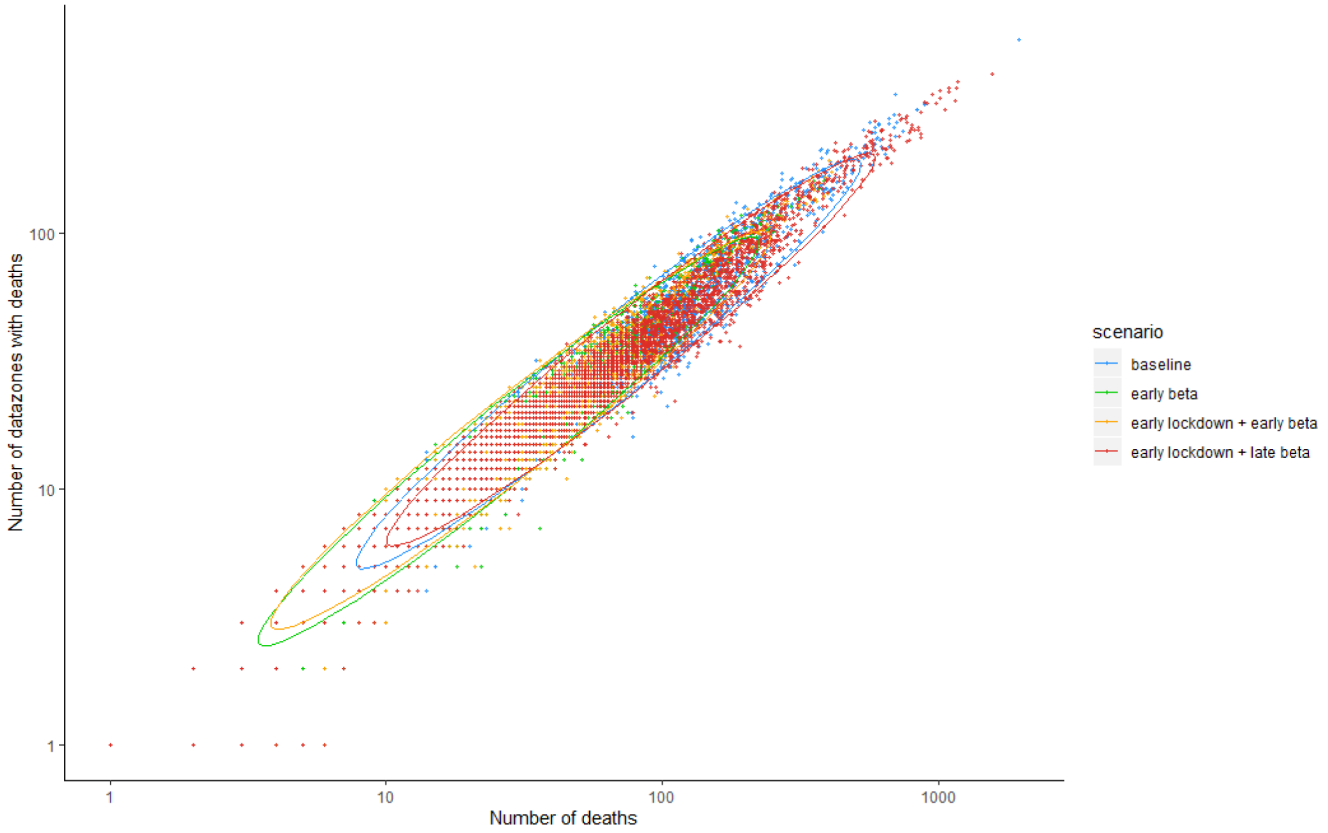

Figure 10. Comparison of number of coronavirus disease 2019 related deaths and number of data zones with deaths as of April 26th 2020, contrasting baseline (imposition of restrictions as they occurred), early (March 9th) imposition of physical distancing measures but without restriction of long distance travel (early beta), early imposition of long distance travel restriction only (early lockdown) and early imposition of both.

carried a greater burden of infections earlier in the pandemic. Deprivation (including health) in Scotland varies substantially at a much finer scale, with zones with the highest deprivation often neighbouring zones with the lowest. This suggests that, were the data available, a deeper interrogation could provide a much more refined assessment of potential health burdens and risks associated with geographical spread.

