SUPPLEMENTARY MATERIALS - Report 41: The 2020 SARS-CoV-2 epidemic in England: key epidemiological drivers and impact of interventions

Edward S. Knock[1]*, Lilith K. Whittles[1]*, John A. Lees[1]*, Pablo N. Perez-Guzman[1]*, Robert Verity[1], Richard G. FitzJohn[1], Katy AM Gaythorpe[1], Natsuko Imai[1], Wes Hinsley[1], Lucy C. Okell[1], Alicia Rosello[4], Nikolas Kantas[5], Caroline E. Walters[1], Sangeeta Bhatia[1], Oliver J Watson[1], Charlie Whittaker[1], Lorenzo Cattarino[1], Adhiratha Boonyasiri[3], Bimandra A. Djaafara[1], Keith Fraser[1], Han Fu[1], Haowei Wang[1], Xiaoyue Xi[5], Christl A. Donnelly[1,6], Elita Jauneikaite[1], Daniel J. Laydon[1], Peter J White[1,2], Azra C. Ghani[1], Neil M. Ferguson[1^], Anne Cori[1^], Marc Baguelin[1,4]

1. MRC Centre for Global Infectious Disease Analysis, Abdul Latif Jameel Institute for Disease and Emergency Analytics (J-IDEA), School of Public Health, Imperial College London; UK
2. National Institute for Health Research Health Protection Research Unit in Modelling and Health Economics, UK
3. Department of Infectious Disease, School of Public Health, Imperial College London; UK
4. Department of Infectious Disease Epidemiology, Faculty of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London, UK
5. Faculty of Natural Sciences, Department of Mathematics, Imperial College London, UK
6. Department of Statistics, University of Oxford, Oxford, UK

*Equal contribution, ^Equal contribution
Correspondence: m.baguelin@imperial.ac.uk, neil.ferguson@imperial.ac.uk

SUGGESTED CITATION
ES Knock, LK Whittles, JA Lees *et al*. The 2020 SARS-CoV-2 epidemic in England: key epidemiological drivers and impact of interventions - SUPPLEMENT. Imperial College London (22-12-2020), doi: https://doi.org/10.25561/85146.

Table of Contents

DOI: https://doi.org/10.25561/85146

1. Materials and Methods

Understanding the transmission of SARS-CoV-2 is challenging. The available data are subject to competing biases, such as dependence on case definition for testing and reporting, as well as being influenced by capacity and logistical constraints. These factors are further complicated by the nature of SARS-CoV-2 transmission, whereby a substantial proportion of infected individuals develop very mild symptoms, or remain asymptomatic, but are nonetheless able to infect others (1). In this section, we describe the data used in our analyses, give details on the dynamic transmission model, and present the methods used for fitting the model to the various data sources, accounting for the inherent biases in those data.

1.1 Data sources

Here we detail the datasets used to calibrate the model to the regional epidemics. We fitted our model to time series data spanning 16th March 2020 to 2nd December 2020 (inclusive), using the data available to us on 14th December 2020, by which point the effect of remaining reporting lags would be minimal.

1.1.1 Hospital admissions and bed occupancy

We use healthcare data for each NHS region from the UK Government Dashboard (supplementary data files: *data_rtm.csv*, columns: *phe_admissions, phe_occupied, phe_patients*) (2).

For admissions data, we use the daily number of confirmed COVID-19 patients admitted to hospital, which includes people admitted to hospital who tested positive for COVID-19 in the 14 days prior to admission and inpatients who tested positive in hospital after admission, with the latter being reported as admitted on the day prior to their diagnosis.

For ICU bed occupancy, we use the daily number of (confirmed) COVID-19 patients in beds which are capable of delivering mechanical ventilation.

For the occupancy in general (i.e. non-ICU) hospital beds, we use the daily number of confirmed COVID-19 patients in hospital beds with ICU occupancy subtracted.

1.1.2 Deaths

We use the number of deaths by date of death for people who had a positive COVID-19 test result and died within 28 days of their first positive test provided Public Health England. These can be found on (2). We also use the number among these deaths occurring in hospital (as reported by NHS England) and consider the remainder to have occurred in care homes. While non-hospital deaths may include deaths in other settings, such as in private residences, comparison with ONS data suggests that care home deaths from COVID-19 may also have been under-reported. As such we consider non-hospital deaths to be an appropriate proxy for care home deaths, and do not expect the margin for under or over- ascertainment to affect our conclusions. These data were provided by PHE and the data we have been using is provided as a supplementary file (supplementary data file: *data_rtm.csv*, columns: *death2, death3*) to allow reproducibility of our analysis.

1.1.3 Pillar 2 testing

We use pillar 2 testing data (see supplementary data files), which covers PCR testing for the general population (as compared with pillar 1 testing, which mainly occurred in hospitals). Since such testing was not available to the whole population for much of the spring wave of the pandemic, we only use this data from June 1st onwards.

DOI: https://doi.org/10.25561/85146

We use the daily number of positives and negative tests by specimen date. Each individual who tested positive was only counted once in the number of positives, on the specimen date of their first positive test. Multiple negatives were allowed per individual, but the negatives of all individuals who ever tested positive had been removed. We only consider PCR tests and thus exclude lateral flow tests, which have been introduced recently in trials of population mass testing. We also only use pillar 2 data for those aged 25 or over, to avoid bias resulting from increased testing of university students around the reopening of universities (supplementary data file: *data_rtm.*csv, columns: *pillar2_negatives_non_lft_over25, pillar2_positives_over25*).

1.1.4 Serology surveys

Serological survey data come from antibody testing by Public Health England of samples from healthy adult blood donors, supplied by NHS Blood and Transplant (NHSBT) (supplementary data file: *data_serology.csv*).

1.1.5 REACT-1 prevalence survey

We use the daily number of positives and negatives by specimen date from the first 7 rounds of the REACT-1 (Real-time Assessment of Community Transmission) infection prevalence survey (supplementary data file: *data_rtm.csv*, columns: *react_positive, react_samples*) (3). Note that results published in REACT preprints use data aggregated using the administrative regions of England, whereas for the purposes of this study the data has been aggregated using NHS regions. Additionally, small changes can occur in the aggregated datasets that were published in real time because of participant withdrawals and additional data cleaning.

1.1.6 Summary of the data used for calibration

Table S 1 details the datasets used to calibrate the model to the regional epidemics.

Table S 1: Data sources and definitions.

Data type	Description	Source	Reference
Hospital deaths	Daily number of COVID-19 deaths reported by NHS England within 28 days of a positive result	PHE	See data supplement. These data underlie what is released on (2)
Care home deaths	Daily number of COVID-19 deaths not reported by NHS England within 28 days of a positive result	PHE	See data supplement. These data underlie what is released on (2)
ICU occupancy	Daily number of confirmed COVID-19 patients in ICU	Gov.uk Dashboard	(2)
General bed occupancy	Daily number of confirmed COVID-19 patients in non-ICU beds	Gov.uk Dashboard	(2)
Admissions	Daily number of confirmed COVID-19 patients admitted to hospital	Gov.uk Dashboard	(2)

DOI: https://doi.org/10.25561/85146

Pillar 2 testing	Daily number of positive and negative PCR test results	PHE	See data supplement. These data underlie what is released on (2)
REACT-1 testing	Daily number of positive and negative PCR test results	REACT	(3)
Serology	Serology survey conducted on blood donors aged 15-65	PHE	See data supplement, these data are collected as part of (4)
Patient progression in hospital	Number of hospital admissions going down each treatment route (e.g. ICU, stepdown care) and length of stay in each ward.	CHESS	(5)

DOI: https://doi.org/10.25561/85146

1.1.7 Other data sources

1.1.7.1 Patient progression in hospital

The COVID-19 Hospitalisation in England Surveillance System (CHESS) data consists of a line list of daily individual patient-level data on COVID-19 infection in persons requiring hospitalisation, including demographic and clinical information on severity and outcomes. We use the individual dates of progression through hospital wards, from admission to eventual death or discharge, to produce age-stratified estimates of hospital progression parameters to be passed to the wider transmission model (see Section 1.9.2 and (supplementary data file: *support_progression.csv, support_severity.csv*).

1.1.7.2 Demographic data

We use data from the Office for National Statistics (ONS (6)) to get the number of individuals in each of the 17 age-groups, i.e. 16 five-year age bands (0-4, 5-9, …, 75-79) and an 80+ group. We get the number of care-home beds in England from (7) giving us the number of care-home beds for each NHS regions. We then got an estimate of the total population of care-home residents in the UK from (8) that we scaled down to the England population size, combined with the estimate of the total number of beds in England, we derived a value of the total occupancy of care-homes of 74.2%. We assumed that the occupancy is the same in all the NHS regions. Care-home residents are subtracted from the 4 oldest age group (5% from age 65-69, 5% age 70-74, 15% age 75-79 and 75% age 80+ (9)). We then assume a 1:1 ratio of care-home residents to care-home workers and assume that the care-home workers population is homogeneously distributed among the 25-65 population in the region.

The contact matrix between the 17 age-groups is based on the POLYMOD contact survey. See parameterisation for more details (10).

1.2 Evidence synthesis

Figure S 1 shows the functional relationships between data sources, modelled outputs and parameters in our study.

DOI: https://doi.org/10.25561/85146

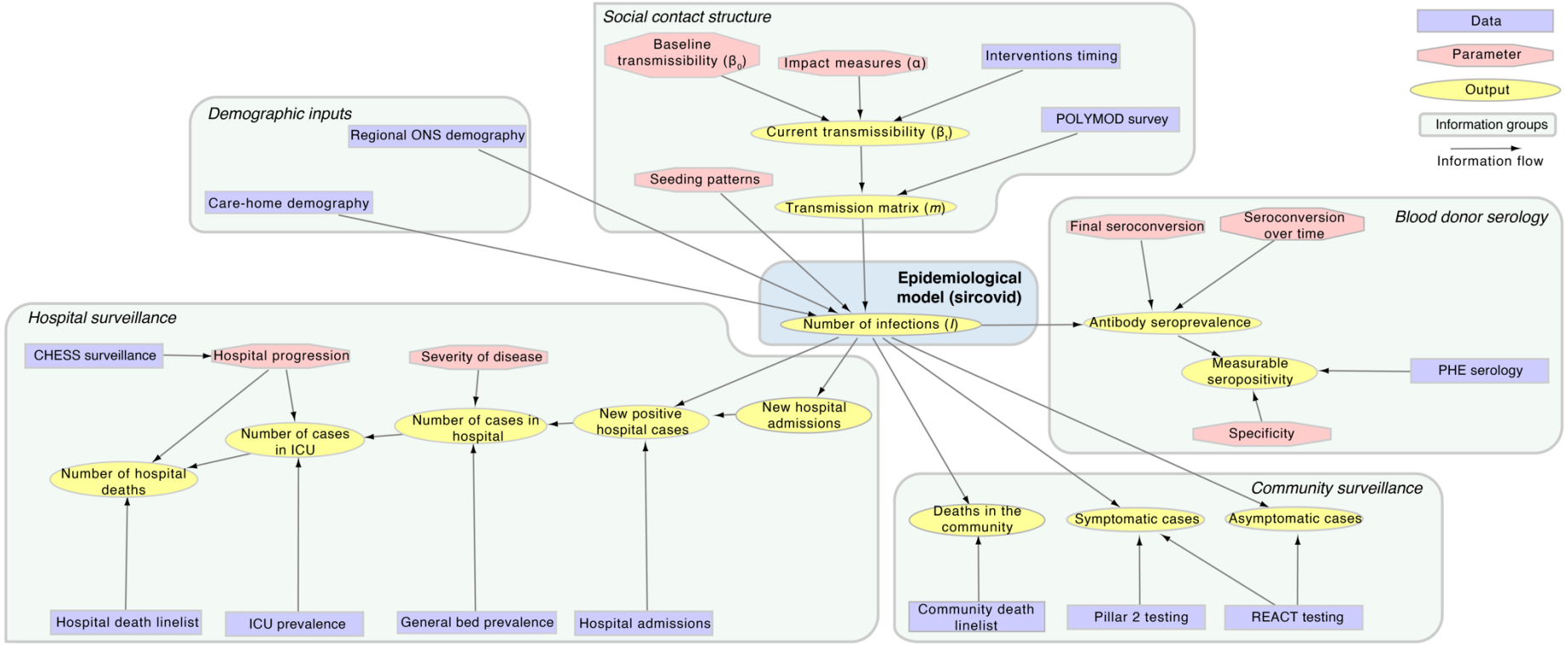

Figure S 1: Graph showing the functional relationships between data sources (rectangles), modelled outputs (ovals) and parameters (hexagons).

DOI: https://doi.org/10.25561/85146

1.3 Model description

We developed a stochastic compartmental transmission-dynamic model incorporating hospital care pathways to reconstruct the course of the COVID-19 epidemic in the seven NHS regions of England (Figure S 2). All analyses were done by regions, and then aggregated somehow if needed (e.g. for national IFR, or cumulative incidence). In the following description we do not mention any index denoting the region and thus all notations refer to the same NHS region.

1.3.1 Stratification of population into groups

We divided each regional population into 19 strata, denoted by the superscript i, 17 strata representing age groups within the general population, and two separate risk groups comprising care home workers (CHW) and care home residents (CHR). The 17 age groups consisted of 16 five-year age bands (0-4, 5-9, …, 75-79) and an 80+ group. The total size of the care home worker and resident groups were calculated assuming that 74.2% of available care home beds are occupied and there is a 1:1 carer to resident ratio (11). The care home workers were then split equally between all 8 age categories in the range 25 – 64-year-old and removed from the corresponding age categories in the general population. Despite the care-home workers being removed from all age categories in the range 25 – 64-year-old, they care-home workers are assumed to constitute one single group in our model for simplicity. The care home residents were drawn from the 65+ year old general population, such that 5% were aged 65-69, 5% aged 70-74, 15% aged 75-79 and 75% aged 80+ (9) and similarly removed from the corresponding age groups in the general population. Again, similarly to care-home workers they do constitute a single group in our model. We thus do not capture specific transmission dynamics within each care home, but rather an average mixing between residents and workers in the regional care home sector as a whole.

1.3.2 Progression of infection and hospitalisation

Prior to the importation of COVID-19, all individuals were assumed equally susceptible to infection (S). Upon infection, individuals pass through a latent period (E) before becoming infectious. A proportion (p_C) of infectious individuals develop symptoms (I_C) while the rest remain asymptomatic (I_A). All asymptomatic individuals are assumed to recover naturally. Those with symptoms may also recover naturally (R), however a proportion (p_H^i, age/care home-dependant as indicated by the *i* superscript) develop severe disease requiring hospitalisation. Of these, a proportion ($p_{G_D}^i$) die at home without receiving hospital care. In practice this proportion is set to zero except among care home residents. Of the patients who are admitted to hospital, a proportion ($p^*(t)$) have their COVID-19 diagnoses confirmed prior to admission, while the remainder may be diagnosed during their inpatient stay. All hospital compartments are divided between suspected (but not yet confirmed) and confirmed diagnoses (indicated by superscript *). A proportion ($p_{ICU}^i(t)$) of new hospital admissions are triaged (ICU_{pre}) before admission to the intensive care unit (ICU), where a fraction ($p_{ICU_D}^i(t)$) die; those who do not die get out of ICU to a ward (W) where a proportion ($p_{W_D}^i(t)$) die, while the remainder recover, following an inpatient care stepdown period. Inpatients not triaged to the ICU are assigned to general hospital beds (H), where a proportion ($p_{H_D}^i(t)$) die, while the remainder recover. Recovered individuals are assumed to be immune to reinfection for at least the duration of the simulation.

DOI: https://doi.org/10.25561/85146

In addition, there are two parallel flows which we use for fitting to testing data: (i) for PCR positivity and (ii) for seropositivity. Upon infection, an individual enters the PCR flow in a pre-positivity compartment ($T_{PCR_{pre}}$) before moving into the PCR positivity compartment ($T_{PCR_{pos}}$) and then ultimately into the PCR negativity compartment ($T_{PCR_{neg}}$). Meanwhile, individuals move into the seropositivity flow upon becoming infectious, entering first into a pre-seropositivity compartment ($T_{sero_{pre}}$). A proportion of individuals ($p_{sero_{pos}}$) then seroconvert and move into the seropositivity compartment ($T_{sero_{pos}}$), while the remainder move into the seronegativity compartment ($T_{sero_{neg}}$).

We calibrated the duration distributions for each hospital compartment, and the age-stratified probabilities of moving between compartments, using the analysis of individual-level patient data (presented below in Section 1.9.2). The required Erlang distributional form was achieved within the constraints of the modelling framework by splitting each model compartment into k sequential sub-compartments (Table S 2).

