

Title: The 2020 SARS-CoV-2 epidemic in England: key epidemiological drivers and impact of interventions

Short title: Epidemiology of SARS-Cov-2 in England

One-sentence summary: We fit a mathematical model of SARS-CoV-2 transmission to surveillance data from England, to estimate transmissibility, severity, and the impact of interventions

Authors: Edward S. Knock^{1*}, Liliith K. Whittles^{1*}, John A. Lees^{1*}, Pablo N. Perez-Guzman^{1*}, Robert Verity¹, Richard G. FitzJohn¹, Katy AM Gaythorpe¹, Natsuko Imai¹, Wes Hinsley¹, Lucy C. Okell¹, Alicia Rosello⁴, Nikolas Kantas⁵, Caroline E. Walters¹, Sangeeta Bhatia¹, Oliver J Watson¹, Charlie Whittaker¹, Lorenzo Cattarino¹, Adhiratha Boonyasiri³, Bimandra A. Djaafara¹, Keith Fraser¹, Han Fu¹, Haowei Wang¹, Xiaoyue Xi⁵, Christl A. Donnelly^{1,6}, Elita Jauneikaite¹, Daniel J. Laydon¹, Peter J White^{1,2}, Azra C. Ghani¹, Neil M. Ferguson^{1^}, Anne Cori^{1^}, Marc Baguelin^{1,4}

Affiliations: 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

Correspondence: m.baguelin@imperial.ac.uk, neil.ferguson@imperial.ac.uk

*Equal contribution, ^Equal contribution

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

Main text: 3728 words

Abstract (125 words)

We fitted a model of SARS-CoV-2 transmission in care homes and the community to regional surveillance data for England. Among control measures implemented, only national lockdown brought the reproduction number below 1 consistently; introduced one week earlier it could have reduced first wave deaths from 36,700 to 15,700 (95%CrI: 8,900–26,800). Improved clinical care reduced the infection fatality ratio from 1.25% (95%CrI: 1.18%–1.33%) to 0.77% (95%CrI: 0.71%–0.84%). The infection fatality ratio was higher in the elderly residing in care homes (35.9%, 95%CrI: 29.1%–43.4%) than those residing in the community (10.4%, 95%CrI: 9.1%–11.5%). England is still far from herd immunity, with regional cumulative infection incidence to 1st December 2020 between 4.8% (95%CrI: 4.4%–5.1%) and 15.4% (95%CrI: 14.9%–15.9%) of the population.

1 Introduction

England is among the countries worst-affected by the global pandemic of COVID-19, caused by the novel *Betacoronavirus* SARS-CoV-2. As of 2nd December 2020, over 51,000 deaths have been reported nationally, or 91 deaths per 100,000 people (1). The impact of the epidemic has varied across the country, with regional epidemics differing in their severity and timing. A key feature in all regions is the burden suffered by older adults living in care homes, where mortality has been high.

We use a mathematical model of SARS-CoV-2 transmission to reproduce the first two waves of the epidemic across England's seven NHS regions and assess the impact of interventions implemented by the UK government. We analyse the epidemic from importation of SARS-CoV-2 into each region to the 2nd December 2020: encompassing the first national lockdown from March – May, the interventions implemented as COVID-19

deaths increased again in the autumn, eventually leading to the second national lockdown in November.

We built an age-structured stochastic transmission model of SARS-CoV-2, representing care homes, hospital clinical pathways and the wider community (Materials and Methods). We developed a Bayesian evidence-synthesis approach to estimate model parameters and to reconstruct regional epidemics using data from daily recorded deaths, PCR testing, hospital admissions, hospital bed occupancy, individual patient outcomes, contact surveys, and serological surveys. We evaluated temporal changes in transmission as new control measures were implemented and then relaxed, and population immunity accrued. Inclusion of serological data allowed us to robustly estimate region- and age-specific disease severity, to compare severity in care home residents to elderly individuals in the community, and estimate the total epidemic size, by calculating the proportion of individuals infected over time in each region. Finally, we examined counterfactual epidemic scenarios, varying the date and duration of the first national lockdown and the effectiveness of restricting care home visits, to quantify the resulting impact on mortality.

Our analysis, which synthesises multiple data sources and parametrically accounts for their biases, provides a comprehensive overview of transmission, hospitalisation, and mortality patterns of SARS-CoV-2 in the first and second waves (up to 2nd December) in all regions of England. Our results provide crucial insights for controlling the epidemic in the future, emphasising the importance of acting fast to save lives.

2 Results

2.1 Epidemic trajectory

We used our evidence-synthesis approach, to infer the COVID-19 epidemic start date in each NHS England region (Figure 1A), then reconstructed epidemic trajectories for hospitalisations (Figure S7) and deaths in care homes and hospitals (Figure 1B-H). We estimated the basic reproduction number, R_0 , defined as the expected number of onward infections from an infectious individual in a fully susceptible population to be 2.9 (95% CrI: 2.8-3.1) nationally. Figure 1I shows how the effective reproduction number R_t^{eff} (the expected number of onward infections from an individual infected at time t) changed in each region over time, in relation to government control measures and accrual of population immunity.

The first COVID-19 death in England occurred on 5th March 2020 (2). Seven days later, in response to the growing epidemic, the government began to introduce control measures, initially requiring individuals with a dry persistent cough and/or fever to self-isolate (3). On 23rd March this escalated to a full national lockdown (4, 5). Irrespective of initial differences, the level of transmission during lockdown was similar across all regions (Figure 1I), consistent with mobility data showing movement during lockdown reduced to a consistent level nationally (6).

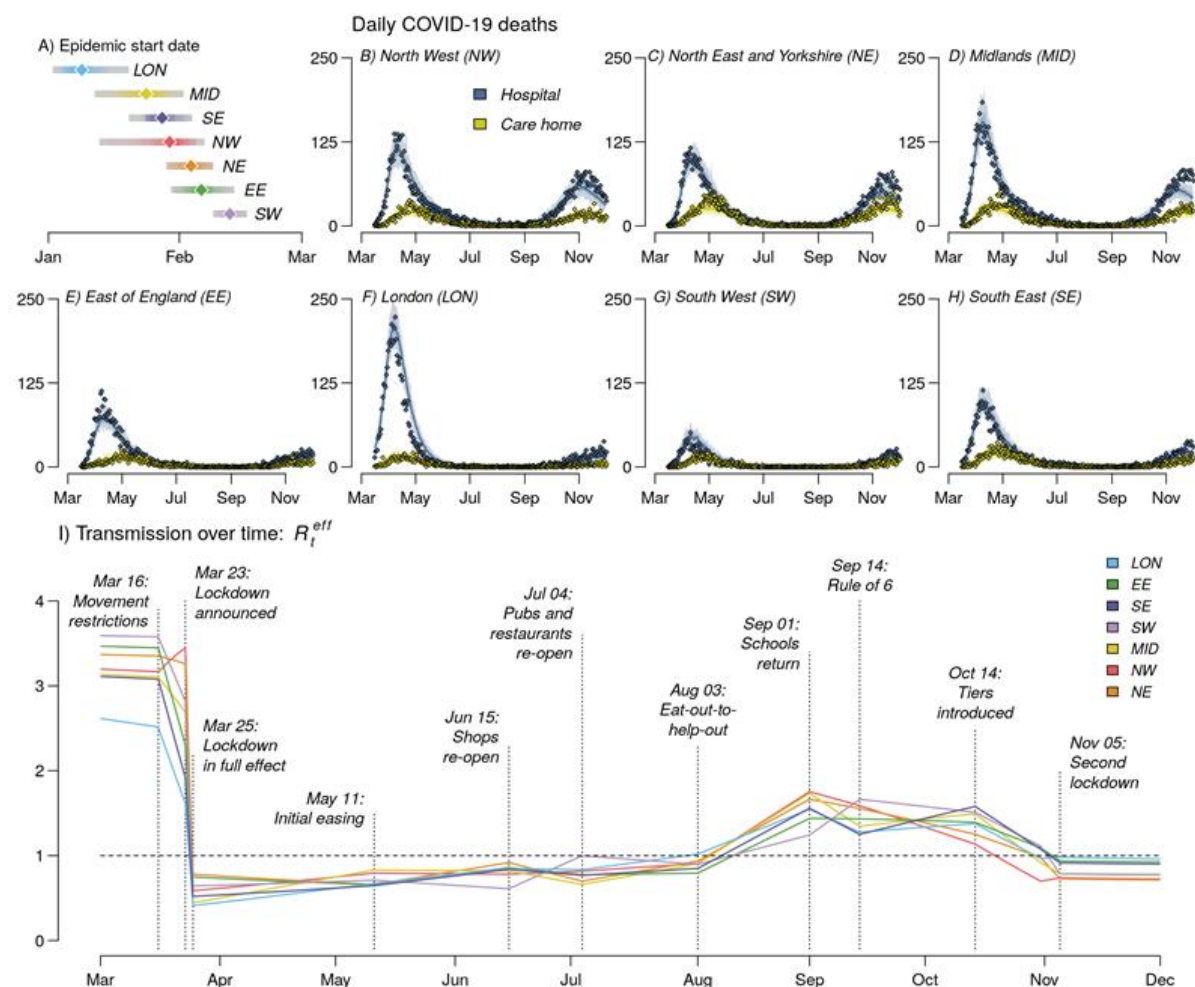


Figure 1: Trajectory of the England COVID-19 epidemic. **A**, The inferred epidemic start date in each NHS England region. **B-H**, The model fit to reported daily deaths from COVID-19 in care homes and hospitals for each NHS England region. The points show the daily data, solid lines the median posterior and the shaded area shows the 95% CrI. **I**, The mean effective reproduction number within the general community (i.e. excluding care homes) in each region from March to December. Vertical lines and labels represent dates of key policy changes, defining the breaking points of the underlying piecewise linear transmission rate. Dashed horizontal line depicts reproduction number of 1.

The epidemic in London began 15 days before (95% CrI: 28 days before, 3 days after) the rest of the country (Figure 1A), meaning the lockdown occurred at a later stage of its epidemic. London experienced a mortality of 88.5 (95% CrI: 79.9–95.3) per 100,000 during the first wave, compared to the national average of 70.7 (95% CrI: 64.6–77.1), despite

having a younger population and a smaller care home population than other regions (296 vs 603 per 100,000 nationally).

A key feature of the first epidemic wave in England, in common with other European countries, was the high death toll within care homes, which accounted for 22.6% laboratory-confirmed COVID-19 deaths in England as of 1st August 2020. Although community transmission rates fell during lockdown, transmission within care homes continued to rise, with infection risk peaking in care home residents, between 26th March in London and 12th April in North East and Yorkshire (Figure 2A). Deaths in care homes peaked on average 13 days later than hospital deaths (Figure 1B-H).

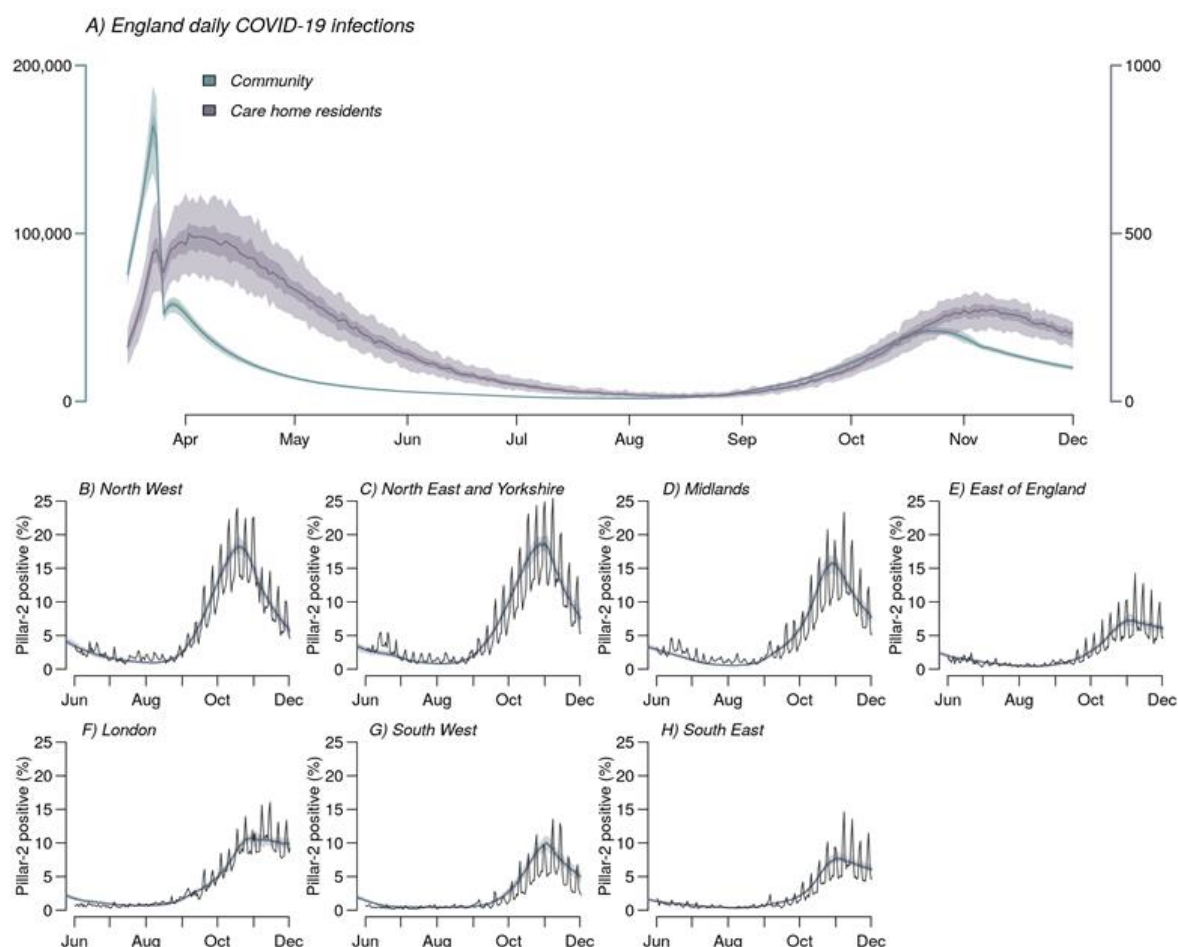


Figure 2: A, Inferred daily SARS-CoV-2 infections in England care home residents (right axis) and the wider community (left axis). **B-H,** Comparison of modelled (shaded bands) and observed (solid line) proportion of PCR tests that are positive, under pillar-2 testing (community swab testing for symptomatic individuals) in >25 year olds. Shaded bands depict 95% CrI, 50% CrI and median model outputs.

The first lockdown in England continued until 11th May, when people unable to work remotely were permitted to resume their jobs. Over the summer restrictions were successively eased, with non-essential shops, pubs and restaurants opening, followed by the government's 'Eat Out to Help Out' restaurant subsidy scheme in August (7). This led to a steady increase in transmission, with R_t^{eff} rising above 1 in all regions by mid-August (Figure 1I).

Increasing PCR test positivity marked the beginning of a second epidemic wave (Figure 2B-H, S6). The accompanying introduction of non-pharmaceutical interventions (NPIs) began with the “Rule of Six” (limiting social gatherings to 6 persons maximum) on 14th September (8), followed by the localised tiered restrictions on 14th October (9). These measures limited transmission in most regions but were not sufficient to reduce R_t^{eff} below 1 (Figure 11). Consequently, on 31st October, the government announced a second national lockdown, which lasted from 5th November to 1st December (10).

Restrictions during the second lockdown were less stringent than the first, with schools and some workplaces remaining open. This was reflected in R_t^{eff} estimates of 0.83 (95% CrI: 0.81–0.85) at the start of the second lockdown, compared to $R_t^{eff} = 0.54$ (95% CrI: 0.50–0.59) at the start of the first. We estimate that without the population immunity accrued during the first wave, contact rates during the second lockdown would have resulted in a reproduction number of $R_t = 0.95$ (95% CrI: 0.93–0.98). Hence, population immunity helped to reduce transmission further below the critical threshold of $R_t^{eff} = 1$.

2.2 Severity and hospitalisation

COVID-19 manifests a broad spectrum of severity, from asymptomatic infection to life-threatening illness requiring intensive care. We estimated age-patterns of clinical progression in people admitted to hospital using individual-level data from 17,702 patients admitted between 18th March and 31st May 2020 (inclusive) in the COVID-19 Hospitalisation in England Surveillance System (CHESS, (11)) (Materials and Methods). We derived estimates of the time spent in each stage of the hospital pathway (including general wards, ICU and post-ICU stepdown care), as well as age-stratified probabilities of progression through that pathway (Figure 3 and Figure S8). We accounted for differing length of stays given different outcomes; there were marked differences in average length of ICU stay for

those who died in ICU, those who later died in stepdown care and those who were discharged following stepdown care (Figure 3F). Among patients over 65, we found the probability of admission to ICU decreased with increasing age. Severity of COVID-19 increases with age, but for older patients and those with most severe illness, the benefit of ICU admission, ventilation and the corresponding prognosis may not be better than with oxygen therapy in a general ward (12). Thus, older and more severely infected patients may be directed to care on a general ward rather than admitted to ICU.

We used estimates of clinical progression to parametrise the transmission model, enabling us to infer temporal and regional differences in disease severity, informed by local demography, observed daily hospital admissions, bed occupancy and deaths. We measured severity of disease by the infection fatality ratio (IFR) and the infection hospitalisation ratio (IHR).