The effectiveness of lockdown will vary in space and time, due to differences in human behaviour, and also because of non-linear relationships between the numbers of cases, probability of spread, and logistical burdens on care homes, hospitals and ICUs reduced. These potentially counterbalancing factors would of course have to be considered in more detail for a full understanding of lockdown effectiveness. Despite these caveats, our simple approach is useful to strategically examine trade-offs between travel related restrictions, and physical distancing when evaluating future releases from lockdown. Here however, most of the protection from new infections is a result of transmission reduction; long distance restrictions only have a minimal impact because in our model, as overall infection pressure is conserved—i.e. if it does not occur in work locations, it is assumed to occur at home. Thus only extreme travel restrictions are likely to have an impact, at least until build-up of immunity levels is more substantial than observed at this early stage.

Since this analysis, the successful deployment of COVID vaccines in the UK and elsewhere has meant that the response to the current pandemic need not include such extreme measures. However, where vaccine escape mutants become a problem of sufficient concern, and booster vaccines not be readily available, the need for future extreme controls cannot be ruled out, with the recent spread of the Omicron variant one example, although the most extreme measures have thus far proven unnecessary. In situations where such measures are considered, our results indicate that even extreme lockdown measures do not entirely prevent geographical spread between regions; thus any restrictions must include measures to not just reduce mobility but also reduce transmission if, in the future, we are to prevent the spread of COVID-19 to areas that have successfully eradicated local COVID-19 cases.