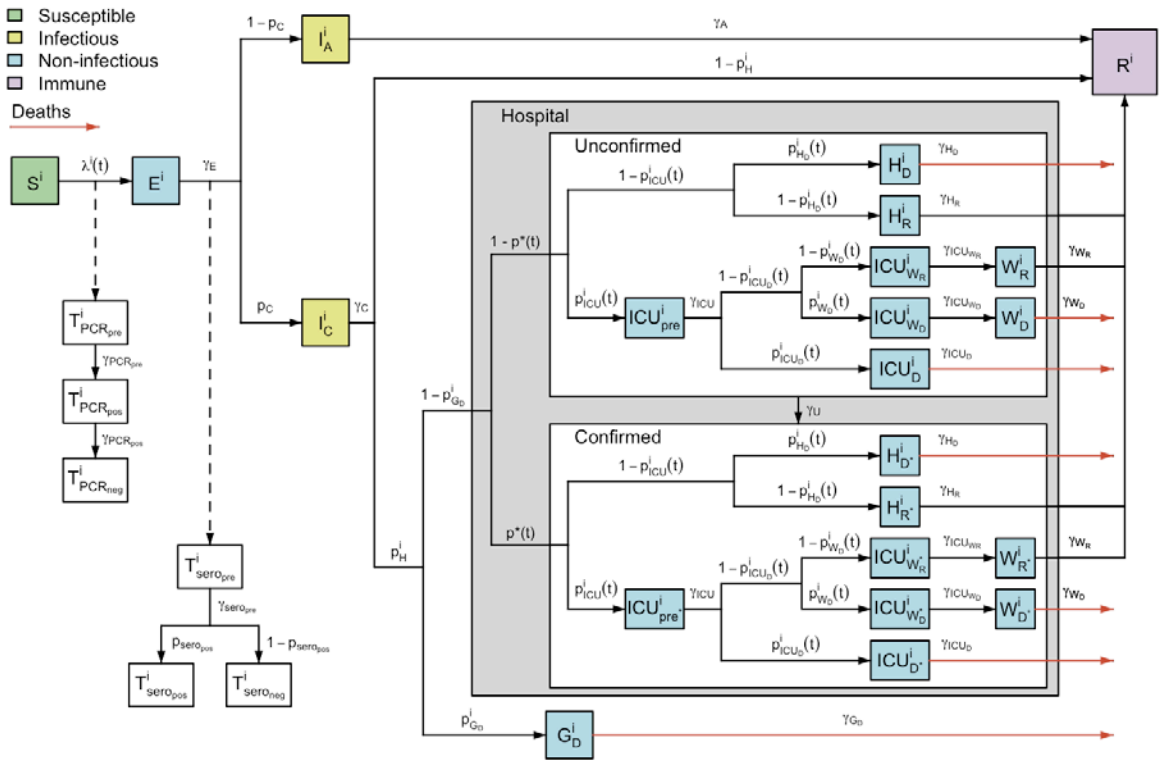

Figure S 2: Model structure flow diagram with rates of transition between infection states. Variable names defined in text.

Table S 2: Description of model compartments and distribution of time spent in each. For each named compartment, we give the associated duration. Due to the Markovian structure these are model Erlang-like distributions with k_j the number of exponential-like compartments and γ_j the rate of the exponential-like compartment. $\mathbb{E}[\tau_j]$ gives the mean duration in days spent in the corresponding compartment. The structure and duration of each stage was assumed to be the same for unconfirmed and confirmed cases in hospital (see Figure S2). For length of stays related to hospital pathways, more detail is given in section 1.9.2.

Compartment	**Description**	**Duration**		*(days)*	**Source**
		$\tau_j \sim \text{Erlang}(k_j, \gamma_j)$		$\mathbb{E}[\tau_j] = k_j/\gamma_j$	
j		k_j	γ_j	*(95% CI)*	
S	Susceptible to infection	*Determined by transmission dynamics*			
E	Latent infection	2	0.44	4.6 (0.6, 12.8)	*Lauer et al.*(12)
I_A	Asymptomatic infection	1	0.48	2.1 (0.1, 7.7)	*Bi et al.*(13)
I_C	Symptomatic infection	1	0.25	4.0 (0.1, 14.8)	*Docherty et al.*(14)
G_D	Severe illness leading to death in the general population	2	0.40	5.0	*Bernabeu-Wittel et al.* (15)
H_R	Hospitalised on general ward leading to recovery	1	0.09	10.7 (0.3, 39.4)	*Fitted to CHESS*
H_D	Hospitalised on general ward leading to death	2	0.19	10.3 (1.3, 28.8)	*Fitted to CHESS*
ICU_{pre}	Triage to ICU	1	0.40	2.5 (0.1, 9.2)	*Fitted to CHESS*
$I_{ICU_{W_R}}$	Hospitalised in ICU, leading to recovery	1	0.06	15.6 (0.4, 57.6)	*Fitted to CHESS*
$I_{ICU_{W_D}}$	Hospitalised in ICU, leading to death in step-down following ICU	1	0.14	7.0 (0.2, 25.7)	*Fitted to CHESS*
I_{ICU_D}	Hospitalised in ICU, leading to death	2	0.17	11.8 (1.4, 32.9)	*Fitted to CHESS*
W_R	Stepdown recovery period after leaving ICU	2	0.16	12.2 (1.5, 34.0)	*Fitted to CHESS*

DOI: https://doi.org/10.25561/85146

W_D	Stepdown period before death after leaving ICU	1	0.12	8.1 (0.2, 29.7)	*Fitted to CHESS*
R	Recovered	-	-	-	-
T_{pre}^{PCR}	Pre-PCR positive	1	0.33	3.0 (0.1, 11.1)	*Omar et al.* (16)
T_{pos}^{PCR}	True PCR positive	1	0.06	17.5 (0.4, 64.5)	-
T_{neg}^{PCR}	True PCR negative	-	-	-	
T_{pre}^{sero}	Pre-seroconversion	1	0.08	13.0 (0.3, 48.0)	*Benny et al.* (17)
T_{pos}^{sero}	True seropositive	-	-		-
T_{neg}^{sero}	True seronegative	-	-		-

Values of fitted parameters are set out in Table S 6.

1.3.3 Progression of infection and hospitalisation

The force of infection, $\lambda^i(t)$, for individuals in group $i \in \{[0,5), \dots, [75,80), [80+), CHW, CHR\}$ depends on time-varying social mixing between age groups and prevalence in all age/care home groups:

$$\lambda^i(t) = \sum_j m_{i,j}(t)\Pi_j(t) \quad (1)$$

where $m_{i,j}(t)$ is the (symmetric) time-varying person-to-person transmission rate from group j to group i, and $\Pi_j(t)$ is the number of infectious individuals in group *j*, given by:

$$\Pi_j(t) = I_A^j(t) + I_C^j(t) \quad (2)$$

Broadly, to parameterise $m_{i,j}(t)$, we informed mixing in the general population, and between the general population and care home workers using POLYMOD (10) via the R package *socialmixr* using age-structured regional demography (18).

Transmission between different age groups $(i, j) \in \{[0,5), \dots, [75,80), [80+)\}^2$ was parameterised as follows:

$$m_{i,j}(t) = \beta(t)c_{i,j} \quad (3)$$

Here $c_{i,j}$ is the (symmetric) person-to-person contact rate between age group i and j, derived from pre-pandemic data (10). $\beta(t)$ is the time-varying transmission rate *w*hich encompasses both changes over time in transmission efficiency (e.g. due to temperature) and temporal changes in the overall level of contacts in the population (due to changes in policy and behaviours).

We assumed $\beta(t)$ to be piecewise linear:

$$\beta(t) = \begin{cases} \beta_1, & \text{if } t \le t_1 \\ \dfrac{t_i - t}{t_i - t_{i-1}}\beta_{i-1} + \dfrac{t - t_{i-1}}{t_i - t_{i-1}}\beta_i, & \text{if } t_{i-1} < t \le t_i, \quad i = 2, \dots, 13 \\ \beta_{13}, & \text{if } t > t_{13} \end{cases} \tag{4}$$

with 12 change points $\boldsymbol{t_i}$ corresponding to major announcements and changes in COVID-19 related policy, as detailed in Table S 3.

Table S 3: Changepoints for $\beta(t)$

Changepoint	Value of $\boldsymbol{\beta(t)}$ at changepoint	Date	Description
$\boldsymbol{t_1}$	β_1	16/03/20	PM makes speech advising working from home, against non-essential travel (19)
$\boldsymbol{t_2}$	β_2	23/03/20	PM announces lockdown 1 (20)
$\boldsymbol{t_3}$	β_3	25/03/20	Lockdown 1 into full effect (21)
$\boldsymbol{t_4}$	β_4	11/05/20	Initial easing of lockdown 1 (22)
$\boldsymbol{t_5}$	β_5	15/06/20	Non-essential shops can re-open (23)
$\boldsymbol{t_6}$	β_6	04/07/20	Restaurants, pubs etc can re-open (24)
$\boldsymbol{t_7}$	β_7	03/08/20	“Eat out to help out” scheme starts (25)
$\boldsymbol{t_8}$	β_8	01/09/20	Schools and universities re-open (26)
$\boldsymbol{t_9}$	β_9	14/09/20	“Rule of six” introduced (27)
$\boldsymbol{t_{10}}$	β_{10}	14/10/20	Tiered system introduced (28)
$\boldsymbol{t_{11}}$	β_{11}	31/10/20	Lockdown 2 announced (29)
$\boldsymbol{t_{12}}$	β_{12}	05/11/20	Lockdown 2 starts (29)

DOI: https://doi.org/10.25561/85146

The contact matrix $c_{i,j}$ between different age groups $(i,j) \in \{[0,5), \dots, [75,80), [80+)\}^2$ is derived from the POLYMOD survey (10) for the United Kingdom using the *socialmixr* package (18,30), scaling by the local population demography to yield the required person-to-person daily contact rate matrix.

We defined parameters representing transmission rates within care homes (between and among workers and residents), which were assumed to be constant over time. Parameter m_{CHW} represents the person-to-person transmission rate among care home workers and between care home workers and residents; m_{CHR} represents the person-to-person transmission rate among care home residents. Hence,

$$m_{CHW,CHW}(t) = m_{CHW,CHR}(t) = m_{CHW} \quad (5)$$

$$m_{CHR,CHR}(t) = m_{CHR} \quad (6)$$

Transmission between the general population and care home workers was assumed to be similar to that within the general population, accounting for the average age of care home workers, with, for $i \in \{[0,5), \dots, [75,80), [80+)\}$,

$$m_{i,CHW}(t) = \beta(t) c_{i,CHW} \quad (7)$$

where $c_{i,CHW}$ is the mean of $c_{i,[25,30)}, c_{i,[30,35)}, \dots, c_{i,[60,65)}$ (i.e. of the age groups that the care home workers are drawn from).

Transmission between the general population and care home residents was assumed to be similar to that between the general population and the 80+ age group, adjusted by a reduction factor (ϵ, which was estimated), such that, for $i \in \{[0,5), \dots, [75,80), [80+)\}$,

$$m_{i,CHR}(t) = \epsilon \beta(t) c_{i,80+} \quad (8)$$

These represent contact between visitors from the general community and care home residents. This might involve a slightly different age profile than the age profile of the contact made by people in the 80+ age group.

1.3.4 Age-varying and time-varying infection progression probabilities

Various probabilities of clinical progression within the model are assumed to vary across age groups to account for severity of infection varying with age, and some are assumed to vary in time in order to model improvements in clinical outcomes, such as those achieved through the use of dexamethasone (31).

Two probabilities are age-varying but not time-varying, the probability of admission to hospital for symptomatic cases, and the probability of death for severe symptomatic cases in care homes. These were modelled as follows:

$$p_H^i = \psi_H^i p_H^{max} \quad (9)$$

$$p_{G_D}^i = \psi_{G_D}^i p_{G_D}^{max} \quad (10)$$

where for probability p_X^i, p_X^{max} is the maximum across all groups and ψ_X^i is the age scaling such that $\psi_X^i = 1$ for the group corresponding to the maximum, against which all other groups are scaled.

As well as varying with age, four probabilities also vary with time: the probability of admission to ICU for hospitalised cases, the probability of death in ICU, the probability of death for hospitalised cases not admitted to ICU, and the probability of death in hospital after discharge from ICU:

$$p^i_{ICU}(t) = \psi^i_{ICU} p^{max}_{ICU} h(\mu_{ICU}, t) \quad (11)$$

$$p^i_{ICU_D}(t) = \psi^i_{ICU_D} p^{max}_{ICU_D} h(\mu_D, t) \quad (12)$$

$$p^i_{H_D}(t) = \psi^i_{H_D} p^{max}_{H_D} h(\mu_D, t) \quad (13)$$

$$p^i_{W_D}(t) = \psi^i_{W_D} p^{max}_{W_D} h(\mu_D, t) \quad (14)$$

where here for probability p^i_X, p^{max}_Xgives the maximum *initial* value across groups and $h(\mu, t) = 1$ before April 1st, $h(\mu, t) = \mu < 1$ after June 1st, with a linear reduction in between.

Care home residents with severe disease leading to death are assumed to remain in compartment $\boldsymbol{G_D}$ for 5 days on average before dying (modelled with $k_{\boldsymbol{G_D}} = 2$ and $\gamma_{\boldsymbol{G_D}} = 0.4$), 95% range 0.6-13.9 days broadly consistent with durations in (15) and with duration about half the length observed in hospital streams (see Figure S *5*).

For care home workers, the age scaling ψ^{CHW}_X is taken as the mean of the age scalings ψ^i_X for $i \in \{[25,30), [30,35), \dots, [60,65)\}$. For care home residents, we assume that $\psi^{CHR}_X = \psi^{[80+)}_X$, with the exception of the probability of individual with severe disease requiring hospitalisation dying at home (without receiving hospital care), where we assume $\psi^{CHR}_{G_D} = 1$ and $\psi^i_{G_D} = 0$ for all other groups, to effectively allow death outside hospital only for care home residents.

1.4 Reproduction number $\boldsymbol{R_t}$ and effective reproduction number $\boldsymbol{R^{eff}_t}$

We calculated the reproduction number over time, R_t, and effective reproduction number over time, R^{eff}_t, using next generation matrix methods (32). The reproduction numbers are calculated for the general population, i.e. excluding care home workers and residents. We define R_t as the average number of secondary infections a case infected at time t would generate in a large entirely susceptible population, and R^{eff}_t as the average number of secondary infections generated by a case infected at time t would accounting for the finite population size and potential immunity in the population.

To compute the next generation matrix, we calculated the mean duration of infectiousness Δ_I, as

$$\Delta_I = (1 - p_C)\mathbb{E}[\tau_{I_A}] + p_C\mathbb{E}[\tau_{I_C}] \quad (15)$$

where parameter and model compartment notations are defined in Table S ***2*** - Table S 8.

For this calculation, the expected durations of stay in compartments were adjusted to account for the discrete-time nature of the model, via calculating the expected number of time-steps (of length dt) spent in a given compartment. Note that if in continuous-time a compartment duration is $\tau \sim \text{Erlang}\,(k, \gamma)$, then the corresponding discrete-time mean duration is:

$$E[\tau] = \frac{k\,dt}{(1 - e^{-\gamma dt})} \quad (16)$$

The next generation matrix was calculated as, for $(i, j) \in \{[0,5), \ldots, [75,80), [80+)\}^2$,

$$\text{NGM}_{ij}(t) = m_{ij}(t)\Delta_I N^i \quad (17)$$

where N^i is the total population of group i and R_t is taken to be the dominant eigenvalue of NGM(t), while the effective next generation matrix was calculated as:

$$\text{NGM}_{ij}^{eff}(t) = m_{ij}(t)\Delta_I S^i(t) \quad (18)$$

with R_t^{eff} taken to be the dominant eigenvalue of $\text{NGM}^{eff}(t)$.

1.5 Infection severity

Posterior estimates of severity, namely the infection hospitalisation and infection fatality ratios, were calculated in each group i as follows:

$$\text{IHR}^i = p_C p_H^i (1 - p_{G_D}^i) \quad (19)$$

$$\text{IFR}^i(t) = p_C p_H^i \left\{ p_{G_D}^i + \left(1 - p_{G_D}^i\right) \left(p_{ICU}^i(t) \left(p_{ICU_D}^i(t) + \left(1 - p_{ICU_D}^i(t)\right) p_{W_D}^i(t) \right) + \left(1 - p_{ICU}^i(t)\right) p_{H_D}^i(t) \right) \right\} \quad (20)$$

Note that for simplicity the notation we use do refer explicitly to the NHS region of interest. We calculated age-aggregated estimates for each region by weighting the age-specific severity estimates by the cumulative incidence in that age group. Aggregate estimates for England were then calculated by weighting the region-specific estimates by the regional attack rates.

1.6 Compartmental model equations

To clearly illustrate the model dynamics, we describe a deterministic version of the model in differential equations (21)-(56), followed by the stochastic implementation used in the analysis. Each compartment is stratified by mixing category $i \in \{[0,5), \dots, [75,80), [80+), CHW, CHR\}$. Full definitions of compartments and model parameters are set out in Table S 2 - Table S 8.

