Understanding the impact of modelling assumptions and population heterogeneity on the robustness of outputs of different epidemiological models in the context of the COVID-19 pandemic and beyond

Ioana Bouros

St Peter's College

Department of Computer Science

University of Oxford

Supervisors:

Prof. David Gavaghan

Prof. Ben Lambert

Prof. Robin Thompson

Dr Annabelle Lemenuel-Diot

Hilary 2025

This thesis is dedicated
to my MUM and DAD

Acknowledgements

I want to thank first and foremost my supervisors, David Gavaghan, Ben Lambert, Robin Thompson and Annabelle Lemenuel-Diot for their continued support and encouragement over the years. Your guidance and advice, and the many hours spent in meetings discussing and providing feedback on my work and writing have been indispensable in getting this project over the finish line. Thank you for helping shape me into the researcher I am today and for nurturing my curiosity and love for research.

Further acknowledgements go to the EPSRC and Roche for providing the funding to facilitate the completion of my DPhil. I would also like to thank the members of the various research groups, whose meetings I had the pleasure to join over the years, for providing ample opportunities to discuss and to dissect a variety of topics pertaining to infectious disease modelling, which have inspired the main goals of this work.

Finally, I would like to thank my parents, Paveluta and Cornel, for their unwavering guidance and love all these years. You have instilled in me the principles of hard-work and perseverance. I would not be here today without your support.

Abstract

The rapid emergence of the SARS-CoV-2 virus in late 2019 and its high virulence generated extensive public interest from both specialists and non-specialists alike and has highlighted the importance of mathematical modelling in epidemiology.

In this DPhil project, we explore how differing modelling approaches used in epidemiology can impact the recommendations made by scientists and scientific advisory groups to policy makers in Government, with a particular focus on the decisions made in the UK during the COVID-19 pandemic. The first major focus of this work revolves around a class of models popularly used in epidemiology, known as renewal models. These models require fewer assumptions to be made compared to alternative modelling techniques, yet still output quantities of relevance to policy. Classic renewal modelling approaches treat a population as a homogeneous group and ignore differences in contact patterns that may vary systematically across groups (e.g. that teenagers typically have more contacts per day than those aged 65+). We develop *de novo* renewal models to incorporate the effects of population heterogeneity for better prediction of public health outcomes, and derive straightforward and entirely new mathematical results to assess the long-term behaviour of a pathogen within a multi-layered population.

Finally, we compare three of the key compartmental models that were used by the UK government and the industrial research sector to model the COVID-19 UK epidemic. We reimplement these models in a unified framework using industrial-strength software engineering practices by creating an open-source package called 'epimodels' [Bouros, 2021]. This work highlights the pitfalls of relying on individual models to inform policy responses for future epidemics and pandemics, as well as the need for a more in-depth study of the impact of modelling assumptions on the quality of model outputs.

Table of Contents

Introduction **1**

Model types, methods & software **11**

0.1 Introduction to renewal equation models . . . 12
0.2 Introduction to compartmental models . . . 15
0.2.1 The SEIR Model . . . 17
0.3 Data . . . 19
0.3.1 Death Data . . . 19
0.3.2 Serology Data and Large-scale Infection Surveys . . . 20
0.3.3 Susceptibles Data – Counts, Fatality Rates & Mobility . . . 21
0.3.4 Interventions Data . . . 21
0.4 Bayesian Inference . . . 21
0.4.1 General Principles of Bayesian Inference . . . 22
0.4.2 Computational Challenges . . . 23
0.5 Software tools for inference and integration with PINTS . . . 25
0.5.1 Branchpro . . . 25
0.5.2 Epimodels . . . 26

1 Stability criterion for multiple population groups renewal processes **28**

1.1 Introduction . . . 29
1.2 Methods . . . 31
1.2.1 Renewal equations used to infer R_t . . . 31
1.2.2 Derivation of the renewal equation for the M'Kendrick–von Foerster model . . . 31
1.2.3 The M'Kendrick–von Foerster model of an epidemic in a homogeneous population . . . 34
1.2.4 Reproducing our simulation results . . . 38
1.3 Results . . . 38
1.3.1 A renewal equation for new infections in a structured M'Kendrick-von Foerster model . . . 38

1.3.2 The time-varying reproduction number for structured populations . 41
1.3.3 The long-run epidemic fate does not depend on the generation time interval . . . 46
1.3.4 Stochastic models have the same R_t threshold governing their mean behaviour . . . 47
1.3.5 The relationship between R_t and growth rate r for structured populations . . . 49
1.4 Discussion . . . 52

2 Impact of generation times on the accuracy of R_t estimation for multiple-group population renewal models 56
2.1 Introduction . . . 57
2.2 Methods . . . 59
2.2.1 Renewal equation models for structured populations . . . 59
2.2.2 Overall and group-specific R_t for multiple-group populations . . . 61
2.2.3 Reproducing our results . . . 63
2.3 Results . . . 63
2.3.1 Posterior distribution comparison: the one group versus the multiple group model . . . 64
2.3.2 Differences in R_t estimates between the one-group and the multiple-group models . . . 66
2.4 Application to real-world outbreaks . . . 81
2.5 Discussion . . . 83

3 Limitations caused by data availability and better policy prediction outcomes using structured renewal processes 86
3.1 Introduction . . . 87
3.2 Methods . . . 87
3.2.1 The baseline scenario – daily total number of infections and no contact information . . . 90
3.2.2 Daily infection data population group with exact daily contact data . 90
3.2.3 Daily infection data by population group with average contact information . . . 93

3.2.4 Daily total number of infections with average contact information . 93
3.2.5 Modelling Infrastructure 94
3.2.6 Reproduction Number Inference 95
3.3 Results 96
3.3.1 Estimated reproduction number trajectories in the multiple-group model versus one-group model 96
3.3.2 Improved targetting of interventions using the group-specific reproduction number 101
3.4 Discussion 110

4 A retrospective analysis of the robustness of existing compartmental ODE models for modelling future pandemics 114
4.1 Introduction 115
4.2 Methods 118
4.2.1 Three models for the COVID-19 epidemic: the PHE, Roche and Warwick-Household models 118
4.2.2 Software package 121
4.2.2.1 epimodels 122
4.2.2.2 Software Properties 122
4.2.2.3 Software Architecture 123
4.3 Results 126
4.3.1 Model equality and equivalence 127
4.3.2 Structural Identifiability 127
4.3.2.1 The PHE Model 129
4.3.2.2 The Roche Model 131
4.3.2.3 The Warwick-Household Model 134
4.3.3 Parameter Sensitivity Analysis 137
4.3.3.1 The PHE Model 138
4.3.3.2 The Roche Model 140
4.3.3.3 The Warwick-Household Model 143
4.3.4 Parameter Optimisation & Inference 145
4.3.5 Parameters values used for the model fitting 147
4.3.5.1 PHE model parameters 147

4.3.5.2 Roche model parameters . . . 148
4.3.5.3 Warwick-Household model parameters . . . 150
4.3.6 Parameter Optimisation . . . 153
4.3.6.1 The PHE Model . . . 154
4.3.6.2 The Roche Model . . . 156
4.3.6.3 The Warwick-Household Model . . . 158
4.3.7 Parameter Inference . . . 159
4.3.7.1 The PHE Model . . . 160
4.3.7.2 The Roche Model . . . 161
4.3.7.3 The Warwick-Household Model . . . 163
4.3.8 Scenario Model Comparison . . . 165
4.3.8.1 Model comparison: NPIs – on or off . . . 167
4.3.8.2 Model comparison: NPIs are applied at different times . . 172
4.4 Discussion . . . 174

Conclusion **177**

Bibliography **187**

Appendix A Appendix Chapter 1 **205**
A1 Thieme's renewal theorem . . . 205
A2 Renewal equations and reproduction numbers for a discrete-time structured population model . . . 208
A3 The typical generation time interval affects epidemic growth rates in calendar time . . . 210
A4 R_t for a discrete-time stochastic model . . . 211
A5 Doubling times . . . 212
A6 The relationship between $R(t)$ and the growth rate for the discrete-time deterministic model . . . 213
A7 Empirical estimates of the growth rate from simulations of the stochastic model . . . 213
A8 Identical rows matrix spectral radius . . . 214
A9 Exact overall reproduction number for a two-group population . . . 214
A10 Varying generation time distributions across groups. . . . 216

Appendix B Appendix Chapter 3 **218**

B1 The impact of the initial fraction of infection in the population on long-term R_t predictions . . . 218

B2 Poisson-Binomial renewal equation model – A different semi-mechanistic approach . . . 219

Appendix C Appendix I Chapter 4 **224**

C1 PHE Model . . . 224

C1.1 PHE Model Structure . . . 224

C1.1.1 Contact Matrices . . . 227

C1.1.2 Regional Matrices and UniGen Matrices . . . 230

C1.1.3 PHE Model Class . . . 231

C2 Roche Model . . . 234

C2.1 Roche Model Structure . . . 234

C2.1.1 Stringency Index . . . 239

C2.1.2 Roche Model Class . . . 243

C3 Warwick-Household Model . . . 245

C3.1 Warwick-Household Model Structure . . . 245

C3.1.1 Social distancing modelling . . . 249

C3.1.2 Warwick-Household Model Class . . . 251

C4 Example model implementations . . . 254

C4.1 PHE Model Implementation . . . 255

C4.2 Roche Model Implementation . . . 256

C4.3 Warwick Model Implementation . . . 259

C4.4 Comparison of the trajectories of daily infection incidence . . . 263

C4.4.1 The PHE Model . . . 263

C4.4.2 The Roche Model . . . 266

C4.4.3 The Warwick-Household Model . . . 268

C4.5 Comparison of the trajectories of daily number of deaths . . . 270

C4.5.1 The PHE Model . . . 270

C4.5.2 The Roche Model . . . 271

C4.5.3 The Warwick-Household Model . . . 272

C4.6 Comparison of the trajectories of time-dependent reproduction number 274

C4.6.1 The PHE Model 274
C4.6.2 The Roche Model 275
C4.6.3 The Warwick-Household Model 277
C5 Shortcomings of the existing form of the models 278

Appendix D Appendix II Chapter 4 **280**
D1 The PHE model 280
D2 The Roche model 283
D3 The Warwick-Household model 288
D4 Parameter relationships across the models 294

Appendix E Appendix III Chapter 4 **297**
E1 Changing the time-step of the Euler method 297
E2 Changing the tolerance of the Scipy solver 298

Appendix F Appendix IV Chapter 4 **300**
F1 Methods for loading the data 300
F2 Numerical and Computational Considerations 301

Introduction

Background

The year 2025 marks the fifth anniversary of the declaration of the COVID-19 pandemic. By the end of 2024 more than 750 million infections and 7 million deaths had been reported worldwide [WHO, 2025], though the true number of deaths is estimated to be up to three times higher [WHO, 2024a]. The early days of the epidemic were marked by strict lockdowns, mask mandates and other interventions of containment [Flaxman et al., 2020; Brauner et al., 2021], which brought the world to a halt, causing immeasurable long-term economic losses [World Bank, 2024]. At the same time, governments and the scientific communities all over the world joined together, in an extensive collaborative effort the likes of which were rarely witnessed before, to produce a number of efficacious vaccines against this highly contagious pathogen, with results being produced less than 12 months from the start of the pandemic [Tregoning et al., 2021]. Overall, the rapid emergence of the novel SARS-CoV-2 virus and its high virulence has generated extensive public interest from both specialists and non-specialists alike and has highlighted the importance of mathematical and computational modelling in epidemiology.

However, due to its novel nature at the time, no dedicated models of the SARS-CoV-2 virus transmission dynamics existed at the start of the pandemic, and scientists had to quickly repurpose existing models for other air-borne diseases (mostly influenza), or create new ones, to prospectively study the progression of the epidemic and therefore assess the effects of various proposed interventions. Consequently, many such models were developed, displaying a wide variety of model architectures. Just in the UK, eight separate research groups, each with their own developed model, were engaged by the Scientific Pandemic Influenza Group on Modelling consortium to consult with the responsible governmental bodies on the right package of interventions to be applied for mitigating severe epidemic outcomes. However, due to time constraints, these models suffered from rushed development, often with poorly engineered designs that provoked criticism of their reproducibility [Singh Chawla, 2020]. Additionally, there was little emphasis on validating the findings of these models at a large scale.

Now that the global emergency state has been declared over [WHO, 2023], it is important to take a step back and tackle the irreproducibility of current epidemiological models, through a rigorous assessment of the field at large. This is an essential step to increase future pandemic preparedness [United Nation, 2024], especially as the risk of new pandemics occurring has progressively hastened over the last century [Jones et al., 2008].

Epidemiological Modelling – Definition & Key Concepts

Epidemiological modelling is a branch of mathematical biology that focuses on creating models that replicate the dynamics of infectious diseases in populations [Anderson and May, 1991; Payne, 2017]. By predicting how the epidemic evolves under different assumptions, these models can be used to inform public health interventions by government bodies [Anderson et al., 2020; Thompson et al., 2020; Flaxman et al., 2020; Brauner et al., 2021] (e.g. vaccination campaigns [Hogan et al., 2021; Watson et al., 2022], partial or full lockdown of cities or regions [Ferguson et al., 2020]).

There are many biologically relevant quantities which dictate the present and future dynamics of an epidemic. For example, one such quantity is the time-dependent effective reproduction number. The effective reproduction number is the expected number of secondary infections caused by one infective host over the duration of their infection if transmission remains the same [Fraser, 2007; Nishiura and Chowell, 2009; Flaxman et al., 2020; Anderson et al., 2020]. A reproduction number below 1 indicates that the number of new infections will decrease, while a value above 1 indicates that the number of cases is expected to rise. There are multiple ways in which the reproduction number can be calculated: the case reproduction number measures the average number of secondary cases arising from a primary case infected at time t throughout their infectious period; the instantaneous reproduction number is the average number of secondary cases that would be generated by an infected case at time t assuming that future transmission remains the same. The instantaneous reproduction number (or instantaneous R_t) is a measure of the 'real-time' epidemic behaviour, as it is sensitive to changes in the behaviour of the population modelled or of the biology of the virus itself [Fraser, 2007; Nishiura and Chowell, 2009; Thompson et al., 2020], as well as to the number of remaining susceptible individuals. Estimating the instantaneous reproduction number is therefore crucial for evaluating the effects of different measures such as vaccinations, treatments or non-pharmaceutical interventions [Flaxman et al., 2020; Brauner

et al., 2021]. Additionally, the instantaneous R_t can be further used to assess whether stricter measures need to be imposed in order to control an epidemic or whether those already in place can be relaxed – usually by comparing the predicted number of people who will get infected, hospitalised or die under different interventional scenarios.

Types of epidemiological models

Two key types of models used in the modelling of epidemics are renewal models and compartmental models.

In practice, renewal models assume a stochastic framework in which we solely model the number of new infections per day [Fraser, 2007; Nishiura and Chowell, 2009; Thompson et al., 2019; Flaxman et al., 2020; Abbott et al., 2020]. Each new case arising in the community is assumed to have been generated by a previously occurring infection, where the time lags between parent and daughter infections are characterised by the *generation time distribution* w [Cori et al., 2013; Thompson et al., 2019] – which is defined such that the s^{th} element of the vector, w_s, is equal to the probability that the daughter case is generated by an infection occurring s days ago. The number of new infections on day t, I_t, depends on the instantaneous reproduction number and the historic timeline of the cases, which is then perturbed by a carefully-chosen noise distribution, usually Poisson – that is,

$$I_t \sim \text{Pois}(R_t \Lambda_t), \text{ where } \Lambda_t = \sum_{s=1}^{t-1} w_s I_{t-s}$$

where Λ_t is the transmission potential.

The second major class of models, compartmental models, track all individuals in the population by grouping them into compartments based on a defining characteristic – usually their current status in the infection cycle [Kermack et al., 1927; Wang et al., 2021; Keeling et al., 2021]. Compartmental models can be both stochastic and deterministic and include the 'SEIR' ordinary differential equations models in which individuals are grouped into the following compartments: S for susceptibles, E for exposed, I for infectives and R for recovered. These individuals move between the compartments with prescribed rates.

Compartmental models can be further refined for our purposes, by adding more compartment labels to the population classification (e.g. consider presymptomatic, asymptomatic

and superspreader individuals [Lemenuel-Diot et al., 2020; Keeling et al., 2021] or hospitalised cases [Bosetti et al., 2021]), adding age-structure [Birrell et al., 2021], or accounting for vaccination [Bosetti et al., 2021; Bouros et al., 2024]. Models with a vaccinated compartment are often paired with waning immunity – in the case of COVID-19, immunity lasts for a period of a few months up to a year [Leidi et al., 2021; Abu-Raddad et al., 2021; Gallais et al., 2021]. For long-term simulations, we can also include ageing of the population in our compartmental model framework; deaths caused by other factors apart from the infections and births [Ballard et al., 2017] can also be integrated. We can also account for other types of heterogeneity in the population – for example, by allowing heterogeneous numbers of interactions across age groups through the use of contact matrices. Contact matrices are matrices with elements, C_{ij}, where each element denotes the average number of daily contacts between an individual in age class i with those of age class j [van der Vegt et al., 2022; Prem et al., 2021]. While the inclusion of contact matrices in the compartmental model framework can allow more realistic modelling of contact patterns between age groups, it also requires more assumptions to be made about the pathogen effects on each group.

Challenges of modelling of COVID-19

A great number of new models were developed [Birrell et al., 2021; Keeling et al., 2021; Bosetti et al., 2021; Lemenuel-Diot et al., 2020], updated [Moore et al., 2022; Bouros et al., 2024] or repurposed for modelling the COVID-19 epidemic – for example, the models described in [Ferguson et al., 2020] and [Danon et al., 2021] were both previously constructed to study the impact of different interventions for an outbreak of influenza. Contributors to the UK's Scientific Pandemic Influenza Group on Modelling (SPI-M) advisory committee used some of these independently-developed modelling methodologies to answer the same policy questions [Anderson et al., 2020; Brooks-Pollock et al., 2021].

However, despite the emergence of multiple models that could explain the evolution of the pandemic in different regions of the globe, no unified framework exists; moreover, conflicting results often arise when the same data are used with different modelling assumptions [Purkayastha et al., 2021]. In addition, these models are often described only at a high level in the published literature, hence making them difficult to reproduce. Additionally, the software tools used in these studies are often not publicly available or, even when they are, the code is not developed using robust software development practices. For example,

the now well-known model developed by the group at Imperial College [Ferguson et al., 2020; Adam, 2020] came under scrutiny early in the pandemic precisely for its software engineering approach [Chawla, 2020].

The lack of transparency of the software makes it particularly difficult to compare models and properly assess the quality of their outputs. For example, in this thesis we look into the Public Health England (PHE) model, which is a model that was developed in partnership with the University of Cambridge. Though the architecture of the model is discussed in an article from May 2021 [Birrell et al., 2021], the data sources it utilizes for fitting are either poorly described or cannot be accessed online due to privacy concerns; in addition to this, the initial software is not publicly available. Moreover, no updates to the model structure have been publicly released, although it continued to be used for policy making for some time afterwards. Additionally, even when access is provided to the private repository in which the code for the forward-simulation and inference of the model are stored, as was the case with the Warwick-Household model [Keeling et al., 2021], the details of the original paper and the current version of the code may not align; paired with a lack of clear documentation of the code, the process of reimplementation of the model can become excruciatingly laborious, if not impossible.

It is important to note, however, that the majority of these models have been developed or altered at a time when the scientific community was required to provide a swift response to the responsible authorities on the course of action that would most effectively contain the spread of the UK COVID-19 epidemic. Now that these time-constraints are no longer an issue, it is important to reflect on the pros and cons of these implemented methodologies and consider what lessons might be learned to improve our response to a future pandemic.

Modelling vaccination

In late 2021 the first COVID-19 vaccine was deployed [Tregoning et al., 2021]; with this development, previously used models were redesigned accordingly. By using age-specific contact data and information about the efficacy of a vaccine for different age categories, we can develop pharmaceutical interventions which have a greater impact on the spread of the virus: e.g. we might vaccinate first those groups most at risk, or those with highest expected number of contacts, depending on which scenario yields the best results in terms of number of infectives, deaths and/or hospitalised cases [Hogan et al., 2021; Moore et al.,

2022; Bouros et al., 2024]. However, waning immunity, breakthrough reinfections [Abu-Raddad et al., 2021; Leidi et al., 2021; Gallais et al., 2021], differences in access [Watson et al., 2022] and receptiveness to a vaccine [Lin et al., 2021] make this problem non-trivial, as the models which aim to account for these features become more complex.

Data used to calibrate the models

The outputs of any epidemiological model depend on the choices of parameters [Anderson et al., 2020] and must be fitted to the data available. There are multiple sources of data typically used to parameterise epidemiological models of COVID-19. Most commonly used are the daily number of deaths [Birrell et al., 2021; Hogan et al., 2021; Wang et al., 2021; Keeling et al., 2021] and the daily number of cases [Birrell et al., 2021; Hogan et al., 2021; Wang et al., 2021]. However, asymptomatic and presymptomatic transmission are not uncommon in COVID-19 [Nishiura et al., 2020; Mizumoto et al., 2020; Koelle et al., 2022], which means that daily cases data are not equivalent to the actual total number of individuals infected daily [Flaxman et al., 2020]. In this case, the numbers of positive results of antigen and antibody tests [Birrell et al., 2021] can be used as indicators of the proportion of the population currently infected (in the case of antigen-testing), or the proportion of the population currently and previously infected by the virus (in the case of antibody-testing). Hospitalisation counts [Bosetti et al., 2021; Pellis et al., 2021; Keeling et al., 2021] also provide additional information which may help to parameterise those model features relating to disease severity.

Other data requirements

In the previous section we introduced some example of data sources that can be used in the parametrisation process of a chosen epidemiological model. However, all of them are by-products and direct epidemic outputs, and are thus entirely dependent on the underlying population structure, interactions and environmental conditions that have given rise to the epidemic. Therefore, we require additional data sources that place our model conditions into the right epidemic context. For example, we use demographic data, such as the population structure by ages from census data [ONS], to specify the unalterable properties of the population we are aiming to study. Behavioural data, conveyed through social studies such as contact surveys [Prem et al., 2021; Google], can also be used to prescribe the manner in

which different population groups interact with each other and how these mixing patterns vary over time. Finally, we can also use clinical studies data, e.g. on the average duration of immunity after recovery [Leidi et al., 2021; Abu-Raddad et al., 2021; Gallais et al., 2021] or on the infection and case fatality [Verity et al., 2020], when available and reliable, to impose realistic values for those parameters governing the pathogen-specific transmission mechanisms.

Aims

In the epidemiological literature, there exists a plethora of modelling frameworks that can be applied to study the transmission dynamics of diseases (compartmental ODEs, branching-processes, agent-based, etc.). Due to the vastness of choice and flexibility in setting up these models, it can be very difficult to choose a particular model configuration to apply to a specific epidemic problem, especially if the epidemic in question is caused by a less-studied pathogen. In this thesis, we explore the extent to which model misspecification impacts the accuracy of epidemic outcomes, using two distinct modelling frameworks:

1. **Renewal equation models**: in which we keep track only of the daily incidence of infections. In the renewal equation models, the incidence on the current day is given by the weighted average of the transmission contribution of the previous days' incidence;

2. **Compartmental ODE models**: which assume the population is split into compartments based on their infection status. Each individual in the population belongs to exactly one compartment at each time point. The compartmental model uses ordinary differential equations to track the changes in the overall number of individuals (or the fraction of the total population) in each compartment at any given simulation time.

For each of the two proposed modelling frameworks we will assess the overall suitability of current epidemiological modelling practices in the context of a major infectious disease threat – with a focus on COVID-19.

In the first three chapters, we explore a novel approach to integrating heterogenous patterns of population mixing and infection risk in the renewal equation model structure. To this end, starting from first principles of structured population modelling, we compute and validate a definition for the overall reproduction number as an indicator for the long-term behaviour

of an epidemic in the context of a renewal equation process with multiple interacting population groups. We then go on to compare how the estimated overall reproduction number trajectory generated using the multiple-group population model compares to the classical one-group approach. This will help us determine in which circumstances the two modelling frameworks are equivalent. Additionally, we develop an approach for weighting the overall generation time distribution of the population in terms of the respective group-specific generation time distributions. We then go on to use synthetic and real epidemic data from the A/H1N1 outbreak in Japan from 2009 [Statistics Bureau of Japan] to visually validate the results. Finally, we compare four different renewal modelling frameworks with different levels of knowledge of the population contact matrix in order to assess whether there are any advantages of incorporating population heterogeneity, how great of an improvement collecting more detailed epidemic data brings to the prediction of public health policy outcomes, and how targeting different population groups impacts the effectiveness of interventions. These analyses will help us to:

- determine whether there are any circumstances in which we need to account for multiple population groups when using renewal equation models;
- assess the improvement in accuracy in the outputs of interests, i.e. the reproduction number trajectory, resulting from incorporating heterogenous mixing patterns;
- identify what level of heterogeneity information is sufficient for improved predictions of the impact of targeted interventions;
- demonstrate that the careful targeting of population groups can ensure similar reductions in transmission as blanket interventions would; and
- extend the application of current modelling tools for renewal equation models, such as 'branchpro' [SABS-R3-Epidemiology, 2025].

In the final chapter of the thesis, we turn our focus to compartmental models for COVID-19 transmission and how differences in the modelling assumptions impact the conclusions that can be drawn when the epidemic context is changed, for example when different non-pharmaceutical interventions (NPIs) such as lock-downs are applied. By doing so, we aim to:

- determine which prominent models used for policy-making during the pandemic pro-

vided the closest representation of the available data;

- understand the limitations of different types of models;
- explore the differences in the predicted trajectories of the outcomes of interest of different compartmental ODE models;
- identify best practices for fitting these models to data and for summarising their sensitivity to the underlying assumptions;
- produce robust open-source epidemiological modelling software in the form of a Python package called 'epimodels' [Bouros, 2021] for compartmental ODE modelling. This package can be used by others to model future pandemics of respiratory pathogens.

In the following, we present the four main steps we take to address the aims set out for chapter 4:

1. **Use professional software development practices to create an open-source Python package which allows simulation of the models.**
    - All models are coded up using the same template and are included in 'epimodels'. This Python package works as a library of different epidemiological models, and includes example notebooks to showcase the code functionalities.
    - The three models included in 'epimodels' are:
        - the PHE model [Birrell et al., 2021],
        - the Roche model [Lemenuel-Diot et al., 2020] and
        - the Warwick-Household model [Keeling et al., 2021].
2. **Fit the models to available data.**
    - For all our models and analyses we use the UK epidemic data collected from various sources:
        - *Cases and Death data*: age-structured daily number of deaths from the Office of National Statistics (ONS) [GOV.UK];

  - *Serology and Infection survey data*: antigen and antibody positive test results from the REACT1 and REACT2 [Nicholson et al., 2021] studies;
  - *Mobility and Contact data*: age-structured contact matrices extracted from Google mobility [Google] and POLYMOD matrices [Prem et al., 2021];
  - *Intervention data*: levels of severity of the different non-pharmaceutical interventions implemented in the UK and their time of application, as recorded by the OxCGRT group [Hale et al., 2021].

- By solving the inverse problem, we aim to understand the breadth and limitations of the information these types of epidemic datasets convey when it comes to parameter identifiability.
- This will also provide information on how to communicate uncertainty in large models of epidemiology, particularly when those models are likely misspecified [Anderson et al., 2020].

3. **Perform structural identifiability and sensitivity analyses on the fitted models to identify which parameters can be identified from the data.**
   - We will first identify candidate model parameters that are theoretically retrievable when fitting the data to the models.
   - We then go on to assess practically the identifiability of these parameters, by comparing model outputs from synthetic datasets generated for a range of parameter values.
   - Once these parameters are identified, we will validate our findings by using the UK epidemic data with the fitted models.
4. **Assess the models in terms of their predictive accuracy and sensitivity to simulation assumptions, such as implemented interventions.**
   - We compare different intervention scenarios for the first wave of the UK COVID-19 epidemic wave for all three of the models in order to determine the impact of non-pharmaceutical interventions in mitigating epidemic transmission and severe outcomes.
   - All models compared in this step are fitted using the same data and similar fixed

parameter values. Previous studies showed that differences in model assumptions can lead to significant changes in the model outputs [Purkayastha et al., 2021], so we expect to observe a similar degree of variability in the range of estimates produced by our models.

5. **Identify recommended courses of action for future pandemics.**

   - From a policy-making viewpoint, we hope our findings will convince the scientific community at large of the unreliability of some of the current compartmental ODE models and of the need to shift our focus from developing more-and-more complex compartmental ODE models to instead ensuring the accuracy of the existing ones.

   - Additionally, through this work we aim to persuade that providing robust, well-engineered software is essential in ensuring the reproducibility of newly developed models of infectious disease dynamics.

Model types, methods & software

In this preliminary chapter, we will provide a more in-depth description of the methods, underlying justifications and assumptions governing the two model formalisms we will focus as part of the thesis – that is renewal equation models and compartmental ODE models, as well as a comprehensive overview of the software tools we developed as part of this DPhil project to help us analyse and answer the questions posed back in the Aims section of the introductory chapter.

0.1 Introduction to renewal equation models

Due to the small number of parameters that need to be fitted and the small amount of data required to do so, renewal equation models have been increasingly used in the last ten years to perform early-epidemic predictions of the trajectory of the reproduction number. The instantaneous reproduction number R_t is defined as the average number of secondary cases that would be generated by an infected case at time t assuming that future transmission remains the same as at time t. A reproduction number value smaller than one is indicative of a declining epidemic, while a value greater than one suggests long-term epidemic growth. Therefore, the temporal trajectory of R_t can be used to identify when stricter measures need to be imposed in order to control the epidemic, or when those already in place can be relaxed. Hence, the estimation of the R_t number is ubiquitously now one of the first analyses performed when studying the progression of an emergent epidemic.

Infectious individuals transmit the disease to other individuals in the population – which will then become what we call secondary infections, on the days following infection. By the definition of the reproduction number, each infection recorded at time t will lead to R_t secondary cases if conditions remain unchanged. In its classic definition, the renewal equation model assumes that each new case registered today is a consequence of a case seen in previous days. The parent infection is selected from among the previous days' incidences according to a probability vector w, known as the generation time distribution [Cori et al., 2013; Thompson et al., 2019] which represents the duration between the primary and offspring cases. Therefore, the number of infections to occur on day t, I_t, can be summarised

by the following renewal equation describing the infector-infectee dynamics within the population:

$$I_t = R_t \Lambda_t, \text{ where } \Lambda_t = \sum_{s=1}^{t-1} w_s I_{t-s},$$

where Λ_t is the transmission potential, given by the weighted average contribution of previous infection incidence.

This deterministic modelling framework can be then made stochastic by adding some noise distribution on the number of new infections per day [Fraser, 2007; Nishiura and Chowell, 2009; Thompson et al., 2019; Flaxman et al., 2020; Abbott et al., 2020] to reflect the variation in the recorded incidence. Most commonly, this noise distribution is assumed to follow a Poisson probability distribution; this is because Poisson processes are frequently used to model the number of new arrivals in a stochastic process. As the parameter of the Poisson distribution equates to the mean number of arrivals, this choice would seem sensible:

$$I_t \sim \text{Pois}(R_t \Lambda_t), \text{ where } \Lambda_t = \sum_{s=1}^{t-1} w_s I_{t-s}, \tag{1}$$

where Λ_t is the transmission potential. However, the arrivals in a Poisson stochastic process are independent of each other and occur at the same rate – that is, all susceptible individuals are equally as likely to become infected. The Poisson-noise distribution of the subsequent cases in the model above hinges on the assumption that there is no variation in the individual number of offspring on day t associated with each previous infection recorded on day $t - s$, that is:

$$\nu_{t-s} = R_t w_s.$$

Summing across all previous infections leads us to eq. (1). This is not always the case; in practice, 80% of outcomes are the result of 20% of all possible causes [Woolhouse et al., 1997]. Indeed, similar to [Lloyd-Smith et al., 2005], a more realistic framework is to add a gamma-noise distribution on the individual number of offspring ν_{t-s}, such that instead:

$$\nu_{t-s} \sim \text{Gamma}(R_t w_s, \phi),$$

where $\text{Gamma}(a, b)$ denotes a Gamma-distribution with mean a and dispersion b. This allows us to account for super-spreader transmission events, which leads to the alternative

approach of modelling the number of new infections recorded at time t with a negative binomial noise distribution:

$$I_t \sim \text{NegBin}(R_t\Lambda_t, \phi), \text{ where } \Lambda_t = \sum_{s=1}^{t-1} w_s I_{t-s},$$

where Λ_t is the transmission potential and ϕ represents the 'overdispersion' parameter. The mean μ and variance σ^2 of the negative binomial are thus expressed by $\mu = R_t\Lambda_t$ and $\sigma^2 = \mu + \phi\mu^2$, respectively. Alternatively, we can instead start from first principles with a semi-mechanistic renewal model by adding noise to the number of contacts individuals accrue each day, as detailed in Appendix B2.

However, all of these models share the same weakness: they assume each individual in the population to mix at similar rates and be equally at risk of becoming infected, which is usually an unrealistic restriction. Indeed, in most cases there exist discrepancies in the infection risk and contact behaviour of different demographic groups, which may impact the suitability of using standard renewal equation models where these sources of population heterogeneity are discounted. Moreover, using a modelling framework which lumps everyone together instead of a more detailed multiple-group approach can suggest blanket interventions where in fact more targeted approaches are more efficient and place a lower burden on the population. Therefore, it is important to develop new and improved methodologies for assessing the overall reproduction number in the context of a highly heterogeneous population, which preserve the simple and easy-to-compute nature of renewal-equation models.

In the first chapter of this thesis, we present a novel renewal equation framework that accounts for the heterogeneity in infectious risks and contact patterns that does not rely on the computation of any additional quantities, such as the next-generation matrix [Green et al., 2022] or the vector of ratios of active infections in each population group [Parag and Obolski, 2023]. We derive from first principles the expression for the time-dependent reproduction number R_t for both a homogeneous (one-group) and heterogeneous (multiple-group) population, for both the deterministic and stochastic renewal equation model. We then go on to illustrate how this definition of R_t relates to the growth rate of the epidemic and the doubling time.

Over the second chapter, we analyse the effect of the group-specific generation time interval on the predicted trajectories of the overall R_t, and identify in which contexts the

multiple-group and the one-group renewal equation models are equivalent. We then go on to demonstrate an appropriate methodology for computing the overall generation time interval of a multiple-group population so that the estimated overall R_t profiles generated by the one-group and the multiple-group renewal equation models align in the long run. We finish the chapter with a case study of the 2009 A/H1N1 outbreak in Japan; we use this example to showcase the changes in the predicted R_t and in the time of the epidemic decline when we do not take into consideration the heterogeneous structure of the population.

Finally, in chapter three, we aim to get a better understanding of the degree of heterogeneous mixing we need to incorporate in the multiple-group renewal models to bring about an improvement in the accuracy of R_t. To do so, we compare four different instances of a multiple-group renewal equation model, with different levels of information about the transmission heterogeneity, in particular contact behaviour and case data aggregation. We also investigate using synthetic data how careful implementation of interventions in those population groups carrying the largest epidemic load – that is, those groups in which the group-specific reproduction number exceeds one – induces similar reductions in the number of predicted cases as overall blanket interventions; additionally, we determine that targetting interventions to those population groups carrying low epidemic burdens instead reduces significantly the efficacy of the implemented measures.

0.2 Introduction to compartmental models

In the Introduction to this thesis, we discussed the importance of epidemiological modelling and its potential direct impact in influencing health policy decisions for mitigating disease spread. However, there is currently no consensus on a clear modelling framework to be used to produce rigorous forecasts and intervention analyses for any type of disease. Indeed, there exists a plethora of available models developed by different scientific research groups from around the globe which are used to study the transmission dynamics of the same pathogen. For example, just in the UK, eight different models have been used to study the effects of a range of contact-reduction measures on the progression of the COVID-19 epidemic during the first wave in 2020. All of them rely on a compartmental model framework, which relies on the assumption that the population can be split into multiple groups (called 'compartments'), according to a defining feature. Each individual in the population belongs to one

and only one compartment at any given time and all individuals in each compartment behave identically and have exactly the same response to infection. Additionally, they can migrate from one compartment to another if certain conditions are met with prescribed transition rates.

There are many ways in which the population can be split – for example by differentiating infectious individuals based on their age, regional affiliation, or whether or not they display symptoms. This flexibility in setting up the compartments of the model leads to a high number of possible models to choose from. This becomes especially relevant when studying the transmission of a novel pathogen (as was the case with SARS-CoV-2) for which no prior knowledge of these dynamics exists. In this case, the choice of the appropriate model becomes non-trivial; indeed, during the early days of the COVID-19 global pandemic, scientists often repurposed models previously used to study the transmission dynamics of other air-borne diseases such as influenza, e.g. the Ferguson model [Ferguson et al., 2020] or the Roche model [Lemenuel-Diot et al., 2020]. Others simply extended the basic SEIR model – the simplest model capable of capturing the dynamics of an air-borne disease with a significant latency period from the infection event until the onset of infectiousness or symptoms, and incorporated particular transmission features as was the case with the PHE model [Birrell et al., 2021] and the Warwick-Household model [Keeling et al., 2021]. However, it is not immediately apparent whether or not these models are interchangeable – that is, whether we arrive at similar conclusions and insights into the effect of non-pharmaceutical interventions when substituting one model for the other.

An additional source of variation arises from the fact that the compartmental models commonly used in epidemiology tend to have a large number of parameters. In most cases, we can use clinical studies to provide information on the expected values of these parameters, e.g. for the infectious period, which is defined as the average length of time for which an individual remains in an infectious state. However, this is not always possible. Instead, we use observed epidemiological data such as the number of daily cases, or the number of deaths, to infer the parameter values of the model that produce outputs similar to the observed data. In this thesis, we take a Bayesian approach to the parameter inference, as we will discuss later in section 0.4 and we will use the pre-existing software package PINTS [Clerx et al., 2019] to run our inference algorithms, as described in section 0.5.1.

To this end, in the final chapter of the thesis we reproduce and analyse three compartmental

epidemiological models for the transmission dynamics of the SARS-CoV-2 virus; two of which were developed and used by separate research teams from the Universities of Cambridge (the PHE model [Birrell et al., 2021]) and Warwick (the Warwick-Household model [Keeling et al., 2021]) respectively to advise the UK's Scientific Pandemic Influenza Group on Modelling (SPI-M) committee on a set of preset policy topics – such as the effect of different interventions schemes. The third model (the Roche model [Lemenuel-Diot et al., 2020]) was developed by the Roche pharmaceutical company to study the transmission dynamics of influenza-type pathogens and then repurposed for SARS-CoV-2 once the 2020 COVID-19 epidemic took off. All of these models have in common their SEIR-type model structure (we introduce the basic SEIR model in the section immediately below) which is then extended to better represent the transmission features of SARS-CoV-2 deemed most important in driving the overall dynamics.

In Appendix C, we provide a more in-depth description of the structure of the three models, and for each of them we present example outputs of interest (daily deaths, cases and the time-dependent reproduction number). We reproduce all three models in a common, purpose-designed framework using the software package 'epimodels' [Bouros, 2021], for which we give a brief overview in section 4.2.2.1. In chapter 4 and Appendix D, we explore how interchangeable these three models are compared to one another and where this is not the case, whether this is a universal trait, or there exist certain parameter regimes for which all three models lie at the same spot in the 'hyper-model' space i.e. the space of all possible compartmental models that can be chosen. We will cover the second topic in Appendix D, where we investigate under which parameter conditions each of the PHE, Roche and Warwick-Household models collapse into the SEIR model and use the equivalence in order to determine how the parameters of these three models relate to one another. We cover parameter inference, as well as the comparison of model outcomes when the intervention conditions are changed, in Chapter 4.

0.2.1 The SEIR Model

A typical compartmental representation of a population affected by a non-lethal respiratory pathogen with delayed onset of symptoms and no possibility of reinfection is the so-called SEIR model. In this characterisation, an individual can either:

- be not yet infected, making them *susceptible* to the infection, i.e. they are part of the

S compartment;

- have been exposed to the virus, but are yet to manifest symptoms or transmit the virus further (hence, they are part of the E compartment);
- are currently infectious (part of I) and can spread the virus, or
- have already been infected with the virus and have since recovered (R).

The infection of a new individual, the onset of symptoms / infectiousness or recovery are the events prompting the individual to change their assigned compartment from S to E, E to I, and I to R respectively. These events occur with a given frequency, or 'rate'. The model tracks the differences in the number of individuals with similar labels, i.e. the compartment counts throughout the epidemic.

The SEIR model can be visualised using the following diagram from [Anderson et al., 2020]:

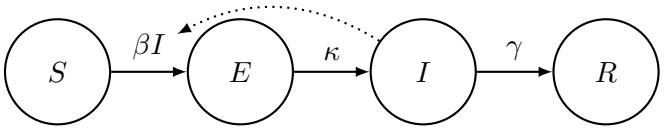


Figure 1: SEIR Model Structure

The changes in compartment counts are described by a system of ordinary-differential equations (ODEs) as follows, where $S(t)$, $E(t)$, $I(t)$ and $R(t)$ are the number of individuals in the S, E, I and R compartment at time t and β, κ and γ are the rates of transmission, incubation and recovery [Anderson and May, 1991; Anderson et al., 2020; Kermack et al., 1927; Keeling et al., 2021]:

$$\begin{aligned} \frac{dS}{dt} &= -\beta SI \\ \frac{dE}{dt} &= \beta SI - \kappa E \\ \frac{dI}{dt} &= \kappa E - \gamma I \\ \frac{dR}{dt} &= -\gamma I \end{aligned} \tag{2}$$

The influx into the exposed compartment βSI depends linearly on both the current number of susceptibles and that of infectious individuals. This is because in order to infect a susceptible individual, they must come in contact with an already infectious individual; that is

if we do not have any individual in either of the S or I compartments, no infection event can take place. The dotted arrow in Figure 1 illustrates the dependency on the number of infectious individuals of the transmission rate between the S and E compartments.

One of the main assumptions of the SEIR model is that the population is well mixed, i.e. any susceptible can come into contact with any infectious individual. For large spatially inhomogeneous populations this assumption will clearly not hold [Keeling et al., 2021]. Moreover, the model assumes that no one is dying either due to the epidemic or of natural causes, there are no births replenishing the susceptibles, and no movement of individuals from the recovered compartment back into S as a consequence of immunity loss. Therefore, longer-term predictions using this particular type of model will not be reliable.

0.3 Data

Some formalisms of the models included in our 'epimodels' package require that the population is not only split in terms of their infection status (susceptible, exposed, infectious, etc.) but also in terms of their ages and the region in which they are located. This requirement means that the datasets that we use in our parameter inference have to account for this finer scaling of the population. Throughout chapter 4, we will be studying the UK COVID-19 epidemic; hence, the origin of the collected data will reflect that.

0.3.1 Death Data

In the Aims section of the introduction chapter, we indicated that for the number of deaths and the number of cases we use data collected by the Office of National Statistics (ONS). These data are available on the UK's government's website [GOV.UK] in a '.csv' format. The daily number of deaths are broken down into nine PHE regions: East of England, North East, North West, Yorkshire and the Humber, South East, South West, East Midlands, West Midlands, and London. For each of these regions, we have separate counts for the number of deaths and cases for each of the 19 5-year-wide age groups: $0-4$, $5-9$ and so on with all instances over 90 in their own category; as well as aggregated counts for all those below 60 and those over 60 years old.

0.3.2 Serology Data and Large-scale Infection Surveys

Due to under-reporting and delays in recording of realised infections, case number datasets often present an unreliable image of the true fraction of the population that is infected at any given time. In practice we often use instead the results of antigen and antibody tests to draw an alternative image of the current infection landscape – a positive antigen test result indicates that the test subject is currently infected with the pathogen, while a positive antibody-test result suggests that the subject has been recently infected. These data are what is collectively known as serological (based on antibody-testing) and infection survey (based on antigen-testing) data. Using the proportion of positive tests results in any given test batch, we are able to arrive at the proportion of the population currently and previously infected by the virus in any given geographical area, which can then be extrapolated to give a good proxy for the true incidence of infection in that region on any given day.

For the serological and infection survey data required for the parameter inference in our analyses, we consult the antigen and antibody positive test results count from the REACT1 and REACT2 [Nicholson et al., 2021] studies. The REACT (REal-time Assessment of Community Transmission) studies rely on self-testing of participants and they aim to unravel further information about the transmission mechanisms of COVID-19. These are blind testing; subjects are chosen at random in order to increase the result's accuracy – the percentage of the tests that come back positive in the study is indicative of the percentage in the overall population, if the chosen sample is representative of the studied population. In the REACT1 study, antigen testing was rolled out to assess the prevalence levels of the SARS-CoV-2 virus, i.e. what the percentage of the population was infected at any given time. In the REACT2 study, antibody testing was used to find out the additional proportion of the population that has previously been infected [Nicholson et al., 2021].

We aggregate only the results of 19 rounds of the REACT1 study as these are the only ones for which we found a regional and age-classification. For each round, two distinct sets of data are available [Eales, 2020]: one from the government (the 'go' extension), the other from the NHS (the 'nhs' extension). The results were generated weekly for each of the rounds of the REACT1 study. The age groups differ from the ONS death data: $5-12$, $13-17$, $18-24$, $25-34$, $35-44$, $45-54$, $55-64$, and $65+$. These counts are then transformed into predictions of the numbers of infections in the population at any given

time-point by multiplying the percentage of positive tests recorded in each round of the study with the total number of individuals present in that population group – for which we use 2019 UK census data [ONS].

0.3.3 Susceptibles Data – Counts, Fatality Rates & Mobility

The total number of individuals of different ages in each of the PHE regions of England represent the initial number of susceptibles in each of these categories. We extract these data from the 2019 census data [ONS]. Mobility and contact data are extracted from Google mobility datasets [Google] and age-structured POLYMOD matrices [Prem et al., 2021] with 5-year intervals e.g. $0-5$, $5-10$, etc. Additional information sources such as the infection fatality ratio (IFR) and the case fatality ratio (CFR) which indicate the probability of an individual and, respectively, of an acknowledged case to die due to COVID-19 are referenced from previous studies of the virus in a contained environment e.g. Verity's paper [Verity et al., 2020] bases these percentages on individual-case data collected mainly from the Hubei region of mainland China. Each of the previously mentioned age-structured datasets has a different age-distribution breakdown. We provide in 'epimodels' Python scripts that process the data into a standardised age-structure: $0-1$, $1-5$, $5-15$, $15-25$, $25-45$, $45-65$, $65-75$, and $75+$. As a general principle, we average over the data fields that are clustered into the same new category, or linearly divide the data field value up if the new age group's edge is within the old age-interval.

0.3.4 Interventions Data

Some of the models account for the changing severity of the non-pharmaceutical interventions put in place by the government using a system of stringency levels developed by the Oxford COVID-19 Government Response Tracker group (OxCGRT). We use the provided '.csv' datafiles on the levels of these NPIs which are provided as additional resources in [Hale et al., 2021].

0.4 Bayesian Inference

With renewal models we use trajectories of the time-dependent instantaneous reproduction number R_t and generation time distribution w to produce timelines of the number of realised

new infections; however, in most cases, we have no prior knowledge of the trajectory of the reproduction number, but rather we want to use the available incidence data to estimate the underlying values of R_t that have generated the data in the first place. In contrast, a large number of parameters are required to run each of these three compartmental ODE models we analyse as part of the project: the PHE [Birrell et al., 2021], the Roche [Lemenuel-Diot et al., 2020] and the Warwick-Household [Keeling et al., 2021] models, which are all formalised as extensions to the basic SEIR structure. However, not all parameter values are known *a priori*, but rather need to be deduced from the noisy case or death data we have available. This process is what is known as *parameter inference* and we choose to use a Bayesian approach to retrieve the R_t values that best fit the infection data in the case of the renewal equation models.

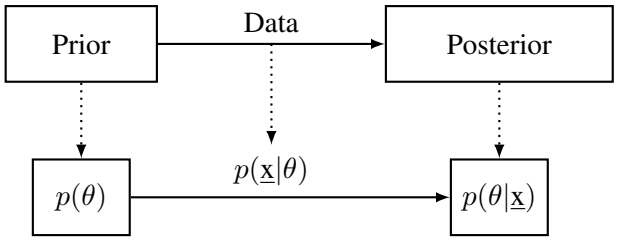
0.4.1 General Principles of Bayesian Inference

Before any data are introduced, we summarise our beliefs about the parameter values in a probability measure known as the *prior distribution*. Great care must be taken when choosing a prior: in some cases, we want to choose a non-informative prior that will not significantly affect the shape of the posterior; in others, the distribution of the data motivates the use of a conjugate prior for ease of computation. The prior can be used to regularise the results of inference, or encode our favoured behaviour when evidence is weak. For example, when seeking to infer the instantaneous reproduction number R_t, we normally choose a prior with a mean significantly greater than 1 in order to avoid learning that $R_t < 1$, unless there is evidence from the data that the epidemic is under control.

As noisy data are collected and analysed, these prior beliefs suffer incremental changes; this generates a new probability measure function called the *posterior distribution*, which specifies our beliefs about the parameter values conditional on the observed data. The effects of the data on the posterior distribution are contained in the *likelihood function*. The likelihood is usually treated as a function of the parameters of the chosen model [Lambert, 2018; Berger and Casella, 2001] and measures how probable it is to observe the data under the current parameter values.

The aim of parameter inference is to find the best choice of parameters, e.g., the parameter values for which the observed data are most likely (maximum likelihood estimation), or the parameter values which maximise the posterior distribution (maximum *a posteriori*

estimation). However, the data cannot inform the precise parameter values by itself due to noise. Therefore, we need to incorporate uncertainty information for reliable parameter estimates; this is provided by the posterior distribution. Bayes' rule ties everything together; the following schematic shows how we generate the posterior distribution:



where θ and $\underline{x}$ are the parameters of the model and the observed data respectively. $p(\theta|\underline{x})$ is the posterior, $p(\theta)$ is the prior and $p(\underline{x}|\theta)$ is the likelihood. From Bayes' rule we have

$$p(\theta|\underline{x}) = \frac{p(\theta) \times p(\underline{x}|\theta)}{p(\underline{x})} \tag{3}$$

with $p(\underline{x})$ being a function of the observed data only, independent of the choice of parameters θ.

0.4.2 Computational Challenges

Sometimes, we can deduce the shape of the posterior distribution $p(\theta|\underline{x})$ (normal, gamma, etc.), and express its parameters as closed form expressions of the data, e.g. when we use conjugate priors for the likelihood function. Knowing the exact shape of the posterior leads to easy computation of its moments and quantiles e.g. mean, variance, centred 90% confidence intervals – these can be calculated using known formulae instead of integrating $p(\theta|\underline{x})$. However, for many problems this is impossible, as these integrals do not have analytical solutions and calculating them via numerical integration is often prohibitively expensive. In these cases we require algorithms to search for the maximum point in the parameter space in a systematic way: optimisation and sampling using Markov Chain Monte Carlo routines are the two mechanisms that we will utilise.

Bayesian optimisation can be used to evaluate the maximum posterior distribution in a sequential way. On the other hand, Markov chain Monte Carlo (MCMC) sampling algorithms explore the parameter space through chains of model evaluations and computing summary statistics on the sampled values. Each algorithm iteration depends on the previous one, gen-

erating correlation between consecutive samples. This correlation results in an informational loss per algorithm run compared to an independent sampling algorithm; therefore, we need more dependent samples to characterise the posterior distribution to a similar degree [Lambert, 2018]. We run multiple Markov chains in parallel to properly assess the convergence of the algorithm to the posterior distribution – this is reached when the chains are well-mixed, i.e. the initial position of the chain cannot be traced back from the histogram of the samples. The convergence of the chain is assessed using Gelman's and Rubin's [Gelman and Rubin, 1992] $\hat{R}$ statistic, defined as

$$\hat{R} = \sqrt{\frac{W + \frac{1}{n}(B - W)}{W}},$$

where n is the total number of samples in the chain, and W and B are the within-path variation and the between-path variation of the samples, respectively. As we run our Markov chains for infinitely many samples and the empirical probability distributions described by the individual chains approach the true underlying posterior (that is, they approach convergence), the between-chain variation B becomes indistinguishable from the within-chain variation W (and hence equal) – which is gives us $\hat{R} = 1$; thus in practice we require a value $\hat{R}$ close to 1 in order to conclude that our algorithm has well-mixing chains which converged to the posterior distribution; otherwise, we need to run the chains for longer.

Since Gelman's and Rubin's 1992 paper [Gelman and Rubin, 1992], more accurate ways of calculating the mixing chains factor have emerged – the inference software which we use in this thesis adheres to this newer definition of $\hat{R}$, with the recommended threshold for good convergence having decreased from $\hat{R} \lesssim 1.1$ to now $\hat{R} \lesssim 1.01$ [Vehtari et al., 2021]. By generating samples from the posterior, and then approximating the moments and quantiles that we require from these samples, MCMC methods stand to be much more efficient than numerical integration and in particular the algorithm scales much more favourably with the number of parameters.

From eq. (3) we observe that computing the posterior is equivalent to computing the value of the product of the prior and the likelihood up to a constant (*posterior* $\sim$ *prior* $\times$ *likelihood*), which is what we use in practice. Our goal has thus shifted from maximising the posterior to maximising the value of the prior and likelihood product. In practice, we work with large

numbers of parameters, which makes the evaluation value of each data point very small. This can also happen when we deal with large numbers of data points. To avoid numeric overflow, we apply the log-function to the maximising product, i.e. we actually search for the maximum of the sum of the log-prior and the log-likelihood functions.

0.5 Software tools for inference and integration with PINTS

PINTS (Probabilistic Inference on Noisy Time-Series) is a framework for optimisation and Bayesian inference on models of noisy time-series [Clerx et al., 2019] and is the main pre-requisite package when installing and using the 'branchpro' and 'epimodels' software packages – which we use to run all the analyses described in this thesis. In the following we describe which of the functionalities on the PINTS package we employ and how they are integrated in the structure of 'branchpro' and 'epimodels'.

0.5.1 Branchpro

For all forward-simulations and inference procedures undertaken in the first three chapters of the thesis, we extend the previously-published renewal equation model framework of 'branchpro' [Creswell et al., 2022] – which was developed during the first year of my DPhil studies as part of a year-long collaborative group software project[1], to allow multiple-group population modelling, as well as to allow for other noise distributions apart from the typical Poisson-distributed noise (e.g. negative binomial). By extending the types of models included in the 'branchpro' suite, we ensure we have the tools to hand to study the effect of population heterogeneity on the overall epidemic trend.

Using the PINTS package, we create custom likelihood and posterior classes for each of the models incorporated in the 'branchpro' package. PINTS (Probabilistic Inference on Noisy Time-Series) is a framework for optimisation and Bayesian inference on models of noisy time-series [Clerx et al., 2019] and is the main prerequisite package when installing and using 'branchpro' and 'epimodels'. It contains a large variety of functionalities which enable us to do parameter inference in a cohesive way: we can create personalised classes for

[1] As part of the 'branchpro' software project, I was heavily involved in the conceptualisation and investigation process including, but not limited to, writing of the main code, creation and running of analyses and writing of the original paper summarising our findings. CRediT author contributions for [Creswell et al., 2022] for IB: formal analysis, investigation, software, validation, visualization and writing—review and editing.

our models' log-prior, log-likelihood, log-posterior distributions, as well as running optimisation and MCMC sampling routines. It also includes methods to analyse the quality of the output of these routines such as plotting of the chain outputs, $\hat{R}$ values and so on. For our inference approach we use the implementation of the No U Turns Sampling (NUTS) algorithm included in PINTS. The NUTS algorithm proposed by [Hoffman and Gelman, 2011] is a recursive sampler which computes at each step the derivative of the likelihood function at the current position, and it uses this potential to propose new points to explore in the hyper-parameter space, and thus exploits the geometry of the posterior surface to its advantage.

In addition to our in-house classes for running the inference using PINTS, we also provide simple alternative inference frameworks for each of the models, using the 'stan' [Lee et al., 2017] package, based on the same NUTS inference algorithm. Through example notebooks that showcase the functionalities of the 'branchpro' package, we demonstrate the correctness of the inference algorithm implemented, as we are able to produce the same inferred R_t trajectories using either of the PINTS-based or Stan-based inference pipelines.

0.5.2 Epimodels

In addition to 'branchpro', we have also independently developed the 'epimodels' package, using the same strict software engineering standards as with 'branchpro' – which we discuss in more depth in section 4.2.2.2. This software package is organised as a library of epidemiological models with highly modularised model setups which can then be used as a template for users to easily add new compartmental models for disease transmission to the suite of models, run parameter inference analyses, and predict trajectories of epidemic outcomes. In 'epimodels' we use ready-made optimisers and samplers from PINTS. For our purposes, we use the *CMA-ES* optimiser [Hansen et al., 2003; Hansen, 2006] and the *Haario-Bardenet Adaptive Covariance Monte Carlo (ACMC)* sampler [Haario et al., 2001; Johnstone et al., 2016] routines.

CMA-ES uses covariance matrix adaptation of a multi-variate normal search distribution [Hansen et al., 2003; Hansen, 2006] and is designed for non-linear, non-convex function optimisation. At each step in the algorithm, we evaluate the covariance matrix C using a single population i.e. parameter instance of one generation. The main advantage of CMA-ES is that it does not require gradient evaluation at each step, making it a fast and inexpensive

optimisation algorithm [Hansen et al., 2003].

Haario-Bardenet ACMC on the other hand is a type of adaptive Metropolis in which the proposal distribution is adapted continuously using previous knowledge of the chain states. The proposal distribution function at time t is a Gaussian centred at the current chain location X_{t-1} and a covariance matrix given by the covariance of all the previous and current chain states X_0, X_1, ..., X_{t-1} [Haario et al., 2001].

In the next chapter, starting from first principles, we will derive analytically an expression for the overall reproduction number for both the one-group and multiple-group population in the contexts of a renewal equation model, and demonstrated that the identified definition of the time-dependent R_t of the whole population provides a necessary and sufficient criterion for summarising the long-term epidemic dynamics in the context of the deterministic renewal equation model.

1 | Stability criterion for multiple population groups renewal processes

Here, we present a theoretical derivation of a general criterion for the long-term fate of an epidemic modelled using a renewal equation process with one or more population categories, with examples that showcase its application both in deterministic and stochastic renewal models. The following forms Chapter 1 of the thesis and has been published in the *Journal of the Royal Society Interface*. This paper is also available as a preprint [Bouros et al., 2025]. CRediT author contributions for this work are as follows. Conceptualisation: All authors. Methodology: IB, BL. Investigation: IB. Writing – original draft: IB, BL. Writing – review and editing: All authors. Supervision: DG, BL, RT.

Summary

The time-dependent reproduction number R_t can be used to track pathogen transmission and to assess the efficacy of interventions. This quantity can be estimated by fitting renewal equation models to time series of infectious disease case counts. These models almost invariably assume a homogeneous population. Individuals are assumed not to differ systematically in the rates at which they come into contact with others. It is also assumed that the typical time that elapses between one case and those it causes (known as the generation time distribution) does not differ across groups. But contact patterns are known to widely differ by age and according to other demographic groupings, and infection risk and transmission rates have been shown to vary across groups for a range of directly transmitted diseases. Here, we derive from first principles a renewal equation framework which accounts for these differences in transmission across groups. We use a generalisation of the classic M'Kendrick–von Foerster equation to handle populations structured into interacting groups. This system of partial differential equations allows us to derive a simple analytical expression for R_t which involves only group-level contact patterns and infection risks. We show that the same expression emerges from both deterministic and stochastic discrete-time versions of the model and demonstrate via simulations that our R_t expression governs the long-run fate of epidemics. Our renewal equation model provides a basis from which to ac-

count for more realistic, diverse populations in epidemiological models and opens the door to inferential approaches which use known group characteristics to estimate R_t.

1.1 Introduction

In the long run, an infectious disease epidemic grows if, on average, the number of new infections generated by each infectious individual exceeds one; contrastingly, it will subside if this value is below one. The (instantaneous) time-dependent reproduction number, R_t, is rigorously defined as the average number of secondary cases generated by an infected case at time t assuming that transmission conditions remain the same in the future. R_t is an emergent property of a pathogen spreading through a specific population, and it depends on the biology of the pathogen and the characteristics of the population, including any measures taken to limit its spread [Flaxman et al., 2020; Brauner et al., 2021]. Accordingly, determining R_t is crucial for public health policymaking during epidemics.

Renewal equation models have, over the past two decades, become the predominant models used when R_t is estimated from epidemiological time series, most commonly using time series of infectious disease case counts [Fraser, 2007; Cori et al., 2013; Thompson et al., 2019]. Their success is due to their relative simplicity: these models assume that new infections are caused by previous infections with time lags between parent and daughter infections; these time lags are assumed to be characterised by a *generation time distribution* allowing for variation in this quantity across infector-infectee pairs [Svensson, 2007]. In these models, the number of new infections caused by a typical new case at time t is given by R_t. Usually, these models are fitted to infectious disease case counts (rather than infection counts, which are harder to observe) and, in so doing, R_t is estimated. There is a large literature and range of software tools devoted to R_t inference via renewal equations [Fraser, 2007; Cori et al., 2013; Thompson et al., 2019; Gostic et al., 2020; Abbott et al., 2020; Creswell et al., 2023].

These inferential frameworks are built upon an implicit assumption of uniformity in the population, i.e. that disease transmission rates are the same across the population. This ignores what is widely known – that disease transmission differs systematically across groups, and is explicitly modelled in other types of models of disease transmission dynamics. For example, compartmental models (e.g. [Como-DTC-Collaboration, 2021]) are frequently structured by age, largely because, in many populations, the numbers and types of contact that individ-

uals typically have with others depend strongly on their age. This is so widely recognised that there is a literature devoted to estimating so-called *contact matrices* that capture this information (e.g. [Prem et al., 2021]). Infection risk can also differ according to other demographics [Pijls et al., 2021; Theodore et al., 2023]. The time period between one case and those it causes can also vary according to demographics: pathogens may undergo distinct dynamics within individuals from specific groups and be spread differently, and members of different groups may also adapt their behaviour to varying degrees if symptoms show. These differences can result in large variations between generation times across groups (e.g. for COVID-19, [Kim et al., 2022]).

This shortcoming of the standard renewal framework has been recognised in previous work. It is possible to estimate a group-specific R_t from infection time series by making assumptions about the rate at which pathogens spread within and between groups [Glass et al., 2011], although differences in generation time distributions according to group have not been considered; in such frameworks, an overall reproduction number can be determined through a weighted average of within- and between-group reproduction numbers. An alternative approach is to model the population using infection-age-structured partial differential equation (PDE) systems, with one PDE for each of the groups [Green et al., 2022]. This approach follows the classic M'Kendrick–von Foerster equation framework which is central in the study of population dynamics in areas such as demography [M'kendrick, 1925; von Foerster, 1959; Murray, 2002], where birth and death processes vary according to age, and we follow this approach here.

We suppose there are groups distinct in terms of their contact behaviour, infection risks and generation time distributions. Like [Green et al., 2022], we determine an analytical expression for R_t, but ours is far simpler and does not involve the generation time distribution. We also provide a rigorous derivation starting from an age-structured system of coupled partial differential equations which makes clear the assumptions underpinning the renewal equation dynamics for a structured population, and crucially this indicates special scenarios when the dynamics deviate from those shown in [Green et al., 2022]. Using a discrete-time version of the model (as most inferential routines for R_t use discrete-time renewal models), we arrive at the same expression for R_t. We use simulation to demonstrate that our expression for R_t behaves as expected: $R_t = 1$ defines the boundary between long-run epidemic growth (if $R_t > 1$) and long-run decline (if $R_t < 1$). We show that this R_t expression holds also

for a stochastic version of our model, where, on average, the behaviour of these systems is delineated by the $R_t = 1$ boundary. We also derive a relationship between R_t and the calendar time growth rate of the epidemic, r_t, for a population organised into groups, and we show that r_t is particularly sensitive to changes in the generation time of groups with the most contacts. Our framework naturally allows known characteristics of different groups to be incorporated into renewal equation models of epidemics and into associated estimates of epidemiological quantities.

1.2 Methods

1.2.1 Renewal equations used to infer R_t

Renewal models are primarily used to estimate R_t, and we now describe the most basic, yet indicative, form of these models. This model is stochastic and discrete-time (typically with time-steps of one day) and assumes that the population is homogeneous and takes the form:

$$I_t \sim \text{Poisson}(R_t \Lambda_t), \text{ where } \Lambda_t = \sum_{a=1}^{t-1} w_a I_{t-a}. \tag{1.1}$$

In this expression, $I_t \geq 0$ represents counts of infections arising at time t; $\Lambda_t \geq 0$ is known as the *transmission potential* at time t, which is a weighted sum of past infections where the weights are determined by a *generation time distribution*. This distribution is a discrete probability distribution, $\{w_a\}_{a=1}^{\infty}$, such that $\sum_{a=1}^{\infty} w_a = 1$, where $w_a \geq 0$ gives the probability that the time elapsing between a past infection and a daughter case is a days.

1.2.2 Derivation of the renewal equation for the M'Kendrick–von Foerster model

Eq. (1.1) is a stochastic renewal equation that can be motivated by the classic M'Kendrick–von Foerster model [M'kendrick, 1925; von Foerster, 1959], which uses an age-structured partial differential equation to model population dynamics where there are birth and death processes that depend on age (e.g. [Murray, 2002, Chapter 1.7]). This model is deterministic and, typically when used in the context of epidemics, models the numbers of those infected continuously and is continuous in time and age of infection. We walk through a derivation

of the renewal equation for this classic model since our approach for modelling structured populations extends it. In section 1.2.3, we make explicit our assumptions in using this framework to model epidemic dynamics for a homogeneous population.

Following closely the derivation and notation used in [Murray, 2002, Chapter 1.7], we denote the *density* of infections still present at time t which began a days ago by $n(t, a)$; the total *number* of infections is then given by $\int_0^\infty n(t, a)da$. Note that, in our version of this classic model, a denotes the infection age and *not* ages of infected individuals. In a small time increment dt, the conservation law for the population dictates that [Murray, 2002, Chapter 1.7]:

$$dn(t, a) = \frac{\partial n}{\partial t}dt + \frac{\partial n}{\partial a}da = \mu(a)dt, \tag{1.2}$$

where $\mu(a)$ is the rate at which infections end. In our work, we do not keep track of the number of recovered or dead cases, essentially assuming that $\mu(a) = 0$. Since infection age changes at the same rate as calendar time, $dt/da = 1$, and eq. (1.2) can then be simplified to the following partial differential equation (PDE):

$$\frac{\partial n}{\partial t} + \frac{\partial n}{\partial a} = 0. \tag{1.3}$$

We specify a boundary condition which dictates the rate at which new infections arise:

$$n(t, 0) = \int_0^\infty b(t, a)n(t, a)da, \tag{1.4}$$

where $b(t, a) \geq 0, \forall t, a$ is the rate at which infections arising a days ago generate new infections. In addition, we also assume that the birth rate of new infections arising from very old infections is effectively zero, that is for large positive values of a, $b(t, a) = 0, \forall t$. In what follows, we explicitly model only the population dynamics for $t \geq 0$, and, in order to close the system, we assume that at $t = 0$, there is a density of infections given by

$$n(0, a) = f(a). \tag{1.5}$$

We can consider two solution classes corresponding to distinct groups of infected individuals

at time t, each of which has a different relationship between a and t:

$$a = \begin{cases} t + a_0, \text{ for } a > t, \text{ i.e. for infections arising before } t = 0, \\ t - t_0, \text{ for } a \leq t, \text{ i.e. for infections arising from } t = 0 \text{ onwards}, \end{cases} \tag{1.6}$$

where $a_0 > 0$ is the age of an infection at time $t = 0$ for the first group, and t_0 is the time at which an infection arises for the second (where $0 \leq t_0 \leq t$). For each of the cases in eqs. (1.6), we can write $n(t, a) = n(t, a(t)) := n(t)$, i.e. a function of t only, which we can substitute into eq. (1.3) to yield an ordinary differential equation (ODE):

$$\frac{dn}{dt} = 0, \tag{1.7}$$

which is valid only along one of the so-called *characteristic* lines defined in eqs. (1.6). In this context, the characteristics correspond to infections arising at a particular point in time, τ: if $a > t$, a unique value of a_0 gives $\tau < 0$; if $a \leq t$, $\tau = t_0$. For each infection onset time then eq. (1.7) has the solution:

$$n(t, a(t)) = \text{constant}, \tag{1.8}$$

meaning that the number of individuals infected at a given time t remains forever constant: this makes sense because, in our framework, we do not allow for cessation of an infection.