Patient progression in hospital

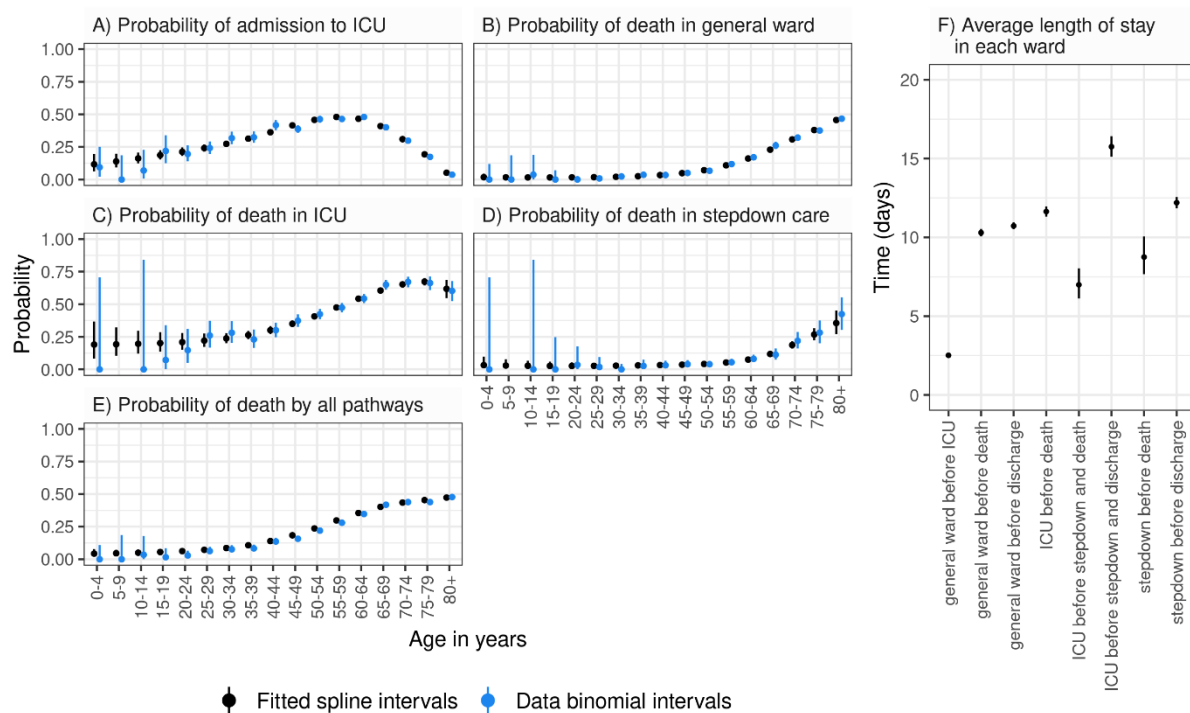


Figure 3: Age-dependent probabilities of progression through hospital pathways. A, Probability of admission to ICU. B, Probability of death in a general ward. C, Probability of death in ICU. D, Probability of death in stepdown care. E, Probability of death through all hospital pathways. Black circles and vertical segments show posterior mean and 95% credible intervals of splines fitted to data, blue circles and vertical segments show raw mean values and 95% confidence intervals (exact binomial) for each 5-year age group. F, Average length of stay in each ward (posterior mean and 95% credible intervals).

The severity of disease increased with age in all regions with the steepest increase above 65 years (Figure 4A-C), in line with observations worldwide (5). Regional estimates of age-aggregated disease severity depend on the population age distribution, which is similar in most regions of the country, except London, where the median age is 34.6 years (vs 39.5 years nationally). At the start of the first wave, London experienced an IFR (respectively IHR) of 0.91% (95% CrI: 0.82%–1.00%) (resp. 3.02%; 95% CrI: 2.82%–3.19%) compared to the national average of 1.25% (95% CrI: 1.18%–1.33%) (resp. 3.52%; 95% CrI: 3.29%–3.72%) (Figure 4D-E).

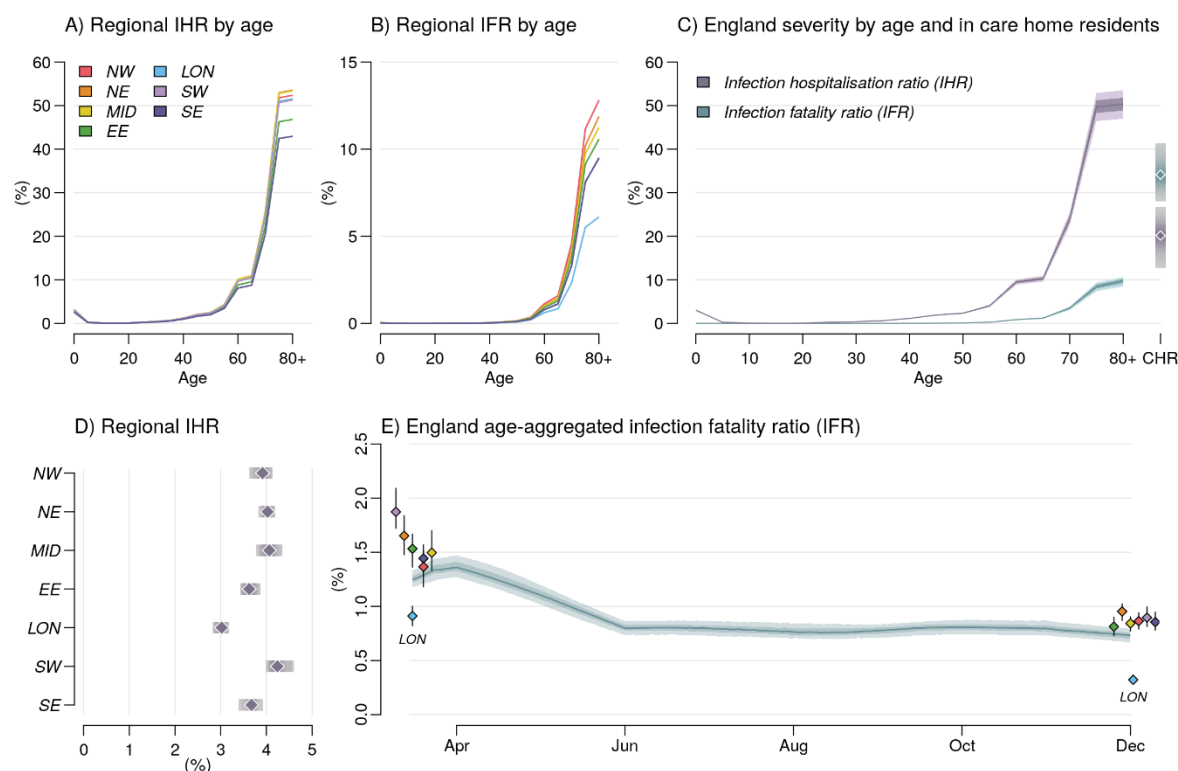


Figure 4: Relative severity of disease by age group and region. A, and B, Variation in the Infection fatality ratio (IFR) and Infection Hospitalisation Ratio (IHR) by age group in each region. Ages 80+ were modelled as a single risk group, care home residents are not included. C, The England IFR and IHR by age group and in care home residents (estimates denoted CHR at the right-hand side of the panel). National severity estimates are produced by aggregating regional estimates based on infection incidence. D, The regional IHR, aggregated over age and risk group by infection incidence. Plots a-d use parameter estimates, and incidence weightings calculated as of 1st December 2020. E, The England IFR over time, coloured dots show regional estimates of IFR at the start of the epidemic and on 1st December 2020 (clusters each correspond to one time-point, LON: London). In plots C-E Shaded bands depict 95% CrI and interquartile ranges, points depict medians.

Regional variation in the population age distribution did not fully account for differences in severity, with London still experiencing lower mortality when stratified by age (Figure 4A-B). The oldest age group (80+) in London had an IFR of 6.1% (95% CrI: 5.2%–6.8%) compared to 12.7% (95% CrI: 10.8%–14.3%) in the North West.

We estimated temporal trends in the IFR for England, by weighting regional estimates by incidence and population demographics. At the start of the first wave, the national IFR was 1.25% (95% CrI: 1.18%–1.33%) (Figure 4E), consistent with earlier reports from serology data alone (13). The national IFR initially appeared to increase, as transmission widened from London to regions with older populations and greater disease severity. Over the first wave, the proportion of hospital admissions resulting in death decreased, due to improvements in clinical management and alleviation of capacity constraints (14), leading to a national IFR of 0.77% (95% CrI: 0.71%–0.84%) by the end of the first wave. The magnitude of the relative reduction in IFR over time varied between regions, from 36.5% (95% CrI: 26.5%–47.5%) in the North West to 64.6% (95% CrI: 58.6%–68.8%) in London.

The IFR was greater among care home residents (35.9%, 95% CrI: 29.1%–43.4%) than in the 80+ in the community (10.4%, 95% CrI: 9.1%–11.5%, Figure 4C). Many care home residents did not transfer into hospital, and instead died in the facilities where they lived, so conversely the IHR was lower in care home residents (19.1%, 95% CrI: 11.5%–26.8%) than in those aged 80+ in the community (51.1%, 95% CrI: 47.6%–54.3%). We present national estimates of severity at the end of the second wave, stratified by age and care home residency in Table S9.

2.3 Epidemic size

Data from repeated serological surveys of blood donors aged 15–65 informed our estimation of the total regional epidemic size (Figure 5A–G), accounting for imperfect sensitivity and specificity of serological tests (Materials and Methods) (15). The cumulative proportion of the population ever infected with SARS-CoV-2 ranged from 4.8% (95% CrI: 4.4%–5.1%) in the South West to 15.4% (95% CrI: 14.9%–15.9%) in London (Figure 5H). Predicted seropositivity was initially greater than cumulative incidence, due to imperfect test specificity.

The increase in seropositivity lagged cumulative infections by two weeks, reflecting the time from infection to seroconversion.

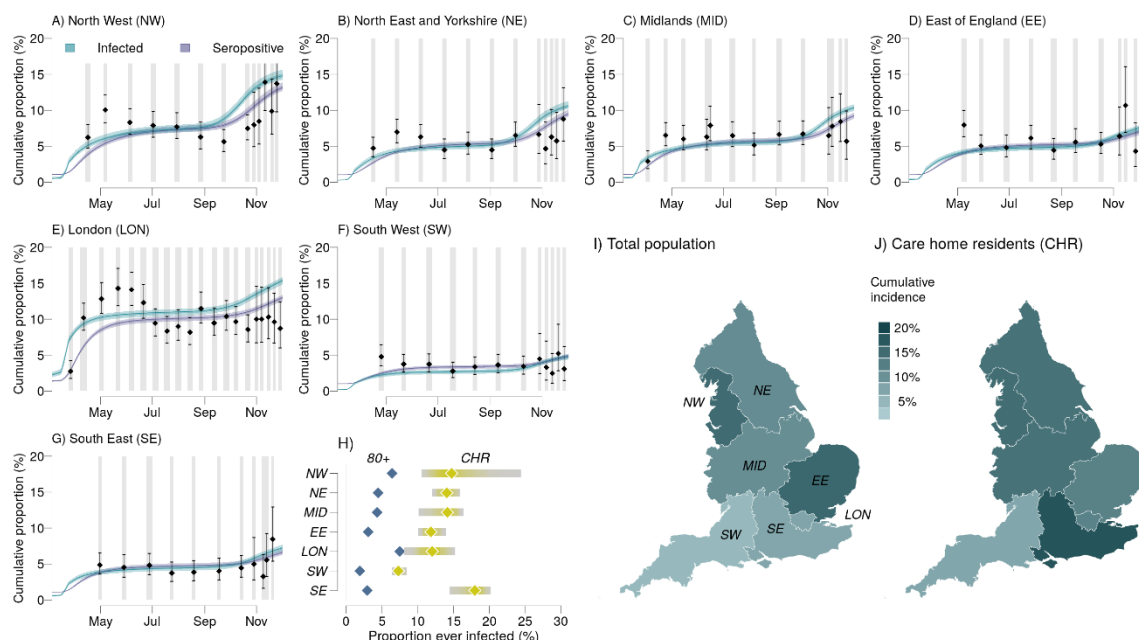


Figure 5: Cumulative incidence and seropositivity by region. A-G, Comparison of the estimated proportion of the population testing seropositive with observations from serological surveys. Vertical grey shaded bands show serological survey timings, black points the observed seroprevalence (bars: 95% exact confidence intervals), blue and purple lines show estimated proportion of the population infected and seropositive respectively (shaded bands the 95% CrI, 50% CrI and median). **H,** Comparison by region of the estimated cumulative attack rate in care home residents vs in the 80+ age group in the community (median, 95% CrI). The final epidemic size in each England NHS region **I)** in total and **J)** in care home residents.

Seropositivity notably declined following the first wave in some regions (Figure 5A-G). This may reflect antibody waning (16), or temporal trends in the composition of the surveyed population. Lockdown restrictions made attending blood donation centres difficult for all except key workers, who were more likely to have been infected (17), and may therefore be overrepresented in the sample of blood donors during the two lockdowns.

The proportion of care home residents ever infected with SARS-CoV-2 was 13.7% (95% CrI: 10.7%–16.7%), much higher than the 4.2% (95% CrI: 4.0%–4.4%) estimated in >80-year olds living in the community. This difference was consistently observed across all regions (Figure 5H). Regional differences in care home attack rates mirrored the patterns seen in the general community, with regions with larger community epidemics also experiencing larger care home epidemics (Figure 5I,J).

2.4 Impact of non-pharmaceutical interventions (NPIs)

We explored counterfactual intervention scenarios and examined the potential impact on mortality of initiating the first national lockdown one week earlier or later; ending that lockdown two weeks earlier or later; and 50% more or less restricted care home visits throughout the epidemic (Figure 6).

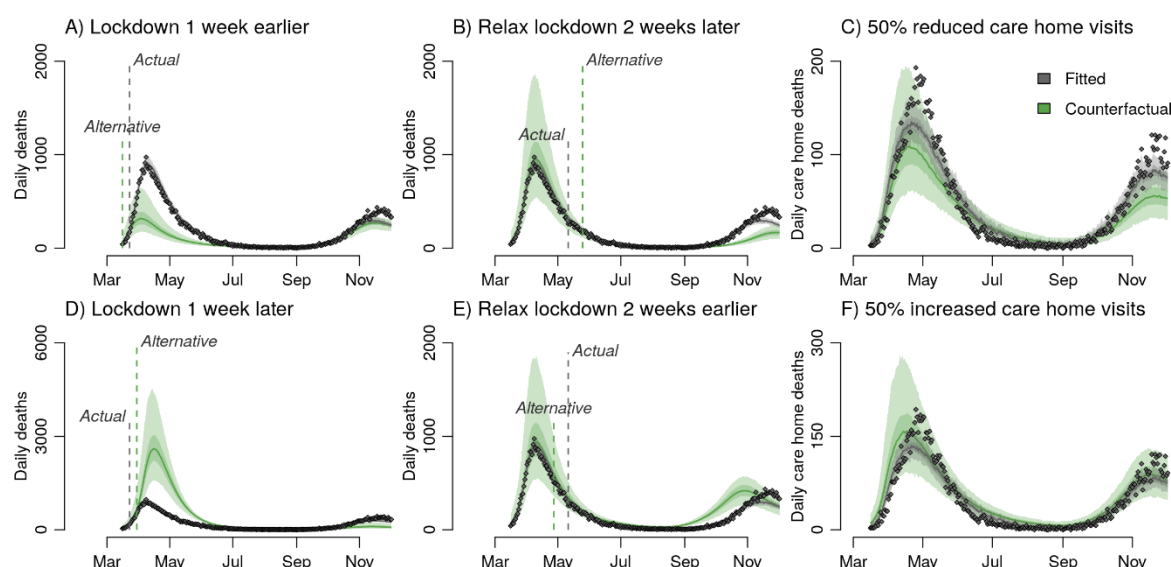


Figure 6: Counterfactual analysis of the impact on mortality aggregated across NHS England regions of A, D, initiating lockdown one week earlier / later, B, E Relaxing lockdown two weeks earlier / later, and C, F 50% more / less restricted care home visits. Panels A, B, D and E all present counterfactual outcomes for daily deaths in England but have different y-axis scales to better highlight differences between the observed data and each alternative lockdown scenario.

275

276 The timing of the initial national lockdown was crucial in determining the eventual epidemic
 277 size in England. Locking down a week earlier could have reduced the first wave death toll
 278 (up to 1st July 2020) from 36,700 to 15,700 (95% CrI: 8,900–26,800) while delaying
 279 lockdown by a week would have increased the deaths to 102,600 (95% CrI: 66,400–
 280 154,800) (Figure 6A, D). The impact varied by region, with regions with less established
 281 epidemics at the time of the first lockdown more sensitive to the timing of the intervention
 282 (Figure S10). Locking down a week later may have increased deaths, with large variability
 283 by region, from 105% in London to 274% in the Midlands but with very large uncertainty
 284 (Figure S9). Initiating a lockdown to interrupt the exponential growth phase of an epidemic
 285 has a much greater impact on reducing total mortality than extending an existing lockdown.
 286 Due to this asymmetry, relaxing the lockdown measures two weeks earlier (respectively
 287 later), could have increased deaths by 9,300 (95% CrI: 700–17,000) (respectively prevented
 288 9,800 (95% CrI: 7,400–12,100) deaths) prior to 2nd December (Figure 6B, D).

289

290 We also explored counterfactual scenarios varying the level of visit restriction in care homes
 291 and estimated that reducing contact between the general population and care home
 292 residents by 50% could have reduced care home deaths by 44% (95% CrI: 17%–64%)
 293 (Figure 6C).

294

295 **3 Discussion**

296

297 We present a comprehensive overview of SARS-CoV-2 transmission, hospitalisation,
 298 mortality and intervention impact in the first two epidemic waves across all regions of
 299 England between March and December 2020. We successfully reproduce the transmission
 300 dynamics of the two epidemic waves, in terms of cases, PCR prevalence, seroprevalence,
 301 hospitalised cases (general wards and ICU), and deaths in hospitals and in care homes.