$$dS^i(t)/dt = -\lambda^i(t)S^i(t) \quad (21)$$

$$dE^{i,1}(t)/dt = \lambda^i(t)S^i(t) - \gamma_E E^{i,1}(t) \quad (22)$$

$$dE^{i,2}(t)/dt = \gamma_E E^{i,1}(t) - \gamma_E E^{i,2}(t) \quad (23)$$

$$dI_A^i(t)/dt = (1-p_C)\gamma_E E^{i,2}(t) - \gamma_A I_A^i(t) \quad (24)$$

$$dI_C^i(t)/dt = p_C\gamma_E E^{i,2}(t) - \gamma_C I_C^i(t) \quad (25)$$

$$dG_D^{i,1}(t)/dt = p_H^i p_{G_D}^i \gamma_C I_C^i(t) - \gamma_{G_D} G_D^{i,1}(t) \quad (26)$$

$$dG_D^{i,2}(t)/dt = \gamma_{G_D} G_D^{i,1}(t) - \gamma_{G_D} G_D^{i,2}(t) \quad (27)$$

$$dICU_{pre}^i(t)/dt = p_H^i\left(1-p_{G_D}^i\right)\left(1-p^*(t)\right) p_{ICU}^i(t)\,\gamma_C I_C^i(t) - \left(\gamma_{ICU_{pre}} + \gamma_U\right) ICU_{pre}^i(t) \quad (28)$$

$$dICU_{pre^*}^i(t)/dt = p_H^i\left(1-p_{G_D}^i\right)p^*(t)\, p_{ICU}^i(t)\gamma_C I_C^i(t) - \gamma_{ICU_{pre}} ICU_{pre^*}^i(t) + \gamma_U ICU_{pre}^i(t) \quad (29)$$

$$dICU_{W_R}^i(t)/dt = \left(1-p_{ICU_D}^i(t)\right)\left(1-p_{W_D}^i(t)\right)\gamma_{ICU_{pre}} ICU_{pre}^i(t) - \left(\gamma_{ICU_{W_R}} + \gamma_U\right) ICU_{W_R}^i(t) \quad (30)$$

$$dICU_{W_{R^*}}^i(t)/dt = \left(1-p_{ICU_D}^i(t)\right)\left(1-p_{W_D}^i(t)\right)\gamma_{ICU_{pre}} ICU_{pre^*}^i(t) - \gamma_{ICU_{W_R}} ICU_{W_{R^*}}^i(t) + \gamma_U ICU_{W_R}^i(t) \quad (31)$$

$$dICU_{W_D}^i(t)/dt = \left(1-p_{ICU_D}^i(t)\right)p_{W_D}^i(t)\gamma_{ICU_{pre}} ICU_{pre}^i(t) - \left(\gamma_{ICU_{W_D}} + \gamma_U\right) ICU_{W_D}^i(t) \quad (32)$$

$$dICU_{W_{D^*}}^i(t)/dt = \left(1-p_{ICU_D}^i(t)\right)p_{W_D}^i(t)\gamma_{ICU_{pre}} ICU_{pre^*}^i(t) - \gamma_{ICU_{W_D}} ICU_{W_{D^*}}^i(t) + \gamma_U ICU_{W_D}^i(t) \quad (33)$$

$$dICU_D^{i,1}(t)/dt = p_{ICU_D}^i(t)\gamma_{ICU_{pre}} ICU_{pre}^i(t) - \left(\gamma_{ICU_D} + \gamma_U\right) ICU_D^{i,1}(t) \quad (34)$$

$$dICU_D^{i,2}(t)/dt = \gamma_{ICU_D} ICU_D^{i,1}(t) - \left(\gamma_{ICU_D} + \gamma_U\right) ICU_D^{i,2}(t) \quad (35)$$

$$dICU_{D^*}^{i,1}(t)/dt = p_{ICU_D}^i(t)\gamma_{ICU_{pre}} ICU_{pre^*}^i(t) - \gamma_{ICU_D} ICU_{D^*}^{i,1}(t) + \gamma_U ICU_D^{i,1}(t) \quad (36)$$

$$dICU_{D^*}^{i,2}(t)/dt = \gamma_{ICU_D} ICU_{D^*}^{i,1}(t) - \gamma_{ICU_D} ICU_{D^*}^{i,2}(t) + \gamma_U ICU_D^{i,2}(t) \quad (37)$$

$$dW_R^{i,1}(t)/dt = \gamma_{ICU_{W_R}} ICU_{W_R}^i(t) - \left(\gamma_{W_R} + \gamma_U\right) W_R^{i,1}(t) \quad (38)$$

$$dW_R^{i,2}(t)/dt = \gamma_{W_R} W_R^{i,1}(t) - \left(\gamma_{W_R} + \gamma_U\right) W_R^{i,2}(t) \quad (39)$$

$$dW_{R^*}^{i,1}(t)/dt = \gamma_{ICU_{W_R}} ICU_{W_{R^*}}^i(t) - \gamma_{W_R} W_{R^*}^{i,1}(t) + \gamma_U W_R^{i,1}(t) \quad (40)$$

$$dW_{R^*}^{i,2}(t)/dt = \gamma_{W_R} W_{R^*}^{i,1}(t) - \gamma_{W_R} W_{R^*}^{i,2}(t) + \gamma_U W_R^{i,2}(t) \quad (41)$$

$$dW_D^i(t)/dt = \gamma_{ICU_{W_D}} ICU_{W_D}^i(t) - \left(\gamma_{W_D} + \gamma_U\right) W_D^i(t) \quad (42)$$

$$dW^i_{D^*}(t)/dt = \gamma_{ICU_{W_D}} ICU^i_{W_{D^*}}(t) - \gamma_{W_D} W^i_{D^*}(t) + \gamma_U W^i_D(t) \quad (43)$$

$$dH^i_R(t)/dt = p^i_H\left(1-p^i_{G_D}\right)\left(1-p^*(t)\right)\left(1-p^i_{ICU}(t)\right)\left(1-p^i_{H_D}(t)\right)\gamma_C I^i_C(t) - \left(\gamma_{H_R}+\gamma_U\right)H^i_R(t) \quad (44)$$

$$dH^i_{R^*}(t)/dt = p^i_H\left(1-p^i_{G_D}\right)p^*(t)\left(1-p^i_{ICU}(t)\right)\left(1-p^i_{H_D}(t)\right)\gamma_C I^i_C(t) + \gamma_U H^i_R(t) - \gamma_{H_R} H^i_{R^*}(t) \quad (45)$$

$$dH^{i,1}_D(t)/dt = p^i_H\left(1-p^i_{G_D}\right)\left(1-p^*(t)\right)\left(1-p^i_{ICU}(t)\right)p^i_{H_D}(t)\gamma_C I^i_C(t) - \left(\gamma_{H_D}+\gamma_U\right)H^{i,1}_D(t) \quad (46)$$

$$dH^{i,2}_D(t)/dt = \gamma_{H_D} H^{i,1}_D(t) - \left(\gamma_{H_D}+\gamma_U\right)H^{i,2}_D(t) \quad (47)$$

$$dH^{i,1}_{D^*}(t)/dt = p^i_H\left(1-p^i_{G_D}\right)p^*(t)\left(1-p^i_{ICU}(t)\right)p^i_{H_D}(t)\,\gamma_C I^i_C(t) + \gamma_U H^{i,1}_D(t) - \gamma_{H_D} H^{i,1}_{D^*}(t) \quad (48)$$

$$dH^{i,2}_{D^*}(t)/dt = \gamma_{H_D} H^{i,1}_{D^*}(t) - \gamma_{H_D} H^{i,2}_{D^*}(t) + \gamma_U H^{i,2}_D(t) \quad (49)$$

$$dR^i(t)/dt = \gamma_A I^i_A(t) + \left(1-p^i_H\right)\gamma_C I^i_C(t) + \gamma_{H_R}\left(H^i_R(t)+H^i_{R^*}(t)\right) + \gamma_{W_R}\left(W^i_R(t)+W^i_{R^*}(t)\right) \quad (50)$$

$$dT^i_{sero_{pre}}(t)/dt = \gamma_E E^{i,2}(t) - \gamma_{sero_{pre}} T^i_{sero_{pre}}(t) \quad (51)$$

$$dT^i_{sero_{pos}}(t)/dt = p_{sero_{pos}}\gamma_{sero_{pre}} T^i_{PCR_{pre}}(t) \quad (52)$$

$$dT^i_{sero_{neg}}(t)/dt = \left(1-p_{sero_{pos}}\right)\gamma_{sero_{pre}} T^i_{PCR_{pre}}(t) \quad (53)$$

$$dT^i_{PCR_{pre}}(t)/dt = \lambda^i(t) S^i(t) - \gamma_{PCR_{pre}} T^i_{PCR_{pre}}(t) \quad (54)$$

$$dT^i_{PCR_{pos}}(t)/dt = \gamma_{PCR_{pre}} T^i_{PCR_{pre}}(t) - \gamma_{PCR_{pos}} T^i_{PCR_{pos}}(t) \quad (55)$$

$$dT^i_{PCR_{neg}}(t)/dt = \gamma_{PCR_{pos}} T^i_{PCR_{pos}}(t) \quad (56)$$

We used the tau-leap method (33) to create a stochastic, time-discretised version of the model described in equations (60-(164), taking four update steps per day. The process was initialised with ten asymptomatic infectious individuals aged 15-19 on the epidemic start date t_0, a parameter we estimate. For each time step, the model iterated through the procedure described below. In the following, we introduce a small abuse of notation: for transitions involving multiple onward compartments (e.g transition from compartment E to compartments I_A or I_C), for conciseness, we write

$$\left(d^i_{E,I_A},\ d^i_{E,I_C}\right) \sim \text{Multinom}\left(E^{i,2}(t),\ q^i_{E,I_A},\ q^i_{E,I_C}\right)$$

instead of

$$\left(d^i_{E,I_A},\ d^i_{E,I_C}, d^i_{nomove}\right) \sim \text{Multinom}\left(E^{i,2}(t),\ q^i_{E,I_A},\ q^i_{E,I_C},\ 1-\textstyle\sum_{x\in\{I_A,I_C\}} q^i_{E,x}\right)$$

where d^i_{nomove} is a dummy variable counting the number of individuals remaining in compartment $E^{i,2}$.

Using this convention, transition variables are drawn from the following distributions, with probabilities defined below:

$$d^i_{S,E} \sim \text{Binom}\left(S^i(t),\ 1-e^{-\lambda^i(t)dt}\right) \quad (57)$$

$$d^i_{E,E} \sim \text{Binom}\left(E^{i,1}(t),\ 1-e^{-\gamma_E dt}\right) \quad (58)$$

$$(q^i_{E,I_A},\ q^i_{E,I_C}) = \left((1-p_C)\left(1-e^{-\gamma_E dt}\right), p_C\left(1-e^{-\gamma_E dt}\right)\right) \quad (59)$$

$$\left(d^i_{E,I_A},\ d^i_{E,I_C}\right) \sim \text{Multinom}\left(E^{i,2}(t),\ q^i_{E,I_A},\ q^i_{E,I_C}\ \right) \quad (60)$$

$$d^i_{I_A,R} \sim \text{Binom}\left(I^i_A(t), 1-e^{-\gamma_A dt}\right) \quad (61)$$

$$q^i_{I_C,G_D} = p^i_H p^i_{G_D}\left(1-e^{-\gamma_C dt}\right) \quad (62)$$

$$q^i_{I_C,R} = \left(1-p^i_H\right)\left(1-e^{-\gamma_C dt}\right) \quad (63)$$

$$q^i_{I_C,ICU_{pre}} = p^i_H\left(1-p^i_{G_D}\right)\left(1-p^*(t)\right)p^i_{ICU}(t)\left(1-e^{-\gamma_C dt}\right) \quad (64)$$

$$q^i_{I_C,ICU_{pre^*}} = p^i_H\left(1-p^i_{G_D}\right)p^*(t)\,p^i_{ICU}(t)\left(1-e^{-\gamma_C dt}\right) \quad (65)$$

$$q^i_{I_C,H_R} = p^i_H\left(1-p^i_{G_D}\right)\left(1-p^*(t)\right)\left(1-p^i_{ICU}(t)\right)\left(1-p^i_{H_D}(t)\right)\left(1-e^{-\gamma_C dt}\right) \quad (66)$$

$$q^i_{I_C,H^*_R} = p^i_H\left(1-p^i_{G_D}\right)p^*(t)\left(1-p^i_{ICU}(t)\right)\left(1-p^i_{H_D}(t)\right)\left(1-e^{-\gamma_C dt}\right) \quad (67)$$

$$q^i_{I_C,H_D} = p^i_H\left(1-p^i_{G_D}\right)\left(1-p^*(t)\right)\left(1-p^i_{ICU}(t)\right)p^i_{H_D}(t)\left(1-e^{-\gamma_C dt}\right) \quad (68)$$

$$q^i_{I_C,H_{D^*}} = p^i_H\left(1-p^i_{G_D}\right)p^*(t)\left(1-p^i_{ICU}(t)\right)p^i_{H_D}(t)\left(1-e^{-\gamma_C dt}\right) \quad (69)$$

$$\left(d^i_{I_C,G_D},\dots,d^i_{I_C,H_{D^*}}\right) \sim \text{Multinom}\left(I^i_C(t), q^i_{I_C,G_D},\dots,q^i_{I_C,H_{D^*}}\right) \quad (70)$$

$$d^i_{G_D,G_D} \sim \text{Binom}\left(G^{i,1}_D(t),\ 1-e^{-\gamma_{G_D} dt}\right) \quad (71)$$

$$d^i_{G_D,D} \sim \text{Binom}\left(G^{i,2}_D(t),\ 1-e^{-\gamma_{G_D} dt}\right) \quad (72)$$

$$q^i_{ICU_{pre},ICU_{W_R}} = \left(1-p^i_{ICU_D}(t)\right)\left(1-p^i_{W_D}(t)\right)\left(1-e^{-\gamma_{ICU_{pre}} dt}\right)e^{-\gamma_U dt} \quad (73)$$

$$q^i_{ICU_{pre},ICU_{W_{R^*}}} = \left(1-p^i_{ICU_D}(t)\right)\left(1-p^i_{W_D}(t)\right)\left(1-e^{-\gamma_{ICU_{pre}} dt}\right)\left(1-e^{-\gamma_U dt}\right) \quad (74)$$

$$q^i_{ICU_{pre},ICU_{W_D}} = \left(1-p^i_{ICU_D}(t)\right)p^i_{W_D}(t)\left(1-e^{-\gamma_{ICU_{pre}} dt}\right)e^{-\gamma_U dt} \quad (75)$$

$$q^i_{ICU_{pre},ICU_{W_{D^*}}} = \left(1-p^i_{ICU_D}(t)\right)p^i_{W_D}(t)\left(1-e^{-\gamma_{ICU_{pre}} dt}\right)\left(1-e^{-\gamma_U dt}\right) \quad (76)$$

$$q^i_{ICU_{pre},ICU_D} = p^i_{ICU_D}(t)\left(1-e^{-\gamma_{ICU_{pre}} dt}\right)e^{-\gamma_U dt} \quad (77)$$

$$q^i_{ICU_{pre},ICU_{D^*}} = p^i_{ICU_D}(t)\left(1-e^{-\gamma_{ICU_{pre}} dt}\right)\left(1-e^{-\gamma_U dt}\right) \quad (78)$$

$$q^i_{ICU_{pre},ICU_{pre^*}} = e^{-\gamma_{ICU_{pre}} dt}\left(1-e^{-\gamma_U dt}\right) \quad (79)$$

$$\left(d^i_{ICU_{pre},ICU_{W_R}},\dots,d^i_{ICU_{pre},ICU^*_{pre}}\right) \sim \text{Multinom}\left(ICU^i_{pre}(t), q^i_{ICU_{pre},ICU_{W_R}},\dots,q^i_{ICU_{pre},ICU^*_{pre}}\right) \quad (80)$$

$$q^i_{ICU^*_{pre},ICU_{W_{R^*}}} = \left(1-p^i_{ICU_D}(t)\right)\left(1-p^i_{W_D}(t)\right)\left(1-e^{-\gamma_{ICU_{pre}} dt}\right) \quad (81)$$

$$q^i_{ICU^*_{pre},ICU_{W_{D^*}}} = \left(1-p^i_{ICU_D}(t)\right)p^i_{W_D}(t)\left(1-e^{-\gamma_{ICU_{pre}} dt}\right) \quad (82)$$

$$q^i_{ICU^*_{pre},ICU_{D^*}} = p^i_{ICU_D}(t)\left(1-e^{-\gamma_{ICU_{pre}} dt}\right) \quad (83)$$

$$\left(d^i_{ICU^*_{pre},ICU_{W_{R^*}}},\dots,d^i_{ICU^*_{pre},ICU_{D^*}}\right) \sim \text{Multinom}\left(ICU^i_{pre^*}(t), q^i_{ICU^*_{pre},ICU_{W_{R^*}}},\dots,q^i_{ICU^*_{pre},ICU_{D^*}}\right) \quad (84)$$

$$q^{i}_{H_D,H_D} = \left(1-e^{-\gamma_{H_D}dt}\right)e^{-\gamma_U dt} \quad (85)$$

$$q^{i,1,1}_{H_D,H_{D^*}} = e^{-\gamma_{H_D}dt}\left(1-e^{-\gamma_U dt}\right) \quad (86)$$

$$q^{i,1,2}_{H_D,H_{D^*}} = \left(1-e^{-\gamma_{H_D}dt}\right)\left(1-e^{-\gamma_U dt}\right) \quad (87)$$

$$\left(d^{i}_{H_D,H_D}, d^{i,1,1}_{H_D,H_{D^*}}, d^{i,1,2}_{H_D,H_{D^*}}\right) \sim \text{Multinom}\left(H^{i,1}_{D}(t), q^{i}_{H_D,H_D}, q^{i,1,1}_{H_D,H_{D^*}}, q^{i,1,2}_{H_D,H_{D^*}}\right) \quad (88)$$

$$d^{i}_{H_{D^*},H_{D^*}} \sim \text{Binom}\left(H^{i,1}_{D^*}(t), 1-e^{-\gamma_{H_D}dt}\right) \quad (89)$$

$$\left(d^{i}_{H_D,D}, d^{i,2,2}_{H_D,H_{D^*}}\right) \sim \text{Multinom}\left(H^{i,2}_{D}(t),\ 1-e^{-\gamma_{H_D}dt}, e^{-\gamma_{H_D}dt}\left(1-e^{-\gamma_U dt}\right)\right) \quad (90)$$

$$d^{i}_{H_{D^*},D} \sim \text{Binom}\left(H^{i,2}_{D^*}(t),\ 1-e^{-\gamma_{H_D}dt}\right) \quad (91)$$

$$\left(d^{i}_{H_R,R}, d^{i}_{H_R,H_{R^*}}\right) \sim \text{Multinom}\left(H^{i}_{R}(t),\ 1-e^{-\gamma_{H_R}dt}, e^{-\gamma_{H_R}dt}\left(1-e^{-\gamma_U dt}\right)\right) \quad (92)$$

$$d^{i}_{H_{R^*},R} \sim \text{Binom}\left(H^{i}_{R^*}(t),\ 1-e^{-\gamma_{H_R}dt}\right) \quad (93)$$

$$q^{i}_{ICU_{W_R},W_R} = \left(1-e^{-\gamma_{ICU_{W_R}}dt}\right)e^{-\gamma_U dt} \quad (94)$$

$$q^{i}_{ICU_{W_R},ICU_{W_{R^*}}} = e^{-\gamma_{ICU_{W_R}}dt}\left(1-e^{-\gamma_U dt}\right) \quad (95)$$

$$q^{i}_{ICU_{W_R},W_{R^*}} = \left(1-e^{-\gamma_{ICU_{W_R}}dt}\right)\left(1-e^{-\gamma_U dt}\right) \quad (96)$$

$$\left(d^{i}_{ICU_{W_R},W_R}, \dots, d^{i}_{ICU_{W_R},W_{R^*}}\right) \sim \text{Multinom}\left(ICU^{i}_{W_R}(t), q^{i}_{ICU_{W_R},W_R}, \dots, q^{i}_{ICU_{W_R},W_{R^*}}\right) \quad (97)$$

$$d^{i}_{ICU_{W_{R^*}},W_{R^*}} \sim \text{Binom}\left(ICU^{i}_{W_{R^*}}(t),\ 1-e^{-\gamma_{ICU_{W_R}}dt}\right) \quad (98)$$