We now derive these constant population sizes for each of the cases in eq. (1.6) by considering these at conveniently chosen moments. When $a > t$, we can write $a = t + a_0$ meaning $n(t, t + a_0) = \text{constant}$, and when $t = 0$, the population size is:

$$n(0, a_0) = f(a_0) = f(a - t). \tag{1.9}$$

When $a \leq t$, we can write $a = t - t_0$ and $n(t, t - t_0) = \text{constant}$, and, when $t = t_0$:

$$n(t_0, 0) = n(t - a, 0). \tag{1.10}$$

Collecting eqs. (1.9) & (1.10), we have the solutions for the population sizes for all infection

onset times:

$$n(t,a) = \begin{cases} f(a-t), a > t, \\ n(t-a,0), a \leq t. \end{cases} \tag{1.11}$$

Substituting eqs. (1.11) into eq. (1.4) results in:

$$n(t,0) = \underbrace{\int_0^t b(t,a)n(t-a,0)da}_{a \leq t} + \underbrace{\int_t^\infty b(t,a)f(a-t)da}_{a>t}. \tag{1.12}$$

To avoid having to specify the initial conditions at $t = 0$ (and therefore specifying the values of $b(0,a)$ for all $a > 0$ when computing $n(0,0)$), we generally assume that t is large such that $b(t,a') \approx 0, \forall a' > t$. This assumption does not impact the early outbreak dynamics, as it simply determines that all active infections prior to the start time of the epidemic are 'reclocked' to have occurred at time zero. With this eq. (1.12) reduces to:

$$n(t,0) = \int_0^t b(t,a)n(t-a,0)da. \tag{1.13}$$

Eq. (1.13) gives a recurrence relation for our system solution, which is an implicit equation for $n(t,0)$. In section 1.2.3, we show how this can be explicitly solved.

We now define $I(t) := n(t,0)$ to represent a measure of new infections arising at time t, and eq. (1.13) then becomes:

$$I(t) = \int_0^t b(t,a)I(t-a)da. \tag{1.14}$$

Eq. (1.14) is a continuous and deterministic renewal equation.

1.2.3 The M'Kendrick–von Foerster model of an epidemic in a homogeneous population

We now make explicit our assumptions around how infections arise in a homogeneous population and adapt eq. (1.14) accordingly. We then derive long-term solutions for this model. We suppose that there are two distinct time-dependent processes that modulate the birth rate of new infections: (i) the total number of individuals an infected person comes into contact with over the course of their infectious period if their infection started at time t, $C(t) \geq 0$; and, (ii), the probability, $0 \leq \gamma(t) \leq 1$, that each contact results in an infection; this im-

plicitly assumes that individuals do not change their behaviour as their infection progresses; for example, it assumes that when symptoms appear infected individuals do not change how they socialise. We assume that there is also an infection-age-dependent process which governs how much time typically elapses between one infection starting and those it causes: we use a continuous probability distribution for this, with probability density function, $w(a)$. The birth rate of new infections is then given by: $b(t,a) = C(t)\gamma(t)w(a)$, and we can substitute this into eq. (1.4) to produce a renewal equation for new infections:

$$I(t) = \int_0^t C(t)\gamma(t)w(a)I(t-a)da. \tag{1.15}$$

The rate at which new infections are generated from past infections is, in this model, not explicitly contingent on the numbers of those infected in the past. This differs from, say, an SIR compartmental model, where the rate at which new infections are generated from past infections depends on the availability of susceptible individuals which diminishes as the epidemic spreads through a population. In addition, we do not consider the termination of an infection; that is the death rate of new infections is assumed to be $\mu(a) = 0$.

If we consider time-constant $C = C(t)$ and $\gamma = \gamma(t)$, the function $I(t)$ is a convolution of past values of itself with a time-invariant kernel $w(a)$. The only function which has this property is an exponential; this means that $I(t)$ can only grow or decline exponentially over time. An alternative way to see that only unregulated growth or decline is allowed is because eq. (1.15) lacks self-regulation, for example, due to the removal of susceptible individuals. This means that the system has no non-zero infection equilibrium.

We then look for a similarity solution [Murray, 2002, Chapter 1.7] of our system (for large t) of the form:

$$n(t,a) = \exp(rt)p(a), \tag{1.16}$$

where r is the growth rate of the epidemic (which can be negative if there is epidemic decline), and $p(a)$ is an age-dependent distribution. Substituting eq. (1.16) into the $a \leq t$ solution in eq. (1.11) which has $n(t,a) = n(t-a,0)$ gives:

$$p(a) = p(0)\exp(-ra), \tag{1.17}$$

which we substitute back into eq. (1.16) to give us the solution:

$$n(t, a) = p(0) \exp(r(t - a)). \tag{1.18}$$

Consequently, the number of new infections grows (or declines) exponentially over time:

$$I(t) = p(0) \exp(rt), \tag{1.19}$$

as desired for our intended solution. Substituting this into eq. (1.15) and dividing both sides by common terms gives:

$$1 = \int_0^t C(t)\gamma(t)w(a)e^{-ra}da, \tag{1.20}$$

which is known as the Euler-Lotka equation and was first derived by Lotka at the start of the 20th century; this extended a more specific result obtained by Euler in the 1700s [Bacaër and Ouifki, 2007].

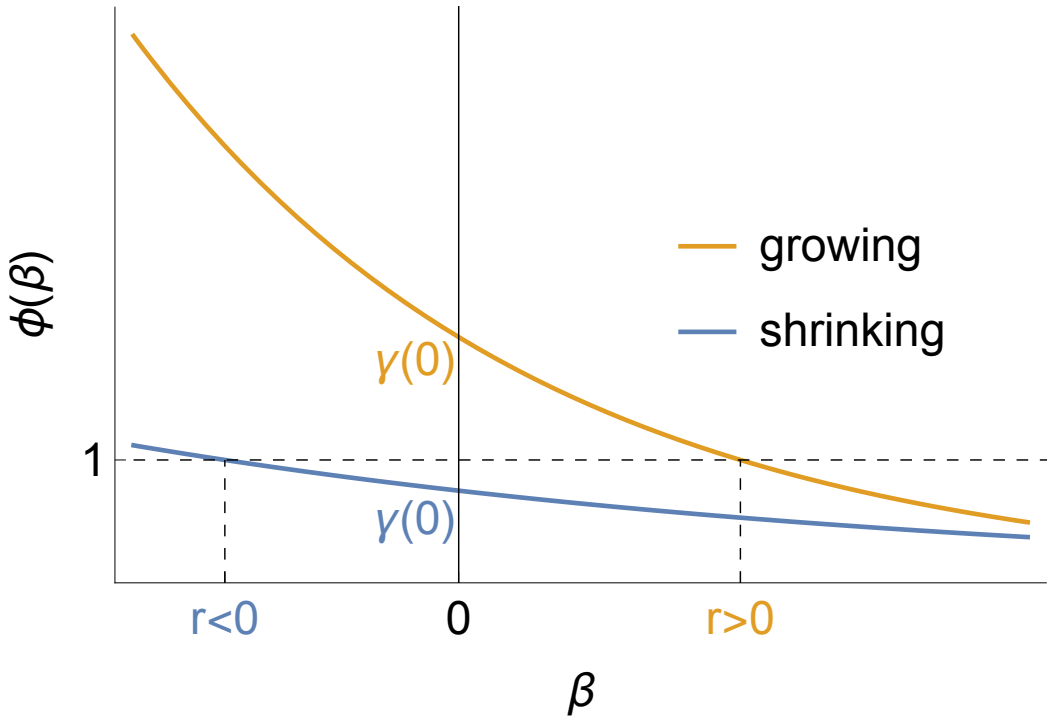


Figure 1.1: *How the growth rate of an epidemic is determined by* $\phi(0)$. *Here,* $\phi(\beta) := \int_0^t C(t)\gamma(t)w(a)e^{-\beta a}da$. *This figure is a reproduction of a figure from [Murray, 2002, Chapter 1.7].*

Whether the infected population grows or declines exponentially and the rate at which it

does so is determined by the value of r which solves eq. (1.20). There are no simple closed-form solutions to eq. (1.20); instead, we show how the sign of r is determined by particular terms in this equation.

We define the function $\phi(\beta) := \int_0^t C(t)\gamma(t)w(a)e^{-\beta a}da$; we note that, using this notation, $r := \{\beta, \text{ s.t. } \phi(\beta) = 1\}$. Crucially, this function is monotonically decreasing in β. This means that if $\phi(0) > 1$, then $r > 0$; if $\phi(0) = 1$, $r = 0$; and if $\phi(0) < 1$, $r < 0$ (see Figure 1.1).

The sign of r is so determined by the value of $\phi(0)$:

$$\phi(0) = \int_0^t C(t)\gamma(t)w(a)da, \tag{1.21}$$

$$= C(t)\gamma(t)\int_0^t w(a)da, \tag{1.22}$$

$$\approx C(t)\gamma(t), \text{ for large } t. \tag{1.23}$$

Specifically, the population grows if $C(t)\gamma(t) > 1$, is maintained if $C(t)\gamma(t) = 1$ and shrinks if $C(t)\gamma(t) < 1$. This behaviour is exactly the type of thresholding behaviour that is expected from a reproduction number, so we define:

$$R(t) := C(t)\gamma(t). \tag{1.24}$$

This time-varying reproduction number acts intuitively: $C(t)\gamma(t)$ is a measure of infections generated over the course of an infection (if transmission remains the same). If it exceeds 1, then each infection more than replaces itself; and analogously for the other two cases.

Formally, the reproduction number is defined as the largest (and here only) eigenvalue of the *next-generation operator* [Diekmann et al., 1990]; here, this is an integral operator, $\mathcal{K}$:

$$\mathcal{K}(g(a)) := \int_0^t C(t)\gamma(t)w(a)g(a)da. \tag{1.25}$$

If $g(a) = 1$, eq. (1.25) yields the expected number of infections generated by an infection, i.e.

$$\mathcal{K}(1) = C(t)\gamma(t), \tag{1.26}$$

meaning $R(t) = C(t)\gamma(t)$ as required. In practice, $R(t)$ is written as R_t.

1.2.4 Reproducing our simulation results

Throughout this Chapter, we use simulations of unfolding outbreaks to support our mathematical derivations. To allow others to rerun these simulations, we make our Python code available through a public GitHub repository available here: [SABS-R3-Epidemiology, 2025].

1.3 Results

In our results, we use a mix of mathematically derived logic complemented with simulations to back up our theory. Our simulations aim to demonstrate support for our derivations rather than closely matching real outbreak scenarios.

1.3.1 A renewal equation for new infections in a structured M'Kendrick-von Foerster model

We now extend the model presented in section 1.2 and assume that the population is structured into interacting groups. In this section, we derive the renewal equation for new infections in this model.

We assume that the groups within the population differ in two ways that affect disease transmission:

Differences in contact patterns. We take inspiration from compartmental ODE models where contact patterns within and between groups are incorporated through contact matrices (e.g. [Prem et al., 2021]).

Differences in within-host pathogen dynamics and behaviour during infection. Within-person pathogen dynamics and transmission patterns can differ across individuals. Individuals can also differ in terms of their behaviour throughout infection: when symptoms show, more risk-averse individuals may choose to avoid socialising to prevent infecting others. Each of these characteristics could differ systematically across, for example, demographic groups, leading to differences in the generation time across these groups.

We define $n^i(t, a)$ as the density of current infections in population group $i \in \{1, 2, ..., N\}$

present at time t that arose a time periods ago. We can stack these densities into a vector containing this information for all groups:

$$\underline{n}(t,a) = \begin{pmatrix} n^1(t,a) \\ \vdots \\ n^N(t,a) \end{pmatrix}.$$

As for the homogeneous population model, applying population conservation yields the following system of PDEs:

$$\frac{\partial \underline{n}}{\partial t} + \frac{\partial \underline{n}}{\partial a} = 0, \tag{1.27}$$

where $\partial \underline{n}/\partial x := [\partial n^1/\partial x, \partial n^2/\partial x, ..., \partial n^N/\partial x]'$ is the vector-partial derivative. We denote the birth rate of infections in group i which were caused by infections in group j by $b^{j\to i}(t,a)$. This means that the total density of new infections in group i at time t is given by those caused by infections across all groups:

$$n^i(t,0) = \sum_{j=1}^{N} \int_0^\infty b^{j\to i}(a) n^j(t,a) da. \tag{1.28}$$

As for the homogeneous model (section 1.2.3), we assume that the birth rate of new infections is the product of three terms:

$$b^{j\to i}(t,a) = \gamma(t) w^j(a) C^{(ji)}(t), \tag{1.29}$$

where $C^{(ji)}(t)$ now represents an element from a contact matrix and indicates the total number of contacts in group i made by an individual in group j over the course of their infection; $w^j(a)$ indicates a probability density function representing the generation time of group j. We make the additional assumption that the probability that each contact results in an infection $\gamma(t)$ is the same over all combinations of population groups. An alternative way to view this assumption is when setting $\gamma(t) = 1$ meaning $C^{(ji)}(t)$ represents the total number of infections caused by group i in those of category j throughout their infection – this implicitly assumes that the contact matrix accounts for variations in infectiousness and susceptibility across infector/infectee groups.

We can then use eqs. (1.28) & (1.29) to write an expression giving the rate of new infection generation across all groups:

$$\begin{aligned}
\underline{n}(t,0) &= \begin{pmatrix} \sum_{j=1}^{N} \int_0^\infty \gamma(t) w^j(a) C^{(1j)}(t) n^j(t,a) da \\ \vdots \\ \sum_{j=1}^{N} \int_0^\infty \gamma(t) w^j(a) C^{(Nj)}(t) n^j(t,a) da \end{pmatrix} \\
&= \int_0^\infty \gamma(t) \begin{pmatrix} \sum_{j=1}^{N} C^{(1j)}(t) w^j(a) n^j(t,a) \\ \vdots \\ \sum_{j=1}^{N} C^{(Nj)}(t) w^j(a) n^j(t,a) \end{pmatrix} da \\
&= \int_0^\infty \gamma(t) C(t) \begin{pmatrix} w^1(a) n^1(t,a) \\ \vdots \\ w^N(a) n^N(t,a) \end{pmatrix} da \\
&= \int_0^\infty \gamma(t) C(t) \underbrace{\begin{bmatrix} w^1(a) & & \\ & \ddots & \\ & & w^N(a) \end{bmatrix}}_{W(a)} \begin{pmatrix} n^1(t,a) \\ \vdots \\ n^N(t,a) \end{pmatrix} da \\
&= \int_0^\infty \gamma(t) C(t) W(a) \begin{pmatrix} n^1(t,a) \\ \vdots \\ n^N(t,a) \end{pmatrix} da. \qquad (1.30)
\end{aligned}$$

As $\int_0^\infty w^i(a) da = 1, \forall i$, the matrix $W(a)$ satisfies the following property:

$$\int_0^\infty W(a) da = \mathbb{I}_N. \qquad (1.31)$$

In order to close the system, we assume that at $t = 0$, there is a density of infections given by

$$\underline{n}(0,a) = \underline{f}(a). \qquad (1.32)$$

We use the same approach as for the homogeneous population model to arrive at the analogous form of eq. (1.12) for the structured population model:

$$\underline{n}(t,0) = \int_0^\infty \gamma(t) C(t) W(a) \underline{n}(t,a) da$$

$$= \int_0^t \gamma(t)C(t)W(a)\underline{n}(t-a,0)da + \int_t^\infty \gamma(t)C(t)W(a)\underline{f}(a-t)da. \quad (1.33)$$

As before, we assume that t is large such that infections with age exceeding this time contribute negligibly to onward transmission, and eq. (1.33) then reduces to:

$$\underline{n}(t,0) = \int_0^t \gamma(t)C(t)W(a)\underline{n}(t-a,0)da, \quad (1.34)$$

which can be written as a renewal equation for new infections:

$$\underline{I}(t) = \int_0^t \gamma(t)C(t)W(a)\underline{I}(t-a)da. \quad (1.35)$$

1.3.2 The time-varying reproduction number for structured populations

The multiple-group renewal equation, eq. (1.35), appears of the same form as the equivalent expression for a single group, eq. (1.14). It is natural then to suppose that the long-term solution for the multiple-group model should be of analogous form:

$$\underline{I}(t) = e^{rt}\underline{\Phi}, \quad (1.36)$$

where $\underline{\Phi}$ is a constant non-negative vector.

In Appendix A1, we outline the set of conditions for eq. (1.36) to hold. We then prove that these conditions hold in general for realistic population structures where groups interact; the only exceptions are populations in which not all population groups mix (as seen in eq. (A.1)) and diseases where the generation-time interval w has only one non-zero entry (that is subsequent infections occur exactly j days after the parent infection for $w(j) \neq 0$).

In the case of non-interacting populations, for example, it is not generally true that the growth rate of new infections should be identical across groups. For example, if groups have differing generation times and they do not interact, the infections will generally grow at different calendar rates in each group. This type of situation could arise if intervention policies were put in place which completely isolated parts of the population; our theory requires only that the contact rate between these parts of the population and others is non-

zero, however, and it is unlikely that such a policy would entirely isolate these groups.

We demonstrate that eq. (1.36) holds using simulations of a discrete deterministic renewal model; we set parameters of the model such that there is epidemic growth. In these simulations, we assume that the population is structured into three age groups: 0-20 year-olds, 20-65 year-olds and those aged $65+$. We use a contact matrix derived from [Prem et al., 2021] as proxy for the total contact matrix for this population:

$$C(t) = \begin{pmatrix} 6.33 & 1.61 & 0.1 \\ 4.65 & 7.41 & 0.51 \\ 0.49 & 1.17 & 1.13 \end{pmatrix}, \tag{1.37}$$

where the rows correspond to infectors aged 0-20 (first row), 20-65 (second row) and $65+$ (third row); the columns correspond to infectees in groups with the corresponding ordering as for the rows. The values in the matrix represent the average daily contacts an individual in the row age group has with those in the column age group. Those aged 0-20 have the vast majority of their contacts with members of the same group; those aged 20-65 have high contact rates with the two youngest age groups; those aged 65+ have comparably low contact rates across all age groups.

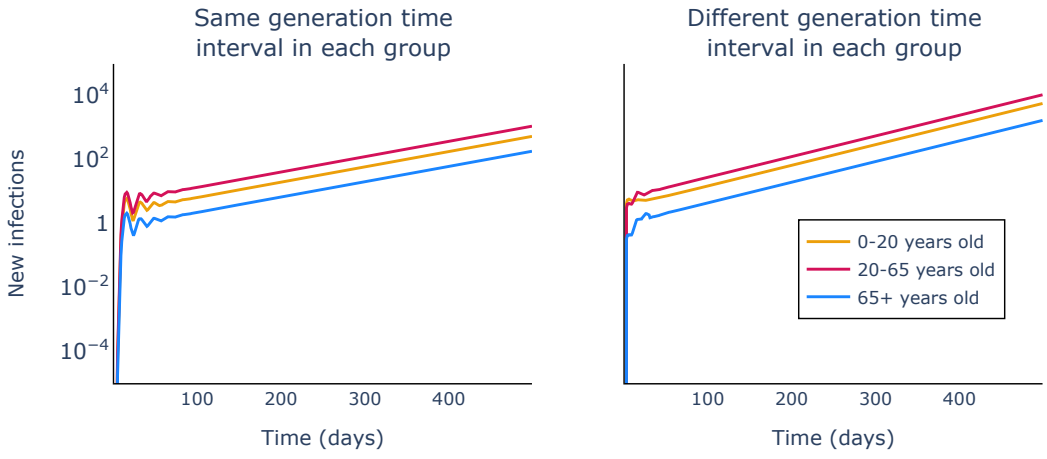


Figure 1.2: *Epidemic growth occurs at the same rate across interacting groups. For both panels, we use the same reproduction number $R(t) = 1.2$ and simulate using a discrete deterministic renewal equation. In the left-hand panel, the generation time distributions are given by scenario 2 in Table 1.1; in the right-hand panel, we assume the generation time distributions differ across groups and are given by scenario 1 of the same table. The contact matrix assumed is given in eq. (1.37).*

In Table 1.1, we summarise different generation interval scenarios that we will use in our following analyses. All generation intervals presented follow a Gamma distribution, with stated means and standard deviation. While the values we utilised in our analyses may not be taken from a clinical study, they mean to demonstrate theoretically the importance of a correct assessment of $R(t)$, when not all population groups follow identical generation times.

In Figure 1.2 we show the results of these simulations; the left-hand panel shows the evolution of the outbreaks in each group when the groups all have the same generation time distribution; the right-hand panel shows the equivalent when the generation time distributions differ across groups. In both cases, the rate of outbreak growth is the same across the three populations, as is predicted by eq. (1.36).

Because eq. (1.36) holds for realistic populations, we assume a solution of this form and substitute it into eq. (1.35). Simplifying both sides, the resulting equation reads as

$$\underline{\Phi} = \Big(\int_0^\infty \gamma(t)C(t)W(a)e^{-ra}da \Big)\underline{\Phi} := \overline{K}(r)\underline{\Phi}, \tag{1.38}$$

where $\overline{K}(\beta) := \Big(\int_0^\infty \gamma(t)C(t)W(a)e^{-\beta a}da \Big)$ is the Laplace integral transform of the positive matrix $\gamma(t)C_tW(a)$. Eq. (1.38) implies that the vector $\underline{\Phi}$ is an eigenvector of the positive matrix $\int_0^\infty \gamma(t)C(t)W(a)e^{-ra}da$, with associated eigenvalue 1.

We now use a linear algebra result called the *Perron-Frobenius* theorem [Bacaer et al., 2022] to convert eq. (1.38) into a form to which we can apply the same logic as for the single population model in section 1.2.3. This theorem states that for a positive matrix, in our case $\overline{K}(\beta)$, the following three results hold:

1. defining the eigenvalues of $\overline{K}(\beta)$ as $\nu_1, \nu_2, ..., \nu_N$, the spectral radius $\rho(\overline{K}(\beta)) := \max(\{|\nu_1|, |\nu_2|, ..., |\nu_N|\})$, is an eigenvalue of $\overline{K}(\beta)$;
2. $\rho(\overline{K}(\beta)) \in \mathbb{R}$ and $\rho(\overline{K}(\beta)) > 0$;
3. the eigenvector $\underline{v}$ of $\overline{K}(\beta)$ associated to the spectral radius $\rho(\overline{K}(\beta)$ is a positive vector and the one and only non-negative eigenvector of $\overline{K}(\beta)$ (all other non-negative eigenvectors are multiples of $\underline{v}$).

All values of the vector $\underline{I}(t)$ are non-negative meaning that $\underline{\Phi}$ must also be a non-negative

		Mean generation time			Standard deviation generation time		
Scenario	**Description**	0-20	20-65	65+	0-20	20-65	65+
1	mixture of long and short generation times, varied across groups	5.6	15.3	25	3.3	3.3	3.3
2	medium length generation times, equal across groups	15.3	15.3	15.3	3.3	3.3	3.3
3	mixture of long and short generation times, swapped and varied across groups	25	15.3	5.6	3.3	3.3	3.3
4	short generation times, varied across groups	3.3	5	7	3.3	3.3	3.3
5	short generation times, equal across groups	5.1	5.1	5.1	3.3	3.3	3.3
6	mixture of long and short generation times, varied across groups with different standard deviations	15.3	3.5	7	5.3	2.5	3
7	long generation times, varied across groups	30	25.3	15	3.3	3.3	3.3

Table 1.1: *Generation time distributions across age groups used in our simulation scenarios. Scenario 7 is only used in producing Figure A1 in Appendix A3.*

vector. Additionally, $\underline{\Phi}$ is an eigenvector of $\overline{K}(r)$, which corresponds to an eigenvalue of 1 and thus satisfies condition 2. Then condition 3 in the Perron-Frobenius theorem indicates

that the spectral radius of $\overline{K}(r)$ is equal to 1, i.e.

$$1 = \rho(\overline{K}(r)). \tag{1.39}$$

Eq. (1.39) is of a similar form to eq. (1.20) for the one-group model. Like the right-hand side of eq. (1.20), the function $\rho(\overline{K}(r))$, the spectral radius of the Laplace integral transform of a positive matrix, is also monotonically decreasing in r [Thieme, 1984] – this means that we can apply the same reasoning as for the single-group model.

Since $\overline{K}(0)$ is the next generation operator [Diekmann et al., 1990] yielding the expected numbers of newly infected in each of the groups assuming a single infected individual in each group, we define the overall reproduction number as:

$$\begin{aligned} R(t) &= \rho(\overline{K}(0)) \\ &= \rho\Big(\int_0^\infty \gamma(t)C(t)W(a)e^{0a}da\Big) \\ &= \rho(\gamma(t)C(t)\int_0^\infty W(a)da) \\ &= \rho(\gamma(t)C(t)) \\ &= \gamma(t)\rho(C(t)). \end{aligned} \tag{1.40}$$

Since the spectral radius of the Laplace integral transform $\overline{K}(r)$ of a positive matrix is monotonically decreasing [Thieme, 1984], we find that:

$$\text{Epidemic growth} \Longleftrightarrow r > 0 \Longleftrightarrow \rho(\overline{K}(0)) > \rho(\overline{K}(r)) \Longleftrightarrow R(t) = \gamma(t)\rho(C(t)) > 1, \tag{1.41}$$

and similarly so with $R(t) < 1$ indicating long-term outbreak declines. We then define the time-varying reproduction number for a structured population, $R(t)$, by

$$R(t) = \gamma(t)\rho(C(t)).$$

As discussed in section 1.2, renewal equations when used in practice are usually discrete; in Appendix A6, we show that the same expression for R_t arises from the analogous discrete renewal model for a structured population. We use this discrete renewal model in the

majority of our simulations, which is given by:

$$I_t^{(j)} = \frac{R_t}{\rho(C_t)} \sum_{i=1}^{N} C_t^{(ji)} \sum_{a=1}^{t-1} w_a^{(i)} I_{t-a}^{(i)}, \tag{1.42}$$

where $C_t^{(ji)}$ represents the $(i, j)^{\text{th}}$ element of a matrix of total contacts made over the course of an infection started at time t[1]; $\rho(C_t)$ is the largest eigenvalue of this contact matrix; $\{w_a^{(i)}\}$ represent elements of a discrete generation time interval distribution for group i.

1.3.3 The long-run epidemic fate does not depend on the generation time interval

Eq. (1.40) indicates that the reproduction number does not depend on the generation times of individual groups, which is a generalisation of results from previous work [Green et al., 2022]. This makes intuitive sense because, in the long-run, an epidemic grows only if each infected case on average causes more than one subsequent infection – it does not depend on *when* those infections occur, which is what the generation time encodes.

We now use simulations from eq. (1.42) to illustrate this. We again assume that the population is structured into three age groups as described in section 1.3.2.

In Figure 1.3, we show the results of our simulations of unfolding outbreaks. This first set of simulations is based on two assumptions around variation in the generation time distribution across the three groups: i) when the three groups have the same generation time distribution (red lines), or ii) each of the groups have different generation time distributions (yellow lines). We explore how the mean of the generation time distribution affects outbreak trajectories: the left-column shows simulations when the mean generation time interval is typically short; the right-hand column shows simulations when the generation time interval is generally longer. Descending the rows corresponds to increasing R_t values.

Figure 1.3 shows that the R_t threshold determines whether, ultimately, an outbreak shrinks (top row: $R_t = 0.8$), stabilises (middle row: $R_t = 1$) or grows (bottom row: $R_t = 1.2$). The assumed form of the generation time distribution across the groups does not affect whether

[1]From now on we will refer to C_t as the transposed of the contact matrix. In our analyses, in order to parameterise C_t we will always use the transposed of the contact matrix from the POLYMOD study [Prem et al., 2021] if we use this type of data.

longer-term infection counts decline, stabilise or grow. However, changes to the generation time distribution do affect the rate in calendar time at which infection counts change in the long term, and this is supported by further simulations (see Appendix A3).

For short generation time intervals, the infection counts reach their smooth long-term trajectories rapidly after an initial period where they respond to the initial conditions. For long generation times, particularly when the generation time distribution varies by group, there are substantial oscillations in the infection counts, which appear to lessen with time.

These oscillations arise at the beginning of the simulations, when the system responds to the initial conditions before reaching the long-run growth trajectories. They are more marked when the generation time is long because this effectively introduces a lag between one clump of infections and those these infections cause.

1.3.4 Stochastic models have the same R_t threshold governing their mean behaviour

We now consider stochastic renewal models. In Appendix A4 we show that the reproduction number in a stochastic, discrete, single-population renewal model governs the long-term fate of the mean of the stochastic process; this proof carries directly over to the multiple-group case. We now illustrate via simulation that the expression derived for R_t for the continuous model case (eq. (1.40)) governs the behaviour of the mean of a stochastic renewal process of the form:

$$I_t^{(j)} \sim \text{Poisson}\Big(\frac{R_t}{\rho(C_t)} \sum_{i=1}^{N} C_t^{(ji)} \sum_{a=1}^{t-1} w_a^{(i)} I_{t-a}^{(i)}\Big),$$

where $I_t^{(j)}$ denotes the incidence in group j. The dynamics of the aggregate incidence across all groups is then given by:

$$I_t \sim \text{Poisson}\Big(\frac{R_t}{\rho(C_t)} \sum_{j=1}^{N} \sum_{i=1}^{N} C_t^{(ji)} \sum_{a=1}^{t-1} w_a^{(i)} I_{t-a}^{(i)}\Big). \tag{1.43}$$

We perform simulations of eq. (1.43) and its deterministic counterpart in a three-group population assuming a contact matrix of the form given by eq. (1.37).

In Figure 1.4 we show that $R_t = \gamma_t \rho(C_t) = 1$ is the boundary that dictates the fate of the

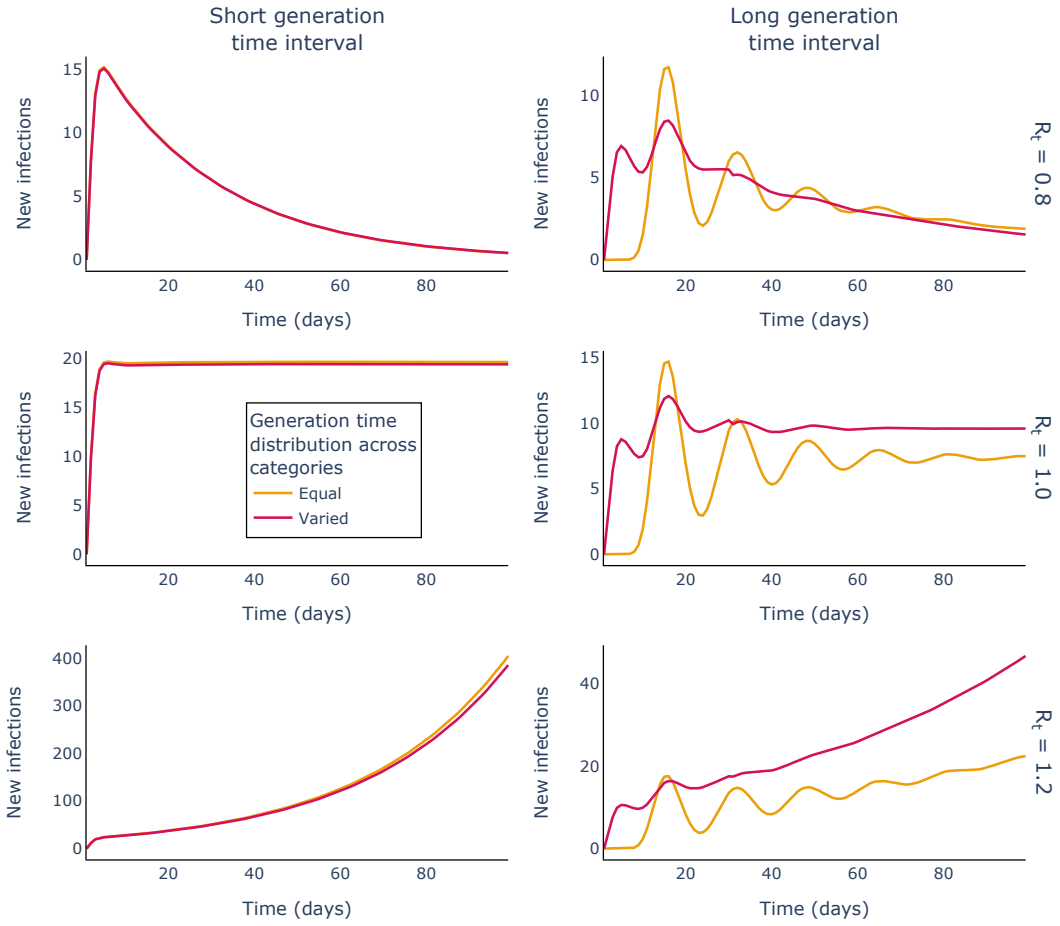


Figure 1.3: *Whether an outbreak shrinks, stabilises or grows in the long-term does not depend on the generation time distribution. The rows correspond to different R_t values, increasing from top to bottom. The columns correspond to short (left) and longer (right) generation times. The yellow lines show aggregate outbreak dynamics (i.e. the evolution of the sum of new infection counts across the three groups) when the generation time distribution is the same across all groups (given by scenarios 5 and 2 in Table 1.1 for the short and long generation time intervals, respectively); the red lines show the equivalent quantity when there is variation in the generation time distributions across the groups (the assumed generation times correspond to scenario 4 for the short generation time intervals and scenario 1 for the long generation time intervals in Table 1.1).*

aggregate incidence for the deterministic model (left-hand column) and, analogously, for the long-run mean aggregate incidence of the stochastic process.

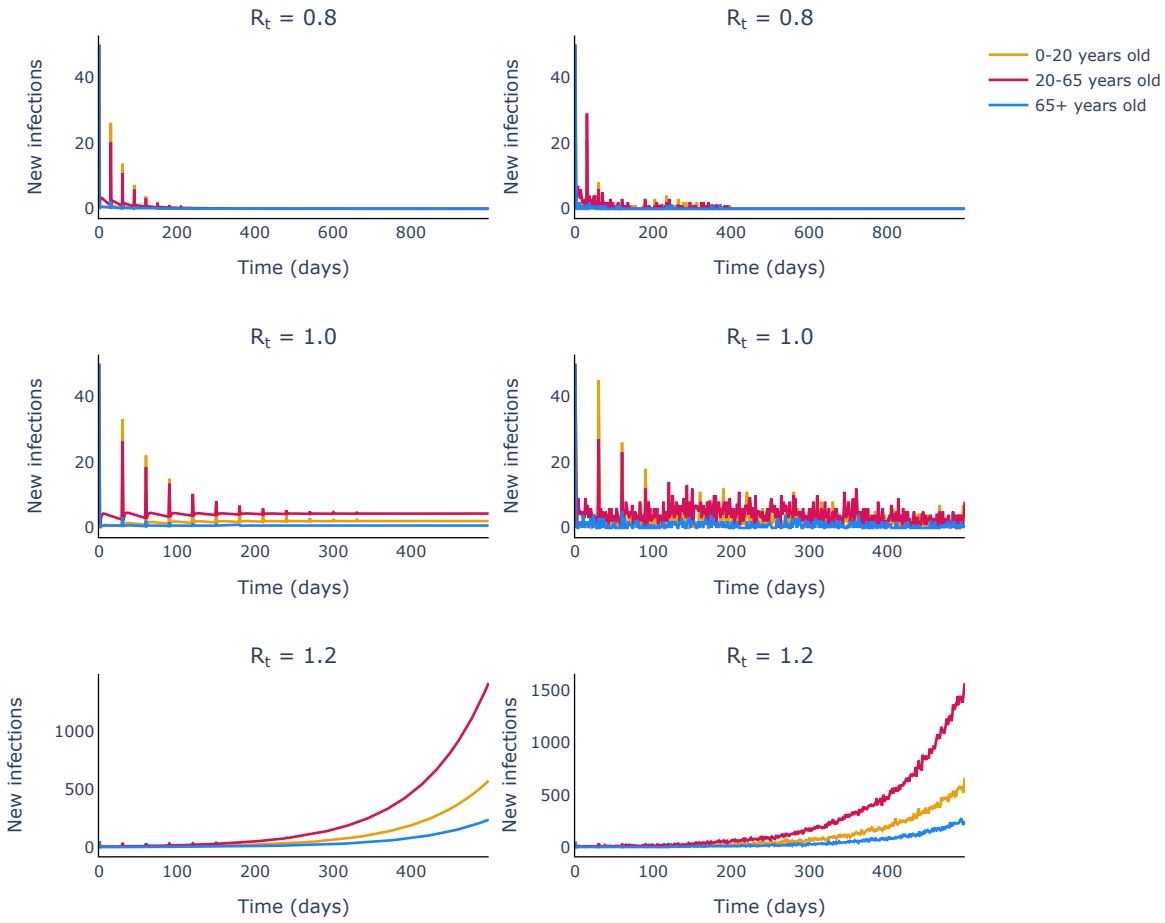


Figure 1.4: *The long-term qualitative behaviour of incidence is determined in stochastic multiple-group renewal models by the same R_t expression as for deterministic models. The left-hand column shows simulations from a deterministic model and the right-hand column shows simulations from the stochastic model. The rows dictate the R_t values assumed. For all the populations modelled we used the same age-specific generation time intervals as given by scenario 6 in Table 1.1.*

1.3.5 The relationship between R_t and growth rate r for structured populations

Considering a homogeneous population, if eq. (1.20) is divided through by $R(t) = C(t)\gamma(t)$, we obtain:

$$\frac{1}{R(t)} = \int_0^t w(a)\exp(-ra)da. \tag{1.44}$$

which previous studies have shown represents the relationship between $R(t)$, the growth rate per infected individual and r, the growth rate of the epidemic size in calendar time [Wallinga and Lipsitch, 2007]. We now derive an analogous relationship for a structured population.

To do so, we take eq. (1.39) and divide this through by $R(t) = \gamma(t)\rho(C(t))$, we obtain:

$$\frac{1}{R(t)} = \frac{\rho(\overline{K}(r))}{\gamma(t)\rho(C(t))} = \frac{\gamma(t)\rho\Big(C(t)\int_0^\infty W(a)\exp(-ra)da\Big)}{\gamma(t)\rho(C(t))} = \frac{\rho\Big(C(t)\int_0^\infty W(a)\exp(-ra)da\Big)}{\rho(C(t))}. \tag{1.45}$$

This shows that the relationship between $R(t)$ and r depends on the contact patterns in the population and the individual group generation times distributions. Interestingly, it does not depend on $\gamma(t)$, the probability that a contact results in an infection (since we have modelled this as not varying by group).

An analogous relationship holds between $R(t)$ and the doubling time of the epidemic (see Appendix A5).

We now demonstrate that the relationship between R_t and r holds using simulations of an equivalent discrete-time model, which has a corresponding relationship between these two quantities:

$$\frac{1}{R_t} = \frac{\rho\Big(C_t\sum_{a=0}^\infty W_a(1+r)^a\Big)}{\rho(C_t)}, \tag{1.46}$$

where W_a is the diagonal matrix of the discretised generation time intervals evaluated at the discrete time interval a.

In Figure 1.5 we show the relationship between r and the mean of the generation time across three values of R_t. The lines show the theoretical quantities predicted by eq. (1.46), which were determined using the SciPy optimiser method (see Appendix A7); the points show the empirical results derived from simulations of the stochastic renewal equation as given by eq. (1.43). This plot shows that the theoretical and empirical results are in reasonable agreement. It also shows that as the generation time mean increases, the rate of change of epidemic size in calendar time declines (for those cases when $R_t \neq 1$).

Same generation time distributions across groups. If the generation time distribution across all groups is the same, we can write $W(a) = w(a)\mathbb{I}_N$ and then use this to simplify eq. (1.45):

$$\frac{1}{R(t)} = \frac{\rho\Big(C(t)\int_0^\infty w(a)\mathbb{I}_N e^{-ra}da\Big)}{\rho(C(t))} = \frac{\rho\Big(C(t)(\int_0^\infty w(a)e^{-ra}da)\Big)}{\rho(C(t))}$$

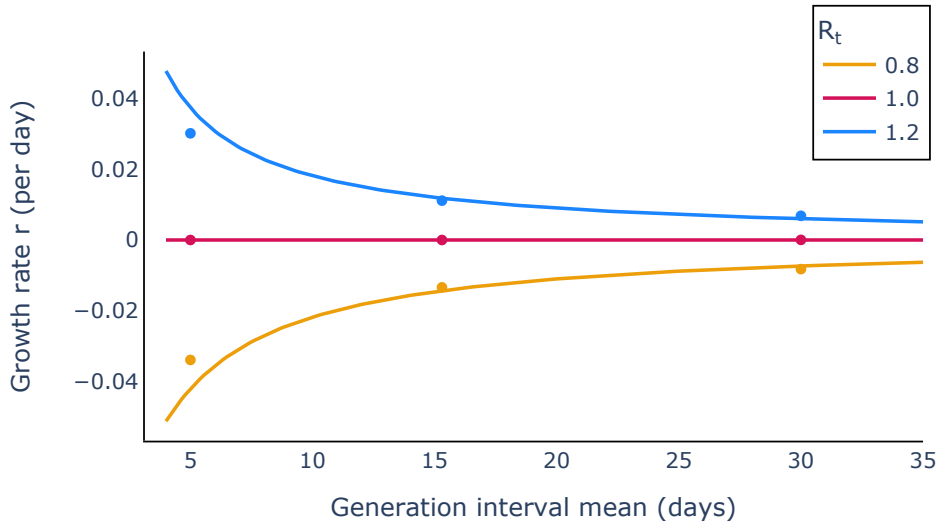


Figure 1.5: *Theoretical and simulation-based relationships between epidemic growth and generation time means largely agree. We represent in different colours the three possible epidemic long-term behaviours: epidemic growth ($R_t = 1.2$, blue), epidemic persistence ($R_t = 1$, red) and decay, respectively ($R_t = 0.8$, yellow). The solid lines indicate the theoretical trajectories given by eq. (1.46) (computed as seen in Appendix A7), and the points represent analogous estimates derived from simulations of the stochastic model.*

$$= \frac{\left(\int_0^\infty w(a)e^{-ra}da\right)\rho(C(t))}{\rho(C(t))} = \int_0^\infty w(a)e^{-ra}da,$$

to recover eq. (1.44). So, when all groups share the same generation time distribution, contact patterns do not affect the relationship between $R(t)$ and r.

Homogeneous mixing We now suppose that all groups intermix at the same rate and derive a relationship between $R(t)$ and r in this limit. Assuming groups mix homogeneously amounts to assuming an $N \times N$ contact matrix of the form:

$$C(t) = c \begin{bmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{bmatrix}.$$

In this case, $\rho(C(t)) = cN$ (see Appendix A8) meaning the reproduction number $R(t) = cN\gamma(t)$. We now consider:

$$C(t)\int_0^\infty W(a)e^{-ra}da = \begin{bmatrix} c\int_0^\infty w^1(a)e^{-ra}da & \dots & c\int_0^\infty w^N(a)e^{-ra}da \\ \vdots & \ddots & \vdots \\ c\int_0^\infty w^1(a)e^{-ra}da & \dots & c\int_0^\infty w^N(a)e^{-ra}da \end{bmatrix},$$

which has spectral radius $\rho\Big(C(t)\int_0^\infty W(a)e^{-ra}da\Big) = c\int_0^\infty \sum_{i=1}^N w^i(a)e^{-ra}da$ (see Appendix A8). This suggests the following relationship between the reproduction number and growth rate:

$$\frac{1}{R(t)} = \frac{c\int_0^\infty \sum_{i=1}^N w_s^i e^{-ra}da}{cN} = \int_0^\infty \Big(\frac{1}{N}\sum_{i=1}^N w^i(a)\Big)e^{-ra}da,$$

which is equivalent to that of a single population with a generation time distribution with elements $\bar{w}(a) := \frac{1}{N}\sum_{i=1}^N w^i(a)$; in other words, with a generation time distribution equal to the average across all groups.

1.4 Discussion

The time-dependent reproduction number, R_t, is a critical measure for tracking pathogen transmission during an infectious disease epidemic. Modelling frameworks for inferring R_t have tended to neglect heterogeneities between host groups in the population under consideration. However, heterogeneous mixing patterns and differences in the susceptibility and progression of a disease across groups can play an important role in shaping the epidemic trajectory. Here, we have developed a renewal equation framework enabling R_t to be determined while accounting for these heterogeneities.

Our analyses produced an expression for R_t which, as expected, is not contingent on the generation time distribution; we show that our R_t expression determines the long-term qualitative dynamics of the epidemic (i.e., whether case numbers eventually grow or decline) if transmission conditions remain the same. This is despite the fact that the generation time sets the relationship between R_t and the epidemic growth rate (Figure 1.5). Finally, we showed that our R_t expression also governs the expected long-term dynamics of stochastic

renewal equation models.

For the stochastic renewal equation underlying parts of our analysis, we assumed that the number of cases each day is drawn from a Poisson distribution. However, more over-dispersed distributions can be used in renewal equation models, including accounting for the possibility of super-spreading events [Ho et al., 2023; Thompson et al., 2024]. If we replace the Poisson distribution with an overdispersed cousin with the same mean, our theory (see section Appendix A4) indicates that the same threshold conditions apply.

The research presented here builds on a substantial amount of literature on the topic of R_t inference. Initial methods for estimating R_t have been extended in multiple directions, including allowing for different strains of a pathogen [Bhatia et al., 2023] and inference of the generation time or serial interval distribution [Thompson et al., 2019]. In recent work, temporal aggregation of disease incidence data (e.g., weekly rather than daily case reporting) has been accounted for in R_t inference methods [Nash et al., 2023; Ogi-Gittins et al., 2024]. Despite these extensions, an exhaustive analysis of R_t quantification in populations consisting of multiple host groups has not been undertaken until now.

As with any modelling study, the approach presented here relies on various assumptions. For robust quantification of R_t using renewal equations, reliable estimates of the generation time distribution are required. Such estimates should account for the possibility that this quantity can change during an ongoing outbreak [Chen et al., 2022; Hart et al., 2022a,b; Parag et al., 2023]. We additionally assumed that all infections are generated within the populations modelled; that is, there are no imported infections, and failing to account for their presence can lead to biases in inferred R_t values [Creswell et al., 2022]. Moreover, the renewal equation model that we have derived is only suitable for modelling directly transmitted pathogens. Extending this work to model heterogeneities that are relevant for other pathogens, such as vector-borne diseases, is another possible target for future research.

Another important limitation to using this general form of the renewal equation model is that an R_t inference method based on our work here requires reliable estimates of the contact matrix and the relative infectiousness and susceptibility for those groups within the population under consideration. However, the collection of such detailed surveillance data is not commonplace, especially at the beginning of an epidemic. Also, if changes in the population dynamics happen faster at the local compared to the global scale, that is the contact matrix

associated with a subpopulation starts differing significantly from the population-level one, this can lead to decoupled epidemic dynamics across groups. A simple workaround for this issue, however, would be to add these divergent subpopulations as separate categories to our overall population group classification.

Additionally, information on the daily variation in the contact numbers and on group-specific case-counts would also provide useful information for R_t quantification; we aim to explore the usefulness of such data in future work.

We also do not account for any differences in the rates of reporting of cases across the population groups. Consistent under-reporting of cases in all population groups will not affect the value of the overall R_t – as in this case all 'correct' infection counts $I_t^{(j)}$ will effectively be multiplied by the same under-reporting rate θ to get the reported infection incidence, which is equivalent to simply multiplying the deterministic renewal eq. (1.42) by θ. However, this no longer applies for populations in which the under-reporting rate changes over time i.e. moving from a under-reporting factor of θ to θ_t (this can happen, for example, when the number of cases rises very quickly, which overwhelms the case-reporting system and leads to longer and longer delays in reporting), or the rate of under-reporting differ between groups (for example, for COVID-19, where due to milder symptoms being expressed in children, infections in the younger population groups were reported less frequently than in the older adults). Therefore, not accounting for differences in reporting could lead to erroneous overall R_t estimation values and introduce larger bias than by using the renewal equation model for a one-group population.

Despite these simplifications, our work provides a theoretical basis for future studies that account for heterogeneity in host populations when inferring R_t. Using our analytically derived expression for R_t, the effects of a range of interventions on the epidemic fate could be explored. Non-pharmaceutical interventions (NPIs) may impact population mixing patterns, which would change the effective contact matrix C_t and, consequently, change R_t. This means that we can assess the impact of some NPIs on R_t by simply changing the contact matrix structure. As an example, the impacts of interventions that affect particular age groups, such as school or workplace closures, could be explored by modifying the contact matrix [Lovell-Read et al., 2022b]. Undertaking such an analysis in a realistic fashion would require real-time estimates of contact matrices, which could be provided from, for example, mobility data [Di Domenico et al., 2024]. These types of analyses are also possi-

ble using next-generation matrix approaches [Green et al., 2022]; by incorporating contact matrix structure information into a renewal equation framework, our approach is a hybrid between mechanistic, transmission dynamics models and more statistical approaches such as the canonical, one-group renewal equation, which require fewer assumptions. This means the analyses it allows are complementary.

Renewal equation models have proved remarkable tools in epidemic surveillance, owing to the relatively few assumptions required to infer R_t. An implicit assumption underpinning the majority of their usage is, however, that there are no important heterogeneities in the population. Here, we show how to account for these heterogeneities in a renewal equation framework for structured populations. With the advent of real-time data on individual contact patterns, these modified renewal equation approaches could form the basis of future methods for inferring epidemiological quantities such as R_t. By enabling public health policy-makers to track transmission more accurately, such approaches have the potential to allow policy-makers to make the most informed possible decisions during future outbreaks of a wide range of pathogens.

So far we have exclusively focused on theoretically deriving an expression for the overall reproduction number for a multiple-group renewal model and how R_t relates to the growth rate r of the process. However, we provided no description of how to actually infer the R_t and how the other model parameters (e.g. the choice of generation time distributions in each population group or contact matrix) might affect the accuracy of these estimates. In chapter 2, we will go on to present how the generation times impact the inferred R_t profile both for the one-group and the multiple-group renewal equation models, and explore the impact on public policy.

2 | Impact of generation times on the accuracy of R_t estimation for multiple-group population renewal models

In Chapter 1, we have developed and validated analytically and through simulations a criterion for the long-term epidemic behaviour of a multiple-group renewal equation model, in the form of the overall reproduction number R_t. Our aim for this chapter is to demonstrate how incorrect estimates of the overall generation time distribution in a population impacts the accuracy of these estimates of the R_t. We also go on to show that in populations which display volatile contact behaviours, a multiple-group population approach to renewal equation modelling is always necessary to ensure correct results of end-of-epidemic analyses. With this work we highlight the need to collect representative demographic-specific generation times, contact pattern and incidence data for improved quantification of pathogen transmission. This forms Chapter 2 of the thesis and is currently being edited in order to be submitted to the *Journal of Theoretical Biology*. CRediT author contributions for this work are as follows. Conceptualisation: All authors. Methodology: IB, BL. Investigation: IB. Writing – original draft: IB, BL. Writing – review and editing: All authors. Supervision: DG, BL, RT.

Summary

Due to its ability to summarise 'real-time' epidemic behaviour, the time-dependent instantaneous reproduction number, R_t, is a powerful tool for quantifying the effects of interventions. Renewal equation models are the predominant tool used to estimate trajectories of R_t, their success likely owing in part to the relatively few assumptions they require. One such assumption is the generation time distribution, which summarises the transmission potential of an infected case over the course of their infection. However, in general this distribution is not identical across different demographics. In this study, we use two R_t inference frameworks based on renewal equation models: one for single-group and another for a structured population. We compare the estimates of R_t generated by the two models for a variety of

population structures and generation times and investigate, both analytically and through simulations, under which conditions the conclusions drawn from these modelling paradigms coincide. We also demonstrate a methodology for selecting the generation time for the one-group model that correctly encapsulates the variations in contact patterns and disease risk across the different population groups and to correctly estimate the overall R_t using the renewal model for a homogeneous population. Finally, we use real epidemic data to showcase the differences in the time of declaration of end-of-epidemic between the one-group and multiple-group models. Our results motivate the need for the rigorous collection of detailed epidemic data that encompass the intrinsic population variations in pathogen transmission in order to improve accuracy of R_t estimates and inform public policy response.

2.1 Introduction

For ease of reference, we begin this chapter with a brief overview of the renewal equation modelling framework and the reproduction number of an epidemic.

The instantaneous time-dependent reproduction number R_t is defined as the average number of secondary infections that would be generated by an infected case at time t assuming that future transmission remains the same as at time t [Fraser, 2007; Thompson et al., 2019; Cori et al., 2013]. A reproduction number value below 1 indicates that the epidemic is expected to slow down, while a value above 1 indicates that the number of infections is expected to rise. Estimates of the time-dependent reproduction number are then used as a measure of the 'real-time' behaviour of an epidemic, which makes it highly sensitive to any changes in epidemic prevention measures, such as vaccinations or non-pharmaceutical interventions [Flaxman et al., 2020; Brauner et al., 2021]. However, fluctuations in environmental conditions, such as variations in behaviour of the population modelled or in the biology of the virus itself [Fraser, 2007; Nishiura and Chowell, 2009; Thompson et al., 2020] may substantially alter the trajectory of R_t.

One of the most popular methods for estimating the time-dependent reproduction number is by considering a renewal equation model framework for the infectious disease dynamics, which is then fitted to time series of infection incidence [Fraser, 2007; Nishiura and Chowell, 2009; Thompson et al., 2019; Flaxman et al., 2020; Abbott et al., 2020]. These models assume that each new infection observed at time t will pass on the disease to other members

in the community, generating on average R_t secondary infections recorded in the following days' incidence counts. The probability that one of these secondary infections occurs s days after the start of infectiousness of the parent infection is given by the s^{th} entry of the *generation time* vector w [Svensson, 2007].

Renewal equation models have become increasing popular in the last decade, likely in part due to the relative simplicity of their setup and reduced computational costs associated with fitting the model. There currently exist extensive software tools for R_t inference which use renewal equation models as their modelling framework of choice [Fraser, 2007; Cori et al., 2013; Thompson et al., 2020; Abbott et al., 2020; Creswell et al., 2022]. However, in general these renewal models implicitly assume a homogeneous mixing population, in which all individuals in the target population interact at equal rates. This assumption represents a very simplistic view of the true underlying transmission dynamics. Indeed, in alternative modelling frameworks, such as compartmental models (e.g [van der Vegt et al., 2022; Bouros et al., 2024]), individuals in the population are often classified based on unique defining features such as age or their vaccination status; this can be partially because age and previous immunity significantly alters the levels of the pathogen present in an infected individual and so the potential of the pathogen to spread in the population. Generation time can then vary across different demographic groups, as seen with COVID-19, where peak infectivity in adults lagged behind the one observed in children by approximately two days [Kim et al., 2022]. Additionally, age has also a strong influence on the number of people individuals will interact with on average each day, as well as on the age-distribution of those contacts; in epidemiological modelling *contact matrices* evaluated from extensive social studies are used to summarise these intricate inter-group dynamics [Prem et al., 2021].

More and more studies in the epidemiological modelling literature have started to recognise and address this weakness in the standard renewal model configuration. However, they hinge on making additional assumptions about disease dynamics (e.g. Glass et al. [2011] account for the relative differences in the number of contacts across the different population groups) or inferring additional epidemic quantities in order to produce estimates of the instantaneous reproduction number (e.g in Green et al. [2022], where the growth rate of the epidemic needs to be inferred before computing the R_t via the next generation matrix approach). Previously in Chapter 1, we analytically demonstrated an alternative methodology for calculating the R_t that does not depend on the generation time. However, as we also demonstrated in Chapter

1, the generation times impact the epidemic growth rate r_t and therefore the speed with which an epidemic increases (for $r_t > 0$ or, equivalently $R_t > 1$) or, respectively, decreases (for $r_t < 0$ or, equivalently $R_t < 1$).

In this chapter, we instead investigate whether one-group renewal equation models can be suitably adapted for heterogeneous population modelling. We prove analytically that, by considering a generation time distribution for the overall population that averages the group-specific generation times according to the long-term fractions of infections in each population category, we can generate estimates of the overall R_t using the standard one-group renewal equation that match the multiple-group population approach. Using synthetic epidemic data, we demonstrate that this no longer holds for an epidemic context characterised by continuously changing contact patterns. We also show that using the one group renewal equation framework to model a heterogenous population leads to shifts in declared epidemic milestones, as illustrated by real epidemic data from the 2009 outbreak of A/H1N1 in Japan [Statistics Bureau of Japan]. These results strengthen our case that detailed and representative collection of epidemic data, such as group-specific generation time distributions, infection incidence and contact matrices, is required for correct overall R_t estimation – any laxer standards of selecting the model parameters would run the risk of advising the lifting of interventions in cases when in fact the epidemic is not under control in one or more population groups.

2.2 Methods

2.2.1 Renewal equation models for structured populations

Back in Chapter 1, we demonstrated how to summarise the epidemic dynamics for both one-group and multiple-group populations using renewal equation models. For completeness and ease of reference, we restate the assumptions associated with this modelling approach.

Standard configurations of the renewal equation model assume a stochastic framework for accounting the number of infections arising on day t. This value, denoted by as $I_t \geq 0$, depends on previous infections according to the following formulation:

$$I_t \sim \text{Poisson}(R_t \Lambda_t), \text{ where } \Lambda_t = \sum_{s=1}^{t-1} w_s I_{t-s}, \tag{2.1}$$

where Λ_t is the *transmission potential* on day t and R_t is the instantaneous reproduction number at time t for the one-group model [Fraser, 2007; Thompson et al., 2019; Cori et al., 2013]. The transmission potential is computed as a weighted average of the number of past infections, where the infection incidence at time $t - s$ is weighted by w_s – that is the s^{th} element of the *generation time distribution* of the overall epidemic. This distribution is discrete with $\sum_{s=1}^{\infty} w_s = 1$ and $w_s \geq 0$ is defined as the probability that the secondary infection occurs s days after the parent infection. This model assumes that all individuals in the population will mix at equal rates, they are equally likely to transmit the infection and they follow the same timeline of the infection progression. This framework, which we refer to as the 'one-group' renewal equation model assumes a homogenously mixing population, in which all individuals share the same infection and transmission risk – that is, identical values of R_t on day t and the same generation time distribution is associated with all infected individuals in the population.

However, these assumptions do not necessarily hold in real-world epidemic settings: for example, children and young adults aged $0 - 20$ interact primarily with other members of their cohort, while the elderly have on average fewer contacts per day than their younger counterparts. Therefore assuming equal interaction behaviours for both of these groups is not only simplistic, but incorrect; an epidemic starting in the young age cohort [Lovell-Read et al., 2022a] will spread mostly among individuals in the same age bracket and will generate larger numbers of daily infections compared to an analogous epidemic of the same pathogen starting in the older age groups instead. We can relax the strong uniformity assumptions of the standard renewal equation model and integrate some of the mechanisms of population heterogeneity by instead considering a 'multiple-group' renewal equation approach, similar to the stochastic discrete-time renewal equation described in Chapter 1. In this framework, we follow separately the number of infections observed in each population category; the number of new infections on day t observed in population group j is given by:

$$I_t^{(j)} \sim \text{Poisson}\left(\gamma_t \sum_{i=1}^{N} C_t^{(ji)} \sum_{s=1}^{t-1} w_s^{(i)} I_{t-s}^{(i)}\right). \tag{2.2}$$

All infections in group j on day t are caused by a previous infection which can belong to any of the N predefined groups in the population. This parent infection has a probability of having been observed in population group i, s days before the daughter infection, equal to

the s^{th} element of the generation time associated with the population group of the infector, that is $w_s^{(i)}$. We use the $C_t^{(ji)}$, representing the $(i,j)^{\text{th}}$ element of a matrix of total effective contacts made over the course of an infection started at time t, to summarise the effects of heterogeneous mixing dynamics of the different population groups. Finally, we assume an identical probability at time t across all individuals in the population for a contact to become an infection, which we denote by γ_t.

2.2.2 Overall and group-specific R_t for multiple-group populations

In the following we identify analytically how the group-specific reproduction numbers of the multiple population group process relate to the overall R_t of the one-group model. We begin by summarising the contribution of previous infections observed in the population group i to the number of new infections in group j occurring at time t, which we denote as $\mathcal{L}_t^{i\to j}$:

$$\mathcal{L}_t^{i\to j} \sim \text{Poisson}\Big(R_t^{i\to j}\sum_{s=1}^{t-1} w_s^{(i)} I_{t-s}^{(i)}\Big). \tag{2.3}$$

We use the notation $R_t^{i\to j}$ to represent the average number of secondary infections in group j occurring as a result of one infection in group i if conditions remain the same as at time t; that is the reproduction number of group i into group j at time t. To get the total instantaneous reproduction number of the average individual in group i, we sum the contributions to subsequent infections across all population categories. The group-specific reproduction number of individuals in group i at time t is then given by $R_t^{(i)} = \sum_{j=1}^{N} R_t^{i\to j}$.

The individual group infection incidence in category j, $I_t^{(j)}$, is similarly computed by summing the number of new infections in population group j generated by infections in group i, $\mathcal{L}_t^{i\to j}$, over all i. Then $I_t^{(j)}$ also follows a Poisson distribution according to:

$$I_t^{(j)} = \sum_{i=1}^{N} \mathcal{L}_t^{i\to j} \sim \text{Poisson}\Big(\sum_{i=1}^{N} R_t^{i\to j}\sum_{s=1}^{t-1} w_s^{(i)} I_{t-s}^{(i)}\Big). \tag{2.4}$$

The total daily incidence I_t is computed by summing $\mathcal{L}_t^{i\to j}$ across all possible population

groups i and j, that is

$$I_t = \sum_{i=1}^{N}\sum_{j=1}^{N} \mathcal{L}_t^{i\to j} \sim \text{Poisson}(\sum_{i=1}^{N}\sum_{j=1}^{N} R_t^{i\to j} \sum_{s=1}^{t-1} w_s^{(i)} I_{t-s}^{(i)}), \tag{2.5}$$

from properties of the Poisson distribution. By comparing eq. (2.4) with eq. (2.2) governing the group-specific incidence, we arrive at an alternative definition of the reproduction number of group i into group j at time t, $R_t^{i\to j}$, which uses instead the total effective number of contacts an infected individual in group i will have in group j over the course of their infection period, $C_t^{(ji)}$, and the probability of a contact becoming an infection γ_t; this between-group reproduction number of group i into group j is given by $R_t^{i\to j} = \gamma_t C_t^{(ji)}$. Hence, eq. (2.5) becomes

$$\begin{aligned} I_t &\sim \text{Poisson}\Big(\gamma_t \sum_{i=1}^{N}\sum_{j=1}^{N} {C_t}^{(ji)} \sum_{s=1}^{t-1} w_s^{(i)} I_{t-s}^{(i)}\Big) \\ &\overset{d}{=} \text{Poisson}\Big(\sum_{i=1}^{N} \gamma_t (\sum_{j=1}^{N} {C_t}^{(ji)}) \sum_{s=1}^{t-1} w_s^{(i)} I_{t-s}^{(i)}\Big) \\ &\overset{!}{=} \text{Poisson}\Big(\sum_{i=1}^{N} R_t^{(i)} \sum_{s=1}^{t-1} w_s^{(i)} I_{t-s}^{(i)}\Big), \end{aligned} \tag{2.6}$$

where for the final step we use the identity $R_t^{(i)} = \sum_{j=1}^{N} R_t^{i\to j} = \sum_{j=1}^{N} \gamma_t C_t^{(ji)}$ to arrive at a definition of the total infection incidence in terms of the group-specific reproduction number $R_t^{(i)}$. Meanwhile, from Chapter 1, we have that the overall reproduction number of the multiple-group renewal model is defined as $R_t = \gamma_t \rho(C_t)$, where $\rho(\cdot)$ determines the spectral radius, or the maximum positive eigenvalue of a matrix. If there are no differences in the transmission across population groups, that is, the group-specific reproduction number $R_t^{(i)}$ and generation times $w_s^{(i)}$ are identical across all groups i, we can reduce the multiple-group model to the one-group Poisson renewal equation model described by eq. (2.1) if the means of the two Poisson processes are equal. The one-group and multiple-group renewal equation models are then equal if

$$\gamma_t \sum_{i=1}^{N}\sum_{j=1}^{N} {C_t}^{(ji)} \sum_{s=1}^{t-1} w_s^{(i)} I_{t-s}^{(i)} = R_t \sum_{s=1}^{t-1} w_s I_{t-s}$$

$$R_t^{(i)} \frac{1}{\sum_{j=1}^{N} C_t^{(ji)}} \Big(\sum_{i=1}^{N} \sum_{j=1}^{N} C_t^{(ji)} \sum_{s=1}^{t-1} w_s^{(i)} I_{t-s}^{(i)} \Big) = R_t \sum_{s=1}^{t-1} w_s^{(i)} \sum_{i=1}^{N} I_{t-s}^{(i)} = R_t \sum_{i=1}^{N} \sum_{s=1}^{t-1} w_s^{(i)} I_{t-s}^{(i)},$$

where R_t is defined as the overall instantaneous reproduction number for the one-group population model. Therefore, the group-specific instantaneous reproduction number $R_t^{(i)}$ in the multiple-group renewal equation model can be defined in terms of the overall time-dependent reproduction number R_t in the one-group population model according to:

$$R_t^{(i)} = R_t \frac{\Big(\sum_{i=1}^{N} \sum_{s=1}^{t-1} w_s^{(i)} I_{t-s}^{(i)} \Big) \Big(\sum_{j=1}^{N} C_t^{(ji)} \Big)}{\sum_{i=1}^{N} \sum_{j=1}^{N} C_t^{(ji)} \sum_{s=1}^{t-1} w_s^{(i)} I_{t-s}^{(i)}}. \tag{2.7}$$

2.2.3 Reproducing our results

We extend the infrastructure of the 'branchpro' software package described in [Creswell et al., 2022] to allow forward simulation of renewal equation models for populations with multiple-groups via a user-implemented contact matrix.

We also provide 'stan' [Lee et al., 2017] scripts for the inference of the group-specific reproduction number trajectories for both the multiple-group and the one-group renewal equation processes using 'stan's default sampling algorithm. All results and code are available through our public repository on Github, see: [SABS-R3-Epidemiology, 2025].

2.3 Results

In the following sections, we compare the accuracy of our estimates for the overall and group-specific R_t generated using the multiple-group renewal equation model instead of the one-group model, for a range of different simulation scenarios – with a particular focus on how the choice of generation time distributions affects the trajectories of the inferred R_t. We use a Bayesian inference approach for estimating R_t(as discussed in section 0.4), and we describe our choice of priors in this section.

2.3.1 Posterior distribution comparison: the one group versus the multiple group model

We first determine the analytical posterior distribution for the overall reproduction numbers for both the one-group and multiple-group renewal equation processes. We use the posterior forms to investigate under which conditions the overall reproduction identified by the two models are identical, that is when we are justified in using the simpler one-group approach instead of the more realistic multiple-group model, which requires finescale epidemic data.

For both the one-group and the multiple-group models we consider a λ sliding-window [Thompson et al., 2019] when inferring the instantaneous reproduction number. This approach assumes that over any interval of length λ consecutive days, the value of the reproduction number remains constant, and so pools infection information from λ days instead of one. For each subsequent R_t value the sliding window is shifted one day to the right. This way, we avoid sharp fluctuations in the reproduction number, which would occur when overfitting to the noise observed in the infection data, leading to smoother R_t trajectories.

For the renewal equation model with one group and a λ sliding window, the likelihood function of R_t is given by [Cori et al., 2013, eq. (2), Web Appendix 1]:

$$\begin{aligned} L(R_t|\underline{I_t}) &= p(I_{t-\lambda}, I_{t-\lambda+1}, \ldots, I_t|\underline{w_s}, R_t, I_0, \ldots, I_{t-\lambda-1}) = \\ &= \prod_{k=t-\lambda}^{t} \text{Poisson}(I_k; R_t \sum_{s=1}^{k-1} w_s I_{k-s}), \end{aligned}$$

where $\underline{I_t}$ represent the vector of all infection incidences until time t – that is $\underline{I_t} = \{I_0, I_1, \ldots I_t\}$.

