Epidemics 48 (2024) 100781

Contents lists available at ScienceDirect

Epidemics

journal homepage: www.elsevier.com/locate/epidemics

COVID-19 in Scottish care homes: A metapopulation model of spread among residents and staff

Matthew Baister [a,1,*], Ewan McTaggart [a,1], Paul McMenemy [a,b], Itamar Megiddo [c], Adam Kleczkowski [a]

[a] Department of Mathematics & Statistics, University of Strathclyde, Glasgow, UK
[b] Department of Computing Science and Mathematics, University of Stirling, Stirling, UK
[c] Department of Management Science, University of Strathclyde, Glasgow, UK

ARTICLE INFO

Dataset link: https://doi.org/10.1016/S2666-7568(20)30012-X, https://github.com/ewanmct/COVID-care-homes

Keywords:
COVID-19
Epidemic modelling
Meta-population
Care home

ABSTRACT

The movement of populations between locations and activities can result in complex transmission dynamics, posing significant challenges in controlling infectious diseases like COVID-19. Notably, networks of care homes create an ecosystem where staff and visitor movement acts as a vector for disease transmission, contributing to the heightened risk for their vulnerable communities. Care homes in the UK were disproportionately affected by the first wave of the COVID-19 pandemic, accounting for almost half of COVID-19 deaths during the period of 6th March – 15th June 2020 and so there is a pressing need to explore modelling approaches suitable for such systems. We develop a generic compartmental Susceptible - Exposed - Infectious - Recovered - Dead (**SEIRD**) metapopulation model, with care home residents, care home workers, and the general population modelled as subpopulations, interacting on a network describing their mixing habits. We illustrate the model application by analysing the spread of COVID-19 over the first wave of the COVID-19 pandemic in the NHS Lothian health board, Scotland. We explicitly model the outbreak's reproduction rate and care home visitation level over time for each subpopulation and execute a data fit and sensitivity analysis, focusing on parameters responsible for inter-subpopulation mixing: staff-sharing, staff shift patterns and visitation. The results from our sensitivity analysis show that restricting staff sharing between homes and staff interaction with the general public would significantly mitigate the disease burden. Our findings indicate that protecting care home staff from disease, coupled with reductions in staff-sharing across care homes and expedient cancellations of visitations, can significantly reduce the size of outbreaks in care home settings.

1. Introduction

The outbreak of the SARS-CoV-2 induced disease (COVID-19) pandemic has had a profound impact, causing 3.7 million deaths by early June 2021 and global economic shocks (World Health Organisation, 2021). In the UK, the care home population suffered a disproportionate amount of COVID-19 related deaths. From the week ending 13th March 2020 to the week ending 26th June 2020 (the "first wave"), 54,510 deaths were associated with COVID-19 in the UK, 40% of which were among care home residents (Bell et al., 2020). The COVID-19 pandemic has highlighted the vulnerability of care homes to epidemics, as their resident population is elderly and often suffers from several co-morbidities (Social Care Working Group (SCWG) for the Scientific Advisory Group for Emergencies, 2021), their systems have not been developed with infection prevention and control (IPC) in mind, and their IPC guidelines have been borrowed from hospitals - a completely different setting (Nguyen et al., 2020b).

Networks of care homes are ecosystems connected by staff working across facilities, and these connections can increase the risk of COVID-19 ingress into care homes, and to protect their vulnerable community, we need to understand the ecosystem dynamics. We find it natural to describe this using a heterogeneous patch size metapopulation model framework.

Very few models explore COVID-19 transmission at a community level and explicitly include the unique dynamics in care homes. For example, in Smith et al. (2020) and Nguyen et al. (2020a) agent-based models (ABMs) of single homes are used to investigate the impact of testing strategies on the disease burden. A report by Nguyen et al. (2021b) uses an ABM to investigate the impact on care home residents of various vaccine coverage, and reducing the weekly testing

* Corresponding author.
E-mail address: matthew.baister@apha.gov.uk (M. Baister).
[1] Contributed equally to this work.

https://doi.org/10.1016/j.epidem.2024.100781
Received 2 August 2023; Received in revised form 28 March 2024; Accepted 18 June 2024
Available online 5 July 2024

of staff. However, the models in Nguyen et al. (2020a) and Nguyen et al. (2021b) do have an external force of infection (FOI) from the community, based on prevalence data, representing staff interaction with the community and visitors. These models (Smith et al., 2020; Nguyen et al., 2020a, 2021b) do not assess the relative impact of the different COVID-19 pathways into care homes. Nguyen et al. (2021a) extend (Nguyen et al., 2020a, 2021b), using a hybrid ABM-System Dynamics model, to explore the conditions under which visitation, heterogeneous care homes sizes, and the cohorting of residents impacts COVID-19 outbreak severity.

Roselló et al. (2021) model an individual care home with a stochastic compartmental model, using multiple forces of infection to capture COVID-19 pathways, including visitors, hospital discharges, staff working at other homes, and staff infections from the community. They find that importations of infections by staff from the community are the main driver of outbreaks, and importation by visitors or from hospitals is rare, but do not explicitly model disease spread throughout a network of care homes. In Overton et al. (2020) individual care homes and the general public are independent, deterministic SEIR models, with a stochastic external FOI connecting the general public to each home. This FOI depends on the prevalence of COVID-19 in the general public, and the size of each home. Transmission rates in homes and in the general public do not vary over time. In Oberhammer (2020), two weakly-coupled SEIR sub-models with time-dependent transmission rates define the dynamics; one sub-model describes the general public and one describes all care home residents in Stockholm as a single homogeneous group. Again, a single FOI acts on the residents to capture infections from staff and visitors. These models (Overton et al., 2020; Oberhammer, 2020) do not differentiate between, and therefore allow comparison of, the COVID-19 pathways into care homes. Bunnik et al. (2021) use a compartmental metapopulation model to explore the trade-offs between increasing protection for a "vulnerable" population and relaxing restrictions for the "non-vulnerable" after the first lockdown in Scotland. They use time-dependent transmission rates with three metapopulation groups; vulnerable, shielders and general public. We extend and apply the methodology of Bunnik et al. (2021) in our model, investigating protection to a vulnerable group (care home residents) in ways other than shielding.

In this paper, we develop a generic **SEIRD** compartmental metapopulation model to capture the transmission dynamics which arise due to population movement between locations and activities. We illustrate the model application by analysing the first wave of COVID-19 in a regional health board in Scotland. The population is divided into groups of care home residents, staff, and the general public, who interact on a network describing their mixing habits. Our care home resident group are not a single homogeneous unit as in Oberhammer (2020) and Bunnik et al. (2021), they are separate units (homes), creating a refined spatial/geographic structure. These homes are not independent as in Overton et al. (2020) but are linked by a visitation and staff-sharing network, which, to our knowledge, is unique. We calibrate this model to 2020 data from the NHS Lothian Health Board and explore the sensitivity of the results to changes in key parameters. We assess the impact of inter-subpopulation mixing patterns on the spread of COVID-19 into and throughout the susceptible care home community to identify potential mitigation strategies to minimise the impact of future outbreaks. We investigate patterns controlling the importation of infections by staff from the community, visitation and staff-sharing.

2. Materials and methods

2.1. Mathematical model

Care homes and their residents are enclosed societies, isolated to some extent from the general population. Their connection to broader society primarily consists of interaction with staff and visits from the general population. Care home staff could play an important role in COVID-19 introduction and spread throughout the care home population. Firstly, staff exposure to infection from the general population can establish an outbreak in a home. Secondly, some staff work across multiple homes - a concept we refer to as staff-sharing. Staff acting as a bridge between care homes and the general population and staff-sharing induces a network, connecting care homes in a given community via their workers. This creates the potential for COVID-19 to spread from one home to another; hence, investigation of this pathway is important.

We develop a deterministic **SEIRD** compartmental metapopulation model with heterogeneous subpopulation sizes. Each subpopulation consists of a host human population, categorised further into five compartments of COVID-19 infection status: Susceptible (**S**), i.e., everyone who is not infected; Exposed (**E**), those exposed to the virus (and infected) but not yet infectious; Infectious (**I**), infected individuals who can infect others; Recovered (**R**), those who had COVID-19 and recovered; and Dead (**D**), those who died from their illness. We do not distinguish between symptomatic and asymptomatic individuals. This model is illustrated in Fig. 1(a).

The metapopulation structure represents the population of the NHS Lothian health board in Scotland, but could be applied to other appropriate scenarios. We distinguish between care home residents, care home workers and the general population, modelling the $m = 109$ care homes for older adults in NHS Lothian (Burton et al., 2020). The jth home has a resident subpopulation, C_j, with a corresponding care home worker subpopulation, W_j. The general population is encapsulated by the subpopulation G (Fig. 1b). Each care home includes the same number of residents, a simplifying assumption made due to lack of publicly available data on care home sizes in Lothian. We also assume the worker and resident subpopulations are the same size (Knock et al., 2021).

Each node of the network, $i \in X := \{C_1, C_2, \dots, C_m, W_1, W_2, \dots, W_m, G\}$ is described in terms of the **SEIRD** compartmental model with equations:

$$\begin{aligned}
\frac{d\mathbf{S}_i}{dt} &= -\mathbf{S}_i \Phi_i \\
\frac{d\mathbf{E}_i}{dt} &= \mathbf{S}_i \Phi_i - \frac{\mathbf{E}_i}{\alpha} \\
\frac{d\mathbf{I}_i}{dt} &= \frac{\mathbf{E}_i}{\alpha} - \frac{\mathbf{I}_i}{\tau} \\
\frac{d\mathbf{R}_i}{dt} &= (1-\mu_i)\frac{\mathbf{I}_i}{\tau} \\
\frac{d\mathbf{D}_i}{dt} &= \mu_i \frac{\mathbf{I}_i}{\tau}
\end{aligned} \tag{1}$$

Susceptibles in subpopulation i ($\mathbf{S}_i$), are introduced to a force of infection Φ_i, and moved to the exposed class ($\mathbf{E}_i$). After a non-infectious latent period of α days, they become infectious, and move to the class $\mathbf{I}_i$. After τ infectious days, a proportion μ_i of the infected population ($\mathbf{I}_i$) die and the rest $(1-\mu_i)$ recover. For simplicity, and considering the model describes a short period of approximately 3 months, non-COVID-related deaths are not considered. For similar reasons, we do not include a birth rate or admission of new residents to care homes from the general population.