302

303 We estimate intense transmission in care homes even during the first national lockdown
 304 when R_t^{eff} in the community was well below one in all regions (Figure 2) (18–20). Combined
 305 with our counterfactual analysis of restricting visits (Figure 6) this suggests that reducing
 306 infection levels in care home residents is challenging. This highlights the difficulty of
 307 protecting care home residents from COVID-19: due to the necessarily close contact
 308 between staff and residents within a care home, once a care home outbreak has begun it is
 309 very difficult to reduce transmission, which overrides any impact of reducing the number of
 310 introductions (21, 22).

311

312 We find that, consistent with existing literature (23), disease severity increases with age.
 313 Assessment of severity is complicated by the broad clinical spectrum of COVID-19 (24–26)
 314 hence, recent published estimates are still based on data from early in the pandemic (27).
 315 Here we provide updated severity estimates based on multiple contemporary data streams.
 316 We estimate considerable regional heterogeneity in severity, broadly consistent in the
 317 general population and in care homes for IFR and IHR. London experienced the lowest
 318 severity even after adjusting for its younger population. The estimated two-fold reduction
 319 over time in IFR (Figure 4) cannot be explained solely by the introduction of dexamethasone
 320 which reduces mortality amongst ICU patients (28), but rather a combination of factors
 321 including improvements in clinical management, greater experience in treating patients in
 322 ICU, and alleviation of capacity constraints (14, 29).

323

324 Our analysis shows large regional variation in burden, especially in the first wave. This is
 325 likely due to the pattern of seeding and the timing of lockdown relative to how advanced
 326 each region's epidemic was (Figure 1A). Our counterfactual scenarios of initiating the first
 327 national lockdown one week earlier or later highlight the importance of early interventions to
 328 reduce overall mortality (Figure 6).

329

330 The extent and duration of infection-induced immunity to SARS-CoV-2 and its relationship to
331 seropositivity remains unclear. Related seasonal coronaviruses induce immunity that wanes
332 in one or two years (30), though antibody titres following SARS-CoV-1 infection appear to
333 decay more slowly (31). Our estimated cumulative incidence over time (Figure 5), strongly
334 supports the hypothesis that the epidemic decline after the first national lockdown was due
335 to NPIs, with immunity playing a minimal role (32). Population-level immunity was insufficient
336 to prevent a second wave of infection in any region (Figure 1), illustrated by the increase in
337 reported cases and deaths which prompted the second national lockdown (33).

338

339 With the authorisation of the first SARS-CoV-2 vaccines in December 2020, we are now
340 entering a new phase in the control of the COVID-19 pandemic. However, our estimates of
341 current population immunity are low, with regional cumulative attack rates ranging from 4.8%
342 to 15.4%, therefore any vaccination campaign will need to achieve high coverage and high
343 levels of protection in vaccinated individuals to allow NPIs to be lifted without a resurgence
344 of transmission. While vaccinating the most vulnerable age and risk groups will considerably
345 reduce the burden of COVID-19, a large proportion of younger age groups may also need to
346 be vaccinated to reach the immunity threshold for control. Our high estimates of
347 transmission in care homes imply that vaccine uptake there will need to be especially high,
348 particularly if vaccine efficacy is lower amongst older age groups.

349

350 We make a number of simplifying assumptions in our analysis. First, due to the
351 compartmental nature of our model, we do not explicitly model individual care homes, rather
352 the regional care home sector as a whole. However, as care home workers may work across
353 multiple facilities leading to within and between care home transmission, we do not expect
354 the simplification to substantially affect our conclusions. Similarly, we do not model individual
355 households or transmission within and between them. When assessing the impact of NPIs

on transmission we therefore capture population averages, rather than the contribution of household and non-household contacts. Second, hospital-acquired infections may have contributed to overall transmission, especially around the peak of the epidemic, and to persistence of infection in England over the summer months (34, 35). Our model does not explicitly represent nosocomial transmission; therefore such effects will be encompassed within our regional R_t^{eff} estimates. Third, each data stream was subject to competing biases, which we statistically accounted for as far as possible (supplement section 1.1.2). A key strength of our evidence-synthesis approach is that we do not rely on any single data source, combining multiple perspectives to provide a robust overall picture of the epidemic. Finally, we model the epidemics in each NHS region in England independently without accounting for spatial effects across regional boundaries, or spatial heterogeneity within regions. This spatial scale was determined by the data and reflects limited movement between regions due to travel restrictions but does allow for movement within regions.

Our analysis provides a comprehensive overview of transmission, hospitalisation, and mortality patterns of COVID-19 in the first and second waves of the epidemic in all regions of England, one of the countries worst-affected by the pandemic. Integration of multiple data streams into a single cohesive modelling framework, enables us to disentangle transmission and severity from features of the surveillance system and provide robust estimates of the epidemiological characteristics of the COVID-19 epidemic in England. As nationwide vaccination programmes are rolled out, our results will help to inform how NPIs are applied in the future.

References

1. GOV.UK, Coronavirus (COVID-19) in the UK (2020), (available at <https://coronavirus.data.gov.uk/details/download>).
2. E. Mahase, Covid-19: UK records first death, as world's cases exceed 100 000. *BMJ*. **368** (2020), doi:10.1136/bmj.m943.
3. GOV.UK, Prime Minister's statement on coronavirus (COVID-19): 12 March 2020 - GOV.UK, (available at <https://www.gov.uk/government/speeches/pm-statement-on-coronavirus-12-march-2020>).
4. GOV.UK, Prime Minister's statement on coronavirus (COVID-19): 20 March 2020 - GOV.UK, (available at <https://www.gov.uk/government/speeches/pm-statement-on-coronavirus-20-march-2020>).
5. GOV.UK, Prime Minister's statement on coronavirus (COVID-19): 22 March 2020 - GOV.UK, (available at <https://www.gov.uk/government/speeches/pm-statement-on-coronavirus-22-march-2020>).
6. B. Jeffrey, C. E. Walters, K. E. C. Ainslie, O. Eales, C. Ciavarella, S. Bhatia, S. Hayes, M. Baguelin, A. Boonyasiri, N. F. Brazeau, G. Cuomo-Dannenburg, R. G. FitzJohn, K. Gaythorpe, W. Green, N. Imai, T. A. Mellan, S. Mishra, P. Nouvellet, H. J. T. Unwin, R. Verity, M. Vollmer, C. Whittaker, N. M. Ferguson, C. A. Donnelly, S. Riley, Anonymised and aggregated crowd level mobility data from mobile phones suggests that initial compliance with covid-19 social distancing interventions was high and geographically consistent across the UK. *Wellcome Open Res.* **5**, 1–10 (2020).
7. HM Revenue & Customs, Get a discount with the Eat Out to Help Out Scheme. *www.gov.uk* (2020).
8. GOV.UK, Rule of six comes into effect to tackle coronavirus - GOV.UK, (available at <https://www.gov.uk/government/news/rule-of-six-comes-into-effect-to-tackle-coronavirus>).

- 406 9. GOV.UK, Prime Minister announces new local COVID Alert Levels - GOV.UK,
407 (available at [https://www.gov.uk/government/news/prime-minister-announces-new-](https://www.gov.uk/government/news/prime-minister-announces-new-local-covid-alert-levels)
408 [local-covid-alert-levels](https://www.gov.uk/government/news/prime-minister-announces-new-local-covid-alert-levels)).
- 409 10. GOV.UK, Prime Minister announces new national restrictions - GOV.UK, (available at
410 [https://www.gov.uk/government/news/prime-minister-announces-new-national-](https://www.gov.uk/government/news/prime-minister-announces-new-national-restrictions)
411 [restrictions](https://www.gov.uk/government/news/prime-minister-announces-new-national-restrictions)).
- 412 11. NHS Digital, SGSS and CHESS data - NHS Digital, (available at
413 [https://digital.nhs.uk/about-nhs-digital/corporate-information-and-](https://digital.nhs.uk/about-nhs-digital/corporate-information-and-documents/directions-and-data-provision-notices/data-provision-notices-dpns/sgss-and-chess-data)
414 [documents/directions-and-data-provision-notices/data-provision-notices-dpns/sgss-](https://digital.nhs.uk/about-nhs-digital/corporate-information-and-documents/directions-and-data-provision-notices/data-provision-notices-dpns/sgss-and-chess-data)
415 [and-chess-data](https://digital.nhs.uk/about-nhs-digital/corporate-information-and-documents/directions-and-data-provision-notices/data-provision-notices-dpns/sgss-and-chess-data)).
- 416 12. NHS, Overview | COVID-19 rapid guideline: critical care in adults | Guidance | NICE.
- 417 13. N. F. Brazeau, R. Verity, S. Jenks, H. Fu, C. Whittaker, P. Winskill, I. Dorigatti, P.
418 Walker, S. Riley, R. P. Schnekenberg, H. Hoeltgebaum, T. A. Mellan, S. Mishra, H. T.
419 Juliette Unwin, O. J. Watson, Z. M. Cucunubá, M. Baguelin, L. Whittles, S. Bhatt, A.
420 C. Ghani, N. M. Ferguson, L. C. Okell, Infection Fatality Ratio: Estimates from
421 Seroprevalence, doi:10.25561/83545.
- 422 14. R. A. Armstrong, A. D. Kane, T. M. Cook, Outcomes from intensive care in patients
423 with COVID-19: a systematic review and meta-analysis of observational studies.
424 *Anaesthesia*. **75**, 1340–1349 (2020).
- 425 15. Public Health England, Sero-surveillance of COVID-19 - GOV.UK, (available at
426 [https://www.gov.uk/government/publications/national-covid-19-surveillance-](https://www.gov.uk/government/publications/national-covid-19-surveillance-reports/sero-surveillance-of-covid-19)
427 [reports/sero-surveillance-of-covid-19](https://www.gov.uk/government/publications/national-covid-19-surveillance-reports/sero-surveillance-of-covid-19)).
- 428 16. F. J. Ibarrondo, J. A. Fulcher, D. Goodman-Meza, J. Elliott, C. Hofmann, M. A.
429 Hausner, K. G. Ferbas, N. H. Tobin, G. M. Aldrovandi, O. O. Yang, Rapid Decay of
430 Anti–SARS-CoV-2 Antibodies in Persons with Mild Covid-19. *N. Engl. J. Med.*, 1–2
431 (2020).
- 432 17. L. H. Nguyen, D. A. Drew, M. S. Graham, A. D. Joshi, C. G. Guo, W. Ma, R. S. Mehta,

- E. T. Warner, D. R. Sikavi, C. H. Lo, S. Kwon, M. Song, L. A. Mucci, M. J. Stampfer, W. C. Willett, A. H. Eliassen, J. E. Hart, J. E. Chavarro, J. W. Rich-Edwards, R. Davies, J. Capdevila, K. A. Lee, M. N. Lochlainn, T. Varsavsky, C. H. Sudre, M. J. Cardoso, J. Wolf, T. D. Spector, S. Ourselin, C. J. Steves, A. T. Chan, C. M. Albert, G. Andreotti, B. Bala, B. A. Balasubramanian, L. E. Beane-Freeman, J. S. Brownstein, F. J. Bruinsma, J. Coresh, R. Costa, A. N. Cowan, A. Deka, S. L. Deming-Halverson, M. Elena Martinez, M. E. Ernst, J. C. Figueiredo, P. Fortuna, P. W. Franks, L. B. Freeman, C. D. Gardner, I. M. Ghobrial, C. A. Haiman, J. E. Hall, J. H. Kang, B. Kirpach, K. C. Koenen, L. D. Kubzansky, J. V Lacey, L. Le Marchand, X. Lin, P. Lutsey, C. R. Marinac, M. E. Martinez, R. L. Milne, A. M. Murray, D. Nash, J. R. Palmer, A. V Patel, E. Pierce, M. M. Robertson, L. Rosenberg, D. P. Sandler, S. H. Schurman, K. Sewalk, S. V Sharma, C. J. Sidey-Gibbons, L. Slevin, J. W. Smoller, C. J. Steves, M. I. Tiirikainen, S. T. Weiss, L. R. Wilkens, F. Zhang, Risk of COVID-19 among front-line health-care workers and the general community: a prospective cohort study. *Lancet Public Heal.* **5**, e475--e483 (2020).
18. GOV.UK, COVID-19: number of outbreaks in care homes - management information, (available at <https://www.gov.uk/government/statistical-data-sets/covid-19-number-of-outbreaks-in-care-homes-management-information#history>).
19. S. N. Ladhani, J. Y. Chow, R. Janarthanan, J. Fok, E. Crawley-Boevey, A. Vusirikala, E. Fernandez, M. S. Perez, S. Tang, K. Dun-Campbell, E. W. Evans, A. Bell, B. Patel, Z. Amin-Chowdhury, F. Aiano, K. Paranthaman, T. Ma, M. Saavedra-Campos, R. Myers, J. Ellis, A. Lackenby, R. Gopal, M. Patel, C. Brown, M. Chand, K. Brown, M. E. Ramsay, S. Hopkins, N. Shetty, M. Zambon, Investigation of SARS-CoV-2 outbreaks in six care homes in London, April 2020. *EClinicalMedicine.* **26**, 100533 (2020).
20. J. K. Burton, G. Bayne, C. Evans, F. Garbe, D. Gorman, N. Honhold, D. McCormick, R. Othieno, J. E. Stevenson, S. Swietlik, K. E. Templeton, M. Tranter, L. Willocks, B. Guthrie, Evolution and effects of COVID-19 outbreaks in care homes: a population

- analysis in 189 care homes in one geographical region of the UK. *Lancet Heal. Longev.* **1**, e21–e31 (2020).
21. L. J. Strausbaugh, S. R. Sukumar, C. L. Joseph, Infectious disease outbreaks in nursing homes: An unappreciated hazard for frail elderly persons. *Clin. Infect. Dis.* **36**, 870–876 (2003).
22. T. Inns, D. Wilson, P. Manley, J. P. Harris, S. J. O'Brien, R. Vivancos, What proportion of care home outbreaks are caused by norovirus? An analysis of viral causes of gastroenteritis outbreaks in care homes, North East England, 2016-2018. *BMC Infect. Dis.* **20**, 2 (2019).
23. P. N. Perez-Guzman, A. Daunt, S. Mukherjee, P. Crook, R. Forlano, M. D. Kont, A. Løchen, M. Vollmer, P. Middleton, R. Judge, C. Harlow, A. Soubieres, G. Cooke, P. J. White, T. B. Hallett, P. Aylin, N. Ferguson, K. Hauck, M. R. Thursz, S. Nayagam, Clinical characteristics and predictors of outcomes of hospitalized patients with COVID-19 in a multi-ethnic London NHS Trust: a retrospective cohort study. *Clin. Infect. Dis.*, 1–11 (2020).
24. A. B. Docherty, E. M. Harrison, C. A. Green, H. E. Hardwick, R. Pius, L. Norman, K. A. Holden, J. M. Read, F. Dondelinger, G. Carson, L. Merson, J. Lee, D. Plotkin, L. Sigfrid, S. Halpin, C. Jackson, C. Gamble, P. W. Horby, J. S. Nguyen-Van-Tam, A. Ho, C. D. Russell, J. Dunning, P. J. M. Openshaw, J. K. Baillie, M. G. Semple, Features of 20 133 UK patients in hospital with covid-19 using the ISARIC WHO Clinical Characterisation Protocol: Prospective observational cohort study. *BMJ.* **369**, 1–12 (2020).
25. S. Tabata, K. Imai, S. Kawano, M. Ikeda, T. Kodama, K. Miyoshi, H. Obinata, S. Mimura, T. Kadera, M. Kitagaki, M. Sato, S. Suzuki, T. Ito, Y. Uwabe, K. Tamura, Clinical characteristics of COVID-19 in 104 people with SARS-CoV-2 infection on the Diamond Princess cruise ship: a retrospective analysis. *Lancet Infect. Dis.* **20**, 1043–1050 (2020).

26. P. Vetter, D. L. Vu, A. G. L'Huillier, M. Schibler, L. Kaiser, F. Jacquerioz, Clinical features of covid-19. *BMJ*. **369**, 1–2 (2020).
27. R. Verity, L. C. Okell, I. Dorigatti, P. Winskill, C. Whittaker, N. Imai, G. Cuomo-Dannenburg, H. Thompson, P. G. T. Walker, H. Fu, A. Dighe, J. T. Griffin, M. Baguelin, S. Bhatia, A. Boonyasiri, A. Cori, Z. Cucunubá, R. FitzJohn, K. Gaythorpe, W. Green, A. Hamlet, W. Hinsley, D. Laydon, G. Nedjati-Gilani, S. Riley, S. van Elsland, E. Volz, H. Wang, Y. Wang, X. Xi, C. A. Donnelly, A. C. Ghani, N. M. Ferguson, Estimates of the severity of coronavirus disease 2019: a model-based analysis. *Lancet Infect. Dis.* (2020), doi:10.1016/S1473-3099(20)30243-7.
28. The RECOVERY Collaborative Group, Dexamethasone in Hospitalized Patients with Covid-19 — Preliminary Report. *N. Engl. J. Med.*, 1–11 (2020).
29. L. I. Horwitz, S. A. Jones, R. J. Cerfolio, F. Francois, J. Greco, B. Rudy, C. M. Petrilli, Trends in COVID-19 Risk-Adjusted Mortality Rates. *J. Hosp. Med.* **23**, 2020 (2020).
30. A. W. D. Edridge, J. Kaczorowska, A. C. R. Hoste, M. Bakker, M. Klein, K. Loens, M. F. Jebbink, A. Matser, C. M. Kinsella, P. Rueda, M. Ieven, H. Goossens, M. Prins, P. Sastre, M. Deijis, L. van der Hoek, Seasonal coronavirus protective immunity is short-lasting. *Nat. Med.* (2020), doi:10.1038/s41591-020-1083-1.
31. S. M. Kissler, C. Tedijanto, E. Goldstein, Y. H. Grad, M. Lipsitch, Projecting the transmission dynamics of SARS-CoV-2 through the postpandemic period. *Science* (80-.). **368**, 860–868 (2020).
32. L. C. Okell, R. Verity, O. J. Watson, S. Mishra, P. Walker, C. Whittaker, A. Katzourakis, C. A. Donnelly, S. Riley, A. C. Ghani, A. Gandy, S. Flaxman, N. M. Ferguson, S. Bhatt, Correspondence Have deaths from COVID-19 in Europe plateaued due to herd. *Lancet*. **395**, e110–e111 (2020).
33. Academy of Medical Science, Preparing for a challenging winter 2020/21, 79 (2020).
34. NHS England and NHS Improvement, “Hospital Onset Covid-19: IPC evidence from recent survey and next steps.”