For each of the R_t, we assume a gamma-prior distribution: $p(R_t) = \text{gamma}(R_t; \alpha, \beta)$, where α is the shape and β is the scale parameter of the distribution, i.e. it has a mean of $\alpha\beta$ and variance of $\alpha\beta^2$. Using Bayes' rule we get the exact distribution for the posterior of R_t:

$$\begin{aligned} p(R_t|\underline{I_t}) &\propto L(R_t|\underline{I_t})p(R_t) \\ &\propto \left(\prod_{k=t-\lambda}^{t} \frac{(R_t \sum_{s=1}^{k-1} w_s I_{k-s})^{I_k}}{I_k!} e^{-R_t \sum_{s=1}^{k-1} w_s I_{k-s}} \right) \frac{R_t^{\alpha-1} e^{-R_t/\beta}}{\Gamma(\alpha)\beta^\alpha} \\ &\propto R_t^{\sum_{k=t-\lambda}^{t} I_k} e^{-R_t \sum_{k=t-\lambda}^{t} \sum_{s=1}^{k-1} w_s I_{k-s}} R_t^{\alpha-1} e^{-R_t/\beta} \end{aligned}$$

$$\propto R_t^{\alpha-1+\sum_{k=t-\lambda}^{t} I_k} \exp\left\{\left(-\frac{R_t}{\beta} - R_t \sum_{k=t-\lambda}^{t}\sum_{s=1}^{k-1} w_s I_{k-s}\right)\right\}$$

$$\propto R_t^{\alpha-1+\sum_{k=t-\lambda}^{t} I_k} \exp\left\{\left(-R_t\left(\frac{1}{\beta} + \sum_{k=t-\lambda}^{t}\sum_{s=1}^{k-1} w_s I_{k-s}\right)\right)\right\}$$

$$R_t|\underline{I_t} \sim \text{gamma}\left(\alpha + \sum_{k=t-\lambda}^{t} I_k, \left(\frac{1}{\beta} + \sum_{k=t-\lambda}^{t}\sum_{s=1}^{k-1} w_s I_{k-s}\right)^{-1}\right). \tag{2.8}$$

We now determine the posterior distribution of the overall reproduction number R_t for a multiple-group renewal model with Poisson noise. In Chapter 1, we derived a definition of the overall reproduction number R_t of with the multiple-group renewal equation model (as defined in eq. (2.2)). According to this previous chapter, the overall reproduction number of the epidemic is given by as $R_t = \gamma_t \rho(C_t)$, where $\rho(C_t)$ represents the spectral radius of the transposed effective contact matrix C_t (i.e. also its maximum real eigenvalue, as the matrix is non-negative). Similar to the standard renewal equation model frameworks, the value of the overall reproduction number uniquely determines the long-term behaviour of the trajectory of new infections: if R_t is greater than 1, the epidemic will grow; alternatively, the epidemic will decay in size if R_t remains below 1. Below, we rewrite the renewal equation governing the number of infections observed in population group j on day t from eq. (2.2) in terms of the overall reproduction number:

$$I_t^{(j)} \sim \text{Poisson}\left(\frac{R_t}{\rho(C_t)} \sum_{i=1}^{N} {C_t}^{(ji)} \sum_{s=1}^{t-1} w_s^{(i)} I_{t-s}^{(i)}\right).$$

Using the properties of the Poisson distribution, we aggregate the number of new infections recorded on day t in each of the N population categories and arrive at an equivalent definition of the multiple-group renewal equation model. In this alternative definition, we provide a definition for the overall number of infections on day t in terms of the group-specific generation times and incidences, the overall R_t and the transposed effective total matrix of contacts at time t:

$$I_t \sim \text{Poisson}(\frac{R_t}{\rho(C_t)} \sum_{j=1}^{N}\sum_{i=1}^{N} {C_t}^{(ji)} \sum_{s=1}^{k-1} w_s^{(i)} I_{k-s}^{(i)}).$$

The likelihood function of R_t when using a sliding-window of length λ is given by

$$\begin{aligned} L(R_t|\underline{I_t}) &= p(I_{t-\lambda}, I_{t-\lambda+1}, \ldots, I_t|\underline{w_s}, R_t, I_0, \ldots, I_{t-\lambda-1}) = \\ &= \prod_{k=t-\lambda}^{t} \text{Poisson}(I_k; \frac{R_t}{\rho(C_t)} \sum_{j=1}^{N} \sum_{i=1}^{N} {C_k}^{(ji)} \sum_{s=1}^{k-1} w_s^{(i)} I_{k-s}^{(i)}). \end{aligned}$$

We use the same gamma(α, β) prior distribution for R_t as in the one-group case. Applying Bayes' rule once again, the posterior distribution for the multiple-group model is also a gamma-distribution with the following form:

$$\begin{aligned} p(R_t|\underline{I_t}) \propto & L(R_t|\underline{I_t})p(R_t) \\ \propto & \left(\prod_{k=t-\lambda}^{t} \frac{(\frac{R_t}{\rho(C_t)} \sum_{j=1}^{N} \sum_{i=1}^{N} {C_k}^{(ji)} \sum_{s=1}^{k-1} w_s^{(i)} I_{k-s}^{(i)})^{I_k}}{I_k!} \right. \\ & \left. e^{-\frac{R_t}{\rho(C_t)} \sum_{j=1}^{N} \sum_{i=1}^{N} {C_k}^{(ji)} \sum_{s=1}^{k-1} w_s^{(i)} I_{k-s}^{(i)}} \right) \frac{R_t^{\alpha-1} e^{-R_t/\beta}}{\Gamma(\alpha)\beta^\alpha} \\ \propto & R_t^{\sum_{k=t-\lambda}^{t} I_k} e^{-\frac{R_t}{\rho(C_t)} \sum_{k=t-\lambda}^{t} \sum_{j=1}^{N} \sum_{i=1}^{N} {C_k}^{(ji)} \sum_{s=1}^{k-1} w_s^{(i)} I_{k-s}^{(i)}} R_t^{\alpha-1} e^{-R_t/\beta} \\ \propto & R_t^{\alpha-1+\sum_{k=t-\lambda}^{t} I_k} \exp\left\{ \left(-\frac{R_t}{\beta} - \frac{R_t}{\rho(C_t)} \sum_{k=t-\lambda}^{t} \sum_{j=1}^{N} \sum_{i=1}^{N} {C_k}^{(ji)} \sum_{s=1}^{k-1} w_s^{(i)} I_{k-s}^{(i)} \right) \right\} \\ \propto & R_t^{\alpha-1+\sum_{k=t-\lambda}^{t} I_k} \exp\left\{ \left(-R_t \left(\frac{1}{\beta} + \frac{1}{\rho(C_t)} \sum_{k=t-\lambda}^{t} \sum_{j=1}^{N} \sum_{i=1}^{N} {C_k}^{(ji)} \sum_{s=1}^{k-1} w_s^{(i)} I_{k-s}^{(i)} \right) \right) \right\} \\ R_t|\underline{I_t} \sim & \text{gamma}\left(\alpha + \sum_{k=t-\lambda}^{t} I_k, \left(\frac{1}{\beta} + \frac{1}{\rho(C_t)} \sum_{k=t-\lambda}^{t} \sum_{j=1}^{N} \sum_{i=1}^{N} {C_k}^{(ji)} \sum_{s=1}^{k-1} w_s^{(i)} I_{k-s}^{(i)} \right)^{-1} \right). \end{aligned} \tag{2.9}$$

2.3.2 Differences in R_t estimates between the one-group and the multiple-group models

In this section, we investigate how selecting the generation time for the overall population when assuming a one-group population renewal equation changes the inferred overall R_t profile. If the estimates of R_t from the standard one-group renewal model match those of the multiple-group one, then we can confidently use the simpler model to study similar epidemic; however, if that is not the case, then we should use the multiple-group approach

instead for the accurate inference of R_t. A naive candidate for the one-group generation time would be to compute the average over the group-specific generation times. As we will see shortly, this choice does not always lead to accurate overall R_t estimates.

We begin by assuming we are modelling a population with two groups. For each of these groups, the generation time distributions can either (i) be identical (as plotted in the top-left panel of Figure 2.3), (ii) partially match (the top-left panel of Figure 2.5), or (iii) be completely different across the two population groups (the top-left panel of Figure 2.7). We investigate how the mean inferred overall R_t trajectory differs between the one-group and the multiple-group models respectively for each of these three generation time distribution scenarios. While we have included dedicated code for the overall R_t estimation using 'stan' in the 'branchpro' software package, we instead use the properties of the analytical form of the posterior distributions computed in eqs. (2.8) & (2.9) to produce the estimated trajectory (and their 95% confidence regions) of the overall R_t for each of the two models studied: the one-group model and the multiple-group renewal model, respectively.

For the case when the generation times between the two groups only partially align, we focus exclusively on the scenario of a highly risk-averse population group. In this situation, we assume the modelled disease will manifest and progress identically in both demographic categories; however, the risk-averse generation time distribution will only match the shape of the generation time distribution associated with the non-risk-averse up until the onset of self-isolation – that is, the time point at which the risk-averse infected case interrupts contact with susceptible individuals and can no longer spread the disease. Hence, in this scenario, the generation times of the risk-averse population group will be zero from the self-isolation time point onwards.

In Tables 2.1 and 2.2, we summarise the chosen parameters for each of the scenarios considered. All presented model parameter values are purely for illustrative purposes only and are not indicative of any particular population or disease – we are more interested in visualising how these changes qualitatively impact the estimates of R_t, rather than modelling any particular epidemic. In Figure 2.1, we plot the group-specific generation times for the different generation times scenario and, respectively, for the case when they match up to a point. These generation time distributions are presented in their discretised version, computed using the [Cori et al., 2013] method and are re-normalised to sum up to 1.

	Mean generation time		Standard deviation generation time	
Scenario	Group 1	Group 2	Group 1	Group 2
Different	5.3	7	2.3	2.3
Same, cropped	5.3	5.3 (cropped after 5 days)	2.3	2.3
Same	5.3	5.3	2.3	2.3

Table 2.1: *Generation times regimes used in our simulation scenarios.*

Parameter	**Original**	**Changed**
Contact matrix (C_t)	$\begin{pmatrix} 6 & 8 \\ 4 & 8 \end{pmatrix}$	$\begin{pmatrix} 0.4 & 4 \\ 6 & 8 \end{pmatrix}$
Probability of infection (γ_t)	[0.35, 0.05, 0.4, 0.05]	[0.35, 0.05, 0.2, 0.1]
Times of changes in γ_t	[0, 37, 60, 160]	[0, 37, 60, 160]
Initial Conditions	[50, 50]	[20, 80]

Table 2.2: *Model parameters used across our simulation scenarios.*

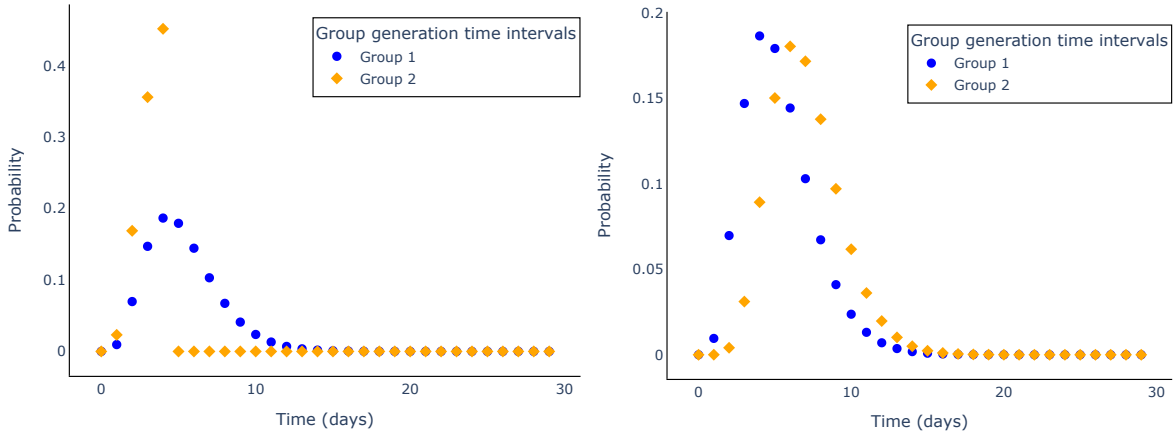


Figure 2.1: *Trajectories of the group-specific generation times for a two-group population when: (left-panel) they match up to day* 5*, (right-panel) they are different throughout. These generation times follow a discretised gamma-distribution with mean and standard deviation as described in Table 2.1.*

Additionally, to determine whether our findings about the impact of the choice of generation times on the accuracy of overall R_t estimates generally hold, we repeat our analyses, changing each of the other model parameters one at a time and plot the results in the remaining three panels of each of the Figures 2.3, 2.5 and 2.7; we consider changes in the contact

matrix C_t (top right panel), the infection rate γ_t (bottom left panel) or initial number of infections in each population group (bottom right panel). In Table 2.2, we record the values of these other model parameters that we utilise. The underlying infections data we fitted the models to in each parameterisation regime considered was generated using a two-group renewal equation model, as described in eq. (2.2), and are plotted in the Figures 2.2, 2.4 and 2.6 respectively.

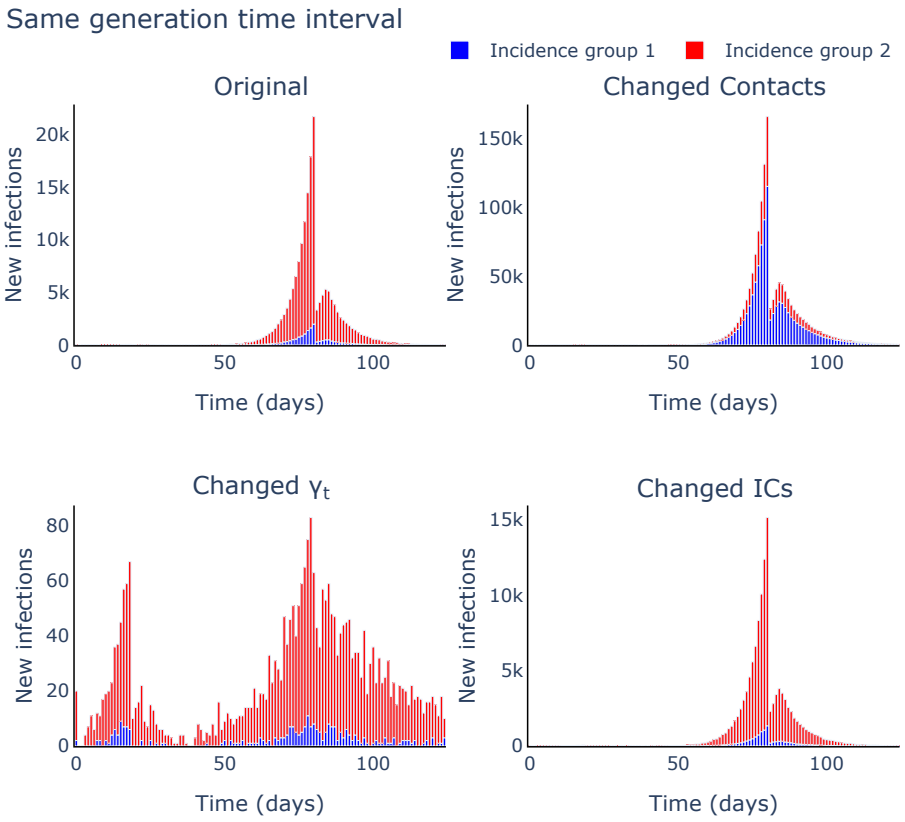


Figure 2.2: *The underlying infection data generated using a two-group population and used to infer the trajectory of the overall R_t in Figure 2.3 for both the one-group and the multiple group renewal models. The disease assumes the same baseline generation time distribution for both population groups.*

When all population groups share the same generation times, the average R_t trajectory inferred using the one-group renewal equation will match that of the overall average and 95% confidence region of R_t of the multiple-group model after an initial time delay (Figure 2.3). However, this is no longer the case if there are any differences in the generation time distributions between the two population groups (both Figure 2.5 and Figure 2.7). Moreover, these differences in the estimates of R_t do not follow a consistent pattern; using the one-group model instead of the multiple-group model can both lead to overestimates (Figure

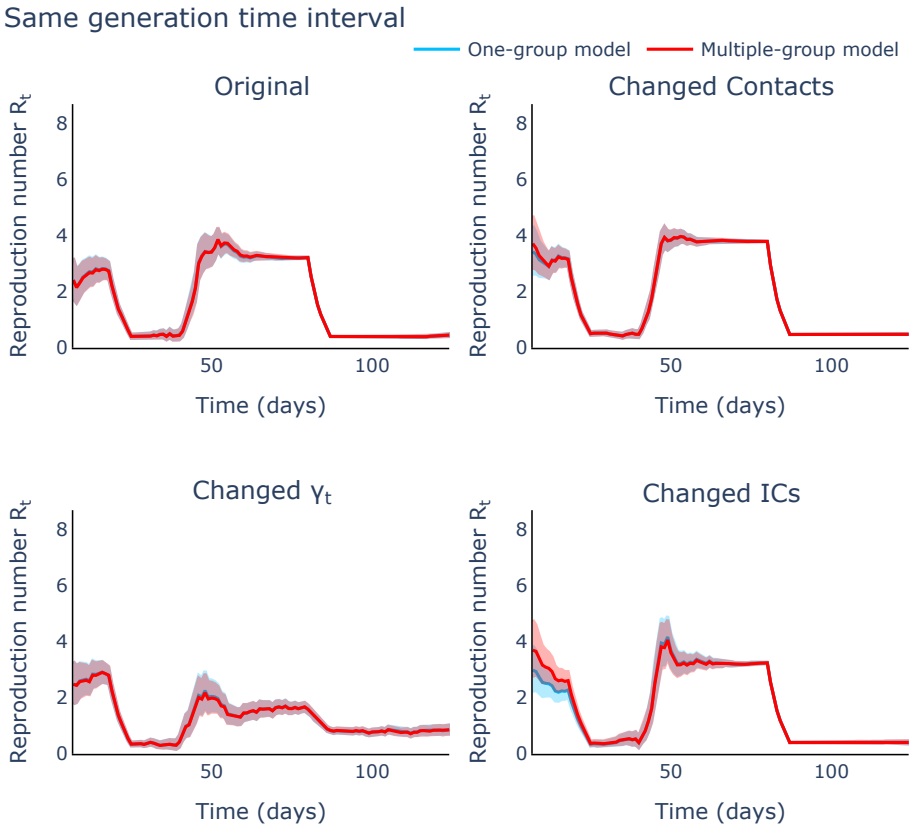


Figure 2.3: *Comparison of the mean and 95% confidence region of the overall reproduction number trajectories inferred using the multiple population group model (red lines) versus the one population group model (light blue lines) for a range of changed model parameters. The underlying infection data are generated using a two-group population. The disease assumes the same baseline generation time distribution for both population groups.*

2.7, top left panel) and underestimates (Figure 2.7, top right panel) in the R_t. This relationship between the two reproduction numbers depends entirely on the choice of contact matrix and other model parameters.

One-group model versus multiple-group model posteriors – general case

We will see in the following that this misalignment of R_t estimates is in fact a more general result: for any population with N groups with the same generation time, the overall reproduction trajectory perfectly matches the one inferred using the one-group population model with the same generation time in the long run. The discordance observed at the beginning disappears as the initial conditions of the process – the number of infections present at time zero, are forgotten – that is, the system has reached a stable age structure.

First, by comparing the forms of the posteriors in eq. (2.8) and eq. (2.9), we observe that both

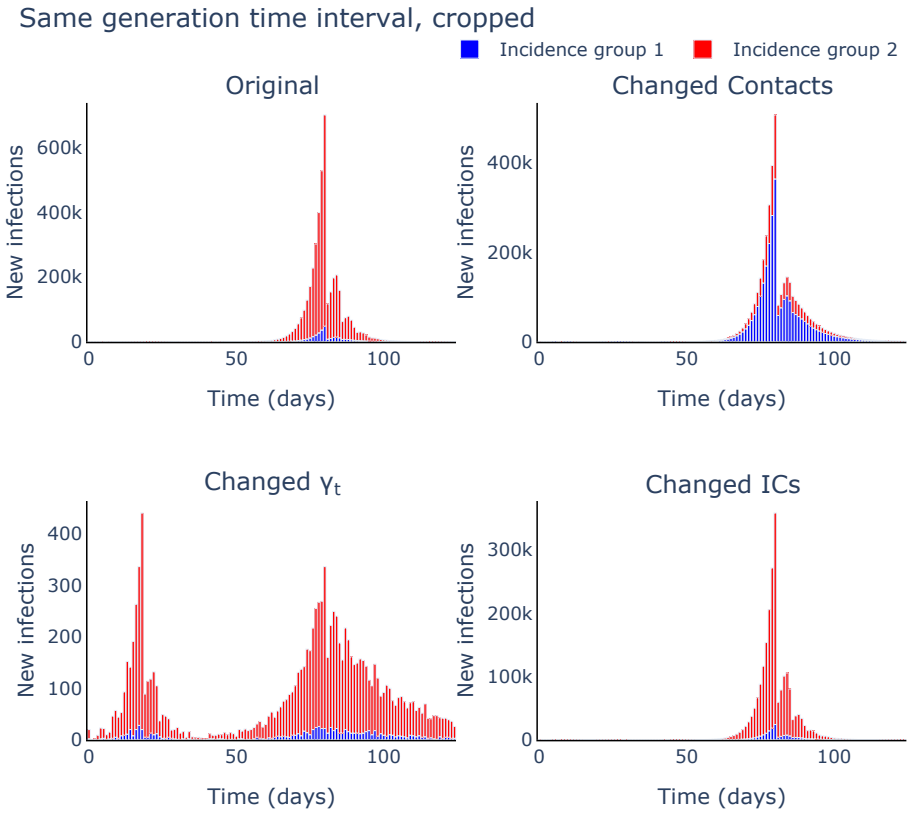


Figure 2.4: *The underlying infection data generated using a two-group population and used to infer the trajectory of the overall R_t in Figure 2.5 for both the one-group and the multiple group renewal models. The disease assumes the same baseline generation times distribution for each of the two different population groups, but in one-group the generation time distribution is cropped to reflect a degree of risk-aversion.*

share the shape parameter. Hence, the overall reproduction number trajectories produced using the one-group and multiple-group models are identical when the scale parameters of the gamma-distributions match. This holds if and only if

$$\left(\frac{1}{\beta}+\sum_{k=t-\lambda}^{t}\sum_{s=1}^{k-1}w_s I_{k-s}\right)^{-1}=\left(\frac{1}{\beta}+\frac{1}{\rho(C_t)}\sum_{k=t-\lambda}^{t}\sum_{j=1}^{N}\sum_{i=1}^{N}{C_k}^{(ji)}\sum_{s=1}^{k-1}w_s^{(i)}I_{k-s}^{(i)}\right)^{-1}$$

$$\frac{1}{\beta}+\sum_{k=t-\lambda}^{t}\sum_{s=1}^{k-1}w_s I_{k-s}=\frac{1}{\beta}+\frac{1}{\rho(C_t)}\sum_{k=t-\lambda}^{t}\sum_{j=1}^{N}\sum_{i=1}^{N}{C_k}^{(ji)}\sum_{s=1}^{k-1}w_s^{(i)}I_{k-s}^{(i)}$$

$$\sum_{k=t-\lambda}^{t}\sum_{s=1}^{k-1}w_s I_{k-s}=\frac{1}{\rho(C_t)}\sum_{k=t-\lambda}^{t}\sum_{j=1}^{N}\sum_{i=1}^{N}{C_k}^{(ji)}\sum_{s=1}^{k-1}w_s^{(i)}I_{k-s}^{(i)}.$$

We consider the ratio between the means of the two posteriors gamma-distributions as a measure of the divergence of the realistic multiple-group renewal equation model from the

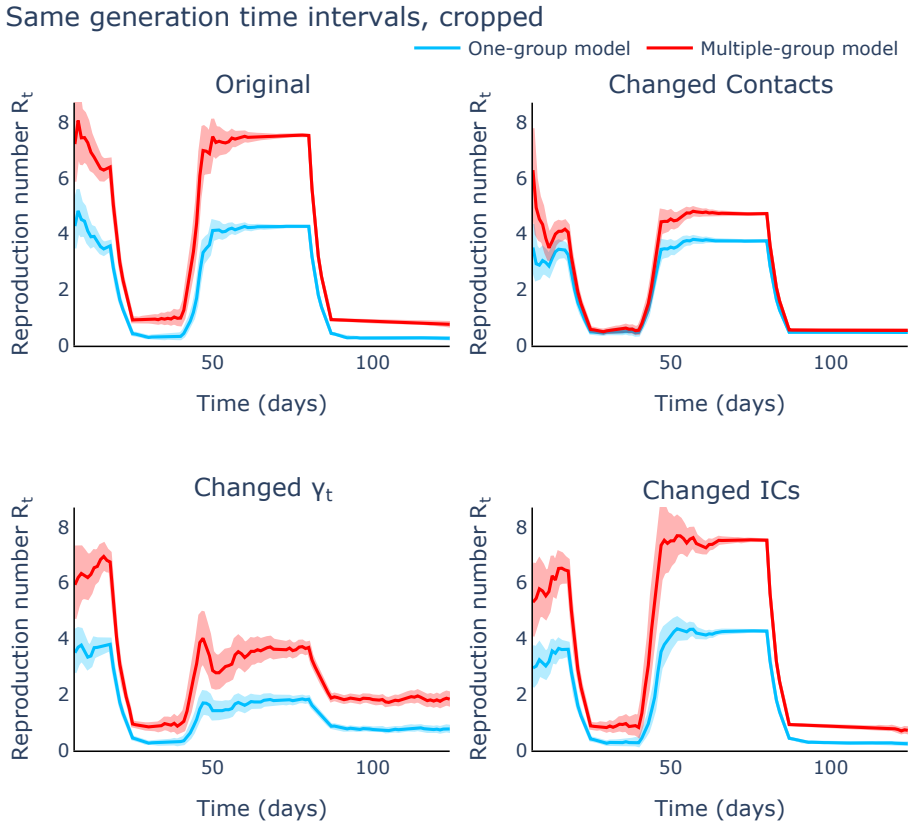


Figure 2.5: *Comparison of the mean and 95% confidence region of the overall reproduction number trajectories inferred using the multiple population group model (red lines) versus one population group model (light blue lines) for a range of parameters. The underlying infection data are generated using a two-group population. The disease assumes the same baseline generation times distribution for each of the two different population groups, but in one-group the generation time distribution is cropped to reflect a degree of risk-aversion.*

simplified one-group model. Since the mean of a gamma(α, β) distribution is equal to the product of the shape α and scale parameter β, and the shape parameter is shared between the two models, the ratio evaluated at time t becomes equal to:

$$\text{Ratio}(t) = \frac{\begin{matrix}\text{multiple-group}\\ \text{mean}\end{matrix}}{\begin{matrix}\text{one-group}\\ \text{mean}\end{matrix}} = \frac{\left(\frac{1}{\beta} + \frac{1}{\rho(C_t)} \sum_{k=t-\lambda}^{t} \sum_{j=1}^{N} \sum_{i=1}^{N} {C_k}^{(ji)} \sum_{s=1}^{k-1} w_s^{(i)} I_{k-s}^{(i)}\right)^{-1}}{\left(\frac{1}{\beta} + \sum_{k=t-\lambda}^{t} \sum_{s=1}^{k-1} w_s I_{k-s}\right)^{-1}}$$

$$= \frac{\frac{1}{\beta} + \sum_{k=t-\lambda}^{t} \sum_{s=1}^{k-1} w_s I_{k-s}}{\frac{1}{\beta} + \frac{1}{\rho(C_t)} \sum_{k=t-\lambda}^{t} \sum_{j=1}^{N} \sum_{i=1}^{N} {C_k}^{(ji)} \sum_{s=1}^{k-1} w_s^{(i)} I_{k-s}^{(i)}}. \quad (2.10)$$

Identical generation times in all population groups imply that $w_s^{(i)} = w_s, \forall i$, for all times s.

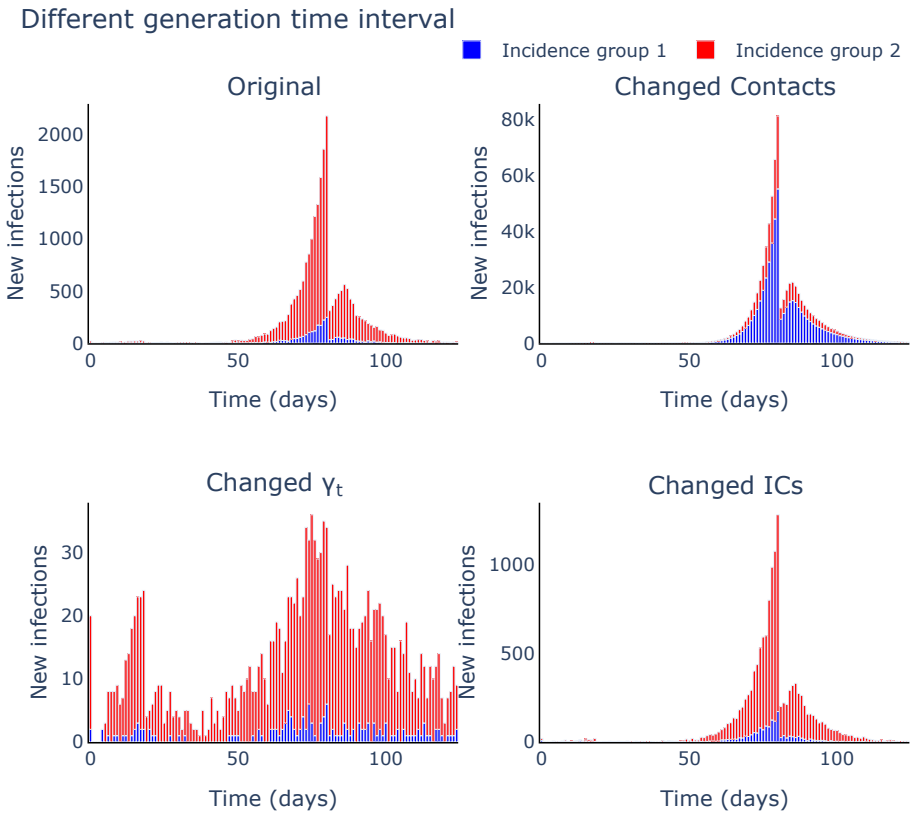


Figure 2.6: *The underlying infection data generated using a two-group population and used to infer the trajectory of the overall R_t in Figure 2.7 for both the one-group and the multiple group renewal models. The disease assumes the different baseline generation time distribution for both population groups.*

If population groups mix at equal rates within and between each other, that is, we have equal numbers of daily contacts in all groups ($C_t = \frac{m}{N}\mathbb{J}_N$ for some m total number of effective contacts per day in each group, and where $\mathbb{J}_N$ denotes the $N \times N$ matrix with all elements equal to one), the ratio defined in equation (2.10) is exactly one for all times t:

$$\frac{1}{\rho(\frac{m}{N}\mathbb{J}_N)} \sum_{k=t-\lambda}^{t} \sum_{j=1}^{N} \sum_{i=1}^{N} \frac{m}{N} \sum_{s=1}^{k-1} w_s I_{k-s}^{(i)} = \frac{1}{m} \sum_{k=t-\lambda}^{t} \sum_{i=1}^{N} \left(\sum_{j=1}^{N} \frac{m}{N} \right) \sum_{s=1}^{k-1} w_s I_{k-s}^{(i)} =$$
$$= \frac{1}{m} \sum_{k=t-\lambda}^{t} \sum_{i=1}^{N} m \sum_{s=1}^{k-1} w_s I_{k-s}^{(i)} = \frac{m}{m} \sum_{k=t-\lambda}^{t} \sum_{s=1}^{k-1} w_s \left(\sum_{i=1}^{N} I_{k-s}^{(i)} \right) = \sum_{k=t-\lambda}^{t} \sum_{s=1}^{k-1} w_s I_{k-s}.$$

However, for a more general form of the transposed effective contact matrix, we can show that the ratio in eq. (2.10) approaches one for large enough time t. At this point, the epidemic will reach a stability state where the fractions of infections in each group which either grow or decay exponentially with the same growth rate in all groups. This is the long-term solution

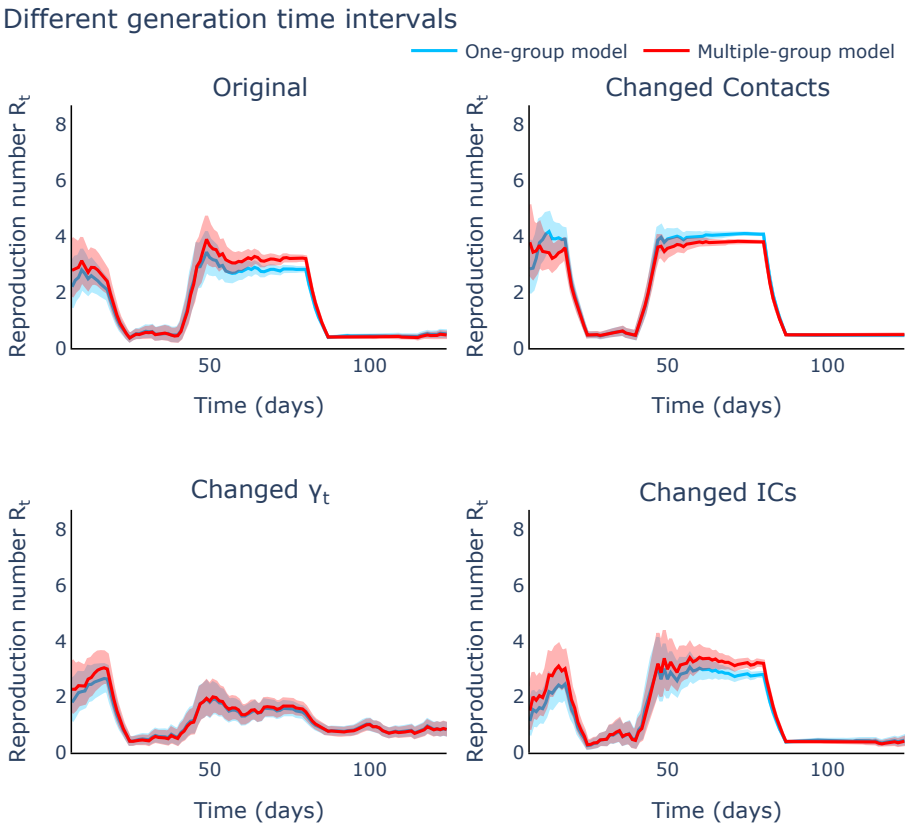


Figure 2.7: *Comparison of the mean and 95% confidence region of the overall reproduction number trajectories inferred using the multiple population group model (red lines) versus one population group model (light blue lines) for a range of parameters. The underlying infection data are generated using a two-group population. The disease assumes different generation time distributions for each of the two different population groups.*

and, for the continuous time renewal equation process, we can write this as follows (from section 1.3.2):

$$\begin{pmatrix} I_t^1 \\ \dots \\ I_t^N \end{pmatrix} = e^{rt}\underline{\Phi},$$

where r is the exponential outbreak growth rate and $\underline{\Phi}$ is the vector of the number of new infections in each population group once the systems reaches stability in terms of the fractions of infections in each group. According to Chapter 1, the equilibrium vector $\underline{\Phi}$ satisfies:

$$\underline{\Phi} = \Big(\int_0^\infty \gamma_t C_t W(s) e^{-rs} ds\Big)\underline{\Phi} = \gamma_t C_t \Big(\int_0^\infty W(s) e^{-rs} ds\Big)\underline{\Phi} := \overline{K}(r)\underline{\Phi}, \qquad (2.11)$$

where $\overline{K}(r)$ is the Laplace Integral transform of $\gamma_t C_t W(s)$ and $W(s)$ is the diagonal matrix of the generation times evaluated at time s, for all population groups – that is any continuous time point s, we have

$$W(s) = \begin{bmatrix} w^1(s) & & \\ & \ddots & \\ & & w^N(s) \end{bmatrix},$$

where $w^i(s)$ is the evaluation at the time s of the continuous time version of the discrete generation interval of group i, w_s^i. We have also shown in Chapter 1 that the spectral radius of the Laplace Integral transform of $\gamma_t C_t W_s$ is equal to one $\rho(\overline{K}(r)) = 1$. Since all population groups share the same generation time , $W(s)$ satisfies $W(s) = w_s \mathbb{I}_N$ for all time points s, where $\mathbb{I}_N$ denotes the $N \times N$ identity matrix. Applying this property to equation (2.11), we get that:

$$\underline{\Phi} = \Big(\int_0^\infty \gamma_t C_t W(s) e^{-rs} ds\Big)\underline{\Phi} = \Big(\int_0^\infty \gamma_t C_t w_s \mathbb{I}_N e^{-rs} ds\Big)\underline{\Phi} = \gamma_t \Big(\int_0^\infty w_s e^{-rs} ds\Big) C_t \underline{\Phi}, \tag{2.12}$$

and, therefore, the spectral radius of the transposed effective contact matrix C_t satisfies:

$$1 = \rho(\overline{K}(\gamma)) = \rho\left(\gamma_t \Big(\int_0^\infty w_s e^{-rs} ds\Big) C_t\right) = \gamma_t \Big(\int_0^\infty w_s e^{-rs} ds\Big)\rho(C_t). \tag{2.13}$$

Substituting the results from equations (2.12) and (2.13) into the ratio coefficient defined in equation (2.10), we conclude that for large times t we have:

$$\begin{aligned}
\text{Ratio}(t) &= \frac{\frac{1}{\beta} + \sum_{k=t-\lambda}^{t} \sum_{s=1}^{k-1} w_s I_{k-s}}{\frac{1}{\beta} + \frac{1}{\rho(C_t)} \sum_{k=t-\lambda}^{t} \sum_{j=1}^{N} \sum_{i=1}^{N} C_t^{(ji)} \sum_{s=1}^{k-1} w_s I_{k-s}^{(i)}} \\
&= \frac{\frac{1}{\beta} + \sum_{k=t-\lambda}^{t} \sum_{s=1}^{k-1} w_s \sum_{j=1}^{N} I_{k-s}^{(j)}}{\frac{1}{\beta} + \frac{1}{\rho(C_t)} \sum_{k=t-\lambda}^{t} \sum_{j=1}^{N} \sum_{i=1}^{N} C_t^{(ji)} \sum_{s=1}^{k-1} w_s I_{k-s}^{(i)}} \\
&= \frac{\frac{1}{\beta} + \sum_{k=t-\lambda}^{t} \sum_{s=1}^{k-1} w_s \sum_{j=1}^{N} e^{r(k-s)} \phi^{(j)}}{\frac{1}{\beta} + \frac{1}{\rho(C_t)} \sum_{k=t-\lambda}^{t} \sum_{j=1}^{N} \sum_{i=1}^{N} C_t^{(ji)} \sum_{s=1}^{k-1} w_s e^{r(k-s)} \phi^{(i)}} \\
&= \frac{\frac{1}{\beta} + (\sum_{k=t-\lambda}^{t} \sum_{s=1}^{k-1} w_s e^{r(k-s)})(\sum_{j=1}^{N} \phi^{(j)})}{\frac{1}{\beta} + \frac{1}{\rho(C_t)} (\sum_{k=t-\lambda}^{t} \sum_{s=1}^{k-1} w_s e^{r(k-s)})(\sum_{j=1}^{N} \sum_{i=1}^{N} C_t^{(ji)} \phi^{(i)})} \\
&= \frac{\frac{1}{\beta} + (\sum_{k=t-\lambda}^{t} \sum_{s=1}^{k-1} w_s e^{r(k-s)})(\sum_{j=1}^{N} \phi^{(j)})}{\frac{1}{\beta} + (\sum_{k=t-\lambda}^{t} \sum_{s=1}^{k-1} w_s e^{r(k-s)})(\frac{1}{\rho(C_t)} \sum_{j=1}^{N} \sum_{i=1}^{N} C_t^{(ji)} \phi^{(i)})}.
\end{aligned} \tag{2.14}$$

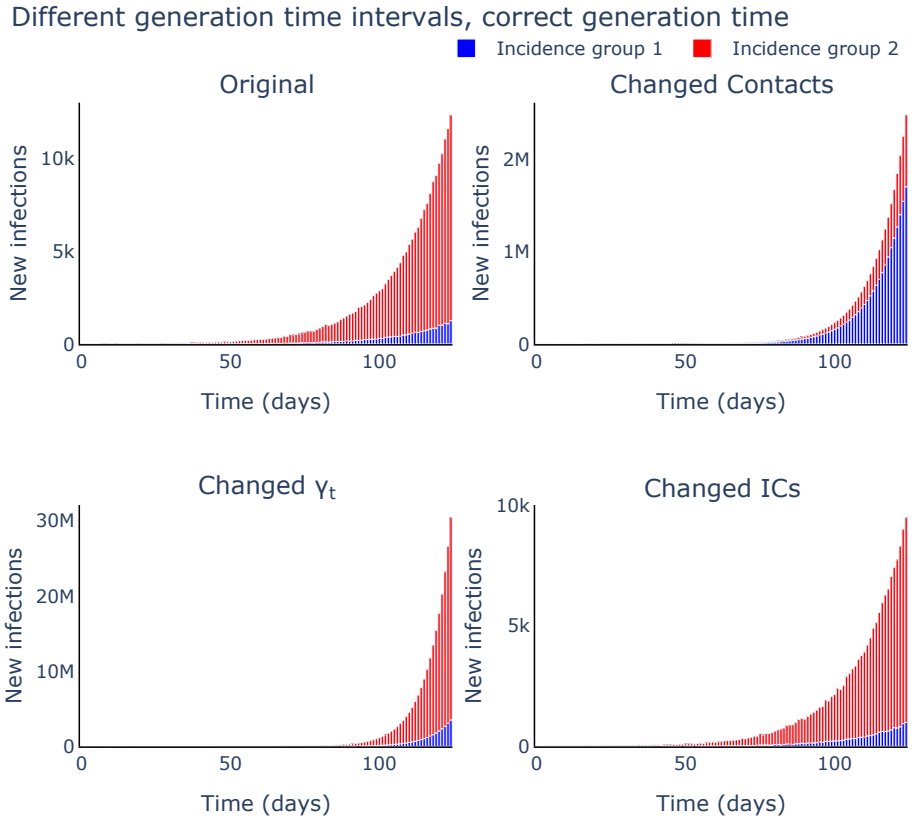


Figure 2.8: *The underlying infection data generated using a two-group population and used to infer the trajectory of the overall R_t in Figure 2.9 for both the one-group and the multiple group renewal models. The disease assumes the different baseline generation time distribution for both population groups.*

Summing across all population groups in equation (2.12) we observe that the total number of infections when the system reaches the equilibrium state vector $\underline{\Phi}$ satisfies:

$$\sum_{j=1}^{N} \phi^{(j)} = \gamma_t \Big(\int_0^{\infty} w_s e^{-rs} ds \Big) \sum_{j=1}^{N} \sum_{i=1}^{N} C_t^{(ji)} \phi^{(i)} = \frac{1}{\rho(C_t)} \sum_{j=1}^{N} \sum_{i=1}^{N} C_t^{(ji)} \phi^{(i)},$$

which when substituted into the equation (2.14) simplifies the fraction to exactly one. Therefore, when all population groups have the same generation times, the one-group and multiple-group estimates are indeed equivalent for large time t.

However, in most real case scenarios, this assumption of equality of generation times across all population groups does not always hold true [Kim et al., 2022]. Instead, we can employ a similar approach to the one above to identify an appropriate weighting of generation times for the one-group renewal equation model with respect to different group-specific generation

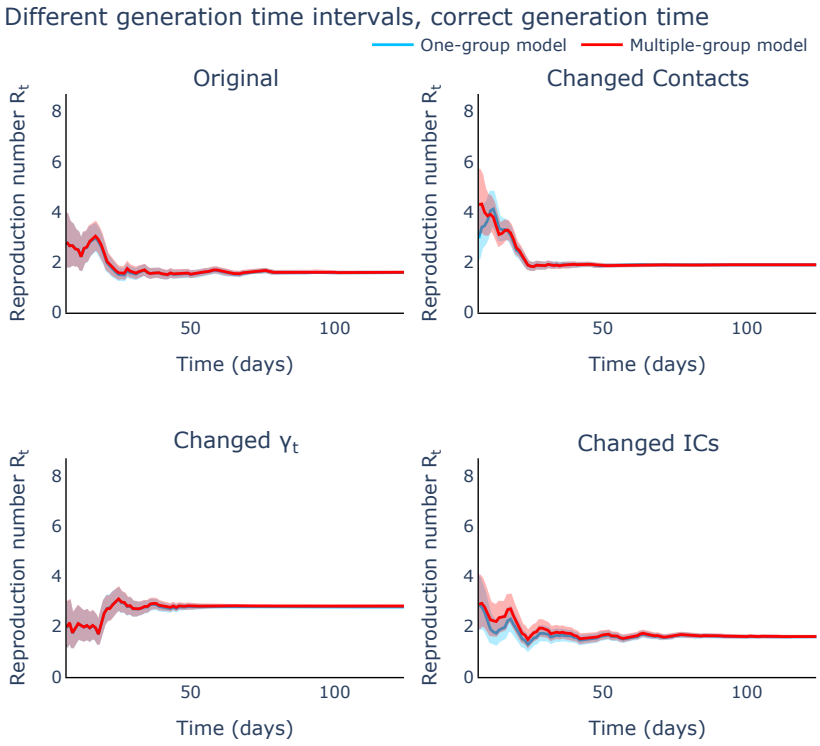


Figure 2.9: *Comparison of the mean and 95% confidence region of the overall reproduction number trajectories inferred using the multiple population group model (red lines) versus one population group model (light blue lines) for a range model parameters as described in Table 2.2, when the generation time for the one-group model is correctly calculated. The underlying infection data are generated using a two-group population.*

times in a multiple-group population.

From eq. (2.10), we have that the posterior distribution and therefore the estimated overall reproduction number at time t for the one-group and multiple-group models match if

$$\sum_{k=t-\lambda}^{t}\sum_{s=1}^{k-1} w_s I_{k-s} = \frac{1}{\rho(C_t)}\sum_{k=t-\lambda}^{t}\sum_{j=1}^{N}\sum_{i=1}^{N} {C_k}^{(ji)}\sum_{s=1}^{k-1} w_s^{(i)} I_{k-s}^{(i)}$$
$$\sum_{k=t-\lambda}^{t}\sum_{s=1}^{k-1} w_s I_{k-s} = \sum_{k=t-\lambda}^{t}\sum_{s=1}^{k-1}\frac{1}{\rho(C_t)}\sum_{j=1}^{N}\sum_{i=1}^{N} {C_k}^{(ji)} w_s^{(i)} I_{k-s}^{(i)}. \tag{2.15}$$

As seen from equation (2.12), the long-run incidence of infection in population group i is given by $I_{k-s}^{(i)} = e^{r(k-s)}\phi^{(i)}$, where $\phi^{(i)}$ satisfies

$$\underline{\Phi} = \Big(\int_0^\infty \gamma_t C_t W(s) e^{-rs} ds\Big)\underline{\Phi}.$$

Substituting this back into eq. (2.15) resolves the dependency in k, if the number of con-

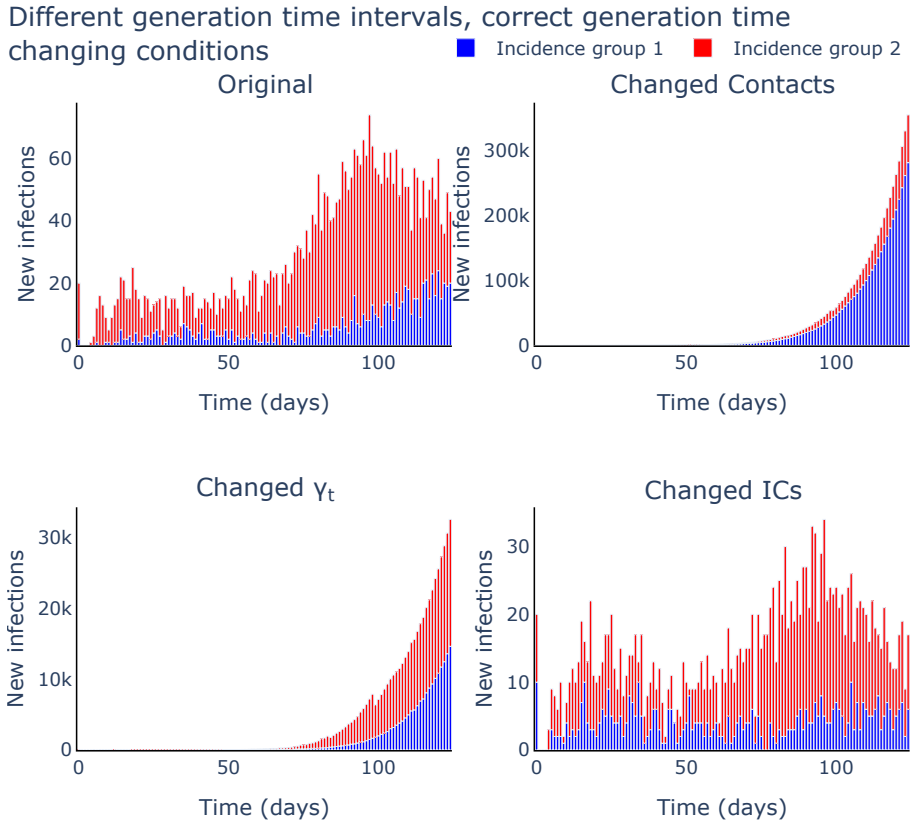


Figure 2.10: *The underlying infection data generated using a two-group population and used to infer the trajectory of the overall R_t in Figure 2.11 for both the one-group and the multiple group renewal models. The disease assumes the different baseline generation time distribution for both population groups.*

tacts (that is the contact matrix C_k) remains constant over time. Hence, a generation times distribution for the one-group model that satisfies

$$w_s = \frac{1}{\rho(C_t)} \frac{\sum_{j=1}^{N} \sum_{i=1}^{N} C_t{}^{(ji)} w_s^{(i)} I_{k-s}^{(i)}}{I_{k-s}} = \sum_{i=1}^{N} \frac{\sum_{j=1}^{N} C_t{}^{(ji)}}{\rho(C_t)} w_s^{(i)} \frac{\phi^{(i)}}{\sum_{l=1}^{N} \phi^{(l)}}. \tag{2.16}$$

will ensure equality between the posteriors of the one-group and the multiple-group renewal equation models. The fraction $\frac{\phi^{(i)}}{\sum_{l=1}^{N} \phi^{(l)}}$ represents the long-term proportion of infections in population group i and is determined by the eigenvector corresponding to the maximum eigenvalue of the matrix $M = \gamma_t C_t \int_0^\infty W(a) e^{-rs} da$, which is 1.

For a two-group population, we demonstrate an exact formula for the overall generation

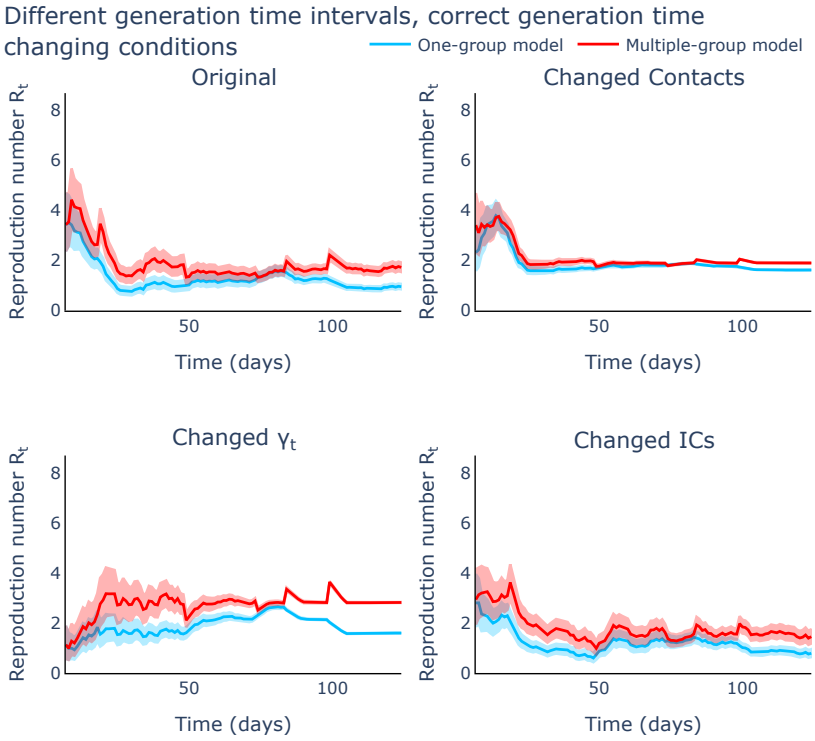


Figure 2.11: *Comparison of the mean and 95% confidence region of the overall reproduction number trajectories inferred using the multiple population group model (red lines) versus one population group model (light blue lines) for a range of changed model parameters, when the generation time for the one-group model is correctly calculated and the average number of contacts is changing frequently. The underlying infection data are generated using a two-group population.*

time distribution. Based on the above, we have that

$$\begin{pmatrix} M^{(11)} & M^{(12)} \\ M^{(21)} & M^{(22)} \end{pmatrix} \begin{pmatrix} \phi^{(1)} \\ \phi^{(2)} \end{pmatrix} = \begin{pmatrix} \phi^{(1)} \\ \phi^{(2)} \end{pmatrix}.$$

This equality can be reinterpreted in terms of the long-term proportion of infections in each of the population groups 1 and 2 as

$$\phi^{(1)} = M^{(11)}\phi^{(1)} + M^{(12)}\phi^{(2)} = C_t^{(11)}\phi^{(1)} \int_0^\infty w^{(1)}(a)e^{-ra}da + C_t^{(12)}\phi^{(2)} \int_0^\infty w^{(2)}(a)e^{-ra}da$$
$$\phi^{(2)} = M^{(21)}\phi^{(1)} + M^{(22)}\phi^{(2)} = C_t^{(21)}\phi^{(1)} \int_0^\infty w^{(1)}(a)e^{-ra}da + C_t^{(22)}\phi^{(2)} \int_0^\infty w^{(2)}(a)e^{-ra}da.$$

By adding, we obtain

$$\phi^{(1)} + \phi^{(2)} = (C_t^{(11)} + C_t^{(21)})\phi^{(1)} \int_0^\infty w^{(1)}(a)e^{-ra}da + (C_t^{(12)} + C_t^{(22)})\phi^{(2)} \int_0^\infty w^{(2)}(a)e^{-ra}da.$$

Hence, the correctly weighted one-group generation time from eq. (2.16) becomes

$$w_s = \frac{(C_t^{(11)} + C_t^{(21)})\phi^{(1)} w_s^{(1)} + (C_t^{(12)} + C_t^{(22)})\phi^{(2)} w_s^{(2)}}{\rho(C_t)(\phi^{(1)} + \phi^{(2)})}. \tag{2.17}$$

Below, we explore how this result holds in practice. We consider the same simulation conditions as before (see Tables 2.1 and 2.2 for the model parameters in each population scenario), with the slight amendment that the probability of infection γ_t only changes once at time 37, as shown in Table 2.3.

Parameter	**Original**	**Changed**
Probability of infection (γ_t)	[0.35, 0.2]	[0.35, 0.38]
Times of changes in γ_t	[0, 37]	[0, 37]

Table 2.3: *Probability of infection used in our simulation scenarios.*

In Figure 2.9, we can visually observe that the estimated value of the overall R_t for both the one-group model and the multiple-group model align for large values of t when the average number of contacts remains unchanged. However, for a contact matrix that frequently changes over the simulation period, caused by either changes in the mixing patterns or the application of social-distancing interventions, the two inferred overall R_t profiles no longer match, as it can be seen in Figure 2.11, where the levels of the contact matrix fluctuate according to the entries of Table 2.4.

Time of change	**Change in Contacts (% Original)**	**Population group affected**
20	75%	2
40	50%	2
100	80%	2
150	100%	2
170	75%	2
200	50%	2

Table 2.4: *Changes in contacts used in our simulation scenarios.*

Therefore, when considering a multiple-group population with dynamic mixing patterns, the overall generation time of the distribution needs to be continuously recomputed for accurate estimation of R_t using the standard one-group renewal equation model.

2.4 Application to real-world outbreaks

We now look at a real-world outbreak of the A/H1N1 strain of influenza in Japan in 2009 [Nishiura et al., 2009] and investigate how using a multiple-group renewal model instead of the standard one-group model changes the estimated profile of the overall reproduction number. The collected infection data are categorised by age, with individuals split into young ($0-19$ years old) and adults ($20+$ years old). In Figure 2.12, we plot the daily incidence of infection in each age group. We observe that the number of infections registered in the young exceeds by far those observed in adults – for example on day 8 of the epidemic, there were 54 infections aged between 0 and 19, with only 10 infections observed in those over 20.

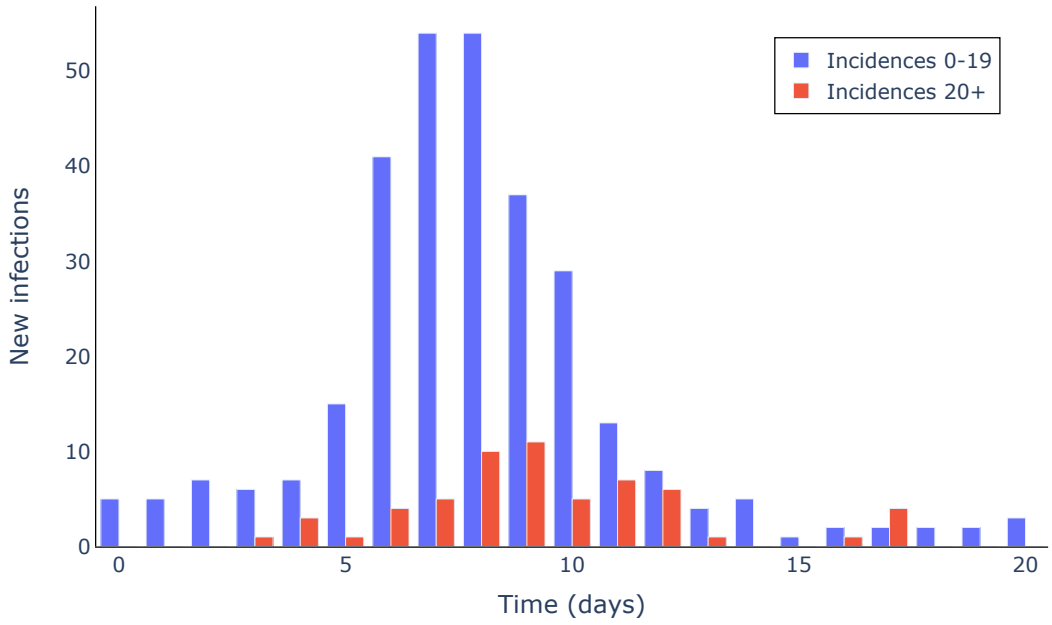


Figure 2.12: *Incidence of infection data for the* A/H1N1 *Japan outbreak in* 2009. *The number of recorded infections are split across two population groups:* $0-19$ *(blue bars) and* $20+$ *years old (red bars). The epidemic runs for* 20 *days.*

In Figure 2.13, we plot the inferred R_t trajectories for both the multiple-group and one-group renewal equation model. For both age groups considered, we assumed identical gamma-distributed generation times with mean 1.9 and standard deviation 0.893 as seen in [Nishiura et al., 2009]. For the transpose of the effective contact matrix C_t (as stipulated by eq. (2.3) in section 2.2), we use the POLYMOD contact matrix for Japan [Prem et al., 2021] and 2010 census data [Statistics Bureau of Japan] to appropriately weight the contacts in the

prescribed age groups, which gives us

$$C = \begin{pmatrix} 6.96 & 6.72 \\ 1.29 & 9.88 \end{pmatrix}.$$

We also assumed that the virus is equally transmissible in both age groups, such that the transposed effective contact matrix of the renewal process is exactly the one computed above.

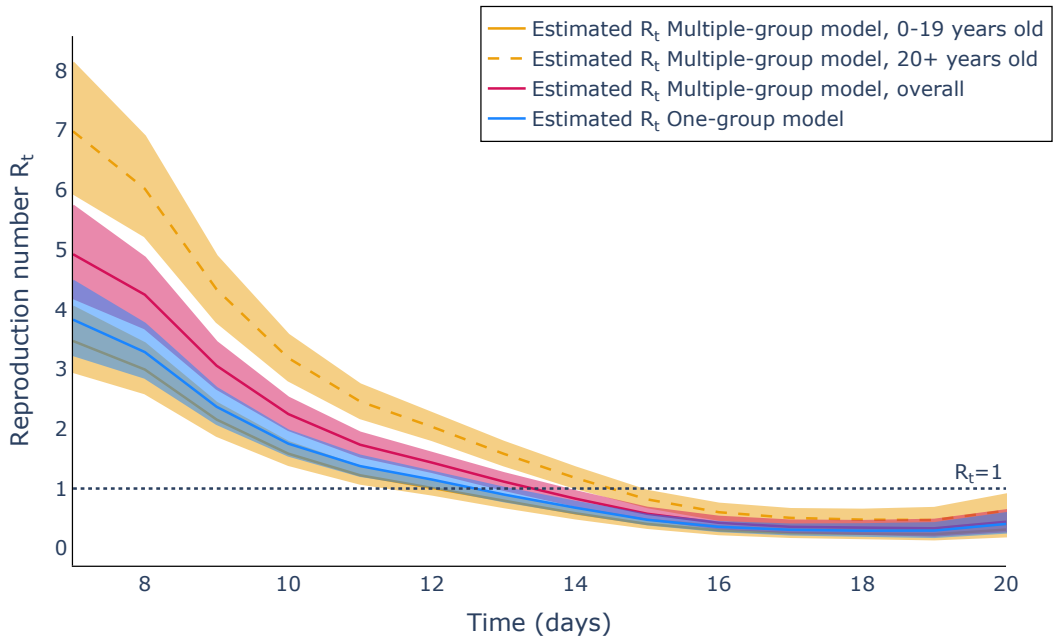


Figure 2.13: *Mean time-dependent trajectories and* 95% *confidence regions of the group-specific reproduction numbers (yellow lines) and overall* R_t *inferred from the multiple-group renewal equation model (red lines) and the one-group model (blue lines), respectively. We use* A/H1N1 *infection data from the* 2009 *Japan outbreak. The number of recorded infections are split across two population groups:* $0-19$ *and* $20+$ *years old. The epidemic runs for* 20 *days, starting from* 9^{th} *of May 2009.*

From Figure 2.13, we observe that in the case of the Japan outbreak the group-specific reproduction number is larger for the $20+$ year-olds ($R_t^{(20+)}$) than that associated with the $0-19$ year-olds ($R_t^{(0-19)}$), despite the fact that more infections were seen in the young than in adults during the outbreak. A possible explanation for this phenomenon is that there were more infections in the young at the start of the time series in Figure 2.12, and so – despite the larger group-specific $R_t^{(20+)}$ in adults - more infections were seen in children. This can also explain why the overall R_t estimated using the one-group population model more strongly resembles the inferred group-specific reproduction number of the $0-19$ year-olds, as the vast

majority of recorded infections originate in this particular age group. Comparing the overall R_t trajectories plotted in Figure 2.13 we can determine using the heterogeneous population approach that the epidemic comes under control later than otherwise expected: we reach a value of the overall $R_t < 1$ on day 14 compared to closer to day 13 in the one-group model. This phenomenon becomes especially important when larger, longer epidemics are analysed and different generation time intervals are associated with each population group considered.

2.5 Discussion

Renewal equation processes offer a fast and straightforward way of estimating the reproduction number of an emergent epidemic. The simplest and most commonly used type of renewal equation models used for epidemic dynamics assume that all individuals in a population behave the same and are equally at risk of being infected with the disease, which is most often not the case. Moreover, using a model in which the population is assumed to behave and react homogeneously to the disease, permits testing only of the impact of blanket interventions (affecting all population groups). This can lead to non-pharmaceutical interventions (NPIs) being applied to groups for which the epidemic trend is already declining, therefore wasting precious resources that can be otherwise redistributed; or we can even arrive at the wrong conclusion in short-term forecasting that the overall epidemic is declining, when it is in fact increasing in a particular population group.

In this study we showed that, in a population with different group-specific generation time distributions, the one-group model and multiple-group model generate different overall R_t estimates, if the generation time distribution used by the one-group model is simply taken as the average of the group-specific generation time distributions. We also showed that if instead we consider a generation time distribution for the one-group model that is a weighted average of the group-specific generation time distributions, which uses the long-term proportions of infections in each group, we arrive at R_t estimates generated by the one-group model and the multiple-group model that match in the long-run. However, this last result only holds for long runs of epidemics, when contact mixing conditions remain unchanged. We demonstrated that for a continuously changing contact matrix, this equivalence no longer holds.

We also used the short 2009 outbreak of A/H1N1 from Japan to illustrate the practical impact of using the one-group renewal equation model to study a multiple group population epidemic. This is further reflected by the differences between the times when the predicted R_t trajectories generated by the one-group and multiple-group model fall below 1, signalling the end of epidemic.

These results indicate that in practice, it would always be best to use the multiple group approach unless working with a homogeneous population. However, we should also acknowledge some of the limitations of the multiple-group renewal modelling approach – primarily, its reliance on far richer data compared to the standard one-group model.

Firstly, we require reliable time-dependent estimates of the contact matrix, group-specific susceptibility and infectiousness and generation times, which are not routinely collected, especially at the beginning of an epidemic, when surveillance systems are just being set up. Additionally, the multiple-group renewal model uses group-specific incidence data; if the available infection data has not been collected with such a population-structure in mind, or it possesses a different classification of the types of infected individuals compared to the groups used by the renewal model, we would not be able to use the multiple-group approach. However, we aim to explore in future work whether we can bypass this limitation by inferring the proportion of infections in each population at the same time as the reproduction number. Also in future work, we will investigate whether we can garner any additional useful information for R_t quantification by exploiting any information on the daily variation in the contact numbers and on group-specific infection counts would we might have.

We also do not account in our extended renewal framework for any importation of infections [Creswell et al., 2022], or differences in reporting across population groups – especially when the rate of under-reporting varies over time especially in the early stages of an outbreak. The latter could happen for a number of reasons: from different symptom development rates across demographies (e.g. COVID-19 leads to milder symptoms being expressed in children, compared in the older adults) to availability of resources (e.g. the reporting system progressively getting overwhelmed, especially in remote areas where testing is limited). By neglecting both migration and under-reporting, we can introduce an even more significant bias in our estimation of R_t compared to simply using the one-group renewal.

However, the theoretical gain in accuracy of R_t that the multiple-group renewal model is able

to produce has great potential in saving up resources in the long-run. This approach helps project a more accurate picture of the ongoing epidemic and hence can help prevent wasting more money on combating the effects of erroneous imposition or lifting of interventions – all we need is to already have in place the mechanisms for collecting the rich surveillance data we discussed above and be ready to scale up these data collection when a new outbreak is detected. This also motivates the need to foster and maintain some measures of low-level surveillance of the population (especially of their patterns of interactions), as we have seen that this type of information becomes indispensable in our quest to accurately model the epidemic trends.

In Chapter 1, we developed from first principles a criterion for the long-term behaviour of an epidemic (the overall reproduction number) in the context of a population with multiple interacting groups, using a renewal equation model. We validated this expression for the overall R_t, both analytically and through simulations, and looked into how it relates to the growth rate. In Chapter 2, we went on to discuss how to infer the trajectory of the overall R_t for both the one-group and multiple-group renewal models, and we explored the impact of the choice of generation times on the accuracy of these estimates. However, up until now, we have assumed perfect knowledge of all data required to parameterise the multiple-group model – primarily, we have assumed that we have access to infection incidence data by population group and to the contact matrix, which is may not be true in some cases. In Chapter 3 we will explore how different levels of knowledge of epidemic data influence the accuracy of the R_t estimation and how the multiple-group renewal framework can be utilised to improve policy response through targeted interventions.

3 | Limitations caused by data availability and better policy prediction outcomes using structured renewal processes

Previously in Chapter 2 we have analysed how the choice of generation times impacts the accuracy of our estimates of the overall R_t in a population with multiple interacting groups. In this chapter we will focus instead on how the availability of detailed epidemic data impacts the quality of model outputs. In order to do this, we conduct a comprehensive analysis of the performance of three renewal-type models for populations structured into multiple groups, with different levels of knowledge of transmission heterogeneity, against the traditional approach which assumes only one population group; we explore this in the context of COVID-19- and Ebola-like outbreaks. We also look into how these models can be used to better assess the effect of targeted interventions and, therefore, how they can be used to better inform public policy. This forms Chapter 3 of the thesis and is currently being edited in order to be submitted to the *BMC Global and Public Health* special issue. CRediT author contributions for this work are as follows. Conceptualisation: All authors. Methodology: IB, BL. Investigation: IB. Writing – original draft: IB, BL. Writing – review and editing: All authors. Supervision: DG, BL, RT.

Summary

The time-dependent reproduction number offers a simple-to-interpret quantifier of epidemic progression; in recent years, fitting case data to renewal equation processes has gained in popularity as a tool for estimating the time-dependent reproduction number, as these models require fewer assumptions to be made compared to other modelling approaches. However, these models assume that all individuals in a population behave the same and are equally at risk of becoming infected – which is not realistic for most diseases. Here, we build on the previous results from Chapters 1 and 2 to develop renewal equation inference methods for structured populations when different levels of knowledge of the time-variation in the contacts and incidence are available, and to compare their performance to the standard

'one-group' model. For each model, we infer the trajectory of the group-specific instantaneous reproduction number, using synthetic data for COVID-19- and Ebola-like pathogens and identify the impact of an imperfect knowledge of the contact matrix on the estimated group-specific reproduction number. We determine that using a renewal equation modelling approach that considers the heterogeneous population structure outperforms the one-group model in terms of R_t accuracy, and showcase how careful targeting of interventions leads to more impactful and cost-effective public health policy measures of epidemic suppression.

3.1 Introduction

In this chapter, we focus on understanding how the richness of available epidemic data, such as detailed recordings of the changes in the contact matrix and the timelines on infection incidence by population group, impacts the quality of inference of the overall reproduction number. We introduce three renewal equation models for a multiple-group population; these paradigms primarily differ by the amount of information we have on the time-dependent contact matrix. We use synthetic infection data to demonstrate how the group-specific reproduction numbers can be used to investigate the behaviour of the epidemic in different groups. This approach offers additional information on the long term behaviour of the epidemic, as in this way we are able to capture the changes in infection profiles. This becomes especially important when the overall epidemic trend shows a decline in the number of infections. In this scenario, we might wrongfully assume that the epidemic is contained, when in truth the number of new infections is increasing in a particular population group, while the epidemic is decaying faster in the other groups. Lastly, we compare the effect of both targeted and blanket interventions on the overall number of cases in a highly heterogeneous population and we discuss strategies for the implementation of cost-effective interventions aimed to reduce transmission in those demographic groups that carry the largest epidemic burden.