The parameters τ and α describe the infectious period and latency period, respectively, and are assumed to be the same across all subpopulations. Mortality rates, μ_i, vary by subpopulation, reflecting the positive association of serious outcomes of COVID-19 with age (McKeigue and Colhoun, 2020). We assume two constant and unique death rates in our model: one for residents (μ_C) and one for the general population (μ_G). We assume care home staff have the same death rate as the general population. As we are modelling over a period of 4 months (approx. first wave), and immunity after COVID-19 infection lasts as long as 5 months (Ripperger et al., 2020; Dan et al., 2021), we do not consider a transition from Recovered to Susceptible.

A number of parameters in our model are time-dependent; we assume they are higher in the pandemic's initial phase than the latter.

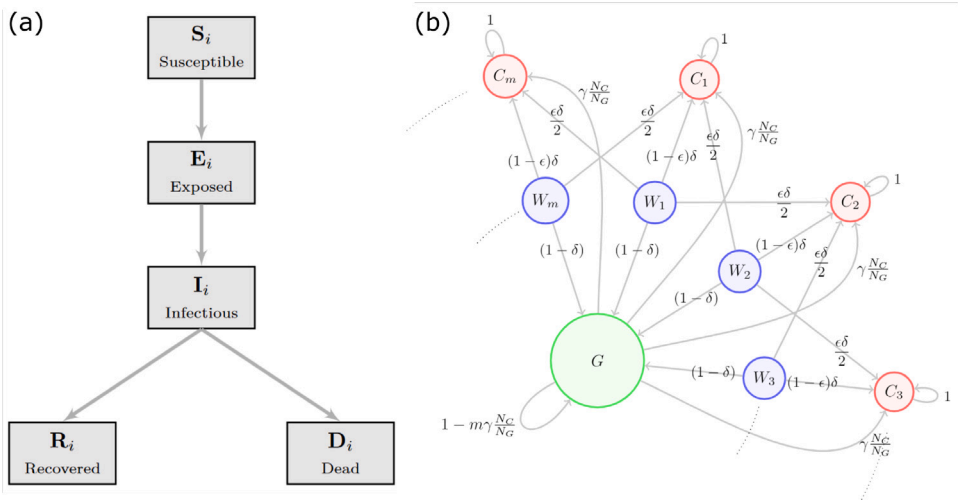

Fig. 1. Schematics for the compartmental and metapopulation structure. (a): **SEIARD** compartmental structure of the model; (b): Time-share network of interaction amongst subpopulations. Directed edge weights are p_{ik}, the proportion of people from subpopulation i who travel to mix at effective population k.

We use the function $f\left(t,\Omega_x\right)$, taking the shape of a sigmoidal logistic curve, to describe their behaviour. The function uses the set of input parameters $\Omega_x = \{\omega^x_{end}, \omega^x_{rate}, \omega^x_{low}, \omega^x_{high}\}$ which are all individually ≥0, and is defined below:

$$f\left(t,\Omega_x\right) = \frac{\left(\omega^x_{high} - \omega^x_{low}\right)}{\left(1+\exp\left(\omega^x_{rate}\left(t-\omega^x_{end}\right)\right)\right)} + \omega^x_{low}, \tag{2}$$

The function $f\left(t,\Omega_x\right)$ is decreasing with $t \geq 0$, and under our assumptions, when $t = 0$, we have $\exp\left(\omega^x_{rate}\left(t-\omega^x_{end}\right)\right) \approx 0$. Therefore, over time the function drops from ω^x_{high} to ω^x_{low} at a time controlled by ω^x_{end}, such that when $t = \omega^x_{end}$, we have that $f\left(t=\omega^x_{end},\Omega_x\right) = (\omega^x_{high}+\omega^x_{low})/2$. The ω_{rate} parameter changes the gradient of the descent at $t = \omega_{end}$.

Interaction across subpopulations is heterogeneous and is described in terms of time-sharing, determining proportions of subpopulations mixing in groups with each other. In the ith subpopulation there are $N_i(t) = \mathbf{S}_i(t) + \mathbf{E}_i(t) + \mathbf{I}_i(t) + \mathbf{A}_i(t) + \mathbf{R}_i(t)$ active individuals who can mix with others. The proportion from subpopulation i who travel to, and mix with, subpopulation k is p_{ik}. The effective population size of subpopulation k, given that others have travelled to it and some people from k have left, is $\hat{N}_k(t) = \sum_{j\in X} p_{jk} N_j(t)$. We assume these effective populations, $\hat{N}_k(t)$, are well mixed, so people who travel to each population can meet all others there.

Our specific time-share assumptions are described in the following paragraphs and represented visually by a directed, weighted network in Fig. 1(b). The corresponding weighted adjacency matrix, the travel/time-share matrix, $\boldsymbol{T}$ is described in Appendix B. There are two types of effective population: the care homes and the general population. Care home j, comprises $\hat{N}_{C_j}(t)$ people: its residents, its working staff, staff from other care homes, and visitors. The general population consists of $\hat{N}_G(t)$ people; this includes all the staff not at work and the non-visiting general population.

We assume that each resident has one visitor per day, up until 13th March 2020 (Nguyen et al., 2020a) when the policy changed to essential visitation only. The visitation rate, the proportion of the general population who travel and mix at the ith home each day, is described by

$$p_{G,C_i} = \frac{N_{C_i}(0)}{N_G(0)}\gamma(t) = \frac{N_{C_i}(0)}{N_G(0)} f(t,\Omega_\gamma). \tag{3}$$

Therefore, we model visitation by multiplying the constant proportion of the general population, $N_{C_i}(0)/N_G(0)$, that visit each care home, and the duration of the visit, $\gamma(t)$, measured as a proportion of a day. The duration of the visit, $\gamma(t)$, is described by Eq. (2) under the input parameters $\Omega_\gamma = \{\omega^\gamma_{end}, \omega^\gamma_{rate}, \omega^\gamma_{low}, \omega^\gamma_{high}\}$. Before 13th March 2020, we assume the duration of the visit, $\gamma(t)$, equals ω^γ_{high} and afterwards it quickly ($\omega^\gamma_{rate} = 3$) drops to zero ($\omega^\gamma_{low} = 0$), reflecting the policy change to essential visitation only (Burton et al., 2020; Cabinet Secretary for Health and Sport, 2020). Furthermore, since there are m care homes, the proportion of the general population not visiting care homes is $p_{G,G} = 1 - m\gamma(t)\frac{N_{C_i}(0)}{N_G(0)}$.

Workers spend a constant proportion of their time, δ, at care homes. With $\delta = 0.5$, a worker compartment, W_i, spends half of their time at care homes, C_i, over the course of a day. This is equivalent to care homes splitting staff into two 12 hour shifts. Workers thus spend the rest of their time, $1-\delta$, mixing in the general population, G. Care homes operate with staff under differing working schedules and require a minimum number of staff to maintain adequate levels of care. This places constraints on feasible values of δ. We assume this minimum value to be $\delta = 0.2$. This value equates to staff being split into five shifts throughout the day. Other possible shifts include care homes having a day and night shift ($\delta = 0.5$) or three 8-h shifts ($\delta = 0.33$).

During the first wave of COVID-19 in Scotland, there was both intra-organisational staff-sharing between homes (i.e., staff who work at multiple homes belonging to the same care provider), as well as inter-organisational staff-sharing (use of bank or agency staff) (Reilly et al., 2020; Office For National Statistics, 2020). Therefore, a constant proportion of each homes' assigned workers, ϵ, were exchanged between homes every day. We refer to this as *staff-sharing*. We have made the simplifying assumption that the staff-sharing network has a topology of a circle, whereby the shared staff for home j are split evenly between homes $j-1$ and $j+1$ (Fig. 1b). We also assume care home residents do not leave their homes ($p_{C_i,C_i} = 1$).

In summary, staff from the ith care homes mixing patterns across the network are described by the proportions $p_{W_i,G} = 1-\delta$, $p_{W_i,C_i} = (1-\epsilon)\delta$, and $p_{W_i,C_{i-1}} = p_{W_i,C_{i+1}} = \frac{\epsilon\delta}{2}$.

Disease transmission in the model is assumed to be frequency-dependent. The FOI integrates which infections occur to whom, from whom and where the infection takes place, as in Rădulescu et al. (2020) and Calvetti et al. (2020). The FOI acting on subpopulation i, $\boldsymbol{\Phi}_i$ (see Eq. (4)), accounts for the mixing that subpopulation i does in a day with all other subpopulations. It is most easily understood by considering $\boldsymbol{\Phi}_i\mathbf{S}_i$:

$$\boldsymbol{\Phi}_i\mathbf{S}_i = \sum_{k\ s.t.\ \hat{N}_k(t)\neq 0} \frac{\beta_k(t)p_{ik}\mathbf{S}_i}{\hat{N}_k(t)} \sum_{j\in X} p_{jk}\mathbf{I}_j \tag{4}$$

At effective population k, $\hat{N}_k(t)$, there is $p_{ik}\mathbf{S}_i$ susceptible individuals from i. At k there will also be $p_{jk}\mathbf{I}_j$ infectious people from j who have

Table 1
Parameter definitions, alongside their base case values and source. Units are given where appropriate in the Value column.