35. G. Iacobucci, Covid-19: Doctors sound alarm over hospital transmissions. *BMJ*. **369**, m2013 (2020).
36. D. Buitrago-Garcia, D. Egli-Gany, M. J. Counotte, S. Hossmann, H. Imeri, A. M. Ipekci, G. Salanti, N. Low, Occurrence and transmission potential of asymptomatic and presymptomatic SARS-CoV-2 infections: A living systematic review and meta-analysis. *PLoS Med.* **17**, e1003346 (2020).
37. S. Riley, C. E. Walters, H. Wang, O. Eales, K. E. C. Ainslie, C. Atchison, C. Fronterre, P. J. Diggle, D. Ashby, C. A. Donnelly, G. Cooke, W. Barclay, H. Ward, A. Darzi, P. Elliott, *medRxiv*, in press, doi:10.1101/2020.12.15.20248244.
38. Office for National Statistics, Office for National Statistics, (available at <https://www.ons.gov.uk/>).
39. Care Quality Commission, [ARCHIVED CONTENT] UK Government Web Archive - The National Archives, (available at <https://webarchive.nationalarchives.gov.uk/20200605160439/https://www.cqc.org.uk/files/cqc-care-directory-filters-1-june-2020>).
40. GOV.UK, “Care Homes Analysis Background” (2020).
41. Age UK, Later Life in the United Kingdom 2019, (available at https://www.ageuk.org.uk/globalassets/age-uk/documents/reports-and-publications/late_life_uk_factsheet.pdf).
42. J. Mossong, N. Hens, M. Jit, P. Beutels, K. Auranen, R. Mikolajczyk, M. Massari, S. Salmaso, G. S. Tomba, J. Wallinga, J. Heijne, M. Sadkowska-Todys, M. Rosinska, W. J. Edmunds, Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS Med.* **5**, 381–391 (2008).
43. S. A. Lauer, K. H. Grantz, Q. Bi, F. K. Jones, Q. Zheng, H. R. Meredith, A. S. Azman, N. G. Reich, J. Lessler, The incubation period of coronavirus disease 2019 (CoVID-19) from publicly reported confirmed cases: Estimation and application. *Ann. Intern. Med.* **172**, 577–582 (2020).

44. Q. Bi, Y. Wu, S. Mei, C. Ye, X. Zou, Z. Zhang, X. Liu, L. Wei, S. A. Truelove, T. Zhang, W. Gao, C. Cheng, X. Tang, X. Wu, Y. Wu, B. Sun, S. Huang, Y. Sun, J. Zhang, T. Ma, J. Lessler, T. Feng, 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), doi:10.1016/s1473-3099(20)30287-5.
45. M. Bernabeu-Wittel, J. E. Ternero-Vega, P. Díaz-Jiménez, C. Conde-Guzmán, M. D. Nieto-Martín, L. Moreno-Gaviño, J. Delgado-Cuesta, M. Rincón-Gómez, L. Giménez-Miranda, . D Navarro-Amuedo, M. M. Muñoz-García, S. Calzón-Fernández, M. Ollero-Baturone, Death risk stratification in elderly patients with covid-19. A comparative cohort study in nursing homes outbreaks. *Arch. Gerontol. Geriatr.* **91**, 104240 (2020).
46. S. Omar, C. Bartz, S. Becker, S. Basenach, S. Pfeifer, C. Trapp, H. Hamm, H. C. Schlichting, M. Friederichs, U. Koch, C. Jestrabek, E. Hilger, M. Vogt, K. Jahn, S. Chen, T. Barnighausen, P. Zanger, Duration of SARS-CoV-2 RNA detection in COVID-19 patients in home isolation, Rhineland-Palatinate, Germany, 2020 - an interval-censored survival analysis. *Eurosurveillance.* **25**, 1–8 (2020).
47. B. Benny, G. Amandine, P. Kc, H. Sarah, M. Abby, C. Caitlin, S. Van, L.-S. James, O. Affiliations, Quantifying antibody kinetics and RNA shedding during early-phase SARS-CoV-2 infection, doi:10.1101/2020.05.15.20103275.
48. S. Funk, Socialmixr: Social Mixing Matrices for Infectious Disease Modelling (2018).
49. GOV.UK, Prime Minister's statement on coronavirus (COVID-19): 25 March 2020 - GOV.UK, (available at <https://www.gov.uk/government/speeches/pm-statement-on-coronavirus-25-march-2020>).
50. GOV.UK, Prime Minister's statement on coronavirus (COVID-19): 11 May 2020 - GOV.UK, (available at <https://www.gov.uk/government/speeches/pm-statement-on-coronavirus-11-may-2020>).
51. GOV.UK, Prime Minister sets out timeline for retail to reopen in June - GOV.UK, (available at <https://www.gov.uk/government/news/prime-minister-sets-out-timeline->

for-retail-to-reopen-in-june).

52. GOV.UK, Pubs, restaurants and hairdressers to reopen from 4 July - GOV.UK, (available at <https://www.gov.uk/government/news/pubs-restaurants-and-hairdressers-to-reopen-from-4-july>).

53. GOV.UK, Eat Out to Help Out launches today – with government paying half on restaurant bills - GOV.UK, (available at <https://www.gov.uk/government/news/eat-out-to-help-out-launches-today-with-government-paying-half-on-restaurant-bills>).

54. GOV.UK, Schools and colleges to reopen in full in September - GOV.UK, (available at <https://www.gov.uk/government/news/schools-and-colleges-to-reopen-in-full-in-september>).

55. O. Diekmann, J. A. P. Heesterbeek, J. A. J. Metz, On the definition and the computation of the basic reproduction ratio R_0 in models for infectious diseases in heterogeneous populations. *J. Math. Biol.* **28**, 365–382 (1990).

56. D. T. Gillespie, Approximate accelerated stochastic simulation of chemically reacting systems. *J. Chem. Phys.* **115**, 1716–1733 (2001).

57. Department of Health and Social Care, COVID-19 testing data: methodology note. *www.gov.uk* (2020).

58. I. M. C. Martin, C. A. Ison, D. M. Aanensen, K. A. Fenton, B. G. Spratt, Rapid Sequence-Based Identification of Gonococcal Transmission Clusters in a Large Metropolitan Area. *J. Infect. Dis.* **189**, 1497–1505 (2004).

59. P. Del Moral, A. Doucet, A. Jasra, Sequential Monte Carlo samplers. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **68**, 411–436 (2006).

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

61. N. J. Gordon, D. J. Salmond, A. F. M. Smith, Novel approach to nonlinear/non-gaussian Bayesian state estimation. *IEE Proceedings, Part F Radar Signal Process.* **140**, 107–113 (1993).

- 595 62. M. Baguelin, E. Knock, L. K. Whittles, R. FitzJohn, J. Lees, sircovid (2020).
- 596 63. E. S. Knock, L. K. Whittles, P. N. Perez-Guzman, S. Bhatia, F. Guntoro, O. J. Watson,
597 C. Whittaker, N. M. Ferguson, A. Cori, M. Baguelin, R. G. FitzJohn, J. A. Lees,
598 Reproducible parallel inference and simulation of stochastic state space models using
599 odin, dust, and mcstate. *Wellcome Open Res.* **5**, 288 (2020).
- 600 64. M. Plummer, N. Best, K. Cowles, K. Vines, CODA: Convergence Diagnosis and
601 Output Analysis for MCMC. *R News.* **6**, 7–11 (2006).
- 602 65. A. Gelman, D. B. Rubin, Inference from iterative simulation using multiple sequences.
603 *Stat. Sci.* **7**, 457–472 (1992).
- 604 66. R. Verity, R. FitzJohn, mrc-ide/markovid at version1.5, (available at
605 <https://github.com/mrc-ide/markovid/tree/version1.5>).
- 606 67. C. I. Jarvis, K. Van Zandvoort, A. Gimma, K. Prem, M. Auzenberg, K. O'Reilly, G.
607 Medley, J. C. Emery, R. M. G. J. Houben, N. Davies, E. S. Nightingale, S. Flasche, T.
608 Jombart, J. Hellewell, S. Abbott, J. D. Munday, N. I. Bosse, S. Funk, F. Sun, A. Endo,
609 A. Rosello, S. R. Procter, A. J. Kucharski, T. W. Russell, G. Knight, H. Gibbs, Q.
610 Leclerc, B. J. Quilty, C. Diamond, Y. Liu, M. Jit, S. Clifford, C. A. B. Pearson, R. M.
611 Eggo, A. K. Deol, P. Klepac, G. J. Rubin, W. J. Edmunds, Quantifying the impact of
612 physical distance measures on the transmission of COVID-19 in the UK. *BMC Med.*
613 **18**, 1–10 (2020).
- 614 68. E. Lavezzo, E. Franchin, C. Ciavarella, G. Cuomo-Dannenburg, L. Barzon, C. Del
615 Vecchio, L. Rossi, R. Manganelli, A. Loregian, N. Navarin, D. Abate, M. Sciro, S.
616 Merigliano, E. De Canale, M. C. Vanuzzo, V. Besutti, F. Saluzzo, F. Onelia, M.
617 Pacenti, S. G. Parisi, G. Carretta, D. Donato, L. Flor, S. Cocchio, G. Masi, A. Sperduti,
618 L. Cattarino, R. Salvador, M. Nicoletti, F. Caldart, G. Castelli, E. Nieddu, B. Labella, L.
619 Fava, M. Drigo, K. A. M. Gaythorpe, K. E. C. Ainslie, M. Baguelin, S. Bhatt, A.
620 Boonyasiri, O. Boyd, L. Cattarino, C. Ciavarella, H. L. Coupland, Z. Cucunubá, G.
621 Cuomo-Dannenburg, B. A. Djafaara, C. A. Donnelly, I. Dorigatti, S. L. van Elsland, R.

FitzJohn, S. Flaxman, K. A. M. Gaythorpe, W. D. Green, T. Hallett, A. Hamlet, D. Haw, N. Imai, B. Jeffrey, E. Knock, D. J. Laydon, T. Mellan, S. Mishra, G. Nedjati-Gilani, P. Nouvellet, L. C. Okell, K. V. Parag, S. Riley, H. A. Thompson, H. J. T. Unwin, R. Verity, M. A. C. Vollmer, P. G. T. Walker, C. E. Walters, H. Wang, Y. Wang, O. J. Watson, C. Whittaker, L. K. Whittles, X. Xi, N. M. Ferguson, A. R. Brazzale, S. Toppo, M. Trevisan, V. Baldo, C. A. Donnelly, N. M. Ferguson, I. Dorigatti, A. Crisanti, Suppression of a SARS-CoV-2 outbreak in the Italian municipality of Vo'. *Nature* (2020), doi:10.1038/s41586-020-2488-1.

69. NHS England and NHS Improvement, Statistics » COVID-19 Hospital Activity, (available at <https://www.england.nhs.uk/statistics/statistical-work-areas/covid-19-hospital-activity/>).

Acknowledgements

We thank all the colleagues at Public Health England (PHE) and frontline health professionals who have driven and continue to drive the daily response to the epidemic, but also for providing the necessary data to inform this study. This work would not have been possible without their dedication and expertise. The use of pillar 2 PCR testing data was made possible thanks to PHE colleagues and we extend our thanks to Dr Nick Gent and Dr André Charlett for facilitation and their insights into these data. The use of serological data was made possible by colleagues at PHE Porton Down, Colindale, and the NHS Blood Transfusion Service. We are particularly grateful to Dr Gayatri Amirthalingam and Prof Nick Andrews for helpful discussions around these data. We also thank the entire Imperial College London Covid-19 Response Team for their support and feedback throughout. This work was supported by the NIHR HPRU in Modelling and Health Economics, a partnership between PHE, Imperial College London and LSHTM (grant code NIHR200908). We acknowledge funding from the MRC Centre for Global Infectious Disease Analysis

(reference MR/R015600/1), jointly funded by the UK Medical Research Council (MRC) and the UK Foreign, Commonwealth & Development Office (FCDO), under the MRC/FCDO Concordat agreement and is also part of the EDCTP2 programme supported by the European Union.

Disclaimer

The views expressed are those of the authors and not necessarily those of the United Kingdom (UK) Department of Health and Social Care, the National Health Service, the National Institute for Health Research (NIHR), Public Health England (PHE), UK MRC, UKRI or European Union.

List of Supplementary Materials

Supplementary materials (Materials and Methods, Supplementary Results)
Supplementary data files: data_rtm.csv, data_serology.csv, support_progression.csv, support_severity.csv

Data availability statement

All code and de-identified regionally aggregated data (see supplementary materials for full details) required to reproduce this analysis are available at <https://github.com/mrc-ide/sarscov2-transmission-england> (<https://zenodo.org/record/4384864>)

Supplementary Materials

Table of Contents

1	Materials and Methods.....	2
1.1	Data sources	2
1.1.1	Hospital admissions and bed occupancy	2
1.1.2	Deaths.....	2
1.1.3	Pillar 2 testing.....	3
1.1.4	Serology surveys	3
1.1.5	REACT-1 prevalence survey	3
1.1.6	Summary of the data used for calibration.....	3
1.1.7	Other data sources	4
1.2	Evidence synthesis	4
1.3	Model description.....	6
1.3.1	Stratification of population into groups	6
1.3.2	Progression of infection and hospitalisation	6
1.3.3	Progression of infection and hospitalisation	8
1.3.4	Age-varying and time-varying infection progression probabilities.....	11
1.4	Reproduction number R_t and effective reproduction number R_{teff}	12
1.5	Infection severity	12
1.6	Compartmental model equations.....	13
1.7	Observation process	18
1.7.1	Notation for distributions used in this section	19
1.7.2	Hospital admissions and new diagnoses in hospital.....	19
1.7.3	Hospital bed occupancy by confirmed COVID-19 cases	19
1.7.4	Hospital and care homes COVID-19 deaths.....	20
1.7.5	Serosurveys	20
1.7.6	PCR testing.....	21
1.8	Bayesian inference and model fitting	22
1.9	Prior distributions and parameter calibration	24
1.9.1	Risk of hospital admission	24
1.9.2	Severity and hospital progression	25
1.9.3	Serosurveys	29
1.9.4	PCR positivity	29
1.9.5	Local start date of the epidemic	29
1.9.6	Time-varying transmission rates	30
1.9.7	Transmission within care homes	30
1.9.8	Parameters relating to Pillar 2 testing.....	30
2	Supplementary Results	35
2.1	Model fitting and validation.....	35
2.2	Severity estimates	36
2.3	Supplementary counterfactual analysis.....	37

References 38

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.

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)
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)

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.

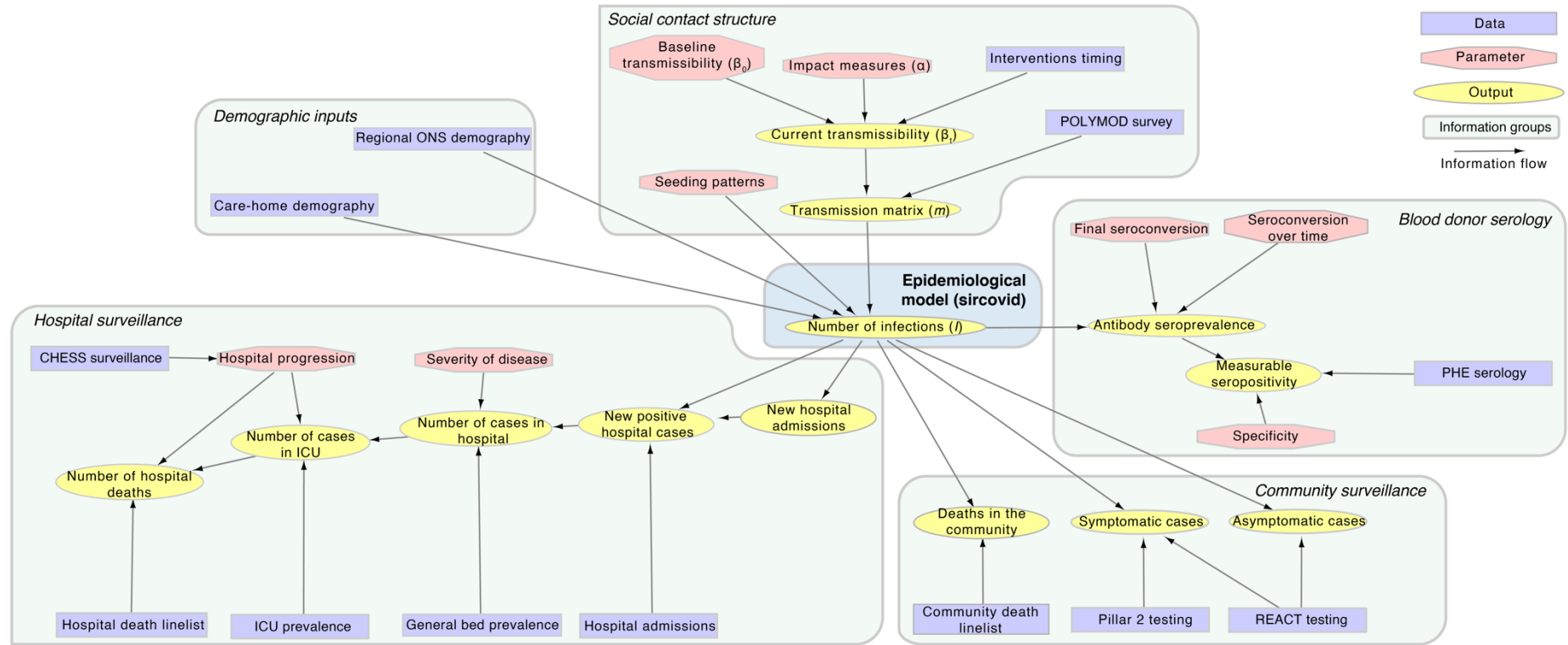


Figure S 1: Graph showing the functional relationships between data sources (rectangles), modelled outputs (ovals) and parameters (hexagons).