3.2 Methods

There are many mechanisms through which population heterogeneity impacts transmission: behavioural heterogeneity, heterogeneity in susceptibility, or heterogeneity in infectiousness to name a few. Here, behavioural heterogeneity refers primarily to heterogeneous patterns

observed in the mixing between the different population groups: for example, young people aged $0-20$ interact primarily with other members of their cohort, while the elderly have on average fewer contacts per day than their younger counterparts. The most common way in which this is incorporated into epidemiological models is by considering a time-dependent contact matrix. The $(i,j)^{\text{th}}$ entry of the contact matrix $C_t^{(ji)}$ is defined as the average number of contacts an individual in population group i has with individuals in group j at time t – that is, C_t is the transposed contact matrix. Contact matrices need not be symmetric; this is because the number of individuals present in each population group can differ drastically: in an aged population, while adults account for a large proportion of young children's daily interactions, the reverse will not be the case. Age is the most usual grouping used to stratify modelled populations. However, occupation, or level of risk-aversion can also be used to define the population-average contact matrix.

Similarly, different population groups will not be equally predisposed to become infected once in contact with an infection, or indeed have the same typical progression of the disease: for example, the mean serial interval observed for COVID-19 in children is on average two days shorter than that of the serial interval distribution observed in adults in South Korea [Kim et al., 2022].

Throughout our analyses, we consider four scenarios based on the amount of information we have at our disposal:

1. Baseline scenario – daily total number of infections and no contact information;
2. Daily infection data by population group with exact daily contact data;
3. Daily infection data by population group with average contact information;
4. Daily total number of infections with average contact information.

We first introduce a baseline renewal model (model 1. above) which assumes that all individuals in the population interact at equal rates and are equally at risk of transmitting and becoming infected – effectively modelling the entire population as a single demographic. In this scenario, we have knowledge only of the total daily incidence of infection; all information on contacts is summarised by the reproduction number, which we aim to estimate.

The last three models (2. – 4.) all assume a structured population with individuals organised in different groups. However, these models differ in the amount of contact information

we have available. For example, in the second scenario, we have access to the exact rates of contacts between the different population groups, as well as the incidence of infection in each population group, at all times during the simulation window. By comparison, in scenarios three and four, we no longer have knowledge of the exact daily variations in the contact matrix, but instead we only have access to the average rates of contact between each population group and the shape of the noise distribution about these means. For our analyses, we assume that the exact contact matrix on day t is drawn from a Poisson distribution, with its mean equal to the POLYMOD contact matrix [Prem et al., 2021]. That is,

$$C_t \sim \text{Poisson}(C), \text{ where } C = \text{average transpose POLYMOD contact matrix}, \tag{3.1}$$

and $C_t^{(ji)}$ (and $C^{(ji)}$, respectively) indicate the exact (average) number of contacts at time t an individual in population group i will have with individuals in group j. This is not an unrealistic assumption, as the average number of contacts a member of a specific population group has can fluctuate from day to day; for example on public holidays or weekends, the true number of contacts may be higher than the expected number of daily contacts.

For these three multiple-group population models, we also assume that we know the group-specific transmissibility vector $(p^{(i)})_{i=1,\dots N}$, which is kept constant at all times – N is defined here as the total number of groups the population is separated into.

For each of the three multiple-group models introduced, we assume a different generation time distribution in each population group; the generation time is defined as the probability vector that a subsequent infection is generated by an infection in group i occurring a certain number of days ago. We use $w_s^{(i)}$ to denote the s^{th} element of the generation time distribution associated with population group i. For the one-group renewal model we use the overall generation time distribution of the population which is computed as a weighted average of the group-specific generation times, as seen in eq. (2.17) in Chapter 2. However, in the example we present in section 3.3.1, we assume identical generation times across all groups – in this case, the overall generation time distribution will be equal to the group-specific one.

In the following, we derive definitions for the group-specific reproduction numbers for each of the three multiple-group models considered and determine how these relate to the overall R_t associated with each of the multiple-group model.

3.2.1 The baseline scenario – daily total number of infections and no contact information

In this model we assume that all members of the population are mixing at an equal rate, are equally susceptible and infectious, and share the same generation times. Therefore, the population behaves as one group. The new infections occurring at time t, I_t, are caused by previously infected individuals; the contribution to the number of new infections of those individuals infected s days ago is given by weighing the total infection incidence s days ago, i.e. I_{t-s}, with the s^{th} element of the generation time distribution, w_s. The typical approach to modelling this type of data is:

$$I_t \sim \text{Poisson}(R_t \Lambda_t), \text{ where } \Lambda_t = \sum_{s=1}^{t-1} w_s I_{t-s}, \tag{3.2}$$

where Λ_t is the transmission potential and R_t is the instantaneous reproduction number at time t for the one-group model.

3.2.2 Daily infection data population group with exact daily contact data

In this model, for each time point t, we have access to the exact rates of contacts across the population groups at time t – that is, the exact time-dependent contact matrix $\left(C_t\right)_{t=1,\dots T}$, where T is the length of the time series of infections. We also know the number of infections occurring in each individual population group at time t, or the aggregated infection data $\left(I_t^{(i)}\right)_{t=1,\dots T}^{i=1,\dots N}$.

We restate here the derivation of the group-specific reproduction number first seen in section 2.2.2 in Chapter 2, for ease of reference. We denote by $R_t^{i\to j}$ the average number of secondary infections in group j occurring as a result of one infection in group i, if conditions remain the same as at time t. The number $\mathcal{L}_t^{i\to j}$ of new infections in group j at time t that have been generated by infections in group i is given by

$$\mathcal{L}_t^{i\to j} \sim \text{Poisson}(R_t^{i\to j} \sum_{s=1}^{t-1} w_s^{(i)} I_{t-s}^{(i)}). \tag{3.3}$$

Since the total instantaneous reproduction number of those in group i is given by $R_t^{(i)} = \sum_{j=1}^{N} R_t^{i \to j}$, then the total daily incidence I_t is computed by summing $\mathcal{L}_t^{i \to j}$ across all possible population groups i and j – that is,

$$I_t = \sum_{i=1}^{N} \sum_{j=1}^{N} \mathcal{L}_t^{i \to j} \sim \text{Poisson}(\sum_{i=1}^{N} \sum_{j=1}^{N} R_t^{i \to j} \sum_{s=1}^{t-1} w_s^{(i)} I_{t-s}^{(i)}), \tag{3.4}$$

from properties of the Poisson distribution. For the individual group incidence $I_t^{(j)}$, we only sum $\mathcal{L}_t^{i \to j}$ over all groups which produce secondary infections in group j on day t, i.e. over all groups i. Therefore, $I_t^{(j)}$ also follows a Poisson probability distribution:

$$I_t^{(j)} = \sum_{i=1}^{N} \mathcal{L}_t^{i \to j} \sim \text{Poisson}(\sum_{i=1}^{N} R_t^{i \to j} \sum_{s=1}^{t-1} w_s^{(i)} I_{t-s}^{(i)}). \tag{3.5}$$

Alternatively, the average number of subsequent infections occurring in group j at time t due to infections in group i, that is $R_t^{i \to j}$, can also be computed by multiplying the effective total number of contacts in group j a case in group i will have throughout their infection period – that is $C_t^{(ji)} p^{(i)}$, with the probability of a contact becoming a case γ_t – as defined in section 1.2.3 and for which we assume that contacts in all population groups are equally likely to become a case once exposed to an infection. Therefore, $R_t^{i \to j}$ is given by $R_t^{i \to j} = \gamma_t C_t^{(ji)} p^{(i)}$. Substituting this back into equation (3.5), the number of new infections in group j, $I_t^{(j)}$, follows:

$$I_t^{(j)} \sim \text{Poisson}(\gamma_t \sum_{i=1}^{N} C_t^{(ji)} p^{(i)} \sum_{s=1}^{t-1} w_s^{(i)} I_{t-s}^{(i)}). \tag{3.6}$$

For ease of notation, from now on we denote the time-dependent transposed effective contact matrix C_t^*, which is calculated as the matrix product of the transposed contact matrix C_t and the transmission vector p, i.e. $C_t^{*(ji)} = C_t^{(ji)} p^{(i)}, \forall i, j \in \{1, \dots N\}$. Hence, $R_t^{i \to j} = \gamma_t C_t^{*(ji)}$ and summing across the values of j, we get the reproduction number of the process in population group i, $R_t^{(i)} = \gamma_t (\sum_{j=1}^{N} C_t^{*(ji)})$. Substituting this back into equation (3.4), we arrive at the following alternative definition of the total number of new infections at time t in terms of the group-specific reproduction number $R_t^{(i)}$:

$$I_t \sim \text{Poisson}(\gamma_t \sum_{i=1}^{N} \sum_{j=1}^{N} C_t^{*(ji)} \sum_{s=1}^{t-1} w_s^{(i)} I_{t-s}^{(i)})$$

$$\sim \text{Poisson}(\sum_{i=1}^{N} \gamma_t (\sum_{j=1}^{N} C_t^{*(ji)}) \sum_{s=1}^{t-1} w_s^{(i)} I_{t-s}^{(i)})$$
$$\sim \text{Poisson}(\sum_{i=1}^{N} R_t^{(i)} \sum_{s=1}^{t-1} w_s^{(i)} I_{t-s}^{(i)}). \tag{3.7}$$

We now look into relating the group-specific reproduction number to the overall R_t of the multiple-group process. As demonstrated back in Chapter 1, the overall reproduction number for a multiple-group population is defined as $R_t = \gamma_t \rho(C_t^*)$, where $\rho(C_t^*)$ is the spectral radius (or maximum eigenvalue) of the transposed matrix of total effective contacts, C_t^*, and γ_t is the probability of a contact becoming an infection. The total number of infections on day t for the multiple-group renewal model follows a Poisson distribution whose mean can be defined in terms of the overall R_t according to eq. (1.43), which we repeat below for ease:

$$I_t \sim \text{Poisson}\Big(\frac{R_t}{\rho(C_t^*)} \sum_{j=1}^{N} \sum_{i=1}^{N} C_t^{*(ji)} \sum_{s=1}^{t-1} w_s^{(i)} I_{t-s}^{(i)}\Big). \tag{3.8}$$

As both this and eq. (3.7) are alternative representations of the same process (modelling the total number of infections in the population on day t) and use the same noise distribution, their means will also be equal, for all times t. That is,

$$\sum_{i=1}^{N} R_t^{(i)} \sum_{s=1}^{t-1} w_s^{(i)} I_{t-s}^{(i)} = \frac{R_t}{\rho(C_t^*)} \sum_{j=1}^{N} \sum_{i=1}^{N} C_t^{*(ji)} \sum_{s=1}^{t-1} w_s^{(i)} I_{t-s}^{(i)},$$
$$R_t^{(i)} \sum_{s=1}^{t-1} w_s^{(i)} I_{t-s}^{(i)} = \frac{R_t}{\rho(C_t^*)} \sum_{j=1}^{N} C_t^{*(ji)} \sum_{s=1}^{t-1} w_s^{(i)} I_{t-s}^{(i)}.$$

Therefore, the group-specific instantaneous reproduction number $R_t^{(i)}$ associated with group i can be defined in terms of the overall time-dependent reproduction number R_t according to:

$$R_t^{(i)} = R_t \frac{\sum_{j=1}^{N} C_t^{*(ji)}}{\rho(C_t^*)}. \tag{3.9}$$

3.2.3 Daily infection data by population group with average contact information

Similar to the previous case in section 3.2.2, for this model we assume we have full knowledge of the daily counts of infections by population group $\left(I_t^{(i)}\right)_{i=1,\ldots,N}$. However, in this situation we are no longer aware of the temporal variations in the average contact matrix; instead we use the average behaviour contact matrix (extracted from population survey data such as the POLYMOD-type matrix [Prem et al., 2021]) as an estimator of the exact time-dependent contact matrix, that is $\hat{C} = \mathbb{E}\big(C_t\big)$.

Following a similar argument to section 3.2.2, we identify that the number of new cases on day t recorded for each group j is distributed according to:

$$\begin{aligned} I_t^{(j)} &\sim \text{Poisson}(\gamma_t \sum_{i=1}^{N} C^{(ji)} p^{(i)} \sum_{s=1}^{t-1} w_s^{(i)} I_{t-s}^{(i)}) \\ &\sim \text{Poisson}(\gamma_t \sum_{i=1}^{N} C^{*(ji)} \sum_{s=1}^{t-1} w_s^{(i)} I_{t-s}^{(i)}) \end{aligned} \tag{3.10}$$

where γ_t is the probability of a contact becoming an infection and $C^{*(ji)}$ is the average effective total number of contacts an infected individual in group i will have in group j.

Similarly, the group-specific reproduction number associated with population group i is given by $R_t^{(i)} = \gamma_t(\sum_{j=1}^{N} C^{(ji)})p^{(i)} = \gamma_t \sum_{j=1}^{N} C^{*(ji)}$, where C^* is the transposed average effective contact matrix, and the group-specific reproduction number, $R_t^{(i)}$ is related to the overall reproduction number R_t of the multiple-group model according to:

$$R_t^{(i)} = R_t \frac{\sum_{j=1}^{N} C^{*(ji)}}{\rho(C^*)}. \tag{3.11}$$

when the multiple-group model and the one-group model are equivalent.

3.2.4 Daily total number of infections with average contact information

This case is very similar to the previous model described in section 3.2.3, with the exception that instead of knowing the daily infection numbers in each population group, we now only

have at our disposal the daily total incidence of infection, $I_t = \sum_{i=1}^{N} I_t^{(i)}, \forall t = 1, \dots T$. However, in order to use the multiple-group population approach we require an agnostic way to break down the total number of cases by each population group; we define $(\pi_t^{(i)})_{t=1,\dots T}^{i=1,\dots N}$ as the proportions in which the daily incidence is divided across the N population groups considered. The number of new infections in each group can than be defined in terms of the total incidence at time t according to

$$I_t^{(1)} = \pi_t^{(1)} I_t,\ I_t^{(2)} = \pi_t^{(2)} I_t, \dots\ I_t^{(N)} = \pi_t^{(N)} I_t \text{ and } \sum_{i=1}^{N} \pi_t^{(i)} = 1, \forall t = 1, \dots T. \tag{3.12}$$

It is important to note that in this scenario we have **no** knowledge of $(\pi_t^{(i)})_{t=1,\dots T}^{i=1,\dots N}$; indeed, as we mention in section 3.2.6, we need to infer these quantities at the same time as the group-specific reproduction number trajectories. Substituting the definition (3.12) of the infection counts in each individual group back into eq. (3.7), we arrive at the following equation governing the dynamics of the total number of new infections at time t:

$$\begin{aligned} I_t &\sim \text{Poisson}(\gamma_t \sum_{i=1}^{N} \sum_{j=1}^{N} C^{*(ji)} \sum_{s=1}^{t-1} w_s^{(i)} \pi^{(i)} I_{t-s}) \\ &\sim \text{Poisson}(\sum_{i=1}^{N} R_t^{(i)} \sum_{s=1}^{t-1} w_s^{(i)} \pi^{(i)} I_{t-s}). \end{aligned} \tag{3.13}$$

Similarly, eq. (1.43) which describes the process in terms of the overall reproduction number becomes now

$$I_t \sim \text{Poisson}\Big(\frac{R_t}{\rho(C^*)} \sum_{j=1}^{N} \sum_{i=1}^{N} C^{*(ji)} \sum_{s=1}^{t-1} w_s^{(i)} \pi^{(i)} I_{t-s}\Big).$$

Hence, the group-dependent instantaneous reproduction number depends on the overall instantaneous reproduction number R_t of the multiple-group model according to eq. (3.11).

3.2.5 Modelling Infrastructure

We extend the infrastructure of the 'branchpro' software package [Creswell et al., 2022] to allow running forward simulations of renewal equation models for populations with multiple-groups. We also implement a method on the multiple-group renewal model class which allows the user to input an average behaviour contact matrix.

3.2.6 Reproduction Number Inference

In order to compute the reproduction number trajectory for each of the four models introduced above, we use a similar approach to the 'epiestim' method [Cori et al., 2013] with a sliding-window [Thompson et al., 2019] of length λ (typically 7 days) to smooth out the inferred R_t curves. This approach identifies the value of R_t that maximises its posterior distribution. For the one-group, homogeneous population, the likelihood function is given by:

$$L(R_t) = \prod_{k=t-\lambda}^{t} \text{Poisson}(I_k; R_t \sum_{s=1}^{k} w_s I_{k-s}),$$

where we use the notation $\text{Poisson}(x;\phi) = \frac{\phi^x e^{-\phi}}{x!}$. For each of the R_t we assume a gamma distribution for the prior with shape and scale parameters α and β respectively. Hence, the prior probability density function (p.d.f) is given by $R_t \sim \text{Gamma}(R_t; \alpha, \beta)$ and the mean and variance of the distribution are $\alpha\beta$ and $\alpha\beta^2$, respectively.

We use 'stan' [Lee et al., 2017] to infer the group-specific reproduction number trajectories for both the multiple-group and the one-group renewal equation processes as follows:

- **Baseline scenario – daily total number of infections and no contact information**: For this model we follow the basic approach of inferring the time-dependent overall R_t profile we described in section 2.3.1.
- **Daily infection data population group with exact daily contact data**: We infer the overall reproduction number, R_t, from equation (3.8), and use the correction terms defined in equation (3.9) to obtain the group-specific reproduction numbers $R_t^{(i)}$.
- **Daily infection data by population group with average contact information**: Similar to above, we infer the overall R_t parameter (using eq. (3.8) and dropping the t-dependence in the transposed effective matrix of total contacts) and use the equivalence relation in eq. (3.11) to determine estimates of the group-dependent reproduction numbers $R_t^{(i)}$.
- **Daily total number of infections with average contact information**: In this case, we have no information on the daily values of the group-specific fractions of the total new infections observed, i.e. we have no knowledge of the (π_t) vector defined as in

(3.12). Therefore, on top of inferring the trajectory of the overall reproduction number R_t, we also infer the fractions $\pi_t^{(i)}$ by which we weight the total incidence of infection at time t to determine the partition on new infections across the different population groups. Since $\sum_{i=1}^{N} \pi_t^{(i)} = 1, \forall t = 1, \dots T$, we assume a Dirichlet prior distribution for the π_t parameter, specifically $\pi_t \sim \text{Dirichlet}(1, 1, \dots 1)$. Similar to above, we use eq. (3.9) to weight the estimated overall R_t and thus obtain the group-dependent reproduction numbers $R_t^{(i)}$.

3.3 Results

3.3.1 Estimated reproduction number trajectories in the multiple-group model versus one-group model

In the case when the population is structured into two or more groups with very different mixing behaviours, we visually demonstrate that under specific conditions the group-specific reproduction number trajectories of the these population groups can exhibit discordant behaviours. To exemplify this phenomenon, we use a two-group population characterised by a highly transmissible demographic group, with a much larger number of the total effective contacts throughout the infection compared to the other less transmissible group (see Appendix B1 for more details).

In the following, we use synthetic infection data to assess how the three multiple-group renewal equation models defined in section 3.2 compare to the one-group model and how these estimates compare to the underlying group-specific R_t profiles used to generate the incidence data. This analysis also allows us to determine the extent of the knowledge of the contact rates across different population groups required by our multiple-group renewal process framework in order to substantially improve the accuracy of our group-specific R_t estimates.

We run a forward simulation for our two-group population over 100 days. We assume that both demographic groups share the same generation time interval, which follows a gamma-distribution with mean 15.3 and standard deviation 9.3, similar to that of an Ebola outbreak [Chowell and Nishiura, 2014]. For the time-dependent probability of a contact becoming a case parameter γ_t, we consider a two-step function: $\gamma_t = 0.5$ until day 30, which drops to

a value of 0.05, where it stays fixed until day 60, when it increases back to 0.2 and remains there until the end of the simulation. We consider the following transposed total contact matrix:

$$C = \begin{pmatrix} 8 & 4.8 \\ 1.2 & 3.84 \end{pmatrix},$$

and assume that the first population group is twice as transmissible as the second one – that is the transmissibility vector $\underline{p}$ is given by $\underline{p} = (1, 0.5)$. Therefore, the transpose of the effective total contact matrix of the population becomes

$$C^* = \begin{pmatrix} 8 & 4.8 \\ 0.6 & 1.92 \end{pmatrix}.$$

To generate our synthetic infection incidence data (which we present in Figure 3.1), we consider Poisson-distributed perturbations about the transposed effective matrix of total contacts C^* – that is, we assume that while the level of contacts between each population group remain constant on average throughout the simulation (C^*), there are daily variations in contacts about these rates which follow a Poisson distribution, i.e. $C_t^* \sim \text{Poisson}(C^*)$. This

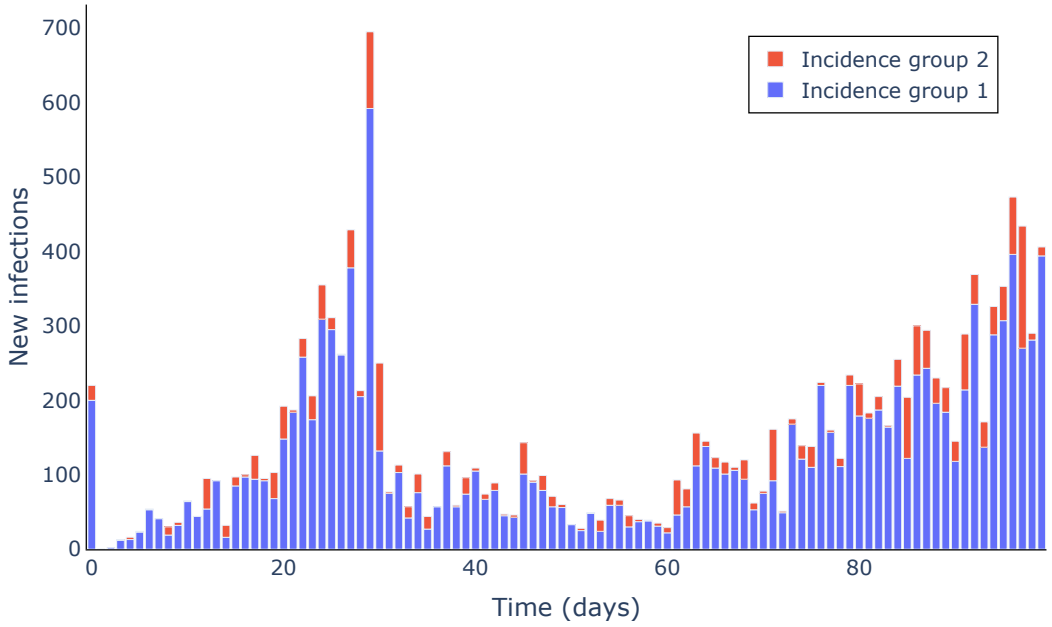


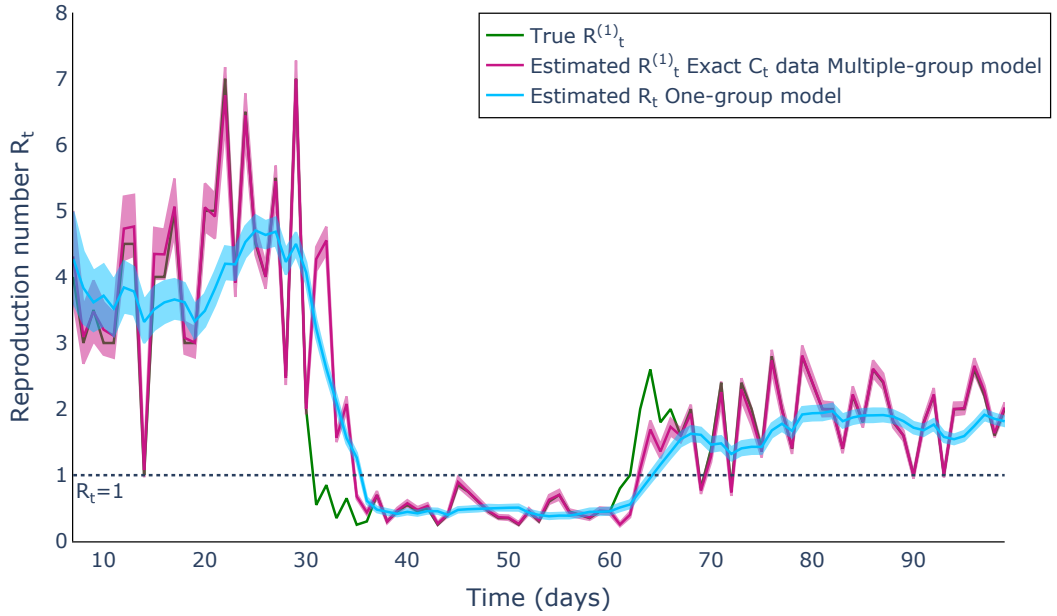
Figure 3.1: *Synthetic infection data generated from a two-group population over* 150 *days using a multiple-group renewal equation process. The model uses a two-step* γ_t *function of the probability of a contact becoming a case, with all other parameters chosen as described in section 3.3.1.*

effect leads to jagged trajectories of the true group-specific R_t, which fluctuate about the same mean unless the general epidemic trends change, such as a change in the average contact matrix C^* or the probability of a contact becoming a case γ_t. For the initial conditions, we assume 200 and, respectively, 20 infections in each of the population groups. All population parameters are chosen for illustrative purposes only – we are simply analysing the effects of the differing amounts of information on the mixing patterns and the infection incidence on the accuracy of R_t estimation for a theoretical disease and population.

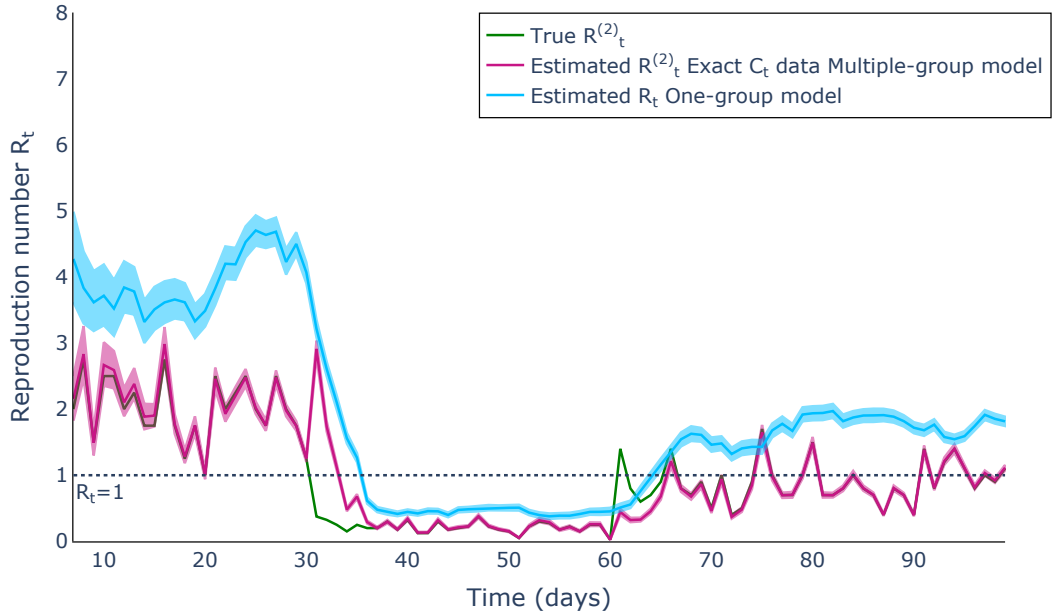
As the average total number of contacts in group 1 is larger than that realised in group 2 (8.6 compared to 6.72), we expect that both the estimated and the underlying group-specific reproduction number of this group – $R_t^{(1)}$ – to sit higher than its equivalent for the other population group, $R_t^{(2)}$. From Figure 3.2, we observe that the first multiple-group model we introduced in section 3.2.2 – which assumes knowledge of the daily variations in the contact rates and the daily incidence of infection broken down by population group, accurately reconstructs the true $R_t^{(1)}$ and $R_t^{(2)}$ profiles. While we expect that the inferred reproduction number profile R_t from the one-group model to lie between the two true group-specific reproduction number trajectories $R_t^{(1)}$ (Figure 3.2(a)) and $R_t^{(2)}$ (Figure 3.2(b)) at all times t, we observe that in the case of this particular synthetic epidemic, the one-group R_t, more strongly resembles the population group 1 dynamics – that is, the one-group R_t is closer in mean to $R_t^{(1)}$ trajectory. This phenomenon is explained by the overwhelmingly larger number of total effective contacts individuals in group 1 will have compared to group 2 on average (12.8 compared to 2.52), which implies that group 1 will completely dominate the overall dynamics of the modelled epidemic.

In Figure 3.3, we plot the trajectories of the inferred group-specific reproduction numbers for the remaining two renewal models for multiple-group populations, which we described in sections 3.2.3 and 3.2.4 respectively, against the overall R_t estimated by the baseline model, when the population is modelled as only one group. The first of these group models – for which we assume knowledge of the daily infection incidence by population group, but lack knowledge of the daily contact rates across the groups apart from the average effective contact matrix, still does reasonably well at estimating the group-specific R_t profiles.

We arrive to similar results when we are instead restricted to only knowing the daily total number of new infections recorded. In this case we also infer the daily fractions of infections in each group, i.e. the (π_t) vector, on top of the overall reproduction number, as we



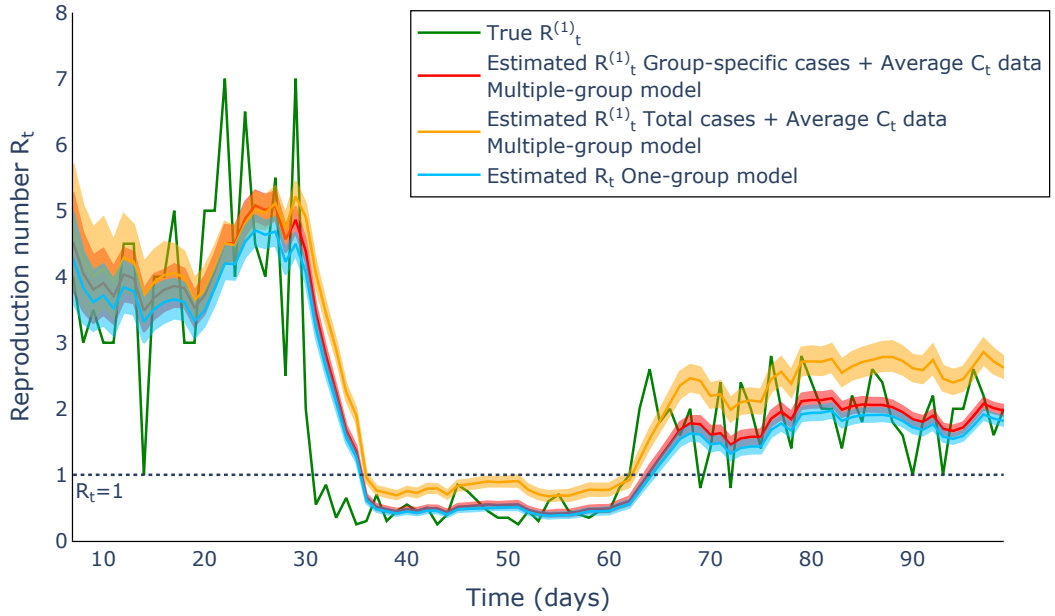
(a) Population group 1



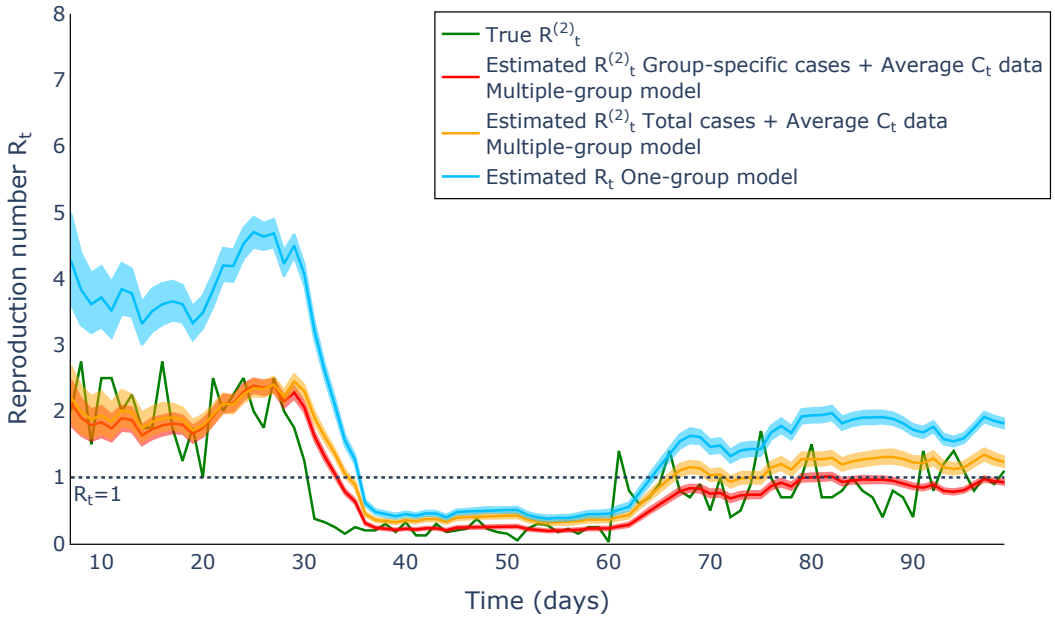
(b) Population group 2

Figure 3.2: *The means and* 95% *credible intervals of the inferred group-specific reproduction numbers for the multiple-group renewal equation process described in section 3.2.2 (MAGENTA) and the baseline one-group model (LIGHT BLUE). We used synthetic infection data generated from a two-group population over* 150 *days (Figure 3.1) using a two-step* γ_t *function of the probability of a contact becoming a case. The true group-specific (a)* $R_t^{(1)}$ *and (b)* $R_t^{(2)}$ *are drawn for reference (GREEN).*

assume no prior knowledge of those. The observation that this particular model does well at estimating the group-specific R_t implies that, regardless of the amount of information we



(a) Population group 1



(b) Population group 2

Figure 3.3: *The means and* 95% *credible intervals of the inferred group-specific reproduction numbers for the multiple-group renewal equation processes described in sections 3.2.3 (RED) and 3.2.4 (YELLOW) and the baseline one-group model (LIGHT BLUE). We used synthetic infection data generated from a two-group population over* 150 *days (Figure 3.1) using a two-step* γ_t *function of the probability of a contact becoming a case. The true group-specific (a)* $R_t^{(1)}$ *and (b)* $R_t^{(2)}$ *are drawn for reference (GREEN).*

have on the exact effective contact matrix each day, incorporating heterogeneous mixing patterns into our modelling framework improves over the standard one-group approach when

a multiple-group structured population is considered. However, it is important to acknowledge that in this particular scenario, the generation time distributions for the two population groups are identical, therefore, the values of the $\pi_t^{(i)}$ remain close to the prior mean for all times t – that is, $\pi_t^{(i)} \approx 0.5, \forall i, t$.

Moreover, analogous to Figure 3.2, the estimates of the time-dependent group-specific reproduction numbers plotted in Figure 3.3 suggest that allocating more resources into interventions targeted to the first, more transmissible population group will have a bigger impact on halting the epidemic growth.

3.3.2 Improved targetting of interventions using the group-specific reproduction number

In the following, we show that the multiple-group renewal model approach can be easily applied to determine the population-group priority in intervention scenarios when a limited number of individuals can benefit from them, e.g. vaccines. We demonstrate the significant impact that the choice of population groups to be targeted by interventions has in ensuring the efficacy of those interventions.

We begin by considering a synthetic population which considers three distinct age groups: $0-20$, $20-60$ and $60+$ years old. Similar to section 3.3.1, all model parameter values used in this analysis are (unless specified) randomly chosen and are used to demonstrate how we would theoretically go on to choose the best interventional scenario based on the group-specific susceptibility, generation time and contact patterns. For each age group we assume a gamma-distributed generation time interval with a mean of 5.3, 7 and 4.2 days, respectively, and identical standard deviations of 2.3 days. Additionally, we assume a different transmissibility of the pathogen in each population group: 1 for those below 20, 0.3 in the $20-60$ year old group, and 0.6 in those over 60 – hence, the transmissibility vector is given by $\underline{p} = (1, 0.3, 0.6)$. Individuals in this population interact at rates prescribed by the POLYMOD contact matrix for Japan [Prem et al., 2021] – transposed to give us the matrix C_t used in eq. (3.7):

$$C = \begin{pmatrix} 6.96 & 5.23 & 1.48 \\ 1.87 & 10.66 & 2.38 \\ 0.34 & 1.29 & 3.37 \end{pmatrix}.$$

Therefore, the transpose of the effective contact matrix of the three-group population system becomes

$$C^* = \begin{pmatrix} 6.96 & 1.57 & 0.9 \\ 1.87 & 3.2 & 1.43 \\ 0.34 & 0.39 & 2.02 \end{pmatrix}.$$

Finally, we fix our default probability of a contact becoming a case to $\gamma_t = 0.14$. We start our simulations with 2, 80 and respectively 13 infected individuals in each age group. We implement a series of ever stricter interventions that reduce the overall number of contacts in the targeted groups: first we consider a reduction of 25% in contacts, implemented at time 20, which further increased to 50% on day 40.

In Figure 3.4, we plot the average trajectory of the daily incidence of infection for three different interventional scenarios, and compare them to the number of infections predicted when no interventions are in place. These trajectories of epidemic progression were averaged over 100 different simulation replicates. Using the definition of the group-specific reproduction numbers given in section 3.2.2 – that is $R_t^{(i)} = \gamma_t(\sum_{j=1}^{N} C_t^{*(ji)})$, we conclude that the population groups carrying the largest transmission burden of the three are in order: $0-20$ years-old, followed by the $20-60$ age group, and finally by the elderly in the $60+$ age bracket (as the total average effective contacts in each of these groups are in order 9.17, 5.15, and respectively 4.35).

In the absence of any interventions, the respective group-specific reproduction numbers for each of the three age groups are: $R_t^{(0-20)} = 1.28$, $R_t^{(20-60)} = 0.72$ and $R_t^{(60+)} = 0.61$. Hence, targeting our interventional resources to the single population group in which the epidemic is predicted to grow, that is for which the group-specific reproduction is above one (in our case the $0-20$ year-olds), makes most sense in terms of effective application of available resources. In the top-right panel of Figure 3.4, we represent the scenario when we wrongly apply the interventions exclusively to the $20-60$ years old group. As stated above, this particular age group, despite having the largest number of net contacts on average, has a group-specific reproduction number below one at all times – implying that the epidemic in this particular demographic is slowing down; this is mainly because the pathogen has the lowest transmissibility in this particular age bracket. Therefore, reducing the transmission in this group does nothing to halt the growing epidemic trend, as infections will keep increasing in the $0-20$ year-olds group unchecked. By comparison, targeting our interventions

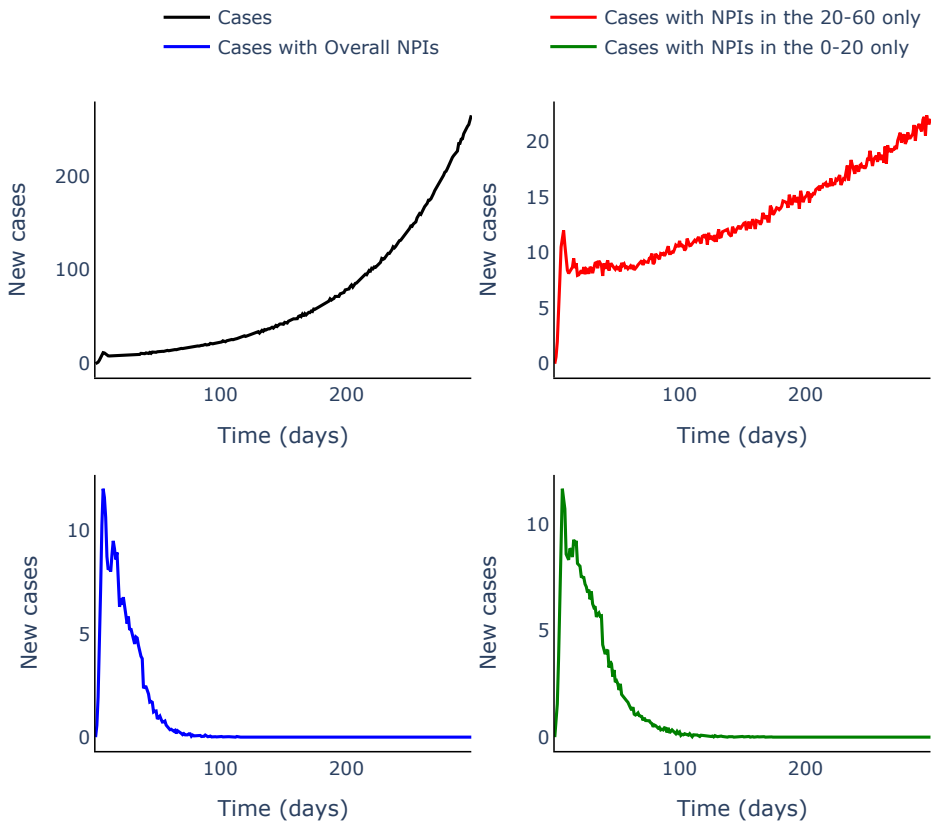


Figure 3.4: *Comparison of the total expected daily incidence of infection for three different interventional scenarios. The underlying infection data are generated using a three-group population and averaged over* 100 *simulations. Each strategy targets different age groups:* $20-60$ *only (RED),* $0-20$ *(GREEN) and all of the groups (BLUE). The effect of interventions is equivalent to an equal reduction in all contacts if an age group is targeted, as described in Table 3.1.*

exclusively on those in the $0-20$ age group (bottom right panel of Figure 3.4), we are able to achieve similar levels of reductions in the number of infections as we would if the same interventions were applied to all groups instead (bottom left panel of the same figure).

However, in practice we are interested in determining what is the least amount of restrictions that we need to impose in order to halt the epidemic, i.e. reach a values of the overall $R_t < 1$ – this is a key analysis needed to ensure the cost-effectiveness of any imposed package of interventions. In order to determine the ideal strength of these NPIs, we plot the changes in overall R_t of the synthetic epidemic from the top left panel of Figure 3.4 in terms of the value of $1 - r$ – the reduction in total contacts in a population group, as shown in Table 3.1.

Indeed, from Figure 3.5 we observe that to in order to get the overall population R_t to go below one, we require similar reductions of contact levels for both interventional scenarios

Type of interventions	Naive NPI effects
No NPIs	$\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$
Overall NPIs	$\begin{pmatrix} ra & rb & rc \\ rd & re & rf \\ rg & rh & ri \end{pmatrix}$
NPIs in the $0-20$ age group only	$\begin{pmatrix} ra & rb & rc \\ d & e & f \\ g & h & i \end{pmatrix}$
NPIs in the $20-60$ age group only	$\begin{pmatrix} a & b & c \\ rd & re & rf \\ g & h & i \end{pmatrix}$

Table 3.1: *The effect of the different interventions on the contact matrix.*

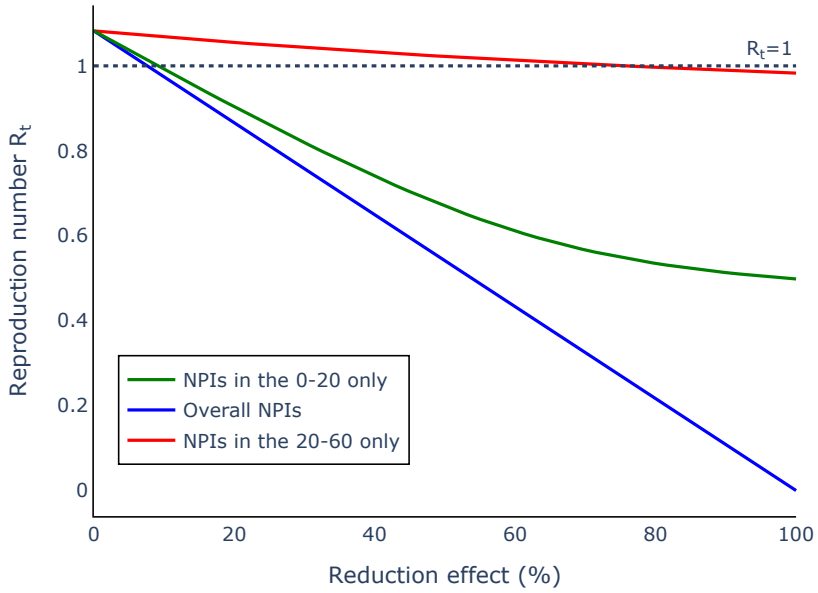


Figure 3.5: *Trajectories of the overall R_t dependency on the severity of imposed NPIs for three different interventional scenarios: $20-60$ only (RED), $0-20$ (GREEN) and all of the groups (BLUE). The effect of interventions is equivalent to an equal reduction in all contacts if an age group is targeted, as described in Table 3.1.*

–when we apply the NPIs to all age groups, and respectively, when we apply them to the $0-20$ year-olds only – around 7% and 9% respectively; this contrast to the interventional scenario in which we apply restrictions exclusively to the $20-60$ years-old, where we would

Type of interventions	**Simple shift NPIs**
No NPIs	$\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$
Overall NPIs	$\begin{pmatrix} ra & 0.95(1-r)a+b & 0.05(1-r)a+c \\ 0.7(1-r)e+d & re & 0.3(1-r)e+f \\ 0.3(1-r)i+g & 0.7(1-r)i+h & ri \end{pmatrix}$
NPIs in the $0-20$ age group only	$\begin{pmatrix} ra & 0.95(1-r)a+b & 0.05(1-r)a+c \\ d & e & f \\ g & h & i \end{pmatrix}$
NPIs in the $20-60$ age group only	$\begin{pmatrix} a & b & c \\ 0.7(1-r)e+d & re & 0.3(1-r)e+f \\ g & (1-0.25r)h & i \end{pmatrix}$

Table 3.2: *The effect of the different interventions on the contact matrix when interventions shift the contacts. The total number of contacts for each age group remains constant.*

require a reduction of over 75% to ensure a decreasing epidemic. Hence, this results suggests that by applying interventions of a similar strength to only the subset of the entire population that has highest effective number of contacts (and therefore group-specific R_t) we can save up considerable resources spent into implementing these NPIs while not compromising on the overall effect.

However, implementing interventions that are reducing all contact of an age group is not always realistic. In practice, reducing the number of contacts in one particular age group will lead to an increase of contact events with the other age groups – for example, implementing school closures, leads to a significant decrease in the number of contacts individuals in the $0-20$ age bracket will have with individuals in the same age group, but at the expense of increasing the number of average daily contacts with the adults (primarily their parents, as they are required to stay home), and a small increase in interactions with the elderly $60+$ (as they may spend more time with their grandparents). Therefore, a more realistic representation of the effect of interventions targeting a specific age demographic is represented by a *shift* of those contacts towards other age-groups. In Figures 3.6 and 3.7 we illustrate the reduction effect that these more realistic interventions will have on the predicted number of new cases. We analyse the same three scenarios as in Figure 3.4: when NPIs are applied

Type of interventions	**Complex shift NPIs**
No NPIs	$\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$
Overall NPIs	$\begin{pmatrix} ra & (1+0.25(1-r))b & (1-0.25r)c \\ (1+0.25(1-r))d & re & (1-0.25r)f \\ (1-0.25r)g & (1-0.25r)h & ri \end{pmatrix}$
NPIs in the $0-20$ age group only	$\begin{pmatrix} ra & (1+0.25(1-r))b & (1-0.25r)c \\ (1+0.25(1-r))d & e & f \\ (1-0.25r)g & h & i \end{pmatrix}$
NPIs in the $20-60$ age group only	$\begin{pmatrix} a & (1+0.25(1-r))b & c \\ (1+0.25(1-r))d & re & (1-0.25r)f \\ g & (1-0.25r)h & i \end{pmatrix}$

Table 3.3: *The effect of the different interventions on the contact matrix when interventions shift the contacts.*

to all population groups, when NPIs are targeted to those in the $20-60$ age group only, and when NPIs are targeted to those in the $0-20$ age group. In Tables 3.1, 3.2 and 3.3 we summarise the effects of the NPIs for all the interventional scenarios considered, where r denotes the strength effect of the chosen intervention.

What we observe is that even when we consider instead interventions that have an effect the contact patterns more closely resembling real NPIs, by targeting those interventions to the age group with the biggest group-specific reproduction number (that is the $0-20$ age group) – which corresponds to changes in the contact matrix as illustrated by the final rows in Tables 3.1, 3.2 and 3.3, we still observe the biggest reduction in the number of predicted new cases, of a similar effect as when applying those same interventions to all population groups – corresponding to changes in the contact matrix as illustrated by the second rows in Tables 3.1, 3.2 and 3.3. This is further confirmed by Figures 3.8 and 3.9, in which we observe similarly to before that the overall NPIs require similar severity levels to the NPIs targeted only in the $0-20$ age group – and therefore we can save resources overall by using targetted interventions instead; meanwhile, no matter how strong we make the NPIs targeted in the $20-60$ age group only, we can never achieve values of the overall R_t less than or equal to one, i.e. the epidemic will only keep increasing.

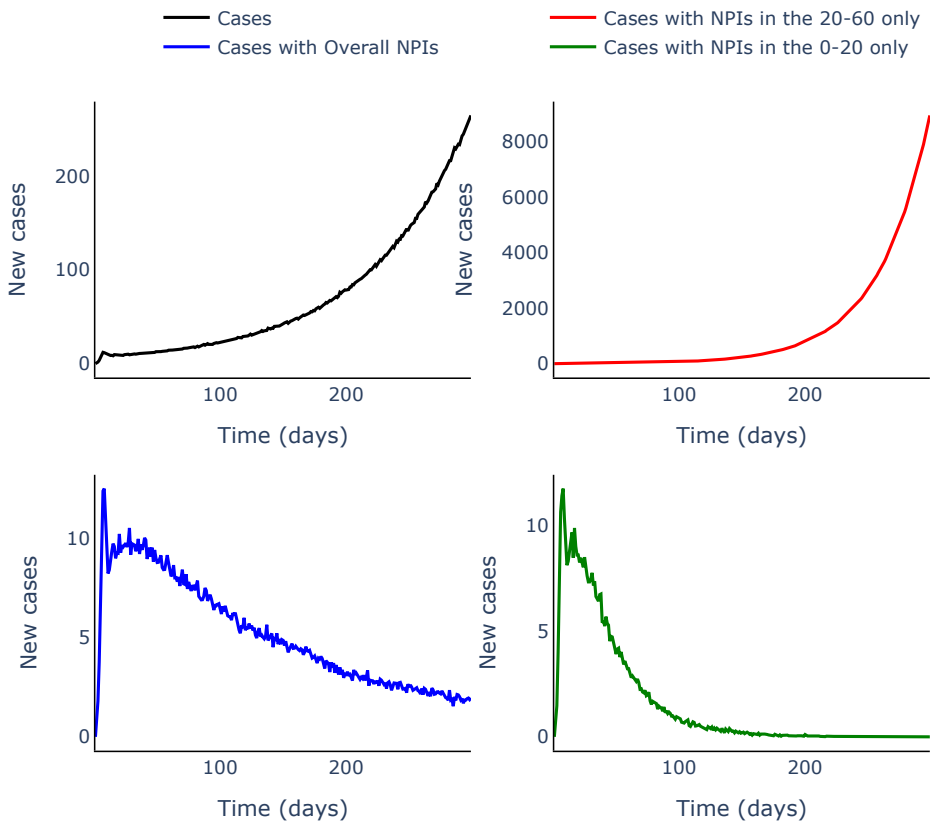


Figure 3.6: *Comparison of the total expected daily incidence of infection for three different interventional scenarios. The underlying infection data are generated using a three-group population and averaged over* 100 *simulations. Each strategy targets different age groups:* $20-60$ *only (RED),* $0-20$ *(GREEN) and all of the groups (BLUE). The effect of interventions is equivalent to a shift of the number of contacts from one group to another, as described in Table 3.2.*

The only time when we observe any differences in the ideal single-group intervention strategy is when the effective contact matrix is changed, either through the underlying contact patterns, as seen in Figure 3.10(a), where we use the UK POLYMOD contact matrix instead of the one for Japan [Prem et al., 2021]:

$$C = \begin{pmatrix} 6.96 & 5.23 & 1.48 \\ 1.87 & 10.66 & 2.38 \\ 0.34 & 1.29 & 3.37 \end{pmatrix}$$

(in this case all interventions lead to a declining epidemic as the overall R_t with no interventions is $R_t = 0.95 < 1$), or when the susceptibility to infection of the different population groups changes, as seen in Figure 3.10(b) where we use instead a susceptibility vector of

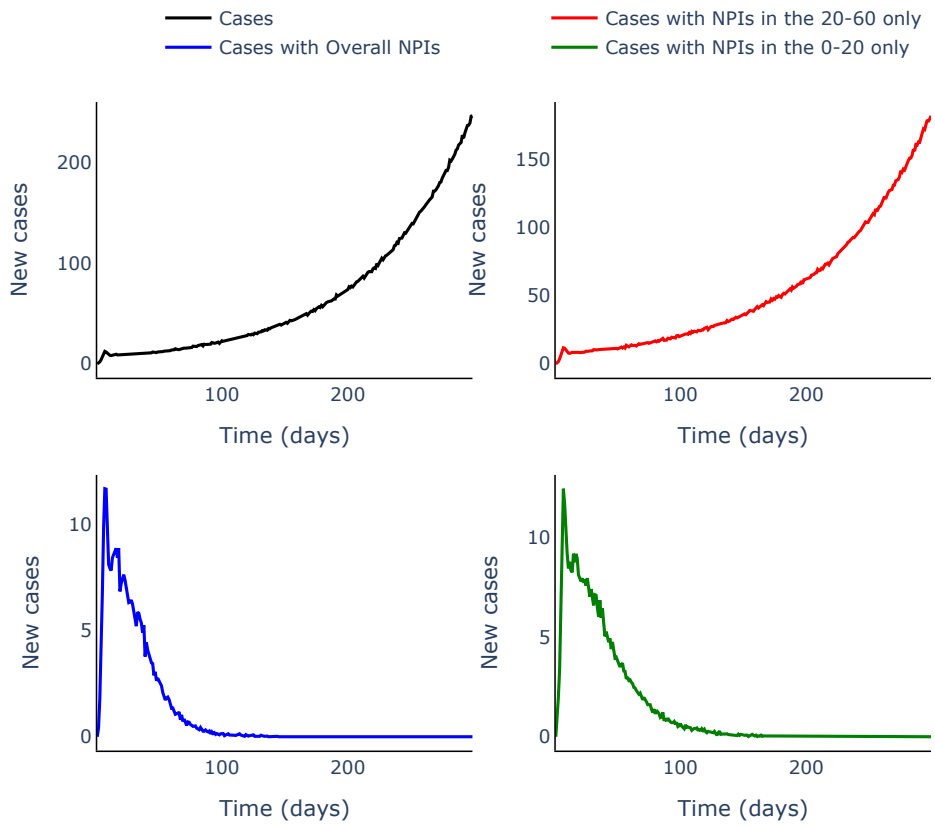


Figure 3.7: *Comparison of the total expected daily incidence of infection for three different interventional scenarios. The underlying infection data are generated using a three-group population and averaged over* 100 *simulations. Each strategy targets different age groups:* $20-60$ *only (RED),* $0-20$ *(GREEN) and all of the groups (BLUE). The effect of interventions is equivalent to a shift of the number of contacts from one group to another, as described in Table 3.3.*

$\underline{p} = (0.3, 0.6, 1)$ (in this case it is the middle age group $20 - 60$ that has the highest number of effective contacts and hence it is the group in which we should focus interventions: $R_t^{(0-20)} = 0.38$, $R_t^{(20-60)} = 1.44$ and $R_t^{(60+)} = 1.01$). For both scenarios depicted in Figure 3.10, we implement the naive contact reduction effect depicted by interventions, as depicted in Table 3.1.

Therefore, by carefully analysing the relative difference in the effective total number of contacts in the groups and therefore their baseline group-specific reproduction numbers, we are able to reliably determine both which population demographics are best suited for targeting interventions to, as well as to identify the level of stringency of those interventions required to successfully slow down and halt the epidemic.

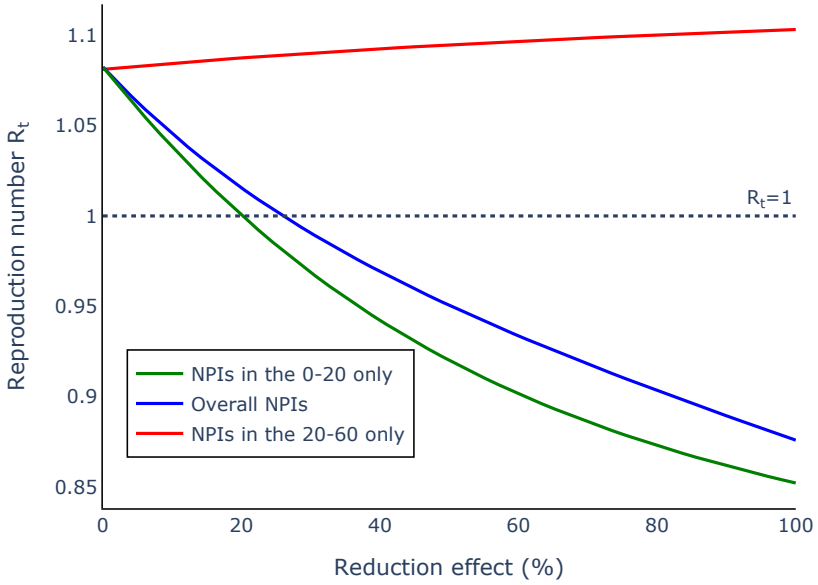


Figure 3.8: *Trajectories of the overall R_t dependency on the severity of imposed NPIs for three different interventional scenarios: $20-60$ only (RED), $0-20$ (GREEN) and all of the groups (BLUE). The effect of interventions is equivalent to a shift of the number of contacts from one group to another, as described in Table 3.2.*

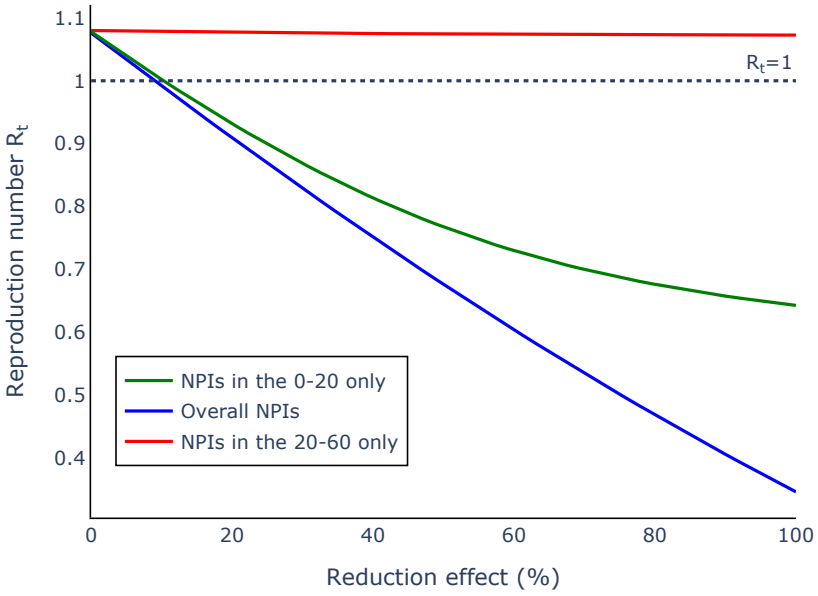
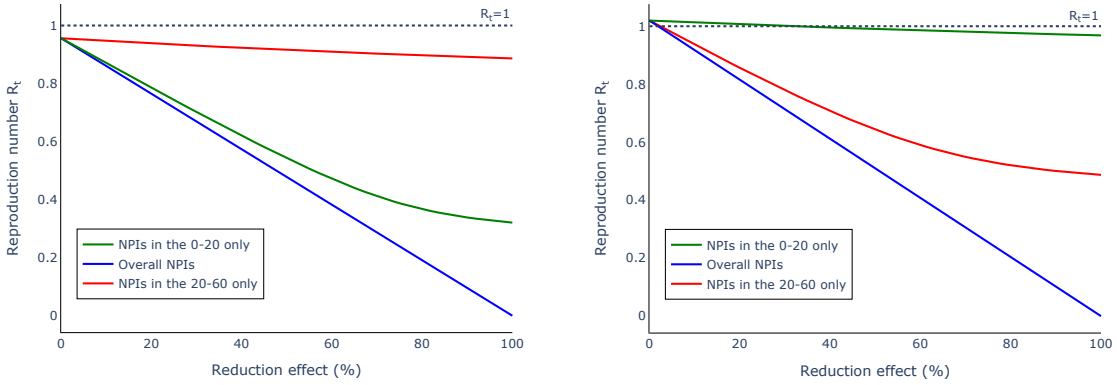


Figure 3.9: *Trajectories of the overall R_t dependency on the severity of imposed NPIs for three different interventional scenarios: $20-60$ only (RED), $0-20$ (GREEN) and all of the groups (BLUE). The effect of interventions is equivalent to a shift of the number of contacts from one group to another, as described in Table 3.3.*



(a) Different contact matrix

(b) Different susceptibility vector

Figure 3.10: *Trajectories of the overall R_t dependency on the severity of imposed NPIs for three different interventional scenarios: $20-60$ only (RED), $0-20$ (GREEN) and all of the groups (BLUE) for different epidemic conditions: (a) a different base contact matrix – in this case the UK POLYMOD contact matrix, and (b) a different vector of infection susceptibility. The effect of interventions is equivalent to a shift of the number of contacts from one group to another, as described in Table 3.1.*

3.4 Discussion

Successful implementation of interventions hinges on accurate predictions of epidemic dynamics in the long run – whether we need to impose stricter interventions if the epidemic is growing, or whether we can relax restrictions if the number of observed infections follows a declining trend. The time-dependent reproduction number is a convenient-to-use measure for the assessment of the long-term epidemic dynamics. The heterogeneous mixing patterns of individuals in a population and the differences in the transmissibility and the disease progression across demographic groups play an important role in shaping the epidemic trajectory; hence, it is important to account for these features in our modelling framework.

Here, we compared three renewal equation processes for multiple-group populations which assume different amounts of knowledge about the daily contact patterns of the different demographic groups and the number of infections observed in each group. For each of these models we presented analytical derivations for the group-specific reproduction number, as well as how these quantities relate to the overall reproduction number associated with each one of these models. We showed that, by having access to the daily variations in the contact rates across the different population groups, as well as to the daily numbers of new infections in each group, we were able to reconstruct the group-specific reproduction num-

ber trajectories with higher accuracy. Moreover, we determined that even in the absence of any knowledge of the exact time-dependent variations in contacts (but maintaining access to the average effective contact matrix), we are still able to reconstruct the group-specific epidemic behaviour if a multiple-group renewal model approach is utilised – which includes the scenario when only the total incidence count data is available and the daily fractions of infections in each population group need to be additionally inferred. These results support our conclusion that, in practice, it is better to include any knowledge of contact rates (and therefore use a multiple-group renewal model) than no knowledge at all when modelling a highly heterogeneous population.

Limitations on the amount of available resources required to implement epidemic mitigating interventions necessitate identifying an intelligent way of prioritising those resources for maximum impact. In this study, we also explored how applying targeted interventions to specific population groups leads to similar epidemic outcomes as when implementing these same measures in all groups. We looked into both naive interventions that reduce the contacts uniformly across the affected age groups, as well as more realistic interventions that simply lead to a partial reshuffling of those contacts – for example, when imposing school closure, we would not only expect a decrease in the number of contacts between children, but also a slight increase in household interactions. We determined that by focusing interventions on a population group carrying less of the epidemic burden – even if they have a higher number of contacts, we are unable to stop epidemic growth. This demonstrates that channelling a larger proportion of resources to suppress the epidemic in those groups carrying most of the transmission burden leads to better intervention outcomes. This also showcases the need to extensively study which population groups are responsible for the largest epidemic growth before deciding where to focus our interventional resources.

We also found that, by using the group-specific reproduction numbers, we can assess the epidemic trend in each population group and hence determine whether a proposed intervention should be applied to only a small proportion of individuals, responsible for most of the transmission, or if this intervention should be implemented indiscriminately across the entire population.

Our analyses however present a number of limitations. Firstly, we have considered only Poisson-distributed perturbations about the observed number of new infections. A possible avenue for future research would be to extend our findings to negative binomial-distributed

noise, and determine if the results still hold. This alternative choice of noise distribution would allow us to better account for the effect of super-spreaders in the transmission of an epidemic; this effect is incorporated through the overdispersion parameter of the negative binomial distribution.

We also assumed independent Poisson-distributed temporal variations in the exact effective matrix of total contacts (as defined in section 3.2.2) in the forward simulation. However, in reality, these variations are much more likely to be autocorrelated from one week to the next [Birrell et al., 2021]; we aim to explore the effect of this type of variation in the contact rates on the accuracy for our R_t estimates in future work.

Moreover, we implicitly assume that we possess reliable estimates for all the additional parameters that come into the structure of the multiple-group renewal equation model – that is group-specific generation time distributions, as well as reliable estimates of the effective contact matrix. However, the continuous collection of such detailed surveillance data is not commonplace. Firstly, estimates of the generation time distribution should account for the possibility that they can vary over the course the outbreak [Chen et al., 2022; Hart et al., 2022a,b; Parag et al., 2023]. Secondly, the way we collect the effective number of contacts is highly dependent on what we define as a contact – usually we scale the average number of interactions in a prescribed time interval by the contactee's probability of being susceptible to the virus (as discussed in section 3.2). In order to obtain a reliable estimate of the susceptibility vector, we would need to run clinical studies from the beginning of an epidemic – however, these types of studies are usually neglected early on in the epidemic, as most resources are instead focused into increasing detection capabilities. Moreover, it is not immediately trivial what is how refined these data need to be – for example how often contact data need to be collected over time and space and if these changes need to be collected daily, weekly or if estimates can reliably work over longer periods of time.

Additionally, we have not accounted for any differences in the reporting of cases observed associated with particular population groups. For example, children infected with SARS-CoV-2 express milder symptoms compared to older adults – therefore, cases in children are more frequently under-reported than in those occurring in older age groups. Also, the rate at which reporting occurs in any given population group can increase as the epidemic takes off very quickly, underwhelming the case reporting systems and increasing reporting delays. Therefore, not accounting for differences in reporting could lead to erroneous overall R_t

estimation values and introduce larger bias than by using the renewal equation model for a one-group population.

Other potential avenues for future exploration include extending the multiple-group renewal equation framework beyond human-to-human transmission dynamics (for example, in future work we could look into adapting the multiple-group renewal equation approach to model vector-borne disease as well) and incorporating migration patterns. At the moment, in our approach we assume that infections are generated exclusively by individuals in the same population. However, we have seen in [Creswell et al., 2022] that by failing to account for imported cases we lead bias the R_t estimates for the one-group renewal model; hence, it is sensible to assume we would observe be a similar effect when dealing with a highly heterogenous population.

In summary, in this study, we conducted a comprehensive analysis of the performance of three renewal equation models for multiple-group populations with different levels of knowledge of the contact matrix, which we then compare against the more traditional one-group population renewal model. We also showed how using the group-specific R_t can be used to identify the population groups which are the biggest drivers of epidemic increase and therefore the prime candidates for targeting interventions aimed at reducing disease transmission in a sustainable manner.

In the first three chapters, we investigated how explicitly accounting for heterogeneities in the population transmission affects our estimates of the reproduction number R_t, in the context of renewal equation models. As stated back in Chapter 2, renewal models are a popular tool for estimating trajectories of R_t, as they only require a small number of assumptions in their setup; however, it is precisely this limited setup that prevents us from accounting for more specific features of transmission, or from analysing the impact of highly targeted interventions. In the remainder of this thesis, we will explore a second widely-used class of epidemiological models called compartmental models – which provide a more flexible framework for incorporating these more precise dynamics and interventions, and we will determine whether we are able to reach similar predictions of epidemic outcomes when compartmental models with different structures are fitted to the same observed datasets.

4 | A retrospective analysis of the robustness of existing compartmental ODE models for modelling future pandemics

In the first three chapters of the thesis, we explored a novel approach for incorporating transmission heterogeneities in the structure of the renewal models – an increasingly popular tool for estimating the reproduction number R_t of an epidemic, which requires only a small number of assumptions to set up. Its streamlined format makes the renewal equation model an ideal candidate for modelling the epidemic dynamics and for evaluating the severity of epidemic outcomes – especially in the early stages of an epidemic of a novel pathogen, when precise transmission features may not be known. However, for better-studied diseases, for which we have a better understanding of the way transmission occurs across the different population groups, the simplicity of the renewal equation structure precludes the integration of these precise epidemic dynamics.