Parameter	Description	Value	Source
ε	Staff-sharing	0.4	Maximum likelihood estimation
δ	Proportion of time workers spend at care homes	0.5	Assumption[a]
$\mu_{i\in\{C_1,\ldots,C_m\}} = \mu_C$	Death rate for residents	0.25	Estimated (Burton et al., 2020; Ladhani et al., 2020; Graham et al., 2020; Byambasuren et al., 2020)
$\mu_{i\in\{G,W_1,\ldots,W_m\}} = \mu_G$	Death rate for general public (and workers)	0.017	Estimated (Dickson et al., 2021; Public Health Scotland, 2020; National Records of Scotland, 2019, 2020)
τ	Infectious period	7 days	He et al. (2020)
α	Latent period	5.8 days	McAloon et al. (2020)
m	Number of care homes	109	Burton et al. (2020)
$N(0) = \sum_i N_i(0)$	Total initial population	907,580	National Records of Scotland (2019)
$N_{C_i}(0)$	Initial resident subpopulation size	48	Burton et al. (2020)
$N_{W_i}(0)$	Initial worker subpopulation size	48	Knock et al. (2021)
$N_G(0)$	Initial general public subpopulation size	897,116	National Records of Scotland (2019), Burton et al. (2020) and Knock et al. (2021)
ω^C_{end}	Timing of R_t descent for care homes	36.086 days	Maximum likelihood estimation
ω^C_{rate}	Rate of descent of R_t for care homes	0.220	Maximum likelihood estimation
ω^C_{low}	Post-descent R_t for care homes	0.6	Assumption[b]
ω^C_{high}	Pre-descent R_t for care homes	5.241	Maximum likelihood estimation
ω^G_{end}	Timing of R_t descent for general population	23.656 days	Maximum likelihood estimation
ω^G_{rate}	Rate of descent of R_t for general population	6.627	Maximum likelihood estimation
ω^G_{low}	Post-descent R_t for general population	0.6	The Scottish Government (2020)
ω^G_{high}	Pre-descent R_t for the general population	4.503	Maximum likelihood estimation
ω^γ_{end}	Timing of descent for visitation	10 days	Burton et al. (2020) and Cabinet Secretary for Health and Sport (2020)
ω^γ_{rate}	Rate of descent of visitation	3	Burton et al. (2020) and Cabinet Secretary for Health and Sport (2020)
ω^γ_{low}	Post-descent value for visitation	0	Assumption[c]
ω^γ_{high}	Pre-descent value for visitation	0.155	Maximum likelihood estimation
H_{seeded}	Number of homes seeded	9	Maximum likelihood estimation
$E_G(0)$	Initial general population infections (exposed)	44.937	Maximum likelihood estimation

[a] We assume workers spend a half day at work, and the other half mixing in the general population. Alternatively, workers do 12-h shifts.
[b] We assume the reproductive rate for every sub-population drops to the Scottish government's (The Scottish Government, 2020) estimated R_t after lockdown (so $\omega^C_{low} = \omega^G_{low}$).
[c] Equals 0 to reflect the policy change to essential visitation only (Burton et al., 2020; Cabinet Secretary for Health and Sport, 2020), and to avoid the complication of modelling end-of-life visitation.

travelled to k. The transmission rate between individuals mixing in effective population k is $\beta_k(t)$, therefore

$$\frac{\beta_k(t)p_{ik}\mathbf{S}_i}{\hat{N}_k}\sum_{j\in X} p_{jk}(\mathbf{I}_j + \mathbf{A}_j)$$

is the number of new daily infections in i which occur in the effective population k.

There are two types of effective populations; the general population, and the care homes. We assume the transmission rate within each effective care home population is the same and equals $\beta_C(t)$. The general population transmission rate equals $\beta_G(t)$. The transmission rates $\beta_k(t)$ allow us to represent heterogeneous interaction patterns of individuals and incorporate the transmission dynamics of COVID-19 changing over time and location, for example, through lockdowns or other changes in behaviour (Rădulescu et al., 2020). We write $\beta_k(t) = \frac{R_k(t)}{\tau}$, describing the transmission rate $\beta_k(t)$ between individuals in effective population k, with the reproduction rate, $R_k(t)$, divided by the infectious period, τ. The reproduction rate $R_k(t)$ is the average number of infections that infected individuals mixing only in effective population k would cause over the duration of their infectiousness, and it captures both contact rate and infection probability between susceptible and infected individuals.

The transmission rates are described by:

$$\beta_C(t) = \frac{f\left(t, \left\{\omega^c_{end}, \omega^c_{rate}, \omega^c_{low}, \omega^c_{high}\right\}\right)}{\tau} = \frac{R_C(t)}{\tau},$$
$$\beta_G(t) = \frac{f\left(t, \left\{\omega^G_{end}, \omega^G_{rate}, \omega^G_{low}, \omega^G_{low}, \omega^G_{high}\right\}\right)}{\tau} = \frac{R_G(t)}{\tau}. \quad (5)$$

The reproduction rates, $R_k(t)$, are modelled using the function $f(t, \{.\})$ (Eq. (2)) with different parameter sets.

2.2. Model calibration process

We used data from the network of care homes in NHS Lothian (Burton et al., 2020) complemented by Public Health Scotland Open Data, breaking down COVID-19 cases and deaths per health board (Public Health Scotland, 2020; National Records of Scotland, 2020), to inform and calibrate our model. Parameters were found using a mixture of methods, as indicated in Table 1. Some were estimated directly from a combination of evidence and assumptions from a literature search, and the remaining free parameters were estimated using maximum likelihood.

2.2.1. Data

NHS Lothian is the second-largest health board in Scotland NHS Lothian (2021), providing public health services to an estimated 907,580 people (2019 mid-year population estimate National Records of Scotland, 2019). The daily confirmed positive tests of COVID-19 cases reported across the entire health board were taken from the Public Health Scotland Open Data (Public Health Scotland, 2020). This data does not delineate which cases occurred in care homes, and thus, we retrieved the subset of cases in care homes from Burton et al. (2020), which reports a 7-day average of confirmed cases in care home residents. Weekly COVID-19 deaths at the NHS Lothian health board level come from National Records Scotland (National Records of Scotland, 2020). Care home resident deaths are a subset of these and are published in Burton et al. (2020). Both death data are weekly counts of registered deaths where COVID-19 is mentioned on the death certificate (either as the underlying cause or as a contributory factor) (National Records of Scotland, 2020).

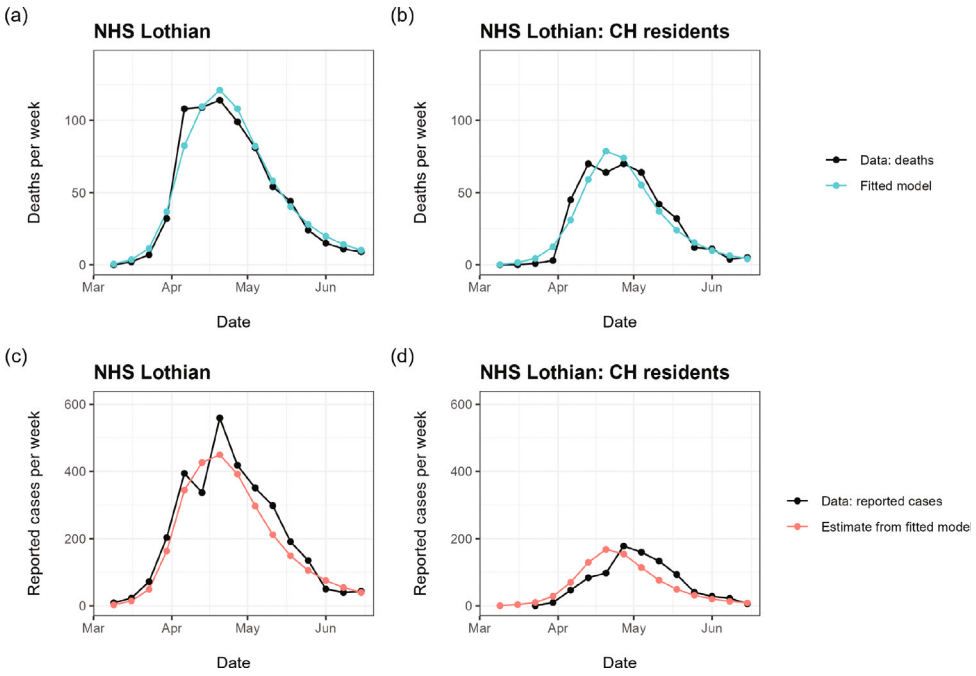

Fig. 2. Surveillance data and fitted model outputs. Surveillance data are shown by black lines: (a) deaths per week for all NHS Lothian inhabitants (care home residents, workers and the general population); (b) deaths per week in NHS Lothian care home residents; (c) reported cases per week for all NHS Lothian inhabitants (care home residents, workers and the general population); (d) reported cases per week in NHS Lothian care home residents. Panels (a) and (b) show the death count data used during maximum likelihood estimation and fitted count estimates from the model, i.e., the model with parameter values in Table 1. The estimate of reported cases in panels (c) and (d) were derived by assuming that a constant proportion of the new individuals that become infectious in each population will test positive — see Appendix A.

2.2.2. Parameters set from evidence and assumptions

In this section, we outline our parameter estimates for some of the parameters responsible for inter-subpopulation mixing ($\omega^{\gamma}_{rate}, \omega^{\gamma}_{low}, \delta$), and seeding infection in the model, as well as the death rates.