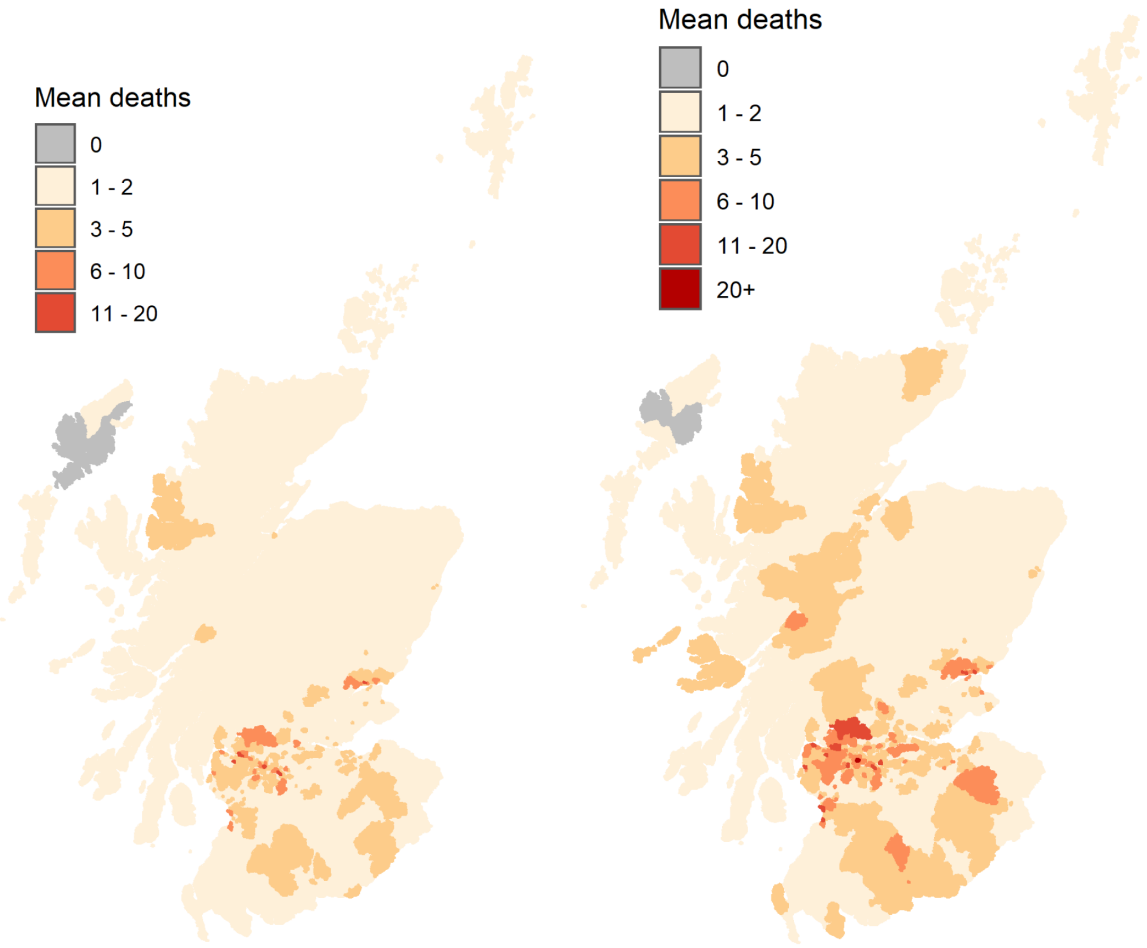

Figure 11. Mean number of deaths per Output Area, averaged over 50 simulations for early lockdown on 16th March 2020 (left) and lockdown as it occurred on 23rd March 2020 (right) considering the total number to 30th April, 2020.

Data availability

Underlying data

Data on underlying population and movements: NRS Census Flows data, Table WU01UK Available from: https://wicid.ukdataservice.ac.uk Access to Output Area level data requires an academic registration.

Data on human mobility over the simulation period: Google Mobility Reports Available from: https://www.google.com/covid19/mobility/

Data on deprivation-related demographics of underlying population: Scottish Index of Multiple Deprivation (SIMD) 2020 Available from: https://www.gov.scot/collections/scottish-index-of-multiple-deprivation-2020

Data on COVID-19 testing and mortality (at Datazone level) were provided by Public Health Scotland (PHS) and made available via the PHS Electronic Data Research and Innovation Service (eDRIS). Available from: https://www.isdscotland.org/Products-and-services/Edris/ The sensitive nature of these data required a Data Sharing Agreement between PHS and University of Edinburgh, and therefore are not publicly available.

Extended data

Model code available from: https://github.com/Kao-Group/SCoVMod.git

Archived model code at time of publication: https://doi.org/10.5281/zenodo.6420991[30]

License: MIT

Acknowledgements

We thank the electronic Data Research and Innovation Service within Public Health Scotland for provision of data. Samantha J. Lycett (Roslin Institute, University of Edinburgh), for helpful advice on the model development. Prof. C. Robertson (Strathclyde University, PHS) for advice on scenarios.