In this final chapter and in the associated appendices, we will instead investigate a different class of epidemiological models called compartmental models which allow for a greater flexibility in the model structure, and are thus able to accommodate more specific dynamics of epidemic transmission and to implement more precisely targeted interventions. We will explore how the modelling assumptions underpinning compartmental ODE models impact the predicted epidemic outcomes and, therefore, our conclusions on the effectiveness of different intervention scenarios. In doing so, we aim to identify whether these compartmental models can be used interchangeably to test transmission dynamics and whether they produce similar epidemic forecasts of the main quantities of interest (such as the daily timelines of infections or deaths or predicted trajectories of the time-dependent reproduction number) for the same non-pharmaceutical interventions, e.g. lock-downs. This constitutes Chapter 4 of the thesis and is currently being edited in order to be submitted to the *BMC Global and Public Health* special issue. CRediT author contributions for this work are as follows. Conceptualisation: All authors. Methodology: IB, BL. Investigation: IB. Writing – original draft: IB, DG, BL, RT. Writing – review and editing: All authors. Supervision: DG, BL, RT, ALD.

Summary

During the early days of the COVID-19 pandemic, the UK government made use of a number of mathematical models of transmission dynamics to help to guide policy response. Several of these epidemiological models use compartments to sort the population, and ODEs to describe the infection dynamics. However, these models rely on a number of modelling assumptions about the disease, which sacrifice accuracy for model tractability. These differences in turn impact the forecasts of the epidemic trajectory and may lead to inconsistent recommendations to policy makers. In this study, we conduct a retrospective analysis of the performance of three models used for modelling the rapid progression of the COVID-19 pandemic in the UK to test the robustness of their results and whether we can these interchangeably to inform policy response: the PHE model, the Roche model and the Warwick-Household model. For each model, we assess the structural identifiability of their parameters, perform sensitivity analyses and fit these models to the same early 2020 UK epidemic death dataset. We then generate predicted trajectories for alternative interventional scenarios such as when no non-pharmaceutical interventions are applied or the timeline of application is changed. We identified that each of the three considered models produced very different death and case trajectories in these counterpart scenarios, which suggest that we cannot substitute the conclusions of any one of these models for another. These results highlight the pitfalls of relying on single models to inform policy response in the context of future epidemics, as well as the need of a more in-depth study of the impact of modelling assumptions of the quality of model outputs.

4.1 Introduction

As especially showcased during the 2020 COVID-19 pandemic, mathematical models of infection disease dynamics can play a vital role in shaping the public and governmental response to the epidemic outbreak and therefore in determining the efficacy of the interventions implemented to mitigate transmission. However, there are many ways in which epidemiological models can be derived based on a wide range of possible modelling assumptions.

Firstly, there are many different types of model formalisms to choose from, based on the

way in which the modeller chooses to represent the population – some examples include: (1) compartmental ODE models in which individuals in the population are grouped into compartments based on certain features such as age, region or infectious status [Birrell et al., 2021; Lemenuel-Diot et al., 2020; Moore et al., 2022; Bouros et al., 2024], (2) renewal equation models which only consider daily infection incidence, place no restrictions on the total population size and disregards the fate of recovered individuals [Thompson et al., 2019; Cori et al., 2013; Fraser, 2007; Bouros et al., 2025] and (3) spatial agent-based models which keep track of every individual in the population, parametrising each interaction between individuals separately [Gallagher et al., 2024; Ferguson et al., 2020].

Additionally, there are also differences in the type of data sources that the modeller can choose to use to parameterise these models; the most common choices are to use death and case incidence data such as those collected by the Office of National Statistics [GOV.UK] or, in order to account for the under-reporting of infections that some of these data sources suffer from, the modeller can choose to use infection surveys like REACT [Nicholson et al., 2021] to get better approximations of current levels of infection within the considered population. Moreover, if the population of interest is structured based on certain characteristics such as age, gender or default activity, then contact matrices evaluated from extensive social studies such as [Prem et al., 2021] can used to summarise the intricate mixing patterns between these groups.

Another source of differences in model setup arise from the assumptions we make about the most important features of disease transmission – for example if we model a pathogen that affects individuals of different ages with varying levels of severity or the population mixes differently across these age groups, then assuming an age-structured population would be a sensible assumption to make and take account of in the model structure [Bouros et al., 2024; Keeling et al., 2021; Birrell et al., 2021; Bouros et al., 2025; Anderson et al., 2020]. Similarly, if asymptomatic transmission is common for the modelled pathogen, as was the case with SARS-CoV-2 [Keeling et al., 2021; Lemenuel-Diot et al., 2020; Moore et al., 2022], then including this feature of disease dynamics in the model would be warranted. Moreover, especially for compartmental ODE and agent-based models, the modeller has much flexibility in differentiating between the different types of infected individuals and how these different categories impact the epidemic dynamics: for example [Lemenuel-Diot et al., 2020] differentiates between the super-spreader and non-super-spreader transmission,

while [Keeling et al., 2021] differentiates between the different types of infectious individuals based on the order of infections within the household they belong to, e.g. first infected, subsequent, etc. However, both of these models consider specific compartments for infectious individuals that are quarantined in order to minimise the spread of the pathogen.

Finally, models can also differ by the types of outputs of interest they can produce predictions for; while most epidemiological models of infectious diseases dynamics are capable of producing predictions of the daily number of new infections [Thompson et al., 2019; Birrell et al., 2021; Cori et al., 2013; van der Vegt et al., 2022], some of them can also be used to forecast more severe disease outcomes such as deaths [Ferguson et al., 2020; Lemenuel-Diot et al., 2020; Birrell et al., 2021; Gallagher et al., 2024; Ferguson et al., 2020; van der Vegt et al., 2022], hospitalisation and ICU occupancy levels [Keeling et al., 2021; Moore et al., 2022; Gallagher et al., 2024; Ferguson et al., 2020; Bosetti et al., 2021], or to indicate the population level of protection against the disease through vaccination [Moore et al., 2022; Bouros et al., 2024; van der Vegt et al., 2022; Bosetti et al., 2021; Hogan et al., 2021].

This plethora of options makes the modeller's task of selecting the ideal approach to modelling the disease dynamics of a specific outbreak non trivial. In this study we explore how the modelling assumptions underpinning compartmental ODE models impact the predicted epidemic outcomes and, therefore, our conclusions on the effectiveness of different intervention scenarios. In doing so, we aim to identify whether these compartmental models can be used interchangeably to test transmission dynamics and whether they produce similar epidemic forecasts of the main quantities of interest (such as the daily timelines of infections or deaths or predicted trajectories of the time-dependent reproduction number) for the same non-pharmaceutical interventions, e.g. lock-downs.

To study this problem, we focus exclusively on models for the population transmission of the SARS-CoV-2 virus, but these results can be easily extended to other air-borne diseases. We consider a beginning-of-epidemic case study, where only non-pharmaceutical interventions (NPIs) are available and there is a delay in their implementation – similar to what happened at the beginning of the COVID-19 pandemic in early 2020; in fact, throughout our analyses we will concentrate exclusively on the first wave of the COVID-19 epidemic in the United Kingdom – between 15th of February 2020 and 25th of June 2020. During this time in the UK, the government relied on scientific advice from the SPI-M (UK's Scientific Pandemic Influenza Group on Modelling) consortium to determine which NPIs were most effective in

mitigating the epidemic transmission, as well as their timeline of implementation.

The majority of the models used – which were either specifically built or repurposed pre-existing Influenza transmission models developed independently from one another, have a compartmental model structure, illustrating their creators' original assumptions about which population groups play a key role or are the biggest drivers of epidemic growth. Another important stakeholder in the goal of understanding COVID-19 transmission were pharmaceutical companies e.g. Roche which use these transmission models to assess the need for anti-viral treatments and their effect in diminishing the size of the epidemic. In the following we focus on only three compartmental ODE models: the PHE [Birrell et al., 2021], the Roche [Lemenuel-Diot et al., 2020] and the Warwick-Household models [Keeling et al., 2021] for each of which we now give a brief introduction to the main transmission features.

4.2 Methods

4.2.1 Three models for the COVID-19 epidemic: the PHE, Roche and Warwick-Household models

We begin this chapter by introducing three ordinary differential equation-based (ODE) compartmental models used to analyse the first-wave of the COVID-19 epidemic in the UK. Two of these models – the PHE model [Birrell et al., 2021] and the Warwick-Household model [Keeling et al., 2021], were developed and used by contributors to the Scientific Pandemic Infections Group on Modelling (SPI-M) advisory committee – a consortium of researchers that advise the UK government on its response to the COVID-19 pandemic, to inform public policy, e.g. estimating the reproduction number, forecasting hospitalisations and deaths. The third model that we considered, the Roche model [Lemenuel-Diot et al., 2020], was initially developed by our industrial partner at Roche[1] to study the transmission dynamics of the Influenza virus in France and subsequently retrofitted to allow for modelling SARS-CoV-2-like pathogens.

We have chosen to refit these three specific models as all of them share similar features (all three models consider an age- and region-structured population with no mixing between

[1]Roche here refers to F. Hoffmann-La Roche Ltd pharmaceutical company who generously sponsored and supported this project.

regions and extend the SEIR model structure that we introduced in section 0.2.1) but also have features that set them apart:

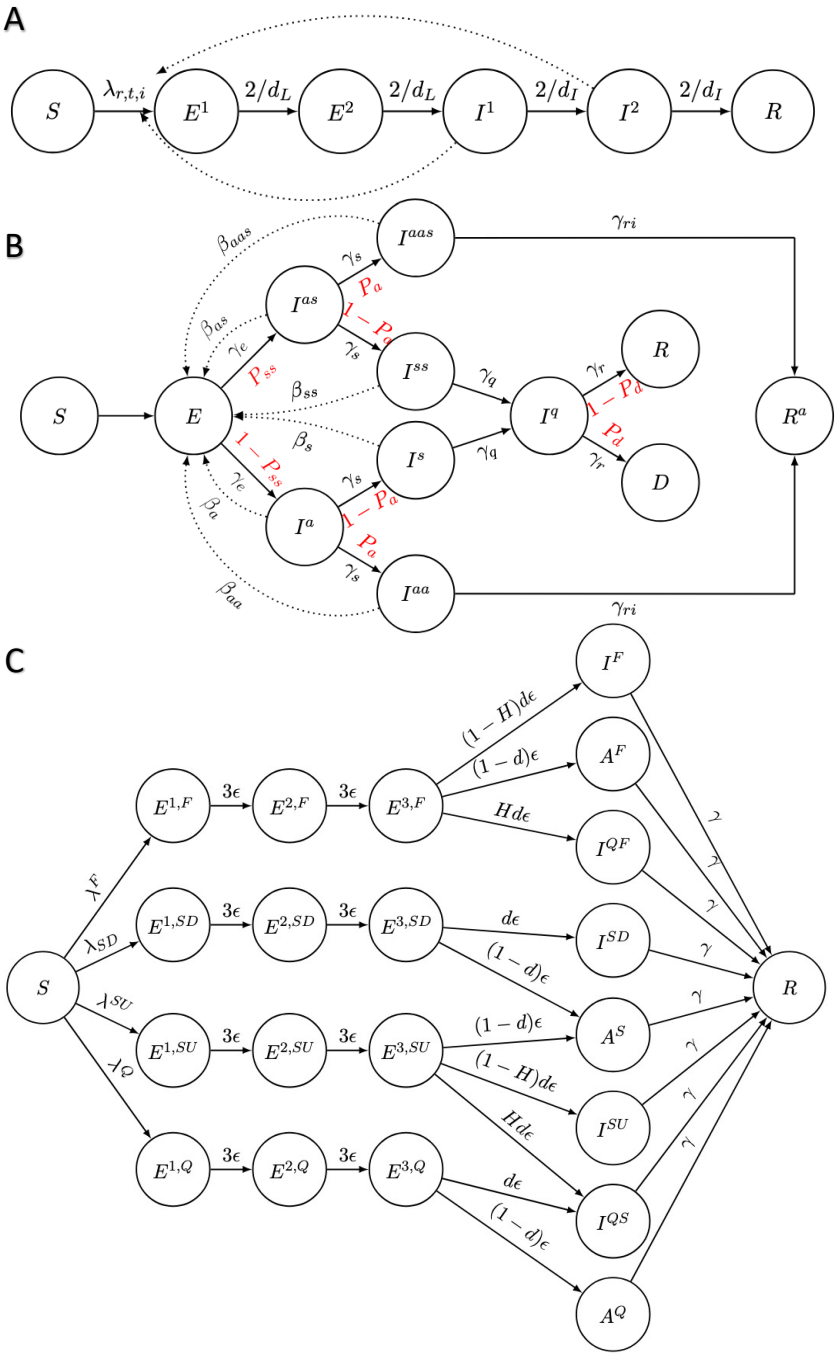


Figure 4.1: Comparison of the epidemic dynamics of the PHE model (panel A), the Roche model (panel B) and the Warwick-Household model respectively (panel C).

1. The **PHE model**: this model assumes there are no asymptomatic infections, and considers two compartments for each of the exposed (E) and infectious (I) infection

states. The effect of non-pharmaceutical interventions is modelled through a time-varying contact matrix computed using POLYMOD [Prem et al., 2021] and Google Mobility data [Google]. Severe outcomes such as deaths are modelled using a separate stochastic process which takes as input the number of new infections.

2. The **Roche model**: unlike the PHE model, the Roche model includes presymptomatic, asymptomatic, super-spreader and quarantined infectious compartments, as well as a compartment for the number of deaths. Non-pharmaceutical interventions are integrated into the model through a stringency index which alters the transmission parameter directly. The stringency index is computed using OxCGRT data [Hale et al., 2021].

3. The **Warwick-Household model**: similarly to the Roche model, this model differentiates between symptomatic and asymptomatic transmission. However, it also takes into account within-household dynamics, with separate compartments for those first infected within the household, subsequent symptomatic and asymptomatic infections, as well as those infections that are quarantined. Severe outcomes are modelled using a separate process, as in the case of the PHE model, with the number of hospitalisations and deaths computed using daily numbers of new symptomatic infections.

In Figure 4.1 we show the significant visual contrast between the compartmental structure of each of these three models. In Appendix C, we provide an even more detailed description of these models as well as showing how interventions are represented in each framework, highlighting the differences in the setup of the PHE, the Roche and Warwick-Household models. In addition, in Appendix C we also present example implementations of a COVID-like epidemic for the three models, using the same parameters values as for the first wave of the epidemic in the UK.

All three models were coded by the original developers in distinct programming languages: C++ for the PHE model, Matlab for the Roche model and Matlab and Julia for the Warwick-Household model. However, to assess their reproducibility, to ensure equal quality of coding, and to allow for model comparison using identical testing conditions, we decided to reimplement these three models in a consistent common framework. Hence, as part of this project, we have developed a Python software package called 'epimodels' [Bouros, 2021] dedicated to systematic compartmental modelling, with a modularised structure and tem-

plate model files which allow users to quickly and efficiently implement new epidemiological models of disease transmission. We have provided more details about the overall structure of 'epimodels' in section 4.2.2.1. The package also provides an easily adjustable parameter inference infrastructure using PINTS [Clerx et al., 2019] which we will discuss further in section 4.3.4 of this chapter.

4.2.2 Software package

As mentioned in the Introduction, most of the epidemiological models previously used by the UK government to study the behaviour of the COVID-19 epidemic suffer from multiple inconsistencies in implementation. These inconsistencies extend to more than just the fact that the models are built on different assumptions, e.g. different compartmentalisations of the population or an altogether different approach to modelling, e.g. a deterministic compartmental model like the PHE [Birrell et al., 2021], Roche [Lemenuel-Diot et al., 2020] or Warwick-Household model [Keeling et al., 2021] versus a stochastic agent-based one as is the case with the Ferguson model [Ferguson et al., 2020]: these models are often coded in different programming languages and very often lack extensive code documentation.

Moreover, a researcher interested in reusing any of these models has to rely on the original paper to understand the code and recode the model from first principles as very often no open-source code is available for the model. This results in increasing time delays in the research process caused by the need to reimplement the same model over and over again as well as possibly introducing errors in the model structure if certain features of the original model are not properly adapted into the newer format.

In the same vein, extending a pre-existing model to include more features becomes very often a challenge without an intrinsic knowledge of the original model, especially if the code contains highly optimised functions with no documentation on their functionality within the code as a whole.

In the following sections, we provide an overview of the main software package we have used to run our model analyses: 'epimodels' [Bouros, 2021], which we developed and will use throughout Chapter 4 to run our compartmental ODE model comparisons.

4.2.2.1 epimodels

The need for a properly documented software infrastructure for epidemiological modelling spurred the creation of 'epimodels' [Bouros, 2021]. As mentioned previously, 'epimodels' is a Python package, created as part of this DPhil project and intended as its main software output; its main aim is to be used as a library of epidemiological models covering an extensive number of features relevant for the disease modelled.

4.2.2.2 Software Properties

In what follows we illustrate some of the key software engineering guidelines for the sustainable development of 'epimodels', by addressing each of the previously raised issues and how we addressed it:

- **Coding language consistency**: All models in the suite are coded in Python. Currently all versions of Python 3.10+ are supported.
- **Code documentation**: We implemented automatically generated API documentation available online, with all methods, including the private ones, fully documented in the source code.
- **Code correctness**: All class files included in 'epimodels' are fully unit tested with 100% coverage on the **main** branch.
- **Code reusability**: In order to allow flexibility in the construction of new models and in the better understanding of the ones already included, the code is highly modularised into many classes. Therefore, constructing new models from scratch using already coded-up features becomes straightforward, with either already existing templates or methods that can be used directly.
- **Code interpretability**: We include example notebooks to illustrate each function of the models included in the Python package. The user can then understand how to use the models with both synthetic and real curated data from the early UK COVID-19 epidemic.
- **Code usefulness**: We include scripts for each model in which we run optimisation, inference and sensitivity analyses using UK death [GOV.UK] and serology and large-scale infection survey data [Nicholson et al., 2021].

4.2.2.3 Software Architecture

In Figure 4.2 we present the overall architecture of the software. More details about the structure of the 'epimodels' package [Bouros, 2021] are publicly available on GitHub at the following link:

https://github.com/I-Bouros/multi-epi-model-cross-analysis.

Details about installing the package and running some of the basic functionalities of the package are indicated in the README file. We also provide direct links (in blue) throughout Appendix C to all Python notebooks, which include the code we run for the various analyses performed in this part of the thesis.

The main pillars are constituted by the model classes. Each of these model classes is constructed using the same template. This template is based upon the *ForwardModel* class in PINTS [Clerx et al., 2019], which is our standard for good software engineering practices. Some of the models included have different transmission dynamics for different regions and age groups. All model classes include methods for adding, altering and retrieving this type of data. Age-dependent compartmental models often require additional information on the way different age groups interact in the form of *contact matrices*, which summarise these mixing patterns as such: the $(i, j)^{\text{th}}$ element of the contact matrix is equal to the average number of contacts in age group j someone in age group i will have over a prescribed time period. We handle contact matrices in a different file – to which we will return later. However, all model classes include four main type of methods:

1. **Simulate:** This method takes as arguments a vector of the times for which we wish to run the model and the parameters of the model fed in the form of a *ParameterController*. It outputs an age-dependent (column-wise) matrix of the number of individuals in each compartment at each time point (row-wise) for the specified region.

2. **New Infections**: This method extracts the age-dependent daily number of new infections given an output matrix of the *simulate* method for the model. The output can than be used as an argument in the computation of the other featured measures which we will discuss below.

3. **Positive Tests & Deaths**: As in the parameter inference the main sources of information are death and infection survey data (i.e. number of positive test results), we would

like to be able to produce such synthetic data from a known model or quantify the likeliness of observing the data under the current parameterisation. Therefore, there are three different methods associated with each of these problems: *log-likelihood*, *mean* computation and *sampling*.

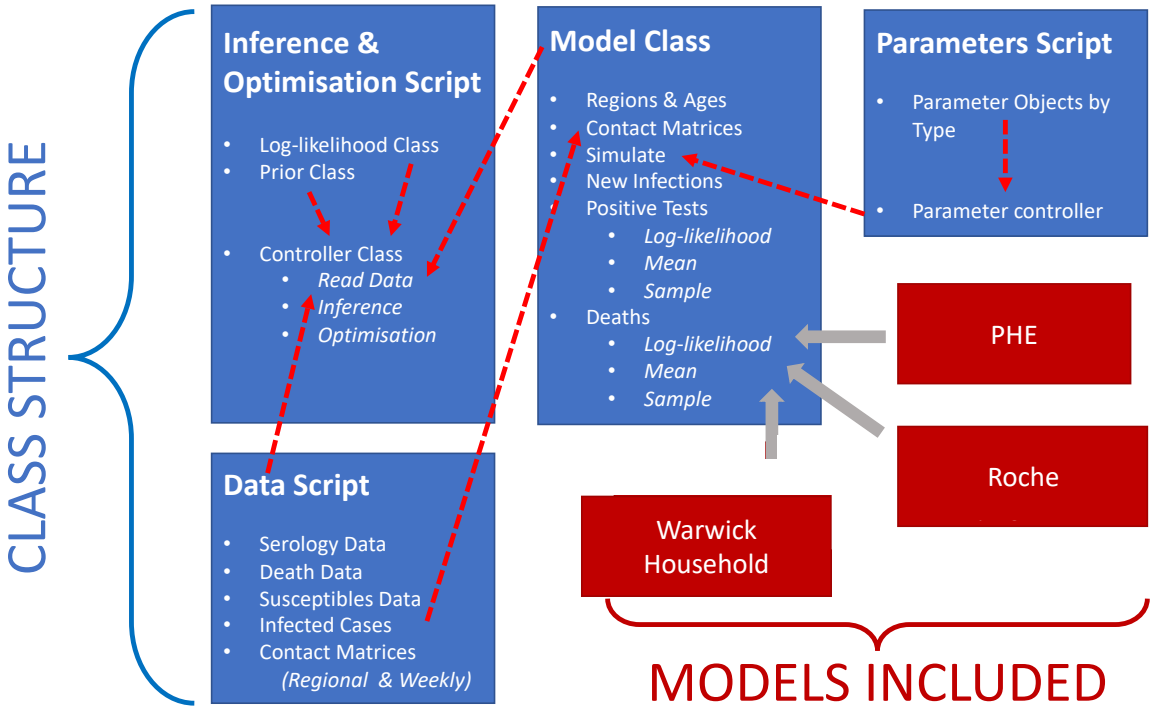


Figure 4.2: *Architecture of* 'epimodels'. *In red we include the names of all the models we have coded up as part of the package, while in blue we represent the different functionality of the files. The red arrows illustrate how each class or method interacts with one another.*

In the _setup_matrices.py file we include other auxiliary classes used in the construction of each of the models. For example, the PHE and Warwick-Household models allow for regional- and time-dependent contact matrices in its setup. Similarly, we will see in section C2 that we have extended the Roche model to include this kind of matrix as well. The object classes responsible for the creation and any alterations of these parts of the model are all included in this script.

The data which are fed into the contact matrices are parsed from *csv* files. These data have been mined and adapted to fit the required format via Python files in the so-called *Data Script* part of the package. Apart from contact data, we include files that contain the raw, the processed data and the parsing files for the age-dependent number of daily deaths, daily positive number of tests and total number of tests conducted in the study, i.e. infection survey

data, number of susceptibles in each PHE region of England by age categories and similarly the number of initial cases.

As mentioned previously, the parameters are fed into the model through a parameter controller class. It gathers several parameter object classes, which are named based on the role each of their own features play in the simulation. We do this to provide better management for the user and easier access to the parameters they wish to alter.

The other main type of scripts which comprise 'epimodels' are those for the parameter inference and optimisation. For each model included in the package, we have a distinct Python file which consists of three classes:

1. **Controller Class**: This is the base class used in the parameter inference or optimisation of any of the models included in the package. There are three key types of methods linked to this class: (1) reading the data – this ranges from the model for which we wish to run the inference/optimisation to the infection survey and death data (from the data script files) used to compute the log-likelihood (see section 0.4 for details); (2) running the optimisation and (3) the inference routines (see section 0.5.2 for details). Both of these last two methods use ready-made optimisers and samplers from PINTS [Clerx et al., 2019]. Therefore, in order to be able to perform our optimisation or inference, we use specific type objects for the log-likelihood and log-prior which follow the PINTS formatting.

2. **Log-likelihood Class**: This class follows, as we have mentioned above, the structure of a PINTS *LogLikelihood* class. It includes methods for reading all the necessary data for the evaluation of the model at a given set of values of the parameters. It returns the value of the total log-likelihood. When instantiating the class we provide two values for the weight which we put on the two main sources of data (the number of deaths and the number of positive tests) in the computation.

3. **Prior Class**: In this class, we compute the value of the log-prior of the model (see section 0.4 for details) and similarly to the previous point, this class follows the structure of a PINTS *LogPrior* class.

The overall structure of the package is a complex one, with many classes depending on others in different modules to work efficiently. As we mentioned previously, all functional

scripts and classes have an associated unit test in the folder 'tests' of the main 'epimodels' submodule. We also include in the folder 'examples' different files that showcase how to run and combine together the different functionalities of the package in order to answer specific questions e.g. "How do we run parameter inference for the PHE model in 'epimodels'?" or "How can we run an instantiation of the Roche model for a given set of non-pharmaceutical interventions?"

4.3 Results

We now go on to study the structural identifiability of these three models for a range of possible input datasets: prevalence data, case incidence data and death data – the last of which we use to fit the models in section 4.3.8.

We then perform parameter sensitivity analyses on the three models. Our goal is to identify those parameters which when their value is varied produce drastic changes in the output of the model. We restrict our simulation window to 15th of February to 15th of May 2020. This is done in order to avoid accounting for the waning immunity against the virus for the recovered individuals in the population as three months is less than the smallest predicted average time to reinfection [Leidi et al., 2021; Abu-Raddad et al., 2021; Gallais et al., 2021].

We finish with detailed descriptions of the optimisation and MCMC sampling algorithms used to infer the model parameters we established as globally identifiable, as a result of the model structural identifiability analysis, for each of the three models from Appendix C. We fit the models to UK death data from the Office of National Statistics [GOV.UK] from the 15th of February 2020 to 25th of June 2020.

We then use these parameter estimates to run alternative NPI intervention scenarios – including if less harsh interventions were imposed. We are interested in determining whether the predicted trajectories for the PHE [Birrell et al., 2021], the Roche [Lemenuel-Diot et al., 2020] and the Warwick-Household [Keeling et al., 2021] models still align when the underlying context of the epidemic is changed from the one the models were fitted to. Otherwise, we can conclude there exist inconsistencies in the policy-advice conclusions that can be drawn using these compartmental models, e.g. the scale of the effect or the peak of the infection and the death curve are heavily reliant on the model structure.

4.3.1 Model equality and equivalence

As discussed in section 4.2.1 and in Appendix C, the PHE [Birrell et al., 2021], the Roche [Lemenuel-Diot et al., 2020] and the Warwick-Household [Keeling et al., 2021] models display a number of common compartments, with each of them extending the SEIR model structure to accommodate a specific set of disease assumptions. In Appendix D, we explore whether we can find combinations of parameters for which each of these three models simplify to the basic SEIR model:

$$\begin{aligned}
\frac{dS(t)}{dt} &= -\beta I(t)S(t) \\
\frac{dE(t)}{dt} &= \beta I(t)S(t) - \kappa E(t) \\
\frac{dI(t)}{dt} &= \kappa E(t) - \gamma I(t) \\
\frac{dR(t)}{dt} &= \gamma I(t).
\end{aligned} \tag{4.1}$$

This will allow us to identify parameter equivalences across these three models and to understand under which conditions the PHE, Roche and Warwick-Household models can be substituted for one another when modelling an epidemic. We first note that the basic SEIR model lacks the age-structure or the regional aspect of the PHE, Roche and Warwick-Household models. Therefore, for these analyses we consider only **1 region** and **1 age group** for each of the three models discussed.

4.3.2 Structural Identifiability

Most of the compartmental ODE systems used in epidemiological modelling include a large number of parameters which regulate either the transmission dynamics of the virus, or the transition of individuals between the different population compartments. In general, but especially in the case of novel or little-studied viruses, we do not know the physiologically- or demographically-correct values of these parameters; indeed, usually we only have access to certain epidemic outputs, such as number of cases, deaths or hospitalisations, in a given time interval. While for some of these parameters (such as the infection period), we can use clinical studies to arrive at realistic parameter values, in the vast majority of cases this is not possible; instead, we use output data to inform us which parameter regimes result in model outputs that closely follow the data observed. Solving the inverse model problem is

non-trivial for most ODE models; there are limited data resources available to help inform a suitable model parametrisation compared to the number of free parameters we need to infer. This underpins the question of how many of these parameter values we can reliably identify for different amounts of output knowledge.

We refer to a model parameter p as being globally identifiable if for any given input dataset there exists a globally unique choice of p which generates the same model output as the inputted one, when all other model parameters are kept fixed. Similarly, a parameter p is considered to be only locally identifiable if there is a unique value for p that generates the same model output in a neighbourhood $V(\underline{p})$ of the parameter space as the inputted one. A neighbourhood $V(\underline{p})$ of a point $\underline{p}$ is defined as the subset V of the space that contains an open subset U, which contains $\underline{p}$; that is, $\underline{p} \in U \subseteq V$ – see [Willard, 2004, Chapter 2.4] for more details. Otherwise, if there exist multiple values for p in any neighbourhood $V(\underline{p})$ of the parameter space, then the parameter p is structurally non-identifiable. Additionally, a model is non-identifiable if one or more of its parameters are non-identifiable.

Previous studies [Dankwa et al., 2022] demonstrated that for the SIR model and its simple extensions, which include exposed individuals (the SEIR model) and asymptomatic infections (the SEAIR model) respectively, the type of input information used to inform the model parameterisation (that is, whether incidence, death and/or prevalence data) has an impact on which parameter combinations are globally identifiable, only locally identifiable or non-identifiable altogether. Since structural identifiability is a necessary but insufficient condition for practical identifiability – as described in [Cobelli and DiStefano, 1980; Dankwa et al., 2022], we will restrict our search of which parameters to infer in each of the PHE, Roche and Warwick-Household models to those parameters which are found to be globally identifiable when running the SIAN structural identifiability algorithm [Hong et al., 2019, 2020].

This algorithm uses both differential algebra and Taylor series to transform the initial ODE system into a single polynomial expression in terms of known input quantities (the observed data); the quantifiers of the polynomial factors are given as expressions of the original model parameters. As the model outputs of interest are the roots of the polynomial and are known, the polynomial, and hence the quantifiers, are unique; this property is then used to determine which parameter combinations of the original ODE model are unique given the data (i.e. are globally identifiable) or are dependent on other parameters (e.g. the sum of two parameters

is uniquely defined for any given input data, but each of the parameters is non-identifiable, as they are linearly dependent on one another). The SIAN algorithm is implemented in Maple with an open-source web application[2]. However, due to server memory limitations caused by the complexity of the models we analysed, we instead used an off-line version of the code. In the following we present step-by-step how we adapted each of the PHE [Birrell et al., 2021], the Roche [Lemenuel-Diot et al., 2020] and the Warwick-Household [Keeling et al., 2021] models to carry out the parameter structural identifiability analysis using the SIAN algorithm.

4.3.2.1 The PHE Model

In order to study the structural identifiability of the PHE model [Birrell et al., 2021], we first translate the ODE system governing the model described in (D.1) into a Maple-friendly framework. The first thing to observe is that in the continuous-time PHE model with one region and one age group the shape of the force of infection parameter $\lambda_t = 1 - (1 - b_t)^{I^1(t)+I^2(t)}$ depends exponentially on the model variables $I^1(t)$ and $I^2(t)$, rather than linearly, as in the SEIR model (eq. (2)), where

$$\lambda(t) = \frac{\beta}{N} I(t) = \frac{\beta}{N}(I^1(t) + I^2(t)). \tag{4.2}$$

However, this exponential dependency on the total number of infections raises a computational issue with our analysis if we were to use the PHE model in its original form. The SIAN algorithm – which we use to study the structural identifiability of the three models discussed in section 4.2.1 is ill-equipped to tackle an ODE system which does not reduce to a polynomial form; that is, the system of equations needs to collapse to a form:

$$P(X_1, X_2, \ldots, X_q) = a_0 + a_1 X_1^{j_1} X_2^{j_2} \ldots X_q^{j_q} + \ldots,$$

where $X_1, X_2, \ldots, X_q$ are model outputs of interest which we have knowledge of, $j_1, j_2, \ldots$ are their respective power coefficients, and $a_0, a_1, \ldots$ are the polynomial coefficients. This is not achievable if any of the ordinary differential equations in eq. (C.1) present non-linear dependencies across the variables of interest. Therefore, we use a proxy for the original PHE model ODE system, where we swap the original force of infection parameter $\lambda(t)$ for

[2]SIAN Web App: https://maple.cloud/app/6509768948056064.

its simplified counterpart in eq. (4.2). Additionally, as the PHE model assumes a constant total number of individuals present at all times – that is $N = S(t) + E^1(t) + E^2(t) + I^1(t) + I^2(t) + R(t)$ is constant, we add this equation to our working simplification of the PHE model, as seen below. We then go on to use this extended system of ODEs to determine the structural identifiability of the three model parameters – the latency period d_L, the infectious period d_I and the transmission rate β:

$$\begin{aligned}
\frac{dS(t)}{dt} &= -\frac{\beta}{N(t)}(I^1(t) + I^2(t))S(t) \\
\frac{dE^1(t)}{dt} &= \frac{\beta}{N(t)}(I^1(t) + I^2(t))S(t) - \frac{2}{d_L}E^1(t) \\
\frac{dE^2(t)}{dt} &= \frac{2}{d_L}E^1(t) - \frac{2}{d_L}E^2(t) \\
\frac{dI^1(t)}{dt} &= \frac{2}{d_L}E^2(t) - \frac{2}{d_I}I^1(t) \\
\frac{dI^2(t)}{dt} &= \frac{2}{d_I}I^1(t) - \frac{2}{d_I}I^2(t) \\
\frac{dR(t)}{dt} &= \frac{2}{d_I}I^2(t) \\
N(t) &= S(t) + E^1(t) + E^2(t) + I^1(t) + I^2(t) + R(t).
\end{aligned} \tag{4.3}$$

In the context of the PHE model, the parameter β is used as a proxy for the effect of a multitude of other parameters that go on to govern the transmission dynamics depicted in section C1.1 – such as the probability of a susceptible individual b_r^t, and therefore, the initial reproduction number R_0, the further temporal and regional fluctuations in the transmission parameters $\beta_{t,r}$ and their variance σ_β. At the beginning of this chapter, we set out to understand how our knowledge of different quantities of interest impacts our ability to recover the true underlying parameter values that generated the observed epidemic data in the first place. Therefore, we extend the system (4.3) to include one or more of the equations governing these inputs of interest:

- Daily number of deaths: $D^{\text{daily}}(t) = \mu\frac{\beta}{N(t)}S(t)(I^1(t) + I^2(t))$,
- Daily prevalence of infections: $P(t) = \frac{1}{N(t)}(I^1(t) + I^2(t))$ and
- Daily incidence of infections: $I(t) = \frac{1}{d_L}E^2(t)$,

where we assume the rate of mortality μ is a known parameter.

Input Data	Globally identifiable parameters	Only locally identifiable parameters	Non-identifiable parameters
Incidence (I)	β, d_I, d_L		
Prevalence (P)	β, d_I, d_L		
Death (D)	β, d_I, d_L		
I & P	β, d_I, d_L		
I & D	β, d_I, d_L		
P & D	β, d_I, d_L		
I & P & D	β, d_I, d_L		

Table 4.1: *The structural identifiability of the parameters of the PHE model for different choices of input data. The initial conditions are also inferred up to the total number of individuals in the population which is fixed. We do not include these in the table for simplicity, and because in practice we provide the initial conditions for the PHE model.*

In Table 4.1 we classify the parameters of the simplified PHE model based on their global, local or non-identifiability. Running the SIAN algorithm, we conclude that all three parameters β, d_I and d_L are globally identifiable – that is, they can be accurately inferred in the absence of noisy data. However, as mentioned in the beginning of this chapter and in [Dankwa et al., 2022; Cobelli and DiStefano, 1980], structural identifiability does not imply that the parameter is practically identifiable. Indeed, as (1) we have clinical information for the d_I and d_L, and (2) β summarises the effect of multiple parameters present in the original formulation of the PHE model, we only perform the parameter sensitivity analysis (section 4.3.3.1) and inference (section 4.3.7.1) for those parameters governing the transmission dynamics: the basic reproduction number R_0, the temporal fluctuations in transmission β_t, and the standard deviation σ_β of the temporal fluctuations.

4.3.2.2 The Roche Model

For identifying the structurally globally identifiable parameters of the Roche model, we proceed in a similar manner to the PHE model – that is, by assuming only one region and one age group. We first translate the existing ODE structure of the model to a Maple-compatible framework; in the original Roche model, the transmission rates vary between each infection type. Indeed, from eq. (C.7) we recall that the transmission rates are related

to each other according to:

$$\begin{aligned}
\beta_a &= \beta_{aa} = \frac{\beta_s}{2} \\
\beta_{as} &= \beta_{aas} = \frac{\beta_{ss}}{2} \\
\beta_s &= \beta_{\text{max}} - (\beta_{\text{max}} - \beta_{\text{min}}) \frac{\Theta^\gamma}{\Theta^\gamma + \Theta_{50}^\gamma} \\
\beta_{as} &= (1 + b_{ss})\beta_a \\
\beta_{aas} &= (1 + b_{ss})\beta_{aa} \\
\beta_{ss} &= (1 + b_{ss})\beta_s.
\end{aligned}$$

From these equations, the following set of parameters to be studied emerges: β_{max}, β_{min}, Θ_{50}, γ and b_{ss} (we do not count Θ among them as this quantity is known and generated at each time point based on the stringency of the non-pharmaceutical interventions in place at that time). However, even if we attempt to infer all of these parameters as is, we would not be able to untangle the effects of β_{max}, β_{min}, Θ_{50} and γ, which all appear in the equation governing β_s. Therefore, we rewrite the ODE system of the Roche model in eq. (C.6) in terms of β_s and b_{ss}:

$$\begin{aligned}
\frac{dS(t)}{dt} &= -C\frac{\beta_s}{N(t)}\Big(\frac{I^a(t)}{2} + \frac{I^{aa}(t)}{2} + I^s(t) + (1 + b_{ss})\frac{I^{as}(t)}{2} + (1 + b_{ss})\frac{I^{aas}(t)}{2} + \\
&\quad (1 + b_{ss})I^{ss}(t)\Big)S(t) \\
\frac{dE(t)}{dt} &= -\gamma_e E + C\frac{\beta_s}{N(t)}\Big(\frac{I^a(t)}{2} + \frac{I^{aa}(t)}{2} + I^s(t) + (1 + b_{ss})\frac{I^{as}(t)}{2} + \\
&\quad (1 + b_{ss})\frac{I^{aas}(t)}{2} + (1 + b_{ss})I^{ss}(t)\Big)S(t) \\
\frac{dI^a(t)}{dt} &= (1 - P_{ss})\gamma_e E(t) - \gamma_s I^a(t) \\
\frac{dI^{aa}(t)}{dt} &= P_a\gamma_s I^a(t) - \gamma_{ra} I^{aa}(t) \\
\frac{dI^s(t)}{dt} &= (1 - P_a)\gamma_s I^a(t) - \gamma_q I^s(t) \\
\frac{dI^{as}(t)}{dt} &= P_{ss}\gamma_e E(t) - \gamma_s I^{as}(t) \\
\frac{dI^{aas}(t)}{dt} &= P_a\gamma_s I^{as}(t) - \gamma_{ra} I^{aas}(t) \\
\frac{dI^{ss}(t)}{dt} &= (1 - P_a)\gamma_s I^{as}(t) - \gamma_q I^{ss}(t)
\end{aligned}$$

$$
\begin{aligned}
\frac{dI^q(t)}{dt} &= \gamma_q I^{ss}(t) + \gamma_q I^s(t) - \gamma_r I^q(t) \\
\frac{dR(t)}{dt} &= (1 - P_d)\gamma_r I^q(t) \\
\frac{dR^a(t)}{dt} &= \gamma_{ra} I^{aas}(t) + \gamma_{ra} I^{aa}(t) \\
\frac{dD(t)}{dt} &= P_d \gamma_r I^q(t) \\
N(t) &= S(t) + E(t) + I^a(t) + I^{aa}(t) + I^s(t) + I^{as}(t) + I^{aas}(t) + I^{ss}(t) + I^q(t) + \\
&\quad + R(t) + R^a(t) + D(t), \qquad (4.4)
\end{aligned}
$$

where we have added the final equation for the total number of individuals in the Roche system $N(t)$ – which is known and assumed constant at all times. When running the SIAN algorithm, if we consider the progression rates γ_e, γ_s, γ_r, γ_q and γ_{ra} among the parameters to be analysed, the algorithm failed to complete; however, when we assumed these parameters to be known, we were able to complete the analyses. Therefore, the results presented in Table 4.2 below were run for an ODE system as in eq. (4.4) with added inputs for the γ_e, γ_s, γ_r, γ_q and γ_{ra} quantities. Finally, to study the structural identifiability of the remaining parameters P_a, P_d, P_{ss}, β_s and b_{ss} in the context of our prior knowledge of different quantities of interest, we supplement the current Roche model with one or more of the following:

- Daily number of deaths: $D^{\text{daily}}(t) = P_d \gamma_r I^q(t)$,
- Daily prevalence of infections: $P(t) = \frac{1}{N(t)}(I^a(t) + I^{aa}(t) + I^s(t) + I^{as}(t) + I^{aas}(t) + I^{ss}(t))$ and
- Daily incidence of infections: $I(t) = \gamma_e E(t)$,

where we assume that the rate of mortality μ is a known parameter. We summarise our findings of the structural identifiability of the Roche model parameters in Table 4.2. We identify that in the case of the Roche model, the choice of input used for model fitting has a significant impact on which parameters can be identified. When using only daily number of deaths to parameterise the Roche model, we find that the proportions of asymptomatic, super-spreader infections and deaths P_a, P_{ss} and P_d are the globally identifiable, while b_{ss} and the symptomatic transmission rate β_s are revealed to be non-identifiable. However, in section 4.3.3.2 of this chapter, we observe that the latter two parameters have in fact a significant effect in determining the trajectories of the number of infections and deaths,

Input Data	Globally identifiable parameters	Only locally identifiable parameters	Non-identifiable parameters
Incidence (I)	P_a, P_d, P_{ss}		b_{ss}, β_s
Prevalence (P)	P_a		b_{ss}, β_s, P_d, P_{ss}
Death (D)	P_a, P_d, P_{ss}		b_{ss}, β_s
I & P	P_a		b_{ss}, β_s, P_d, P_{ss}
I & D	P_a, P_d		b_{ss}, β_s, P_{ss}
P & D	P_a, P_d		b_{ss}, β_s, P_{ss}
I & P & D	P_a, P_d		b_{ss}, β_s, P_{ss}

Table 4.2: *The structural identifiability of the parameters of the Roche model for different choices of input data. The initial conditions are also inferred up to the total number of individuals in the population which is fixed. We do not include these in the table for simplicity, and because in practice we provide the initial conditions for the Roche model.*

and indeed the log-likelihood plotted in Figure 4.4 suggests that for noisy inputs we can identify b_{ss} and β_s in practice. This seems to contradict the statement that structurally non-identifiable parameters are also practically non-identifiable, which brings into question the correctness of the SIAN algorithm used.

4.3.2.3 The Warwick-Household Model

For the third and final compartmental model we analyse using the SIAN algorithm, we use the following ODE system for a one age group, one region population to replicate the transmission dynamics of the Warwick-Household model as closely as possible to eq. (C.11):

$$\begin{aligned}
\frac{dS(t)}{dt} =& - C^N \frac{\sigma}{N(t)} S(t)\big(I^F(t) + I^{SD}(t) + I^{SU}(t) + \tau(A^F(t) + A^S(t))\big) \\
& - C^H \frac{\sigma}{N(t)} S(t)\big(I^F(t) + A^F(t) + I^{QF}(t)\big) \\
\frac{dE^{1,F}(t)}{dt} =& C^N \frac{\sigma}{N(t)} S(t)\big(I^F(t) + I^{SD}(t) + I^{SU}(t) + \tau(A^F(t) + A^S(t))\big) - 3\epsilon E^{1,F}(t) \\
\frac{dE^{1,SD}(t)}{dt} =& C^H \frac{\sigma}{N(t)} S(t) I^F(t) - 3\epsilon E^{1,SD}(t) \\
\frac{dE^{1,SU}(t)}{dt} =& C^H \frac{\sigma}{N(t)} S(t) A^F(t) - 3\epsilon E^{1,SU}(t) \\
\frac{dE^{1,Q}(t)}{dt} =& C^H \frac{\sigma}{N(t)} S(t) I^{QF}(t) - 3\epsilon E^{1,Q}(t) \\
\frac{dE^{2,F}(t)}{dt} =& 3\epsilon E^{1,F}(t) - 3\epsilon E^{2,F}(t)
\end{aligned}$$

$$
\begin{aligned}
\frac{dE^{2,SD}(t)}{dt} &= 3\epsilon E^{1,SD}(t) - 3\epsilon E^{2,SD}(t) \\
\frac{dE^{2,SU}(t)}{dt} &= 3\epsilon E^{1,SU}(t) - 3\epsilon E^{2,SU}(t) \\
\frac{dE^{2,Q}(t)}{dt} &= 3\epsilon E^{1,Q}(t) - 3\epsilon E^{2,Q}(t) \\
\frac{dE^{3,F}(t)}{dt} &= 3\epsilon E^{2,F}(t) - 3\epsilon E^{3,F}(t) \\
\frac{dE^{3,SD}(t)}{dt} &= 3\epsilon E^{2,SD}(t) - 3\epsilon E^{3,SD}(t) \\
\frac{dE^{3,SU}(t)}{dt} &= 3\epsilon E^{2,SU}(t) - 3\epsilon E^{3,SU}(t) \\
\frac{dE^{1,Q}(t)}{dt} &= 3\epsilon E^{2,Q}(t) - 3\epsilon E^{3,Q}(t) \\
\frac{dI^{F}(t)}{dt} &= 3(1-H)\epsilon d E^{3,F}(t) - \gamma I^{F}(t) \\
\frac{dI^{SD}(t)}{dt} &= 3\epsilon d E^{3,SD}(t) - \gamma I^{SD}(t) \\
\frac{dI^{SU}(t)}{dt} &= 3(1-H)\epsilon d E^{3,SU}(t) - \gamma I^{SU}(t) \\
\frac{dI^{QF}(t)}{dt} &= 3H\epsilon d E^{3,F}(t) - \gamma I^{QF}(t) \\
\frac{dI^{QS}(t)}{dt} &= 3H\epsilon d E^{3,SU}(t) + 3\epsilon d E^{3,Q}(t) - \gamma I^{QS}(t) \\
\frac{dA^{F}(t)}{dt} &= 3\epsilon(1-d) E^{3,F}(t) - \gamma A^{F}(t) \\
\frac{dA^{S}}{dt} &= 3\epsilon(1-d)(E^{3,SD}(t) + E^{3,SU}(t)) - \gamma A^{S}(t) \\
\frac{dA^{Q}(t)}{dt} &= 3\epsilon(1-d) E^{3,Q}(t) - \gamma A^{Q}(t) \\
\frac{dR(t)}{dt} &= \gamma(I^{F}(t) + I^{QF}(t) + A^{F}(t) + I^{SD}(t) + A^{S}(t) + I^{SU}(t) + I^{QS}(t) + A^{Q}(t)) \\
N(t) &= S(t) + E^{1,F}(t) + E^{2,F}(t) + E^{3,F}(t) + E^{1,SD}(t) + E^{2,SD}(t) + E^{3,SD}(t) + \\
&\quad + E^{1,SU}(t) + E^{2,SU}(t) + E^{3,SU}(t) + E^{1,Q}(t) + E^{2,Q}(t) + E^{3,Q}(t) + I^{F}(t) \\
&\quad + I^{SD}(t) + I^{SU}(t) + I^{QF}(t) + I^{QS}(t) + A^{F}(t) + A^{S}(t) + A^{Q}(t) + R(t),
\end{aligned}
\tag{4.5}
$$

where similar to the PHE and the Roche models, the total number of individuals in the population is assumed to be a constant and known quantity. The household and non-household contacts matrices (in this case each is a 1×1 contact matrix) are also known at all times and given as input when running the SIAN algorithm. We supplement the existing ODE system

in eq. (4.5) with one or more of the equations governing the trajectories for the:

- Daily number of deaths: $D^{\text{daily}}(t) = 3\epsilon d\mu(E^{3,F}(t) + E^{3,SD}(t) + E^{3,SU}(t) + E^{3,Q}(t))$,
- Daily prevalence of infections: $P(t) = \frac{1}{N(t)}(I^F(t) + I^{SD}(t) + I^{SU}(t) + I^{QF}(t) + I^{QS}(t) + A^F(t) + A^S(t) + A^Q(t))$, and
- Daily incidence of infections: $I(t) = 3\epsilon(E^{3,F}(t) + E^{3,SD}(t) + E^{3,SU}(t) + E^{3,Q}(t))$,

which serve as input data for the Warwick-Household model parametrisation. Similarly to the PHE model case, we used a proxy for the number of daily deaths registered, with a known mortality rate μ, rather than the more accurate renewal equation-type process employed in the Warwick-Household model; this is due to the limitations of the SIAN algorithm which requires us being able to transforms the ODE system into a polynomial representation – which is not possible using a renewal equation for the number of deaths. It is also important to note that when running the structural identifiability algorithm, for a number of choices of input data (represented with an asterisk in Table 4.3) the algorithm does not finish running unless we set an additional parameter as an input – the latency rate, ϵ.

Input Data	Globally identifiable parameters	Only locally identifiable parameters	Non-identifiable parameters
Incidence (I)	H, d, ϵ, γ, σ, τ		
Prevalence (P)*	H, d, γ, σ, τ		
Death (D)*	H, d, γ, σ, τ		
I & P	H, d, ϵ, γ, σ, τ		
I & D	H, d, ϵ, γ, σ, τ		
P & D*	H, d, γ, σ, τ		
I & P & D	H, d, ϵ, γ, σ, τ		

Table 4.3: *The structural identifiability of the parameters of the Warwick-Household model for different choices of input data. The initial conditions are also inferred up to the total number of individuals in the population which is fixed. We do not include these in the table for simplicity, and because in practice we provide the initial conditions for the Warwick-Household model. In the scenarios marked with an asterisk (*), the SIAN algorithm gets stuck unless the latency rate parameter ϵ is set as an additional input of the model, i.e. is known.*

We summarise in Table 4.3 our findings about the structural identifiability of the following parameters of the Warwick-Household model: the household quarantined rate H, the probability of symptom development d, the recovery rate γ, the susceptibility to infection parameter σ, the reduction in transmission of the asymptotic infections τ (and, when it is

possible to assess, the parameter equal to the latency rate, ϵ). According to these findings, all parameters searched are globally identifiable, for all input data combinations. However, in practice, we find that we are not able to accurately infer all parameters – for example, we were unable to infer the d and σ parameters when using the methods for the computation as described in the original papers [Keeling et al., 2021, 2022] and in their GitHub code[3] (more details in section 4.3.7.3). This aligns with our understanding that parameters can be structurally but not practically identifiable.

In this section, we analysed the structural identifiability of the parameters of the PHE, the Roche and the Warwick-Household models (or slightly adjusted versions of these models, when the original model features were incompatible with the SIAN algorithm) for a range of possible model inputs: prevalence data, incidence data, death data and all their possible combinations. However, as we only have access to reliable information about the number of deaths each day in each of the PHE regions of England, going forward we focus exclusively on those model parameters that are structurally identifiable in the process of determining which of the model parameters are also practically identifiable when using only death data for model fitting. In selecting the rest of the model parameters for the PHE, the Roche and the Warwick-Household models, we went through extensive and laborious efforts to identify those appropriate parameter values which are both realistic in terms of the reported disease dynamics, and which generate trajectories of the model outcomes that are as close as possible to the observed data we fit the models to. We will come back to this in the Conclusion chapter.

4.3.3 Parameter Sensitivity Analysis

For the parameter sensitivity analysis, we disregard all of the parameters unrelated to the epidemic dynamics, for all of the three models – for example, for the PHE model we would not consider the method used by the Scipy solver or the step-size δ_t for the home-made solver as possible candidates for parameter inference. Additionally, we have seen previously in Figure E1 that the choice and calibration of the home-made solver for the PHE model has little effect on the shape or amplitude of the plotted compartment count trajectory.

[3]This code is not publicly available. We were granted access to it following some discussions with the original code developers Prof Edward Hill and Prof Matthew Keeling, with whom we collaborated on a separate project.

In the following we go through each of the three models and explore how varying the parameters of interest in each model impacts the trajectories of the predicted number of daily deaths and incidence.

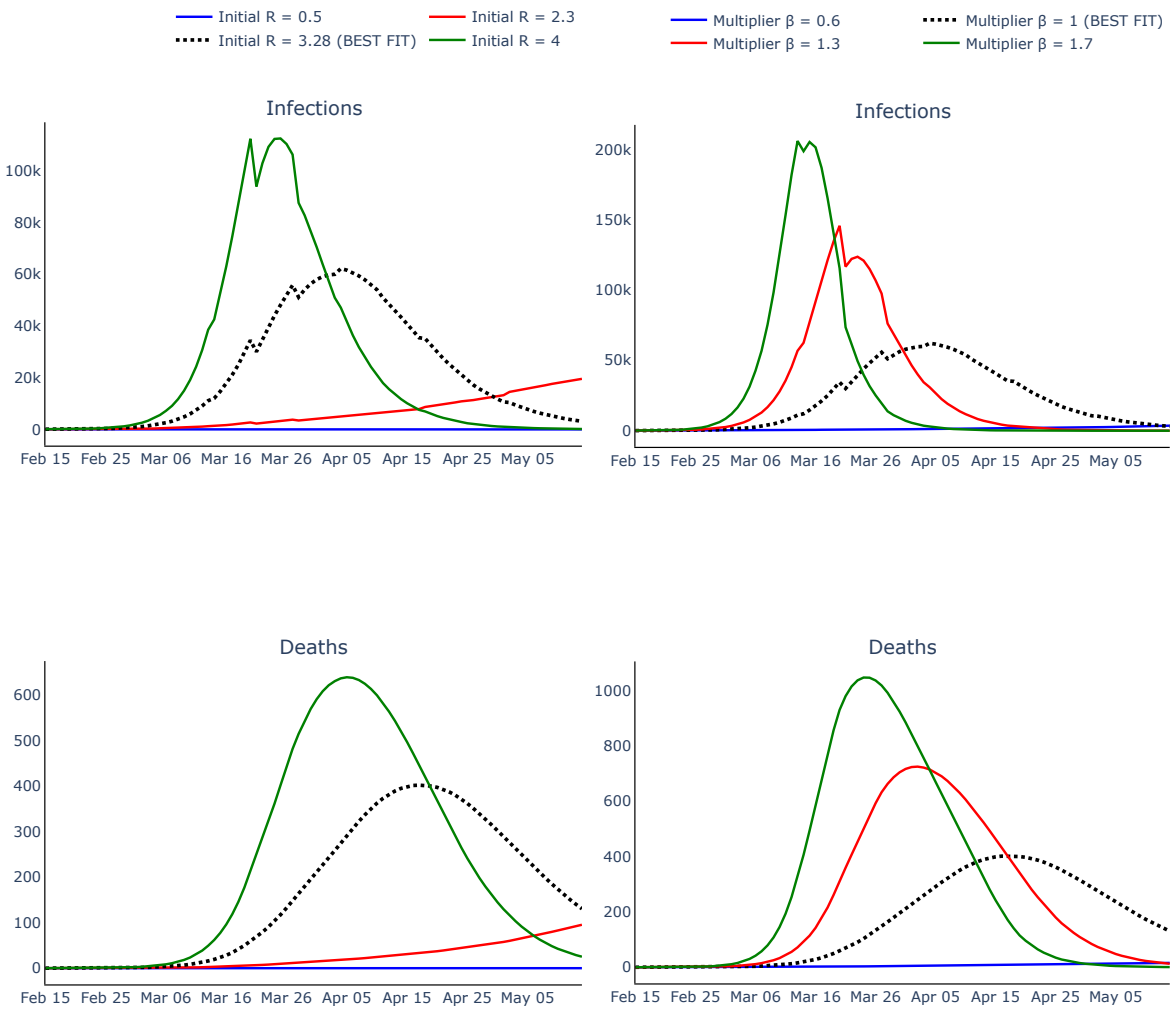
4.3.3.1 The PHE Model

In section 4.3.2.1, we identified that the latent and infectious period d_L and d_I respectively are both globally identifiable. However, these disease-specific parameters are specified in the original paper [Birrell et al., 2021], so there is no need to infer their value; we keep them fixed going forward. The initial conditions cover a large number of compartment counts, therefore, making it very cumbersome to fully explore the effect of incremental changes in initial conditions on the trajectories of the model outputs.

We restrict ourselves to study the effects of the regional-parameters, more precisely, the region-specific initial reproduction number R_0 and that of the regional weekly fluctuation in transmission parameters β. However, the number of regional fluctuation parameters is large: for a single region and the time window chosen (15$^{\text{th}}$ of February to 25$^{\text{th}}$ of June 2020), we would have 13 β_t parameters to vary, which, similarly to the initial conditions, is not feasible. Instead, we consider a model parameterisation in which $\beta_t = \beta, \forall t$ with β (which we refer to as the 'multiplier β') being the quantity with respect to which we assess the model sensitivities to changes.

In Figure 4.3 we plot the trajectories of the daily number of new infections and deaths for different values of the initial reproduction number R_0 and of the multiplier of the profile of temporal fluctuations β. In black, we represent the curves of best fit for the UK epidemic data calculated in section 4.3.7.1 and in colour how the trajectories of the number of new infections and deaths respectively change for different values of the parameter considered. The slight jaggedness we observed in the generated infection trajectories in all of the scenarios considered around 13$^{\text{th}}$ of March 2020 is caused by the application of the first interventions in the UK, which alter the regional contact matrices that were used up to this point in the simulation.

An increase in the overall temporal fluctuations lead to both a higher and earlier peak in both the predicted number of deaths and new infections, as seen in Figure 4.3(b). However, a decrease in transmission of 40% overall ensures that the epidemic either increases extremely



(a) Initial reproduction number R_0 (b) Trajectory multiplier of β_t parameters

Figure 4.3: *Trajectories of the total number of infected individuals and deaths when varying the: (a) initial reproduction number R_0 and (b) overall trajectory multiplier of the β_t parameters (previously fitted to the data according to section 4.3.7.1) of the model. We represent in different colours the curves produced by running the PHE model for a series of parameters and in the dotted black the curve of the best fit to the data.*

slowly over time, or it never takes off. We identify a similar result if the initial reproduction number $R_0 \leq 1$. Moreover, from Figure 4.3(a) we observe that even a rather large value of $R_0 \approx 2.3$ leads to a much slower start of the COVID-19 epidemic than the UK data suggests. Therefore, a value of $R_0 = 3.28$ may be reasonable – see section C10.

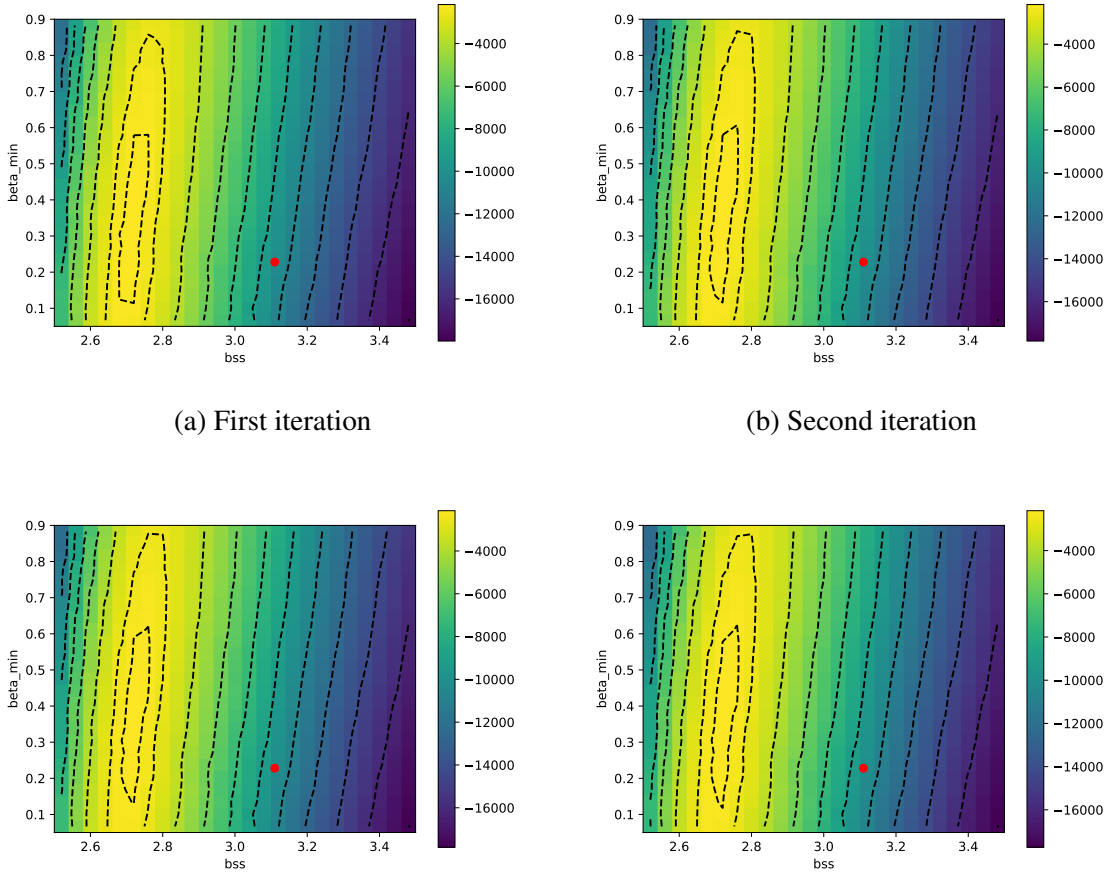
Since there is a clear change in the overall dynamics of the number of deaths and the number of new infections for both of the parameters considered in this analysis, we will proceed with inferring the initial reproduction number R_0 and the temporal fluctuations in transmission profile β_t (and their variance σ_β) as described in section C.3.

4.3.3.2 The Roche Model

Similar to the PHE model case, we look to Table 4.2 to inform our selection of which Roche model parameters to perform our sensitivity analysis on. According to these results, when using exclusively death data, we identified that only the proportions of people that go on to become asymptomatic (P_a), super-spreader (P_{ss}) or dead (P_d) are globally identifiable in the Roche model, while all other transmission parameters are non-identifiable (the increase in super-spreader transmission b_{ss} and the symptomatic transmission rate β_s).

Just as a reminder, to ensure that the structural identifiability algorithm finished running, we assumed that all compartment time parameters were known. The symptomatic transmission rate β_s is not among the parameters of the original Roche model, but rather it is computed using the minimum and maximum transmission rates β_{min} and β_{max}, the stringency index Θ, the stringency index for a 50% reduction in transmission Θ_{50}, and the tuning parameter for the adaptability of interventions γ, as described in eq. (C.7). This allows for only one of these parameters to be inferred, as they only appear in this expression of the transmission parameter β_s; we choose for that reason to assume only β_{min} as a free parameter amongst them, as we can use clinical studies and social studies on the effect of interventions to determine accurate values for the maximum transmission rate β_{max} (a property of the virus), and the stringency indices Θ and Θ_{50} and the adaptability factor γ (which are properties of the population and intervention profiles). Therefore, the only parameters of the Roche model remaining as candidates for inference would be P_a, P_{ss}, P_d, β_{min} and the increase in super-spreader transmission factor b_{ss} – the latter two of which are practically non-identifiable as the structural identifiability analysis in section 4.3.2.2 would suggest.

However, running the Roche model for a variety of values of the β_{min} and, respectively, b_{ss} (as depicted in Figures 4.5(a) and 4.5(b) respectively) while keeping all other parameters fixed produces outputs that change significantly with the value of the parameter varied. This implies that the two parameters are in fact correlated and each can be recovered when the other is known. Indeed, when analysing the changes in the log-likelihood evaluated for variable values of the β_{min} and the b_{ss} parameters, we observe a slight linear correlation of the two parameter estimates – as illustrated in Figure 4.4 in which we plot the log-likelihood surface for four different iterations of the noise distribution on the death data. Running both optimisation and sampling inference algorithms to retrieve the true values of β_{min} and b_{ss}



(a) First iteration

(b) Second iteration

(c) Third iteration

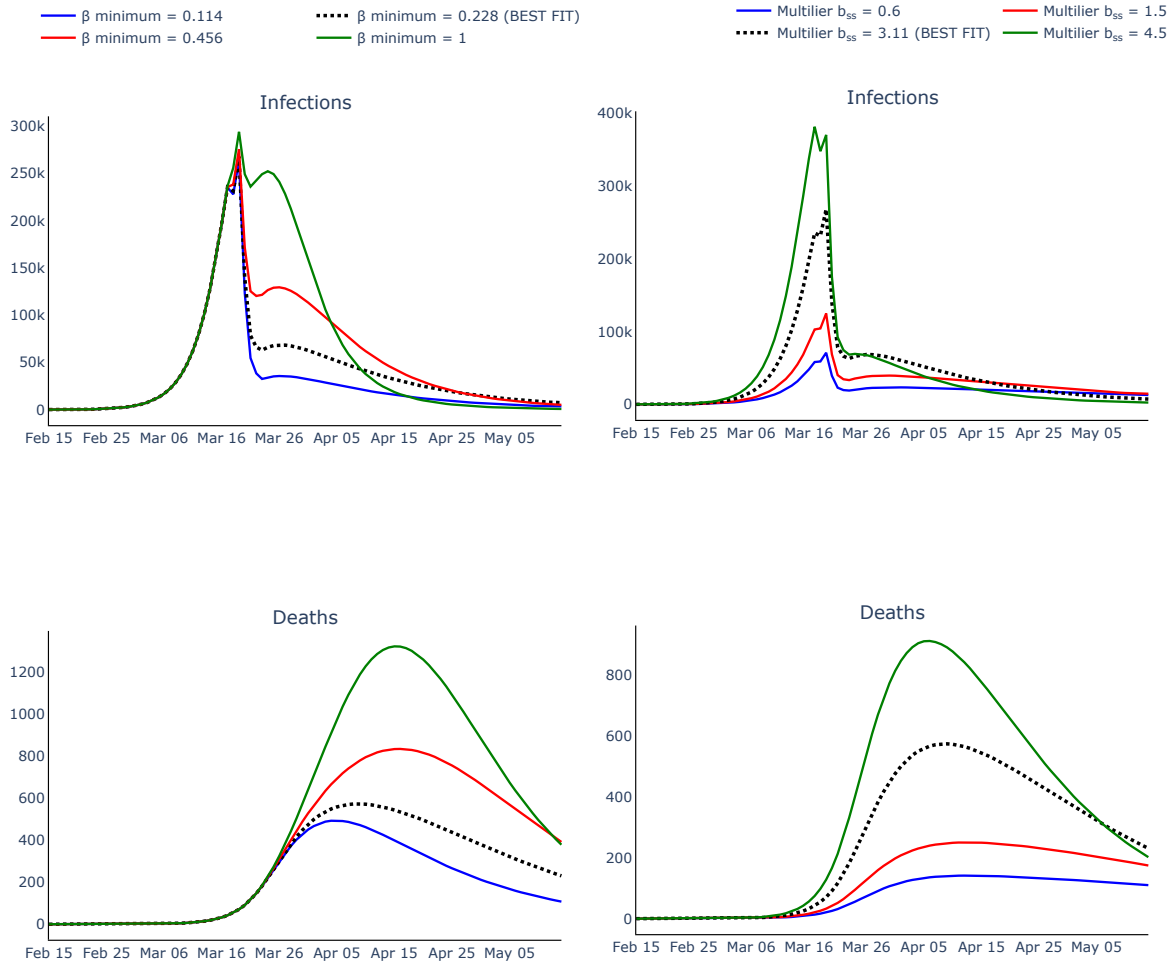
(d) Fourth iteration

Figure 4.4: *Four instantiations of the log-likelihood of the Roche model. The red dot in each log-likelihood plot indicates the true parameter values for β_{min} and b_{ss} that were used to generate daily number of deaths trajectories. These deterministic quantities are then perturbed using a negative binomial noise distribution, according to eq.* (C.9)*, to generate realistic synthetic datasets of daily number of deaths. We then go on to use these synthetic datasets to fit the Roche model and compute the log-likelihood surface.*

we observe that these inference algorithms run correctly (as can be seen in section 4.3.7.2); this seems to directly contradicts our findings in section 4.3.2.2. A reason for this occurring might be that the optimisation algorithm is affected by the noise associated with the death data, making these parameters practically identifiable, even though they are not structurally so; otherwise, this again brings into question the correctness of the SIAN algorithm used.

In Figure 4.5 we present the trajectories for the daily number of new infections and daily number of deaths for four different choices of the minimum transmission rate β_{min} (that is,

when all NPIs included in OxCGRT study [Hale et al., 2021] are maximised – see Table C2) and respectively the relative increase in super-spreader transmission b_{ss}. The black dotted trajectories represent the curves of best fit of the Roche model to UK epidemic data. The jaggedness of all generated infection trajectories around the 13th of March 2020 is caused by the application of the first interventions in the UK, which alter the stringency index and therefore the rates of transmission that were used up to this point in the simulation.