We assume that $t = 0$ in the model corresponds to March 4th, 2020, the first day with a reported COVID-19 case in NHS Lothian. We make several assumptions about the population initially infected with COVID-19. In the general population, we assume an equal amount of exposed and infectious individuals, i.e., $E_G(0) = I_G(0)$. In our model, care homes were seeded with infections via the parameter $H_{seeded} = |\{C_j \in \{C_1, \ldots, C_m\} : E_{C_j}(0) > 0\}|$. These exposed individuals represent any introductions that could have taken place prior to the start of the simulation, such as through visitors or hospital discharges. Since there were reported cases in the data for care home residents by March the 9th, 2020, we assume for all $j \in \{1, \ldots, m\}$, $I_{C_j}(0)$. To account for the delay in infections at the start of the pandemic in care homes compared to the general population, as seen in the data Fig. 2, we assume for all $j \in \{1, \ldots, m\}$, $I_{C_j}(0) = A_{C_j}(0) = 0$. Initially infected homes were seeded equally spaced on the circular staff-sharing structure (see Fig. 1(b)). If a home is seeded then we assume $E_{C_j}(0) = 1$, and if not, $E_{C_j}(0) = 0$.

In terms of the visitation parameters, $\Omega_\gamma = \{\omega^{\gamma}_{end}, \omega^{\gamma}_{rate}, \omega^{\gamma}_{low}, \omega^{\gamma}_{high}\}$, we set the rate that levels fell $\omega^{\gamma}_{rate} = 3$ and the timing $\omega^{\gamma}_{end} = 10$ to reflect the rapid visitation policy changes in care homes (Burton et al., 2020; Cabinet Secretary for Health and Sport, 2020). We have made the simplifying assumption that the post-lockdown visitation level $\omega^{\gamma}_{low} = 0$, to avoid the complications of modelling end-of-life visitation in care homes. We estimate the proportion of each visitor's day (pre-lockdown), ω^{γ}_{high}, spent at homes using maximum likelihood, discussed below. For simplicity, we assume that all homes operate under two 12-h shifts per day, i.e., $\delta = 0.5$. Other shifts are explored in the sensitivity analysis.

Using a combination of data and assumptions, we derive estimates for the constant resident and general population death rate parameters (μ_C, μ_G). The details are contained in Appendix A. We first derive estimates of reporting rates, which we define as the proportion of COVID-19 infections which are identified by a positive test. We assume the reporting rate differs for the care home resident and the general population and is given by r_C and r_G, respectively. Then, using the numbers of reported cases and deaths in each population in NHS Lothian over the study period, we estimate the true number of infections and, therefore, the death rate for each population (μ_C, μ_G).

2.2.3. Maximum likelihood estimation

While some model parameter values can be found based on the external data and literature, as shown in the previous section, the remaining parameters θ were estimated using maximum likelihood methods, Table 1. We attempt to reproduce the dynamics of COVID-19 deaths over the first wave by estimating the set of parameters θ that maximises the log-likelihood (LL) of observing the weekly numbers of reported deaths for the care home residents and total population in NHS Lothian. The set of parameters to be estimated is

$$\theta = \{\omega^{C}_{end}, \omega^{C}_{rate}, \omega^{C}_{high}, \omega^{G}_{end}, \omega^{G}_{rate}, \omega^{G}_{high}, \varepsilon, \omega^{\gamma}_{high}, H_{seeded}, E_G(0)\}. \quad (6)$$

We assume the weekly counts of reported COVID-19 deaths are Poisson distributed, and let $Y_R = \{Y_{R1}, Y_{R2}, \ldots, Y_{Rn}\}$ and $Y_T = \{Y_{T1}, Y_{T2}, \ldots, Y_{Tn}\}$ be the observed death counts for the resident and total population, respectively. There are $n = 15$ weeks of death counts. The first week starts on March the 9th, 2020, and the last week starts on June the 15th, 2020. We let $\Lambda_R(\theta) = \{\lambda_{R1}(\theta), \lambda_{R2}(\theta), \ldots, \lambda_{Rn}(\theta)\}$ and $\Lambda_T(\theta) = \{\lambda_{T1}(\theta), \lambda_{T2}(\theta), \ldots, \lambda_{Tn}(\theta)\}$ be the expected weekly death counts for the resident and total population (i.e., general, staff and residents) produced by our model. The log-likelihood function is given by

$$\log L(\theta|Y_R, Y_T) = \log L_R(\theta|Y_R) + \log L_T(\theta|Y_T), \quad (7)$$

where

$$\log L_R(\theta|Y_R) = \sum_{i=1}^{n} \left(Y_{Ri} \cdot \log(\lambda_{Ri}(\theta)) - \lambda_{Ri}(\theta) - \log(Y_{Ri}!)\right)$$

and

$$\log L_G(\theta|Y_T) = \sum_{i=1}^{n} \left(Y_{Ti} \cdot \log(\lambda_{Ti}(\theta)) - \lambda_{Ti}(\theta) - \log(Y_{Ti}!)\right).$$

Table 2
The search space explored by the genetic algorithm for each parameter in θ.

Parameter	Description	Range
ω^C_{end}	Timing of Rt descent for care homes	[30, 50]
ω^C_{rate}	Rate of descent of Rt for care homes	[0, 10]
ω^C_{high}	Pre-descent Rt for care homes	[3, 5.5]
ω^G_{end}	Timing of Rt descent for general population	[15, 30]
ω^G_{rate}	Rate of descent of Rt for general population	[0, 3]
ω^G_{high}	Pre-descent Rt for general population	[3, 5.5]
ϵ	Staff-sharing	[0, 0.5]
ω^γ_{high}	Pre-descent value for visitation	$\left[0, \frac{4}{24}\right]$
H_{seeded}	Number of homes seeded with infection	[0, 20]
$E_G(0)$	Initial general population infections (exposed)	[0, 200]

We maximise the combined log-likelihood function (Eq. (7)) with respect to the parameter set θ (Eq. (6)) using numerical optimisation methods. Specifically, we implemented a parallelised genetic algorithm using the ga package in RStudio (Flasch, 2012), configured with 4000 iterations, to search for the maximum likelihood estimators. To ensure convergence of the genetic algorithm (GA) to the same maximum, and gain confidence the maximum is global, we performed five separate runs of the GA with randomised initial conditions. The search space explored by the genetic algorithm for each parameter in θ is shown in Table 2.

2.3. Sensitivity analysis

We measured the change in total care home resident deaths for different combinations of pairs of the time-share parameters (ω^γ_{high}, δ, ϵ), keeping all other parameters at the base case. The results were stored in a 50 × 50 grid and visualised using heat maps to assess the mixing patterns that spread COVID-19 into and throughout care homes and identify potential mitigation strategies.

3. Results

In this section, we first show how the model captures the NHS Lothian data for cases and deaths in the period from March to June 2020, and then show how sensitive the results are to changes in key parameters.

3.1. Data fit

The fitted model captures the key features of the COVID-19-related deaths in both the care home resident and total population in NHS Lothian, Fig. 2. The maximum likelihood estimators found by the genetic algorithm search are shown in Table 2. The corresponding maximum log-likelihood was −92.9, with our model predicting 726 total COVID-19-related deaths in total, including 413 in care home residents, compared to the 709 and 423 recorded by the data (Fig. 2). Additionally, investigation of the sensitivity of log-likelihood to perturbations in the parameters estimated by the GA confirms that the algorithm found a maximum, Appendix C Fig. 5.

The recorded number of COVID-19 deaths per week in NHS Lothian rose slowly from the second week in March, reaching over 100 by the first week of April, Fig. 2a. The weekly deaths were similarly high for the next two weeks before gradually falling back towards zero by the third week in June. Recorded COVID-19 care home residents deaths were ≈0 each week in March before rising sharply in the first week of April, which we note is a delayed and sharper rise as well as a later fall than the total death count in NHS Lothian, Fig. 2b. The resident death count was then ≈75 per week from the second week in April until the first week of May, before gradually falling back to none by mid-June. The model estimates a gradual rise in resident deaths per week from mid-March, which peaks in the third week of April, Fig. 2b. The steepness of the rise and width of the peak relative to the data are both underestimated slightly. Similarly, the model underestimates the total deaths in the first week of April, otherwise matching the wave of COVID-19 deaths in NHS Lothian between March and June 2020.

For comparison/validation, we show how our model would reproduce the dynamics of reported COVID-19 cases (positive tests), by assuming that a constant proportion of those infected with COVID-19 are picked up by testing. The total reported cases in NHS Lothian are shown in Fig. 2c and those only in care home residents in Fig. 2d. The constant reporting rates used are derived in Appendix A. The model with constant reporting struggles to reproduce the dynamics of COVID-19 positive tests and deaths concurrently due to the underlying dynamics of the data. For example, the timeseries of resident deaths (Fig. 2b) rose more sharply and peaked earlier than the reported resident tests (Fig. 2d) – most likely due to testing changes. In Scotland, the policy from the start of March to 16th April 2020 was to test only the first few symptomatic residents, and afterwards, was to test all symptomatic residents (Burton et al., 2020).

The fitted reproduction rates change significantly over the period of late April – May 2020, Fig. 3 reflecting the strict social distancing restrictions (lockdown) implemented in mid-March. Both the general and care home resident population reproductive rates begin to drop at the same time (late March). The general population's reproductive rate drops below 1 after a few days, while the care home's higher reproductive rate does not fall below 1 until late April, Fig. 3.

Additionally, we investigated the sensitivity of maximum log-likelihood (Eq. (7)) to perturbations in each fitted parameter, Appendix C. These suggest that the GA had found a maximum, while providing insight into the relative impact a unit change in each parameter has on the quality of the fit.

3.2. Sensitivity analysis

In our model, staff catching infections from the community is controlled by δ (staff time spent at homes); staff spreading infections between homes through staff-sharing by ϵ; and visitors bringing infections into homes from the outside community (prior to the visitation ban) by ω^γ_{high}. We determine which combinations of these time-share parameters have the greatest impact on the total resident deaths over the first wave of COVID-19 using Fig. 4. This assesses mixing patterns and helps identify mitigation strategies for the simulated epidemic.