References

1. Patel JA, Nielsen FBH, Badiani AA, *et al.*: **Poverty, inequality and COVID-19: the forgotten vulnerable.** *Public Health.* 2020; **183**: 110–111.
PubMed Abstract | Publisher Full Text | Free Full Text
2. Liu Y, Morgenstern C, Kelly J, *et al.*: **The impact of non-pharmaceutical interventions on sars-cov-2 transmission across 130 countries and territories.** *BMC Med.* 2021; **19**(1): 40.
PubMed Abstract | Publisher Full Text | Free Full Text
3. Haug N, Geyrhofer L, Londei A, *et al.*: **Ranking the effectiveness of worldwide COVID-19 government interventions.** *Nat Hum Behav.* 2020; **4**(12): 1303–1312.
PubMed Abstract | Publisher Full Text
4. Flaxman S, Mishra S, Gandy A, *et al.*: **Estimating the effects of non-pharmaceutical interventions on covid-19 in europe.** *Nature.* 2020; **584**(7820): 257–261.
PubMed Abstract | Publisher Full Text
5. Arenas A, Cota W, Gómez-Gardeñes J, *et al.*: **A mathematical model for the spatiotemporal epidemic spreading of COVID19.** *MedRxiv.* 2020.
Publisher Full Text
6. Di Domenico L, Pullano G, Sabbatini CE, *et al.*: **Impact of lockdown on COVID-19 epidemic in Île-de-France and possible exit strategies.** *BMC Med.* 2020; **18**(1): 240.
PubMed Abstract | Publisher Full Text | Free Full Text
7. Williamson EJ, Walker AJ, Bhaskaran K, *et al.*: **Factors associated with COVID-19-related death using OpenSAFELY.** *Nature.* 2020; **584**(7821): 430–436.
PubMed Abstract | Publisher Full Text | Free Full Text
8. Treibel TA, Manisty C, Burton M, *et al.*: **COVID-19: PCR screening of asymptomatic health-care workers at London hospital.** *Lancet.* 2020; **395**(10237): 1608–1610.
PubMed Abstract | Publisher Full Text | Free Full Text
9. He X, Lau EH, Wu P, *et al.*: **Temporal dynamics in viral shedding and transmissibility of COVID-19.** *Nat Med.* 2020; **26**(5): 672–675.
PubMed Abstract | Publisher Full Text
10. Li R, Pei S, Chen B, *et al.*: **Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2).** *Science.* 2020; **368**(6490): 489–493.
PubMed Abstract | Publisher Full Text | Free Full Text
11. Bi Q, Wu Y, Mei S, *et al.*: **Epidemiology and transmission of COVID-19 in 391 cases and 1286 of their close contacts in Shenzhen, China: a retrospective cohort study.** *Lancet Infect Dis.* 2020; **20**(8): 911–919.
PubMed Abstract | Publisher Full Text | Free Full Text
12. Zhang J, Litvinova M, Wang W, *et al.*: **Evolving epidemiology and transmission dynamics of coronavirus disease 2019 outside Hubei province, China: a descriptive and modelling study.** *Lancet Infect Dis.* 2020; **20**(7): 793–802.
PubMed Abstract | Publisher Full Text | Free Full Text
13. Linton NM, Kobayashi T, Yang Y, *et al.*: **Incubation Period and Other Epidemiological Characteristics of 2019 Novel Coronavirus Infections with Right Truncation: A Statistical Analysis of Publicly Available Case Data.** *J Clin Med.* 2020; **9**(2): 538.
PubMed Abstract | Publisher Full Text | Free Full Text
14. Lauer SA, Grantz KH, Bi Q, *et al.*: **The incubation period of coronavirus disease 2019 (CoVID-19) from publicly reported confirmed cases: Estimation and application.** *Ann Intern Med.* 2020; **172**(9): 577–582.
PubMed Abstract | Publisher Full Text | Free Full Text
15. Sanche S, Lin YT, Xu C, *et al.*: **The novel coronavirus, 2019-nCoV, is highly contagious and more infectious than initially estimated.** arXiv, 2020.
Publisher Full Text
16. Sanche S, Lin YT, Xu C, *et al.*: **High Contagiousness and Rapid Spread of Severe Acute Respiratory Syndrome Coronavirus 2.** *Emerg Infect Dis.* 2020; **26**(7): 1470–1477.
PubMed Abstract | Publisher Full Text | Free Full Text
17. Wang D, Hu B, Hu C, *et al.*: **Clinical Characteristics of 138 Hospitalized Patients with 2019 Novel Coronavirus-Infected Pneumonia in Wuhan, China.** *JAMA.* 2020; **323**(11): 1061–1069.
PubMed Abstract | Publisher Full Text | Free Full Text
18. Liu T, Hu J, Xiao J, *et al.*: **Time-varying transmission dynamics of Novel Coronavirus Pneumonia in China.** *bioRxiv.* 2020.
Publisher Full Text
19. Han YN, Feng ZW, Sun LN, *et al.*: **A comparative-descriptive analysis of clinical characteristics in 2019-coronavirus-infected children and adults.** *J Med Virol.* 2020; **92**(9): 1596–1602.
PubMed Abstract | Publisher Full Text
20. Pung R, Chiew CJ, Young BE, *et al.*: **Investigation of three clusters of COVID-19 in Singapore: implications for surveillance and response measures.** *Lancet.* 2020; **395**(10229): 1039–1046.
PubMed Abstract | Publisher Full Text | Free Full Text
21. Verity R, Okell LC, Dorigatti I, *et al.*: **Estimates of the severity of coronavirus disease 2019: a model-based analysis.** *Lancet Infect Dis.* 2020; **20**(6): 669–677.
PubMed Abstract | Publisher Full Text | Free Full Text
22. da Silva Filipe A, Shepherd J, Williams T, *et al.*: **Genomic epidemiology of SARS-CoV-2 spread in Scotland highlights the role of European travel in COVID-19 emergence.** *medRxiv.* 2020.
Publisher Full Text
23. Qin J, You C, Lin Q, *et al.*: **Estimation of incubation period distribution of COVID-19 using disease onset forward time: A novel cross-sectional and forward follow-up study.** *Sci Adv.* 2020; **6**(33): eabc1202.
PubMed Abstract | Publisher Full Text | Free Full Text
24. Arentz M, Yim E, Klaff L, *et al.*: **Characteristics and Outcomes of 21 Critically Ill Patients with COVID-19 in Washington State.** *JAMA.* 2020; **323**(16): 1612–1614.
PubMed Abstract | Publisher Full Text | Free Full Text
25. Jarvis CI, Van Zandvoort K, Gimma A, *et al.*: **Quantifying the impact of physical distance measures on the transmission of COVID-19 in the UK.** *BMC Med.* 2020; **18**(1): 124.
PubMed Abstract | Publisher Full Text | Free Full Text
26. Hartig F, Calabrese JM, Reineking B, *et al.*: **Statistical inference for stochastic simulation models--theory and application.** *Ecol Lett.* 2011; **14**(8): 816–827.
PubMed Abstract | Publisher Full Text
27. Toni T, Welch D, Strelkowa N, *et al.*: **Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems.** *J R Soc Interface.* 2009; **6**(31): 187–202.
PubMed Abstract | Publisher Full Text | Free Full Text
28. Gillespie DT: **Approximate accelerated stochastic simulation of chemically reacting systems.** *J Chem Phys.* 2001; **115**(4): 1716–1733.
Publisher Full Text
29. Lycett SJ, Hughes J, Mchugh MP, *et al.*: **Epidemic waves of COVID-19 in Scotland: a genomic perspective on the impact of the introduction and relaxation of lockdown on SARS-CoV-2.** *medRxiv.* 2021; 2021.01.08.20248677.
Publisher Full Text
30. Banks C: **Kao-Group/SCoVMod: SCoVMod Paper Release (v1.0).** *Zenodo.* 2022.
http://www.doi.org/10.5281/zenodo.6420992