(a) Minimum transmission rate β_{min} (b) Increase in super-spreader transmission b_{ss}

Figure 4.5: *Trajectories of the total number of infected individuals and deaths when varying the: (a) the minimum transmission rate of the virus β_{min} when all interventions are maximised and (b) the factor of increase in the super-spreader transmission (previously fitted to the data according to the section 4.3.7.2) of the model. We represent in different colours the curves produced by running the Roche model for a series of parameters and in the dotted black the curve of the best fit to the data.*

From Figure 4.5(a) we observe that mid-way through the peak of infections, when interventions are applied, the subsequent amplitudes of both the daily incidence of infection and

deaths increase as the minimum transmission rate of the virus becomes higher, that is the virus is less responsive to the strongest set of restrictions. Additionally, for lower β_{min}, the observed peak in the number of deaths occurs earlier, while for the number of new infections, there are no perceptible changes in the timing of the occurrence of the secondary peak. In comparison, stronger levels of the super-spreader transmission result in higher and slightly earlier occurring peaks in the trajectories of daily incidence and deaths as it can be seen in Figure 4.5(b). In section 4.3.7.2, we therefore proceed with inferring these two parameters, β_{min} and b_{ss}.

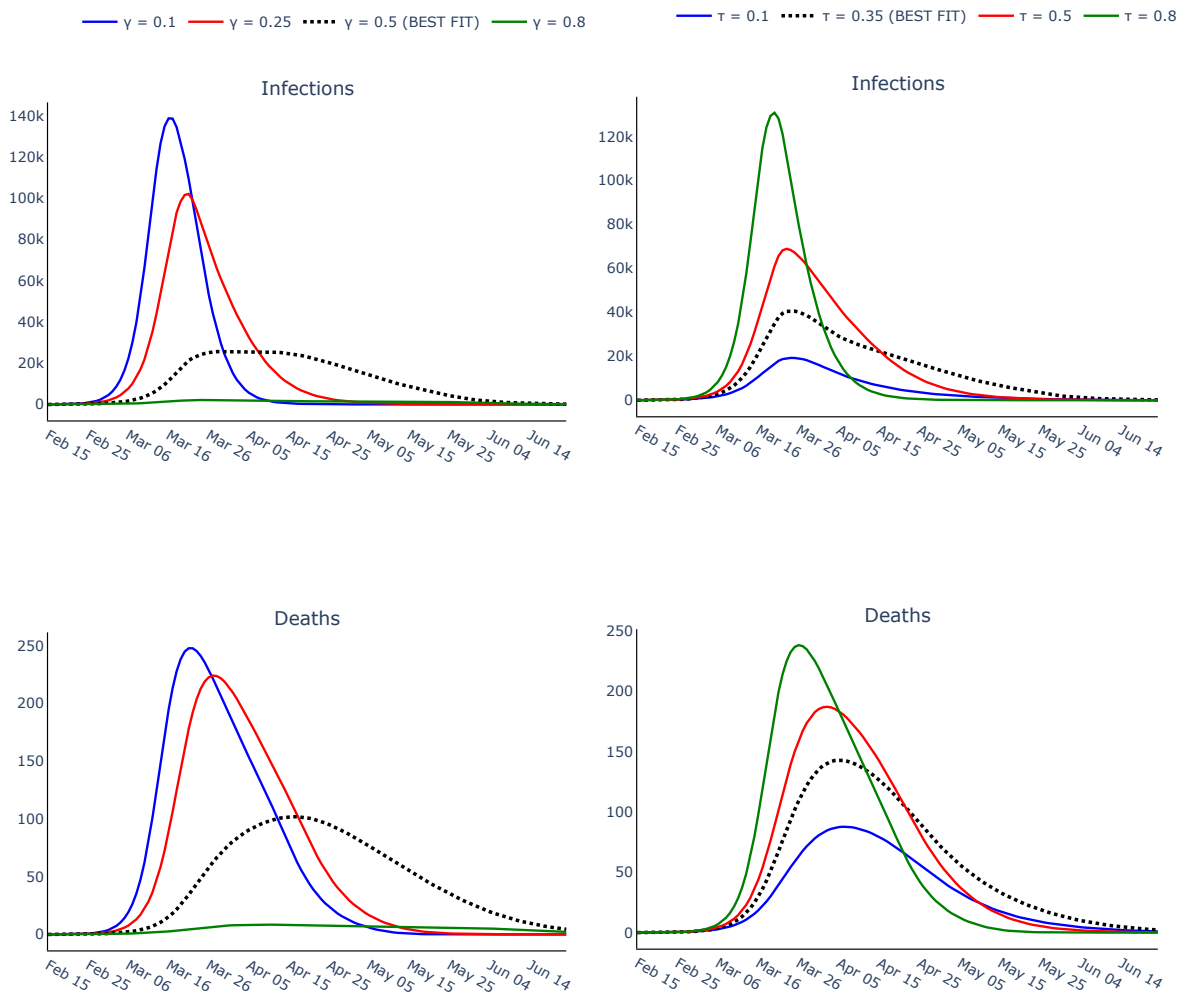
4.3.3.3 The Warwick-Household Model

For the Warwick-Household model, we identified in section 4.3.2.3 that when fitting the model exclusively to daily deaths data, we should be able to infer a large number of parameters, provided we know the latency rate ϵ, the strength of intervention parameters q_H, q_S, q_W, q_O, θ and ϕ, and the contact matrices for household (C^H) and non-household (C^N) contacts. However, out of these identifiable parameters, we are unable to infer the susceptible to infection parameter σ and the probability of symptom development vector d.

Additionally, we do not need to fit the household quarantine rate H, as this is defined in terms of the strength of intervention parameter ϕ according to the following relationship: $H = 0.9\phi$ as prescribed in the technical paper [Keeling et al., 2022]; hence, we are only really left with two parameters which we can practically identify in the context of the Warwick-Household model: the recovery rate γ and the factor of reduction in transmission of asymptomatic infections τ. The initial conditions cover a large number of compartment counts, therefore making it very difficult to study the behaviour of the output of the model over incremental changes. We therefore fix them according to the original [Keeling et al., 2021] paper, where we prescribe a total number of symptomatic infections occurring in the region, distributed across the age groups according to a pre-defined vector of initial infection fractions as given in eq. (4.6):

$$(0.0039, 0.00403, 0.00485, 0.029, 0.17231, 0.28433, 0.16048, 0.34113),$$

corresponding to the $0-1$, $1-5$, $5-15$, $15-25$, $25-45$, $45-65$, $65-75$ and the $75+$ years old, respectively.



(a) Recovery rate γ (b) Reduction in asymptomatic transmission τ

Figure 4.6: *Trajectories of the total number of infected individuals and deaths when varying the: (a) the recovery rate γ and (b) the factor of reduction in transmission of asymptomatic infections τ. We represent in different colours the curves produced by running the Warwick-Household model for a series of parameters and in the dotted black the curve of the best fit to the data.*

In this section, we restrict ourselves to study the effects of the γ and τ parameters on the model outputs. From Figure 4.6(a) we observe that the peak in both the daily incidence of infection and deaths is both higher and occurs at an earlier date as the average length of the infectious period increases, that is as the recovery rate γ decreases. Similarly, stronger reduction in asymptomatic transmission results in earlier and higher peaks in the trajectories of daily incidence and deaths as it can be seen in Figure 4.6(b). In the following sections, we proceed with inferring these two parameters, γ and τ.

4.3.4 Parameter Optimisation & Inference

From the parameter sensitivity analysis detailed in section 4.3.3, we identified the following groupings of parameters as important sources of variability in the outputs of each of the three models described in section 4.2.1 and further in Appendix C:

- *the PHE model*: the initial reproduction number R_0, the values of the temporal fluctuations in transmission β_t, and the standard deviation σ_β of the temporal fluctuations;
- *the Roche model*: the minimum transmission rate of the virus – that is when all NPIs maximised β_{min} and the super-spreader increase in transmission effect b_{ss} ;
- *the Warwick-Household model*: the recovery rate γ and the factor of reduction in transmission of asymptomatic infections τ.

In this section, we use both synthetic datasets and the early UK epidemic data on the number of deaths (collected from the ONS data [GOV.UK]) and infection prevalence (with numbers adapted from the REACT1 study [Nicholson et al., 2021; Eales, 2020]) to inform on the best choices of these parameters. Fitting the model parameters ensures the accuracy of the forecasts on the daily number deaths and new infections. For all other model parameters which we assume known, we either use the values identified from the original papers describing the models [Birrell et al., 2021], or through extensive experimentation, i.e. hand-tuning, with possible realistic values [Lemenuel-Diot et al., 2020; Keeling et al., 2021] to inform the choice of the parameters.

In section 0.4 we introduced two methods for model fitting: optimisation and Markov Chain Monte Carlo (MCMC) sampling. In the following we will present the results for each of these methods, which take as their only parameters a vector of the times for which we wish to run the parameter inference and the weighting values ω_p and ω_d of the contribution of the large-scale infection survey and death data respectively in computing our quantity of interest – the posterior distribution (details in the Bayesian Inference section). Both methods are run using the same controller class object: *PheSEIRInfer* for the PHE model, *RocheSEIRInfer* for the Roche model and *WarwickSEIRInfer* for the Warwick-Household model. In Appendix F1 we include the a list of all the methods for loading the data required for parameter inference (respectively parameter optimisation), which are run prior to the inference (optimisation) routine.

The models are run from 15th of February to 25th of June 2020, which covers the first wave of the UK epidemic as well as the first lockdown. The lockdown was instituted in the UK on the 23rd of March, which is 44 days after the start of the simulation window. Therefore, based on eq. (C.3), we have 13 weekly β_t parameters – and together with the standard deviation σ_β of the temporal fluctuation β_t and the initial reproduction number R_0, 15 per region left to infer for the PHE model. For the Roche model both parameters we identified from section 4.3.2, the minimum transmission rate β_{min} and the relative increase in super-spreader transmission b_{ss} are neither age- nor region-structured; this results in only two parameters left to infer. Finally, for the Warwick-household model, the reduction in transmission of asymptomatic cases τ and the recovery rate γ are also one dimensional and therefore we have only two parameters to infer. Since there are no inter-regional interactions, we run our parameter inference methods decoupled to ensure best results.

The ONS death data provide daily number of deaths for each PHE region, beginning with 12th of March 2020 and can be used directly in the fitting procedure. The REACT1 study results require however additional work. Firstly, the numbers of positive tests do not immediately translate to the number of new infections in the population, as the number of tests conducted is much smaller than the whole population count. Instead, we can use the proportion of positives for each region and age group for each REACT1 round as the percentage of the number of newly infected individuals in the population between two consecutive rounds (as the case study subjects are randomly sampled in the population, results within this sample are indicative of the behaviour of the population as a whole) and multiply this quantity by the number of individuals in the desired region and of the desired age taken from census data [ONS]. Secondly, the regional age-structured counts of the total number of samples tested and their positive results are provided on a weekly basis, starting with the 4th of May 2020, which provides for our analysis time-frame only 8 datapoints occurring much later then the predicted time of the peak in the number of infections. Thus, incorporating the effect of the infection survey data in the computation of the overall likelihood function will induce the model to overfit the tail of the trajectory of infections at the expense of producing worse fits for the death curve. The scarcity of this data source for the initial period of the epidemic together with this biasing effect compels us to ignore the contribution of the infection survey data to the posterior distribution, i.e. $\omega_p = 0, \omega_d = 1$.

4.3.5 Parameters values used for the model fitting

4.3.5.1 PHE model parameters

We assume that initially all the infected individuals of the population are in the I^1 compartment with equal counts in all age groups and regions, with no individuals in either of the E^1, E^2, I^2 or R compartments. We assign the number of individuals in the susceptible age-dependent compartments to be equal to the unaltered census data counts [ONS], as the number of individuals in the infectious compartments is negligible population-wise. We assume both the latent and the infectious periods to be $d_L = d_I = 4$ and the over-dispersion of the negative binomial distribution used in the computation of the log-likelihood of the number of deaths is $\eta = 5$; these values are taken directly from the original paper [Birrell et al., 2021] – we use the mean of the prior distributions to guide our choice of parameter values for d_L, d_I and η. The timestep of the solver is not used in our analysis – we use the 'RK45' Scipy solver [Virtanen et al., 2020; Dormand and Prince, 1980] instead. The rest of the PHE model parameters, i.e. the initial reproduction number R_0, the temporal fluctuations in transmission post-lockdown β_t and the standard deviation of the log-normal distribution σ_β are to be inferred.

Hence, the PHE model parameters are given as follows:

- **the initial conditions**:
  - $S(0) =$ the age-structured London census data [ONS];
  - $I^1(0) = 50$ individuals present in each age compartment;
  - $E^1(0) = E^2(0) = I^2(0) = R(0) = 0$ individuals in each age compartment;
- **the disease-specific parameters**:
  - $d_L = 4$;
  - $d_I = 4$;
- **the simulation-specific parameters**:
  - *Solver method = RK 4(5) SciPy* ODE solver.

4.3.5.2 Roche model parameters

In the parameter sensitivity analysis we studied the impact of the minimum transmission rate β_{min} and the increase in super-spreader transmission factor b_{ss} on two outputs of interest of the Roche models: the daily incidence and daily death counts. Similar to the PHE model, the susceptible individuals compartments are populated using the unaltered population census datasets from the ONS [ONS] as the number of initially infected individuals is negligibly small compared to the entire population size. We concentrate all initial infectious individuals in the symptomatic infectious compartment I^s, leaving all other infectious, exposed and recovered compartments empty. These initial infections are spread equally across all regions and age groups.

For the values of the expected waiting times in each infection state for the Roche model, we use the following fixed values: the latency period $k = 6.8$ days, the presymptomatic infectious period $k_s = 5$ days, the time until quarantining from symptom onset $k_q = 2.5$ days and the recovery period $k_r = 8$ days (and therefore $k_{ri} = 10.5$ days). The symptomatic and asymptomatic infectious periods are age-dependent; we assume equal times to recovery for all age groups, that is $k_{ri} = 8$ days (and $k_{rii} = 10.5$ days) for all age groups i. Moreover, we assume that 60% of infections are asymptomatic ($P_a = 0.6$) for all age groups, that 15% of all infections are super-spreaders ($P_{ss} = 0.15$) and that of those quarantined, the proportion of cases that go on to die (P_d) is given by the age-dependent case fatality ratio (CFR) from [Verity et al., 2020]. In terms of transmission parameters, we settle on a maximum transmission value $\beta_{\text{max}} = 1.8$ – that is the transmission rate of the virus when no interventions are implemented, a half-stringency index $\Theta_{50} = 50$ and flexibility of intervention parameter $\gamma = 12$, as defined in Appendix C2. However, a cursory overview reveals that all these parameter values, apart from P_d, differ from the values reported in the original study [Lemenuel-Diot et al., 2020] (e.g. $k = 3.43$, $k_s = 2.57$, $k_q = 1$, $k_{ri} = 9$, $P_a = 0.658$, $P_{ss} = 0.0955$, $\Theta_{50} = 37.5$ to name a few); this is because when we instantiate the Roche model with these prescribed parameter values we are unable to fit the model to the trajectory of observed daily number of deaths. Therefore, we hand-tuned these parameters until a reasonable fit is achieved, while being mindful that only reasonable parameter regimes were selected – for example, we do not accept a parameterisation in which the latency period $k > 10$ days, as clinical studies indicate average values of around 4 days.

Therefore, for the Roche model we use the following set of fixed parameters:

- **the initial conditions**:
  - $S(0) =$ the age-structured London census data [ONS];
  - $I^s(0) = 40$ individuals present in each age compartment;
  - $E(0) = I^a(0) = I^{ss}(0) = I^{aa}(0) = I^{as}(0) = I^{ss}(0) = I^q(0) = R(0) = R^i(0) = D(0) = 0$ individuals in each age compartment;
- **the compartment times parameters**:
  - $k = 6.8$ days;
  - $k_s = 5$ days;
  - $k_q = 2.5$ days;
  - $k_r = 8$ days;
  - $k_{ri} = 10.5$ days;
- **the proportion parameters**:
  - $P_a = 0.6$ in each age compartment;
  - $P_{ss} = 0.15$ in each age compartment;
  - $P_d =$ the age-structured case fatality ration vector from [Verity et al., 2020];
- **the transmission-specific parameters**:
  - $\beta_{\text{max}} = 1.8$;
  - $\gamma = 12$;
  - $\Theta_{50} = 50$;
- **the simulation-specific parameters**:
  - *Solver method = RK 4(5) SciPy* ODE solver.

4.3.5.3 Warwick-Household model parameters

In section 4.3.2.3 we identified that when fitting exclusively to death data we are able to retrieve the underlying values of the household quarantine rate H, the probability of symptom development d, the recovery rate γ, the susceptibility to infection parameter σ, and the reduction in transmission of the asymptotic infections τ, provided that we know the latency rate parameter ϵ. However, as discussed in the same section, we are unable to practically identify all the parameters at once from the observed daily number of deaths. Instead, we focus our efforts to infer instead two of the key parameters enumerated, which could not be reliably inferred from clinical studies; that is the recovery rate γ and the reduction in transmission of the asymptotic infections τ.

From the sensitivity analyses on these two parameters we performed in section 4.3.3.3, we concluded that both γ and τ have a strong impact on the overall shape of the infection and the death curves, primarily in terms of the location of the trajectory peak and its height. For fitting the Warwick-Household model to the observed daily number of deaths in each of the seven PHE regions of UK during the first wave of the COVID-19 epidemic [GOV.UK], we assume the same number of individuals in each age-dependent susceptible compartment as those accounted for in the 2020 population census [ONS]. As epidemics usually start and spread across multiple households before spreading within the community [Goeyvaerts et al., 2018], we assume that all infected individuals are initially concentrated in the first-within-the-household symptomatic compartments I^F, leaving all other infectious, exposed and recovered compartments empty. We assume a total number of 200 individuals, spread across the different age groups according the the following vector of infection incidence frequencies used in [Bouros et al., 2024; Moore et al., 2022] to correspond to the seven age groups considered in the Warwick-Household:

$$(0.0039, 0.00403, 0.00485, 0.029, 0.17231, 0.28433, 0.16048, 0.34113). \tag{4.6}$$

To run the optimisation and inference routines in the case of the Warwick-Household model, we fix all transmission and progression parameters, apart from τ and γ. Similar to the case of the Roche model, we hand-tune the rest of the parameters of the Warwick-Household model, as in the original paper most of them are claimed to be inferrable from data – however, we still use the prior distributions reported in the original technical paper Keeling et al. [2022] as

a guideline in selecting these parameter values. In the following, we provide the values used for the parameters we fix: for the latency rate we use $\epsilon = 0.2$ (equal to the prior mean for the parameter as reported in [Keeling et al., 2022]), which equates with a 5 day latency period; the vector of age-dependent probabilities of symptom development d is fixed at a value 1.43 bigger than the group-adjusted equivalent quantity used in [Bouros et al., 2024; Moore et al., 2022] for the UK. Similarly, the age-dependent vector of susceptibility parameter σ is fixed at 1.09 times the equivalent group-adjusted quantities from the same studies. That is:

$$d = (0.097, 0.097, 0.026, 0.068, 0.142, 0.308, 0.837, 0.973)$$
$$\sigma = (0.67, 0.67, 0.47, 0.57211056, 0.77, 1.007, 1.468, 1.545).$$

Using the unadjusted values for d and σ, i.e. identical to the references, leads to poor model fitting to the death data. Lastly, we choose the intervention parameters as: the household quarantine parameter H is fixed at a proportion of 0.9 of the strength of interventions parameter ϕ; the intervention parameters for the increase and reduction factors of the intervention matrix and the baseline contact matrix for the different activities are also assumed to vary. We use the values identified in the original code associated with the [Keeling et al., 2021, 2022] submissions as a guide for the timeline of interventions imposed and their respective strengths. The scaling factor for all public-facing interactions θ remains at 0.3 at all times.

The timeline for the adherence to interventions parameter ϕ can be visually retrieved from [Keeling et al., 2022, Figure 8]. The rest of the parameters, which as we mentioned are chosen to follow as closely to those used in the supporting code, are not openly-available; we received special permission from the original authors to consult the private online repository containing the source code. However, using the values included in the inference routine of the original source code did not produce accurate fits of the Warwick-Household model to the death data observed for the first COVID-19 wave in the UK. We, therefore, manually adapted the values until a good enough fit was produced. In Table 4.4 we record the values of the NPIs strengths for each activity and their time of implementation.

However, despite this, we were unable to fully replicate the transmission dynamics presented and their visual results, due to project time constraints, and more importantly because there was no detailed documentation of the code functionality and the features implemented differ significantly from the infrastructure detailed in the associated papers [Keeling et al., 2021,

Definition	Parameter	15th Feb	12th Mar	16th Mar	20th Mar	12th May
Increase in household contacts	q_H	1	1.075	1.4	1.675	1.675
Reduction in school contacts	q_S	1	1	0.2	0.05	0.05
Reduction in work-place contacts	q_W	1	0.7	0.5	0.2	0.2
Reduction in all-other contexts contacts	q_O	1	1	0.7	0.05	0.05
Adherence to interventions	ϕ	0	0.3	0.5	0.5	0.7

Table 4.4: Time-dependent intervention values for the Warwick-Household model

2022]. We will discuss this in more detail in section 4.3.8.1.

Hence, for the Warwick-Household model we use the following fixed parameter values:

- **the initial conditions**:
  - $S(0) =$ the age-structured London census data [ONS];
  - $I^F(0) = 200$ individuals spread across the different age compartments according to the initial infection fraction defined in [Bouros et al., 2024; Moore et al., 2022];
  - $E^{j,F}(0) = E^{j,SD}(0) = E^{j,SU}(0) = E^{j,Q}(0) = I^{QF}(0) = I^{SD}(0) = I^{SU}(0) = I^{QS}(0) = A^F(0) = A^S(0) = A^Q(0) = R(0) = 0$ individuals in each age compartment;
- **the regional parameters**:
  - $h = 0.9$;
- **the disease-specific parameters**:
  - $d = 1.43$ times the age-structured symptom risk vector from [Bouros et al., 2024;

Moore et al., 2022];

- **the transmission-specific parameters**:
  - $\epsilon = 0.2$;
  - $\sigma = 1.09$ times the age-structured susceptibility vector from [Bouros et al., 2024; Moore et al., 2022];
- **the simulation-specific parameters**:
  - *Solver method = RK 4(5) SciPy* ODE solver.
- **the social-distancing parameters**:
  - $\theta = 0.3$;
  - $\phi = (0, 0.3, 0.5, 0.5, 0.7)$;
  - $q_H = (1, 1.075, 1.4, 1.675, 1.675)$,
  - $q_S = (1, 1, 0.2, 0.05, 0.05)$,
  - $q_W = (1, 0.7, 0.5, 0.2, 0.2)$,
  - $q_O = (1, 1, 0.7, 0.05, 0.05)$,
  - $\text{times}^{\text{NPIs}} = (1, 27, 31, 35, 95)$.

4.3.6 Parameter Optimisation

In this section, we use the CMA-ES algorithm to run the parameter optimisation for the remaining free parameters of each of the three models discussed in section 4.2.1: the PHE, the Roche and the Warwick-Household models. The CMA-ES optimisation algorithm was developed by [Hansen et al., 2003; Hansen, 2006] and is integrated in the PINTS software package [Clerx et al., 2019].

This procedure will help us determine for each of the models which values of these free parameters best fits the observed number of deaths. Since in this section we are primarily interested in validating the correctness of our implemented inference algorithm, we fit the PHE model to synthetic number of deaths, rather than real ONS death data, as in the case of the generated data, we know exactly which choice of model parameters have produced the

trajectories of observations, and hence we can assess how close our estimates of the model parameters are to the true values underpinning the data. The death dataset we use to fit the model comprises of the daily number of deaths recorded by the Office of National Statistics (ONS) [GOV.UK] between 15$^{\text{th}}$ of February 2020 and 25$^{\text{th}}$ of June 2020 in the region of London, with similar plots being able to be produced for the other PHE regions. However, if we were to use the number of cases registered by ONS at the start of the simulation window (that is, ONS records only a total 10 cases on the 15$^{\text{th}}$ of February 2020 in all of England) to set the initial conditions of any of the model we would not be able to produce adequate fits to the death curve – instead we carefully hand-tune the value of the total number of initial infections in each of the PHE, the Roche and the Warwick-Household models. The rationale behind this choice is that ONS only accounts for the reported Coronavirus cases at that date – the true incidence of infection was speculated to be much higher [WHO, 2024d]. After laborious trial and errors, we settle on initialising the number of infectives in the population so that:

- **the PHE model**: 50 infected individuals in each age group of the first infectious I^1 compartments;
- **the Roche model**: 40 initial infected individuals in each age group of the symptomatic infectious I^s compartments;
- **the Warwick-Household models**: a total of 200 symptomatic individuals spread across all the age groups of the first infectious individuals within the household, I^F, according t the initial infection frequency vector from [Bouros et al., 2024; Moore et al., 2022].

These values are not too large that the total number of initial infectives is no longer be negligible with respect to the susceptible counts, nor that it is unlikely to have been overlooked early on in the epidemic. All other parameters that are not fitted are fixed as described at the beginning of this section.

4.3.6.1 The PHE Model

We run the parameter optimisation for the 15 parameters we identified previously, i.e. the initial reproduction number R_0, the temporal fluctuation post-lockdown β_t and its standard deviation σ_β.In Figure 4.7, we plot the trajectories of the daily total number of new infec-

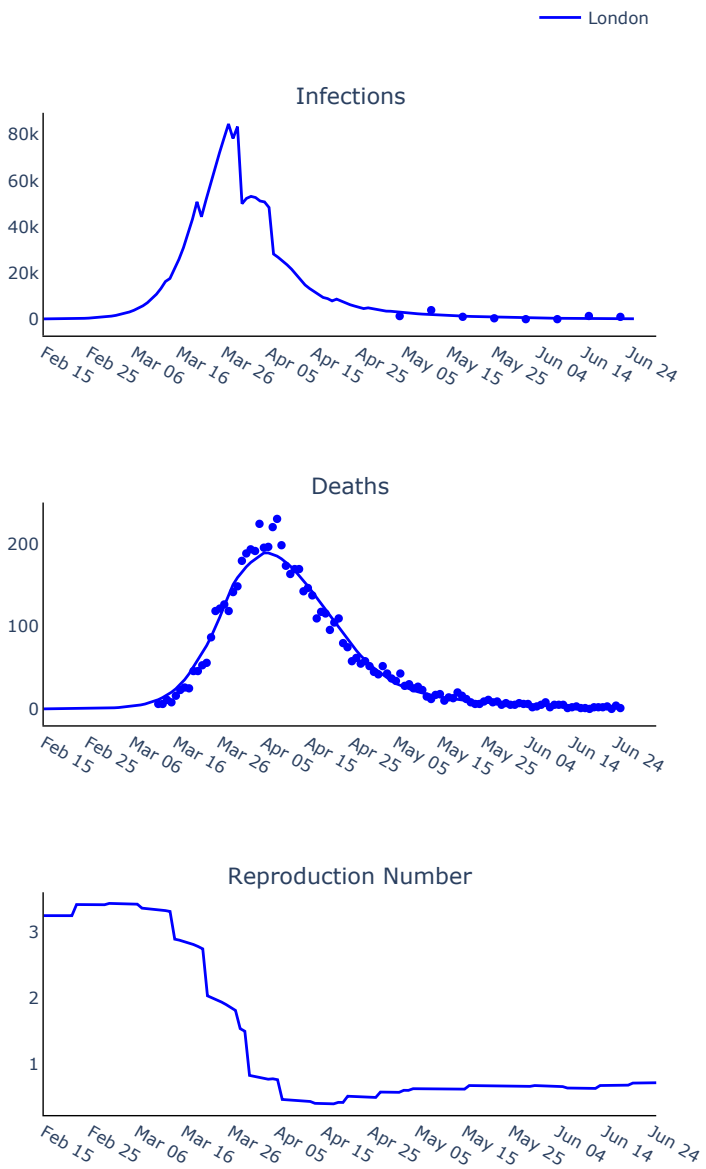


Figure 4.7: *Trajectories of the daily total number of new infections, deaths and instantaneous reproduction number R_t for London produced using the PHE model. The dots represent the observed number of positive tests scaled to the population (first panel) and the observed number of deaths (second panel).*

tions, deaths and instantaneous reproduction number R_t for London. The dots represent the observed number of positive tests scaled to the population (first panel) and the observed number of deaths (second panel). The optimised parameter values generating this figure are $R_0 \approx 3.25$, $\sigma_\beta = 0.15$ and the post-lockdown temporal fluctuation in transmission β_t ranging between 0.346 and 0.565.

We observe that the trajectory of the expected daily number of deaths produced using the

optimised parameters fits the observed data very well. Additionally, the number of new infections validate the infection survey data used to instantiate the optimisation procedure. The reproduction number suffers a sharp decrease around the 23rd of March 2020, which coincides with the institution of the lockdown in the UK, followed by a slow increase towards the end of the model time-frame, which again mirrors the gradual lifting of interventional measures by the government.

4.3.6.2 The Roche Model

Similar to the PHE model, in Figure 4.8, we present the results of our optimisation routine for the London PHE region for the minimum transmission β_{min} and the relative increase in super-spreader transmission b_{ss}.

We plot the trajectory of the daily total number of new infections, daily deaths as well as the time-dependent instantaneous reproduction number R_t. The dots indicate the observed number of positive tests scaled to the population computed from the REACT1 study (first panel) and, respectively, the observed number of deaths according to the ONS dataset (second panel). This figure is generated by forward-simulating the Roche model with the following optimised values for the minimum transmission $\beta_{\text{min}} = 0.032$ and the relative increase in super-spreader transmission $b_{ss} = 0.95$, respectively. This latter value indicates that super-spreader infectives have roughly double the transmission risk compared to their non super-spreader counterparts. We observe that around the date of 13th of March 2020 both the number of daily new infections and the R_t trajectory present a jaggedness – this coincides with the date of the implementation of the first set of interventions in the modelled population, which significantly alter the transmission rates that were used up until that point in time.

We also observe that the inferred curves for both the number of new infections and daily number of deaths fit the observed data well. Unlike the PHE model, the inferred trajectories for infection incidence, deaths and time-dependent R_t are much smoother. Importantly, the sudden drop in the values of the reproduction number, which we associate with intervention coming into effect, occurs around the 16th of March 2020, which is a week earlier compared to the institution of the lockdown in the UK, but around the same time as the first social distancing measures have been applied. This is followed by a slow increase towards the end of the model time-frame, after which it remains low until the end of the simulation window.

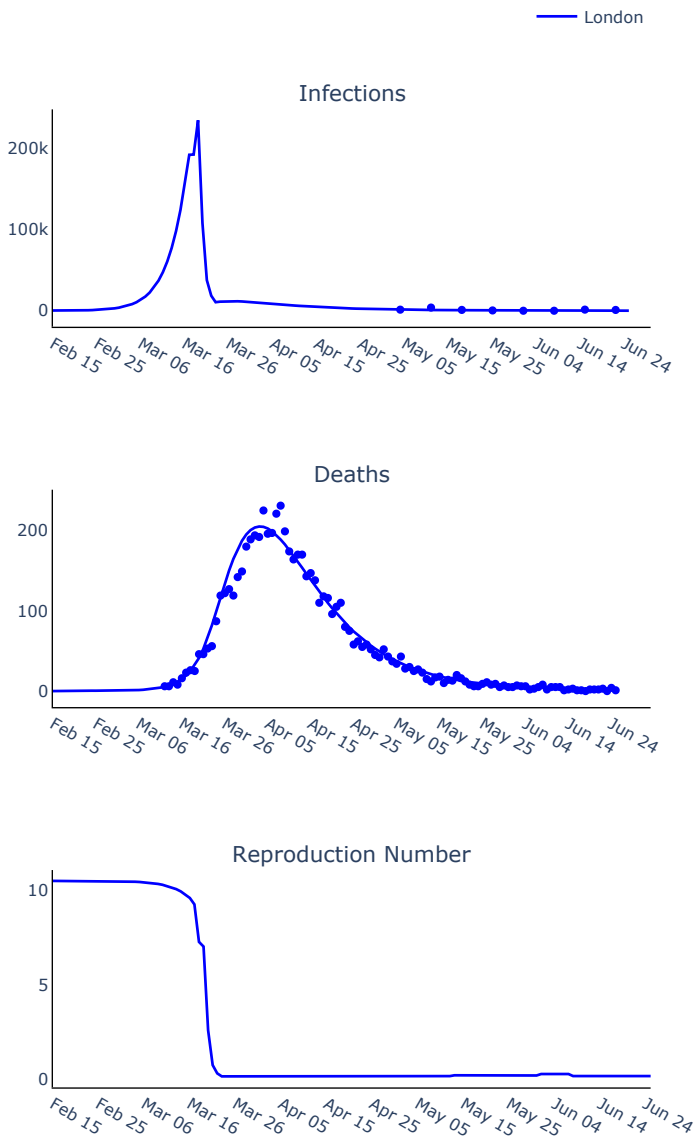


Figure 4.8: *Trajectories of the daily total number of new infections, deaths and instantaneous reproduction number R_t for London produced using the Roche model. The dots represent the observed number of positive tests scaled to the population (first panel) and the observed number of deaths (second panel).*

This implies that the full-lockdown had less of an impact in suppressing the epidemic compared to the earlier imposed measures; this is also supported by the fact that the R_t number does not increase when the lockdown is lifted – contrasting to what we observed in the PHE model case.

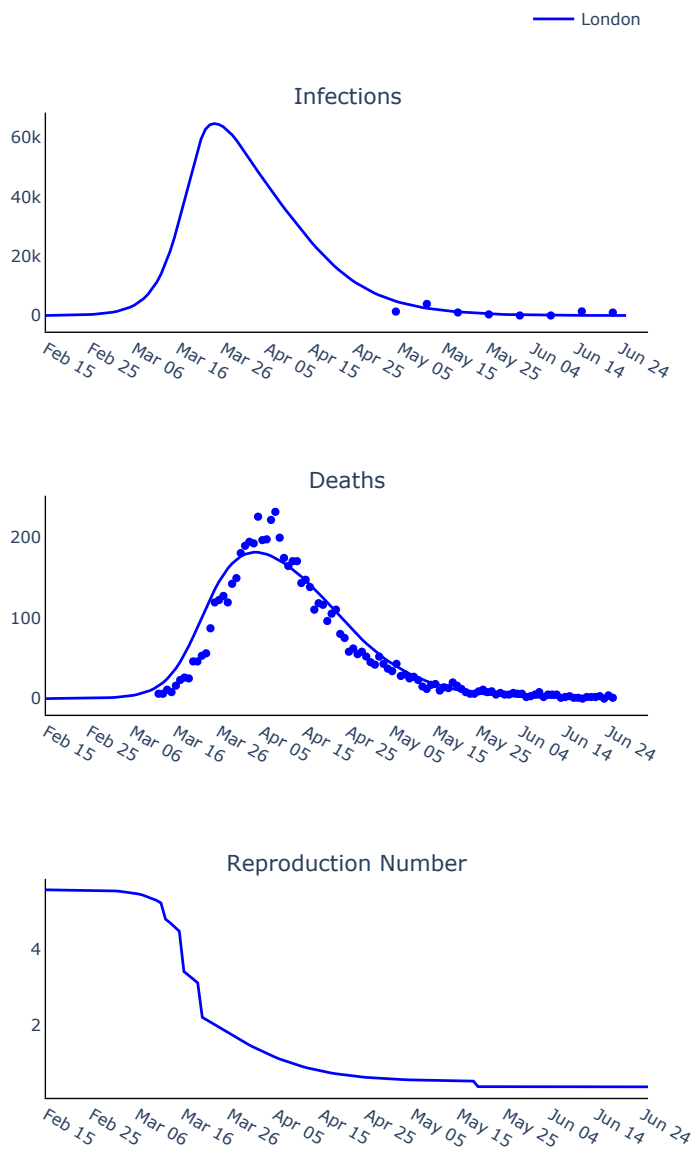


Figure 4.9: *Trajectories of the daily total number of new infections, deaths and instantaneous reproduction number R_t for London produced using the Warwick-Household model. The dots represent the observed number of positive tests scaled to the population (first panel) and the observed number of deaths (second panel).*

4.3.6.3 The Warwick-Household Model

In Figure 4.9, we present the trajectories for the total daily number of new infections (first panel) and deaths (second panel), as well as the trajectory of the time-varying reproduction number of the Warwick-Household model when the recovery rate γ and asymptomatic reduction in transmission parameter τ are fixed at the values obtained from running the CMA-ES algorithm, that is $\gamma = 0.35$ and $\tau = 0.5$. A value for τ of 0.5 indicates that the infectiousness associated with asymptomatic cases is half of that for the symptomatic cases.

The dots in each panel represent respectively:

1. the proxy of the observed number of infections, computed from the number of positive tests recorded in the REACT1 study [Nicholson et al., 2021; Eales, 2020], scaled to the population size, and
2. the observed number of deaths recorded by ONS [GOV.UK].

Similar to the Roche model, the curves of the number of daily new infections deaths and R_t generated by the Warwick-Household model run much smoother than their counterparts for the PHE model. This is mainly due to the fact that unlike the PHE model, both the Roche and the Warwick-Household models account for fewer fluctuations in the level of contacts – they only adapt their parameters with alterations of NPIs strength, compared to natural weekly fluctuations in mixing patterns as was the case in the PHE model. Regardless, the inferred curves form the Warwick-Household model fit the observed data well.

The reproduction number has a sharp drop around the 16$^{\text{th}}$ of March, around a week earlier then the establishment of the lockdown, but around the same time as the first social distancing measures have been applied. Moreover, similar to the Roche model and unlike the PHE model, the R_t remains low even after the interventions are lifted towards the end of the simulation window.

4.3.7 Parameter Inference

We proceed now to validate the optimisation findings using MCMC sampling. The method we employ for the sampler is called the Haario-Bardenet model, which is an adaptive covariance matrix algorithm [Haario et al., 2001; Johnstone et al., 2016]. We use three parallel chains starting at the optimised parameter estimates and run them for 60000 iterations each to ensure algorithm convergence – the value of $\hat{R}$ we obtain for each of the three models is within the desired margins, i.e. ≈ 1.0. Since each new sample depends on the previous position of the chain, we discard the first half of the chain samples as the burn-in period in order to forget the initial conditions. We use the remaining 30000 samples to plot the mean projected trajectories of the daily incidence, deaths and the time-dependent instantaneous reproduction number R_t as well as the centred 95% credible intervals for each of them.

All further auxiliary plots of the chains and histogram of samples show that the three chains

converge and therefore is a sign of a correct inference; in Appendix F2, we include a series diagnostic plots for the PHE model inference framework, e.g. plots for autocorrelation of MCMC chains and histograms of each estimated model parameters, to showcase proper algorithm convergence – for example, as there there is an overlap for all three chains of the PHE inference routine, this indicates there is no codependency for the chosen parameters and therefore the parameters can be correctly inferred.

4.3.7.1 The PHE Model

Figure 4.10 illustrates the uncertainty in the trajectories of the daily total number of new infections, deaths and instantaneous reproduction number R_t for synthetic data. In this case, we run a model for which we set $R_0 = 2.35$ and $\beta_t = 1, \forall t$, , while all other model parameters are chosen as in section 4.3.5.1, to sample the number of deaths, using the PHE model class method for each day of the simulation window – we take it here to be 90 days, between 15$^{\text{th}}$ of February and 15$^{\text{th}}$ of May 2020. Similarly to the optimisation, we indicate through dots the epidemic data used to infer the values of the free parameters of the PHE model:

- the number of daily new infections (first panel) and
- the number of deaths (second panel)

generated by running a forward simulation of the model as described above.

The inference results match the synthetic death data: the estimated mean value of the initial reproduction number is given by $\hat{R}_0 = 2.37$ ($[2.34, 2.39]$ for the 95% credible interval), and the mean estimates for the weekly $\hat{\beta}_t$s parameters range between 0.42 and 0.80 ($[0.35, 0.49]$ and $[0.72, 0.89]$ respectively for the 95% credible interval), with the a mean log-normal variance $\hat{\sigma_\beta} = 0.11$ ($[0.08, 0.15]$ for the 95% credible interval). The uncertainty measure in each of the three model outputs increases as we approach the end of the time window. This is to be expected as small differences in the trajectories increment to much larger deviation the further we go.

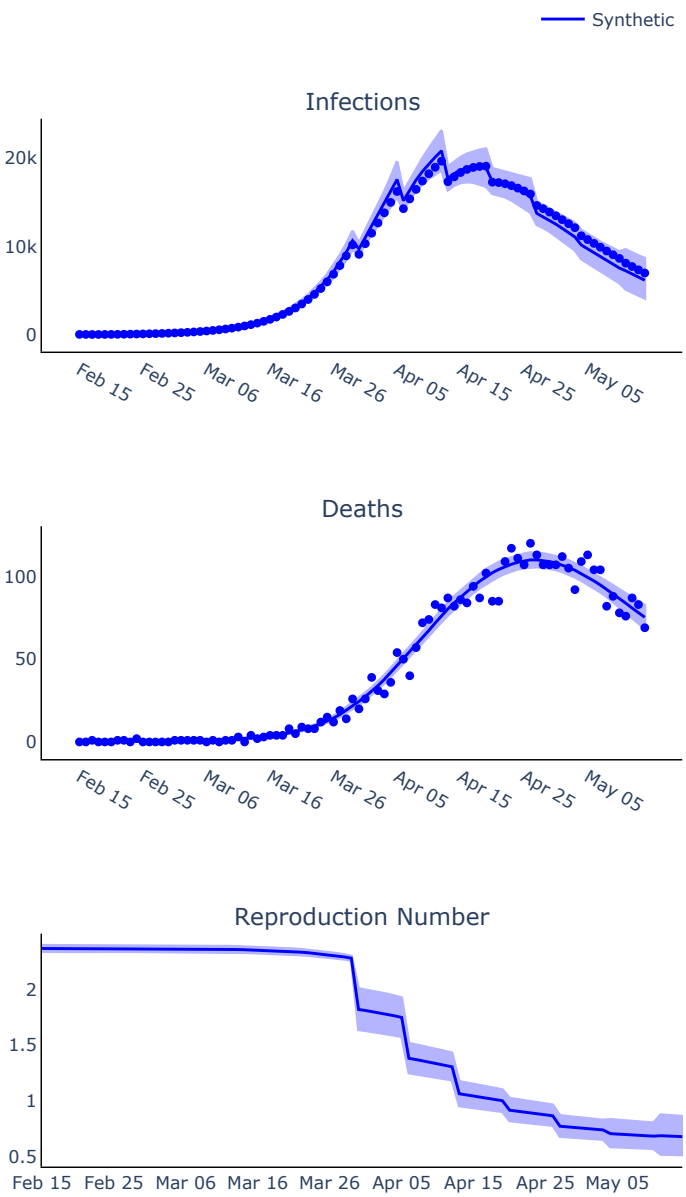


Figure 4.10: *Mean and* 95% *confidence regions of trajectories of the daily total number of new infections, deaths and instantaneous reproduction number* R_t*. These trajectories are produced using the PHE model, which was fitted to synthetic death data. The dots represent the observed number of daily new infections (first panel) and number of deaths (second panel).*

4.3.7.2 The Roche Model

We summarise our results in Figure 4.11, in which we fit the Roche model to synthetic death data. To generated the daily number of deaths used for parameter fitting, we forward-simulate the Roche model for a period of 90 days, between the 15^{th} of February and 15^{th} of May 2020, similar to what we have done in the case of the PHE model. We instantiate

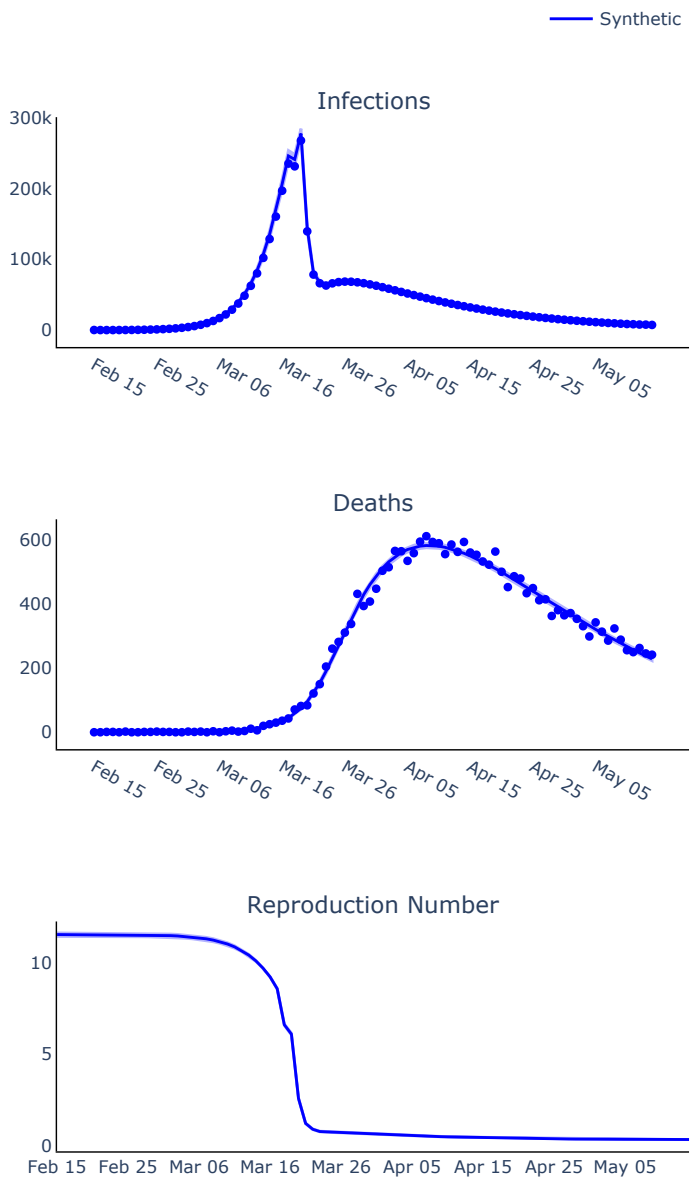


Figure 4.11: *Mean and* 95% *confidence regions of trajectories of the daily total number of new infections, deaths and instantaneous reproduction number* R_t*. These curves were produced using the Roche model, which was fitted to synthetic death data. The dots represent the observed number of daily new infections (first panel) and number of deaths (second panel). Due to the very small number of parameters that we can infer for the Roche model, the confidence regions around the trajectories of the number of infections, deaths and* R_t *are extremely narrow and therefore difficult to discern.*

the Roche model at the minimum transmission rate $\beta_{\min} = 0.228$ and relative increase in super-spreader transmission effect $b_{ss} = 3.11$; all other model parameters are chosen as in section 4.3.5.2. The number of daily new infections (first panel) and deaths (second panel) thus generated are indicated on the plots through dots. The hallmark of an effective

parameter inference routine consists in our ability to accurately retrieve the values of the two parameters we aim to estimate – β_{min} and b_{ss}.

As we can see from Figure 4.11, the inferred trajectory of the daily number of deaths match the synthetic death data we fit the model to. In addition to the diagnostic plots which indicate algorithm convergence, we observe that the mean estimates of the inferred parameters are close to their true values: we obtain the following mean estimates for $\hat{\beta}_{\text{min}} = 0.23$ ($[0.22, 0.23]$ for the 95% credible interval) and respectively $\hat{b}_{ss} = 3.22$) ($[3.13, 3.29]$ for the 95% credible interval) which indicates the correctness of our inference framework. Just like in the case of the optimisation routine, we observe that around the date of 13$^{\text{th}}$ of March 2020 both the number of daily new infections and the R_t trajectory present a jaggedness – this coincides with the date of the implementation of the first set of interventions in the toy population, which significantly alter the transmission rates that were used up until that point in time.

Since we apply the same interventions as those implemented during the first wave of the COVID-19 epidemic in the UK, the trajectory of the reproduction number will resemble in shape the R_t trajectory we identified in the optimisation procedure; that is, there exists a drop in the value of R_t occurring roughly around the 16$^{\text{th}}$ of March 2020 mark, one week before full lockdown is implemented.

4.3.7.3 The Warwick-Household Model

In Figure 4.12, we plot the mean trajectories and their respective 95% credible intervals of the number of new infections, deaths and the reproduction number resulted from fitting the Warwick-Household model to synthetic death data.

The death data we fit the model to were generated by running a forward simulation of the Warwick-Household model for 130 days, between the 15$^{\text{th}}$ of February and 25$^{\text{th}}$ of June 2020; the parameters of the Warwick-Household are taken to be exactly as the ones used in in section 4.3.5.3. The true values of the γ and τ for the toy population are fixed at 0.2 and 0.4, respectively. We indicate through dots the daily number of new infections and (first panel of Figure 4.12) and deaths (second panel of Figure 4.12) generated by running a forward simulation of the Warwick-Household as described above.

Both the summary statistics of the retained samples, as well as the plotted curves in the sec-

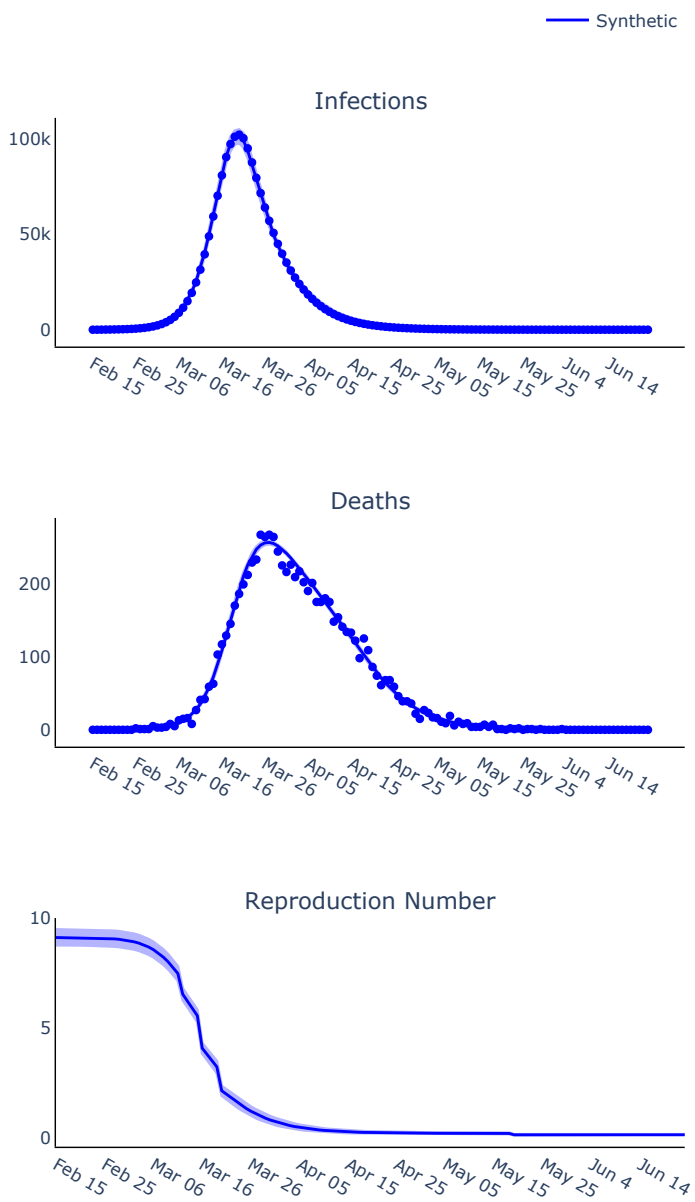


Figure 4.12: *Mean and* 95% *confidence regions of trajectories of the daily total number of new infections, deaths and instantaneous reproduction number* R_t *for synthetic data. These curves were produced using the Warwick-Household model, which was fitted to synthetic death data. The dots represent the observed number of daily new infections (first panel) and number of deaths (second panel). Due to the very small number of parameters that we can infer for the Warwick-Household model, the confidence regions around the trajectories of the number of infections, deaths and* R_t *are extremely narrow.*

ond panel of Figure 4.12 illustrate a good fit, with the mean inferred $\hat{\gamma} = 0.19$ ($[0.17, 0.21]$ for the 95% credible interval) and $\hat{\tau} = 0.38$ ($[0.33, 0.43]$ for the 95% credible interval), which are close to the original values that generated the observed death counts.

As a final point, we observe that for this synthetic dataset, the average initial reproduction

number is determined to be around 10, which is almost twice as large as the one computed by the optimiser. However, this can be explained by the lower underlying value of the recovery period which almost doubles from $\frac{1}{\gamma} = 2.85$ days to $\frac{1}{\gamma^{\text{est}}} = 5$ days.

4.3.8 Scenario Model Comparison

In this section, we analyse whether the compartmental ODE models introduced back in section 4.2.1 are able to accurately determine the impact of the non-pharmaceutical interventions implemented during the first COVID-19 epidemic wave in the UK, when tested against identical counterfactual scenarios. The time-window considered for this analysis is from 15th of February to 25th of June 2020, which covers the entirety of the first wave; in Figure 4.13, we mark the main important dates of the applications of the NPIs, and which models they are used by pertain to.

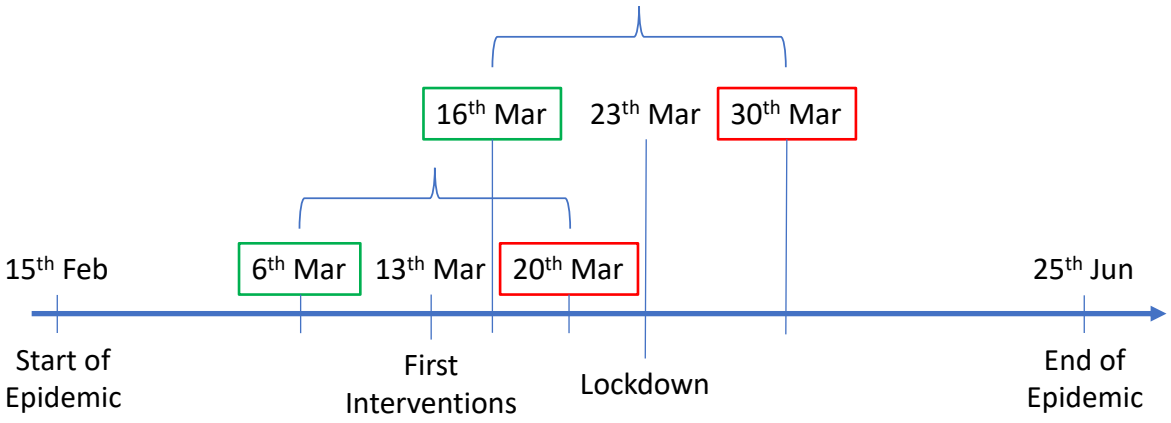


Figure 4.13: *Timeline of Interventions. In green we mark the one-week-earlier onset date for of the NPI, and in red the date of the one-week-later onset of the NPI. The lockdown date is exclusively used by the PHE model; for the Roche and the Warwick-Household models, changes in NPIs occur from onwards of the corresponding date in the 'First Interventions' cluster.*

We use the PHE, Roche and Warwick-Household models to assess four scenarios:

1. **lockdowns and other NPIs are imposed 'on-time', i.e. at the same time as NPIs were deployed during the first UK epidemic wave**: The time-dependent contact matrices are affected by changes in the population's mobility; depending on the model, the fluctuations in the number of contacts are represented through a varying Google mobility data factor [Google] multiplied to the baseline POLYMOD contact matrix

[Prem et al., 2021] (PHE model), or different levels of interventional strength (Roche and Warwick-Household model, though through different mechanisms).

2. **no NPIs are used**: In this case, the contact matrices remain at the level from the beginning of the simulation window; in the case of the PHE and the Warwick-Household model, the contact matrix is selected to be the same as the one registered on 15$^{\text{th}}$ of February 2020, while the Roche model uses instead the baseline POLYMOD contact matrix for the UK [Prem et al., 2021], with adapted age groups.

3. **delayed NPIs**: With this scenario, the same level of interventions are applied as in the on-time scenario, but they are instantiated a week later, e.g. if the lockdown in the UK was mandated on the 30$^{\text{th}}$ of March instead of the 23$^{\text{rd}}$ of March 2020. In the case of the PHE model, that translates to considering the same contact matrix for the first four weeks of the simulation window; then for week 5, we reuse the same contact matrix as in week 4, and continue from then on for the weeks $k > 5$ with the contact matrix for week $k - 1$. For the Roche model and the Warwick-Household model, we instead delay the application time of interventions to a week later after the 13$^{\text{th}}$ of March 2020 (see Figure 4.13 for the NPIs timeline overview).

4. **earlier NPIs**: This is similar to the delayed NPIs case, but instead now interventions are imposed one week earlier than the initial scenario. For the PHE model, we proceed in the same manner as before until week 4, after which we continue to use the contact matrix weighted with the Google mobility factor registered for week $k + 1$ in the original data to represent the number of contacts matrix on week $k \geq 5$ (see Figure 4.13 for the NPIs timeline overview).

The first step in the comprehensive analysis of these four scenarios is to fit each of the three models – the PHE, the Roche and the Warwick-Household model respectively using the same sets of observed data. As these observed values correspond to the original 'on-time' time-frame, we run our parameter inference algorithm according to the first scenario exclusively, that is when the non-pharmaceutical interventions have been implemented at the same times as they were during the actual first wave of the COVID-19 epidemic in the UK. This way, we generate estimates for the model parameters, which we than use to run forward simulations of the models outputs for each of the three remaining scenarios. Finally, we perform side-by-side comparison of the mean and 95% confidence regions of the trajectory

of the number of new infections, deaths or reproduction number for each of the scenarios in order to determine if the three models will predict different epidemic behaviours when NPIs are no longer implemented identically to the original timeline.

As mentioned previously in section 4.3.4, we decided on using exclusively ONS death data [GOV.UK] to fit the models for the on-time intervention scenario for all models. This is because typically cases are under-reported or there exist delays between infection onset and reporting of the case; moreover, while using a frequency of infection measure instead from a large-scale infection survey such as was the case with REACT1 study data [Nicholson et al., 2021] would lead to sensible approximate values of the incidence of infection, the resulting dataset is too sparse (one each week or so) and too late-occurring in the simulation window (well after the peak of infection is predicted) to be informative. Indeed, if we were to fit our models to the positive tests data from the REACT1 study, we would only end up overfitting the tail, and overall arriving at a worse overall fitting of the death curve, which is the output we are most interested in accurately replicating.

We now proceed with the actual model comparison. For each of these analyses, we run the Haario-Bardenet inference routine for 60000 iterations fitted with only death data, discard the first half of them and proceed to thin the resulting MCMC chains out until we are left with 9000 samples for each of the three models in section 4.2.1 – we do this in order to further reduce the autocorrelation of the remaining sampled parameter values form the MCMC routine. For each of the samples, we run our PHE, Roche and Warwick-Household model instantiations to produce timelines of the number of daily new infections, deaths and the instantaneous reproduction number R_t for each of the four scenarios illustrated in section 4.3.8. For the non-inferrable parameter, we use the values in section 4.3.5.

4.3.8.1 Model comparison: NPIs – on or off

In Figure 4.14 we plot the mean trajectory and the 95% confidence intervals for each of the three PHE, Roche and Warwick-Household model outputs. In red we present the results for the hypothetical scenario in which no interventions are implemented to prevent the spread of the virus, keeping the contact matrix equal to that instantiated at simulation time zero. We represent in blue the trajectories of the outcome of interests when NPIs have been implemented as they were between the 15^{th} of February and 25^{th} of June 2020. The grey dots represent the observed number of positive tests scaled to the population (in the first panel)

– we use census [ONS] and REACT1 study data [Eales, 2020; Nicholson et al., 2021] to compute those datapoints as we have done previously – and number of deaths (in the second panel).

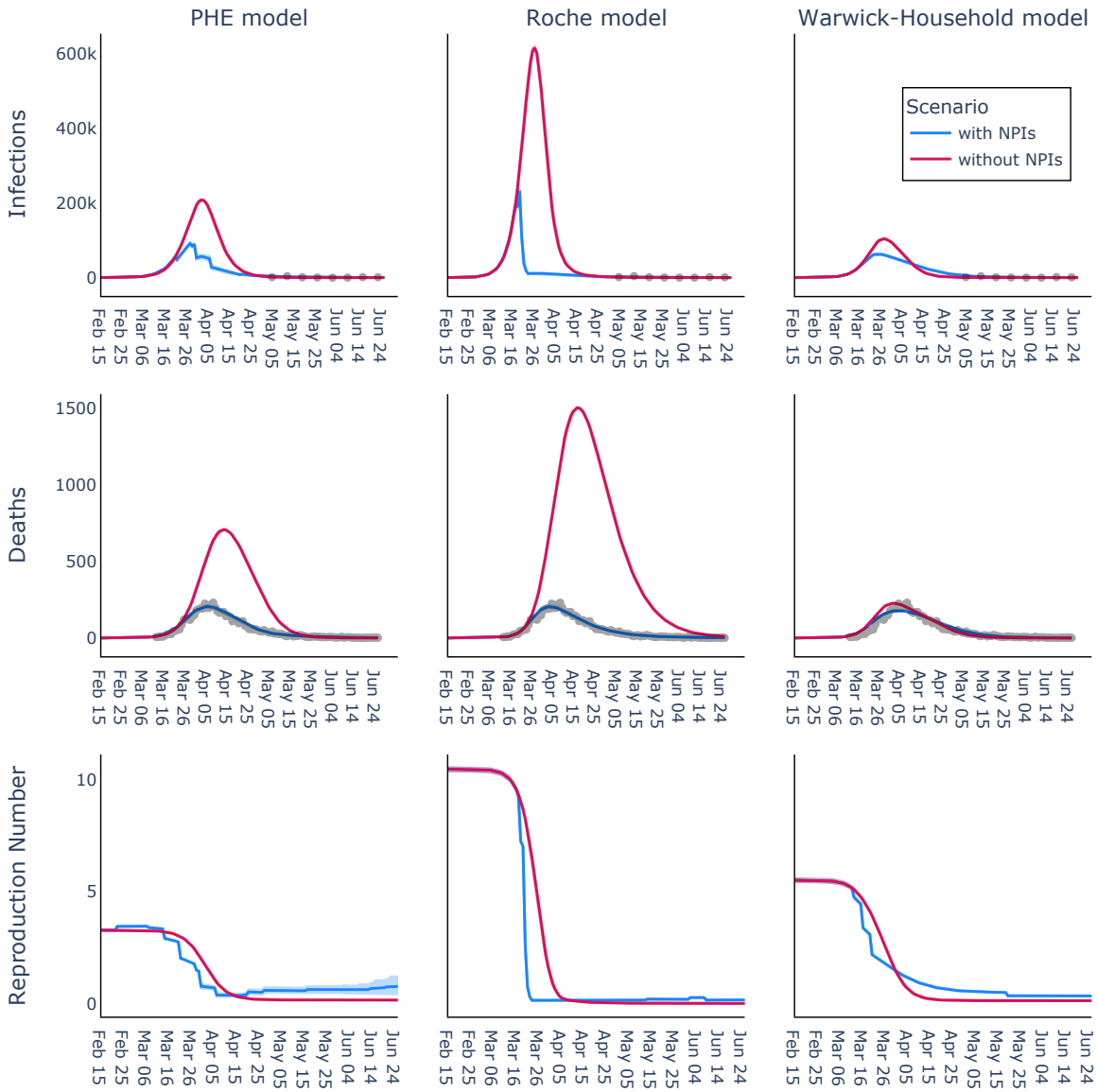


Figure 4.14: *Mean and 95% confidence regions of trajectories of the daily total number of new infections, deaths and instantaneous reproduction number R_t for London in two scenarios: (red) if no non-pharmaceutical interventions were imposed and (blue) what actually happened. The grey dots represent the observed number of positive tests scaled to the population (first row) and number of deaths (second row). Due to the very small number of parameters that we can infer for the models, the confidence regions around the trajectories of the number of infections, deaths and R_t are extremely narrow.*

As expected, the trajectories produced by forward simulating the with-interventions on-time

scenario for each of the three models follow the observed data really well. The red and blue curves match perfectly until the 23rd of March 2020 for the PHE model – which corresponds to the date when lockdown was instituted in the UK, and respectively until the 13th of March 2020 for the Roche and Warwick-Household models – which itself corresponds to the date when the first stringent interventions started to be implemented in the UK. In the no NPIs scenarios, the peak of the daily number of new cases predicted by the PHE model is twice as big compared to the NPIs scenario and occurs slightly later (around 1st of April, compared to 26th of March). Similarly, we observe different behaviours in the number of deaths curve for each of the three compartmental models. For the PHE model, the number of deaths observed when contacts remain identical throughout the simulation time-frame is predicted to be more than three times larger compared to the situation when NPIs have been applied in the UK. We also can observe a lag between when the peak in the number of deaths occurs for the two scenarios considered: while in the on-time with-intervention case the largest number of deaths is observed around late March, in the case of no-interventions, unchanged contacts, the maximal number of deaths is reached in mid-April. For the Roche model a similar situation to the PHE model is outlined: the model predicts that in the counterfactual scenario the number of deaths observed will be nine times larger than the peak in daily number of deaths in the case when NPIs are applied as in the original situation, with a slightly bigger time difference between the location of the peaks. These results suggest that interventions had an important impact in mitigating the spread of the epidemic and reducing the severe disease outcomes during the first wave of the UK COVID-19 pandemic.

However, a completely different situation emerges when we use the Warwick-Household model instead to study the epidemic. In this case, the daily death curves for both the on-time and no intervention scenarios are very close to one another; indeed, the peak of the number of deaths in the case when no NPIs are in place and the contact matrix remains unchanged throughout the simulation occurs a few days earlier, and has a value closer to the value of the peak in the observed number of cases, compared to the fitted death curve. This result would imply that in the context of a population behaving according to the Warwick-Household population dynamics, the interventions have little to no impact in mitigating severe disease outcomes.

This differs significantly from the behaviour of the epidemic reported in the original study [Keeling et al., 2021], where in the case when no interventions are applied a significant in-

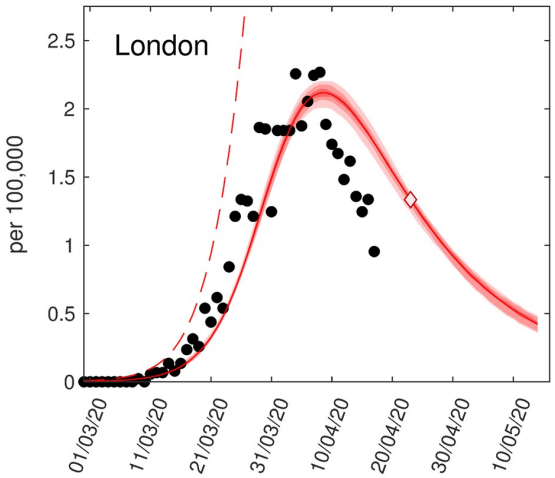


Figure 4.15: *Mean and* 95% *confidence regions of trajectories of the daily total number of deaths for London in two scenarios: (dotted red line) if no non-pharmaceutical interventions were imposed and (solid red line) what actually happened. These trajectories were predicted using the original Warwick-Household model. The black dots represent the observed number number of deaths. This figure corresponds to [Keeling et al., 2021, Figure 2(e)].*

crease in the number of daily reported deaths is predicted – see dotted trajectory in Figure 4.15 which is taken directly from the original paper. An explanation for this discrepancy is that our reimplementation of the Warwick-Household model is not exactly the one that was used to originally generate Figure 4.15. Firstly, the original result and technical paper discussing the Warwick-Household model and the available source code (which is not open-access) describe very different parameter inference mechanisms. For example, in the code and from discussions with the authors, the Warwick-Household model uses a set of pre-defined of parameter relationships to determine the value of most of the parameters of the model – this means that only a small number of parameters need to be inferred as most of them can be deduced from these equivalences; however, these equations are not referenced or mentioned anywhere in the papers introducing the Warwick-Household model. Additionally, these articles clearly state that the parameters q_H, q_S, q_W and q_O that calibrate the full-lockdown contact matrices are kept fixed throughout and have predefined values – this contrasts with the online code where these quantities are varied over the course of the first COVID-19 wave in the UK, which is again a contradiction. These differences essentially give rise to two distinct official versions of the Warwick-Household model, which significantly complicates our task of reimplementing the model, as we cannot be entirely certain which version to adhere to. And while it is clear that our reimplemented version of the

Warwick-Household model is not the same as the original and therefore yield very different results, the issue at hand still remains that the model is not reproducible in the first place – which puts into question the usefulness of the approach altogether in modelling a future similar epidemic.

In the third row of panels in Figure 4.14, we plot the instantaneous reproduction number R_t trajectories for each of the three compartmental models considered (PHE, Roche and Warwick-Household) for both of the scenarios considered in this section. For each of the models we observe that the reproduction number trajectory decreases more slowly for the case when no interventions are implemented and contacts are kept fixed when compared to the factual scenario. We observe that the PHE and the Warwick-Household models predict an initial reproduction number between 3 and 5.5, which aligns with clinical studies [Liu et al., 2020]; this contrasts with the Roche model which instead estimates an initial value of R_t over 10, much larger than predicted by any of these studies. This either implies that either the initial variant of SARS-CoV-2 was much more transmissible than initially thought, or the Roche model is unable to properly model certain aspects of the disease transmission, which in turn lead to an over-estimation of the true force of infection.