Elimination of staff-sharing (a reduction from $\epsilon = 0.38$ to $\epsilon = 0$) while keeping the time staff spend at care homes at baseline ($\delta = 0.5$) causes a ∼40% reduction in resident deaths, Fig. 4(a) (b). Increasing the time staff spend in care homes to a live-in situation ($\delta = 0.5$ to $\delta = 1$), but not restricting staff-sharing (keeping at baseline $\epsilon = 0.38$), would cause a ∼45% increase in resident deaths since it concentrates transmission across the network of isolated units, Fig. 4(a). However, if staff living in care homes were not shared across them ($\delta = 1$, $\epsilon = 0$), this strategy would cause a reduction in resident deaths of ∼60%, compared to the baseline count, Fig. 4(a).

Decreases in staff-sharing (ϵ) cause a more significant reduction of deaths compared to lowering pre-lockdown visitation levels (ω^γ_{high}), while holding all other parameters constant at their respective values in Table 1, Fig. 4(b). Reducing pre-lockdown visitation from 4 visiting hours per resident to 0 h per resident would only reduce our predicted first-wave deaths by about 2% of the death count in the first wave in NHS Lothian, Fig. 4(b). However, cancelling both visitation and staff sharing causes a ∼45% reduction in the observed resident deaths. Additionally, if staff spend all their time in care homes ($\delta = 1$) (e.g., staff live in homes) and there was no staff-sharing ($\epsilon = 0$) or pre-lockdown visitation ($\omega^\gamma_{high} = 0$), there is a significant ∼75% reduction in the observed resident deaths.

4. Discussion

In this paper, we developed a generic compartmental metapopulation model which allowed for the parameterisation of mixing across

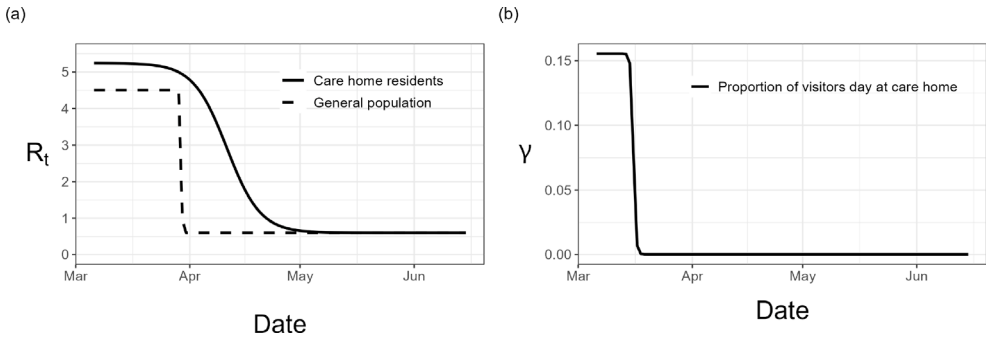

Fig. 3. Fitted time-dependent parameters. (a) Fitted reproductive numbers over time for care home residents, $R_C(t)$, and general population, $R_G(t)$; (b) fitted visitation, γ, over time with drop highlighting the change in policy. These time-dependent parameters are described by the function $f\left(t, \Omega_x\right)$, Eq. (2), with input parameters $\Omega_x = \{\omega^x_{end}, \omega^x_{rate}, \omega^x_{low}, \omega^x_{high}\}$ taken from Table 2.

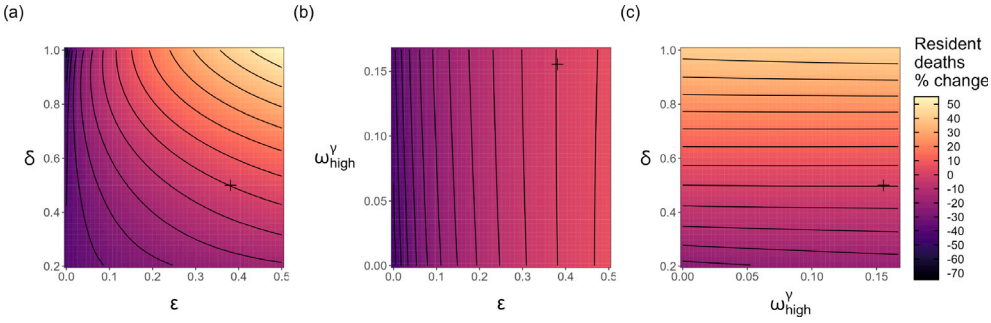

Fig. 4. Sensitivity of the total resident deaths by the end of the first wave to the time-share/mixing parameters $(\delta, \epsilon, \omega^{\gamma}_{high})$. Proportion of CH staff time at work is δ, proportion of staff shared between homes is ϵ, and pre-lockdown visitation is ω^{γ}_{high}. Each panel shows the percentage change in the total resident deaths over the first wave for different combinations of two of the time-share/mixing parameters, relative to baseline, with all other model parameters fixed as the base case (Table 2). The baseline resident death count is 413, which are those observed in the fitted model with parameter values in Table 2. The cross in each panel indicates the base case values for each parameter. The black lines in each panel are isoclines.

subpopulations, specifying where subpopulations spend their time at each location as a proportion of their day. This model can capture the movement of populations between locations and activities and the complex transmission dynamics which arise. Here, it has been used to simulate the propagation of COVID-19 throughout care homes in NHS Lothian, Scotland. However, this methodology can be applied in many other contexts. For example, to model the mixing of individuals across different locations with heterogenous levels of local transmission, such as workplaces or prisons.

To assess the impact of inter-subpopulation mixing between care home residents, care home staff and the general population in relation to the spread of COVID-19 during the first wave of the pandemic, we deployed a combination of modelling, data fit and simulations. With this view, we find that restricting staff-sharing would be an effective disease control measure. Additionally, the modelled strategy of staff living in care homes with no staff-sharing would reduce COVID-related deaths of care home residents by 65% during the first wave of the pandemic. Importantly, this result holds for any contact structure between staff and care homes. Furthermore, a restriction in the movements of all populations (no visitation, no staff-sharing, staff living in care homes) could significantly limit disease spread, reducing cases in residents by approximately 75%.

These results from our model indicate that protecting care home staff from the disease, coupled with reductions in staff-sharing across care homes and expedient cancellations of visitations, can effectively reduce outbreak size in care homes. Additionally, our findings highlight the need for more planning and support for care homes and their staff in organising quick and effective responses to emerging pandemics.

In particular, our results point to a strategy of staff living-in the care home, in conjunction with timely lowering of visitation, as an effective pandemic response. This was implemented in France, where outbreaks in care homes were reduced significantly in care homes where staff self-confined (Belmin et al., 2020). If living within the care home is not possible, this result may imply that the strategy of segmenting the staff away from both care home residents and the general population whilst they are not at work would be effective, i.e., organisation of accommodation for care home workers (Social Care Working Group (SCWG) for the Scientific Advisory Group for Emergencies, 2021).

Eliminating staff-sharing did not completely eliminate outbreaks in our model simulations, suggesting that more restrictive control measures or protection for staff are needed. However, reducing staff-sharing does significantly reduce the outbreak severity, which supports existing literature (McMichael et al., 2020; Reilly et al., 2020). This conclusion is limited due to our assumption of the circular contact structure, which in turn reflects limited data availability regarding the contact structure of the care home in Scottish Health boards (Nguyen and Megiddo, 2022) (most likely due to commercial sensitivity). A different contact structure could result in staff-sharing leading to changes in the exportation of infection from care homes with outbreaks. A more thorough examination on the contact-structure of this system and how that impacts disease spread dynamics would be an important contribution to current literature.

One way to achieve this would be to consider an addition of highly-connected hubs (Barabási and Bonabeau, 2003), which we expect would increase the effect of staff-sharing. In our simulations, staff-sharing has an effect when there exists a non-uniform distribution of infections in worker sub-populations. Worker sub-populations acting as hubs would acquire disease quicker and skew the distribution of infection amongst worker sub-populations. However, the general population strongly connects all nodes in the network, and dominates the impact of the staff-sharing network on disease spread, as we assume a single general population with full mixing. At the geographical scale modelled

here (a health board) this is an appropriate assumption, although it would not hold for larger scales, e.g., the national level (Scotland). Furthermore, our results from considering no staff-sharing ($\varepsilon = 0$) will be the same for any contact structure . Thus, it is important to note that the significant impact of the strategy of staff living in care homes with no staff-sharing, is independent of contact structure.

A reduction in visitation reduces predicted resident deaths, as speculated in Social Care Working Group (SCWG) for the Scientific Advisory Group for Emergencies (2021), and our model supports findings that visitation was not the driving cause of infection in care homes (Comas-herrera, 2020). Since visitation was banned, the evidence for visitation causing outbreaks is limited. Investigation into continuing visitation during lockdown or similar would be necessary to see how the outcome would be different if visitation did not change at all; this was not the focus of our investigation.

One of the model's limitations is that it does not explicitly account for the variation in susceptibility with age (Davies et al., 2020), and is only implicitly addressed by considering different values of β within and outwith of care homes, while keeping the staff and general population homogeneous. Due to the unavailability of data regarding care home worker infections, we assumed the resident-resident and resident-worker transmission rates were equal. However, contacts between care-giving staff and residents are likely more frequent and closer than between residents. On the other hand, there may be more adherence or better knowledge of how to use PPE among staff. Additionally, contact between residents could be reduced more easily during the pandemic (Social Care Working Group (SCWG) for the Scientific Advisory Group for Emergencies, 2021). Furthermore, we do not explicitly model self-isolation or any behavioural change after infection.

In our model, we assume a uniform home size in order to keep the model generic. As a result (and since the model is deterministic) the risk of staff and visitors bringing in infections is the same for all care homes, which may result in an underestimation of the initial rate of spread. An obvious extension of the paper would be to consider various sources of heterogeneity, including size. The size of individual care homes is believed to be the main factor that influences the likelihood of a care home outbreak (Reilly et al., 2020; Burton et al., 2020, 2021). However, larger homes typically have more staff and therefore a higher chance of experiencing an outbreak before the smaller ones. In general, we expect larger care homes to receive an increased force of infection from all sources, proportional to its increased size, and therefore an increased outbreak risk. This in turn could increase risk for smaller homes directly connected to the larger ones through staff-sharing and visitations, and the overall outbreak risk. However, this effect could be balanced by a lowered risk associated with small care homes, with the total population size kept constant.