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_{GD}^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 ($p_{WD}^i(t)$) 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 ($p_{HD}^i(t)$) die, while the remainder recover. Recovered individuals are assumed to be immune to reinfection for at least the duration of the simulation.

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_{PCRneg}). Meanwhile, individuals move into the seropositivity flow upon becoming infectious, entering first into a pre-seropositivity compartment ($T_{seropre}$). A proportion of individuals ($p_{seropos}$) then seroconvert and move into the seropositivity compartment ($T_{seropos}$), while the remainder move into the seronegativity compartment ($T_{seroneg}$).

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).

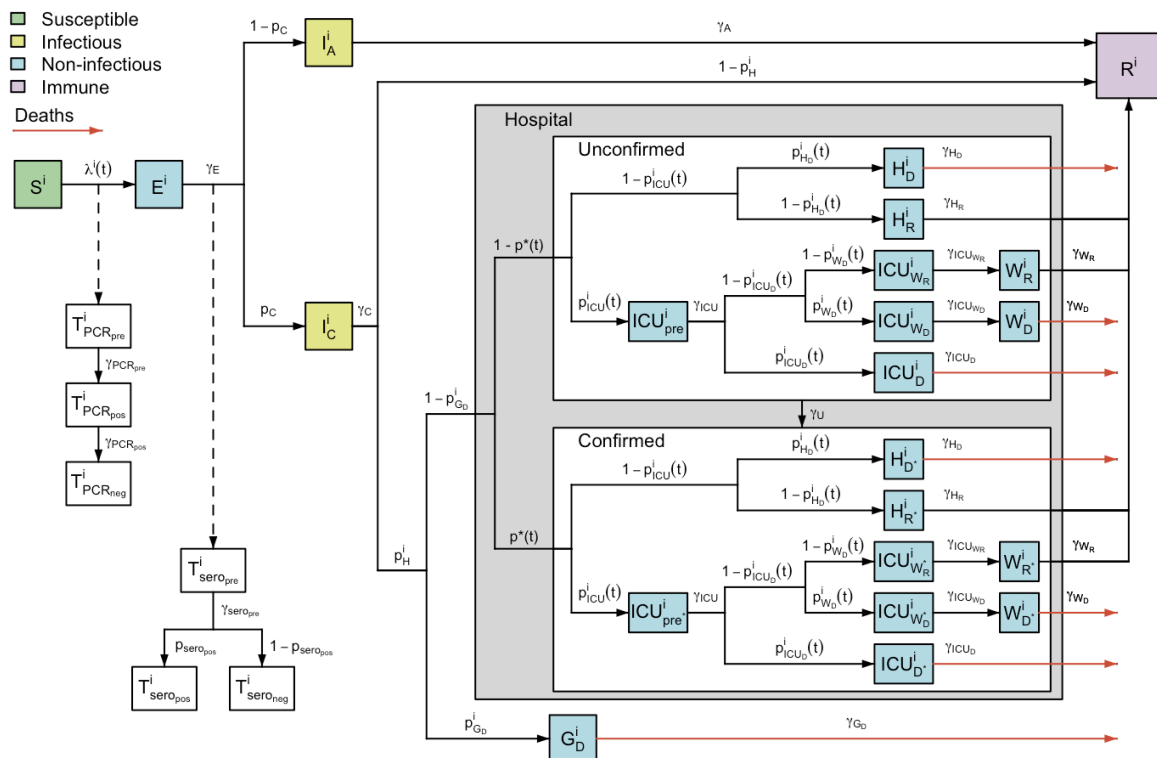


Figure S 2: Model structure flow diagram with rates of transition between infection states. Variable names defined in text.

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 CHES
H_D	Hospitalised on general ward leading to death	2	0.19	10.3 (1.3, 28.8)	Fitted to CHES
ICU_{pre}	Triage to ICU	1	0.40	2.5 (0.1, 9.2)	Fitted to CHES
ICU_{WR}	Hospitalised in ICU, leading to recovery	1	0.06	15.6 (0.4, 57.6)	Fitted to CHES
ICU_{WD}	Hospitalised in ICU, leading to death in step-down following ICU	1	0.14	7.0 (0.2, 25.7)	Fitted to CHES
ICU_D	Hospitalised in ICU, leading to death	2	0.17	11.8 (1.4, 32.9)	Fitted to CHES
W_R	Stepdown recovery period after leaving ICU	2	0.16	12.2 (1.5, 34.0)	Fitted to CHES
W_D	Stepdown period before death after leaving ICU	1	0.12	8.1 (0.2, 29.7)	Fitted to CHES
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 5.

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:

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

220

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

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

223

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

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

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

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

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

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

235

236 with 12 change points t_i corresponding to major announcements and changes in COVID-19
237 related policy, as detailed in Table S 3.

238

239

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

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

241
242

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

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

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

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

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

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

255 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
256 home workers are drawn from).

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

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

260 These represent contact between visitors from the general community and care home
261 residents. This might involve a slightly different age profile than the age profile of the contact
262 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:

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

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

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:

$$p_{ICU}^i(t) = \psi_{ICU}^i p_{ICU}^{max} h(\mu_{ICU}, t) \quad (11)$$

$$p_{ICUD}^i(t) = \psi_{ICUD}^i p_{ICUD}^{max} h(\mu_D, t) \quad (12)$$

$$p_{HD}^i(t) = \psi_{HD}^i p_{HD}^{max} h(\mu_D, t) \quad (13)$$

$$p_{WD}^i(t) = \psi_{WD}^i p_{WD}^{max} h(\mu_D, t) \quad (14)$$

where here for probability p_X^i , p_X^{max} gives 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 G_D for 5 days on average before dying (modelled with $k_{GD} = 2$ and $\gamma_{GD} = 0.4$), 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_X^{CHW} is taken as the mean of the age scalings ψ_X^i for $i \in \{[25,30), [30,35), \dots, [60,65)\}$. For care home residents, we assume that $\psi_X^{CHR} = \psi_X^{(80+)}$, with the exception of the probability of individual with severe disease requiring hospitalisation dying at home (without receiving hospital care), where we assume $\psi_{GD}^{CHR} = 1$ and $\psi_{GD}^i = 0$ for all other groups, to effectively allow death outside hospital only for care home residents.

1.4 Reproduction number R_t and effective reproduction number R_t^{eff}

We calculated the reproduction number over time, R_t , and effective reproduction number over time, R_t^{eff} , 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_t^{eff} 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

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

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 $(i, j) \in \{[0,5), \dots, [75,80), [80+)\}^2$,

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

where N^i is the total population of group i and R_t is taken to be the dominant eigenvalue of $\text{NGM}(t)$, 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:

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

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

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 (22)-(57), 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.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

$$dW_{D^*}^i(t)/dt = \gamma_{ICU_{WD}} ICU_{WD}^i(t) - \gamma_{WD} W_{D^*}^i(t) + \gamma_U W_D^i(t) \quad (43)$$

$$dH_R^i(t)/dt = p_H^i(1 - p_{GD}^i)(1 - p^*(t))(1 - p_{ICU}^i(t))(1 - p_{HD}^i(t)) \gamma_C I_C^i(t) - (\gamma_{HR} + \gamma_U) H_R^i(t) \quad (44)$$

$$dH_{R^*}^i(t)/dt = p_H^i(1 - p_{GD}^i) p^*(t)(1 - p_{ICU}^i(t))(1 - p_{HD}^i(t)) \gamma_C I_C^i(t) + \gamma_U H_R^i(t) - \gamma_{HR} H_{R^*}^i(t) \quad (45)$$

$$dH_D^{i,1}(t)/dt = p_H^i(1 - p_{GD}^i)(1 - p^*(t))(1 - p_{ICU}^i(t)) p_{HD}^i(t) \gamma_C I_C^i(t) - (\gamma_{HD} + \gamma_U) H_D^{i,1}(t) \quad (46)$$

$$dH_D^{i,2}(t)/dt = \gamma_{HD} H_D^{i,1}(t) - (\gamma_{HD} + \gamma_U) H_D^{i,2}(t) \quad (47)$$

$$dH_{D^*}^{i,1}(t)/dt = p_H^i(1 - p_{GD}^i) p^*(t)(1 - p_{ICU}^i(t)) p_{HD}^i(t) \gamma_C I_C^i(t) + \gamma_U H_D^{i,1}(t) - \gamma_{HD} H_{D^*}^{i,1}(t) \quad (48)$$

$$dH_{D^*}^{i,2}(t)/dt = \gamma_{HD} H_{D^*}^{i,1}(t) - \gamma_{HD} H_{D^*}^{i,2}(t) + \gamma_U H_D^{i,2}(t) \quad (49)$$

$$dR^i(t)/dt = \gamma_A I_A^i(t) + (1 - p_H^i) \gamma_C I_C^i(t) + \gamma_{HR} (H_R^i(t) + H_{R^*}^i(t)) + \gamma_{WR} (W_R^i(t) + W_{R^*}^i(t)) \quad (50)$$

$$dT_{seropre}^i(t)/dt = \gamma_E E^{i,2}(t) - \gamma_{seropre} T_{seropre}^i(t) \quad (51)$$

$$dT_{seropos}^i(t)/dt = p_{seropos} \gamma_{seropre} T_{PCRpre}^i(t) \quad (52)$$

$$dT_{seroneg}^i(t)/dt = (1 - p_{seropos}) \gamma_{seropre} T_{PCRpre}^i(t) \quad (53)$$

$$dT_{PCRpre}^i(t)/dt = \lambda^i(t) S^i(t) - \gamma_{PCRpre} T_{PCRpre}^i(t) \quad (54)$$

$$dT_{PCRpos}^i(t)/dt = \gamma_{PCRpre} T_{PCRpre}^i(t) - \gamma_{PCRpos} T_{PCRpos}^i(t) \quad (55)$$

$$dT_{PCRneg}^i(t)/dt = \gamma_{PCRpos} T_{PCRpos}^i(t) \quad (56)$$

339

340 We used the tau-leap method (33) to create a stochastic, time-discretised version of the
 341 model described in equations (58-162), taking four update steps per day. The process was
 342 initialised with ten asymptomatic infectious individuals aged 15-19 on the epidemic start date
 343 t_0 , a parameter we estimate. For each time step, the model iterated through the procedure
 344 described below. In the following, we introduce a small abuse of notation: for transitions
 345 involving multiple onward compartments (e.g transition from compartment E to
 346 compartments I_A or I_C), for conciseness, we write

$$(d_{E,I_A}^i, d_{E,I_C}^i) \sim \text{Multinom}(E^{i,2}(t), q_{E,I_A}^i, q_{E,I_C}^i)$$

347 instead of

$$(d_{E,I_A}^i, d_{E,I_C}^i, d_{nomove}^i) \sim \text{Multinom}(E^{i,2}(t), q_{E,I_A}^i, q_{E,I_C}^i, 1 - \sum_{x \in \{I_A, I_C\}} q_{E,x}^i)$$

348 where d_{nomove}^i is a dummy variable counting the number of individuals remaining in compartment
 349 $E^{i,2}$.

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

352

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

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

$$(q_{E,IA}^i, q_{E,IC}^i) = ((1 - p_C)(1 - e^{-\gamma_E dt}), p_C(1 - e^{-\gamma_E dt})) \quad (59)$$

$$(d_{E,IA}^i, d_{E,IC}^i) \sim \text{Multinom} \left(E^{i,2}(t), q_{E,IA}^i, q_{E,IC}^i \right) \quad (60)$$

$$d_{IA,R}^i \sim \text{Binom} \left(I_A^i(t), 1 - e^{-\gamma_A dt} \right) \quad (61)$$

$$q_{IC,GD}^i = p_H^i p_{GD}^i (1 - e^{-\gamma_C dt}) \quad (62)$$

$$q_{IC,R}^i = (1 - p_H^i)(1 - e^{-\gamma_C dt}) \quad (63)$$

$$q_{IC,ICU_{pre}}^i = p_H^i (1 - p_{GD}^i) (1 - p^*(t)) p_{ICU}^i(t) (1 - e^{-\gamma_C dt}) \quad (64)$$

$$q_{IC,ICU_{pre}^*}^i = p_H^i (1 - p_{GD}^i) p^*(t) p_{ICU}^i(t) (1 - e^{-\gamma_C dt}) \quad (65)$$

$$q_{IC,HR}^i = p_H^i (1 - p_{GD}^i) (1 - p^*(t)) (1 - p_{ICU}^i(t)) (1 - p_{HD}^i(t)) (1 - e^{-\gamma_C dt}) \quad (66)$$

$$q_{IC,HR}^{i*} = p_H^i (1 - p_{GD}^i) p^*(t) (1 - p_{ICU}^i(t)) (1 - p_{HD}^i(t)) (1 - e^{-\gamma_C dt}) \quad (67)$$

$$q_{IC,HD}^i = p_H^i (1 - p_{GD}^i) (1 - p^*(t)) (1 - p_{ICU}^i(t)) p_{HD}^i(t) (1 - e^{-\gamma_C dt}) \quad (68)$$

$$q_{IC,HD}^{i*} = p_H^i (1 - p_{GD}^i) p^*(t) (1 - p_{ICU}^i(t)) p_{HD}^i(t) (1 - e^{-\gamma_C dt}) \quad (69)$$

$$(d_{IC,GD}^i, \dots, d_{IC,HD}^{i*}) \sim \text{Multinom} \left(I_C^i(t), q_{IC,GD}^i, \dots, q_{IC,HD}^{i*} \right) \quad (70)$$

$$d_{GD,GD}^i \sim \text{Binom} \left(G_D^{i,1}(t), 1 - e^{-\gamma_{GD} dt} \right) \quad (71)$$

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

$$q_{ICU_{pre},ICU_{WR}}^i = (1 - p_{ICU_D}^i(t)) (1 - p_{WD}^i(t)) (1 - e^{-\gamma_{ICU_{pre}} dt}) e^{-\gamma_{UD} dt} \quad (73)$$

$$q_{ICU_{pre},ICU_{WR}^*}^i = (1 - p_{ICU_D}^i(t)) (1 - p_{WD}^i(t)) (1 - e^{-\gamma_{ICU_{pre}} dt}) (1 - e^{-\gamma_{UD} dt}) \quad (74)$$

$$q_{ICU_{pre},ICU_{WD}}^i = (1 - p_{ICU_D}^i(t)) p_{WD}^i(t) (1 - e^{-\gamma_{ICU_{pre}} dt}) e^{-\gamma_{UD} dt} \quad (75)$$

$$q_{ICU_{pre},ICU_{WD}^*}^i = (1 - p_{ICU_D}^i(t)) p_{WD}^i(t) (1 - e^{-\gamma_{ICU_{pre}} dt}) (1 - e^{-\gamma_{UD} dt}) \quad (76)$$

$$q_{ICU_{pre},ICU_D}^i = p_{ICU_D}^i(t) (1 - e^{-\gamma_{ICU_{pre}} dt}) e^{-\gamma_{UD} dt} \quad (77)$$

$$q_{ICU_{pre},ICU_D^*}^i = p_{ICU_D}^i(t) (1 - e^{-\gamma_{ICU_{pre}} dt}) (1 - e^{-\gamma_{UD} dt}) \quad (78)$$

$$q_{ICU_{pre},ICU_{pre}^*}^i = e^{-\gamma_{ICU_{pre}} dt} (1 - e^{-\gamma_{UD} dt}) \quad (79)$$

$$(d_{ICU_{pre},ICU_{WR}}^i, \dots, d_{ICU_{pre},ICU_{pre}^*}^i) \sim \text{Multinom} \left(ICU_{pre}^i(t), q_{ICU_{pre},ICU_{WR}}^i, \dots, q_{ICU_{pre},ICU_{pre}^*}^i \right) \quad (80)$$

$$q_{ICU_{pre}^*,ICU_{WR}^*}^i = (1 - p_{ICU_D}^i(t)) (1 - p_{WD}^i(t)) (1 - e^{-\gamma_{ICU_{pre}} dt}) \quad (81)$$

$$q_{ICU_{pre}^*,ICU_{WD}^*}^i = (1 - p_{ICU_D}^i(t)) p_{WD}^i(t) (1 - e^{-\gamma_{ICU_{pre}} dt}) \quad (82)$$

$$q_{ICU_{pre}^*,ICU_D^*}^i = p_{ICU_D}^i(t) (1 - e^{-\gamma_{ICU_{pre}} dt}) \quad (83)$$

$$(d_{ICU_{pre}^*,ICU_{WR}^*}^i, \dots, d_{ICU_{pre}^*,ICU_D^*}^i) \sim \text{Multinom} \left(ICU_{pre}^{i*}(t), q_{ICU_{pre}^*,ICU_{WR}^*}^i, \dots, q_{ICU_{pre}^*,ICU_D^*}^i \right) \quad (84)$$

$$q_{H_D, H_D}^i = (1 - e^{-\gamma_{H_D} dt}) e^{-\gamma_U dt} \quad (85)$$

$$q_{H_D, H_D^*}^{i,1,1} = e^{-\gamma_{H_D} dt} (1 - e^{-\gamma_U dt}) \quad (86)$$

$$q_{H_D, H_D^*}^{i,1,2} = (1 - e^{-\gamma_{H_D} dt}) (1 - e^{-\gamma_U dt}) \quad (87)$$

$$(d_{H_D, H_D}^i, d_{H_D, H_D^*}^{i,1,1}, d_{H_D, H_D^*}^{i,1,2}) \sim \text{Multinom} (H_D^{i,1}(t), q_{H_D, H_D}^i, q_{H_D, H_D^*}^{i,1,1}, q_{H_D, H_D^*}^{i,1,2}) \quad (88)$$

$$d_{H_D^*, H_D^*}^i \sim \text{Binom} (H_D^{i,1}(t), 1 - e^{-\gamma_{H_D} dt}) \quad (89)$$

$$(d_{H_D, D}^i, d_{H_D, H_D^*}^{i,2,2}) \sim \text{Multinom} (H_D^{i,2}(t), 1 - e^{-\gamma_{H_D} dt}, e^{-\gamma_{H_D} dt} (1 - e^{-\gamma_U dt})) \quad (90)$$

$$d_{H_D^*, D}^i \sim \text{Binom} (H_D^{i,2}(t), 1 - e^{-\gamma_{H_D} dt}) \quad (91)$$

$$(d_{H_R, R}^i, d_{H_R, H_R^*}^i) \sim \text{Multinom} (H_R^i(t), 1 - e^{-\gamma_{H_R} dt}, e^{-\gamma_{H_R} dt} (1 - e^{-\gamma_U dt})) \quad (92)$$

$$d_{H_R^*, R}^i \sim \text{Binom} (H_R^i(t), 1 - e^{-\gamma_{H_R} dt}) \quad (93)$$