$$q^{i}_{ICU_{W_D},W_D} = \left(1-e^{-\gamma_{ICU_{W_D}}dt}\right)e^{-\gamma_U dt} \quad (99)$$

$$q^{i}_{ICU_{W_D},ICU_{W_{D^*}}} = e^{-\gamma_{ICU_{W_D}}dt}\left(1-e^{-\gamma_U dt}\right) \quad (100)$$

$$q^{i}_{ICU_{W_D},W_{D^*}} = \left(1-e^{-\gamma_{ICU_{W_D}}dt}\right)\left(1-e^{-\gamma_U dt}\right) \quad (101)$$

$$\left(d^{i}_{ICU_{W_D},W_D}, \dots, d^{i}_{ICU_{W_D},W_{D^*}}\right) \sim \text{Multinom}\left(ICU^{i}_{W_D}(t), q^{i}_{ICU_{W_D},W_D}, \dots, q^{i}_{ICU_{W_D},W_{D^*}}\right) \quad (102)$$

$$d^{i}_{ICU_{W_{D^*}},W_{D^*}} \sim \text{Binom}\left(ICU^{i}_{W_{D^*}}(t),\ 1-e^{-\gamma_{ICU_{W_D}}dt}\right) \quad (103)$$

$$q^{i}_{ICU_D,ICU_D} = \left(1-e^{-\gamma_{ICU_D}dt}\right)e^{-\gamma_U dt} \quad (104)$$

$$q^{i,1,1}_{ICU_D,ICU_{D^*}} = e^{-\gamma_{ICU_D}dt}\left(1-e^{-\gamma_U dt}\right) \quad (105)$$

$$q^{i,1,2}_{ICU_D,ICU_{D^*}} = \left(1-e^{-\gamma_{ICU_D}dt}\right)\left(1-e^{-\gamma_U dt}\right) \quad (106)$$

$$\left(d^{i}_{ICU_D,ICU_D}, d^{i,1,1}_{ICU_D,ICU_{D^*}}, d^{i,1,2}_{ICU_D,ICU}\right. \sim \text{Multinom}\left(ICU^{i,1}_{D}(t), q^{i}_{ICU_D,ICU_D}, q^{i,1,1}_{ICU_D,ICU_{D^*}}, q^{i,1,2}_{ICU_D,ICU_{D^*}}\right) \quad (107)$$

$$d^{i}_{ICU_{D^*},ICU_{D^*}} \sim \text{Binom}\left(ICU^{i,1}_{D^*}(t),\ 1-e^{-\gamma_{ICU_D}dt}\right) \quad (108)$$

$$\left(d^{i}_{ICU_D,D}, d^{i,2,2}_{ICU_D,ICU_{D^*}}\right) \sim \text{Multinom}\left(ICU^{i,2}_{D}(t),\ 1-e^{-\gamma_{ICU_D}dt}, e^{-\gamma_{ICU_D}dt}\left(1-e^{-\gamma_U dt}\right)\right) \quad (109)$$

$$d^{i}_{ICU_{D^*},D} \sim \text{Binom}\left(ICU^{i,2}_{D^*}(t),\ 1-e^{-\gamma_{ICU_D}dt}\right) \quad (110)$$

$$q^i_{W_R,W_R} = \left(1 - e^{-\gamma_{W_R} dt}\right) e^{-\gamma_U dt} \quad (111)$$

$$q^{i,1,1}_{W_R,W_{R^*}} = e^{-\gamma_{W_R} dt}\left(1 - e^{-\gamma_U dt}\right) \quad (112)$$

$$q^{i,1,2}_{W_R,W_{R^*}} = \left(1 - e^{-\gamma_{W_R} dt}\right)\left(1 - e^{-\gamma_U dt}\right) \quad (113)$$

$$\left(d^i_{W_R,W_R}, d^{i,1,1}_{W_R,W_{R^*}}, d^{i,1,2}_{W_R,W_{R^*}}\right) \sim \text{Multinom}\left(W^{i,1}_R(t), q^i_{W_R,W_R}, q^{i,1,1}_{W_R,W_{R^*}}, q^{i,1,2}_{W_R,W_{R^*}}\right) \quad (114)$$

$$d^i_{W_{R^*},W_{R^*}} \sim \text{Binom}\left(W^{i,1}_{R^*}(t),\ 1 - e^{-\gamma_{W_R} dt}\right) \quad (115)$$

$$\left(d^i_{W_R,R}, d^{i,2,2}_{W_R,W_{R^*}}\right) \sim \text{Multinom}\left(W^{i,2}_R(t),\ 1 - e^{-\gamma_{W_R} dt}, e^{-\gamma_{W_R} dt}\left(1 - e^{-\gamma_U dt}\right)\right) \quad (116)$$

$$d^i_{W_{R^*},R} \sim \text{Binom}\left(W^{i,2}_{R^*}(t),\ 1 - e^{-\gamma_{W_R} dt}\right) \quad (117)$$

$$\left(d^i_{W_D,D}, d^i_{W_D,W_{D^*}}\right) \sim \text{Multinom}\left(W^i_D(t),\ 1 - e^{-\gamma_{W_D} dt}, e^{-\gamma_{W_D} dt}\left(1 - e^{-\gamma_U dt}\right)\right) \quad (118)$$

$$d^i_{W_{D^*},D} \sim \text{Binom}\left(W^i_{D^*}(t),\ 1 - e^{-\gamma_{W_D} dt}\right) \quad (119)$$

$$\left(d^i_{W_D,D}, d^i_{W_D,W_{D^*}}\right) \sim \text{Multinom}\left(W^i_D(t),\quad 1 - e^{-\gamma_{W_D} dt}, e^{-\gamma_{W_D} dt}\left(1 - e^{-\gamma_U dt}\right)\right) \quad (120)$$

$$q^i_{T_{sero_{pre}},T_{sero_{pos}}} = p_{sero_{pos}}\left(1 - e^{-\gamma_{sero_{pre}} dt}\right) \quad (121)$$

$$q^i_{T_{sero_{pre}},T_{sero_{neg}}} = \left(1 - p_{sero_{pos}}\right)\left(1 - e^{-\gamma_{sero_{pre}} dt}\right) \quad (122)$$

$$\left(d^i_{T_{sero_{pre}},T_{sero_{pos}}}, d^i_{T_{sero_{pre}},T_{sero_{neg}}}\right) \sim \text{Multinom}\left(T^i_{sero_{pre}}(t), q^i_{T_{sero_{pre}},T_{sero_{pos}}}, q^i_{T_{sero_{pre}},T_{sero_{neg}}}\right) \quad (123)$$

$$d^i_{T_{PCR_{pre}},T_{PCR_{pos}}} \sim \text{Binom}\left(T^i_{PCR_{pre}}(t),\ 1 - e^{-\gamma_{PCR_{pre}} dt}\right) \quad (124)$$

$$d^i_{T_{PCR_{pos}},T_{PCR_{neg}}} \sim \text{Binom}\left(T^i_{PCR_{pos}}(t),\ 1 - e^{-\gamma_{PCR_{pos}} dt}\right) \quad (125)$$

Model compartments were then updated as follows:

$$S^i(t + dt) \coloneqq S^i(t) - d^i_{S,E} \quad (126)$$

$$E^{i,1}(t + dt) \coloneqq E^{i,1}(t) + d^i_{S,E} - d^i_{E,E} \quad (127)$$

$$E^{i,2}(t + dt) \coloneqq E^{i,2}(t) + d^i_{E,E} - d^i_{E,I_A} - d^i_{E,I_C} \quad (128)$$

$$I^i_A(t + dt) \coloneqq I^i_A(t) + d^i_{E,I_A} - d^i_{I_A,R} \quad (129)$$

$$I^i_C(t + dt) \coloneqq I^i_C(t) + d^i_{E,I_C} - d^i_{I_C,G_D} - d^i_{I_C,R} - d^i_{I_C,ICU_{pre}} - d^i_{I_C,ICU_{pre^*}} - d^i_{I_C,H_R} - d^i_{I_C,H^*_R} - d^i_{I_C,H_D} - d^i_{I_C,H_{D^*}} \quad (130)$$

$$G^{i,1}_D(t + dt) \coloneqq G^{i,1}_D(t) + d^i_{I_C,G_D} - d^i_{G_D,G_D} \quad (131)$$

$$G^{i,2}_D(t + dt) \coloneqq G^{i,2}_D(t) + d^i_{G_D,G_D} - d^i_{G_D,D} \quad (132)$$

DOI: https://doi.org/10.25561/85146

$$ICU^i_{pre}(t+dt) := ICU^i_{pre}(t) + d^i_{I_C,ICU_{pre}} - d^i_{ICU_{pre},ICU_{W_R}} - d^i_{ICU_{pre},ICU_{W_D}} - d^i_{ICU_{pre},ICU_D} - d^i_{ICU_{pre},ICU_{pre^*}} - d^i_{ICU_{pre},ICU_{W_{R^*}}} - d^i_{ICU_{pre},ICU_{W_{D^*}}} - d^i_{ICU_{pre},ICU_{D^*}} \quad (133)$$

$$ICU^i_{pre^*}(t+dt) := ICU^i_{pre^*}(t) + d^i_{I_C,ICU_{pre^*}} - d^i_{ICU_{pre},ICU_{W_{D^*}}} - d^i_{ICU_{pre^*},ICU_{W_{R^*}}} - d^i_{ICU_{pre^*},ICU_{D^*}} \quad (134)$$

$$ICU^i_{W_R}(t+dt) := ICU^i_{W_R}(t) + d^i_{ICU_{pre},ICU_{W_R}} - d^i_{ICU_{W_R},W_R} - d^i_{ICU_{W_R},ICU_{W_{R^*}}} - d^i_{ICU_{W_R},W_{R^*}} \quad (135)$$

$$ICU^i_{W_{R^*}}(t+dt) := ICU^i_{W_{R^*}}(t) + d^i_{ICU_{pre^*},ICU_{W_{R^*}}} + d^i_{ICU_{W_R},ICU_{W_{R^*}}} + d^i_{ICU_{pre},ICU_{W_{R^*}}} - d^i_{ICU_{W_{R^*}},W_{R^*}} \quad (136)$$

$$ICU^i_{W_D}(t+dt) := ICU^i_{W_D}(t) + d^i_{ICU_{pre},ICU_{W_D}} - d^i_{ICU_{W_D},W_D} - d^i_{ICU_{W_D},ICU_{W_{D^*}}} - d^i_{ICU_{W_D},W_{D^*}} \quad (137)$$

$$ICU^i_{W_{D^*}}(t+dt) := ICU^i_{W_{D^*}}(t) + d^i_{ICU_{pre^*},ICU_{W_{D^*}}} + d^i_{ICU_{W_D},ICU_{W_{D^*}}} + d^i_{ICU_{pre},ICU_{W_{D^*}}} - d^i_{ICU_{W_{D^*}},W_{D^*}} \quad (138)$$

$$ICU^{i,1}_D(t+dt) := ICU^{i,1}_D(t) + d^i_{ICU_{pre},ICU_D} - d^i_{ICU_D,ICU_D} - d^{i,1,1}_{ICU_D,ICU_{D^*}} - d^{i,1,2}_{ICU_D,ICU_{D^*}} \quad (139)$$

$$ICU^{i,2}_D(t+dt) := ICU^{i,2}_D(t) + d^i_{ICU_D,ICU_D} - d^i_{ICU_D,D} - d^{i,2,2}_{ICU_D,ICU_{D^*}} \quad (140)$$

$$ICU^{i,1}_{D^*}(t+dt) := ICU^{i,1}_{D^*}(t) + d^i_{ICU_{pre^*},ICU_{D^*}} + d^{i,1,1}_{ICU_D,ICU_{D^*}} + d^i_{ICU_{pre},ICU_{D^*}} - d^i_{ICU_{D^*},ICU_{D^*}} \quad (141)$$

$$ICU^{i,2}_{D^*}(t+dt) := ICU^{i,2}_{D^*}(t) + d^i_{ICU_{D^*},ICU_{D^*}} + d^{i,1,2}_{ICU_D,ICU_{D^*}} + d^{i,2,2}_{ICU_D,ICU_{D^*}} - d^i_{ICU_{D^*},D} \quad (142)$$

$$W^{i,1}_R(t+dt) := W^{i,1}_R(t) + d^i_{ICU_{W_R},W_R} - d^i_{W_R,W_R} - d^{i,1,1}_{W_R,W_{R^*}} - d^{i,1,2}_{W_R,W_{R^*}} \quad (143)$$

$$W^{i,2}_R(t+dt) := W^{i,2}_R(t) + d^i_{W_R,W_R} - d^i_{W_R,R} - d^{i,2,2}_{W_R,W_{R^*}} \quad (144)$$

$$W^{i,1}_{R^*}(t+dt) := W^{i,1}_{R^*}(t) + d^i_{ICU_{W_{R^*}},W_{R^*}} + d^{i,1,1}_{W_R,W_{R^*}} + d^i_{ICU_{W_R},W_{R^*}} - d^i_{W_{R^*},W_{R^*}} \quad (145)$$

$$W^{i,2}_{R^*}(t+dt) := W^{i,2}_{R^*}(t) + d^i_{W_{R^*},W_{R^*}} + d^{i,2,2}_{W_R,W_{R^*}} + d^{i,1,2}_{W_R,W_{R^*}} - d^i_{W_{R^*},R} \quad (146)$$

$$W^i_D(t+dt) := W^i_D(t) + d^i_{ICU_{W_D},W_D} - d^i_{W_D,D} - d^i_{W_D,W_{D^*}} \quad (147)$$

$$W^i_{D^*}(t+dt) := W^i_{D^*}(t) + d^i_{ICU_{W_{D^*}},W_{D^*}} + d^i_{W_D,W_{D^*}} + d^i_{ICU_{W_D},W_{D^*}} - d^i_{W_{D^*},D} \quad (148)$$

$$H^{i,1}_D(t+dt) := H^{i,1}_D(t) + d^i_{I_C,H_D} - d^i_{H_D,H_D} - d^{i,1,1}_{H_D,H^*_D} - d^{i,1,2}_{H_D,H^*_D} \quad (149)$$

$$H^{i,2}_D(t+dt) := H^{i,2}_D(t) + d^i_{H_D,H_D} - d^i_{H_D,D} - d^{i,2,2}_{H_D,H^*_D} \quad (150)$$

$$H^{i,1}_{D^*}(t+dt) := H^{i,1}_{D^*}(t) + d^i_{I_C,H_{D^*}} + d^{i,1,1}_{H_D,H^*_D} - d^i_{H_{D^*},H_{D^*}} \quad (151)$$

$$H^{i,2}_{D^*}(t+dt) := H^{i,2}_{D^*}(t) + d^i_{H_{D^*},H_{D^*}} + d^{i,2,2}_{H_D,H^*_D} + d^{i,1,2}_{H_D,H^*_D} - d^i_{H_{D^*},D} \quad (152)$$

$$H^i_R(t+dt) := H^i_R(t) + d^i_{I_C,H_R} - d^i_{H_R,R} - d^i_{H_R,H^*_R} \quad (153)$$

$$H^i_{R^*}(t+dt) \coloneqq H^i_{R^*}(t) + d^i_{I_C,H_{R^*}} + d^i_{H_R,H^*_R} - d^i_{H_{R^*},R} \quad (154)$$

$$R^i(t+dt) \coloneqq R^i(t) + d^i_{I_A,R} + d^i_{I_C,R} + d^i_{H_R,R} + d^i_{H_{R^*},R} + d^i_{W_R,R} + d^i_{W_{R^*},R} \quad (155)$$

$$T^i_{sero_{pre}}(t+dt) \coloneqq T^i_{sero_{pre}}(t) + d^i_{E,I_A} + d^i_{E,I_C} - d^i_{T_{sero_{pre}},T_{sero_{pos}}} - d^i_{T_{sero_{pre}},T_{sero_{neg}}} \quad (156)$$

$$T^i_{sero_{pos}}(t+dt) \coloneqq T^i_{sero_{pos}}(t) + d^i_{T_{sero_{pre}},T_{sero_{pos}}} \quad (157)$$

$$T^i_{sero_{neg}}(t+dt) \coloneqq T^i_{sero_{neg}}(t) + d^i_{T_{sero_{pre}},T_{sero_{neg}}} \quad (158)$$

$$T^i_{PCR_{pre}}(t+dt) \coloneqq T^i_{PCR_{pre}}(t) + d^i_{S,E} - d^i_{T_{PCR_{pre}},T_{PCR_{pos}}} \quad (159)$$

$$T^i_{PCR_{pos}}(t+dt) \coloneqq T^i_{PCR_{pos}}(t) + d^i_{T_{PCR_{pre}},T_{PCR_{pos}}} - d^i_{T_{PCR_{pos}},T_{PCR_{neg}}} \quad (160)$$

$$T^i_{PCR_{neg}}(t+dt) \coloneqq T^i_{PCR_{neg}}(t) + d^i_{T_{PCR_{pos}},T_{PCR_{neg}}} \quad (161)$$

1.7 Observation process

To describe the epidemic in each NHS region, we fitted our model to time series data on hospital admissions, hospital ward occupancy (both in general beds and in ICU beds), deaths in hospitals, deaths in care homes, population serological surveys and PCR testing data (section 1.1 and Table S 1).

1.7.1 Notation for distributions used in this section

If $Y \sim \text{NegBinom}(m, \kappa)$, then Y follows a negative binomial distribution with mean m and shape κ, such that

$$P(Y = y) = \frac{\Gamma(\kappa + y)}{y!\,\Gamma(\kappa)} \left(\frac{\kappa}{\kappa + m}\right)^{\kappa} \left(\frac{m}{\kappa + m}\right)^{y} \quad (162)$$

where $\Gamma(x)$ is the gamma function. The variance of Y is $m + \frac{m^2}{\kappa}$.

If $Z \sim \text{BetaBinom}(n, \omega, \rho)$, then Z follows a beta-binomial distribution with size n, mean probability ω and overdispersion parameter ρ, such that

$$P(Z = z) = \binom{n}{z} \frac{\text{B}(z + a, n - z + b)}{\text{B}(a, b)} \quad (163)$$

where $a = \omega\left(\frac{1-\rho}{\rho}\right)$, $b = (1-\omega)\left(\frac{1-\rho}{\rho}\right)$ and $\text{B}(a, b)$ is the beta function. The mean of Z is $n\omega$ and the variance is $n\omega(1-\omega)[1 + (n-1)\rho]$.