The National Records Scotland death data used were the dates of death registration, not the actual date of death. This is limiting, as we are an average of three days behind in the prediction of deaths (National Records of Scotland, 2020). The data for care home resident deaths includes deaths in hospitals, including nosocomial infections, which we do not take into account in our model. We expect this not to limit the interpretation of our results, as hospital deaths of care home residents were approximately only 5% of the total care home resident deaths (National Records of Scotland, 2020).

We focused our analysis on the mixing patterns of patient-facing care home workers and have not considered the impact of non-care staff. From a study on care homes in Norfolk during April and May 2020, the number of non-care staff in homes was found to be the most statistically significant predictor of COVID-19 entry into homes (Brainard et al., 2020). We do note, however, that this study did not consider the role of staff-sharing. A valuable extension of our model would include non-care worker subpopulations with different contact structures to other homes compared to patient-facing staff.

Data regarding care home outbreaks were limited due to the commercial nature of care home organisations in Scotland. Making this data available would allow for additional modelling approaches. Adding further heterogeneity into the system by including a distribution of home sizes and types would further improve the modelling approach. Including a stochastic component to this model could lead to more insight into "super-spreader" events in care homes (Majra et al., 2021) and their effect on epidemic response.

CRediT authorship contribution statement

Matthew Baister: Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Ewan McTaggart:** Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Paul McMenemy:** Methodology, Project administration, Resources, Software, Supervision, Writing – review & editing. **Itamar Megiddo:** Methodology, Project administration, Supervision, Writing – review & editing. **Adam Kleczkowski:** Conceptualization, Funding acquisition, Methodology, Project administration, Resources, Software, Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Care home data are available on request from Prof B. Guthrie (https://doi.org/10.1016/S2666-7568(20)30012-X). All code is available at GitHub: https://github.com/ewanmct/COVID-care-homes.

Acknowledgements

The authors thank Public Health Scotland for their support, and Prof. Bruce Guthrie (University of Edinburgh) for sharing the Lothian care home data. EMT has been partially supported by the University of Strathclyde Student Excellence Award, UK and MB by Defra, UK. Results were obtained using the ARCHIE-WeSt High Performance Computer (www.archie-west.ac.uk) based at the University of Strathclyde. Results were also obtained using the Pin cluster at the University of Stirling. The authors thank the referees for their constructive input.

Appendix A. Parameter assumptions and estimates

In this section, we derive estimates for the resident and general population death rate parameters. To do so, we first derive estimates of reporting rates, which we define as the proportion of COVID-19 infections which are identified by a positive test. We assume the reporting rate differs for the care home resident and the general population and are given by r_C and r_G, respectively. Then, using the numbers of reported cases and deaths in each population in NHS Lothian over the study period, we estimate the true number of infections and, therefore, the death rate for each population (μ_C, μ_G).

A Scottish population study between 10th April to 15th June (Dickson et al., 2021) estimated a combined adjusted seroprevalence across their study period (first wave = 10th April to 15th June) of 4.3% (95% CI 4.2%–4.5%). As of the week beginning 15th June 2020, there had been 18,077 positive tests (Public Health Scotland, 2020), which as a percentage of Scotland's population (2019 census National Records of Scotland, 2019) is ∼0.33%. We use this information to assume a constant reporting rate in the first wave for the general public of $r_G = 0.33/4.3 \sim 0.077$.

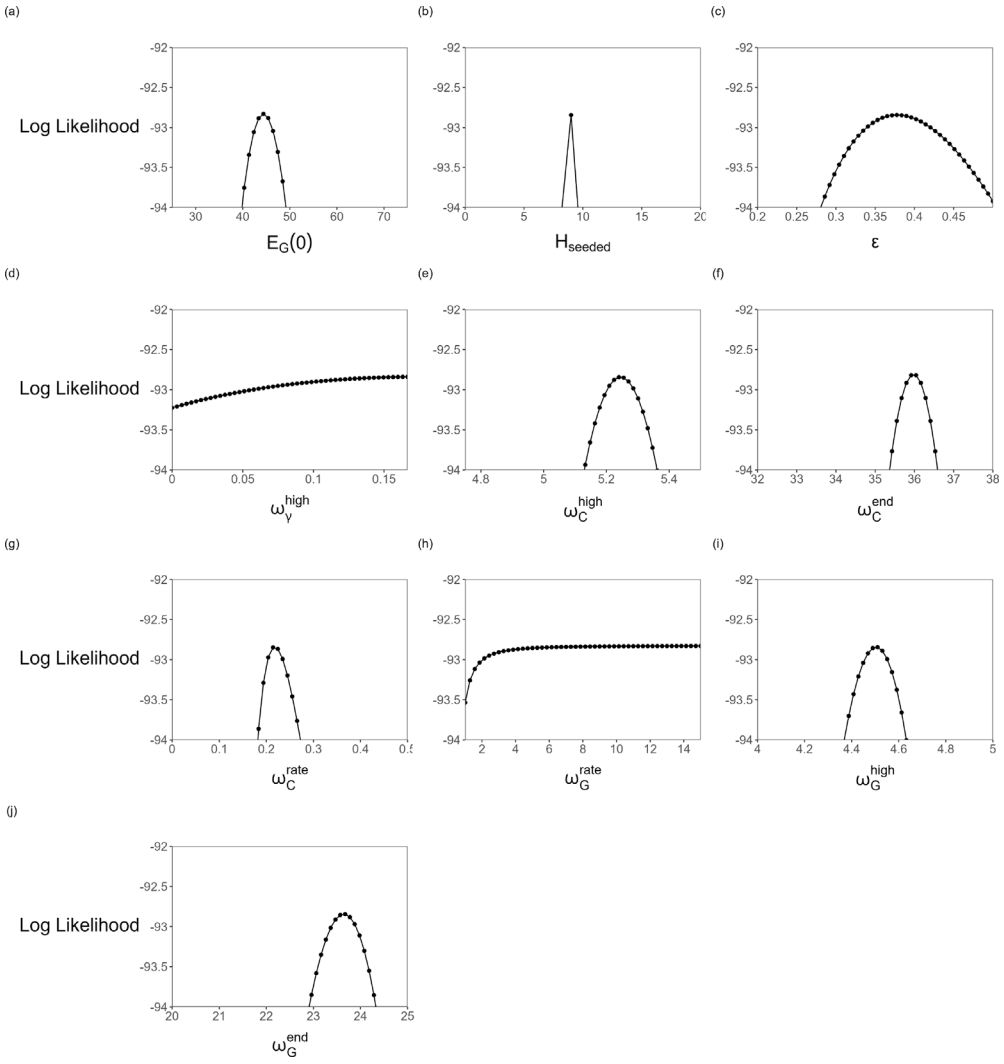

Fig. 5. Log-likelihood profiles. Log-likelihood is plotted as a function of each individual parameter fitted using maximum likelihood estimation, all other parameters are held at the values in Table 2. (a) Initial general population infections (exposed), $E_G(0)$; (b) number of homes seeded with infection, H_{seeded}; (c) staff-sharing, ϵ; (d) pre-descent value for visitation, ω_γ^{high}; (e) pre-descent Rt for care homes, ω_C^{high}; (f) timing of Rt descent for care homes, ω_C^{end}; (g) rate of descent of Rt for care homes, ω_C^{rate}; (h) rate of descent of Rt for general population, ω_G^{rate}; (i) pre-descent Rt for general population, ω_G^{high}; (j) timing of Rt descent for general population, ω_G^{end}.

In Scotland, the policy from the start of March to 16th April 2020 was to test only the first few symptomatic care home residents, and afterwards, was to test all symptomatic residents (Burton et al., 2020). Assuming when there is an outbreak in a home, 40% of the residents end up infected (40% incidence) (Ladhani et al., 2020; Graham et al., 2020). Given 48 residents per care home, until 16th of April we assume a reporting rate of (a few tested)/(total infected) = $3/(0.4 \times 48) = 5/32$. After 16th April, we assume all the symptomatic cases are reported, giving a reporting rate of 4/5 (an estimated symptomatic proportion of COVID-19 cases in long term aged care is 80% Byambasuren et al., 2020). Between the start of our simulation (6th March 2020) and 16th April 2020 is a time difference of 42 days, and between 17th April 2020 and the end of our simulation period (15th June 2020) is a time difference of 60 days. Therefore, for 42 days, we assume a reporting rate of 5/32, and for 60 days, it is 4/5. The weighted average and constant CH reporting rate over the simulation period is $r_C = (5/32)(42/102) + (4/5)(60/102) \sim 0.53$.

There are two constant death rates in our model: a resident death rate (μ_C) and a general population death rate (μ_G). We assume care home staff have the same death rate as the general population. There were $\sim$ 899 positive tests and 423 deaths in NHS Lothian care home residents over the study period. Using our resident reporting rate, we estimate there were $899.1/0.53 \sim 1697$ total residents infected with COVID-19 over the study period. Therefore, we estimate a resident death rate of $\mu_C = 423/1697 \sim 0.25$. Similarly, there were 3123 total positive tests and 709 deaths over the study period in NHS Lothian overall. Using our general reporting rate, r_G, we estimate a general population death rate of $\mu_G = 709/(3123/0.077) \sim 0.017$.

Appendix B. Force of infection, time-share and transmission rate matrix representations

In this section, we derive a matrix interpretation of the model dynamics and show how the force of infection $\boldsymbol{\Phi}_i$, Eq. (4), can be expressed in terms of matrix multiplication. This is useful for programming our model in code and solving the differential equations numerically.