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

$$q_{ICU_{W_R}, ICU_{W_R}^*}^i = e^{-\gamma_{ICU_{W_R}} dt} (1 - e^{-\gamma_U dt}) \quad (95)$$

$$q_{ICU_{W_R}, W_R^*}^i = (1 - e^{-\gamma_{ICU_{W_R}} dt}) (1 - e^{-\gamma_U dt}) \quad (96)$$

$$(d_{ICU_{W_R}, W_R}^i, \dots, d_{ICU_{W_R}, W_R^*}^i) \sim \text{Multinom} (ICU_{W_R}^i(t), q_{ICU_{W_R}, W_R}^i, \dots, q_{ICU_{W_R}, W_R^*}^i) \quad (97)$$

$$d_{ICU_{W_R^*}, W_R^*}^i \sim \text{Binom} (ICU_{W_R^*}^i(t), 1 - e^{-\gamma_{ICU_{W_R}} dt}) \quad (98)$$

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

$$q_{ICU_{W_D}, ICU_{W_D}^*}^i = e^{-\gamma_{ICU_{W_D}} dt} (1 - e^{-\gamma_U dt}) \quad (100)$$

$$q_{ICU_{W_D}, W_D^*}^i = (1 - e^{-\gamma_{ICU_{W_D}} dt}) (1 - e^{-\gamma_U dt}) \quad (101)$$

$$(d_{ICU_{W_D}, W_D}^i, \dots, d_{ICU_{W_D}, W_D^*}^i) \sim \text{Multinom} (ICU_{W_D}^i(t), q_{ICU_{W_D}, W_D}^i, \dots, q_{ICU_{W_D}, W_D^*}^i) \quad (102)$$

$$d_{ICU_{W_D^*}, W_D^*}^i \sim \text{Binom} (ICU_{W_D^*}^i(t), 1 - e^{-\gamma_{ICU_{W_D}} dt}) \quad (103)$$

$$q_{ICU_D, ICU_D}^i = (1 - e^{-\gamma_{ICU_D} dt}) e^{-\gamma_U dt} \quad (104)$$

$$q_{ICU_D, ICU_D^*}^{i,1,1} = e^{-\gamma_{ICU_D} dt} (1 - e^{-\gamma_U dt}) \quad (105)$$

$$q_{ICU_D, ICU_D^*}^{i,1,2} = (1 - e^{-\gamma_{ICU_D} dt}) (1 - e^{-\gamma_U dt}) \quad (106)$$

$$(d_{ICU_D, ICU_D}^i, d_{ICU_D, ICU_D^*}^{i,1,1}, d_{ICU_D, ICU_D^*}^{i,1,2}) \sim \text{Multinom} (ICU_D^{i,1}(t), q_{ICU_D, ICU_D}^i, q_{ICU_D, ICU_D^*}^{i,1,1}, q_{ICU_D, ICU_D^*}^{i,1,2}) \quad (107)$$

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

$$(d_{ICU_D, D}^i, d_{ICU_D, H_D^*}^{i,2,2}) \sim \text{Multinom} (ICU_D^{i,2}(t), 1 - e^{-\gamma_{ICU_D} dt}, e^{-\gamma_{ICU_D} dt} (1 - e^{-\gamma_U dt})) \quad (109)$$

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

$$q_{W_R, W_R}^i = (1 - e^{-\gamma_{W_R} dt}) e^{-\gamma_U dt} \quad (111)$$

$$q_{W_R, W_R^*}^{i,1,1} = e^{-\gamma_{W_R} dt} (1 - e^{-\gamma_U dt}) \quad (112)$$

$$q_{W_R, W_R^*}^{i,1,2} = (1 - e^{-\gamma_{W_R} dt}) (1 - e^{-\gamma_U dt}) \quad (113)$$

$$\left(d_{W_R, W_R}^i, d_{W_R, W_R^*}^{i,1,1}, d_{W_R, W_R^*}^{i,1,2}\right) \sim \text{Multinom}\left(W_R^{i,1}(t), q_{W_R, W_R}^i, q_{W_R, W_R^*}^{i,1,1}, q_{W_R, W_R^*}^{i,1,2}\right) \quad (114)$$

$$d_{W_R^*, W_R^*}^i \sim \text{Binom}\left(W_R^{i,1}(t), 1 - e^{-\gamma_{W_R} dt}\right) \quad (115)$$

$$\left(d_{W_R, R}^i, d_{W_R, W_R^*}^{i,2,2}\right) \sim \text{Multinom}\left(W_R^{i,2}(t), 1 - e^{-\gamma_{W_R} dt}, e^{-\gamma_{W_R} dt}(1 - e^{-\gamma_U dt})\right) \quad (116)$$

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

$$\left(d_{W_D, D}^i, d_{W_D, W_D^*}^i\right) \sim \text{Multinom}\left(W_D^i(t), 1 - e^{-\gamma_{W_D} dt}, e^{-\gamma_{W_D} dt}(1 - e^{-\gamma_U dt})\right) \quad (118)$$

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

$$\left(d_{W_D, D}^i, d_{W_D, W_D^*}^i\right) \sim \text{Multinom}\left(W_D^i(t), 1 - e^{-\gamma_{W_D} dt}, e^{-\gamma_{W_D} dt}(1 - e^{-\gamma_U dt})\right) \quad (120)$$

$$q_{T_{seropre}, T_{seropos}}^i = p_{seropos}(1 - e^{-\gamma_{seropre} dt}) \quad (121)$$

$$q_{T_{seropre}, T_{seroneg}}^i = (1 - p_{seropos})(1 - e^{-\gamma_{seropre} dt}) \quad (122)$$

$$\left(d_{T_{seropre}, T_{seropos}}^i, d_{T_{seropre}, T_{seroneg}}^i\right) \sim \text{Multinom}\left(T_{seropre}^i(t), q_{T_{seropre}, T_{seropos}}^i, q_{T_{seropre}, T_{seroneg}}^i\right) \quad (123)$$

$$d_{T_{PCRpre}, T_{PCRpos}}^i \sim \text{Binom}\left(T_{PCRpre}^i(t), 1 - e^{-\gamma_{PCRpre} dt}\right) \quad (124)$$

$$d_{T_{PCRpos}, T_{PCRneg}}^i \sim \text{Binom}\left(T_{PCRpos}^i(t), 1 - e^{-\gamma_{PCRpos} dt}\right) \quad (125)$$

353

354 Model compartments were then updated as follows:

355

$$S^i(t + dt) := S^i(t) - d_{S,E}^i \quad (126)$$

$$E^{i,1}(t + dt) := E^{i,1}(t) + d_{S,E}^i - d_{E,E}^i \quad (127)$$

$$E^{i,2}(t + dt) := E^{i,2}(t) + d_{E,E}^i - d_{E,IA}^i - d_{E,IC}^i \quad (128)$$

$$I_A^i(t + dt) := I_A^i(t) + d_{E,IA}^i - d_{IA,R}^i \quad (129)$$

$$I_C^i(t + dt) := I_C^i(t) + d_{E,IC}^i - d_{IC,GD}^i - d_{IC,R}^i - d_{IC,ICUpre}^i - d_{IC,ICUpre^*}^i - d_{IC,HR}^i - d_{IC,HR^*}^i - d_{IC,HD}^i - d_{IC,HD^*}^i \quad (130)$$

$$G_D^{i,1}(t + dt) := G_D^{i,1}(t) + d_{IC,GD}^i - d_{GD,GD}^i \quad (131)$$

$$G_D^{i,2}(t + dt) := G_D^{i,2}(t) + d_{GD,GD}^i - d_{GD,D}^i \quad (132)$$

$$ICU_{pre}^i(t + dt) := ICU_{pre}^i(t) + d_{IC,ICUpre}^i - d_{ICUpre,ICUWR}^i - d_{ICUpre,ICUWD}^i - d_{ICUpre,ICUD}^i - d_{ICUpre,ICUpre^*}^i - d_{ICUpre,ICUWR^*}^i - d_{ICUpre,ICUWD^*}^i - d_{ICUpre,ICUD^*}^i \quad (133)$$

$$ICU_{pre^*}^i(t + dt) := ICU_{pre^*}^i(t) + d_{IC,ICUpre^*}^i - d_{ICUpre^*,ICUWR}^i - d_{ICUpre^*,ICUWR^*}^i - d_{ICUpre^*,ICUD}^i \quad (134)$$

$$ICU_{WR}^i(t + dt) := ICU_{WR}^i(t) + d_{ICUpre,ICUWR}^i - d_{ICUWR,WR}^i - d_{ICUWR,ICUWR^*}^i - d_{ICUWR,WR^*}^i \quad (135)$$

$$ICU_{WR^*}^i(t + dt) := ICU_{WR^*}^i(t) + d_{ICUpre^*,ICUWR^*}^i + d_{ICUWR,ICUWR^*}^i + d_{ICUpre,ICUWR^*}^i - d_{ICUWR^*,WR^*}^i \quad (136)$$

$$ICU_{WD}^i(t + dt) := ICU_{WD}^i(t) + d_{ICUpre,ICUWD}^i - d_{ICUWD,WD}^i - d_{ICUWD,ICUWD^*}^i - d_{ICUWD,WD^*}^i \quad (137)$$

$$ICU_{W_D^*}^i(t+dt) := ICU_{W_D^*}^i(t) + d_{ICU_{pre^*,ICU_{W_D^*}}^i} + d_{ICU_{W_D^*,ICU_{W_D^*}}^i} + d_{ICU_{pre,ICU_{W_D^*}}^i} - d_{ICU_{W_D^*,W_D^*}}^i \quad (138)$$

$$ICU_D^{i,1}(t+dt) := ICU_D^{i,1}(t) + d_{ICU_{pre,ICU_D}^i} - d_{ICU_D,ICU_D}^i - d_{ICU_D,ICU_D^*}^{i,1,1} - d_{ICU_D,ICU_D^*}^{i,1,2} \quad (139)$$

$$ICU_D^{i,2}(t+dt) := ICU_D^{i,2}(t) + d_{ICU_D,ICU_D}^i - d_{ICU_D,D}^i - d_{ICU_D,ICU_D^*}^{i,2,2} \quad (140)$$

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

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

$$W_R^{i,1}(t+dt) := W_R^{i,1}(t) + d_{ICU_{W_R,W_R}^i} - d_{W_R,W_R}^i - d_{W_R,W_R^*}^{i,1,1} - d_{W_R,W_R^*}^{i,1,2} \quad (143)$$

$$W_R^{i,2}(t+dt) := W_R^{i,2}(t) + d_{W_R,W_R}^i - d_{W_R,R}^i - d_{W_R,W_R^*}^{i,2,2} \quad (144)$$

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

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

$$W_D^i(t+dt) := W_D^i(t) + d_{ICU_{W_D,W_D}^i} - d_{W_D,D}^i - d_{W_D,W_D^*}^i \quad (147)$$

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

$$H_D^{i,1}(t+dt) := H_D^{i,1}(t) + d_{I_C,H_D}^i - d_{H_D,H_D}^i - d_{H_D,H_D^*}^{i,1,1} - d_{H_D,H_D^*}^{i,1,2} \quad (149)$$

$$H_D^{i,2}(t+dt) := H_D^{i,2}(t) + d_{H_D,H_D}^i - d_{H_D,D}^i - d_{H_D,H_D^*}^{i,2,2} \quad (150)$$

$$H_{D^*}^{i,1}(t+dt) := H_{D^*}^{i,1}(t) + d_{I_C,H_{D^*}}^i + d_{H_D,H_{D^*}}^{i,1,1} - d_{H_{D^*,H_{D^*}}^i} \quad (151)$$

$$H_{D^*}^{i,2}(t+dt) := H_{D^*}^{i,2}(t) + d_{H_{D^*,H_{D^*}}^i} + d_{H_D,H_{D^*}}^{i,2,2} + d_{H_D,H_{D^*}}^{i,1,2} - d_{H_{D^*,D}}^i \quad (152)$$

$$H_R^i(t+dt) := H_R^i(t) + d_{I_C,H_R}^i - d_{H_R,R}^i - d_{H_R,H_R^*}^i \quad (153)$$

$$H_{R^*}^i(t+dt) := H_{R^*}^i(t) + d_{I_C,H_{R^*}}^i + d_{H_R,H_{R^*}}^i - d_{H_{R^*,R}}^i \quad (154)$$

$$R^i(t+dt) := R^i(t) + d_{I_A,R}^i + d_{I_C,R}^i + d_{H_R,R}^i + d_{H_{R^*,R}}^i + d_{W_R,R}^i + d_{W_{R^*,R}}^i \quad (155)$$

$$T_{seropre}^i(t+dt) := T_{seropre}^i(t) + d_{E,I_A}^i + d_{E,I_C}^i - d_{T_{seropre},T_{seropos}}^i - d_{T_{seropre},T_{seroneg}}^i \quad (156)$$

$$T_{seropos}^i(t+dt) := T_{seropos}^i(t) + d_{T_{seropre},T_{seropos}}^i \quad (157)$$

$$T_{seroneg}^i(t+dt) := T_{seroneg}^i(t) + d_{T_{seropre},T_{seroneg}}^i \quad (158)$$

$$T_{PCRpre}^i(t+dt) := T_{PCRpre}^i(t) + d_{S,E}^i - d_{T_{PCRpre},T_{PCRpos}}^i \quad (159)$$

$$T_{PCRpos}^i(t+dt) := T_{PCRpos}^i(t) + d_{T_{PCRpre},T_{PCRpos}}^i - d_{T_{PCRpos},T_{PCRneg}}^i \quad (160)$$

$$T_{PCRneg}^i(t+dt) := T_{PCRneg}^i(t) + d_{T_{PCRpos},T_{PCRneg}}^i \quad (161)$$

356

357 1.7 Observation process

358 To describe the epidemic in each NHS region, we fitted our model to time series data on
 359 hospital admissions, hospital ward occupancy (both in general beds and in ICU beds),
 360 deaths in hospitals, deaths in care homes, population serological surveys and PCR testing
 361 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

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

where $a = \omega \left(\frac{1-\rho}{\rho}\right)$, $b = (1 - \omega) \left(\frac{1-\rho}{\rho}\right)$ and $B(a, b)$ 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:

$$X_{adm}(t) := \sum_i \left(\sum_{j \in \{H_R^*, H_D^*, ICU_{pre}^*\}} d_{i,j}^i + \sum_{j \in \{H_R, ICU_{pre}, ICU_{WR}, ICU_{WD}, W_D\}} d_{j,j^*}^i + d_{H_D, H_D^*}^{i,1,1} + d_{H_D, H_D^*}^{i,1,2} + d_{H_D, H_D^*}^{i,2,2} + d_{ICU_D, ICU_D^*}^{i,1,1} + d_{ICU_D, ICU_D^*}^{i,1,2} + d_{ICU_D, ICU_D^*}^{i,2,2} + d_{WR, WR^*}^{i,1,1} + d_{WR, WR^*}^{i,1,2} + d_{WR, WR^*}^{i,2,2} + d_{ICU_{pre}, ICU_{WR}^*}^i + d_{ICU_{pre}, ICU_{WD}^*}^i + d_{ICU_{pre}, ICU_D^*}^i + d_{ICU_{WD}^*, W_D^*}^i + d_{ICU_{WR}^*, W_R^*}^i \right) \quad (164)$$

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:

$$X_{hosp}(t) := \sum_i \left(I_{H_R}^i(t) + I_{H_D}^i(t) + I_{ICU_{pre}}^i(t) + I_{W_D}^i(t) + I_{W_R}^i(t) \right) \quad (166)$$

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

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

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

393

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

$$X_{ICU}(t) := \sum_i \left(I_{ICU_{W_R}^*}^i(t) + I_{ICU_{W_D}^*}^i(t) + I_{ICU_{D^*}}^i(t) \right) \quad (168)$$

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

397

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

398

399 with $\kappa_{ICU} = 2$.