1.7.2 Hospital admissions and new diagnoses in hospital

We represented the daily number of confirmed COVID-19 hospital admissions and new diagnoses for existing hospitalised cases, $Y_{adm}(t)$, as the observed realisations of an underlying hidden Markov process, $X_{adm}(t)$, defined as:

$$X_{adm}(t) \coloneqq \sum_i \Big(\sum_{j \in \{H_R^*, H_D^*, ICU_{pre}^*\}} d^i_{I_C,j} + \sum_{j \in \{H_R, ICU_{pre}, ICU_{W_R}, ICU_{W_D}, W_D\}} d^i_{j,j^*} \Big) + d^{i,1,1}_{H_D,H_{D^*}} + d^{i,1,2}_{H_D,H_{D^*}} + d^{i,2,2}_{H_D,H_{D^*}} + d^{i,1,1}_{ICU_D,ICU_{D^*}} + d^{i,1,2}_{ICU_D,ICU_{D^*}} + d^{i,2,2}_{ICU_D,ICU_{D^*}} + d^{i,1,1}_{W_R,W_{R^*}} + d^{i,1,2}_{W_R,W_{R^*}} + d^{i,2,2}_{W_R,W_{R^*}} + d^i_{ICU_{pre},ICU_{W_{R^*}}} + d^i_{ICU_{pre},ICU_{W_{D^*}}} + d^i_{ICU_{pre},ICU_{D^*}} + d^i_{ICU_{W_{D^*}},W_{D^*}} + d^i_{IC} \quad (164)$$

Which was related to the data via a reporting distribution:

$$Y_{adm}(t) \sim \text{NegBinom}(X_{adm}(t), \kappa_{adm}) \quad (165)$$

We allow for overdispersion in the observation process to account for noise in the underlying data streams, for example due to day-of-week effects on data collection. We adopt $\kappa = 2$ for all NHSE data streams, so that they contribute equal weight to the overall likelihood.

1.7.3 Hospital bed occupancy by confirmed COVID-19 cases

The model predicted general hospital bed occupancy by confirmed COVID-19 cases, $X_{hosp}(t)$ as:

$$X_{hosp}(t) \coloneqq \sum_i \left(I^i_{H_R^*}(t) + I^i_{H_D^*}(t) + I^i_{ICU_{pre}^*}(t) + I^i_{W_D^*}(t) + I^i_{W_R^*}(t) \right) \quad (166)$$

Which was related to the observed daily general bed-occupancy via a reporting distribution:

$$Y_{hosp}(t) \sim \text{NegBinom}\left(X_{hosp}(t), \kappa_{hosp}\right) \quad (167)$$

with $\kappa_{hosp} = 2$ as above.

Similarly, the model predicted ICU bed occupancy by confirmed COVID-19 cases, $X_{ICU}(t)$ as:

$$X_{ICU}(t) \coloneqq \sum_i \left(I^i_{ICU_{W_{R^*}}}(t) + I^i_{ICU_{W_{D^*}}}(t) + I^i_{ICU_{D^*}}(t) \right) \quad (168)$$

Which was related to the observed daily ICU bed-occupancy via a reporting distribution:

$$Y_{ICU}(t) \sim \text{NegBinom}(X_{ICU}(t), \kappa_{ICU}) \quad (169)$$

with $\kappa_{ICU} = 2$.

1.7.4 Hospital and care homes COVID-19 deaths

The reported number of daily COVID-19 deaths in hospitals, $Y_{hosp_D}(t)$ was considered as the observed realisation of an underlying hidden Markov process, $X_{hosp_D}(t)$, defined as:

$$X_{hosp_D}(t) \coloneqq \sum_i \left(d^i_{H_D,D} + d^i_{H^*_D,D} + d^i_{ICU_D,D} + d^i_{ICU^*_D,D} + d^i_{W_D,D} + d^i_{W^*_D,D} \right) \quad (170)$$

Which was related to the data via a reporting distribution:

$$Y_{hosp_D}(t) \sim \text{NegBinom}\left(X_{hosp_D}(t),\ \kappa_{hosp_D}\right) \quad (171)$$

with $\kappa_{hosp_D} = 2$.

Similarly, we represented the reported number of daily COVID-19 deaths in care homes, $Y_{G_D}(t)$, as the observed realisations of an underlying hidden Markov process, $X_{G_D}(t)$, defined as:

$$X_{G_D}(t) \coloneqq d^{CHR}_{G_D,D} \quad (172)$$

Which was related to the data via a reporting distribution:

$$Y_{G_D}(t) \sim \text{NegBinom}\left(X_{G_D}(t),\ \kappa_{G_D}\right) \quad (173)$$

with $\kappa_{G_D} = 2$.

1.7.5 Serosurveys

We model serological testing of all individuals aged 15-65, and define the resulting number of seropositive and seronegative individuals (were all individuals aged 15-65 to be tested), as:

$$X_{sero_{pos}}(t) \coloneqq \sum_{i=[15,20)}^{[60,65)} T^i_{sero_{pos}}(t) \quad (174)$$

$$X_{sero_{neg}}(t) \coloneqq \left(\sum_{i=[15,20)}^{[60,65)} N^i \right) - X_{sero_{pos}}(t) \quad (175)$$

We compared the observed number of seropositive results, $Y_{sero_{pos}}(t)$, with that predicted by our model, allowing for i) the sample size of each serological survey, $Y_{sero_{pos}}(t)$ and ii) imperfect sensitivity ($p_{sero_{sens}}$) and specificity ($p_{sero_{spec}}$) of the serological assay:

$$Y_{sero_{pos}}(t) \sim \text{Binom}\left(Y_{sero_{test}}(t), \omega_{sero_{pos}}(t)\right) \quad (176)$$

Where:

$$\omega_{sero_{pos}}(t) := \frac{p_{sero_{sens}} X_{sero_{pos}}(t) + \left(1 - p_{sero_{spec}}\right) X_{sero_{neg}}(t)}{X_{sero_{pos}}(t) + X_{sero_{neg}}(t)} \quad (177)$$

1.7.6 PCR testing

As described in the data section (1.1), we fitted the model to PCR testing data from two separate sources:

- pillar 2: the government testing programme, which recommends that individuals with COVID-19 symptoms are tested (34),
- the REACT-1 study, which aims to quantify the prevalence of SARS-CoV-2 in a random sample of the England population on an ongoing basis (35).

We only use Pillar 2 PCR test results for individuals aged 25 and over (we assume this includes all care home workers and residents). We assume that individuals who get tested through Pillar 2 PCR testing are either newly symptomatic SARS-CoV-2 cases (who will test positive):

$$X_{P2_{pos}}(t) \coloneqq \sum_{i=[25,30)}^{CHW} d^i_{E,I_C} \tag{178}$$

or non-SARS-CoV-2 cases who have symptoms consistent with COVID-19 (who will test negative):

$$X_{P2_{neg}}(t) \coloneqq p_{NC}\left(\left(\sum_{i=[25,30)}^{CHW} N^i\right) - X_{P2_{pos}}(t)\right) \tag{179}$$

where p_{NC} is the probability of non SARS-CoV-2 cases having symptoms consistent with COVID-19 leading them to seek a PCR test.

We compared the observed number of positive PCR tests, $Y_{P2_{pos}}(t)$ with that predicted by our model, accounting for the number of PCR tests conducted each day under pillar 2, $Y_{P2_{test}}(t)$, by calculating the probability of a positive PCR result (assuming perfect sensitivity and specificity of the PCR test):

$$\omega_{P2_{pos}}(t) \coloneqq \left(X_{P2_{pos}}(t)\right) / \left(X_{P2_{pos}}(t) + X_{P2_{neg}}(t)\right) \tag{180}$$

People may seek PCR tests for many reasons and thus the pillar 2 data are subject to competing biases. We therefore allowed for an over-dispersion parameter $\rho_{P2_{test}}$, which we fitted separately for each region in the modelling framework:

$$Y_{P2_{pos}}(t) \sim \text{BetaBinom}\left(Y_{P2_{test}}(t), \omega_{P2_{pos}}(t), \rho_{P2_{test}}\right) \tag{181}$$

DOI: https://doi.org/10.25561/85146

We incorporated the REACT-1 PCR testing data into the likelihood analogously to the serology data, by considering the model-predicted number of PCR-positives $X_{R1_{pos}}(t)$ and PCR-negatives $X_{R1_{neg}}(t)$, were all individuals aged over five and not resident in a care home to be tested:

$$X_{R1_{pos}}(t) \coloneqq \sum_{i=[5,10),\ldots,[80+),CHW} T^{i}_{PCR_{pos}}(t) \quad (182)$$

$$X_{R1_{neg}}(t) \coloneqq \sum_{i=[5,10),\ldots,[80+),CHW} N^{i}(t) - X_{R1_{pos}}(t) \quad (183)$$

We compared the daily number of positive results observed in REACT-1, $Y_{R1_{pos}}(t)$, given the number of people tested on that day, $Y_{R1_{test}}(t)$, to our model predictions, by calculating the probability of a positive result, assuming perfect sensitivity and specificity of the REACT-1 assay:

$$\omega_{R1_{pos}}(t) := \left(X_{R1_{pos}}(t)\right) / \left(X_{R1_{pos}}(t) + X_{R1_{neg}}(t)\right) \quad (184)$$

$$Y_{R1_{pos}}(t) \sim \text{Binom}\left(Y_{R1_{test}}(t), \omega_{R1_{pos}}(t)\right) \quad (185)$$

The overall likelihood function was then calculated as the product of the likelihoods of the individual observations.

1.8 Bayesian inference and model fitting

A closed-form expression of the likelihood of the observed data given the model and its parameters was not analytically tractable, so we used particle filtering methods to obtain an unbiased estimate of the likelihood which can be efficiently sampled from (36). Where appropriate, we used estimates from the literature to set model parameters at fixed values. We limited the parameters being inferred to just those with particular epidemiological interest, or with large uncertainty in existing literature.

The model was fitted independently to each NHS region. For each NHS region, we aimed to infer the values of 26 model parameters:

- the local epidemic start-date, t_0;
- thirteen transmission rates at different time points $\beta_1, \ldots, \beta_{12}$;
- three parameters governing transmission to and within care homes $m_{CHW}, m_{CHR}, \epsilon$;
- the probability of symptomatic individuals developing serious disease requiring hospitalisation, p_H^{max}, for the group with the largest probability;
- the probability of a care home resident dying in a care home if they have severe disease requiring hospitalisation, $p_{G_D}^{CHR}$;
- the probability, at the start of the pandemic, of a patient being admitted to ICU after hospitalisation, p_{ICU}^{max}, for the group with the largest probability;

DOI: https://doi.org/10.25561/85146

- the probabilities, at the start of the pandemic, of dying in a hospital general ward, $p_{H_D}^{max}$, in the ICU, $p_{ICU_D}^{max}$, and in a stepdown ward following ICU, $p_{W_D}^{max}$, for the groups with the largest probability;
- the multiplier for hospital mortality after improvement in care, μ_H;
- the multiplier for probability of admission to ICU after improvement in care, μ_{ICU};
- the daily proportion p_{NC}, of the population seeking to get tested for an infection of SARS-Cov-2 following COVID-19 like symptoms and the overdispersion of the corresponding observation distribution $\rho_{P2_{test}}$.

We used particle Monte Carlo Markov Chain (pMCMC) methods (37), implementing a particle marginal Metropolis-Hastings algorithm with a bootstrap particle filter (38) with 96 particles (for sufficient variance in likelihood and a convenient multiple of number of available CPU cores for efficiency), to obtain a sample from the posterior distribution of the model parameters given the observed data. If the expected values of count distributions are zero when observed values are non-zero, this results in particles of zero weight, which can lead to the particle filter estimating the marginal likelihood to be 0. Therefore, to get a small but non-zero weight for each particle at every observation, within our particle filter likelihood we add a small amount of noise (exponentially distributed with mean 10^{-6}) to count values from the model.

Within our particle filter we add small amounts of exponentially-distributed noise (with mean 10^{-6}) to model outputs prior to calculating likelihood weights to avoid particles of zero weight, instead resulting in small but non-zero weights.

We implemented our model and parameter inference in an R package, *sircovid* (39), available at https://mrc-ide.github.io/sircovid, which uses two further R packages, *dust* to run the model in efficient compiled code and *mcstate* to implement the pMCMC sampler using Metropolis-Hastings sampling (40).

At each iteration, the sampler proposes an update to the joint distribution of parameters. These proposals are generated from multivariate Gaussian densities centred on the current parameter values, and with a covariance structure chosen to facilitate efficient mixing of the Markov chain. We specified reflecting boundaries for the proposal kernel to ensure that the proposed parameters are both epidemiologically and mathematically plausible and retain symmetry in the proposals.

For each regional fit, eight parallel chains of the pMCMC were run for 11,000 iterations, with the first 1,000 discarded as burn-in, and a 1/80 thinning. We assessed convergence visually.

DOI: https://doi.org/10.25561/85146

1.9 Prior distributions and parameter calibration

1.9.1 Risk of hospital admission

In our Bayesian inference framework, we estimate p_H^{max}, the probability of hospital admission for symptomatic cases in the group (across all ages and CHW and CHR) with the largest probability of hospital admission. However, we fix the relative probability of hospital admission for the other age groups, ψ_H^i, defined so that $p_H^i = p_H^{max}\,\psi_H^i$, with $\psi_H^i = 1$ in the group with largest probability of hospital admission.

In this section we explain how the values of ψ_H^i were chosen. We used two sources of information, an individual-level and an aggregated dataset. On the one hand, the COVID-19 Hospitalisation in England Surveillance System (CHESS) is a daily, confidential line list containing highly detailed information on patients admitted to hospital with confirmed COVID-19 (see following section 1.9.2 for further details). On the other hand, the Government's Coronavirus Dashboard is an aggregated, publicly available situation report updated daily. Amongst other data, it provides updates on the number of daily admissions and hospital occupancy by devolved nation and, for England, by NHS region. We found the demography of hospitalisation in CHESS to be biased toward older patients compared to Dashboard data (Figure S3). We thus undertook a two-step approach to infer the demographic composition of COVID-19 hospitalisations across England.

Firstly, we derived an initial approximation of ψ_H^i by dividing the total number of hospital admissions for age group i in CHESS over the total number of positive PCR tests (Pillar 2) for i. Both data sources were censored to include patients admitted to hospital or with a specimen data (i.e. the date the test was taken), respectively, between March 1 and December 2, 2020. We ran our full inference framework using this initial approximation for ψ_H^i and observed its fit to the demographic composition of admissions from the data.

As a second step, we refined our initial approximations of ψ_H^i over a series of iterations of our inference framework, by drawing the modelled ($p_{H\,Model}^i$) and observed ($p_{H\,Dashboard}^i$) proportion of admissions for each age group (i.e. admissions in age group i divided by all admissions) and using it to derive a re-scaling factor for a new proposal for ψ_H^i as follows:

$$New\ \psi_H^i = Initial\ \psi_H^i * \frac{p_{H\,Dashboard}^i}{p_{H\,Model}^i} \tag{186}$$

This process was repeated to obtain a close approximation to the observed proportion of admissions by age and region (Figure S3). A key strength of our approach is that we did not overfitted demography by individual regions. Rather, by assuming ψ_H^i to be independent of geographic region, we allowed our inference framework to derive the number of admissions for each five-year age band i solely based on ψ_H^i, the demographic composition of the NHS region and inferred epidemic parameters, such as R_t.

DOI: https://doi.org/10.25561/85146

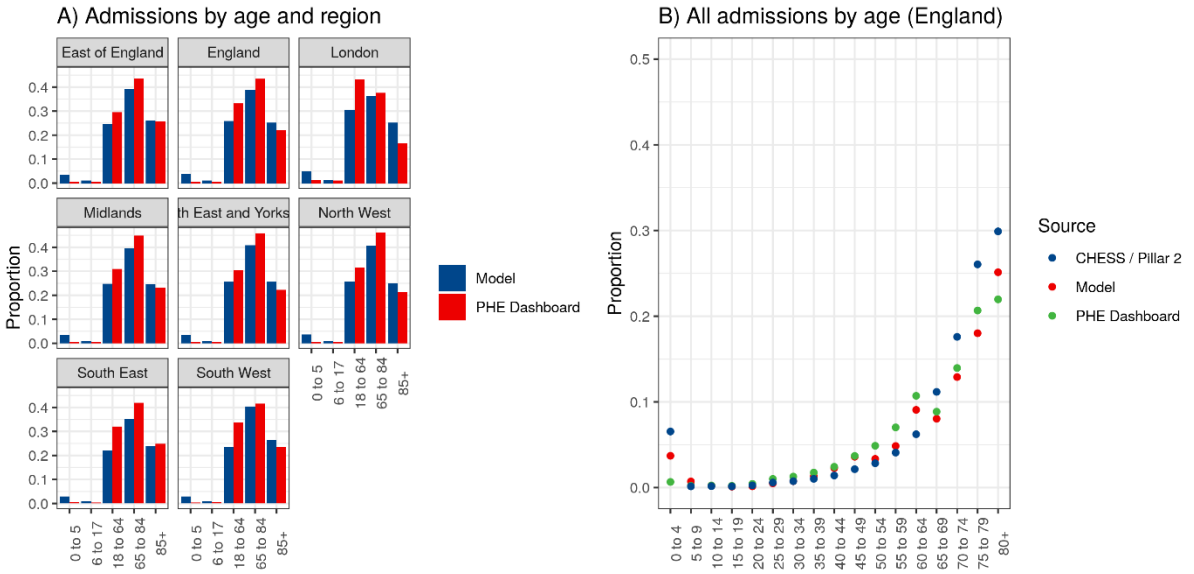

Figure S 3: Proportion of admissions by age. a) Comparison of model outputs to data from the Government's Coronavirus Dashboard, aggregated to five broad age categories. b) Age spline fitted (red) to Government's Coronavirus Dashboard, with age categories disaggregated to five-year bands. The fitted spline (red) was used as input parameters for the probability of hospitalisation by age.