Our specific time-share assumptions are represented visually as a directed, weighted network in Fig. 1(b). The corresponding weighted adjacency matrix, the travel/time-share matrix, is $\boldsymbol{T} \in \mathbb{R}^{n \times n}$, whose

$[i,j]^{th}$ element is p_{ij}. The are n nodes (subpopulations) in the network and and each row of the matrix sums $\boldsymbol{T}$ to 1. The rows and columns of $\boldsymbol{T}$ are in the order of $\{C_1, C_2, \ldots, C_m, W_1, W_2, \ldots, W_m, G\}$. $\boldsymbol{T}$ consists of the partitions $\{T_{CC}, T_{CW}, T_{WC}, T_{WW}, T_{CG}, T_{WG}, T_{GC}, T_{GW}\}$. For example, the submatrix T_{WC} defines the proportion of time that each worker subpopulation spends mixing in each care home. To clarify notation: matrix $\boldsymbol{I}_m$ indicates the identity matrix of dimension m, matrix $[a]_{m\times m}$ indicates a matrix of dimension $m \times m$ with all entries a. Hence $\boldsymbol{T}$ and the subsequent sub-matrices are as follows:

$$\boldsymbol{T} = \begin{bmatrix} \boldsymbol{T_{CC}} & \boldsymbol{T_{CW}} & \boldsymbol{T_{CG}} \\ \boldsymbol{T_{WC}} & \boldsymbol{T_{WW}} & \boldsymbol{T_{WG}} \\ \boldsymbol{T_{GC}} & \boldsymbol{T_{GW}} & \boldsymbol{T_{GG}} \end{bmatrix}_{n\times n}, \tag{8}$$

$$\boldsymbol{T_{CC}} = \boldsymbol{I}_m, \qquad \boldsymbol{T_{CW}} = \boldsymbol{T_{WW}} = \begin{bmatrix} 0 \end{bmatrix}_{m\times m}, \qquad \boldsymbol{T_{CG}} = \begin{bmatrix} 0 \end{bmatrix}_{m\times 1},$$

$$\boldsymbol{T_{GW}} = \begin{bmatrix} 0 \end{bmatrix}_{1\times m},$$

$$\boldsymbol{T_{WG}} = \begin{bmatrix} 1-\delta \end{bmatrix}_{m\times 1}, \qquad \boldsymbol{T_{GC}} = \begin{bmatrix} \frac{N_C(0)}{N_G(0)}\gamma(t) \end{bmatrix}_{1\times m},$$

$$\boldsymbol{T_{GG}} = \begin{bmatrix} 1 - m\frac{N_C(0)}{N_G(0)}\gamma(t) \end{bmatrix}_{1\times 1},$$

$$\boldsymbol{T_{WC}} = \begin{bmatrix} (1-\varepsilon)\delta & \frac{\varepsilon\delta}{2} & 0 & \ldots & 0 & \frac{\varepsilon\delta}{2} \\ \frac{\varepsilon\delta}{2} & (1-\varepsilon)\delta & \frac{\varepsilon\delta}{2} & \ldots & 0 & 0 \\ 0 & \frac{\varepsilon\delta}{2} & (1-\varepsilon)\delta & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ldots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & (1-\varepsilon)\delta & \frac{\varepsilon\delta}{2} \\ \frac{\varepsilon\delta}{2} & 0 & 0 & \ldots & \frac{\varepsilon\delta}{2} & (1-\varepsilon)\delta \end{bmatrix}_{m\times m}. \tag{9}$$

Furthermore, the transmission rates can be arranged in a matrix $\boldsymbol{\beta} \in \mathbb{R}^{1\times n}$ whose kth element is $\beta_k(t)$. The elements of $\boldsymbol{\beta}$ are in the order of $\{C_1, C_2, \ldots, C_m, W_1, W_2, ..., W_m, G\}$. The matrix $\boldsymbol{\beta}$ can be written in terms of block sub-matrices, as shown below;

$$\boldsymbol{\beta} = \begin{bmatrix} \begin{bmatrix} \beta_C(t) \end{bmatrix}_{1\times m} & \begin{bmatrix} 0 \end{bmatrix}_{1\times m} & \begin{bmatrix} \beta_G(t) \end{bmatrix}_{1\times 1} \end{bmatrix}, \tag{10}$$

Recall that the force of infection $\boldsymbol{\Phi}_i$, Eq. (4) is given by

$$\boldsymbol{\Phi}_i = \sum_{k\ s.t.\ \hat{N}_k(t)\neq 0} \frac{\beta_k(t) p_{ik}}{\hat{N}_k(t)} \sum_{j\in X} p_{jk} \mathbf{I}_j$$

Let $\mathbf{S} \in \mathbb{R}^{1\times n}$, $\mathbf{E} \in \mathbb{R}^{1\times n}$, $\mathbf{I} \in \mathbb{R}^{1\times n}$, and $\mathbf{R} \in \mathbb{R}^{1\times n}$ be the vectors of susceptible, exposed, infected, and recovered individuals. The vector of effective populations is given by $\hat{\mathbf{N}} \in \mathbb{R}^{1\times n}$, defined below

$$\hat{\mathbf{N}} = (\mathbf{S} + \mathbf{E} + \mathbf{I} + \mathbf{R})\,\mathbf{T}.$$

Where $\mathbf{T}$ is the time-share matrix discussed above. The kth element of $\mathbf{IT} \in \mathbb{R}^{1\times n}$ gives us the total number of infectious individuals at effective population k. Now, we let the symbols $\odot$ and $\oslash$ denote element-wise multiplication, and element-wise division between two matrices of the same dimension.

The force of infection $\boldsymbol{\Phi}_i$ is given by the ith element of

$$(\mathbf{T})\left(\boldsymbol{\beta} \odot (\mathbf{IT}) \oslash \hat{\mathbf{N}}\right)^T,$$

where we set the NaN elements of $\boldsymbol{\beta} \odot (\mathbf{IT}) \oslash \hat{\mathbf{N}}$ equal to zero before taking the transpose. The NaN elements arise since individuals in our model mix only in the care home or general effective populations, so the staff effective populations have size zero ($\hat{\mathbf{N}}$ has elements taking value zero).

Appendix C. Sensitivity of log-likelihood to perturbations in maximum likelihood estimates

In this section, we show the sensitivity of the estimated maximum log-likelihood (Eq. (7)) to perturbations in each fitted parameter, Fig. 5. These suggest that the genetic algorithm found a maximum, while providing insight into the relative impact a unit change in each parameter has on the quality of the fit. Log-likelihood took an open downwards parabolic shape with respect to changes in almost all the fitted (using MLE) parameters. The only exception was the rate of descent of Rt for general population, ω_G^{rate}: log-likelihood is initially an increasing function of this parameter but then levels out (an elbow shape). Increased ω_G^{rate} causes the general population Rt to behave closer to a step function. When $\omega_G^{rate} > 5$ (and all other parameters fixed), the general population Rt is a step function where the drop in Rt is instantaneous, i.e., less than 1 day.

References

Barabási, A., Bonabeau, E., 2003. Scale-free networks. Sci. Am. 288, 60–69. http://dx.doi.org/10.2307/26060284.

Bell, D., Comas-Herrera, A., Henderson, D., Jones, S., Lemmon, E., Moro, M., et al., 2020. COVID-19 mortality and long-term care: a UK comparison. LTCcovid. Available from: https://ltccovid.org/wp-content/uploads/2020/08/COVID-19-mortality-in-long-term-care-final-Sat-29-1.pdf.

Belmin, J., Um-Din, N., Donadio, C., Magri, M., Nghiem, Q.D., Oquendo, B., et al., 2020. Coronavirus disease 2019 outcomes in french nursing homes that implemented staff confinement with residents. JAMA Netw. Open 3 (8), http://dx.doi.org/10.1001/jamanetworkopen.2020.17533.

Brainard, J., Rushton, S., Winters, T., Hunter, P.R., 2020. Introduction to and spread of COVID-19-like illness in care homes in Norfolk, UK. J. Public Health 43 (2), 228–235. http://dx.doi.org/10.1093/pubmed/fdaa218.

Bunnik, BADv., Morgan, A.L.K., Bessell, P.R., Calder-Gerver, G., Zhang, F., Haynes, S., et al., 2021. Segmentation and shielding of the most vulnerable members of the population as elements of an exit strategy from COVID-19 lockdown. Philos. Trans. R. Soc. B 376 (1829), http://dx.doi.org/10.1098/RSTB.2020.0275.

Burton, J.K., Bayne, G., Evans, C., Garbe, F., Gorman, D., Honhold, N., et al., 2020. Evolution and effects of COVID-19 outbreaks in care homes: a population analysis in 189 care homes in one geographical region of the UK. Lancet Healthy Longev. 1 (1), E21–E31. http://dx.doi.org/10.1016/s2666-7568(20)30012-x.

Burton, J.K., McMinn, M., Vaughan, J.E., Fleuriot, J., Guthrie, B., 2021. Care-home outbreaks of COVID-19 in Scotland March to 2020: National linked data cohort analysis. Age Ageing http://dx.doi.org/10.1093/ageing/afab099.

Byambasuren, O., Cardona, M., Bell, K., Clark, J., McLaws, M.L., Glasziou, P., 2020. Estimating the extent of asymptomatic COVID-19 and its potential for community transmission: Systematic review and meta-analysis. J. Assoc. Med. Microbiol. Infect. Dis. Can. 5 (4), 223–234. http://dx.doi.org/10.3138/jammi-2020-0030.

Cabinet Secretary for Health and Sport, 2020. Coronavirus (COVID-19) social care response - 13 2020. Available from: https://www.careinspectorate.com/images/COVID-19_-_Letter_from_Cabinet_Secretary_for_Health_and_Sport_-_Social_care_guidance_-_13_March_2020.pdf.

Calvetti, D., Hoover, A.P., Rose, J., Somersalo, E., 2020. Metapopulation network models for understanding, predicting, and managing the coronavirus disease COVID-19. Front. Phys. 8, http://dx.doi.org/10.3389/fphy.2020.00261.