Moreover, we observe that for the PHE model the trajectory of the R_t in the counterfactual scenario diverges from the one corresponding to the scenario with NPIs at an earlier date compared to when interventions are first implemented; this is because contact patterns started shifting in the UK before any official interventions were implemented on the 13$^{\text{th}}$ of March 2020 – compared to the Roche and Warwick-household models, where differences in the two R_t trajectories begin from that date onwards. Additionally, we notice that for the Roche model the reproduction number plummets below one even in the absence of interventions roughly around the same date as when interventions were applied – a similar behaviour can also be observed for the other two models, but to a lesser degree. This phenomenon can be explained by the rapid depletion of the number of susceptibles that accompanies the high force of infection predicted by the Roche model; that is, all available susceptible individuals go on to become exposed and then infected very quickly and we are left with very few other individuals on pass the diseases onto.

4.3.8.2 Model comparison: NPIs are applied at different times

In this section we aim to understand how the implementation time of the different non-pharmaceutical interventions impact their effectiveness. In Figure 4.16 we plot the trajectories of the number of new infections, daily deaths and the reproduction number for each of the three models presented in section 4.2.1 for three choices of implementation times: in yellow we represent the 'on-time' scenario, when interventions have been implemented on the same dates as during the first wave of the COVID-19 epidemic in the UK – this also denotes the trajectories for the model parameterisation that the observed death data has been fitted to; the red lines indicate the predicted trajectories when we run the fitted models with NPIs occurring 7 days latter compared to their timing in the original scenario; we proceed similarly for the blue lines, which indicate a similar change in the timing of NPIs, this time a week earlier compared to the original timeline.

Similar to Figure 4.14, the simulation captures the overall behaviour of the seven PHE regions in the UK, between 15th of February and 25th of June 2020. The grey dots indicate the observed data used for fitting the model: - death data from the Office of National Statistics [GOV.UK] (second panel) and incidence of new infections identified from the REACT1 study data [Nicholson et al., 2021; Eales, 2020] and then scaled to population size using census data [ONS].

In the first panel of Figure 4.16 we compare the trajectories of the daily incidence of infection for each of the three models. As expected, for the Roche model we observe that the later NPIs are introduced in the population the larger the overall number of infections becomes – hence, an earlier implementation of lock-down and similar social-distancing measures is predicted to have a strong impact in reducing the number of infections. However, this conclusion is not ubiquitous. According to the PHE model, while the earlier implementation of interventions would have lead to a smaller initial increase, the curve of incidence for this case is flattened and by March 26th the earlier-NPIs scenario overtakes the factual 'on-time' one, leading to a larger number of infections registered overall in the alternative implementation; a possible explanation for this phenomenon would be that the earlier implementation of NPIs has lead to a slowdown of infections in the adult population – the population group with the most number of daily contacts and therefore the biggest drivers of transmission. The lower number of infections in this group also translates into fewer recovered adults

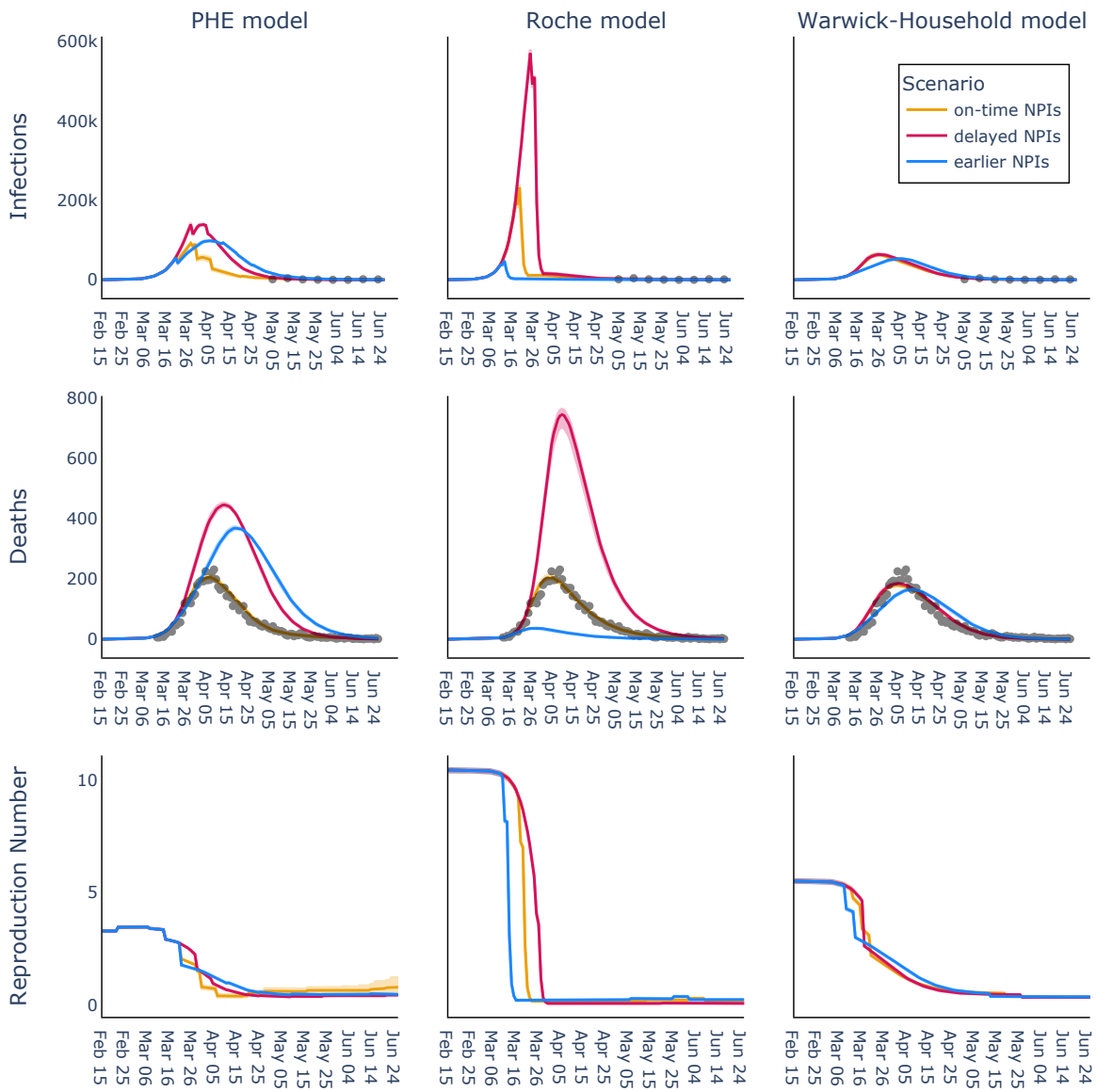


Figure 4.16: *Mean and* 95% *confidence regions of trajectories of the daily total number of new infections, deaths and instantaneous reproduction number* R_t *for London in three scenarios: (red) if non-pharmaceutical interventions are imposed* 7 *day later than in the original timeline, (blue) if non-pharmaceutical interventions are imposed* 7 *day earlier than in the original timeline, and (yellow) what actual happened, i.e. the original timeline of NPIs. The grey dots represent the observed number of positive tests scaled to the population (first panel) and number of deaths (second panel). Due to the very small number of parameters that we can infer for the models, the confidence regions around the trajectories of the number of infections, deaths and* R_t *are extremely narrow.*

(and hence unable to transmit the virus later on), which in turn may lead to higher levels of infections overall, especially in the elder, more vulnerable age groups; this can also explain the higher number of deaths observed in the earlier-NPIs scenario for the PHE model

compared to the factual one. As expected, for the delayed intervention case there is an increase in the predicted number of infections registered daily compared to the original 'on-time' scenario. These results suggest that if interventions had been applied either earlier or later the impact these NPIs would not have been as potent. For the Warwick-Household model, the trajectories for the 'on-time' and delayed are overlapping, while the 'earlier-NPIs' scenario produces a slightly flatter curve, but accounting for less infections overall – this suggests that in terms of mitigating epidemic growth, an earlier implementation of the non-pharmaceutical interventions would have reduced this output of interest which, while similar to the effect observed in the Roche model, is much less pronounced.

A similar situation occurs when considering the number of deaths. The Roche model predicts that a one-week difference in the application of social-distancing and other non-pharmaceutical interventions has a four-fold effect in the trajectory of the daily numbers of deaths: an earlier peak of less than 50 individuals for the 'earlier-NPIs' scenario, around at most 200 deaths a day in the original timeline and a later peak of up to 800 deaths per day in the case of a one week delay in lockdown. The Warwick-Household model paints a similar picture – that expediting the application of NPIs by a week leads to a decrease in overall number of deaths; however, there are almost no differences compared to the 'on-time NPIs' prediction when interventions are applied one week later. This suggests that if the epidemic followed the same dynamics as those described by the Warwick-Household model, by only applying the first NPIs from the 13^{th} of March, we have missed the window of opportunity which would have led to larger reductions in the number of realised deaths.

Lastly, the trajectories of the reproduction number for each of the models indicate roughly similar overall behaviour: earlier NPIs lead to earlier drops in R_t. However, due to the noisy and rapid changes in the strength of interventions, for the PHE and the Warwick-Household model the yellow line indicating the original scenario trajectory drops below the blue one corresponding to the earlier application of NPIs, which can explain the unexpected results of the corresponding infection incidence and death panels in rows 1 and 2 of Figure 4.16.

4.4 Discussion

In this study we have focused on understanding how the different compartmental models compare to one another in terms of their outputs of interest. Our aim in doing so was to

determine whether we arrive at analogous conclusions about epidemic trends and the effect of mitigating non-pharmaceutical interventions, when considering the predictions of each of the three models. Significant differences in the trajectories of deaths and infection would indicate that we cannot determine which, if any, of these models can accurately model the true epidemic dynamics; further work is therefore needed in determining which of these models might be suitable to be used in the decision making process. This in turn has implications for how we might approach modelling the epidemic dynamics when the next pandemic strikes.

Indeed, as we have seen in both sections 4.3.8.1 and 4.3.8.2 analyses, there is no consensus among the three models on what the effect of NPIs on the predicted model outcomes really is. While both the PHE and Roche models predict there is a substantial reduction in both number of deaths and infections due to interventions, we arrive at very different conclusions on the impact of changing the time of NPI implementation on their effectiveness. If the epidemic followed the dynamics of the PHE model, we predicted that starting applying the lockdown and all other NPIs either a week earlier or a week later lead to both more deaths and more new infections, compared to the original scenario; this is very much unlike the Roche model, for which expediting NPIs by one week leads to a four-fold effect in reducing severe disease outcomes.

These observations completely change when using the Warwick-Household model. For an epidemic which follows the dynamics of this model, the earlier application of interventions leads to only a small reduction in the number of deaths and cases; however, if we instead delay NPIs by a week, we again observe an increase in the number of new infections or deaths. This surprising result can be either explained by interventions having been implemented too late in the original timeline to be effective, or by the Warwick-Household model being unable to properly capture the key transmission features of the epidemic.

It is important to remember that all of these models are used to study and analyse the same disease at the same snapshot in time; therefore, we would expect to arrive at similar conclusions using each one of these three models. Therefore, it becomes clear that in the case of these three compartmental models the conclusions we are able to draw using these models are very much model-dependent and cannot all be indicative of the true behaviour of the epidemic. This result is by no means specific to these three chosen models: for example, [Purkayastha et al., 2021; Silal et al., 2016] looks at five (respectively six) other compartmental ODE type models and similarly demonstrates their non-equivalence, while other

studies such as [Siegenfeld et al., 2022] look at precisely how incorporating population heterogeneity in the different model parameter assumption affects the predicted epidemic outcomes, with similar findings being observed for other modelling formalisms such as renewal equations [Brockhaus et al., 2023], individual-based [Gibson et al., 2018] and agent-based models [Özmen et al., 2016]. However, such field-wide validation and assessment of modelling practices are not common. This perpetual oversight leads to the over-saturation of the field with new published models that are difficult to replicate and which produce divergent results when models share similar assumptions. These lax community practices undermine the already fragile trust in modelling among policymakers and can cause concern to the general public. This ultimately leads to smaller and smaller levels of intervention compliance and overall reduces the effectiveness of these interventions – which itself leads to increasingly stronger interventions being put in place to mitigate this effect and propagating this vicious cycle.

Indeed, this model non-equivalence phenomenon is not particularly uncommon in the general field of mathematical biology; when choosing a statistical model for any particular problem we have to keep in mind that *'all models are approximations'* [Box and Draper, 1986]. Coupled with the inherent flexibility of this modelling framework, the search for an ODE model structure that more precisely encapsulates the transmission dynamics could prove futile, especially as it is not immediately clear which epidemic features should or should not be accounted for. In other fields, such as machine learning, weak predictors can be used speculatively to produce more robust predictions through the use of adaptive boosting, or *adaboost* [Schapire, 2013]. In adaboost, the imprecise predictions of these weak but cheap-to-run predictors are averaged to produce better results. Going forward, extensive research needs to be done to improve our current methodologies for compartmental modelling of epidemic dynamics, either by (i) identifying an approach that lets us aggregate the results of simple, vetted ODE models, in the manner resembling adaboost, to increase the accuracy of epidemic outcome predictions or (ii) by developing a check-list of standard analyses and validation datasets that we need to test models against before using these models in any decision making process.

In conclusion, we should employ greater care when using ODE compartmental models to study epidemic dynamics, especially in the case of novel pathogens for which there is no extensive knowledge of the true underlying population behaviour or the role that household

transmission plays in the overall epidemic dynamics. Models should not be assumed to be automatically substitutable for one another, especially when changing the conditions under which the models have been fitted; additionally, while there can exist some overlap in the predictions of outcomes-of-interest across different compartmental models, these are more a consequence of parameterisation choice (that is, we chose parameter values which satisfy the equivalence relations existing between the parameters of the models), rather than due to the compared models being equivalent to one another.

Conclusion

Key research questions

Over the course of this thesis, we have explored the impact that including different assumptions about the heterogeneity of disease transmission has on the predicted epidemic outcomes produced by the models, whether they refer to the number of new infections, deaths or the trajectory of the observed reproduction number. We looked at both renewal equation and compartmental ODE modelling frameworks.

In the first three chapters of the thesis, we focused on renewal-type models and how integrating different sources of heterogeneity in transmission dynamics impacts the predicted reproduction number trajectory. We were specifically interested in determining:

- whether we can derive a simple criterion for the long-term fate of an epidemic for multiple-groups population models that does not require any additional quantity to be inferred, as seen in similar studies within the literature [Green et al., 2022; Parag and Obolski, 2023];
- if there are any circumstances in which we need to account for multiple population groups when using renewal equation models instead of using the simpler one-group approach;
- what is a sufficient level of heterogeneity information to observe an improvement in predictions;
- how to identify in which population groups to target interventions in order to ensure their most efficacious implementation.

In comparison, the analyses presented in Chapter 4 concern compartmental ODE models, which are characterised by their flexible structure, on top of which different features of disease transmission can easily be added (for example asymptomatic infections [Lemenuel-Diot et al., 2020; Bosetti et al., 2021] or vaccination [Moore et al., 2022]). Indeed, in the case of this model formalism our goals centre on identifying how much we can rely on the conclusions to be drawn from fitting compartmental models. To this end, we aimed to

answer the following three questions:

- If we use different compartmental models to analyse the same epidemic, by fitting to identical datasets, do we arrive at similar conclusions on the efficacy of different non-pharmaceutical interventions (NPIs)? That is, are these compartmental models equivalent to one another?
- If that is not the case, how different are the predicted trajectories of the outputs of interest when the implemented NPIs context is changes?
- Can we identify under which parametrisations these compartmental models become equivalent? That is, when can we substitute the conclusions of one for another as they will essentially act at the same model?

Main results

We started the thesis by tackling the first aim of our renewal equation model list of goals, as this builds up the framework on which we are able to answer the other three points. In Chapter 1, we derived from first principles the expression for a definition for the overall reproduction number as an indicator for the long-term behaviour of an epidemic in the context of a renewal equation process with multiple population groups for both the deterministic and stochastic versions of the model [Thompson et al., 2019; Creswell et al., 2022], and demonstrated its correctness through both an analytical and a simulation-based approach. This method for calculating the overall R_t completely sidesteps the need for computing the next-generation matrix – a popular [Green et al., 2022; Munday et al., 2023], but laborious and computationally expensive approach for deriving the reproduction number of an epidemic – and also does provides an exact expression for the overall R_t in terms of the group-specific reproduction numbers we defined in Chapter 2, rather and an estimator, as done in [Parag and Obolski, 2023].

In Chapter 2, we continued our journey in demonstrating the true usefulness of integrating heterogeneous population patterns into the renewal equation model framework. In this chapter, we showed that in highly variable epidemic context, that is populations in which the mixing patterns change often, using the standard one group renewal equation produces inaccurate estimates of the overall R_t if the different population groups have different associated

generation time distributions, even if the correct form of the average population generation time distribution is used. This is because the weighting of each individual group's generation time in the computation of the population average depends on the proportion of cases in each of these groups, proportions which will change significantly over time if contact patterns are constantly shifting, i.e. the population never reaches equilibrium in the fractions of infections in each case. This result has far-reaching consequences, as we have seen in the A/H1N1 Japan outbreak case study from 2009 which we used to visually highlight the difference between the predicted overall R_t trajectories when the population age-structure is assumed, compared to when mixing patterns are considered homogeneous, and how this would impact the policy making decision process.

To round off the final two points on our list, in Chapter 3 we assessed what is the minimum amount of heterogenous information – that is how detailed the contact matrix and the infection incidence data – that needs to be provided to the multiple group renewal equation model in order to obtain better predictions of the epidemic dynamics then by using the standard one group approach. In order to do so, we introduced four different renewal modelling frameworks, characterised by different levels of contact information and incidence data aggregation required to run the models and identified that the accuracy of R_t inference increases when population heterogeneity is incorporated into the modelling framework through the use of estimates of the contact matrix, even when no age-structured infection data is available, which answers the third goal set out for our meta-study of renewal equation models. Using synthetic epidemic data, we also showcased for a number of different interventional scenarios that targeting said interventions to those population groups with the highest number of effective contacts leads to the biggest reductions in transmission, on par with applying the same interventions to all population groups. This particular result is primarily of interest to policy decision makers that can therefore implement more efficient interventions by cutting down costs accumulated by applying interventions in demographics where the epidemic is not growing and instead redirecting these financial resources to improving the epidemic data collection process necessary to inform these models.

In Chapter 4, we focused on compartmental models for COVID-19 transmission and how differences in the modelling assumptions impact the conclusions that can be drawn when the epidemic context is changed, for example when different non-pharmaceutical interventions (NPIs) such as lock-downs are applied. We first introduce three compartmental ODE models

used by both:

- **scientific research consortia**: the PHE [Birrell et al., 2021] and the Warwick-Household model [Keeling et al., 2021], which were used by SPI-M consortium [Anderson et al., 2020; Brooks-Pollock et al., 2021] to advise governmental health agencies;
- and **pharmaceutical companies**: the Roche model [Lemenuel-Diot et al., 2020] used initially to study the infectious transmission dynamics of flu, and subsequently retrofitted for COVID-19;

in modelling the first wave of the UK COVID-19 epidemic in early 2020. All models were coded within a unified software framework, using the 'epimodels' [Bouros, 2021] Python package – an open-source package for building epidemiological models created as part of this DPhil project and which adheres to strict software engineering practices. For all of our models, we used age-structured daily number of deaths registered in London during the first 2020 epidemic wave which were collected by the Office of National Statistics (ONS) [GOV.UK] to fit the structurally identifiable model parameters, the rest of which we fixed through extensive hand-tuning. Similar analyses can also be performed for the other six different administrative regions of England. We then went on to compare the predictive power of these three models when we change the interventions away from those that were implemented in reality between February and June 2020 for two hypothetical scenarios: firstly, if no interventions were applied and individuals kept the same number of contacts as at the beginning of the simulation, and secondly how a seven day delay or advance in the introduction of the NPIs impacts the number of deaths, cases, and the reproduction number trajectory. We observed that none of the three models considered reached a consensus with one another when the intervention regime is changed from the one they were fitted to, suggesting that we cannot substitute any of these models for one another, thus answering our main research task for this part. Moreover, some of the model results (primarily those obtained from the Warwick-Household model) differ significantly from those presented in the original paper, or suggest improbable epidemic dynamics, e.g. the Roche model predicts an unrealistically high initial reproduction number of round 10 for the London region, while the Warwick-Household comparative analyses would suggest that interventions had little effect in reducing the number of deaths in the same region. These dissimilarities can be attributed to any of the two possibilities: either some of these models are unrealistic and fail to represent the actual epidemic dynamics – which is highly problematic as the recommendations

made to policy makers using these models will be biased by the modeller's assumptions on which key transmission features the model needs to exhibit – or the reimplementation process did not prove possible with the information available, in the attempt to reproduce the model – which is equally problematic, as these models cannot be hence validated and subsequently used to model future outbreaks.

Significance, study limitations and possible extensions

Throughout this project, we were able to highlight some of the discrepancies which we encounter when reproducing existing epidemiological models; how by simply changing the intervention regime to which the models have been fitted we arrive at epidemic outcomes that no long align with each other, or that indicate unrealistic epidemic behaviours such as a too fast depletion of the number of susceptible individuals. This result underlines a key result of this thesis: either the original models are not reproducible despite being correct, or they are reproducible but yield vastly different outcomes, even when standardising the data used and, to some extent, the assumptions made. This message is very much in line with those of other similar studies in the literature [Purkayastha et al., 2021; Silal et al., 2016; Siegenfeld et al., 2022; Özmen et al., 2016; Brauner et al., 2021; Gibson et al., 2018] that look at comparing different modelling frameworks. Moreover, this also implies that current complex compartmental models are not always reliable when it comes to quantifying the effects of pharmaceutical and non-pharmaceutical interventions, and in fact the model results might be biased by integrating these more complicated transmission assumptions – that is, a more complex modelling framework is not guaranteed to perform better than a simpler one.

Hence, more emphasis needs to be put on ensuring the reproducibility of these methods; crafting, fitting and implementing a new model that each time an epidemic outbreak occurs because previous models and their results cannot be replicated is not only a waste of time and resources, but also greatly undermines the reliability of infectious disease dynamics modelling for evidence-based policymaking. Therefore, developing and validating more robust models should be a top priority for the epidemiological scientific community to tackle as part of future pandemic preparedness efforts, to dispel overconfidence in models and to counteract the evermore prevalent dismissal of epidemiological models as opaque black

boxes.

At the same time, we showcased the importance of incorporating transmission heterogeneity in the renewal equation modelling framework, in particular regarding contact behaviour and the effect of case data aggregation, and demonstrated how this information can be used to improves the robustness of overall and group-specific reproduction number inference and implementation of interventions. Therefore, a more routine collection of contact and infection surveillance data used to inform those models is needed that will ultimately help us get a better understanding of the true underlying epidemic dynamics and equip us with the means to more effectively mitigate the severe outcomes of the epidemic in question. Another important result of this work is the development of an improved methodology for computing R_t which does not hinge on the computation of any additional quantities, such as the growth rate of an epidemic [Green et al., 2022].

The other consistent output of the project is the creation and maintenance of dedicated software for compartmental ODE ('epimodels') and renewal equation ('branchpro') modelling, which adhere to high standards of software engineering, with well documented and fully tested code. Additionally, due to the transparency of the coding, with many notebook examples that showcase their functionality, new users can easily get up to speed with using these tools to run these models or use the existing codebase as a template to add their own models to the package.

However, we should acknowledge some of the limitations of our findings. Firstly, due to time and data access constraints, we were only able to compare the results of a small sample of compartmental models. While the significance of the work done so far is by no means diminished, including more models in our comparison can only strengthen the message of this meta-analysis. Each of the three models had to to be reimplemented *de novo*, often based on outdated material compared to the available software. Moreover, some of these modelling approaches relied on sensitive clinical data such as the number of hospitalisations, to which we did not have access, or they used large-scale infection surveys such as the REACT1 study [Nicholson et al., 2021; Eales, 2020], which only started including snapshots of the population structure after the epidemic wave passed.

Secondly, all compartmental models included in these analyses focus on COVID-19 transmission and predate the development and deployment of effective vaccines. Moreover, these

models do not account for any waning in the immunity of the recovered individuals, therefore making them suitable for reuse only when analysing early epidemic short-term dynamics. Integrating additional models with these additional features in the open-source epimodels library will increase the impact and applicability of the software and provide future users with even more tools to develop their desired modelling framework in a rigorous, well-tested manner. Another immediate extension of this project is to develop a comprehensive approach to produce more accurate predictions of epidemic outcomes. For this, we could take inspiration from other fields such as machine learning, where weak predictors can be speculated to produce more robust predictions through the use of adaptive boosting, or adaboost [Schapire, 2013]; the imprecise predictions of these weak but cheap-to-run predictors are then averaged to produce better results. However, it is not clear that this approach would lead to similar improvements in epidemiological modelling. An alternative approach to solve this issue would be for the scientific community to refocus their efforts into identifying a set of key quality control checks that proposed new compartmental models need to pass before their conclusions are used to determine recommended courses of actions; for example, by comparing the outputs produced by fitting these compartmental models to 'realistic' synthetic data generated by complex agent-based models such as COVIDsim [Ferguson et al., 2020] or its reimplementation Epiabm [Gallagher et al., 2024].

With regard to our work on renewal equation models, we assumed that we have access to the effective contact information at any point in space and time; however, it is important to acknowledge that the collection of such detailed data is a huge challenge, both logistically – e.g. not always collected from the start of the outbreak, not enough human resources to do the collection, especially in isolated regions, or not the right scale of detail – and theoretically – e.g. how a contact is defined impacts the definition of the probability of a contact becoming a case, or which is the best way in which to split the population and therefore which are the groups for which we need to collect the contact and infection incidence data. Additionally, all the conclusions and analyses we have performed throughout this project have been applied and validated both theoretically and experimentally to renewal-type models for COVID-19 and Influenza-like diseases. However, there are no immediately apparent reasons that preclude the use of this modelling framework beyond human-to-human transmission; indeed given the additional heterogeneities that come from interactions of multiple organisms in vector-borne disease, there is every reason to imagine that this trend may be

exaggerated in this case.

Vector-borne disease (VBD) is an umbrella term for any human illness caused by pathogens that are transmitted indirectly via other living organisms, mostly insects, that can transmit infectious pathogens between humans, or from animals to humans (also known as vectors): some examples of such diseases include malaria, dengue, yellow fever and Zika virus fever (all transmitted by mosquitoes). Overall, the WHO estimates that vector-borne disease accounts for more than 17% of total disease infections worldwide, with more than 700,000 people dying every year; 400,000 of these deaths are due to malaria alone, with the majority occurring in those under five [WHO, 2024c]. Apart from air-borne diseases caused by pathogens transmitted through the air (COVID-19, influenza, measles amongst others) [WHO, 2024b], vector-borne diseases also pose a great risk as the precursor of the next pandemic. Hence, we can extend the applications of our newly developed methodology for a more accurate R_t estimation to a whole new class of pathogens and provide new insights on the efficacy of targeted interventions used to mitigate vector-to-host transmission.

Some final thoughts

In the process of choosing the best modelling framework to study the growth of an epidemic in a population of interest, the budding epidemiologist finds themself often spoilt for choice. They could choose to consider a renewal equation model, in order to minimise the amount of data and parameters that they need to contend with; or perhaps an agent-based model, as they instead focus on integrating as many of the features of transmission of the pathogen of interest; or indeed going for a compartmental ODE model framework instead, aiming to find an equilibrium between realistic diseases assumptions and model identifiability. However, they invariably fall short of the mark: either the model is too complex, and they can only retrieve a handful of the many parameters of the model from the data – thus significantly altering the usefulness of the resulting parameterisation when a new epidemic of the same pathogen comes about; or the model is too simple and it cannot fully reproduce the effects of behavioural variations in the population and in the effects of the pathogen on different demographics.

Therefore, going forward, we as a scientific community should demonstrate more rigour when developing new models for the study of infectious disease transmission. Indeed, by

taking a step back now and concentrating our collective effort to produce well-vetted models for air-borne (or for that matter vector-borne) diseases, or simply reach a consensus on a standard set of a key analyses that mathematical models of disease transmission need to meet before being used in policy management, we will be sparing the time and resources currently being used to continue churning out new models and 'reinventing the wheel'. Additionally, we need to make sure that these newly identified modelling tools are available through well-engineered open-source software. This endeavour has the potential to streamline the process of selecting a model in the long run. Moreover, it will ensure the correctness of the advice given by researchers in persuading governmental health agencies on the best method for mitigating epidemic growth, not if, but when the next pandemic hits.

Bibliography

S. Abbott, J. Hellewell, R. N. Thompson, K. Sherratt, H. P. Gibbs, N. I. Bosse, J. D. Munday, S. Meakin, E. L. Doughty, J. Y. Chun, Y.-W. D. Chan, F. Finger, P. Campbell, A. Endo, C. A. B. Pearson, A. Gimma, T. Russell, S. Flasche, A. J. Kucharski, R. M. Eggo, and S. F. and. Estimating the time-varying reproduction number of SARS-CoV-2 using national and subnational case counts. *Wellcome Open Research*, 5:112, Dec. 2020. doi: 10.12688/wellcomeopenres.16006.2. URL https://doi.org/10.12688/wellcomeopenres.16006.2.

L. J. Abu-Raddad, H. Chemaitelly, P. Coyle, J. A. Malek, A. A. Ahmed, Y. A. Mohamoud, S. Younuskunju, H. H. Ayoub, Z. Al Kanaani, E. Al Kuwari, et al. SARS-CoV-2 antibody-positivity protects against reinfection for at least seven months with 95% efficacy. *EClinicalMedicine*, 35:100861, 2021.

D. Adam. Special report: The simulations driving the world's response to COVID-19. *Nature*, 580(7803):316–318, Apr. 2020. doi: 10.1038/d41586-020-01003-6. URL https://doi.org/10.1038/d41586-020-01003-6.

R. Anderson, C. Donnelly, D. Hollingsworth, M. Keeling, C. Vegvari, R. Baggaley, and R. Maddren. Reproduction number (R) and growth rate (r) of the COVID-19 epidemic in the UK: methods of estimation, data sources, causes of heterogeneity, and use as a guide in policy formulation. *The Royal Society*, Sep, 2020. URL https://royalsociety.org/-/media/policy/projects/set-c/set-covid-19-R-estimates.pdf.

R. M. Anderson and R. M. May. *Infectious diseases of humans : dynamics and control*. Oxford science publications. Oxford University Press, Oxford, 1991. ISBN 9780198545996.

O. D. B. Boldin and J. A. J. Metz. Population growth in discrete time: a renewal equation oriented survey. *Journal of Difference Equations and Applications*, pages 1–29, 2023. doi: 10.1080/10236198.2023.2265499. URL https://doi.org/10.1080/10236198.2023.2265499.

N. Bacaër and R. Ouifki. Growth rate and basic reproduction number for population models with a simple periodic factor. *Mathematical Biosciences*, 210(2):647–658, 2007.

N. Bacaer, D. Maxin, F. Munteanu, F. Avram, P. Georgescu, I. Stoleriu, and A. Halanay.

Matematica si epidemii. N. Bacaer, Paris, 2022. ISBN 979-10-343-0320-5. URL https://www.documentation.ird.fr/hor/fdi:010084975.

P. Ballard, N. Bean, and J. Ross. Intervention to maximise the probability of epidemic fade-out. *Mathematical Biosciences*, 293:1–10, 2017. ISSN 0025-5564. doi: https://doi.org/10.1016/j.mbs.2017.08.003. URL https://www.sciencedirect.com/science/article/pii/S0025556416302541.

R. Berger and G. Casella. *Statistical Inference*. Duxbury Press, Florence, AL, 2 edition, 2001. ISBN 9780534243128.

S. Bhatia, J. Wardle, R. K. Nash, P. Nouvellet, and A. Cori. Extending EpiEstim to estimate the transmission advantage of pathogen variants in real-time: SARS-CoV-2 as a case-study. *Epidemics*, 44:100692, 2023.

P. Birrell, J. Blake, E. van Leeuwen, P. J. M. Cell, N. Gent, and D. De Angelis. Real-time Nowcasting and Forecasting of COVID-19 Dynamics in England: the first wave? *Philosophical Transactions of the Royal Society B*, 376(1829):20200279, 2021. doi: 10.1101/2020.08.24.20180737. URL https://doi.org/10.1098/rstb.2020.0279.

P. Bosetti, C. Tran Kiem, A. Andronico, V. Colizza, Y. Yazdanpanah, A. Fontanet, D. Benamouzig, and S. Cauchemez. Epidemiology and control of SARS-CoV-2 epidemics in partially vaccinated populations: a modeling study applied to France. Working paper or preprint at https://hal-pasteur.archives-ouvertes.fr/pasteur-03272638, Sep 2021. URL https://hal-pasteur.archives-ouvertes.fr/pasteur-03272638.

I. Bouros. Multi Epidemiology Model Cross Analysis. https://github.com/I-Bouros/multi-epi-model-cross-analysis, 2021.

I. Bouros, E. M. Hill, M. J. Keeling, S. Moore, and R. N. Thompson. Prioritising older individuals for COVID-19 booster vaccination leads to optimal public health outcomes in a range of socio-economic settings. *PLOS Computational Biology*, 20(8):1–19, 08 2024. doi: 10.1371/journal.pcbi.1012309. URL https://doi.org/10.1371/journal.pcbi.1012309.

I. Bouros, R. Thompson, D. Gavaghan, and B. Lambert. The time-dependent reproduction number for epidemics in heterogeneous populations, 2025. ISSN 1742-5662. URL http://dx.doi.org/10.1098/rsif.2025.0095.

G. E. P. Box and N. R. Draper. *Empirical Model-Building and Response Surface*. John Wiley & Sons, Inc., USA, 1986. ISBN 0471810339. doi: 10.5555/17317.

J. M. Brauner, S. Mindermann, M. Sharma, D. Johnston, J. Salvatier, T. Gavenčiak, A. B. Stephenson, G. Leech, G. Altman, V. Mikulik, A. J. Norman, J. T. Monrad, T. Besiroglu, H. Ge, M. A. Hartwick, Y. W. Teh, L. Chindelevitch, Y. Gal, and J. Kulveit. Inferring the effectiveness of government interventions against COVID-19. *Science*, 371(6531): eabd9338, 2021. doi: 10.1126/science.abd9338. URL https://www.science.org/doi/abs/10.1126/science.abd9338.

E. K. Brockhaus, D. Wolffram, T. Stadler, M. Osthege, T. Mitra, J. M. Littek, E. Krymova, A. J. Klesen, J. S. Huisman, S. Heyder, L. M. Helleckes, M. an der Heiden, S. Funk, S. Abbott, and J. Bracher. Why are different estimates of the effective reproductive number so different? A case study on COVID-19 in Germany. *PLOS Computational Biology*, 19 (11):e1011653, Nov. 2023. ISSN 1553-7358. doi: 10.1371/journal.pcbi.1011653. URL http://dx.doi.org/10.1371/journal.pcbi.1011653.

E. Brooks-Pollock, L. Danon, T. Jombart, and L. Pellis. Modelling that shaped the early COVID-19 pandemic response in the UK. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 376(1829):20210001, 2021. doi: 10.1098/rstb.2021.0001. URL https://royalsocietypublishing.org/doi/abs/10.1098/rstb.2021.0001.

D. S. Chawla. Critiqued coronavirus simulation gets thumbs up from code-checking efforts. *Nature*, 582(7812):323–324, June 2020. doi: 10.1038/d41586-020-01685-y. URL https://doi.org/10.1038/d41586-020-01685-y.

D. Chen, Y.-C. Lau, X.-K. Xu, L. Wang, Z. Du, T. K. Tsang, P. Wu, E. H. Lau, J. Wallinga, B. J. Cowling, and S. T. Ali. Inferring time-varying generation time, serial interval, and incubation period distributions for COVID-19. *Nature Communications*, 13(1):7727, 2022.

G. Chowell and H. Nishiura. Transmission dynamics and control of Ebola virus disease (EVD): a review. *BMC Medicine*, 12(1), Oct. 2014. ISSN 1741-7015. doi: 10.1186/s12916-014-0196-0. URL http://dx.doi.org/10.1186/s12916-014-0196-0.

M. Clerx, M. Robinson, B. Lambert, C. L. Lei, S. Ghosh, G. R. Mirams, and D. J. Gavaghan. Probabilistic Inference on Noisy Time Series (PINTS). *Journal of Open Research Software*, 7(1):23, 2019. doi: 10.5334/jors.252.

C. Cobelli and J. J. DiStefano. Parameter and structural identifiability concepts and ambiguities: a critical review and analysis. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, 239(1):R7–R24, 1980. doi: 10.1152/ajpregu.1980.239.1.R7. URL https://doi.org/10.1152/ajpregu.1980.239.1.R7. PMID: 7396041.

Como-DTC-Collaboration. como-models. https://github.com/Como-DTC-Collaboration/como-models, 2021.

A. Cori, N. M. Ferguson, C. Fraser, and S. Cauchemez. A new framework and software to estimate time-varying reproduction numbers during epidemics. *American Journal of Epidemiology*, 178(9):1505–1512, 09 2013. ISSN 0002-9262. doi: 10.1093/aje/kwt133. URL https://doi.org/10.1093/aje/kwt133.

R. Creswell, D. Augustin, I. Bouros, H. Farm, S. Miao, A. Ahern, M. Robinson, A. Lemenuel-Diot, D. Gavaghan, B. Lambert, and R. Thompson. Heterogeneity in the onwards transmission risk between local and imported cases affects practical estimates of the time-dependent reproduction number. *Philosophical Transactions of the Royal Society A*, 380(2233):20210308, 2022.

R. Creswell, M. Robinson, D. Gavaghan, K. V. Parag, C. L. Lei, and B. Lambert. A Bayesian nonparametric method for detecting rapid changes in disease transmission. *Journal of Theoretical Biology*, 558:111351, 2023.

E. A. Dankwa, A. F. Brouwer, and C. A. Donnelly. Structural identifiability of compartmental models for infectious disease transmission is influenced by data type. *Epidemics*, 41:100643, 2022. ISSN 1755-4365. doi: https://doi.org/10.1016/j.epidem.2022.100643. URL https://www.sciencedirect.com/science/article/pii/S1755436522000834.

L. Danon, L. Lacasa, and E. Brooks-Pollock. Household bubbles and COVID-19 transmission: insights from percolation theory. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 376(1829):20200284, 2021. doi: 10.1098/rstb.2020.0284. URL https://royalsocietypublishing.org/doi/abs/10.1098/rstb.2020.0284.

L. Di Domenico, P. Bosetti, C. E. Sabbatini, L. Opatowski, and V. Colizza. Mobility-driven synthetic contact matrices: a scalable solution for real-time pandemic response modeling. *medRxiv*, pages 2024–12, 2024.

O. Diekmann, J. Heesterbeek, and J. Metz. *On the definition and the computation of the basic reproduction ratio R0 in models for infectious diseases in heterogeneous populations*, volume 28. Springer Science and Business Media LLC, June 1990. doi: 10.1007/bf00178324. URL http://dx.doi.org/10.1007/BF00178324.

O. Diekmann, J. A. P. Heesterbeek, and M. G. Roberts. The construction of next-generation matrices for compartmental epidemic models. *Journal of The Royal Society Interface*, 7(47):873–885, 2010. doi: 10.1098/rsif.2009.0386. URL https://royalsocietypublishing.org/doi/abs/10.1098/rsif.2009.0386.

J. Dormand and P. Prince. A family of embedded Runge-Kutta formulae. *Journal of Computational and Applied Mathematics*, 6(1):19–26, 1980. doi: 10.1016/0771-050x(80)90013-3.

O. Eales. reactidd. https://github.com/mrc-ide/reactidd, 2020.

N. Ferguson, D. Laydon, G. Nedjati Gilani, N. Imai, K. Ainslie, M. Baguelin, S. Bhatia, A. Boonyasiri, Z. Cucunuba Perez, G. Cuomo-Dannenburg, A. Dighe, I. Dorigatti, H. Fu, K. Gaythorpe, W. Green, A. Hamlet, W. Hinsley, L. Okell, S. Van Elsland, H. Thompson, R. Verity, E. Volz, H. Wang, Y. Wang, P. Walker, P. Winskill, C. Whittaker, C. Donnelly, S. Riley, and A. Ghani. Report 9: Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand. Preprint at Spiral https://doi.org/10.25561/77482, 2020.

S. Flaxman, S. Mishra, A. Gandy, H. J. T. Unwin, T. A. Mellan, H. Coupland, C. Whittaker, H. Zhu, T. Berah, J. W. Eaton, M. Monod, P. N. Perez-Guzman, N. Schmit, L. Cilloni, K. E. C. Ainslie, M. Baguelin, A. Boonyasiri, O. Boyd, L. Cattarino, L. V. Cooper, Z. Cucunubá, G. Cuomo-Dannenburg, A. Dighe, B. Djaafara, I. Dorigatti, S. L. van Elsland, R. G. FitzJohn, K. A. M. Gaythorpe, L. Geidelberg, N. C. Grassly, W. D. Green, T. Hallett, A. Hamlet, W. Hinsley, B. Jeffrey, E. Knock, D. J. Laydon, G. Nedjati-Gilani, P. Nouvellet, K. V. Parag, I. Siveroni, H. A. Thompson, R. Verity, E. Volz, C. E. Walters, H. Wang, Y. Wang, O. J. Watson, P. Winskill, X. Xi, P. G. T. Walker, A. C. Ghani, C. A. Donnelly, S. Riley, M. A. C. Vollmer, N. M. Ferguson, L. C. Okell, S. Bhatt, and I. Team. Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature*, 584 (7820):257–261, Aug 2020. ISSN 1476-4687. doi: 10.1038/s41586-020-2405-7. URL https://doi.org/10.1038/s41586-020-2405-7.

C. Fraser. Estimating Individual and Household Reproduction Numbers in an Emerging Epidemic. *PLOS ONE*, 2(8):1–12, 08 2007. doi: 10.1371/journal.pone.0000758. URL https://doi.org/10.1371/journal.pone.0000758.

K. Gallagher, I. Bouros, N. Fan, E. Hayman, L. Heirene, P. Lamirande, A. Lemenuel-Diot, B. Lambert, D. Gavaghan, and R. Creswell. Epidemiological Agent-Based Modelling Software (Epiabm). *Journal of Open Research Software*, 12(1), Mar. 2024. doi: 10.5334/jors.449.

F. Gallais, P. Gantner, T. Bruel, A. Velay, D. Planas, M.-J. Wendling, S. Bayer, M. Solis, E. Laugel, N. Reix, et al. Evolution of antibody responses up to 13 months after SARS-CoV-2 infection and risk of reinfection. *EBioMedicine*, 71:103561, 2021.

A. Gelman and D. B. Rubin. Inference from iterative simulation using multiple sequences. *Statistical science: a review journal of the Institute of Mathematical Statistics*, 7(4):457–472, 1992. ISSN 0883-4237. doi: 10.1214/ss/1177011136. URL http://dx.doi.org/10.1214/ss/1177011136.

G. J. Gibson, G. Streftaris, and D. Thong. Comparison and assessment of epidemic models. *Statistical Science*, 33(1):19–33, 2018. ISSN 08834237, 21688745. URL https://www.jstor.org/stable/26770976.

K. Glass, G. N. Mercer, H. Nishiura, E. S. McBryde, and N. G. Becker. Estimating reproduction numbers for adults and children from case data. *Journal of The Royal Society Interface*, 8(62):1248–1259, 2011. doi: 10.1098/rsif.2010.0679. URL https://royalsocietypublishing.org/doi/abs/10.1098/rsif.2010.0679.

N. Goeyvaerts, E. Santermans, G. Potter, A. Torneri, K. Van Kerckhove, L. Willem, M. Aerts, P. Beutels, and N. Hens. Household members do not contact each other at random: implications for infectious disease modelling. *Proceedings of the Royal Society B: Biological Sciences*, 285(1893):20182201, Dec. 2018. ISSN 1471-2954. doi: 10.1098/rspb.2018.2201. URL http://dx.doi.org/10.1098/rspb.2018.2201.

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

K. M. Gostic, L. McGough, E. B. Baskerville, S. Abbott, K. Joshi, C. Tedijanto, R. Kahn, R. Niehus, J. A. Hay, P. M. De Salazar, J. Hellewell, S. Meakin, J. D. Munday, N. I. Bosse, K. Sherrat, R. N. Thompson, L. F. White, J. S. Huisman, J. Scire, S. Bonhoeffer,

T. Stadler, J. Wallinga, S. Funk, M. Lipsitch, and S. Cobey. Practical considerations for measuring the effective reproductive number, R_t. *PLOS Computational Biology*, 16(12): e1008409, 2020.

GOV.UK. Daily summary. https://coronavirus.data.gov.uk/details/cases. [Online; accessed 26-May-2021].

W. D. Green, N. M. Ferguson, and A. Cori. Inferring the reproduction number using the renewal equation in heterogeneous epidemics. *Journal of The Royal Society Interface*, 19(188):20210429, 2022. doi: 10.1098/rsif.2021.0429. URL https://royalsocietypublishing.org/doi/abs/10.1098/rsif.2021.0429.

H. Haario, E. Saksman, and J. Tamminen. An Adaptive Metropolis Algorithm. *Bernoulli: official journal of the Bernoulli Society for Mathematical Statistics and Probability*, 7(2): 223, 2001. ISSN 1350-7265. doi: 10.2307/3318737. URL http://dx.doi.org/10.2307/3318737.

T. Hale, N. Angrist, R. Goldszmidt, B. Kira, A. Petherick, T. Phillips, S. Webster, E. Cameron-Blake, L. Hallas, S. Majumdar, and H. Tatlow. A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker). *Nature Human Behaviour*, 5(4):529–538, Mar. 2021. doi: 10.1038/s41562-021-01079-8. URL https://doi.org/10.1038/s41562-021-01079-8.

N. Hansen. *The CMA Evolution Strategy: A Comparing Review*, pages 75–102. Springer Berlin Heidelberg, Berlin, Heidelberg, 2006. ISBN 978-3-540-32494-2. doi: 10.1007/3-540-32494-1_4. URL https://doi.org/10.1007/3-540-32494-1_4.

N. Hansen, S. D. Müller, and P. Koumoutsakos. Reducing the Time Complexity of the Derandomized Evolution Strategy with Covariance Matrix Adaptation (CMA-ES). *Evolutionary Computation*, 11(1):1–18, 03 2003. ISSN 1063-6560. doi: 10.1162/106365603321828970. URL https://doi.org/10.1162/106365603321828970.

W. S. Hart, S. Abbott, A. Endo, J. Hellewell, E. Miller, N. Andrews, P. K. Maini, S. Funk, and R. N. Thompson. Inference of the SARS-CoV-2 generation time using UK household data. *ELife*, 11:e70767, 2022a.

W. S. Hart, E. Miller, N. J. Andrews, P. Waight, P. K. Maini, S. Funk, and R. N. Thomp-

son. Generation time of the alpha and delta SARS-CoV-2 variants: an epidemiological analysis. *The Lancet Infectious Diseases*, 22(5):603–610, 2022b.

F. Ho, K. V. Parag, D. C. Adam, E. H. Lau, B. J. Cowling, and T. K. Tsang. Accounting for the potential of overdispersion in estimation of the time-varying reproduction number. *Epidemiology*, 34(2):201–205, 2023.

M. D. Hoffman and A. Gelman. The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo, 2011. URL https://arxiv.org/abs/1111.4246.

A. B. Hogan, P. Winskill, O. J. Watson, P. G. Walker, C. Whittaker, M. Baguelin, N. F. Brazeau, G. D. Charles, K. A. Gaythorpe, A. Hamlet, E. Knock, D. J. Laydon, J. A. Lees, A. Løchen, R. Verity, L. K. Whittles, F. Muhib, K. Hauck, N. M. Ferguson, and A. C. Ghani. Within-country age-based prioritisation, global allocation, and public health impact of a vaccine against SARS-CoV-2: A mathematical modelling analysis. *Vaccine*, 39(22):2995–3006, 2021. ISSN 0264-410X. doi: https://doi.org/10.1016/j.vaccine.2021.04.002. URL https://www.sciencedirect.com/science/article/pii/S0264410X21004278.

H. Hong, A. Ovchinnikov, G. Pogudin, and C. Yap. Sian: software for structural identifiability analysis of ode models. *Bioinformatics*, 35(16):2873–2874, 01 2019. ISSN 1367-4803. doi: 10.1093/bioinformatics/bty1069. URL https://doi.org/10.1093/bioinformatics/bty1069.

H. Hong, A. Ovchinnikov, G. Pogudin, and C. Yap. Global identifiability of differential models. *Communications on Pure and Applied Mathematics*, 73(9):1831–1879, 2020. doi: https://doi.org/10.1002/cpa.21921. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/cpa.21921.

R. H. Johnstone, E. T. Y. Chang, R. Bardenet, T. P. de Boer, D. J. Gavaghan, P. Pathmanathan, R. H. Clayton, and G. R. Mirams. Uncertainty and variability in models of the cardiac action potential: Can we build trustworthy models? *Journal of molecular and cellular cardiology*, 96:49–62, 2016. ISSN 0022-2828. doi: 10.1016/j.yjmcc.2015.11.018. URL http://dx.doi.org/10.1016/j.yjmcc.2015.11.018.

K. E. Jones, N. G. Patel, M. A. Levy, A. Storeygard, D. Balk, J. L. Gittleman, and P. Daszak. Global trends in emerging infectious diseases. *Nature*, 451(7181):990–993,

Feb. 2008. ISSN 1476-4687. doi: 10.1038/nature06536. URL http://dx.doi.org/10.1038/nature06536.

M. J. Keeling, E. M. Hill, E. E. Gorsich, B. Penman, G. Guyver-Fletcher, A. Holmes, T. Leng, H. McKimm, M. Tamborrino, L. Dyson, and M. J. Tildesley. Predictions of COVID-19 dynamics in the UK: Short-term forecasting and analysis of potential exit strategies. *PLOS Computational Biology*, 17(1):1–20, 01 2021. doi: 10.1371/journal.pcbi.1008619. URL https://doi.org/10.1371/journal.pcbi.1008619.

M. J. Keeling, L. Dyson, G. Guyver-Fletcher, A. Holmes, M. G. Semple, I. Investigators, M. J. Tildesley, and E. M. Hill. Fitting to the uk covid-19 outbreak, short-term forecasts and estimating the reproductive number. *Statistical Methods in Medical Research*, 31 (9):1716–1737, 2022. doi: 10.1177/09622802211070257. URL https://doi.org/10.1177/09622802211070257. PMID: 35037796.

W. O. Kermack, A. G. McKendrick, and G. T. Walker. A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 115(772):700–721, 1927. doi: 10.1098/rspa.1927.0118. URL https://royalsocietypublishing.org/doi/abs/10.1098/rspa.1927.0118.

D. Kim, S. T. Ali, S. Kim, J. Jo, J. Lim, S. Lee, and S. Ryu. Estimation of serial interval and reproduction number to quantify the transmissibility of SARS-CoV-2 omicron Variant in South Korea. *Viruses*, 14(3), 2022. ISSN 1999-4915. doi: 10.3390/v14030533. URL https://www.mdpi.com/1999-4915/14/3/533.

K. Koelle, M. A. Martin, R. Antia, B. Lopman, and N. E. Dean. The changing epidemiology of SARS-CoV-2. *Science*, 375(6585):1116–1121, 2022. doi: 10.1126/science.abm4915. URL https://www.science.org/doi/abs/10.1126/science.abm4915.

B. Lambert. *A student's guide to Bayesian statistics*. SAGE Publications, London, England, 2018. ISBN 9781473916364.

D. Lee, B. Carpenter, P. Li, M. Morris, M. Betancourt, Maverickg, M. Brubaker, R. Trangucci, M. Inacio, A. Kucukelbir, S. Buildbot, Bgoodri, Seantalts, J. Arnold, D. Tran, M. Hoffman, C. Margossian, M. Modrák, A. Adler, K. Sakrejda, A. Stukalov,

M. Lawrence, R. J. Goedman, K. S. Van Horn, A. Vehtari, J. Gabry, J. S. Casallas, and B. Bales. stan-dev/stan: v2.17.1, 2017. URL https://zenodo.org/record/1101116.

A. Leidi, F. Koegler, R. Dumont, R. Dubos, M.-E. Zaballa, G. Piumatti, M. Coen, A. Berner, P. D. Farhoumand, P. Vetter, et al. Risk of reinfection after seroconversion to SARS-CoV-2: A population-based propensity-score matched cohort study. *MedRxiv*, 2021.

A. Lemenuel-Diot, B. Clinch, A. C. Hurt, P. Boutry, J. Laurent, M. Leddin, S. Frings, and J. E. Charoin. A COVID-19 transmission model informing medication development and supply chain needs. *medRxiv*, 2020. doi: 10.1101/2020.11.23.20237404. URL https://www.medrxiv.org/content/early/2020/12/02/2020.11.23.20237404.

C. Lin, P. Tu, and L. M. Beitsch. Confidence and Receptivity for COVID-19 Vaccines: A Rapid Systematic Review. *Vaccines*, 9(1), 2021. ISSN 2076-393X. doi: 10.3390/vaccines9010016. URL https://www.mdpi.com/2076-393X/9/1/16.

Y. Liu, A. A. Gayle, A. Wilder-Smith, and J. Rocklöv. The reproductive number of COVID-19 is higher compared to SARS coronavirus. *Journal of Travel Medicine*, 27(2):taaa021, 02 2020. ISSN 1708-8305. doi: 10.1093/jtm/taaa021. URL https://doi.org/10.1093/jtm/taaa021.

J. O. Lloyd-Smith, S. J. Schreiber, P. E. Kopp, and W. M. Getz. Superspreading and the effect of individual variation on disease emergence. *Nature*, 438(7066):355–359, 2005. ISSN 0028-0836. doi: 10.1038/nature04153. URL http://dx.doi.org/10.1038/nature04153.

F. A. Lovell-Read, S. Shen, and R. N. Thompson. Estimating local outbreak risks and the effects of non-pharmaceutical interventions in age-structured populations: Sars-cov-2 as a case study. *Journal of Theoretical Biology*, 535:110983, 2022a. ISSN 0022-5193. doi: https://doi.org/10.1016/j.jtbi.2021.110983. URL https://www.sciencedirect.com/science/article/pii/S0022519321004033.

F. A. Lovell-Read, S. Shen, and R. N. Thompson. Estimating local outbreak risks and the effects of non-pharmaceutical interventions in age-structured populations: SARS-CoV-2 as a case study. *Journal of Theoretical Biology*, 535:110983, 2022b.

K. Mizumoto, K. Kagaya, A. Zarebski, and G. Chowell. Estimating the asymptomatic proportion of coronavirus disease 2019 (COVID-19) cases on board the Diamond Princess cruise ship, Yokohama, Japan, 2020. *Eurosurveillance*, 25(10"):2000180,

2020. doi: https://doi.org/10.2807/1560-7917.ES.2020.25.10.2000180. URL https://www.eurosurveillance.org/content/10.2807/1560-7917.ES.2020.25.10.2000180.

A. M'kendrick. Applications of mathematics to medical problems. *Proceedings of the Edinburgh Mathematical Society*, 44:98–130, 1925.

S. Moore, E. M. Hill, L. Dyson, M. J. Tildesley, and M. J. Keeling. Retrospectively modeling the effects of increased global vaccine sharing on the COVID-19 pandemic. *Nature Medicine*, 28(11):2416–2423, Nov 2022. ISSN 1546-170X. doi: 10.1038/s41591-022-02064-y. URL https://doi.org/10.1038/s41591-022-02064-y.

J. D. Munday, S. Abbott, S. Meakin, and S. Funk. Evaluating the use of social contact data to produce age-specific short-term forecasts of sars-cov-2 incidence in england. *PLOS Computational Biology*, 19(9):1–22, 09 2023. doi: 10.1371/journal.pcbi.1011453. URL https://doi.org/10.1371/journal.pcbi.1011453.

J. D. Murray. *Mathematical Biology I. An Introduction*, volume 17 of *Interdisciplinary Applied Mathematics*. Springer, New York, 3 edition, 2002. doi: 10.1007/b98868.

R. K. Nash, S. Bhatt, A. Cori, and P. Nouvellet. Estimating the epidemic reproduction number from temporally aggregated incidence data: a statistical modelling approach and software tool. *PLOS Computational Biology*, 19(8):e1011439, 2023.

G. Nicholson, B. Lehmann, T. Padellini, K. B. Pouwels, R. Jersakova, J. Lomax, R. E. King, A.-M. Mallon, P. J. Diggle, S. Richardson, M. Blangiardo, and C. Holmes. Improving local prevalence estimates of SARS-CoV-2 infections using a causal debiasing framework. *Nature Microbiology*, 7(1):97–107, Dec. 2021. doi: 10.1038/s41564-021-01029-0. URL https://doi.org/10.1038/s41564-021-01029-0.

H. Nishiura and G. Chowell. *The Effective Reproduction Number as a Prelude to Statistical Estimation of Time-Dependent Epidemic Trends*. Springer Netherlands, Dordrecht, 2009. ISBN 978-90-481-2313-1. doi: 10.1007/978-90-481-2313-1_5. URL https://doi.org/10.1007/978-90-481-2313-1_5.

H. Nishiura, C. Castillo-Chavez, M. Safan, and G. Chowell. Transmission potential of the new influenza A(H1N1) virus and its age-specificity in Japan. *Eurosurveillance*, 14(22):19227, 2009. doi: https://doi.org/10.2807/ese.14.22.19227-en. URL https://www.eurosurveillance.org/content/10.2807/ese.14.22.19227-en.

H. Nishiura, T. Kobayashi, T. Miyama, A. Suzuki, S.-m. Jung, K. Hayashi, R. Kinoshita, Y. Yang, B. Yuan, A. R. Akhmetzhanov, and N. M. Linton. Estimation of the asymptomatic ratio of novel coronavirus infections (COVID-19). *International Journal of Infectious Diseases*, 94:154–155, May 2020. ISSN 1201-9712. doi: 10.1016/j.ijid.2020.03.020. URL https://doi.org/10.1016/j.ijid.2020.03.020.

I. Ogi-Gittins, W. Hart, J. Song, R. Nash, J. Polonsky, A. Cori, E. M. Hill, and R. Thompson. A simulation-based approach for estimating the time-dependent reproduction number from temporally aggregated disease incidence time series data. *Epidemics*, 47:100773, 2024.

ONS. Dataset: Estimates of the population for the UK, England and Wales, Scotland and Northern Ireland. https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/datasets/populationestimatesforukenglandandwalesscotlandandnorthernireland. [Online; accessed 16-April-2021].

O. Özmen, J. J. Nutaro, L. L. Pullum, and A. Ramanathan. Analyzing the impact of modeling choices and assumptions in compartmental epidemiological models. *SIMULATION*, 92(5):459–472, 2016. doi: 10.1177/0037549716640877. URL https://doi.org/10.1177/0037549716640877.

K. V. Parag and U. Obolski. Risk averse reproduction numbers improve resurgence detection. *PLOS Computational Biology*, 19(7):1–25, 07 2023. doi: 10.1371/journal.pcbi.1011332. URL https://doi.org/10.1371/journal.pcbi.1011332.

K. V. Parag, B. J. Cowling, and B. Lambert. Angular reproduction numbers improve estimates of transmissibility when disease generation times are misspecified or time-varying. *Proceedings of the Royal Society B*, 290(2007):20231664, 2023.

S. Payne. Chapter 5 - Virus Transmission and Epidemiology. In S. Payne, editor, *Viruses*, pages 53–60. Academic Press, 2017. ISBN 978-0-12-803109-4. doi: https://doi.org/10.1016/B978-0-12-803109-4.00005-2. URL https://www.sciencedirect.com/science/article/pii/B9780128031094000052.

L. Pellis, F. Scarabel, H. B. Stage, C. E. Overton, L. H. K. Chappell, E. Fearon, E. Bennett, K. A. Lythgoe, T. A. House, I. Hall, and n. null. Challenges in control of COVID-19:

short doubling time and long delay to effect of interventions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 376(1829):20200264, 2021. doi: 10.1098/rstb.2020.0264. URL https://royalsocietypublishing.org/doi/abs/10.1098/rstb.2020.0264.

B. G. Pijls, S. Jolani, A. Atherley, R. T. Derckx, J. I. Dijkstra, G. H. Franssen, S. Hendriks, A. Richters, A. Venemans-Jellema, S. Zalpuri, and M. P. Zeegers. Demographic risk factors for COVID-19 infection, severity, ICU admission and death: a meta-analysis of 59 studies. *BMJ Open*, 11(1):e044640, 2021.

K. Prem, K. v. Zandvoort, P. Klepac, R. M. Eggo, N. G. Davies, C. for the Mathematical Modelling of Infectious Diseases COVID-19 Working Group, A. R. Cook, and M. Jit. Projecting contact matrices in 177 geographical regions: An update and comparison with empirical data for the COVID-19 era. *PLOS Computational Biology*, 17 (7):1–19, 07 2021. doi: 10.1371/journal.pcbi.1009098. URL https://doi.org/10.1371/journal.pcbi.1009098.

S. Purkayastha, R. Bhattacharyya, R. Bhaduri, R. Kundu, X. Gu, M. Salvatore, D. Ray, S. Mishra, and B. Mukherjee. A comparison of five epidemiological models for transmission of SARS-CoV-2 in India. *BMC Infectious Diseases*, 21(1):533, Jun 2021. ISSN 1471-2334. doi: 10.1186/s12879-021-06077-9. URL https://doi.org/10.1186/s12879-021-06077-9.

SABS-R3-Epidemiology. branchpro. https://github.com/SABS-R3-Epidemiology/branchpro, 2025.

R. E. Schapire. Explaining adaboost. In B. Schölkopf, Z. Luo, and V. Vovk, editors, *Empirical Inference: Festschrift in Honor of Vladimir N. Vapnik*, pages 37–52. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013. ISBN 978-3-642-41136-6. doi: 10.1007/978-3-642-41136-6_5.

A. F. Siegenfeld, P. K. Kollepara, and Y. Bar-Yam. Modeling complex systems: A case study of compartmental models in epidemiology. *Complexity*, 2022(1):3007864, 2022. doi: https://doi.org/10.1155/2022/3007864. URL https://onlinelibrary.wiley.com/doi/abs/10.1155/2022/3007864.

S. P. Silal, F. Little, K. I. Barnes, and L. J. White. Sensitivity to model structure: a com-

parison of compartmental models in epidemiology. *Health Systems*, 5(3):178–191, 2016. doi: 10.1057/hs.2015.2. URL https://doi.org/10.1057/hs.2015.2. PMID: 36061953.

D. Singh Chawla. Critiqued coronavirus simulation gets thumbs up from code-checking efforts. *Nature*, 582(7812):323–324, June 2020. ISSN 1476-4687. doi: 10.1038/d41586-020-01685-y. URL http://dx.doi.org/10.1038/d41586-020-01685-y.

Statistics Bureau of Japan. Summary of the results of Population Census of Japan 2010. https://www.stat.go.jp/english/data/kokusei/2010/final_en/pdf/summary.pdf. [Online; accessed 25-August-2024].

Å. Svensson. A note on generation times in epidemic models. *Mathematical Biosciences*, 208(1):300–311, 2007.

D. A. Theodore, A. R. Branche, L. Zhang, D. S. Graciaa, M. Choudhary, T. J. Hatlen, R. Osman, T. M. Babu, S. T. Robinson, P. B. Gilbert, D. Follmann, H. Janes, J. G. Kublin, L. R. Baden, P. Goepfert, G. E. Gray, B. Grinsztejn, K. L. Kotloff, C. L. Gay, B. Leav, J. Miller, I. Hirsch, J. Sadoff, L. M. Dunkle, K. M. Neuzil, L. Corey, A. R. Falsey, H. M. E. Sahly, M. E. Sobieszczyk, and Y. Huang. Clinical and demographic factors associated with COVID-19, severe COVID-19, and SARS-CoV-2 infection in adults: a secondary cross-protocol analysis of 4 randomized clinical trials. *JAMA network open*, 6(7):e2323349–e2323349, 2023.

H. Thieme. Renewal theorems for linear discrete Volterra equations. *Journal fur die Reine und Angewandte Mathematik*, 1984(353):55–84, 1984. ISSN 0075-4102. doi: 10.1515/crll.1984.353.55.

R. Thompson, J. Stockwin, R. van Gaalen, J. Polonsky, Z. Kamvar, P. Demarsh, E. Dahlqwist, S. Li, E. Miguel, T. Jombart, J. Lessler, S. Cauchemez, and A. Cori. Improved inference of time-varying reproduction numbers during infectious disease outbreaks. *Epidemics*, 29:100356, 2019. ISSN 1755-4365. doi: https://doi.org/10.1016/j.epidem.2019.100356. URL https://www.sciencedirect.com/science/article/pii/S1755436519300350.

R. Thompson, W. Hart, M. Keita, I. Fall, A. Gueye, D. Chamla, M. Mossoko, S. Ahuka-Mundeke, J. Nsio-Mbeta, T. Jombart, and J. Polonsky. Using real-time modelling to

inform the 2017 Ebola outbreak response in DR Congo. *Nature Communications*, 15(1): 5667, 2024.

R. N. Thompson, T. D. Hollingsworth, V. Isham, D. Arribas-Bel, B. Ashby, T. Britton, P. Challenor, L. H. K. Chappell, H. Clapham, N. J. Cunniffe, A. P. Dawid, C. A. Donnelly, R. M. Eggo, S. Funk, N. Gilbert, P. Glendinning, J. R. Gog, W. S. Hart, H. Heesterbeek, T. House, M. Keeling, I. Z. Kiss, M. E. Kretzschmar, A. L. Lloyd, E. S. McBryde, J. M. McCaw, T. J. McKinley, J. C. Miller, M. Morris, P. D. O'Neill, K. V. Parag, C. A. B. Pearson, L. Pellis, J. R. C. Pulliam, J. V. Ross, G. S. Tomba, B. W. Silverman, C. J. Struchiner, M. J. Tildesley, P. Trapman, C. R. Webb, D. Mollison, and O. Restif. Key questions for modelling COVID-19 exit strategies. *Proceedings of the Royal Society B: Biological Sciences*, 287(1932):20201405, 2020. doi: 10.1098/rspb.2020.1405. URL https://royalsocietypublishing.org/doi/abs/10.1098/rspb.2020.1405.

M. J. Tildesley and M. J. Keeling. Is R0 a good predictor of final epidemic size: Foot-and-mouth disease in the UK. *Journal of Theoretical Biology*, 258(4):623–629, 2009. ISSN 0022-5193. doi: https://doi.org/10.1016/j.jtbi.2009.02.019. URL https://www.sciencedirect.com/science/article/pii/S0022519309000824.

J. S. Tregoning, K. E. Flight, S. L. Higham, Z. Wang, and B. F. Pierce. Progress of the COVID-19 vaccine effort: viruses, vaccines and variants versus efficacy, effectiveness and escape. *Nature reviews. Immunology*, 21(10):626–636, 2021. ISSN 1474-1733. doi: 10.1038/s41577-021-00592-1. URL http://dx.doi.org/10.1038/s41577-021-00592-1.

United Nation. Global solidarity key to future pandemic preparedness, says UN chief. https://news.un.org/en/story/2024/12/1158586, 2024. Accessed: 03 February 2024.

S. A. van der Vegt, L. Dai, I. Bouros, H. J. Farm, R. Creswell, O. Dimdore-Miles, I. Cazimoglu, S. Bajaj, L. Hopkins, D. Seiferth, F. Cooper, C. Lok Lei, D. Gavaghan, and B. Lambert. Learning transmission dynamics modelling of COVID-19 using comomodels. *Mathematical Biosciences*, 349:108824, 2022.

E. van Leeuwen, P. Klepac, D. Thorrington, R. Pebody, and M. Baguelin. fluEvidenceSynthesis: An R package for evidence synthesis based analysis of epidemiological outbreaks. *PLOS Computational Biology*, 13(11):1–12, 11 2017. doi: 10.1371/journal.pcbi.1005838. URL https://doi.org/10.1371/journal.pcbi.1005838.

A. Vehtari, A. Gelman, D. Simpson, B. Carpenter, and P.-C. Bürkner. Rank-Normalization, Folding, and Localization: An Improved $\widehat{R}$ for Assessing Convergence of MCMC (with Discussion). *Bayesian Analysis*, 16(2):667 – 718, 2021. doi: 10.1214/20-BA1221. URL https://doi.org/10.1214/20-BA1221.

R. Verity, L. C. Okell, I. Dorigatti, P. Winskill, C. Whittaker, N. Imai, G. Cuomo-Dannenburg, H. Thompson, P. G. 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, and N. M. Ferguson. Estimates of the severity of COVID-19 disease. *medRxiv*, 2020. doi: 10.1101/2020.03.09.20033357. URL https://www.medrxiv.org/content/early/2020/03/13/2020.03.09.20033357.

P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, İ. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, and SciPy 1.0 Contributors. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272, 2020. doi: 10.1038/s41592-019-0686-2.