400 1.7.4 Hospital and care homes COVID-19 deaths

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

$$X_{hosp_D}(t) := \sum_i \left(d_{H_D,D}^i + d_{H_D^*,D}^i + d_{ICU_{D,D}}^i + d_{ICU_{D^*,D}}^i + d_{W_D,D}^i + d_{W_D^*,D}^i \right) \quad (170)$$

403

404 Which was related to the data via a reporting distribution:

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

405 with $\kappa_{hosp_D} = 2$.

406 Similarly, we represented the reported number of daily COVID-19 deaths in care homes,
407 $Y_{G_D}(t)$, as the observed realisations of an underlying hidden Markov process, $X_{G_D}(t)$, defined
408 as:

$$X_{G_D}(t) := d_{G_D,D}^{CHR} \quad (172)$$

409 Which was related to the data via a reporting distribution:

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

410 with $\kappa_{G_D} = 2$.

411

412 1.7.5 Serosurveys

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

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

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

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:

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

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):

$$X_{P2_{pos}}(t) := \sum_{i=[25,30)}^{CHW} d_{E,IC}^i \quad (178)$$

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

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

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):

$$\omega_{P2_{pos}}(t) := \left(X_{P2_{pos}}(t) \right) / \left(X_{P2_{pos}}(t) + X_{P2_{neg}}(t) \right) \quad (180)$$

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:

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

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:

$$X_{R1_{pos}}(t) := \sum_{i=[5,10), \dots, [80+), CHW} T_{PCR_{pos}}^i(t) \quad (182)$$

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

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 ;

- 476 • thirteen transmission rates at different time points $\beta_1, \dots, \beta_{12}$;
- 477 • three parameters governing transmission to and within care homes
- 478 $m_{CHW}, m_{CHR}, \epsilon$;
- 479 • the probability of symptomatic individuals developing serious disease
- 480 requiring hospitalisation, p_H^{max} , for the group with the largest probability;
- 481 • the probability of a care home resident dying in a care home if they have
- 482 severe disease requiring hospitalisation, p_{GD}^{CHR} ;
- 483 • the probability, at the start of the pandemic, of a patient being admitted to ICU
- 484 after hospitalisation, p_{ICU}^{max} , for the group with the largest probability;
- 485 • the probabilities, at the start of the pandemic, of dying in a hospital general
- 486 ward, p_{HD}^{max} , in the ICU, p_{ICUD}^{max} , and in a stepdown ward following ICU, p_{WD}^{max} ,
- 487 for the groups with the largest probability;
- 488 • the multiplier for hospital mortality after improvement in care, μ_H ;
- 489 • the multiplier for probability of admission to ICU after improvement in care,
- 490 μ_{ICU} ;
- 491 • the daily proportion p_{NC} , of the population seeking to get tested for an infection
- 492 of SARS-Cov-2 following COVID-19 like symptoms and the overdispersion of
- 493 the corresponding observation distribution ρ_{P2test} .

494

495 We used particle Monte Carlo Markov Chain (pMCMC) methods (37), implementing a

496 particle marginal Metropolis-Hastings algorithm with a bootstrap particle filter (38) with 96

497 particles (for sufficient variance in likelihood and a convenient multiple of number of

498 available CPU cores for efficiency), to obtain a sample from the posterior distribution of the

499 model parameters given the observed data. If the expected values of count distributions are

500 zero when observed values are non-zero, this results in particles of zero weight, which can

501 lead to the particle filter estimating the marginal likelihood to be 0. Therefore, to get a small

502 but non-zero weight for each particle at every observation, within our particle filter likelihood

503 we add a small amount of noise (exponentially distributed with mean 10^{-6}) to count values

504 from the model.

505

506 Within our particle filter we add small amounts of exponentially-distributed noise (with mean

507 10^{-6}) to model outputs prior to calculating likelihood weights to avoid particles of zero

508 weight, instead resulting in small but non-zero weights.

509 We implemented our model and parameter inference in an R package, *sircovid* (39),

510 available at <https://mrc-ide.github.io/sircovid>, which uses two further R packages, *dust* to run

511 the model in efficient compiled code and *mcstate* to implement the pMCMC sampler using

512 Metropolis-Hastings sampling (40).

513 At each iteration, the sampler proposes an update to the joint distribution of parameters.

514 These proposals are generated from multivariate Gaussian densities centred on the current

515 parameter values, and with a covariance structure chosen to facilitate efficient mixing of the

516 Markov chain. We specified reflecting boundaries for the proposal kernel to ensure that the

517 proposed parameters are both epidemiologically and mathematically plausible and retain

518 symmetry in the proposals.

519 For each regional fit, eight parallel chains of the pMCMC were run for 11,000 iterations, with

520 the first 1,000 discarded as burn-in, and a 1/80 thinning. We assessed convergence visually.

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 (CHES) 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 CHES 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 CHES 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:

$$New \psi_H^i = Initial \psi_H^i * \frac{p_{H_{Dashboard}}^i}{p_{H_{Model}}^i} \quad (186)$$

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 .

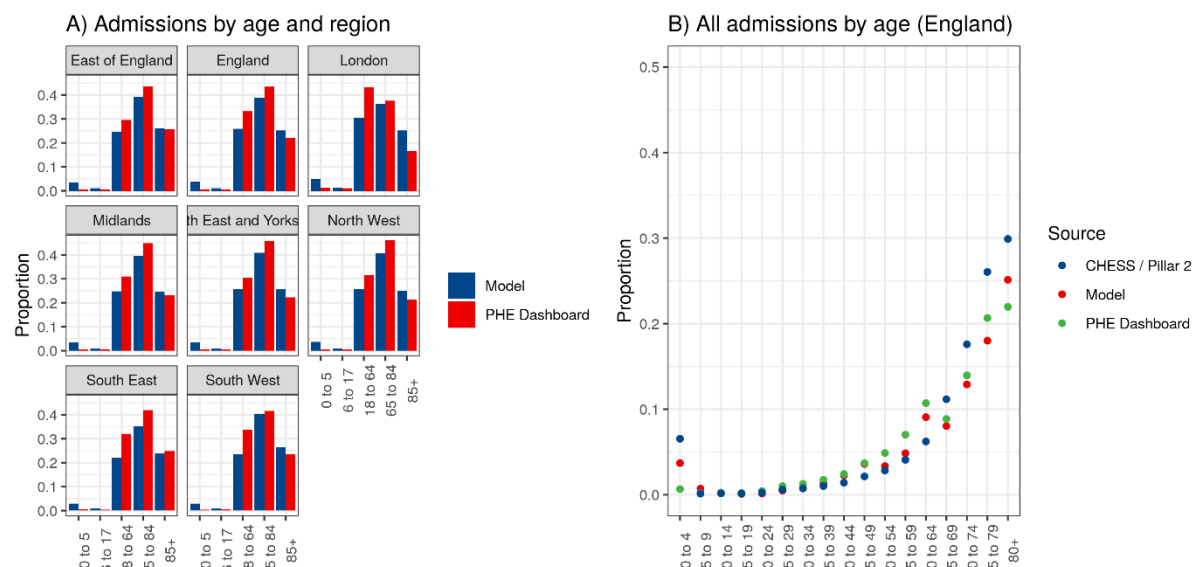


Figure S 3: Proportion of admissions by age. a) Comparison of model outputs to data from the Government's Coronavirus Dashboard, aggregated to five broad age categories. b) Age spline fitted (red) to Government's Coronavirus Dashboard, with age categories disaggregated to five-year bands. The fitted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

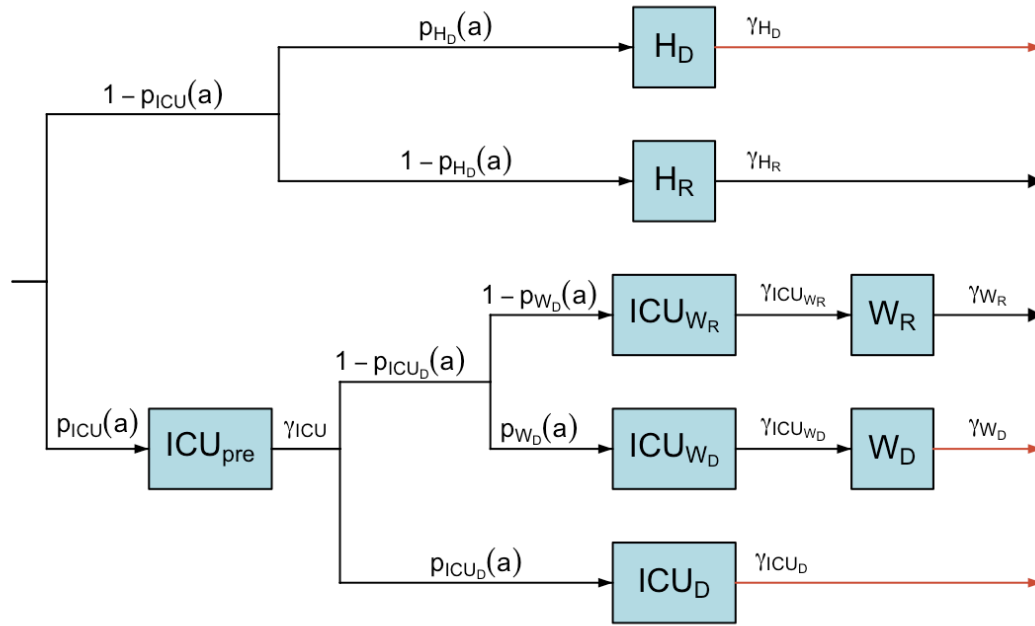


Figure S 4: Directed Acyclic Graph of the hospital pathways fitted to CHES data, which mirror the model structure described in Figure S 2, but with all parameters varying with age and not over time.

CHES 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 :

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

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_2^1 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
$m_X = \frac{1}{\gamma_X}$	Mean of Erlang length of stay distribution	$m_X \sim Uniform_{cont}(0, 20)$

k_X	Shape parameter of Erlang length of stay distribution	$k_X \sim \text{Uniform}_{\text{disc}}(1,10)$
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 $y_Z^1 \in (-10, 10)$
y_Z^j for $j \in 2:7$	Subsequent spline nodes of (transformed) transition probability	$y_Z^j \sim \text{Normal}(y_Z^{j-1}, 0.25)$

620

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

630 For the wider mechanistic transmission model, we used Beta distributions for the priors of
631 the various fitted probabilities regarding hospitalisation. The priors for p_{ICU}^{max} , $p_{H_D}^{max}$, $p_{ICU_D}^{max}$ and
632 $p_{W_D}^{max}$ were all informed by the fitting to CHES data by taking the median fitted value for the
633 prior mean, which we halve in the case of p_{ICU}^{max} to account for CHES being biased to more
634 severe patients. The prior distributions are then calibrated so that the lower bound of the
635 95% confidence interval is 0.1 lower than the prior mean. For p_H^{max} and $p_{G_D}^{CHR}$, we assume
636 prior means of 0.75 and calibrate the prior so that the lower bound of the 95% confidence
637 interval is 0.2 lower than the mean. For the multipliers for hospital mortality after
638 improvement in care, μ_H , and for probability of admission to ICU after improvement in care,
639 μ_{ICU} , we used uninformative $U[0,1]$ priors.

640

641

642

Length of stay in each hospital state

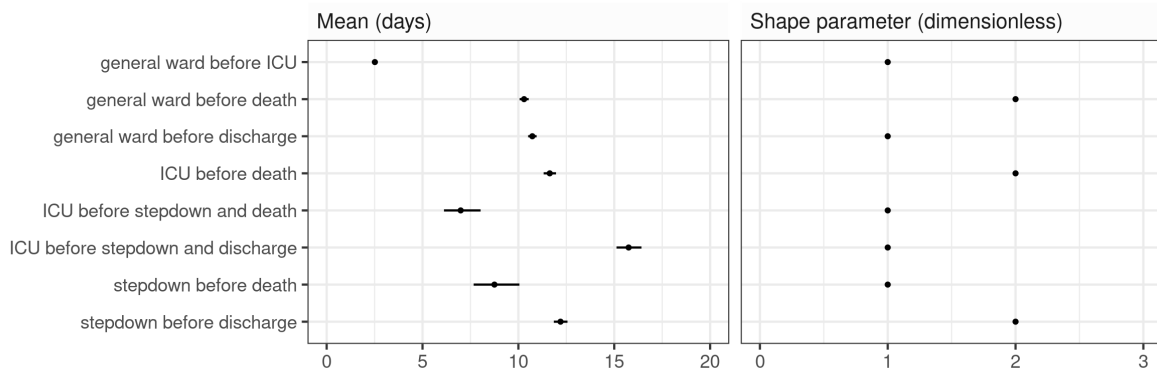


Figure S 5: Posterior 95% credible intervals of length of stay mean (left) and shape parameters (right).

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 $T_{seropre}^i$ with a proportion $p_{seropos}$ ultimately seroconverting and moving to $T_{seropos}^i$ and the remaining staying negative and moving to $T_{seroneg}^i$.

We fixed $p_{seropos}$ 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 $p_{serospec}$ is fixed to 0.99 also from (42). Finally, the sensitivity of serology test $p_{serosens}$ 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 $p_{P2_{test}}$, we used uninformative $U[0,1]$ priors.

697 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
698 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 $a/(a + b)$).
699

	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)											
$\beta(t)$	β_1	-	$\Gamma(136, 0.0006)$	$\frac{0.07}{(0.06, 0.08)}$	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	-	$\Gamma(21.9, 0.0026)$	$\frac{0.06}{(0.04, 0.08)}$	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	-	$\Gamma(4.25, 0.0079)$	$\frac{0.03}{(0.01, 0.07)}$	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	-	$\Gamma(4.25, 0.0079)$	$\frac{0.03}{(0.01, 0.07)}$	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	-	$\Gamma(4.25, 0.0079)$	$\frac{0.03}{(0.01, 0.07)}$	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	-	$\Gamma(4.25, 0.0079)$	$\frac{0.03}{(0.01, 0.07)}$	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	-	$\Gamma(4.25, 0.0079)$	$\frac{0.03}{(0.01, 0.07)}$	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	-	$\Gamma(4.25, 0.0079)$	$\frac{0.03}{(0.01, 0.07)}$	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	-	$\Gamma(4.25, 0.0079)$	$\frac{0.03}{(0.01, 0.07)}$	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)
	β_{10}	-	$\Gamma(4.25, 0.0079)$	$\frac{0.03}{(0.01, 0.07)}$	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}	-	$\Gamma(4.25, 0.0079)$	$\frac{0.03}{(0.01, 0.07)}$	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}	-	$\Gamma(4.25, 0.0079)$	$\frac{0.03}{(0.01, 0.07)}$	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	-	$U[0,1]$	$\frac{0.5}{(0.03, 0.98)}$	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		$\Gamma(5, 4.3 \times 10^{-7})$	$\Gamma(5, 3.7 \times 10^{-7})$	$\Gamma(5, 2.9 \times 10^{-7})$	$\Gamma(5, 5.2 \times 10^{-7})$	$\Gamma(5, 7.6 \times 10^{-7})$	$\Gamma(5, 4.9 \times 10^{-7})$	$\Gamma(5, 3.1 \times 10^{-7})$
					$2.2 \times 10^{-6} (7.0 \times 10^{-7}, 4.4 \times 10^{-6})$	$1.8 \times 10^{-6} (5.9 \times 10^{-7}, 3.7 \times 10^{-6})$	$1.5 \times 10^{-6} (4.7 \times 10^{-7}, 2.9 \times 10^{-6})$	$2.6 \times 10^{-6} (8.4 \times 10^{-7}, 5.3 \times 10^{-6})$	$3.8 \times 10^{-6} (1.2 \times 10^{-6}, 7.8 \times 10^{-6})$	$2.5 \times 10^{-6} (8.0 \times 10^{-7}, 5.0 \times 10^{-6})$	$1.6 \times 10^{-6} (5.1 \times 10^{-7}, 3.2 \times 10^{-6})$
			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)
p_{GD}^{max}	Probability of death in care home if requiring hospitalisation	ψ_{GD}^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_{HD}^{max}	Initial probability of death for general inpatients	ψ_{HD}^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_{ICUD}^{max}	Initial probability of death for ICU inpatients	ψ_{ICUD}^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_{WD}^{max}	Initial probability of death for stepdown inpatients	ψ_{WD}^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)
ρ_{P2test}	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

704 *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	CHES (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

705

706

707 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_{HD})	Death in ICU (p_{ICUD})	Death in stepdown (p_{WD})	Death in the community (p_{GD})
Population group	ψ_H^i	ψ_{ICU}^i	ψ_{HD}^i	ψ_{ICUD}^i	ψ_{WD}^i	ψ_{GD}^i
<i>Age</i>						
[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
[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
<i>Care home</i>						
CHW	0.104	0.784	0.134	0.519	0.114	0
CHR	0.873	0.107	1.000	0.918	1.000	1

708

2 Supplementary Results

2.1 Model fitting

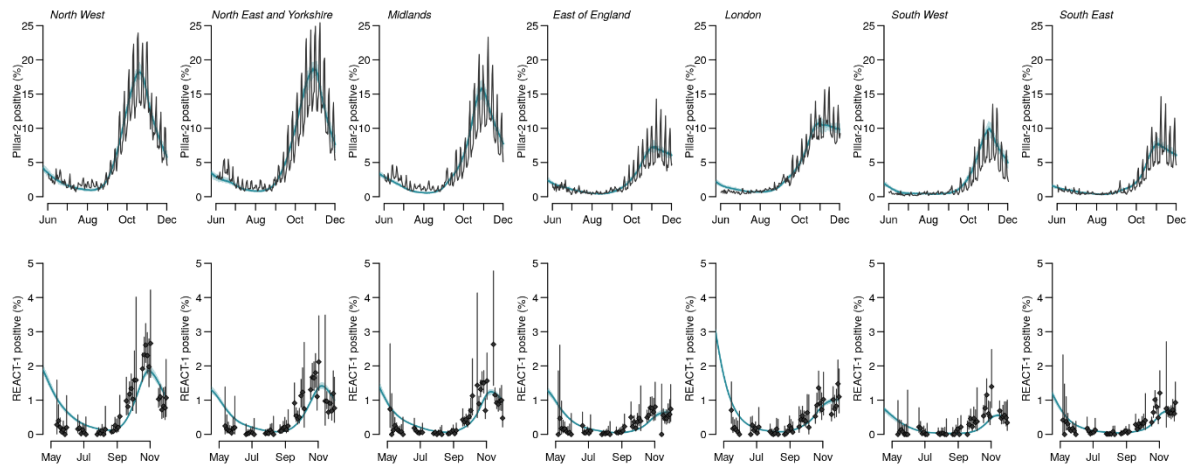


Figure S 6: Model fits by region to PCR positivity for individuals aged >25 years (top row) and PCR positivity from the REACT-1 study (bottom row). The points show the data and bars the 95% CI. The solid line the median model fit and the shaded area the 95% CrI.