Open Peer Review

Current Peer Review Status: ✔ ✔

Version 1

Reviewer Report 11 July 2022

https://doi.org/10.21956/wellcomeopenres.19604.r50789

✔ **Cameron Zachreson**
School of Computing and Information Systems, The University of Melbourne, Melbourne, Vic, Australia

This paper presents a spatially explicit model of COVID-19 in Scotland. The population model incorporates medium to long-range mobility patterns as well as local contact rates and the study investigates the effectiveness of lockdown measures with respect to these two model components. The study makes the case that, in Scotland, mobility restrictions were relatively ineffective compared to the influence of local transmission mitigation (i.e., reduction in contact rates within regions). The results indicate that this is because mobility restrictions were implemented too late, and were not stringent enough to prevent the spread of SARS-CoV-2 between areas. The conceptual case is clear and the argument is reasonable, but there are some strong assumptions that could be tested to make the conclusions more robust. There are also some minor clarity issues with the model description, and the presentation of data could be more thorough/professional. Overall I found the paper well-written, the arguments clearly articulated, and the conclusions relatively convincing. Furthermore, the code provided appears to have been developed according to a high professional standard (though it could benefit from additional explanatory documentation and inline comments).

Main issues: strong assumption about initial conditions; posterior correlations not presented.

Suggest analysis of sensitivity to strong assumptions:

- (i) distribution of initial cases in proportion to observed cumulative deaths (this one is the most important):
  Models like this are very sensitive to initial conditions (location and number of seed cases). Since fitting is performed based on the spatial distribution of deaths, the assumption that the initial conditions are themselves strongly dependent on the observed outcomes could be an artificially strong constraint. The true initial conditions were not observed with high levels of accuracy, so some well-reasoned relaxation of the strong constraint used in the design of initial conditions should be explored. For example, a null model could be compared in which the seed cases are distributed with respect to population only (which

could be age-stratified etc.).