1.9.2 Severity and hospital progression

We also performed extensive preliminary analysis to inform the age-structure of progression parameters within hospital. Data from the COVID-19 Hospitalisation in England Surveillance System (CHESS) were used to fit a simple model of patient clinical progression in hospital. The model structure was designed to mirror the within-hospital component of the wider mechanistic transmission model, but without the complexities arising from unknown admission dates and with greater detail on trends with age (Figure S 4).

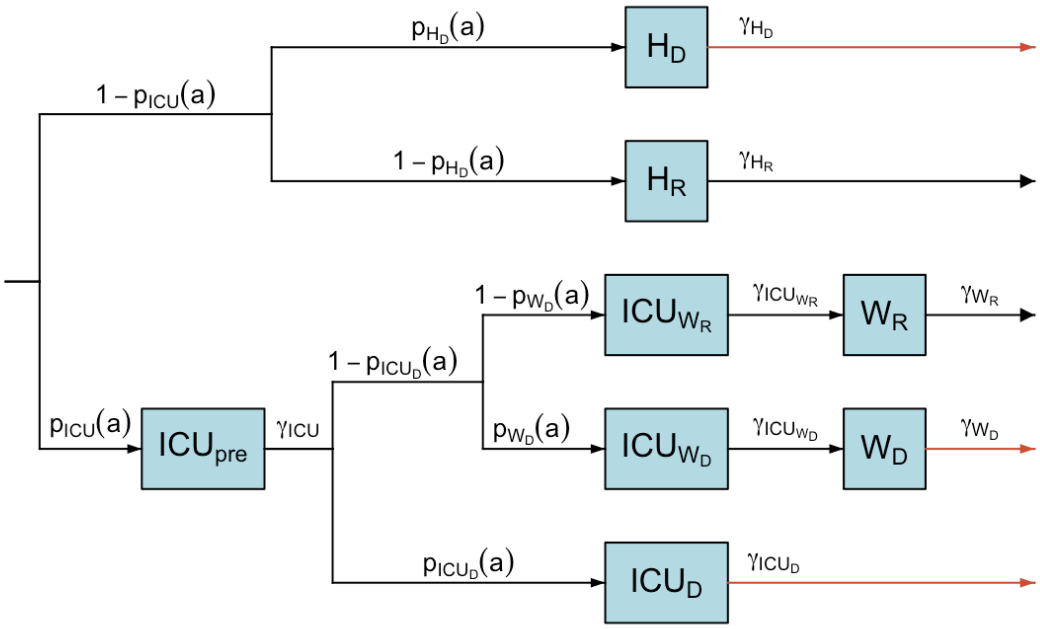

Figure S 4: Directed Acyclic Graph of the hospital pathways fitted to CHESS data, which mirror the model structure described in described in Figure S 2, but with all parameters varying with age and not over time.

CHESS data consists of a line list of daily individual patient-level data on COVID-19 infection in persons requiring hospitalisation, including demographic and clinical information on severity and outcomes. Data were filtered to patients admitted between 18th March and 31st May 2020 (inclusive), with subsequent progression events possible up until 25th Nov 2020. This gave >5 months for outcomes to complete, and hence justified filtering to patients with resolved outcomes only. The length of stay in each state was taken as the difference between the registered dates of entering and leaving each hospital ward. Lengths of stay were assumed to follow Erlang distributions, as in the wider model, with a distinct mean and shape parameter for each state. Specifically, the probability of being in state $X \in \{pre, H_D, H_R, ICU_D, ICU_{W_R}, ICU_{W_D}, W_R, W_D\}$ for $n \in \mathbb{N}_0$ days was taken as the integral over day n of the Erlang distribution with mean m_X and shape s_X:

$$\Pr(\text{in state } X \text{ for } n \text{ days}) = \int_n^{n+1} \frac{\left(\frac{s_X}{m_X}\right)^{s_X} t^{s_X-1} e^{\frac{-s_X t}{m_X}}}{(s_X-1)!}\, dt. \tag{187}$$

For a patient of age a, this was combined with the probability of their path through the hospital progression model, taken as the product of the individual transition probabilities at each bifurcation, i.e. values taken from $p_Z(a)$ for $Z \in \{ICU, H_D, ICU_D, W_D\}$. Transition probabilities were modelled as functions of age using logistic-transformed cubic splines. Knots were defined at coordinates $[x^i, y_Z^i]$, where x^i values were fixed at $\{0, 20, 40, 60, 80, 100, 120\}$ and y_Z^i were free parameters to be estimated. The complete spline, $y_Z(a)$ for $a \in 0{:}\,120$, was obtained from these knots using standard expressions for cubic spline interpolation. Finally, transition probabilities were obtained from the raw $y_Z(a)$ values using the logistic transformation: $p_Z(a) = 1/(1 + e^{-y_Z(a)})$.

In total there were 44 free parameters in the within-hospital progression model: 8 mean length of stay parameters, 8 length of stay shape parameters and 4×7 transition probability spline nodes (Figure S 4, Table S 4).

Table S 4: Descriptions of all states and transitions in the simplified hospital progression model fitted to CHESS data.

State (X)	Description
pre	General admission before step-up to ICU
H_D	General ward before death in general ward
H_R	General ward before discharge from general ward
ICU_D	ICU before death in ICU
ICU_{W_D}	ICU before step-down and eventual death in step-down care
ICU_{W_R}	ICU before step-down and eventual discharge from step-down care
W_D	Step-down (general) ward before death
W_R	Step-down (general) ward before discharge
Transition (Z)	**Description**
ICU	Admission to ICU from general ward
H_D	Death in general ward
ICU_D	Death in ICU
W_D	Death in step-down care

All parameters of the hospital progression model were given priors (Table S 5) and estimated within a Bayesian framework. All length of stay parameters were given uniform priors over a plausible range of values. For transition probabilities, the first spline node y_Z^1 was given a prior that corresponded to a uniform distribution after logistic transformation, and subsequent spline nodes were given a multivariate normal prior to apply a smoothing constraint to the spline. Parameters were estimated jointly via MCMC using the custom package *markovid* v1.5.0 (41), which uses the random-walk Metropolis-Hastings algorithm to draw from the joint posterior distribution. MCMC was run for 1000 burn-in iterations and 100,000 sampling iterations replicated over 10 independent chains. Convergence was assessed via the Gelman-Rubin diagnostic (all parameters had potential scale reduction factor <1.1) and sampling sufficiency was assessed by visualising posterior distributions and by effective sample size (ESS) calculations (all parameters had ESS >100,000).

Table S 5: Priors on all length of stay distributions and transition probability splines. $Uniform_{cont}(a,b)$ denotes the continuous uniform distribution, and $Uniform_{disc}(a,b)$ the discrete uniform distribution between a and b (inclusive).

Parameter	Description	Prior
$\boldsymbol{m_X = \frac{1}{\gamma_X}}$	Mean of Erlang length of stay distribution	$m_X \sim \text{Uniform}_{\text{cont}}(0,20)$
$\boldsymbol{k_X}$	Shape parameter of Erlang length of stay distribution	$k_X \sim \text{Uniform}_{\text{disc}}(1,10)$
$\boldsymbol{y_Z^1}$	First spline node of (transformed) transition probability	$f(y_Z^1) \propto \frac{e^{-y_Z^1}}{(1+e^{-y_Z^1})^2}$, for $y_Z^! \in (-10,10)$
$\boldsymbol{y_Z^j}$ for $\boldsymbol{j \in 2:7}$	Subsequent spline nodes of (transformed) transition probability	$y_Z^j \sim \text{Normal}(y_Z^{j-1}, 0.25)$

Parameter estimates (posterior medians) were passed to the wider mechanistic transmission model as fixed values (Figure S *5*). For transition probabilities, the full age-spline (Figure 3, main text) was aggregated to 5-year age groups and normalised by the largest value to define the relative risk with age. The absolute risk in the mechanistic transmission model was obtained by multiplying the relative risk by region-specific scaling factors that were fitted as free parameters in the pMCMC. Hence, the preliminary analysis of CHESS data was used to inform trends of severity with age, but not the absolute probability of progression through the hospital states, which was informed by the Government's Coronavirus Dashboard data.

For the wider mechanistic transmission model, we used Beta distributions for the priors of the various fitted probabilities regarding hospitalisation. The priors for p_{ICU}^{max}, $p_{H_D}^{max}$, $p_{ICU_D}^{max}$ and $p_{W_D}^{max}$ were all informed by the fitting to CHESS data by taking the median fitted value for the prior mean, which we halve in the case of p_{ICU}^{max} to account for CHESS being biased to more severe patients. The prior distributions are then calibrated so that the lower bound of the 95% confidence interval is 0.1 lower than the prior mean. For p_H^{max} and $p_{G_D}^{CHR}$, we assume prior means of 0.75 and calibrate the prior so that the lower bound of the 95% confidence interval is 0.2 lower than the mean. For the multipliers for hospital mortality after improvement in care, μ_H, and for probability of admission to ICU after improvement in care, μ_{ICU}, we used uninformative $U[0,1]$ priors.

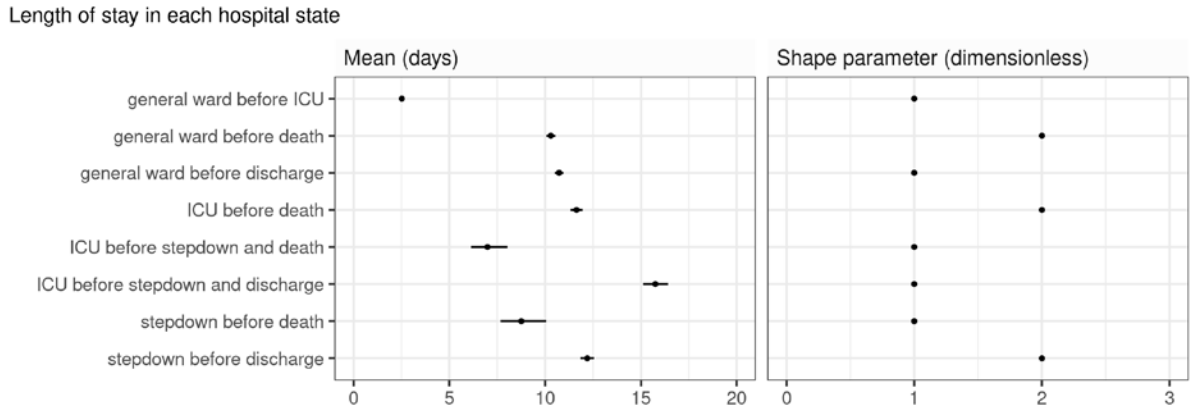

Figure S 5*: Posterior 95% credible intervals of length of stay mean (left) and shape parameters (right).*

1.9.3 Serosurveys

To keep serology parameters consistent between all regions we used estimates from the literature to fix the parameters of the seroconversion process. An alternative would have been to use these estimates as priors within a hierarchical model where some parameters would be shared between regions, but this would be much more involved computationally.

As described in section 1.3.2, the time to seroconversion from leaving the E^i compartment is modelled by an exponential distribution time spent in $T^i_{sero_{pre}}$ with a proportion $p_{sero_{pos}}$ ultimately seroconverting and moving to $T^i_{sero_{pos}}$ and the remaining staying negative and moving to $T^i_{sero_{neg}}$.

We fixed $p_{sero_{pos}}$ to 0.85 based on the estimate of 85% of infections becoming detectably seropositive with the EUROIMMUN assay used in the NHSBT serological surveys (42). The specificity of the serology test $p_{sero_{spec}}$is fixed to 0.99 also from (42). Finally, the sensitivity of serology test $p_{sero_{sens}}$ is assumed to be 1 as it is non-distinguishable from the time varying seroconversion process (Table S7).

1.9.4 PCR positivity

As for other compartments, we modelled the duration of SARS-CoV-2 PCR-positivity after symptom onset using an Erlang distribution $\tau \sim \text{Erlang}(k, \gamma)$, with k successive compartments and a total mean time spent of $\frac{k}{\gamma}$ and variance $\frac{k}{\gamma^2}$.

We estimated the parameters of this distribution from Omar et al. (16), which reported the cumulative distribution of duration of PCR positivity in 523 individuals with mild COVID-19 disease in home quarantine in a German region. We performed a survival analysis using a gamma-accelerated failure time model fitted to their data, from which we estimated the mean and variance of the time from symptom onset to PCR negativity. This was used to derive values of k and γ shown in Table S 2.

1.9.5 Local start date of the epidemic

The start date of the epidemic for each region is assumed to have a uniform prior on the dates from 1st January 2020 to 15th March 2020, inclusive – with the latter date corresponding to the last date before the data begin.

1.9.6 Time-varying transmission rates

We set priors for the transmission rates $\beta_1, \dots, \beta_{12}$ to reflect a Gamma distribution for the reproduction number R_t with a reasonable 95% confidence interval a priori. To obtain a prior for the corresponding β_k, we then scale by a factor of 0.0241 (given other parameter values, $\beta_k = 0.0241$ would correspond approximately to $R_t = 1$). The 95% ranges for R_0 we used are (i) (2.5, 3.5) at the onset of the epidemic (corresponding to β_1); and then R_t (ii) (0.4, 3.5) at announcement of the first lockdown (corresponding to β_2); and (iii) (0.4, 3) from the implementation of the first lockdown onwards (corresponding to $\beta_3, \dots, \beta_{12}$). The values are consistent with the values of the COMIX study (43).

1.9.7 Transmission within care homes

For the transmission between care home workers and residents, m_{CHW}, and transmission among care home residents, m_{CHR}, we used a prior distributions reflecting that these are person-to-person infectious contact rates and thus should be scaled according to regional care home demography. We then used a Gamma distribution with shape 5 and mean $\frac{0.1}{N^{CHR}}$ for both of these parameters (recall that we assume there is a 1-to-1 ratio of care home workers to residents in each region, so $N^{CHW} = N^{CHR}$).

For the parameter governing the reduction in contacts between the general population and care home residents, ϵ, we used an uninformative $U[0,1]$ prior.

1.9.8 Parameters relating to Pillar 2 testing

For both the parameters p_{NC} and $\rho_{P2_{test}}$, we used uninformative $U[0,1]$ priors.

DOI: https://doi.org/10.25561/85146

Table S 6: Inferred model parameter notations, prior and posterior distributions. Note that $\Gamma(a, b)$ here refers to a Gamma distribution with shape a and scale b (such that the mean is ab), and $B(a, b)$ refers to a Beta distribution with shape parameters a and b (such that the mean is $a/(a+b)$).

	Description	Group scaling	Prior	*Mean (95% CI)*	Posterior NW**	Mean (95% CrI) NEY	MID	EE	LON	SW	SE
t_0	Start date of regional outbreak *(dd/mm/2020)*	-	$U[01/01, 15/03]$	-	29-01 (13/01, 07/02)	03-02 (29/01, 09/02)	24-01 (12/01, 02/02)	06-02 (30/01, 14/02)	08-01 (02/01, 20/01)	12-02 (09/02, 17/02)	27-01 (20/01, 04/02)
	Transmission rate *(pp)*										
$\boldsymbol{\beta}(\mathbf{t})$	β_1	-	$\Gamma(136,0.00($	*0.07 (0.06, 0.08)*	0.08 (0.06, 0.09)	0.08 (0.07, 0.09)	0.08 (0.07, 0.09)	0.08 (0.08, 0.09)	0.06 (0.06, 0.07)	0.09 (0.08, 0.09)	0.08 (0.07, 0.09)
	β_2	-	$\Gamma(21.9,0.00$	*0.06 (0.04, 0.08)*	0.09 (0.06, 0.11)	0.08 (0.05, 0.1)	0.07 (0.05, 0.09)	0.06 (0.04, 0.07)	0.04 (0.03, 0.06)	0.07 (0.04, 0.09)	0.05 (0.03, 0.06)
	β_3	-	$\Gamma(4.25,0.00$	*0.03 (0.01, 0.07)*	0.01 (0.01, 0.02)	0.02 (0.02, 0.02)	0.01 (0.01, 0.01)	0.02 (0.02, 0.02)	0.01 (0.01, 0.01)	0.02 (0.01, 0.02)	0.01 (0.01, 0.01)
	β_4	-	$\Gamma(4.25,0.00$	*0.03 (0.01, 0.07)*	0.02 (0.02, 0.02)	0.02 (0.01, 0.02)	0.02 (0.02, 0.02)	0.02 (0.02, 0.02)	0.02 (0.02, 0.02)	0.02 (0.02, 0.02)	0.02 (0.01, 0.02)
	β_5	-	$\Gamma(4.25,0.00$	*0.03 (0.01, 0.07)*	0.02 (0.02, 0.02)	0.02 (0.02, 0.03)	0.02 (0.02, 0.02)	0.02 (0.02, 0.02)	0.02 (0.02, 0.03)	0.02 (0.01, 0.02)	0.02 (0.02, 0.02)
	β_6	-	$\Gamma(4.25,0.00$	*0.03 (0.01, 0.07)*	0.02 (0.02, 0.03)	0.02 (0.01, 0.02)	0.02 (0.01, 0.02)	0.02 (0.02, 0.02)	0.02 (0.02, 0.03)	0.02 (0.02, 0.03)	0.02 (0.02, 0.02)
	β_7	-	$\Gamma(4.25,0.00$	*0.03 (0.01, 0.07)*	0.02 (0.02, 0.03)	0.02 (0.02, 0.03)	0.02 (0.02, 0.03)	0.02 (0.02, 0.02)	0.03 (0.02, 0.03)	0.02 (0.02, 0.03)	0.02 (0.02, 0.03)
	β_8	-	$\Gamma(4.25,0.00$	*0.03 (0.01, 0.07)*	0.05 (0.04, 0.05)	0.04 (0.04, 0.05)	0.05 (0.04, 0.05)	0.04 (0.03, 0.04)	0.04 (0.04, 0.05)	0.03 (0.02, 0.04)	0.04 (0.03, 0.05)
	β_9	-	$\Gamma(4.25,0.00$	*0.03 (0.01, 0.07)*	0.04 (0.04, 0.05)	0.04 (0.04, 0.04)	0.04 (0.03, 0.04)	0.04 (0.03, 0.04)	0.04 (0.03, 0.04)	0.04 (0.04, 0.05)	0.03 (0.03, 0.04)