Comas-herrera, A., 2020. Rapid review of the evidence on impacts of visiting policies in care homes during the COVID-19 pandemic. LTCcovid. Available from: https://ltccovid.org/wp-content/uploads/2020/11/Rapid-review-of-evidence-on-impacts-of-visiting-policies-in-care-homes-during-the-COVID-pandemic-LSE068110.pdf.

Dan, J.M., Mateus, J., Kato, Y., Hastie, K.M., Yu, E.D., Faliti, C.E., et al., 2021. Immunological memory to SARS-CoV-2 assessed for up to 8 months after infection. Science 371 (6529), http://dx.doi.org/10.1126/science.abf4063.

Davies, N.G., Klepac, P., Liu, Y., Prem, K., Jit, M., Pearson, C.A.B., et al., 2020. Age-dependent effects in the transmission and control of COVID-19 epidemics. Nat. Med. 26, 1205–1211. http://dx.doi.org/10.1038/s41591-020-0962-9.

Dickson, E., Palmateer, N.E., Murray, J., Robertson, C., Waugh, C., Wallace, L.A., et al., 2021. Enhanced surveillance of COVID-19 in Scotland: population-based seroprevalence surveillance for SARS-CoV-2 during the first wave of the epidemic. Public Health 190, 132–134. http://dx.doi.org/10.1016/j.puhe.2020.11.014.

Flasch, O., 2012. Ga: Genetic algorithms. r package version 3.2. URL: https://CRAN.R-project.org/package=ga.

Graham, N.S.N., Junghans, C., Downes, R., Sendall, C., Lai, H., McKirdy, A., et al., 2020. SARS-CoV-2 infection, clinical features and outcome of COVID-19 in United Kingdom nursing homes. J. Infect. 81 (3), 411–419. http://dx.doi.org/10.1016/j.jinf.2020.05.073.

He, X., Lau, E.H.Y., Wu, P., Deng, X., Wang, J., Hao, X., et al., 2020. Temporal dynamics in viral shedding and transmissibility of COVID-19. Nat. Med. 26, 672–675. http://dx.doi.org/10.1038/s41591-020-0869-5.

Knock, E.S., Whittles, L.K., Lees, J.A., Perez-Guzman, P.N., Verity, R., FitzJohn, R.G., et al., 2021. Key epidemiological drivers and impact of interventions in the 2020 SARS-CoV-2 epidemic in England. Sci. Transl. Med. 13 (602), http://dx.doi.org/10.1126/scitranslmed.abg4262.

Ladhani, S.N., Chow, J.Y., Janarthanan, R., Fok, J., Crawley-Boevey, E., Vusirikala, A., et al., 2020. Investigation of SARS-CoV-2 outbreaks in six care homes in London, 2020. EClinicalMedicine 26, http://dx.doi.org/10.1016/j.eclinm.2020.100533.

Majra, D., Benson, J., Pitts, J., Stebbing, J., 2021. SARS-CoV-2 (COVID-19) super-spreader events. J. Infect. 82 (1), 36–40. http://dx.doi.org/10.1016/j.jinf.2020.11.021.

McAloon, C., Collins, Á., Hunt, K., Barber, A., Byrne, A.W., Butler, F., et al., 2020. Incubation period of COVID-19: A rapid systematic review and meta-analysis of observational research. BMJ Open. 10, http://dx.doi.org/10.1136/bmjopen-2020-039652.

McKeigue, P.M., Colhoun, H.M., 2020. Evaluation of stratify and shield as a policy option for ending the COVID-19 lockdown in the UK. medRxiv [Preprint]; medRxiv 2020.04.25.20079913. Available from: https://www.medrxiv.org/content/early/2020/04/30/2020.04.25.20079913.

McMichael, T.M., Currie, D.W., Clark, S., Pogosjans, S., Kay, M., Schwartz, N.G., et al., 2020. Epidemiology of Covid-19 in a long-term care facility in king county, Washington. New England. J. Med. 382, 2005–2011. http://dx.doi.org/10.1056/NEJMoa2005412.

National Records of Scotland, 2019. Population estimates time series data. Available from: https://www.nrscotland.gov.uk/statistics-and-data/statistics/statistics-by-theme/population/population-estimates/mid-year-population-estimates/population-estimates-time-series-data.

National Records of Scotland, 2020. Deaths involving coronavirus (COVID-19) in Scotland. Available from: https://www.nrscotland.gov.uk/statistics-and-data/statistics/statistics-by-theme/vital-events/general-publications/weekly-and-monthly-data-on-births-and-deaths/deaths-involving-coronavirus-covid-19-in-scotland.

Nguyen, L.K.N., Howick, S., McLafferty, D., Anderson, G.H., Pravinkumar, S.J., Van Der Meer, R., et al., 2020a. Evaluating intervention strategies in controlling COVID-19 spread in care homes: An agent-based model. Infect. Control Hosp. Epidemiol. 1–11. http://dx.doi.org/10.1017/ice.2020.1369.

Nguyen, L.K.N., Howick, S., McLafferty, D., Anderson, G.H., Pravinkumar, S.J., Van Der Meer, R., et al., 2021a. Impact of visitation and cohorting policies to shield residents from covid-19 spread in care homes: An agent-based model. Am. J. Infect. Control http://dx.doi.org/10.1016/j.ajic.2021.07.001.

Nguyen, L.K.N., Megiddo, I., 2022. Howick s hybrid simulation modelling of networks of heterogeneous care homes and the inter-facility spread of Covid-19 by sharing staff. PLoS Comput. Biol. 18 (1), http://dx.doi.org/10.1371/journal.pcbi.1009780.

Nguyen, L.K.N., Megiddo, I., Howick, S., 2020b. Challenges of infection prevention and control in Scottish long-term care facilities. Infect. Control Hosp. Epidemiol. 41 (8), 943–945. http://dx.doi.org/10.1017/ICE.2020.113.

Nguyen, L.K.N., Megiddo, I., Howick, S., 2021b. Report 3: Impact of various vaccination coverages on the spread of Covid-19 and deaths in care homes, LTCcovid. Available from: https://ltccovid.org/wp-content/uploads/2020/12/Report-3_Care-homes_Vaccination_Strathclyde.pdf.

NHS Lothian, 2021. NHS lothian: Our services. Available from: https://services.nhslothian.scot/Pages/default.aspx.

Oberhammer, J., 2020. Social-distancing effectiveness tracking of the COVID-19 hotspot stockholm. medRxiv [Preprint]; medRxiv 2020.06.30.20143487. Available from: https://www.medrxiv.org/content/10.1101/2020.06.30.20143487v2.

Office For National Statistics, 2020. Impact of Coronavirus in Care Homes in England: 26 May To 19 2020. Office For National Statistics, Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/articles/impactofcoronavirusincarehomesinenglandvivaldi/26mayto19june2020.

Overton, C.E., Stage, H.B., Ahmad, S., Curran-Sebastian, J., Dark, P., Das, R., et al., 2020. Using statistics and mathematical modelling to understand infectious disease outbreaks: COVID-19 as an example. Infect. Dis. Model. 5, 409–441. http://dx.doi.org/10.1016/j.idm.2020.06.008.

Public Health Scotland, 2020. Scottish health and social care open data: Daily case trends by health board. Available from: https://www.opendata.nhs.scot/dataset/covid-19-in-scotland/resource/2dd8534b-0a6f-4744-9253-9565d62f96c2.

Rădulescu, A., Williams, C., Cavanagh, K., 2020. Management strategies in a SEIR-type model of COVID 19 community spread. Sci. Rep. 10 (1), http://dx.doi.org/10.1038/s41598-020-77628-4.

Reilly, J., Crawford, D., Boyle, D.O., 2020. CARE HOME REVIEW: A Rapid Review of Factors Relevant To the Management of COVID-19 in the Care Home Environment in Scotland. Cabinet Secretary for Health and Sport, The Scottish Government, Available from: https://www.gov.scot/publications/root-cause-analysis-care-home-outbreaks/.

Ripperger, T.J., Uhrlaub, J.L., Watanabe, M., Wong, R., Castaneda, Y., Pizzato, H.A., et al., 2020. Orthogonal SARS-CoV-2 serological assays enable surveillance of low-prevalence communities and reveal durable humoral immunity. Immunity 53 (5), 925–933. http://dx.doi.org/10.1016/j.immuni.2020.10.004.

Roselló, A., Barnard, R.C., Smith, D.R.M., Evans, S., Grimm, F., Davies, N.G., et al., 2021. Impact of non-pharmaceutical interventions on SARS-CoV-2 outbreaks in english care homes: A modelling study. medRxiv [Preprint]; medRxiv 2021.05.17.21257315. Available from: https://www.medrxiv.org/content/early/2021/05/18/2021.05.17.21257315.

Smith, D.R.M., Duval, A., Pouwels, K.B., Guillemot, D., Fernandes, J., Huynh, B.T., et al., 2020. Optimizing COVID-19 surveillance in long-term care facilities: A modelling study. BMC Med. 18, 1–16. http://dx.doi.org/10.1186/S12916-020-01866-6.

Social Care Working Group (SCWG) for the Scientific Advisory Group for Emergencies, 2021. Commission: What are the Appropriate Layers of Mitigation To Deploy for Care Homes in the Context of Post Vaccination Risk Landscape?. UK Government, Available from: https://www.gov.uk/government/publications/scwg-what-are-the-appropriate-mitigations-to-deploy-in-care-homes-in-the-context-of-the-post-vaccination-risk-landscape-26-may-2021.

The Scottish Government, 2020. Coronavirus (COVID-19): Modelling the Epidemic in Scotland (Issue No. 29). The Scottish Government, Available from: https://www.gov.scot/publications/coronavirus-covid-19-modelling-epidemic-issue-no-29/.

World Health Organisation, 2021. WHO coronavirus (COVID-19) dashboard. Available from: https://covid19.who.int/.