- (ii) Distribution of mobility reduction uniformly in proportion to global mobility reduction: The model defines mobility patterns based on a statistical snapshot. The mobility dynamics under lockdown are treated as uniform adjustments to the snapshot of mobility patterns in the base model. This is a reasonable treatment given the limitations of the data, but represents a strong assumption against which the robustness of the results should be tested. For example, a perturbation approach could be used to systematically randomise the pairwise reductions in mobility, or an alternate assumption could be tested, in which relative mobility reductions depend on distance between origin and destination or some other set of region pair properties.

I suggest additional results:

- (i) Correlations in joint posterior (for parameters in Figure 5) would help informed readers understand model uncertainty more clearly.

- (ii) Full posterior and correlations for estimated efficacy of physical distancing with epi parameters would help quantify uncertainty in the conclusions. Currently, only the best-fit value is reported (0.12).

Minor issues:

- I found the model software well organised, but it could benefit from more informative documentation and commenting, as the design is relatively complex.

- Figures need a thorough proof-read and should be presented in high-resolution format (preferably vector graphics)

- The results in Figure 10 could be presented more clearly (a 1-d histogram for each observable may be preferable given the very high correlation).

- Equations should be numbered.

- Model could be more clearly described as an agent-based model, and specified according to agent state variables (Note: this is especially confusing in the discussion of how individuals are assigned work locations and how the mobility flows are computed - it needs to be clear that each working adult is assigned only one location, I had to read the section four times to understand that origins were not stochastically assigned each day).

Is the work clearly and accurately presented and does it cite the current literature?

Yes

Is the study design appropriate and is the work technically sound?

Yes

Are sufficient details of methods and analysis provided to allow replication by others?

Yes

If applicable, is the statistical analysis and its interpretation appropriate?

Yes

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Computational modelling of disease transmission, with a focus on agent-based models of pandemics in structured populations.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Reviewer Report 07 July 2022

https://doi.org/10.21956/wellcomeopenres.19604.r51159

Christopher Davis
Mathematics Institute, University of Warwick, Coventry, UK

The authors present a spatial agent-based transmission model for COVID-19 in Scotland, which is applied to the early months of the pandemic in 2020. The model parameters are obtained by a combination of using values from previous literature, values taken directly from data, and estimation using ABC-SMC. Mortality in the model is adjusted spatially using data from the Scottish Index of Multiple Deprivation.

The model is used to estimate the impact of lockdown restrictions, the difference in mortality by deprivation status, and, given the model is spatially structured, assess the separate impact of long-distance movement restrictions and local transmission reduction.

Overall, the manuscript is well written and explicit in the assumptions made to arrive at the presented conclusions. The message is clear and the methods are appropriate to tackle the questions asked. The model is shown to have a good fit to the death data. All data used is listed and the model code is available to access.

Some minor comments are listed below:

- In the first paragraph “lockdown” and “key workers” are italicised, without being defined anywhere else.

- Table 1 shows the fitted parameters, but the parameter symbols do not completely match

those used for the posteriors in Figure 5 (e.g. the transmission rate multiplier for mild infectors is given by both “y” and “phi_beta” and differences in capitalisation for day and night).

- For the equation for the number of deaths, should it also include the number of susceptible people?
- The definition of the degree in Figure 4 is not completely clear.
- In Figure 10, the y axis label should have “data zones” as two words. The caption should be split into two sentences, and the number of simulations done also stated here. The methodology for drawing the rings is not obviously explained. In addition, most of the points look red (all the red points were plotted last?), and so the figure is difficult to read. Could some jittering be applied or the order in which the points are plotted randomised to provide a more representative picture?
- In Figure 11, the first of the two figure legends should be removed.
- Also, Figure 3 and Figure 4 are currently presented as low resolution images.

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Mathematical epidemiology

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.