DOI: https://doi.org/10.25561/85146

	β_{10}	-	$\Gamma(4.25,0.00$	*0.03 (0.01, 0.07)*	0.03 (0.03, 0.03)	0.03 (0.03, 0.04)	0.04 (0.04, 0.04)	0.04 (0.03, 0.04)	0.04 (0.04, 0.04)	0.04 (0.03, 0.04)	0.04 (0.04, 0.04)
	β_{11}	-	$\Gamma(4.25,0.00$	*0.03 (0.01, 0.07)*	0.02 (0.02, 0.02)	0.03 (0.02, 0.03)	0.03 (0.02, 0.03)	0.03 (0.02, 0.03)	0.03 (0.02, 0.03)	0.03 (0.02, 0.03)	0.03 (0.03, 0.03)
	β_{12}	-	$\Gamma(4.25,0.00$	*0.03 (0.01, 0.07)*	0.02 (0.02, 0.02)	0.02 (0.02, 0.02)	0.02 (0.02, 0.02)	0.02 (0.02, 0.03)	0.03 (0.03, 0.03)	0.02 (0.02, 0.02)	0.02 (0.02, 0.03)
ϵ	Relative reduction in contacts between CHR and the general population	-	$U[0,1]$	*0.5 (0.03, 0.98)*	0.43 (0.03, 0.95)	0.75 (0.51, 0.98)	0.77 (0.37, 0.97)	0.79 (0.51, 0.96)	0.28 (0.03, 0.49)	0.82 (0.74, 0.91)	0.89 (0.77, 0.99)
m_{CHW}	Transmission rate between care home residents and staff	-	*Regional Prior*		$\Gamma(5, 4.3 \times 10^{-7})$ 2.2×10^{-6} *(*7.0×10^{-7}*,* 4.4×10^{-6}*)*	$\Gamma(5, 3.7 \times 10^{-7})$ 1.8×10^{-6} *(*5.9×10^{-7}*,* 3.7×10^{-6}*)*	$\Gamma(5, 2.9 \times 10^{-7})$ 1.5×10^{-6} *(*4.7×10^{-7}*,* 2.9×10^{-6}*)*	$\Gamma(5, 5.2 \times 10^{-7})$ 2.6×10^{-6} *(*8.4×10^{-7}*,* 5.3×10^{-6}*)*	$\Gamma(5, 7.6 \times 10^{-7})$ 3.8×10^{-6} *(*1.2×10^{-6}*,* 7.8×10^{-6}*)*	$\Gamma(5, 4.9 \times 10^{-7})$ 2.5×10^{-6} *(*8.0×10^{-7}*,* 5.0×10^{-6}*)*	$\Gamma(5, 3.1 \times 10^{-7})$ 1.6×10^{-6} *(*5.1×10^{-7}*,* 3.2×10^{-6}*)*
			Posterior:		2.1e-06 (1.4e-06, 2.7e-06)	1.7e-06 (1.3e-06, 2.2e-06)	1.5e-06 (1.1e-06, 1.9e-06)	2.7e-06 (2.1e-06, 3.1e-06)	3.8e-06 (3.1e-06, 4.7e-06)	1.8e-06 (1.3e-06, 2.2e-06)	1.5e-06 (1.1e-06, 1.8e-06)
m_{CHR}	Transmission rate among care home residents	-	*Regional Prior*		$\Gamma(5, 4.3 \times 10^{-7})$ 2.2×10^{-6} *(*7.0×10^{-7}*,* 4.4×10^{-6}*)*	$\Gamma(5, 3.7 \times 10^{-7})$ 1.8×10^{-6} *(*5.9×10^{-7}*,* 3.7×10^{-6}*)*	$\Gamma(5, 2.9 \times 10^{-7})$ 1.5×10^{-6} *(*4.7×10^{-7}*,* 2.9×10^{-6}*)*	$\Gamma(5, 5.2 \times 10^{-7})$ 2.6×10^{-6} *(*8.4×10^{-7}*,* 5.3×10^{-6}*)*	$\Gamma(5, 7.6 \times 10^{-7})$ 3.8×10^{-6} *(*1.2×10^{-6}*,* 7.8×10^{-6}*)*	$\Gamma(5, 4.9 \times 10^{-7})$ 2.5×10^{-6} *(*8.0×10^{-7}*,* 5.0×10^{-6}*)*	$\Gamma(5, 3.1 \times 10^{-7})$ 1.6×10^{-6} *(*5.1×10^{-7}*,* 3.2×10^{-6}*)*
			Posterior:		2.2e-06 (1e-06, 3.4e-06)	2.5e-06 (1.4e-06, 3.6e-06)	1.6e-06 (7e-07, 2.4e-06)	3.4e-06 (2.1e-06, 4.3e-06)	2.8e-06 (5e-07, 4.8e-06)	4.2e-06 (3.8e-06, 4.6e-06)	3.3e-06 (2.9e-06, 3.6e-06)
p_H^{max}	Probability of hospitalisation if symptomatic	ψ_H^i	B(15.8, 5.28)	*0.75 (0.55, 0.91)*	0.87 (0.8, 0.92)	0.9 (0.85, 0.94)	0.89 (0.83, 0.95)	0.78 (0.73, 0.84)	0.85 (0.79, 0.9)	0.86 (0.81, 0.93)	0.73 (0.68, 0.79)

DOI: https://doi.org/10.25561/85146

$p_{G_D}^{max}$	Probability of death in care home if requiring hospitalisation	$\psi_{G_D}^i$	B(15.8, 5.28)	*0.75 (0.55, 0.91)*	0.66 (0.37, 0.85)	0.77 (0.64, 0.88)	0.53 (0.41, 0.69)	0.58 (0.52, 0.63)	0.66 (0.5, 0.91)	0.64 (0.6, 0.69)	0.36 (0.32, 0.43)
p_{ICU}^{max}	Probability of triage to ICU for new hospital admissions	ψ_{ICU}^i	B(13.9, 43.9)	*0.24 (0.14, 0.36)*	0.15 (0.11, 0.18)	0.15 (0.11, 0.18)	0.17 (0.13, 0.21)	0.25 (0.21, 0.31)	0.31 (0.26, 0.37)	0.12 (0.11, 0.13)	0.23 (0.2, 0.25)
$p_{H_D}^{max}$	Initial probability of death for general inpatients	$\psi_{H_D}^i$	B(42.1, 50.1)	*0.46 (0.36, 0.56)*	0.42 (0.35, 0.5)	0.46 (0.39, 0.53)	0.43 (0.38, 0.47)	0.47 (0.44, 0.51)	0.37 (0.32, 0.46)	0.5 (0.47, 0.53)	0.41 (0.35, 0.46)
$p_{ICU_D}^{max}$	Initial probability of death for ICU inpatients	$\psi_{ICU_D}^i$	B(60.2, 29.3)	*0.67 (0.57, 0.77)*	0.66 (0.6, 0.72)	0.71 (0.66, 0.77)	0.69 (0.58, 0.77)	0.69 (0.62, 0.75)	0.61 (0.51, 0.69)	0.71 (0.64, 0.77)	0.63 (0.54, 0.75)
$p_{W_D}^{max}$	Initial probability of death for stepdown inpatients	$\psi_{W_D}^i$	B(28.7, 52.1)	*0.35 (0.25, 0.46)*	0.35 (0.25, 0.46)	0.35 (0.3, 0.4)	0.36 (0.3, 0.5)	0.37 (0.3, 0.43)	0.34 (0.24, 0.43)	0.51 (0.44, 0.59)	0.37 (0.29, 0.45)
μ_{ICU}	ICU admission multiplier after improvement in care	-	$U[0,1]$	*0.5 (0.03, 0.98)*	0.79 (0.59, 0.93)	0.76 (0.62, 0.93)	0.72 (0.56, 0.94)	0.51 (0.37, 0.64)	0.62 (0.51, 0.75)	0.83 (0.73, 0.99)	0.44 (0.3, 0.54)
μ_D	Hospital mortality multiplier after improvement in care	-	$U[0,1]$	*0.5 (0.03, 0.98)*	0.58 (0.45, 0.69)	0.47 (0.41, 0.53)	0.49 (0.42, 0.61)	0.47 (0.42, 0.56)	0.32 (0.27, 0.38)	0.35 (0.28, 0.43)	0.53 (0.44, 0.63)
p_{NC}	Prevalence of non-COVID symptomatic leading to test	-	$U[0,1]$	*0.5 (0.03, 0.98)*	0.0031 (0.0029, 0.0033)	0.0022 (0.0021, 0.0024)	0.0025 (0.0023, 0.0027)	0.0028 (0.0026, 0.0031)	0.0028 (0.0027, 0.003)	0.0019 (0.0018, 0.002)	0.003 (0.0028, 0.0031)

DOI: https://doi.org/10.25561/85146

$\rho_{P2_{test}}$	Overdispersion of PCR positivity	-	$U[0,1]$	*0.5 (0.03, 0.98)*	0.0052 (0.0042, 0.0062)	0.0076 (0.0064, 0.0086)	0.0072 (0.0058, 0.0088)	0.0033 (0.0029, 0.0037)	0.0026 (0.0021, 0.0031)	0.0091 (0.0079, 0.0103)	0.0032 (0.0027, 0.0037)

*Age-specific scaling factors for each parameter are set out in Table S 8.** Region codes: NW = North West, NEY = North East and Yorkshire, MID = Midlands, EE = East of England, LON = London, SW = South West, SE = South East. N.B. when the prior is region specific the prior is shown in the same columns as the posterior distributions*

DOI: https://doi.org/10.25561/85146

Table S 7: Fixed model parameters (age / care home scaling factors are shown separately in Table S 8).

Parameter	Description	Value	Source
p_C	Probability of developing symptoms after becoming infectious	0.6	Lavezzo et al. (44)
p^*	Probability of arriving at hospital with a confirmed diagnosis	0.25	NHS (45)
$1/\gamma_U$	Mean time to confirmation of diagnosis within hospital (days)	2 days	CHESS (5)
$p_{sero_{pos}}$	Probability of seroconversion	0.85	Brazeau et al. (42)
$1/\gamma_{sero_{pos}}$	Mean time to seroconversion from infectiousness (days)	13	Benny et al. (17)
$p_{sero_{spec}}$	Specificity of serology test	0.99	Brazeau et al. (42)
$p_{sero_{sens}}$	Sensitivity of serology test	1	Assumption as non-distinguishable from time varying seroconversion

DOI: https://doi.org/10.25561/85146

Table S 8: Age / care-home scaling factors

	Age / care home -stratified scaling to probability of:					
	Hospitalisation if symptomatic (p_H)	*Triage to ICU* (p_{ICU})	*Death for general inpatients* (p_{H_D})	*Death in ICU* (p_{ICU_D})	*Death in stepdown* (p_{W_D})	*Death in the community* (p_{G_D})
Population group	ψ_H^i	ψ_{ICU}^i	$\psi_{H_D}^i$	$\psi_{ICU_D}^i$	$\psi_{W_D}^i$	$\psi_{G_D}^i$
Age						
[0, 5)	0.039	0.243	0.039	0.282	0.091	0
[5, 10)	0.001	0.289	0.037	0.286	0.083	0
[10, 15)	0.006	0.338	0.035	0.291	0.077	0
[15, 20)	0.009	0.389	0.035	0.299	0.074	0
[20, 25)	0.026	0.443	0.036	0.310	0.074	0
[25, 30)	0.040	0.503	0.039	0.328	0.076	0
[30, 35)	0.042	0.570	0.045	0.353	0.080	0
[35, 40)	0.045	0.653	0.055	0.390	0.086	0
[40, 45)	0.050	0.756	0.074	0.446	0.093	0
[45, 50)	0.074	0.866	0.107	0.520	0.102	0
[50, 55)	0.138	0.954	0.157	0.604	0.117	0
[55, 60)	0.198	**1.000**	0.238	0.705	0.148	0
[60, 65)	0.247	0.972	0.353	0.806	0.211	0
[65, 70)	0.414	0.854	0.502	0.899	0.332	0
[70, 75)	0.638	0.645	0.675	0.969	0.526	0

DOI: https://doi.org/10.25561/85146

[75, 80)	**1.000**	0.402	0.832	**1.000**	0.753	0
80+	0.873	0.107	**1.000**	0.918	**1.000**	0
Care home						
CHW	0.104	0.784	0.134	0.519	0.114	0
CHR	0.873	0.107	**1.000**	0.918	**1.000**	**1**

DOI: https://doi.org/10.25561/85146

2 Supplementary Results

2.1 Model fitting

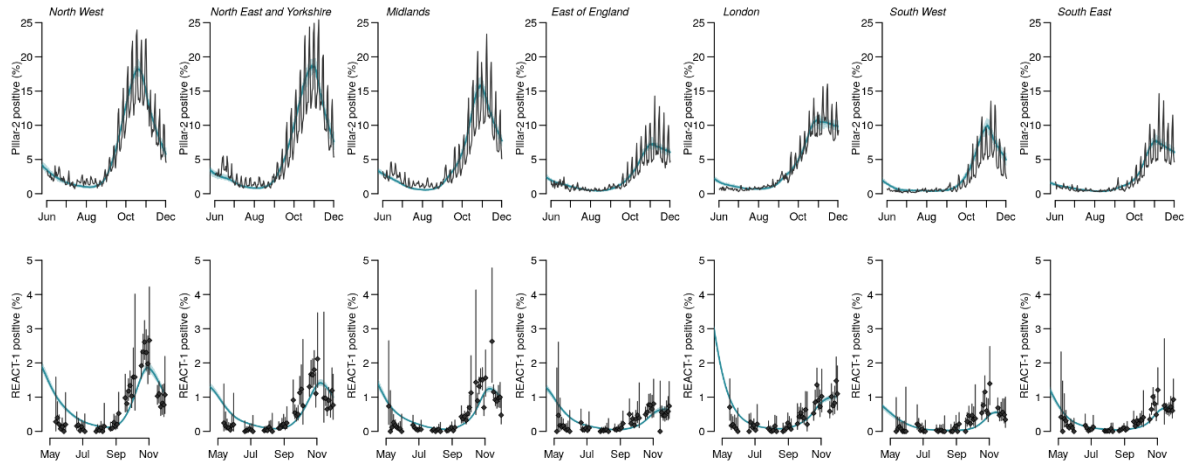

Figure S 6: Model fits by region to PCR positivity for individuals aged >25 years (top row) and PCR positivity from the REACT-1 study (bottom row). The points show the data and bars the 95% CI. The solid line the median model fit and the shaded area the 95% CrI.

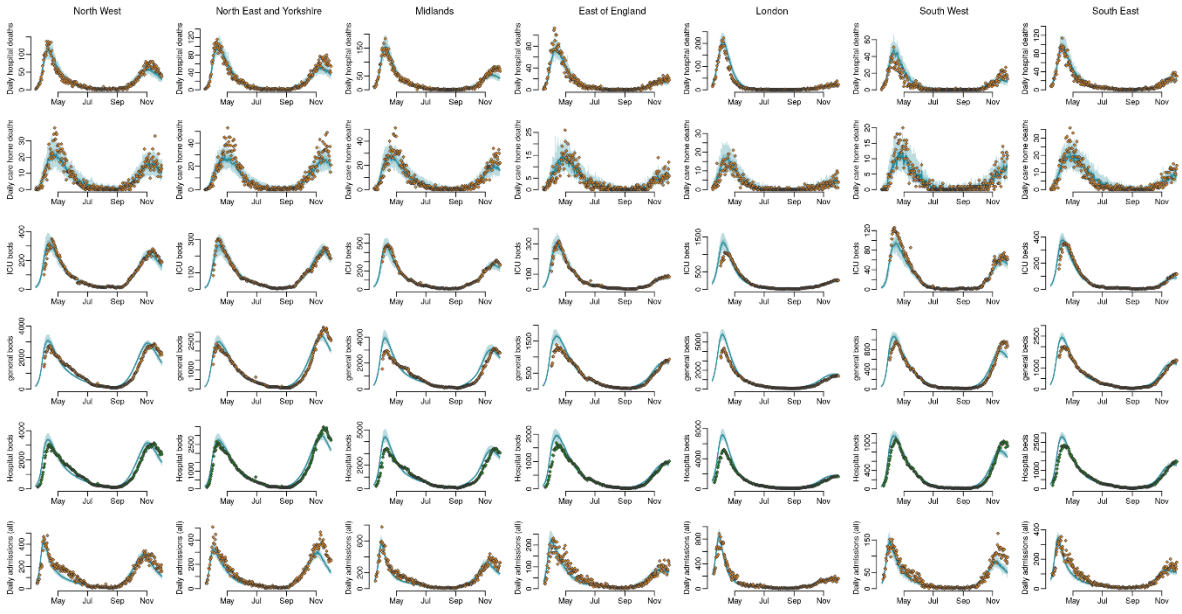

Figure S 7: Model fits to daily hospital deaths (top row), daily care home deaths (second row), ICU bed occupancy (third row), general bed occupancy (fourth row), all hospital beds (fifth row), and all daily admissions (bottom row) by region (columns). The points show the data, the solid line the median model fit and the shaded area the 95% CrI.