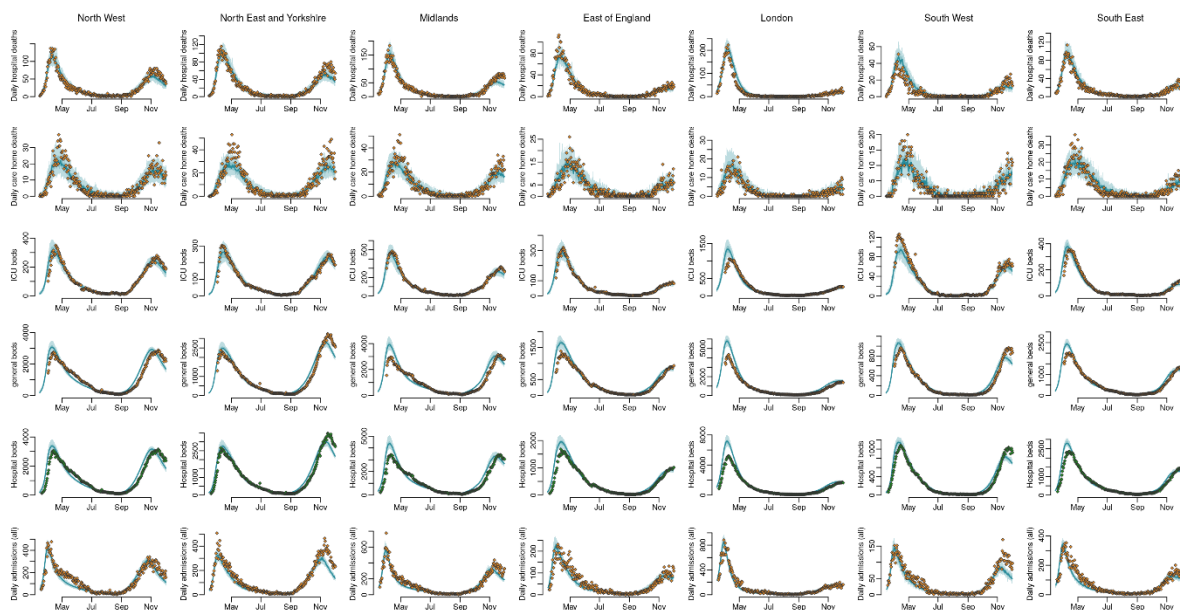
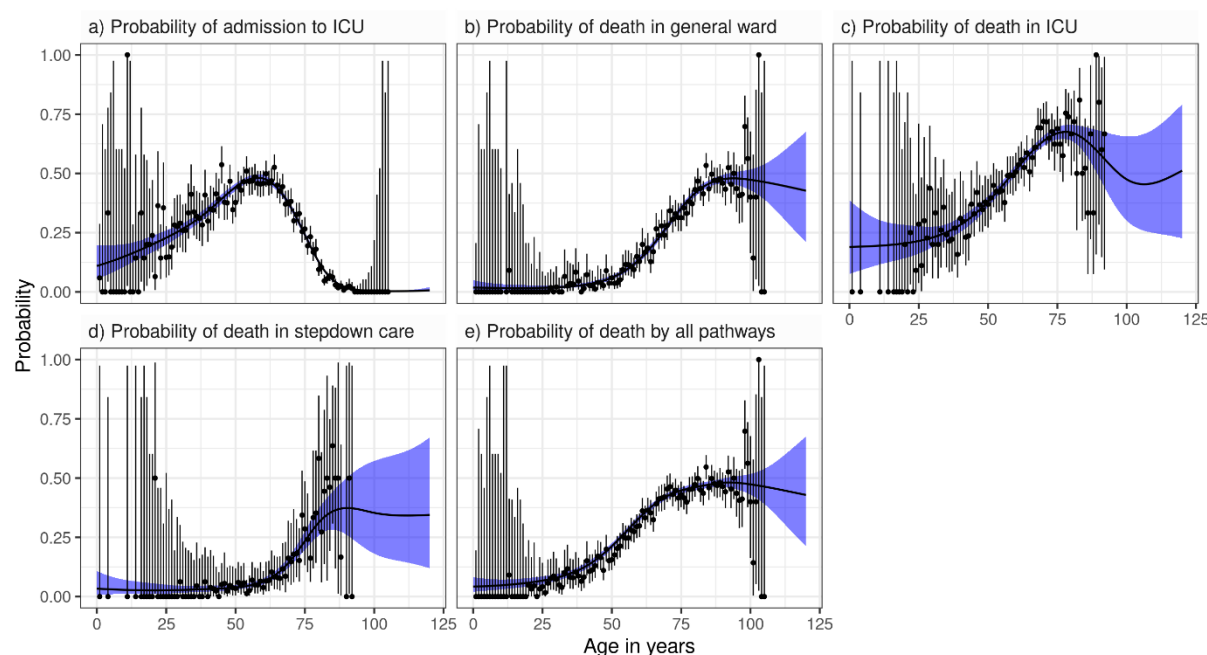


Figure S 7: Model fits to daily hospital deaths (top row), daily care home deaths (second row), ICU bed occupancy (third row), general bed occupancy (fourth row), all hospital beds (fifth row), and all daily admissions (bottom row) by region (columns). The points show the data, the solid line the median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

723 2.2 Severity estimates

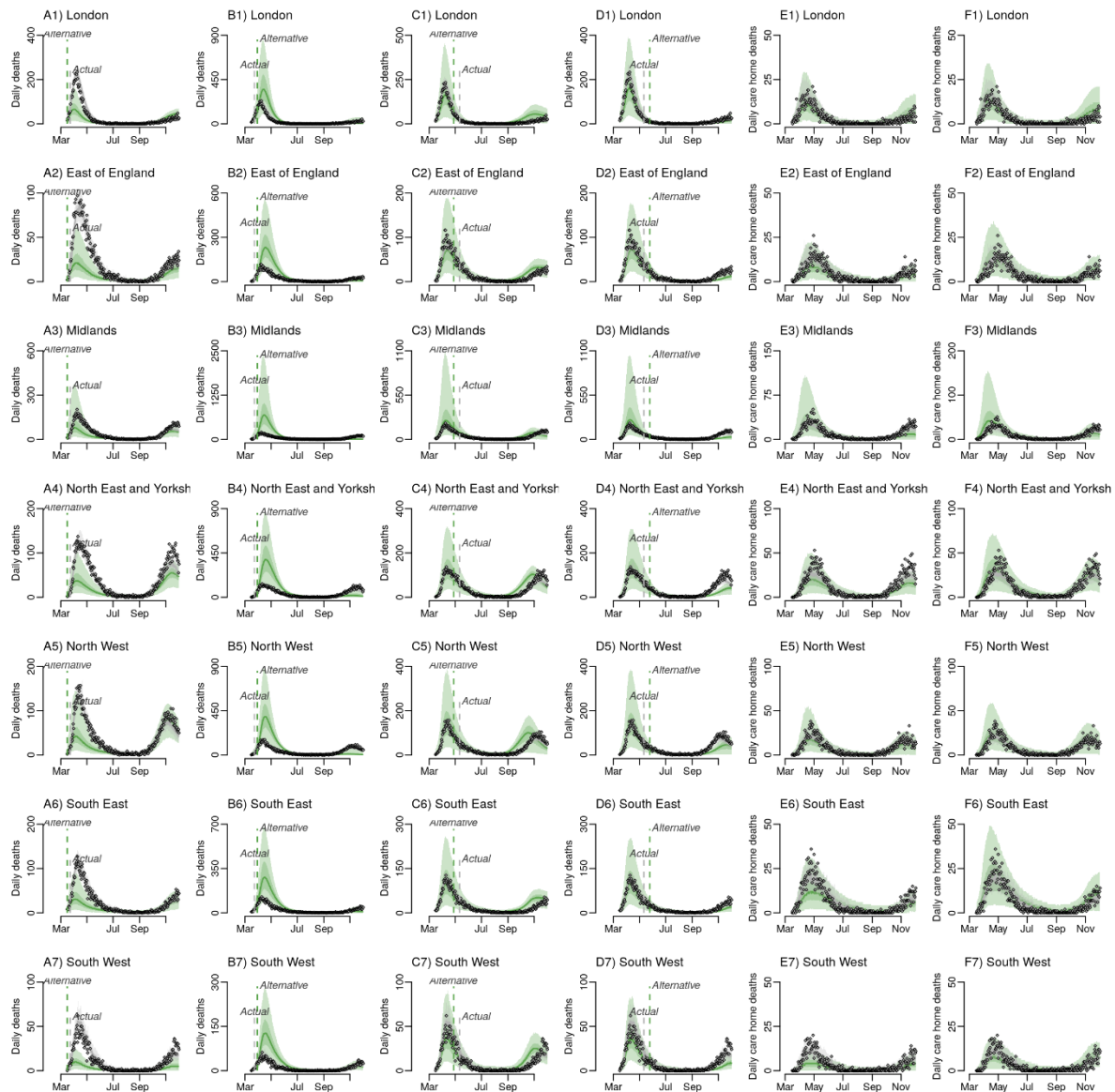


724
725 *Figure S 8: Fits to CHES data broken down into one-year age bands. Blue ribbons show the 95%*
726 *Crl of the fitted spline, black circles and vertical segments give the raw mean and 95% CI from the*
727 *data (exact binomial).*

728
729 *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%)

731 Supplementary counterfactual analysis



732
733 *Figure S 9: Counterfactual intervention scenarios in each England NHS Region: Panel A1-7 impact of*
734 *locking down one-week earlier Panel B1-7 impact of locking down one week later; Panel C1-7 impact*
735 *of relaxing lockdown restrictions two weeks earlier. Panel D1-7 impact of relaxing lockdown*
736 *restrictions two weeks later; Panel E1-7 impact of 50% less contact between care home residents and*
737 *the general population; Panel F1-7 impact of 50% more contact between care home residents and the*
738 *general population.*

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 CHES 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

- 786 covid-19. A comparative cohort study in nursing homes outbreaks. Arch Gerontol
787 Geriatr. 2020;91:104240.
- 788 16. Omar S, Bartz C, Becker S, Basenach S, Pfeifer S, Trapp C, et al. Duration of SARS-
789 CoV-2 RNA detection in COVID-19 patients in home isolation, Rhineland-Palatinate,
790 Germany, 2020 - an interval-censored survival analysis. Eurosurveillance.
791 2020;25(30):1–8.
- 792 17. Benny B, Amandine G, Kc P, Sarah H, Abby M, Caitlin C, et al. Quantifying antibody
793 kinetics and RNA shedding during early-phase SARS-CoV-2 infection.
- 794 18. Funk S. Socialmixr: Social Mixing Matrices for Infectious Disease Modelling. The
795 Comprehensive R Archive Network; 2018.
- 796 19. GOV.UK. Prime Minister’s statement on coronavirus (COVID-19): 12 March 2020 -
797 GOV.UK [Internet]. [cited 2020 Dec 3]. Available from:
798 [https://www.gov.uk/government/speeches/pm-statement-on-coronavirus-12-march-](https://www.gov.uk/government/speeches/pm-statement-on-coronavirus-12-march-2020)
799 [2020](https://www.gov.uk/government/speeches/pm-statement-on-coronavirus-12-march-2020)
- 800 20. GOV.UK. Prime Minister’s statement on coronavirus (COVID-19): 22 March 2020 -
801 GOV.UK [Internet]. [cited 2020 Dec 3]. Available from:
802 [https://www.gov.uk/government/speeches/pm-statement-on-coronavirus-22-march-](https://www.gov.uk/government/speeches/pm-statement-on-coronavirus-22-march-2020)
803 [2020](https://www.gov.uk/government/speeches/pm-statement-on-coronavirus-22-march-2020)
- 804 21. GOV.UK. Prime Minister’s statement on coronavirus (COVID-19): 25 March 2020 -
805 GOV.UK [Internet]. [cited 2020 Dec 16]. Available from:
806 [https://www.gov.uk/government/speeches/pm-statement-on-coronavirus-25-march-](https://www.gov.uk/government/speeches/pm-statement-on-coronavirus-25-march-2020)
807 [2020](https://www.gov.uk/government/speeches/pm-statement-on-coronavirus-25-march-2020)
- 808 22. GOV.UK. Prime Minister’s statement on coronavirus (COVID-19): 11 May 2020 -
809 GOV.UK [Internet]. [cited 2020 Dec 16]. Available from:
810 <https://www.gov.uk/government/speeches/pm-statement-on-coronavirus-11-may-2020>
- 811 23. GOV.UK. Prime Minister sets out timeline for retail to reopen in June - GOV.UK
812 [Internet]. [cited 2020 Dec 16]. Available from:
813 [https://www.gov.uk/government/news/prime-minister-sets-out-timeline-for-retail-to-](https://www.gov.uk/government/news/prime-minister-sets-out-timeline-for-retail-to-reopen-in-june)
814 [reopen-in-june](https://www.gov.uk/government/news/prime-minister-sets-out-timeline-for-retail-to-reopen-in-june)
- 815 24. GOV.UK. Pubs, restaurants and hairdressers to reopen from 4 July - GOV.UK
816 [Internet]. [cited 2020 Dec 16]. Available from:
817 [https://www.gov.uk/government/news/pubs-restaurants-and-hairdressers-to-reopen-](https://www.gov.uk/government/news/pubs-restaurants-and-hairdressers-to-reopen-from-4-july)
818 [from-4-july](https://www.gov.uk/government/news/pubs-restaurants-and-hairdressers-to-reopen-from-4-july)
- 819 25. GOV.UK. Eat Out to Help Out launches today – with government paying half on
820 restaurant bills - GOV.UK [Internet]. [cited 2020 Dec 16]. Available from:
821 [https://www.gov.uk/government/news/eat-out-to-help-out-launches-today-with-](https://www.gov.uk/government/news/eat-out-to-help-out-launches-today-with-government-paying-half-on-restaurant-bills)
822 [government-paying-half-on-restaurant-bills](https://www.gov.uk/government/news/eat-out-to-help-out-launches-today-with-government-paying-half-on-restaurant-bills)
- 823 26. GOV.UK. Schools and colleges to reopen in full in September - GOV.UK [Internet].
824 [cited 2020 Dec 16]. Available from: [https://www.gov.uk/government/news/schools-](https://www.gov.uk/government/news/schools-and-colleges-to-reopen-in-full-in-september)
825 [and-colleges-to-reopen-in-full-in-september](https://www.gov.uk/government/news/schools-and-colleges-to-reopen-in-full-in-september)
- 826 27. GOV.UK. Rule of six comes into effect to tackle coronavirus - GOV.UK [Internet].
827 [cited 2020 Dec 7]. Available from: [https://www.gov.uk/government/news/rule-of-six-](https://www.gov.uk/government/news/rule-of-six-comes-into-effect-to-tackle-coronavirus)
828 [comes-into-effect-to-tackle-coronavirus](https://www.gov.uk/government/news/rule-of-six-comes-into-effect-to-tackle-coronavirus)
- 829 28. GOV.UK. Prime Minister announces new local COVID Alert Levels - GOV.UK
830 [Internet]. [cited 2020 Dec 7]. Available from:
831 [https://www.gov.uk/government/news/prime-minister-announces-new-local-covid-](https://www.gov.uk/government/news/prime-minister-announces-new-local-covid-alert-levels)
832 [alert-levels](https://www.gov.uk/government/news/prime-minister-announces-new-local-covid-alert-levels)

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

836 30. Funk S. socialmixr @ github.com.

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

839 32. Diekmann O, Heesterbeek JAP, Metz JAJ. On the definition and the computation of
840 the basic reproduction ratio R_0 in models for infectious diseases in heterogeneous
841 populations. *J Math Biol*. 1990;28(4):365–82.

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

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

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

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

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

853 38. Gordon NJ, Salmond DJ, Smith AFM. Novel approach to nonlinear/non-gaussian
854 Bayesian state estimation. *IEE Proceedings, Part F Radar Signal Process*.
855 1993;140(2):107–13.

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

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

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

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

864 43. Jarvis CI, Van Zandvoort K, Gimma A, Prem K, Auzenberg M, O'Reilly K, et al.
865 Quantifying the impact of physical distance measures on the transmission of COVID-
866 19 in the UK. *BMC Med*. 2020;18(1):1–10.

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

870 45. NHS England and NHS Improvement. Statistics » COVID-19 Hospital Activity
871 [Internet]. [cited 2020 Dec 17]. Available from:
872 <https://www.england.nhs.uk/statistics/statistical-work-areas/covid-19-hospital-activity/>
873