2.2 Severity estimates

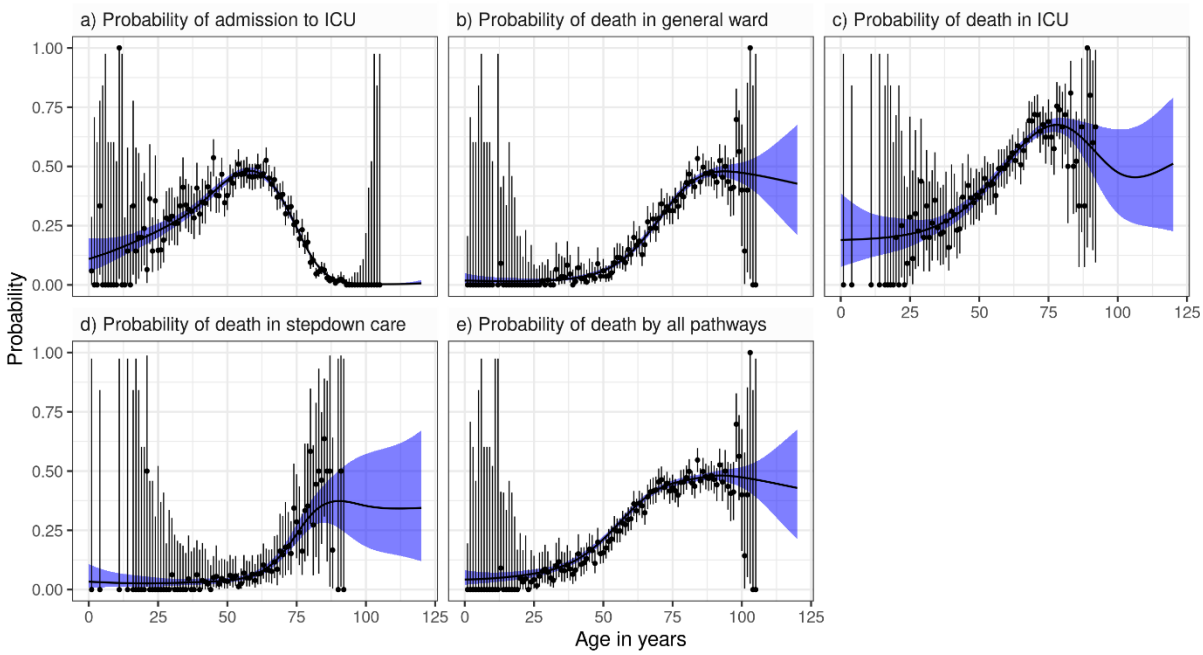

Figure S 8: Fits to CHESS data broken down into one-year age bands. Blue ribbons show the 95% CrI of the fitted spline, black circles and vertical segments give the raw mean and 95% CI from the data (exact binomial).

Table S 9: Age-stratified estimates of disease severity (*to 2sf, ^to 3dp)

	Age stratified estimate of:	
Age group	***Proportion of infections who were hospitalised****	***Proportion of infections who died^***
[0, 5)	3.0% (2.8%, 3.2%)	0.031% (0.027%, 0.034%)
[5, 10)	0.26% (0.24%, 0.28%)	0.003% (0.002%, 0.003%)
[10, 15)	0.084% (0.078%, 0.089%)	0.001% (0.001%, 0.001%)
[15, 20)	0.042% (0.039%, 0.045%)	0.000% (0.000%, 0.001%)
[20, 25)	0.080% (0.075%, 0.085%)	0.001% (0.001%, 0.001%)
[25, 30)	0.26% (0.24%, 0.28%)	0.004% (0.003%, 0.004%)
[30, 35)	0.40% (0.37%, 0.42%)	0.006% (0.006%, 0.007%)
[35, 40)	0.63% (0.58%, 0.67%)	0.013% (0.011%, 0.014%)
[40, 45)	1.2% (1.1%, 1.2%)	0.031% (0.026%, 0.035%)
[45, 50)	1.9% (1.8%, 2.1%)	0.070% (0.061%, 0.080%)
[50, 55)	2.3% (2.2%, 2.5%)	0.116% (0.101%, 0.133%)
[55, 60)	4.0% (3.8%, 4.3%)	0.276% (0.242%, 0.315%)
[60, 65)	9.6% (8.9%, 10%)	0.867% (0.762%, 0.971%)
[65, 70)	10% (9.6%, 11%)	1.215% (1.070%, 1.352%)
[70, 75)	24% (22%, 26 %)	3.512% (3.083%, 3.900%)
[75, 80)	50% (46%, 53%)	8.430% (7.407%, 9.338%)
80+	50% (47%, 54%)	9.696% (8.501%, 10.640%)
Combined	20% (13%, 27%)	34.132% (28.020%, 41.359%)

DOI: https://doi.org/10.25561/85146

2.3 Supplementary counterfactual analysis

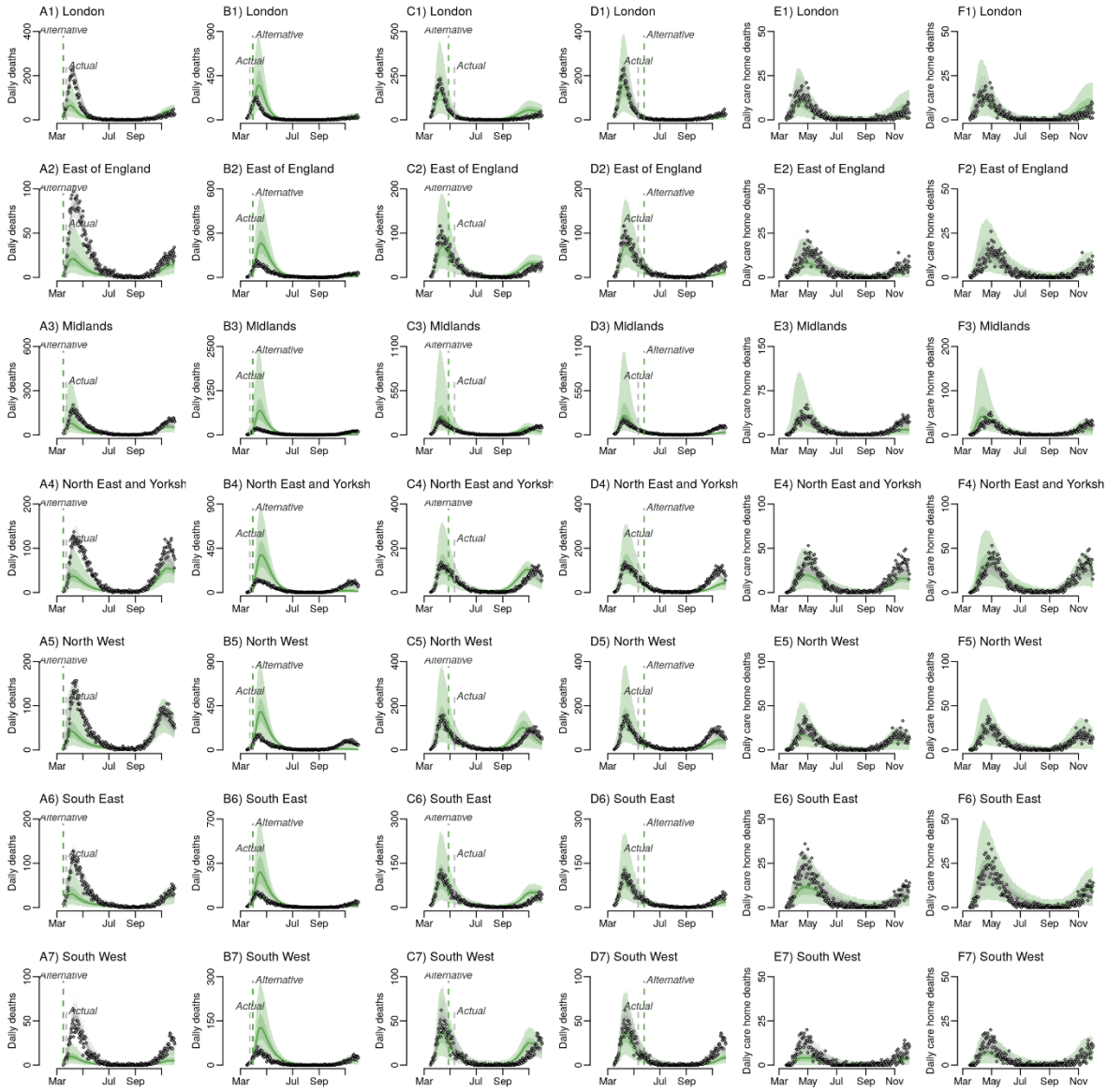

Figure S 9: Counterfactual intervention scenarios in each England NHS Region: Panel **A1**-7 impact of locking down one-week earlier Panel **B**1-7 impact of locking down one week later; Panel **C**1-7 impact of relaxing lockdown restrictions two weeks earlier. Panel **D**1-7 impact of relaxing lockdown restrictions two weeks later; Panel **E**1-7 impact of 50% less contact between care home residents and the general population; Panel **F**1-7 impact of 50% more contact between care home residents and the general population.

3. References

1. Buitrago-Garcia D, Egli-Gany D, Counotte MJ, Hossmann S, Imeri H, Ipekci AM, et al. Occurrence and transmission potential of asymptomatic and presymptomatic SARS-CoV-2 infections: A living systematic review and meta-analysis. PLoS Med. 2020;17(9):e1003346.
2. GOV.UK. Coronavirus (COVID-19) in the UK [Internet]. 2020 [cited 2020 Dec 3]. Available from: https://coronavirus.data.gov.uk/details/download
3. Riley S, Walters CE, Wang H, Eales O, Ainslie KEC, Atchison C, et al. REACT-1 round 7 updated report: regional heterogeneity in changes in prevalence of SARS-CoV-2 infection during the second national COVID-19 lockdown in England. medRxiv. 2020 Dec 16;2020.12.15.20248244.
4. Public Health England. Sero-surveillance of COVID-19 - GOV.UK [Internet]. [cited 2020 Dec 16]. Available from: https://www.gov.uk/government/publications/national-covid-19-surveillance-reports/sero-surveillance-of-covid-19
5. NHS Digital. SGSS and CHESS data - NHS Digital [Internet]. [cited 2020 Dec 4]. Available from: https://digital.nhs.uk/about-nhs-digital/corporate-information-and-documents/directions-and-data-provision-notices/data-provision-notices-dpns/sgss-and-chess-data
6. Office for National Statistics. Office for National Statistics [Internet]. [cited 2020 Dec 16]. Available from: https://www.ons.gov.uk/
7. Care Quality Commission. [ARCHIVED CONTENT] UK Government Web Archive - The National Archives [Internet]. [cited 2020 Dec 16]. Available from: https://webarchive.nationalarchives.gov.uk/20200605160439/https://www.cqc.org.uk/files/cqc-care-directory-filters-1-june-2020
8. GOV.UK. Care Homes Analysis Background. 2020.
9. Age UK. Later Life in the United Kingdom 2019 [Internet]. [cited 2020 Dec 16]. Available from: https://www.ageuk.org.uk/globalassets/age-uk/documents/reports-and-publications/later_life_uk_factsheet.pdf
10. Mossong J, Hens N, Jit M, Beutels P, Auranen K, Mikolajczyk R, et al. Social contacts and mixing patterns relevant to the spread of infectious diseases. PLoS Med. 2008;5(3):381–91.
11. Ladhani SN, Chow JY, Janarthanan R, Fok J, Crawley-Boevey E, Vusirikala A, et al. Investigation of SARS-CoV-2 outbreaks in six care homes in London, April 2020. EClinicalMedicine. 2020 Sep 1;26:100533.
12. Lauer SA, Grantz KH, Bi Q, Jones FK, Zheng Q, Meredith HR, 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):577–82.
13. Bi Q, Wu Y, Mei S, Ye C, Zou X, Zhang Z, 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;(PG-).
14. Docherty AB, Harrison EM, Green CA, Hardwick HE, Pius R, Norman L, et al. Features of 20 133 UK patients in hospital with covid-19 using the ISARIC WHO Clinical Characterisation Protocol: Prospective observational cohort study. BMJ. 2020;369(March):1–12.
15. Bernabeu-Wittel M, Ternero-Vega JE, Díaz-Jiménez P, Conde-Guzmán C, Nieto-Martín MD, Moreno-Gaviño L, et al. Death risk stratification in elderly patients with covid-19. A comparative cohort study in nursing homes outbreaks. Arch Gerontol Geriatr. 2020;91:104240.
16. Omar S, Bartz C, Becker S, Basenach S, Pfeifer S, Trapp C, et al. Duration of SARS-CoV-2 RNA

detection in COVID-19 patients in home isolation, Rhineland-Palatinate, Germany, 2020 - an interval-censored survival analysis. Eurosurveillance. 2020;25(30):1–8.

17. Benny B, Amandine G, Kc P, Sarah H, Abby M, Caitlin C, et al. Quantifying antibody kinetics and RNA shedding during early-phase SARS-CoV-2 infection.

18. Funk S. Socialmixr: Social Mixing Matrices for Infectious Disease Modelling. The Comprehensive R Archive Network; 2018.

19. GOV.UK. Prime Minister's statement on coronavirus (COVID-19): 12 March 2020 - GOV.UK [Internet]. [cited 2020 Dec 3]. Available from: https://www.gov.uk/government/speeches/pm-statement-on-coronavirus-12-march-2020

20. GOV.UK. Prime Minister's statement on coronavirus (COVID-19): 22 March 2020 - GOV.UK [Internet]. [cited 2020 Dec 3]. Available from: https://www.gov.uk/government/speeches/pm-statement-on-coronavirus-22-march-2020

21. GOV.UK. Prime Minister's statement on coronavirus (COVID-19): 25 March 2020 - GOV.UK [Internet]. [cited 2020 Dec 16]. Available from: https://www.gov.uk/government/speeches/pm-statement-on-coronavirus-25-march-2020

22. GOV.UK. Prime Minister's statement on coronavirus (COVID-19): 11 May 2020 - GOV.UK [Internet]. [cited 2020 Dec 16]. Available from: https://www.gov.uk/government/speeches/pm-statement-on-coronavirus-11-may-2020

23. GOV.UK. Prime Minister sets out timeline for retail to reopen in June - GOV.UK [Internet]. [cited 2020 Dec 16]. Available from: https://www.gov.uk/government/news/prime-minister-sets-out-timeline-for-retail-to-reopen-in-june

24. GOV.UK. Pubs, restaurants and hairdressers to reopen from 4 July - GOV.UK [Internet]. [cited 2020 Dec 16]. Available from: https://www.gov.uk/government/news/pubs-restaurants-and-hairdressers-to-reopen-from-4-july

25. GOV.UK. Eat Out to Help Out launches today – with government paying half on restaurant bills - GOV.UK [Internet]. [cited 2020 Dec 16]. Available from: https://www.gov.uk/government/news/eat-out-to-help-out-launches-today-with-government-paying-half-on-restaurant-bills

26. GOV.UK. Schools and colleges to reopen in full in September - GOV.UK [Internet]. [cited 2020 Dec 16]. Available from: https://www.gov.uk/government/news/schools-and-colleges-to-reopen-in-full-in-september

27. GOV.UK. Rule of six comes into effect to tackle coronavirus - GOV.UK [Internet]. [cited 2020 Dec 7]. Available from: https://www.gov.uk/government/news/rule-of-six-comes-into-effect-to-tackle-coronavirus

28. GOV.UK. Prime Minister announces new local COVID Alert Levels - GOV.UK [Internet]. [cited 2020 Dec 7]. Available from: https://www.gov.uk/government/news/prime-minister-announces-new-local-covid-alert-levels

29. GOV.UK. Prime Minister announces new national restrictions - GOV.UK [Internet]. [cited 2020 Dec 7]. Available from: https://www.gov.uk/government/news/prime-minister-announces-new-national-restrictions

30. Funk S. socialmixr @ github.com.

31. The RECOVERY Collaborative Group. Dexamethasone in Hospitalized Patients with Covid-19 — Preliminary Report. N Engl J Med. 2020;1–11.

32. Diekmann O, Heesterbeek JAP, Metz JAJ. On the definition and the computation of the basic reproduction ratio R0 in models for infectious diseases in heterogeneous populations. J Math Biol. 1990;28(4):365–82.

33. Gillespie DT. Approximate accelerated stochastic simulation of chemically reacting systems. J Chem Phys. 2001;115(4):1716–33.

34. Department of Health and Social Care. COVID-19 testing data: methodology note. www.gov.uk. 2020.

35. Martin IMC, Ison CA, Aanensen DM, Fenton KA, Spratt BG. Rapid Sequence-Based Identification of Gonococcal Transmission Clusters in a Large Metropolitan Area. J Infect Dis. 2004;189(8):1497–505.

36. Del Moral P, Doucet A, Jasra A. Sequential Monte Carlo samplers. J R Stat Soc Ser B Stat Methodol. 2006;68(3):411–36.

37. Andrieu C, Doucet A, Holenstein R. Particle Markov chain Monte Carlo methods. J R Stat Soc Ser B Stat Methodol. 2010;72(3):269–342.

38. Gordon NJ, Salmond DJ, Smith AFM. Novel approach to nonlinear/non-gaussian Bayesian state estimation. IEE Proceedings, Part F Radar Signal Process. 1993;140(2):107–13.

39. Baguelin M, Knock E, Whittles LK, FitzJohn R, Lees J. sircovid. 2020.

40. Knock ES, Whittles LK, Perez-Guzman PN, Bhatia S, Guntoro F, Watson OJ, et al. Reproducible parallel inference and simulation of stochastic state space models using odin, dust, and mcstate. Wellcome Open Res. 2020 Dec 11;5:288.

41. Verity R, FitzJohn R. mrc-ide/markovid at version1.5 [Internet]. [cited 2020 Dec 4]. Available from: https://github.com/mrc-ide/markovid/tree/version1.5

42. Brazeau NF, Verity R, Jenks S, Fu H, Whittaker C, Winskill P, et al. Infection Fatality Ratio: Estimates from Seroprevalence.

43. Jarvis CI, Van Zandvoort K, Gimma A, Prem K, Auzenbergs M, O'Reilly K, et al. Quantifying the impact of physical distance measures on the transmission of COVID-19 in the UK. BMC Med. 2020;18(1):1–10.

44. Lavezzo E, Franchin E, Ciavarella C, Cuomo-Dannenburg G, Barzon L, Del Vecchio C, et al. Suppression of a SARS-CoV-2 outbreak in the Italian municipality of Vo'. Nature. 2020;

45. NHS England and NHS Improvement. Statistics » COVID-19 Hospital Activity [Internet]. [cited 2020 Dec 17]. Available from: https://www.england.nhs.uk/statistics/statistical-work-areas/covid-19-hospital-activity/

DOI: https://doi.org/10.25561/85146