H. von Foerster. Some Remarks on Changing Populations. In *The Kinetics of Cellular Proliferation*, pages 382–407. Grune and Stratton, 1959.

J. Wallinga and M. Lipsitch. How generation intervals shape the relationship between growth rates and reproductive numbers. *Proceedings of the Royal Society B: Biological Sciences*, 274(1609):599–604, 2007. doi: 10.1098/rspb.2006.3754. URL https://royalsocietypublishing.org/doi/abs/10.1098/rspb.2006.3754.

L. Wang, Y. Zhou, J. He, B. Zhu, F. Wang, L. Tang, M. Kleinsasser, D. Barker, M. C. Eisenberg, and P. X. Song. An epidemiological forecast model and software assessing interventions on the COVID-19 epidemic in China. *Journal of Data Science*, 18(3):409–432, 2021. ISSN 1680-743X. doi: 10.6339/JDS.202007_18(3).0003.

J. Watson, P. F. Whiting, and J. E. Brush. Practice pointer: Interpreting a covid-19 test

result. *BMJ*, 2020, may 2020. ISSN 0959-8138. doi: 10.1136/bmj.m1808. URL https://www.medrxiv.org/content/early/2020/03/13/2020.03.09.20033357.

O. J. Watson, G. Barnsley, J. Toor, A. B. Hogan, P. Winskill, and A. C. Ghani. Global impact of the first year of COVID-19 vaccination: a mathematical modelling study. *Lancet Infect. Dis.*, June 2022.

H. J. Wearing, P. Rohani, and M. J. Keeling. Appropriate Models for the Management of Infectious Diseases. *PLOS Medicine*, 2(7), 07 2005. doi: 10.1371/journal.pmed.0020174. URL https://doi.org/10.1371/journal.pmed.0020174.

WHO. Coronavirus disease (covid-19) pandemic. https://www.who.int/europe/emergencies/situations/covid-19#:~:text=On%205%20May%202023%2C%20more,the%20definition%20of%20a%20PHEIC., 2023. Accessed: 03 February 2024.

WHO. Who director-general's opening remarks at the media briefing. https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing---10-december-2024, 2024a. Accessed: 03 February 2025.

WHO. Leading health agencies outline updated terminology for pathogens that transmit through the air. https://www.who.int/news/item/18-04-2024-leading-health-agencies-outline-updated-terminology-for-pathogens-that-transmit-through-the-air, 2024b. Accessed: 11 February 2024.

WHO. Vector-borne diseases. https://www.who.int/news-room/fact-sheets/detail/vector-borne-diseases#:~:text=Vector%2Dborne%20diseases%20account%20for,either%20parasites%2C%20bacteria%20or%20viruses, 2024c. Accessed: 11 February 2024.

WHO. WHO Director-General's opening remarks at the media briefing – 10 December 2024 — who.int. https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing---10-december-2024, 2024d. [Accessed 02-04-2025].

WHO. WHO Coronavirus (COVID-19) Dashboard. https://covid19.who.int/, 2025. [Online; accessed 10-Jan-2025].

S. Willard. *General Topology*. Dover Books on Mathematics. Dover Publications, Mineola, NY, Feb 2004.

M. Woolhouse, C. Dye, J.-F. Etard, T. Smith, J. Charlwood, G. Garnett, P. Hagan, J. Hii, P. Ndhlovu, R. Quinnell, C. Watts, S. Chandiwana, and R. Anderson. Heterogeneities in the transmission of infectious agents: Implications for the design of control programs. *Proceedings of the National Academy of Sciences*, 94(1):338–342, Jan. 1997. ISSN 1091-6490. doi: 10.1073/pnas.94.1.338. URL http://dx.doi.org/10.1073/pnas.94.1.338.

World Bank. Chapter 1. the economic impacts of the covid-19 crisis. https://www.worldbank.org/en/publication/wdr2022/brief/chapter-1-introduction-the-economic-impacts-of-the-covid-19-crisis#1, 2024. Accessed: 03 February 2024.

J. T. Wu, K. Leung, M. Bushman, N. Kishore, R. Niehus, P. M. de Salazar, B. J. Cowling, M. Lipsitch, and G. M. Leung. Estimating clinical severity of COVID-19 from the transmission dynamics in Wuhan, China. *Nature medicine*, 26(4):506–510, 2020. ISSN 1078-8956. doi: 10.1038/s41591-020-0822-7. URL http://dx.doi.org/10.1038/s41591-020-0822-7.

A | Appendix Chapter 1

A1 Thieme's renewal theorem

Here, we closely follow the more general approach presented in [B. Boldin and Metz, 2023] for a continuous-time renewal equation model with multiple groups.

We reframe our modelling framework as follows:

$$\begin{pmatrix} I^1(t) \\ \vdots \\ I^N(t) \end{pmatrix} = \int_{a=0}^{\infty} K(a) \begin{pmatrix} I^1(t-a) \\ \vdots \\ I^N(t-a) \end{pmatrix} da,$$

where the kernel of the renewal process is given by $K(a) = \gamma(t)C(t)W(a)$. If the kernel satisfies a set of conditions (which we give below), then the long-term solution can be shown to follow:

$$\begin{pmatrix} I_t^1 \\ \vdots \\ I_t^N \end{pmatrix} = e^{rt}\underline{\Phi}.$$

The kernel function must then satisfy the following five conditions:

1. $K(a)$ is a non-negative matrix for all $a > 0$. This is true, since $\gamma(t) > 0$, and $W(a)$ and $C(t)$ are non-negative matrices.

2. An element of $K(a)$ exceeds zero for some $a > 0$. Similar to before, this is also satisfied as $\int_{a=0}^{\infty} K(a)da = \gamma(t)C(t)\int_{a=0}^{\infty} W(a)da = \gamma(t)C(t)$ has at least two non-zero elements (otherwise all contacts are 0).

3. $\int_{a=0}^{\infty} \|K(a)\|da < \infty$, where we define the L_1-norm of a matrix M as $\|M\| = \max_{1\leq i\leq N} \sum_{j=1}^{N} \left|M^{(ji)}\right|$ and $\left|M^{(ji)}\right|$ is the absolute value of the $(j,i)^{\text{th}}$ element of matrix M. This condition is also met: by replacing the kernel function evaluated at a

with $\gamma(t)C(t)W(a)$, the conditions becomes

$$\int_{a=0}^{\infty} \|K(a)\|da = \gamma(t) \int_{a=0}^{\infty} \max_{1\leq i\leq N} \sum_{j=1}^{N} \left|C^{(ji)}(t)w^i(a)\right|da =$$
$$= \gamma_t\Big(\int_{a=0}^{\infty} w^i(a)da\Big) \max_{1\leq i\leq N} \sum_{j=1}^{N} \left|C^{(ji)}(t)\right| =$$
$$= \gamma_t \max_{1\leq i\leq N} \sum_{j=1}^{N} C^{(ji)}(t) < \infty$$

as required. For the last two conditions, Boldin et al [B. Boldin and Metz, 2023] uses a discrete time renewal process. We follow with similar assumptions which we assume to hold in the continuous time case as well.

4. There exists $a_0 \in \mathbb{N}$ such that $R(a_0)$ is a strictly positive matrix, where $R(a) = R * K(a) + K(a) = K * R(a) + K(a), \forall a > 0$ is defined as the resolvent of the kernel matrix $K(a)$. We use $R * K(a)$ to mean $R * K(a) = \sum_{t=0}^{a} R(a-t)K(t)$ as described in [Thieme, 1984]; also a strictly positive matrix implies that all its elements are greater than 0; that is $m_{ij} > 0, \forall i, j$ for a strictly positive matrix $M = (m_{ij})_{i,j}$.

5. Additionally, there exists $a_1 \in \mathbb{N}$ and $a_2 \in \mathbb{N}$ with greatest common divisor $(a_1, a_2) = 1$ such that both $R(a_1)$ and $R(a_2)$ have all elements greater than 0 and at least one element exceeding 0.

For the fifth condition, we recursively compute the values of resolvent matrix $R(a)$. For $a = 0$, we have that $R(0) = K(0) = \mathbb{O}_N$ as seen in [B. Boldin and Metz, 2023]. For the next value $a = 1$, we have that $R(1) = K(1) + W(0)R(1) + W(1)R(0) = K(1)$. As $K(1) = \gamma(t)C(t)W(1)$ is a non-negative matrix, then $R(1)$ is a non-negative matrix. From the definition of $R(a)$, by induction we conclude that $R(a)$ is a non-negative matrix for all $a \in \mathbb{N}$. We consider the first time point a' for which $K(a')$ has an element bigger than zero. From the second condition we have that such an a' exists.

If $a' = 1$, we have that both $R(2) = K(2) + K(1)R(1) = K(2) + K(1)^2$ and $R(3) = K(3) + K(1)R(2) + K(2)R(1) = K(3) + K(1)K(2) + K(2)K(1) + K(1)^3$ will have at least one element greater than zero and are non-negative matrices, and as $(2, 3) = 1$, the desired fifth condition is satisfied.

If $a' \geq 2$, however, $R(a') = K(a') + R * K(a')$ has an element bigger than zero as $R * K(a')$ is itself a non-negative matrix. Additionally, $R(a'+1) = K(a'+1) + K(a')R(1) + \ldots K(1)R(a') = K(a'+1) + K(a')K(1) = K(a'+1)$ as the condition on the value of a' implies that $K(a)$ is the identical zero matrix $\forall a < a'$ and $R(1) = K(1)$.

In terms of the fourth condition, we saw from before that $R(a)$ can be written an a sum of powers of the kernel matrices $K(b)$ for some bs with $b \leq a$ for all $a \in \mathbb{N}$. The kernel matrices are given by $K(a) = \gamma(t)C(t)W(a)$, where $W(a)$ is a diagonal matrix. If all population groups mix between each other, the contact matrix $C(t)$ will be strictly positive. Therefore, $K(a')$ will also be strictly positive, satisfying the fourth condition.

Otherwise, if the modelled population includes demographic groups that do not interact with any other groups, then the contact matrix $C(t)$ is a diagonal matrix and can be written as

$$C(t) = \left(\begin{array}{c|c} \begin{matrix} c_{11} & & \\ & \ddots & \\ & & c_{mm} \end{matrix} & \mathbf{0} \\ \hline \mathbf{0} & C^*(t) \end{array}\right). \tag{A.1}$$

As $W(a)$ is a diagonal matrix itself, then we would have that the kernel matrix evaluated for any time a is equal to

$$K(a) = \gamma(t)\left(\begin{array}{c|c} \begin{matrix} c_{11}w^1(a) & & \\ & \ddots & \\ & & c_{mm}w^m(a) \end{matrix} & \mathbf{0} \\ \hline \mathbf{0} & C^*(t)W_{m:,m:}(a) \end{array}\right).$$

Regardless of the number of how many times this type of matrix is raised to any power and how many times we sum it to compute $R(a)$, the top block of the resulting matrix will remain a diagonal matrix and therefore not satisfying the fourth condition imposed. Hence, our findings are confined to only fully mixing populations.

For example, suppose that the total contact matrix, $C(t)$, is diagonal, meaning that the groups do not interact. In this case, it is not generally true that the growth rate of new infections should be identical across groups, as eq. (1.36) supposes. For example, if groups have differing generation times and they do not interact, the infections will generally grow at different calendar rates in each group.

A2 Renewal equations and reproduction numbers for a discrete-time structured population model

When inferring R_t, the renewal equation frameworks typically used are discrete in time (usually with time steps of 1 day; see section 1.2). We now demonstrate that for a discrete-time version of the model presented so far, we obtain an analogous renewal equation as for the continuous-time model and the same expression for R_t.

When time and age are discrete, the analogous conservation equation for the population is a difference equation:

$$\frac{n(t+\Delta t, a+\Delta a) - n(t, a+\Delta a)}{\Delta t} + \frac{n(t, a+\Delta a) - n(t, a)}{\Delta a} = 0,$$

where $\Delta t = \Delta a$ represents a discrete unit of time. The density of new infections occurring at a time $t = T\Delta t$ is:

$$n(T\Delta t, 0) = \sum_{A=0}^{\infty} C_t \gamma_t w(A\Delta t) n(T\Delta t, A\Delta t). \tag{A.2}$$

Similar to the continuous-time population model, we only consider infections occurring after time $t = T\Delta t = 0$. The initial number of infections at time $T = 0$ can be written as $n(0, A\Delta t) = f(A\Delta t)$. The analogous characteristics for this discrete-time model are:

$$a = A\Delta t = \begin{cases} T\Delta t + A_0\Delta t, & \text{for } a > t \text{ i.e. } A > T \\ T\Delta t - T_0\Delta t, & \text{for } a \leq t \text{ i.e. } A \leq T. \end{cases} \tag{A.3}$$

where $A_0 > 0$ is the number of age units of an infection at time $t = 0$ and $T_0 \geq 0$ is the number of units of time at which an infection occurred. This allows us to define the number of infections as a function of t only, that is we can write $n(T\Delta t, A\Delta t) = n(T\Delta t, A(T)\Delta t) :=$

$n(T\Delta t)$, for a given choice of A_0, or T_0 respectively. Along each such defined *characteristic* lines, the number of infections is determined by a singular point in time τ, with $\tau = A_0\Delta t$ uniquely for $A > T$ and $\tau = T_0\Delta t$ for $A \leq T$. Therefore, the number of infections is conserved along each characteristic line, i.e.

$$n(t, a(t)) = \text{const.}$$

Using eqs. (A.3), we derive the constant population sizes: when $A > T$, A can be written $a = A\Delta t = T\Delta t + A_0\Delta t$ which implies $n(T\Delta t, T\Delta t + A_0\Delta t) = \text{const.}$; similarly when $T = 0$, the population size is:

$$n(0, A_0\Delta t) = f(A_0\Delta t) = f((A - T)\Delta t). \tag{A.4}$$

When $A \leq T$, we use instead the fact that $A\Delta t = T\Delta t - T_0\Delta t$; as $n(t, t - T_0\Delta t) = \text{const.}$, and, when $T = T_0$:

$$n(T_0\Delta t, 0) = n((T - A)\Delta t, 0). \tag{A.5}$$

Collecting eqs. (A.4) & (A.5), the solutions for the discrete population sizes for all infection onset times are given by either of the following two forms:

$$n(T\Delta t, A\Delta t) = \begin{cases} f((A - T)\Delta t), A > T, \\ n((T - A)\Delta t, 0), A \leq T. \end{cases} \tag{A.6}$$

Substituting the eqs. (A.6) back into (A.2), we obtain that

$$\begin{aligned} n(T\Delta t, 0) &= \sum_{A=0}^{T} b(A\Delta t) n(T\Delta t, A\Delta t) + \sum_{A=T+1}^{\infty} b(A\Delta t) n(T\Delta t, A\Delta t) = \\ &= \sum_{A=0}^{T} b(A\Delta t) n((T - A)\Delta t, 0) + \sum_{A=T+1}^{\infty} b(A\Delta t) f((A - T)\Delta t). \end{aligned}$$

Since similarly to the continuous-time case, for any $A' > T$, $b(A\Delta t) \approx 0$, the number of new infections observed at time $t = T\Delta t$ becomes:

$$n(T\Delta t, 0) = \sum_{A=0}^{T} b(A\Delta t) n((T - A)\Delta t, 0) + \sum_{A=T+1}^{\infty} 0 f((A - T)\Delta t) =$$

$$= \sum_{a=0}^{t} \gamma_t C_t w_a n(t-a,0) = \sum_{a=0}^{t} \gamma_t C_t w_a I_{t-a},$$

which is indeed identical to the deterministic form of the renewal equation process as described in eq. (1.1). The rest of the proof for the definition of the reproduction number and epidemic stability criterion $R_t = \sum_{a=0}^{\infty} \gamma_t C_t w_a = \gamma_t C_t \sum_{a=0}^{\infty} w_a = \gamma_t C_t$ follows according to the methodology presented in [B. Boldin and Metz, 2023].

For the multiple group population, a similar argument emerges [B. Boldin and Metz, 2023] – we arrive at the conclusion that the overall reproduction number $R_t = \rho(\sum_{a=0}^{\infty} \gamma_t C_t W_a) = \rho(\gamma_t C_t \sum_{a=0}^{\infty} W_a) = \rho(\gamma_t C_t) = \gamma_t \rho(C_t)$ is a correct criterion for the long-term behaviour of an epidemic.

A3 The typical generation time interval affects epidemic growth rates in calendar time

In Figure A1, we simulate the trajectories of the total number of new infections for a range of values of the generation time intervals means:

$$\overline{W}_1 = \begin{pmatrix} 3.3 \\ 5 \\ 7 \end{pmatrix}, \overline{W}_2 = \begin{pmatrix} 15.3 \\ 15.3 \\ 15.3 \end{pmatrix} \text{ and } \overline{W}_3 = \begin{pmatrix} 30 \\ 25.3 \\ 15 \end{pmatrix}.$$

In each case, the standard deviation of the generation time distribution was assumed to be 3.3 days.

We use the contact matrix given by eq. (1.37). The infection rate $\gamma(t)$ is also kept constant for all plotted scenarios. In the beginning of the simulation we see the different oscillatory patterns for each of our choices of the generation time intervals. However, as the time increases, for all $R(t)$ values and all selected $\overline{W}$s, the oscillations die out and the number of new infections approach 0 (panel A; epidemic decay), reach stable values (panel B; epidemic persistence) or continue to increase exponentially at a constant rate (panel C; epidemic growth).

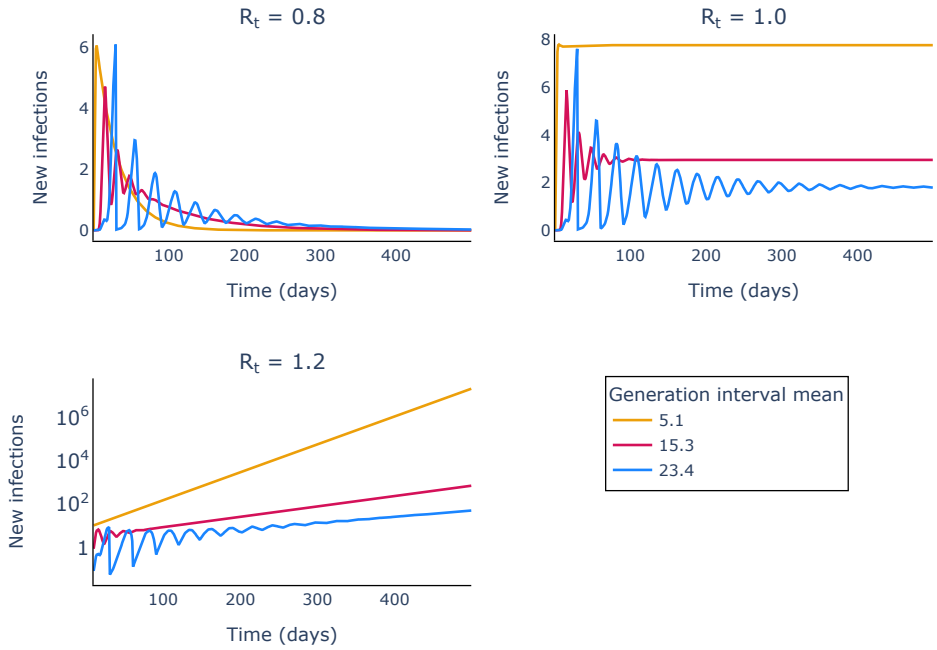


Figure A1: *The typical generation time interval affects epidemic growth rates in calendar time. Each panel represents a different $R(t)$ value, as indicated. Each line represents aggregate new infection counts (i.e. the aggregate across three populations) for a different simulation scenario: the line colours represent these scenarios and correspond to differing average generation time distributions across the three groups as indicated in the legend (in yellow, scenario 4; in red, scenario 2 and in blue, scenario 7 in Table 1.1).*

A4 R_t for a discrete-time stochastic model

We now derive R_t for a stochastic version of a renewal model. We consider only a renewal equation for a population comprising members of a single group since these results directly translate to the multiple group case. The renewal equation we assume is of the form:

$$I_t \sim \text{Poisson}\left(R\sum_{s=1}^{p}\omega_s I_{t-s}\right). \tag{A.7}$$

This means that for an arbitrary time $t+\tau$, we have:

$$I_{t+\tau} \sim \text{Poisson}\left(R\sum_{s=1}^{p}\omega_s I_{t+\tau-s}\right). \tag{A.8}$$

Taking expectations at time t yields:

$$\mathbb{E}_t\left[I_{t+\tau}\right] = R\sum_{s=1}^{p}\omega_s\mathbb{E}_t\left[I_{t+\tau-s}\right]. \tag{A.9}$$

Then denoting $M_j := \mathbb{E}_t\left[I_{t+\tau}\right]$ as the expectation at time t of the infection count at time $t+\tau$ yields a renewal equation in terms of this quantity:

$$M_j = R\sum_{s=1}^{p}\omega_s M_{j-s}. \tag{A.10}$$

We then suppose a long-run solution of the form $M_j \approx \lambda^j$, which results in:

$$\lambda^j = R\sum_{s=1}^{p}\omega_s\lambda^{j-s} \Rightarrow 1 = R\sum_{s=1}^{p}\omega_s\lambda^{-s}. \tag{A.11}$$

We can then define:

$$K(\lambda) = R\sum_{s=1}^{p}\omega_s\lambda^{-s}. \tag{A.12}$$

By construction, $K(1) = R$ and $K(\lambda)$ is a decreasing function of λ. This means that:

- If $R > 1$, $\lambda^* = \lambda$ s.t. $K(\lambda) = 1$ must satisfy $\lambda^* > 1$,
- If $R \leq 1$, $\lambda^* \leq 1$,

and similarly so for complex λ (see [B. Boldin and Metz, 2023] section 2.1). This means that we can consider R to satisfy the threshold conditions governing the long-run fate of the stochastic mean in the stochastic renewal model.

A5 Doubling times

We derive the relationship between the reproduction number and the doubling time $\tau_0 := \frac{1}{r}\ln(2)$ – the time taken for an epidemic to double in size. Rearranging this expression and substituting it in place of r in eq. (1.45) results in

$$\frac{1}{R(t)} = \frac{\rho\Big(C(t)\int_0^\infty W(a)2^{-\tau_0^{-1}a}da\Big)}{\rho(C(t))}.$$

A6 The relationship between $R(t)$ and the growth rate for the discrete-time deterministic model

We derive the relationship between $R(t)$ and the growth rate r for structured populations in the case of the discrete-time deterministic model. From [B. Boldin and Metz, 2023] section 2.2 and appendix the overall reproduction number of the system is defined by $R(t) = \rho(K(1))$, where $K(\lambda)$ is defined as

$$K(\lambda) = \gamma(t)C(t)\sum_{a=0}^{\infty} W(a)\lambda^{-a};$$

this definition for $K(\lambda)$ is simply the multivariate equivalent of eq. (A.12). Similar to appendix A4, λ is defined in terms of the long-term solution of the infection vector count, that is $\underline{I}_j = \lambda^j \underline{\Phi}$, where $\underline{\Phi}$ is a constant non-negative vector for large values of j. For the discrete-time deterministic model, the overall growth rate r is defined such that $\underline{I}_{t+1} = (1+r)\underline{I}_t$, therefore implying that $\lambda = 1 + r$. Hence, we have the following system of identities:

$$1 = \rho(K(\lambda)) = \gamma(t)\rho\Big(C(t)\sum_{a=0}^{\infty} W(a)\lambda^{-a}\Big) = \gamma(t)\rho\Big(C(t)\sum_{a=0}^{\infty} W(a)(1+r)^{-a}\Big) \quad \text{(A.13)}$$

$$R_t = \rho(K(1)) = \gamma(t)\rho\Big(C(t)\sum_{a=0}^{\infty} W(a)\Big) = \gamma(t)\rho(C(t)), \quad \text{(A.14)}$$

which when divided produce eq. (1.46).

A7 Empirical estimates of the growth rate from simulations of the stochastic model

In order to determine the empirical growth rate r of a population for which we know its overall reproduction number $R(t)$, we use the *scipy.integrate* and *scipy.optimize.minimize* functions [Virtanen et al., 2020]. The goal function for the optimiser is given by the follow-

ing equation in terms of r derived from eq. (1.45):

$$f(r) := \left| \frac{\rho(C(t))}{\rho\Big(C(t)\int_0^\infty W(a)\exp(-ra)da\Big)} - R(t) \right|.$$

The optimiser function will then return the value of r for which $f(r)$ is closest to zero and thus approximates the empirical growth rate of the population thus simulated.

A8 Identical rows matrix spectral radius

We are computing the spectral radius of a positive matrix with identical rows, which is defined as the maximum real eigenvalue of the matrix, i.e. the maximum value of λ where

$$\begin{bmatrix} a_1 & \dots & a_N \\ \vdots & \dots & \vdots \\ a_1 & \dots & a_N \end{bmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix} = \begin{pmatrix} a_1x_1 + \dots a_Nx_N \\ \dots \\ a_1x_1 + \dots a_Nx_N \end{pmatrix} = \lambda \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix}$$

which implies $a_1x_1 + \dots a_Nx_N = \lambda x_i, \forall i$. One solution is for $\lambda = 0$, which implies that $a_1x_1 + \dots a_Nx_N = 0$. Otherwise, if $\lambda \neq 0$ then that implies that $x_1 = \dots x_N$, therefore $a_1x_1 + \dots a_Nx_N = (\sum_{i=1}^N a_i)x_1 = \lambda x_1$. Since we cannot have $x_1 = 0$ (otherwise the eigenvector is the zero-vector, contradiction), then the only other eigenvalue of the matrix is given by $\lambda = \sum_{i=1}^N a_i > 0$ by the definition of the positive matrix. Therefore the spectral radius of the matrix

$$\rho\Big(\begin{bmatrix} a_1 & \dots & a_N \\ \vdots & \dots & \vdots \\ a_1 & \dots & a_N \end{bmatrix}\Big) = \sum_{i=1}^N a_i.$$

A9 Exact overall reproduction number for a two-group population

For the more general behaviour of a two group population we use the property that the spectral radius of a general 2x2 matrix $M = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is given by the maximum eigenvalue,

i.e. $\max\left(\frac{a+d\pm\sqrt{(a-d)^2+4bc}}{2}\right) = \frac{a+d+\sqrt{(a-d)^2+4bc}}{2}$. Therefore, if the contact matrix of the population is given by

$$C(t) = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}$$

then we have that $\rho(C(t)) = \frac{c_{11}+c_{22}+\sqrt{(c_{11}-c_{22})^2+4c_{12}c_{21}}}{2}$ and similarly, using the fact that

$$C(t)\int_0^\infty w(a)e^{-ra}da = \begin{bmatrix} c_{11}\int_0^\infty w^1(a)e^{-ra}da & c_{12}\int_0^\infty w^2(a)e^{-ra}da \\ c_{21}\int_0^\infty w^1(a)e^{-ra}da & c_{22}\int_0^\infty w^2(a)e^{-ra}da \end{bmatrix}$$

implies that

$$\rho\Big(C(t)\int_0^\infty w(a)e^{-ra}da\Big) = \frac{\sqrt{\Big(\int_0^\infty (c_{11}w^1(a)-c_{22}w^2(a))e^{-ra}da\Big)^2 + 4\Big(c_{12}\int_0^\infty w^1(a)e^{-ra}da\Big)\Big(c_{21}\int_0^\infty w^2(a)e^{-ra}da\Big)} + \int_0^\infty (c_{11}w^1(a)+c_{22}w^2(a))e^{-ra}da}{2}$$

Therefore, the relationship between the overall reproduction number $\gamma(t)$ and epidemic growth rate r is given by

$$\frac{1}{R(t)} = \frac{\sqrt{\Big(\int_0^\infty (c_{11}w^1(a)-c_{22}w^2(a))e^{-ra}da\Big)^2 + 4\Big(c_{12}\int_0^\infty w^1(a)e^{-ra}da\Big)\Big(c_{21}\int_0^\infty w^2(a)e^{-ra}da\Big)} + \int_0^\infty (c_{11}w^1(a)+c_{22}w^2(a))e^{-ra}da}{c_{11}+c_{22}+\sqrt{(c_{11}-c_{22})^2+4c_{12}c_{21}}}$$

More particularly if

$$C_t = c\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$$

then this relationship becomes

$$\frac{1}{R(t)} = \frac{\int_0^\infty (2cw^1(a)+cw^2(a))e^{-ra}da + \sqrt{4c^2\Big(\int_0^\infty w^1(a)e^{-ra}da\Big)^2 + c^2\Big(\int_0^\infty w^2(a)e^{-ra}da\Big)^2}}{3c+c\sqrt{5}}$$

$$= \frac{\int_0^\infty (2w^1(a) + w^2(a))e^{-ra}da + \sqrt{4\Big(\int_0^\infty w^1(a)e^{-ra}da\Big)^2 + \Big(\int_0^\infty w^2(a)e^{-ra}da\Big)^2}}{3+\sqrt{5}}$$

A10 Varying generation time distributions across groups.

We now explore how the relationship between r and $R(t)$ is affected by the composition of generation time distributions across the groups.

We consider three scenarios making different assumptions about the generation time distributions in each age group. For each scenario and for each group, we assume the generation time distribution is governed by a (discretised) gamma distribution with the same standard deviation, of 3.3 days; the mean generation time is allowed to vary by group. In the first scenario, we assume the mean generation time is lowest in the youngest age group and highest in the eldest; in the second, the mean generation time is the same in each group; in the third scenario, the generation time decreases with age. The means in each scenario are given in Table 1.1.

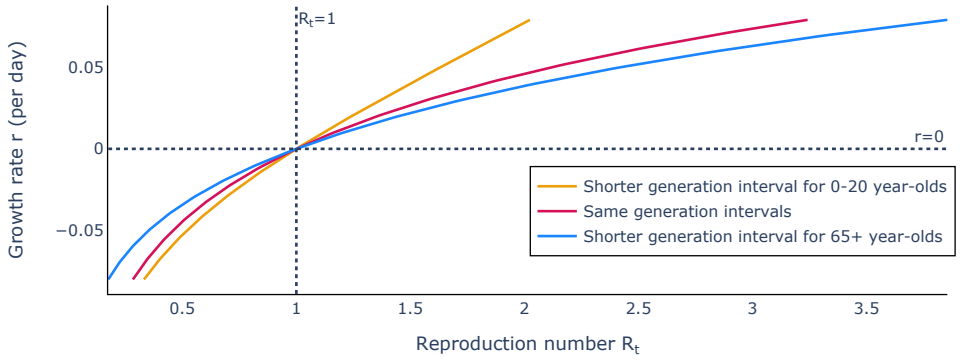


Figure A2: *The relationship between r and $R(t)$ is shaped by subpopulation generation time distributions. The yellow line corresponds to a population with a shorter generation time interval in the young population, equivalent to scenario 1. The red line corresponds to a simulation for which the generation time interval is identical in all population groups, with a mean of 15.3 as seen in scenario 2. Lastly, the blue line corresponds to a population with a shorter generation time interval in the old, equivalent to scenario 3.*

In Figure A2 we show the relationship between epidemic growth rate r and $R(t)$ implied by

eq. (1.45) according to each of the three scenarios. In all cases, $r = 0$ implies $R(t) = 1$. Since the 0-20 age group has the highest contact rates, most infections are generated from this group and it has the single greatest contribution to $R(t)$. Because of this, changes in the generation time of this group skew the relationship between r and $R(t)$: when the generation time interval is short in this group, the corresponding growth rate is high; and this decreases as their generation time is increased.

To further highlight this sensitivity, in Figure A3 we show how the epidemic growth rate responds to changes in the mean generation time in those aged 0-20 (left-hand panel) and those aged 65+ (right-hand panel); in each case, the mean generation time in the other groups remains fixed at 15.3 days. We consider three values of $R(t)$ for each panel (corresponding to the coloured lines).

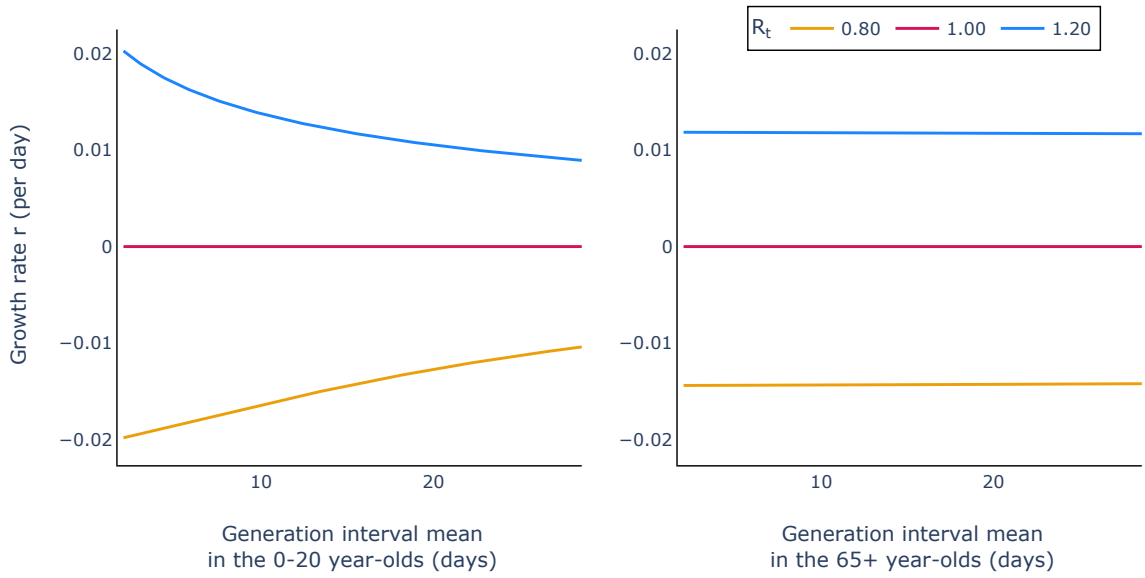


Figure A3: *The epidemic growth rate is sensitive to changes in the mean generation time in high contact groups. In both panels we show the relationship between the epidemic growth rate in calendar time and the mean generation time interval in a particular group: left, those aged 0-20; right, those aged over 65. When varying the generation time mean in one group (between* 2 *and* 28.6 *days, we hold the means in the other groups constant at* 15.3 *days. For all age groups, we considered an equal standard deviation of* 3.3*.*

B | Appendix Chapter 3

B1 The impact of the initial fraction of infection in the population on long-term R_t predictions

In the following we present an example population for which the short-term predicted behaviour of an epidemic, given by the overall reproduction number is less informative than using the group-specific reproduction number trajectories instead.

We consider a two-group population, in which group one is highly infectious (twice as likely to transmit compared to the second group), but there are fewer initial cases in this group (50 against 150 in group two). We fix the probability of a contact becoming an infection $\gamma_t = 0.088$ and simulate the number of new daily infections in each group over 30 days, using a multiple group renewal equation model; for this model, we assume Poisson-distributed temporal variations in the contact matrix (as seen in 3.2.2). The average contact matrix we used for this example is equal to

$$C = \begin{pmatrix} 12 & 0.6 \\ 6 & 4.8 \end{pmatrix}$$

and since the transmissibility vector $\underline{p}$ is given by $\underline{p} = (1, 0.5)$, then the average effective contact matrix is equal to

$$C^* = \begin{pmatrix} 12 & 0.6 \\ 3 & 2.4 \end{pmatrix}.$$

We use the same Ebola-like generation time interval we used to generate the previous synthetic dataset for both population groups, which is Gamma-distributed with mean 15.3 and standard deviation 9.3.

In Figure B1, we plot the overall reproduction number trajectories inferred using the synthetic infection data generated above against the known group-specific reproduction numbers. We observe that $R_t^{(1)}$ is consistently greater than one, that is the epidemic is always growing in the first population group, while $R_t^{(2)}$ is less than one, i.e. the epidemic in pop-

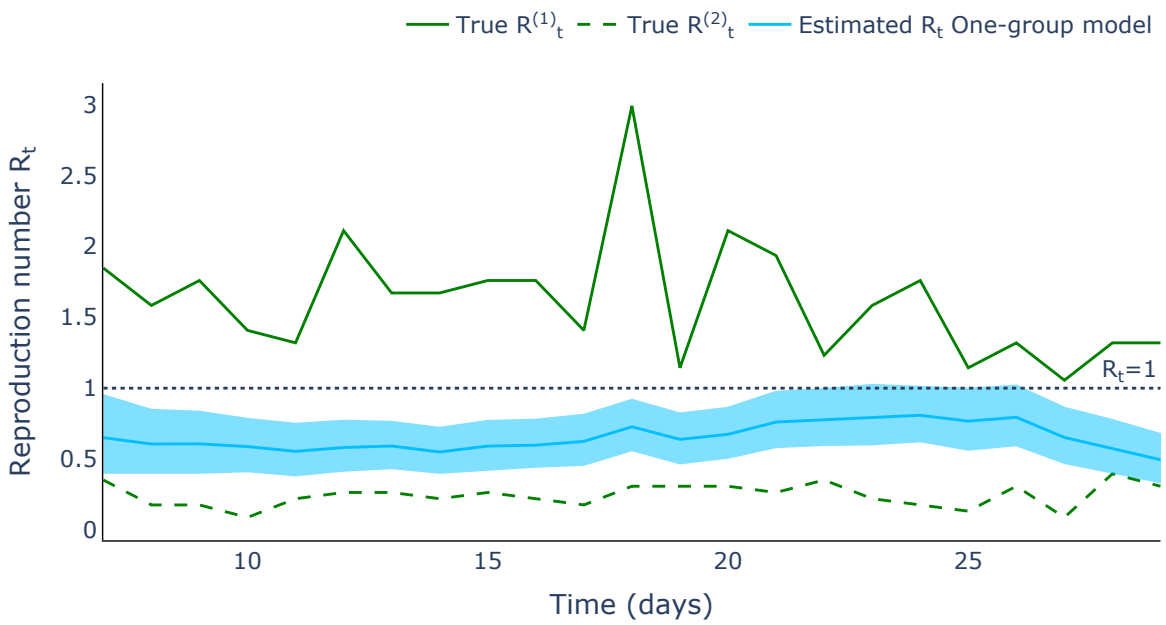


Figure B1: *Mean and* 95% *credible interval of the time-dependent reproduction number* R_t *trajectory of the Poisson renewal model with only one group (LIGHT BLUE) and the true* $R_t^{(a)}$ *trajectories of the underlying 2-population group model (GREEN).*

ulation group two is declining overtime. Additionally, the mean trajectory of the inferred overall R_t, as well as most (or the entirety) of the 95% credible interval lie below the $R_t = 1$ line for the entire simulation period; this would suggest erroneously that the epidemic is under control and will die down. Instead, there is an overall shift in the provenance of the new daily infections – from occurring primarily in group two, to being predominately from group one. The increasing long-term behaviour of the epidemic is reinforced by the value of the average overall reproduction number calculated using the known infection rate and average contact matrix when we account for the multiple group population structure; this reproduction number is given by $R_t = \gamma_t \rho(C^*) = 1.072 > 1$, which indicates that the epidemic will persist in the long run, as expected.

B2 Poisson-Binomial renewal equation model – A different semi-mechanistic approach

In the case of the conventional use of both Poisson and negative binomial processes, the arrival events are independent; however, we know that the number of cases observed on a

day depends on the number of past arrivals, i.e. not independent. This unsuitability is evident when trying to recover reproduction number profiles R_t that exhibit sharp drops in the daily values.

An option would be then to construct a semi-mechanistic renewal model, starting bottom-up. We consider first the distribution of the number of individuals encountered by an individual i on day t, $H_{t,i}$. In the absence of any temporal impact of external factors, i.e. non-pharmaceutical interventions or other changes in overall population behaviour, this value will be a positive integer centred around the mean; therefore, using a Poisson distribution is a good initial assumption:

$$H_{t,i} \sim \text{Pois}(\mu_{t,i})$$

where the mean $\mu_{t,i}$ depends on the individual interaction dynamics: i.e. use the mean number of contacts per day of an individuals in the same age group as individual i. This is also the point where we could include any effect of NPIs on the overall population mobility. We could go a step further and assume $\mu_{t,i}$ is stochastic as well to account for heterogeneity in individual behaviour within the same age category, e.g $\mu_{t,i} \sim \text{Gamma}(a, b)$ for some a and b depending on the contact matrix and NPIs-dependent mobility coefficient.

The next step in the procedure of finding a distribution of the number of daily new infections I_t is to write down a formula for computing the number of new cases that the infected individual i will go on to infect on day t. We will denote this quantity as $l_{t,i}$ and this is computed by going through each individual j who interacted with i on day t and assume it becomes infected with probability θ_{t-t_i}. That is

$$l_{t,i} = \sum_{j=1}^{H_{t,i}} Z_{t,i,j}, \text{ where } Z_{t,i,j} \sim \text{Bernoulli}(\theta_{t-t_i}).$$

The θ_{t-t_i} parameter is the probability of an infected individual with characteristics like i, infected on day t_i of infecting someone on day t; this will depend on both the generation time distribution w mentioned above and the infection parameter in the usual compartmental models. Due to the nature of Bernoulli i.i.d. random variables, we can compact the definition of $l_{t,i} \sim \text{Bin}(H_{t,i}, \theta_{t-t_i})$.

Finally, to account for all the number of new cases occurring on day t, I_t, we will sum the number of new infections $l_{t,i}$ generated by any individual i identified as a case before day t, that is

$$I_t = \sum_{i=1}^{I_{<t}} \sum_{j=1}^{H_{t,i}} Z_{t,i,j}, \text{ where } Z_{t,i,j} \sim \text{Bernoulli}(\theta_{t-t_i}).$$

Since this is a sum of independent Bernoulli random variables with non-equal parameters, the underlying distribution of I_t becomes a Poisson-binomial distribution with the following p.d.f

$$\mathbb{P}(I_t = k) = \sum_{i=1}^{I_{<t}} \sum_{j=1}^{H_{t,i}} \theta_{t-t_i}^{\mathbb{I}(\text{outcome of } j = \text{infection})} (1 - \theta_{t-t_i})^{1-\mathbb{I}(\text{outcome of } j = \text{infection})}$$

which is equivalent to

$$\mathbb{P}(I_t = k) = \sum_{A \in F_k} \prod_{i \in A} \theta_{t-t_i} \prod_{i \in A^c} (1 - \theta_{t-t_i}),$$

where F_k is the set of all subsets of size k of $\{1, 2, \ldots, \sum_{i=1}^{I_{<t}} \sum_{j=1}^{H_{t,i}}\}$, which matches the definition of a Poisson-binomial with $\sum_{i=1}^{I_{<t}} H_{t,i}$ Bernoulli trials with the following vector of success probabilities:

$$p = (\underbrace{\theta_{t-t_1}}_{\text{for } H_{t,1} \text{ times}}, \underbrace{\theta_{t-t_2}}_{\text{for } H_{t,2} \text{ times}}, \ldots, \underbrace{\theta_{t-t_{I_{<t}}}}_{\text{for } H_{t,I_{<t}} \text{ times}}).$$

$$p = (\underbrace{\theta_{t-1}}_{\text{for } \sum H_{t,j} | t_j = 1 \text{ times}}, \underbrace{\theta_{t-2}}_{\text{for } \sum H_{t,j} | t_j = 2 \text{ times}}, \ldots, \underbrace{\theta_{1}}_{\text{for } \sum H_{t,j} | t_j = t-1 \text{ times}}).$$

The mean and variance of the new process will thus become

$$\mathbb{E}[I_t | H_t] = \sum_{i=1}^{I_{<t}} H_{t,i} \theta_{t-t_i}$$

$$\mathbb{V}[I_t|H_t] = \sum_{i=1}^{I_{<t}} H_{t,i}\theta_{t-t_i}(1-\theta_{t-t_i}).$$

We assume that $\mu_{t,i}$ is known and fixed across individuals, i.e. μ_t. To compute the unconditional mean and probability of I_t, we use the Laws of Total Variance and Expectation and the independence of the $H_{t,i}$ random variables:

$$\begin{aligned}
\mathbb{E}[I_t] &= \mathbb{E}\Big[\mathbb{E}[I_t|H_t]\Big] = \mathbb{E}\Big[\sum_{i=1}^{I_{<t}} H_{t,i}\theta_{t-t_i}\Big] \\
&= \sum_{i=1}^{I_{<t}} \theta_{t-t_i}\mathbb{E}[H_{t,i}] = \sum_{i=1}^{I_{<t}} \theta_{t-t_i}\mu_t \\
\mathbb{V}[I_t] &= \mathbb{E}\Big[\mathbb{V}[I_t|H_t]\Big] + \mathbb{V}\Big[\mathbb{E}[I_t|H_t]\Big] \\
&= \mathbb{E}\Big[\sum_{i=1}^{I_{<t}} H_{t,i}\theta_{t-t_i}(1-\theta_{t-t_i})\Big] + \mathbb{V}\Big[\sum_{i=1}^{I_{<t}} H_{t,i}\theta_{t-t_i}\Big] \\
&= \sum_{i=1}^{I_{<t}} \theta_{t-t_i}(1-\theta_{t-t_i})\mathbb{E}[H_{t,i}] + \sum_{i=1}^{I_{<t}} \theta_{t-t_i}^2\mathbb{V}[H_{t,i}] \\
&= \sum_{i=1}^{I_{<t}} \theta_{t-t_i}(1-\theta_{t-t_i})\mu_t + \sum_{i=1}^{I_{<t}} \theta_{t-t_i}^2\mu_t = \sum_{i=1}^{I_{<t}} \theta_{t-t_i}\mu_t.
\end{aligned}$$

Instead, if we assume there is some uncertainty in μ_t, e.g. it is drawn from a distribution $\mu_t \sim \text{Gamma}(a,b)$ such that $\mathbb{V}[\mu_t] = \frac{\mathbb{E}[\mu_t]^2}{\phi}$ for some overdisperssion factor ϕ, then

$$\begin{aligned}
\mathbb{E}[I_t] &= \sum_{i=1}^{I_{<t}} \theta_{t-t_i}\mathbb{E}\Big[\mathbb{E}[H_{t,i}|\mu_t]\Big] = \sum_{i=1}^{I_{<t}} \theta_{t-t_i}\mathbb{E}[\mu_t] \\
\mathbb{V}[I_t] &= \sum_{i=1}^{I_{<t}} \theta_{t-t_i}(1-\theta_{t-t_i})\mathbb{E}[H_{t,i}] + \sum_{i=1}^{I_{<t}} \theta_{t-t_i}^2\mathbb{V}[H_{t,i}] \\
&= \sum_{i=1}^{I_{<t}} \theta_{t-t_i}(1-\theta_{t-t_i})\mathbb{E}\Big[\mathbb{E}[H_{t,i}|\mu_t]\Big] + \sum_{i=1}^{I_{<t}} \theta_{t-t_i}^2\left(\mathbb{E}\Big[\mathbb{V}[H_{t,i}|\mu_t]\Big] + \mathbb{V}\Big[\mathbb{E}[H_{t,i}|\mu_t]\Big]\right) \\
&= \sum_{i=1}^{I_{<t}} \theta_{t-t_i}(1-\theta_{t-t_i})\mathbb{E}[\mu_t] + \sum_{i=1}^{I_{<t}} \theta_{t-t_i}^2(\mathbb{E}[\mu_t] + \mathbb{V}[\mu_t])
\end{aligned}$$

$$= \sum_{i=1}^{I_{<t}} \theta_{t-t_i} \mathbb{E}[\mu_t] + \sum_{i=1}^{I_{<t}} \theta_{t-t_i}^2 \frac{\mathbb{E}[\mu_t]}{\phi}$$
$$= \mathbb{V}_{\text{Pois}}[I_t] + \sum_{i=1}^{I_{<t}} \theta_{t-t_i}^2 \frac{\mathbb{E}[\mu_t]^2}{\phi}.$$

There is no Poisson-binomial distribution in either of the Python **NumPy** or **SciPy** packages, so in order to simulate this distribution, we can use the off-the shelf-package poibin for forward simulations.

One of the usual byproducts of using renewal models for epidemiology is a simple way in which to compute the timelines of the instantaneous reproduction number, which is defined as the as the average number of secondary cases that would be generated by an infected case at time t assuming that future transmission remains the same as at time t. That is, if we start with the individual reproduction number of i

$$R_{t,i} = \sum_{t'=t}^{\infty} l_{t',i},$$

then

$$R_t = \mathbb{E}[R_{t,i}] = \sum_{t'=t}^{\infty} \mathbb{E}[l_{t',i}] = \sum_{t'=t}^{\infty} \theta_{t'} \mathbb{E}[H_{t',i}] = \mu_i \sum_{t'=t}^{\infty} \theta_{t'}.$$

For the inference, because we still work in a hierarchical model, we can use an inference algorithm such as Approximate Bayesian Computation ('ABC').

C | Appendix I Chapter 4

C1 PHE Model

In this section, we reconstruct a model previously used by Public Health England (PHE) and developed in partnership with the University of Cambridge to model the transmission dynamics of COVID-19 in early 2020 in the UK. To faithfully reimplement this model, which from now on we refer to as the PHE model, we follow the model structure detailed in the original paper [Birrell et al., 2021].

C1.1 PHE Model Structure

The PHE model has at its core a discretization of an ODE-based SEIR model. However, as we can see in the model diagram Figure C1, the E and I compartments are split into two compartments each; this adds a measure of lag between individuals moving through these two epidemiological states. The reason for this choice is not explained in the original paper, but according to [van Leeuwen et al., 2017] and [Anderson et al., 2020], this is done so that in the individual-based, stochastic version of this model, the corresponding waiting time of an individual in either of the exposed or infectious states is gamma-distributed, instead of exponential as it would have been when using only one compartment for each of E and I. This is a more realistic assumption for general respiratory disease – and therefore also applicable in the modelling of COVID-19, and follows as the sum of two independently and identically distributed exponential random variables with mean μ will be gamma$(2, \mu)$-distributed.

We can visualise the flow of individuals across the different compartments as:

where $\lambda_{r,t,i}$, $2/d_L$ and $2/d_I$ are the rates of progression between the compartments – more precise definitions for how each of these are computed will be given in the following sections. The dotted arrows illustrate the effect of the infectious individual counts on the transmission rate between the S and E^1 compartment.

The population is further stratified into age groups and regions in order to account for its heterogeneity and differences in the behaviour of individuals. The system is numerically

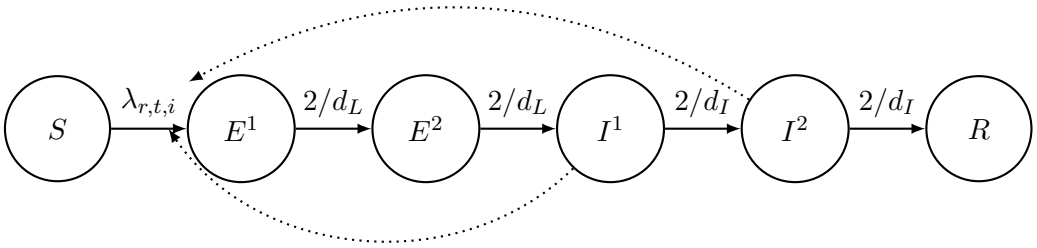


Figure C1: PHE Model Structure

solved using Euler's explicit method with a hard-wired time-step of half a day:

$$
\begin{aligned}
S_{r,t_k,i} &= S_{r,t_{k-1},i}(1 - \lambda_{r,t_{k-1},i}\delta t) \\
E^1_{r,t_k,i} &= E^1_{r,t_{k-1},i}\left(1 - \frac{2\delta t}{d_L}\right) + S_{r,t_{k-1},i}\lambda_{r,t_{k-1},i}\delta t \\
E^2_{r,t_k,i} &= E^2_{r,t_{k-1},i}\left(1 - \frac{2\delta t}{d_L}\right) + E^1_{r,t_{k-1},i}\frac{2\delta t}{d_L} \\
I^1_{r,t_k,i} &= I^1_{r,t_{k-1},i}\left(1 - \frac{2\delta t}{d_I}\right) + E^2_{r,t_{k-1},i}\frac{2\delta t}{d_L} \\
I^2_{r,t_k,i} &= I^2_{r,t_{k-1},i}\left(1 - \frac{2\delta t}{d_I}\right) + I^1_{r,t_{k-1},i}\frac{2\delta t}{d_I} \\
R_{r,t_k,i} &= I^2_{r,t_{k-1},i}\frac{2\delta t}{d_I},
\end{aligned}
\tag{C.1}
$$

where $S_{r,t_k,i}$ indicate the number of susceptibles in age group i present in region r at the discrete time t_k (and similarly for $E^1_{r,t_k,i}$, $E^2_{r,t_k,i}$, $I^1_{r,t_k,i}$, $I^2_{r,t_k,i}$, and $R_{r,t_k,i}$), and δt is the length of the time-step used by the Euler's solver, i.e. $\delta t = 0.5$ days.

In this model no ageing or population migration are considered. The number of deaths is negligible relative to the population size, and so the total numbers of individuals in each region and age category remains constant over time. No initial conditions are given for any of the compartments in the original paper [Birrell et al., 2021].

To reconstruct the model in our 'epimodels' Python library, we begin by first creating an infrastructure for storing timelines of contact and regional matrices which are then used in the computation of the time-dependent force of infection $\lambda_{r,t_k,i}$ parameter of the model,

according to the following schema:

$$\lambda_{r,t_k,i} = 1 - \prod_{j=1}^{n_A} \Big[(1 - b_{r,i,j}^{t_k})^{I_{r,t_k,j}^1 + I_{r,t_k,j}^2} \Big],$$

where $b_{r,i,j}^{t_k} \in (0,1)$ is the probability of a susceptible individual in region r of age group i being infected by an individual in age group j at time t_k. This probability is given by:

$$b_{r,i,j}^{t_k} = \beta_{t_k,r} \frac{R_{0,r} \tilde{C}_{r,i,j}^{t_k}}{R_{0,r}^*}, \tag{C.2}$$

where $\beta_{t_k,r} > 0$ is a parameter encapsulating further temporal fluctuations in transmission for all ages in region r at time t_k, $R_{0,r}$ is the basic reproduction number of the virus in region r and $R_{0,r}^*$ is the dominant (maximum) eigenvalue of the initial next generation matrix $\Lambda_{0,r}$ whose elements are defined as

$$\Lambda_{0,r,i,j} = N_{r,i} \tilde{C}_{r,i,j}^{t_0} d_I.$$

$N_{r,i}$ is the total population size in region r in age group i, d_I is the mean infectious period and is common across all age groups, and $\tilde{C}_{r,i,j}^{t_0}$ is the expected number of new infections in age group i caused by an infectious individual in age group j in region r. This last quantity can be expressed in terms of the contact matrix and relative susceptibility to the disease matrix as follows

$$\tilde{C}_{r,i,j}^{t_k} = C_{i,j}^{t_k} M_{r,i,j}^{t_k},$$

where $C_{i,j}^{t_k}$ prescribes the average number of contacts in age group i per person in age group j, while $M_{r,i,j}^{t_k}$ gives the relative susceptibility to infection of a person in age group i from an individual in age group j assuming contact.

The temporal fluctuation parameters $\beta_{t_k,r}$ are captured by introducing weekly change-points for each region, that is at time t_k, the β parameter corresponds to that of the week w_k that contains the time t_k. The weekly values change smoothly according to a log-normal distribution

$$\log(\beta_{w_k,r}) \sim N(\log\big(\beta_{w_{k-1},r}\big), \sigma_\beta^2), \tag{C.3}$$

where $\beta_{w_{\text{lock}},r} = 1$ for the week w_{lock} and all weeks after the institution of the UK lockdown.

C1.1.1 Contact Matrices

Contact matrices are used to keep track of the expected number of contacts between people of different ages at various stages of the epidemic. The idea behind their use is that it is reasonable to assume that a retired person will interact differently with other people than a working adult, i.e. in this example, fewer close contacts per day on average. Contact matrices account for the quantitative behavioural difference in the population given a particular age-structure. The $(i,j)^{\text{th}}$ element of the matrix at time point t_k is:

$$C_{i,j}^{t_k} = \mathbb{E}\left(\begin{array}{l}\text{\# contacts of an individual in age group} \\ i \text{ with an individual in age group } j\end{array}\right).$$

Contact matrices can also vary depending on the time and the geographical location of the population they reference. Measures implemented by the authorities can prohibit interactions between certain groups, leading to a decrease in the elements of the contact matrix for the age groups that the measure affects. Additionally, the proportion of rural versus urban population in a geographical area can also impact the average contact rates for the region, as the profile of contacts between a rural and urban community may differ significantly.

The PHE model requires the data for the COVID-19 pandemic be split according to the main 7 historical and administrative regions in England, i.e. East of England (EE), North East (NE), North West (NW), South East (SE), South West (SW), the Midlands (Mid) and London, as seen in Figure C2. In the available governmental data sources, all information is provided at district-level; hence, we make the following assumptions:

- The region of Yorkshire and the Humber is included in the North East region.
- East Midlands and West Midlands are collapsed into a single region, the Midlands.

Preventive measures affect people's behaviour [Flaxman et al., 2020; Brauner et al., 2021]. Therefore, contact matrices are used in the PHE model to compute the $\lambda_{r,t_k,i}$ parameter from equations (C.1); we produce timecourses of these matrices. In the original paper the changes in contacts are recorded weekly, which we also do in our implementation of the model. The baseline is represented by POLYMOD [Prem et al., 2021] contact matrices for different activities (school, home, work, others). The sum of these matrices gives the

Figure C2: *Geographical tessellation of England into PHE regions.*

overall contact matrix of the population. This matrix corresponds to the scenario when no interventions for the prevention of the spread of the epidemic are made. For the case when interventions are applied, we multiply all elements of each of the activity-specific contact matrices by the same scalar, which is given by the percentage difference in daily contacts associated with that activity – the current total contact matrix is then computed as the sum of the newly computed activity-specific contact matrices, just as before. In the PHE model, the population is age-structured as follows:

$$\begin{array}{ll} [0,1) & [1,5) \\ [5,15) & [15,25) \\ [25,45) & [45,65) \\ [65,75) & [75,+). \end{array}$$

In Figure C3 we plot the total number of individuals by age groups in each of the seven regions considered. London differentiates itself from all other regions, since there are significantly more people aged between 25 to 45 than between 45 and 65, unlike any of the other PHE regions considered.

If a lockdown or other similar measures are in place, a reduction in the number of contacts across different age groups is to be expected. The way we take into account such changes

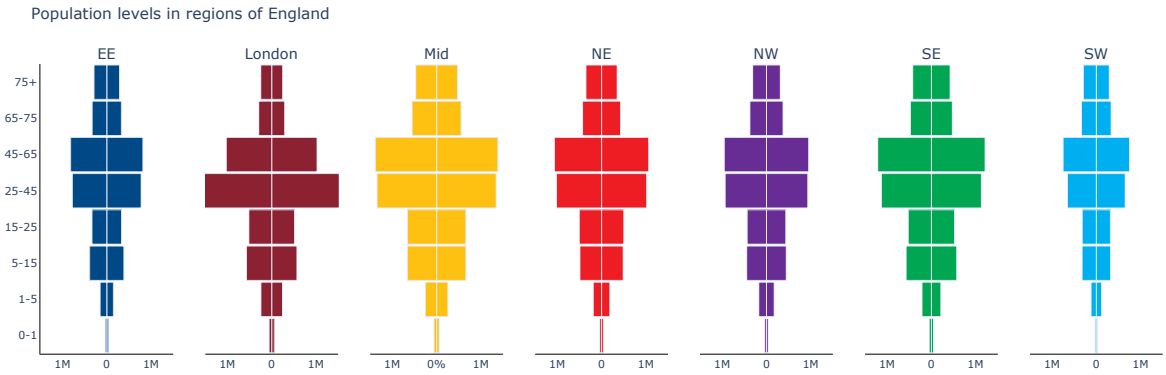


Figure C3: *The age-structure of population of the regions of England, represented by total number of individuals in each age group. The left to right separation in each bar plot indicates the male (left) to female (right) number of individuals in each subpopulation, if an exact* $50\% - 50\%$ *split across genders is assumed.*

is using Google mobility data [Google] collected across the whole of the UK and separated into subregions. We match each type of activity with the ones for the correct activity contact matrix (i.e. school and work matrices fitted according to workplace data, home with residential data, and others with the average of the other activities data).

To account for the weekly aspect of the changes in baseline activity levels, we take averages over 7-day intervals for the available mobility data. If there are missing data, then the previously recorded difference is maintained. To collect the activity level change over the whole region in the given week, weighted averages with respect to the population of each local authority in the Google data are computed. The parsing of all these necessary data is done via a Python script in the 'epimodels' *data* submodule which produces separate scripts with the computed contact matrix organised by region and week.

Figure C4 shows how the total expected number of contacts of any particular individual fluctuates with respect to time across different age groups and regions. We see a sharp decrease in the number of contacts for all age groups during the first lockdown period in the UK (23^{rd} of March – 10^{th} of May 2020), followed by a slight increase in the levels once the measure has been lifted. For the more active age groups, we see that this fluctuation is more pronounced and, with the easing of the lockdown, the levels of activity return approximately to the baseline. For the first two and especially the last two age groups though, the number of contacts remain visibly below the baseline level throughout the time period considered. This can be explained by the fact that school has been interrupted for most of this time interval and the government strongly encouraged the elderly to minimise the number of people they

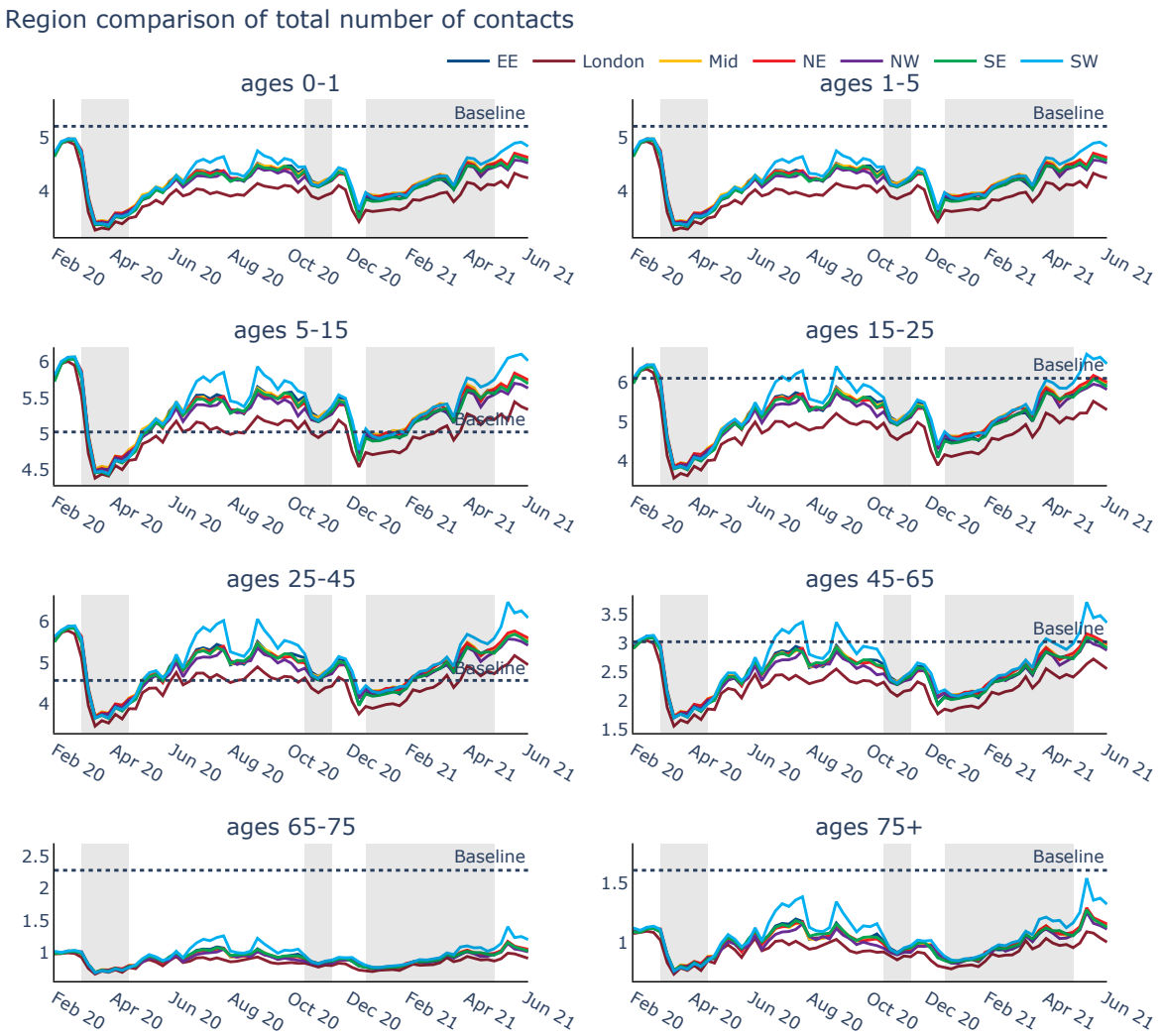


Figure C4: *Mean number of contacts per individual, in terms of regions and age groups. Dotted lines represent the levels corresponding to the baseline POLYMOD contact matrix for all locations. The coloured trajectories represent how this number changes with respect to time across each region. The dashed areas indicate the periods of lockdown in England.*

come into contact with.

C1.1.2 Regional Matrices and UniGen Matrices

A different type of matrix referenced in the model is what we call regional matrices in the 'epimodels' framework. These matrices account for the relative susceptibility of people in different age groups to be infected by another if contact is assumed. These matrices are both

region-specific and time-dependent:

$$M_r^{t_k} = \{M_{r,i,j}^{t_k}\}, \text{ where } M_{r,i,j}^{t_k} = \begin{array}{l}\text{susceptibility to infection of a person in age group i}\\ \text{from an individual in age group j assuming contact.}\end{array}$$

This matrix class is coded similarly to that for the contact matrices, with the added region feature, which is absent in the 'epimodels' implementation of the contact matrix. Element-wise multiplication of the regional with the appropriate contact matrix creates a next generation matrix (*UniNextGen*), which is then used to compute other quantities relevant to the PHE model, such as the instantaneous reproduction number R_t.

As our data preparation process already incorporates the time-varying, regional aspect of contacts [Google] directly on the POLYMOD contact matrix [Prem et al., 2021], for the reimplemented PHE model, we switch the contact matrix and the regional matrix around in the computation of the next generation matrix $\tilde{C}_r^{t_k}$. We assume that regional matrices have all elements equal to 1 – that is, regardless of age group or region, if there is contact between two individuals, the virus will be transmitted with odds equal to $m_r = 1$. We choose this value because, according to [Birrell et al., 2021],

$$M_{r,i,j}^{t_k} = 1 - 1_{t_k \geqslant t_{\text{loc}}}(1 - m_r),$$

where the prior for the m_r parameters is a gamma$(4, 4)$-distribution (as stated in the original paper [Birrell et al., 2021]), which has a mean value of 1. The odds of transmission m_r can be easily changed to depend on both region and age group should we desire a more realistic profiling of the epidemic.

These matrices, as well as the contact matrix class of objects, are all hosted in the submodule *_setup_matrix.py* of 'epimodels', which contains all auxiliary methods used for setting up the epidemiological models in the package.

C1.1.3 PHE Model Class

Our reimplementation of the PHE model (the *PheSEIRModel* class in the *phe_model.py* file) is structured according to the SEIR example model included in PINTS[1] [Clerx et al., 2019].

[1] PINTS (Probabilistic Inference on Noisy Time-Series) is a framework for optimisation and Bayesian inference on ODE models of noisy time-series. More details about this package can be found in [Clerx et al., 2019]

This choice provides coding robustness and consistency.

We provide two distinct solver methods of the PHE model:

- a prescribed discrete time-step implementation of the ODE system (C.1) (*my_solver*), as presented in the original PHE paper [Birrell et al., 2021] which uses Euler's method to solve for each compartment (as presented in Appendix E1), and
- a continuous-time version of the system which uses *SciPy* [Virtanen et al., 2020] ODE solvers (*scipy_solver*). We select the explicit Runge-Kutta method of order 5(4) [Dormand and Prince, 1980] as the default solver for computing the values in each compartment at specific time points.

The main method of the PHE model class is *simulate* which takes as arguments the parameters of the model fed in the form of a *PheParameterController*. The *PheParameterController* parameters are classified into:

- **the initial conditions**: the initial values of the compartments;
- **the regional parameters**: the regional base reproduction number R_0, the region for which we choose to simulate, the daily regional fluctuation parameters β, and the times we wish to run the model;
- **the disease parameters**: the latent d_L and infectious d_I period;
- **the simulation-specific parameters**: the method of the solver used (the *SciPy* ODE solver, or 'my-solver' for the PHE paper solver), and the discrete time-step of the homemade solver method.

The time-dependent regional and contact matrices together with the time points of their changes are implemented beforehand; similarly for fitting the model with the age groups and the regions targeted by the model.

Apart from the forward simulation method, which returns timelines of the numbers of individuals at specified time points for a given region, we implement three additional types of methods in the *PheSEIRModel* class. The first one is for computing the number of new infections in a specified region at the simulated time points for the different age categories, according to

$$\Delta^{\text{infec}}_{r,t_k,i} = S_{r,t_k,i} p^{\lambda}_{r,t_k,i},$$

where the $p^{\lambda}_{r,t_k,i}$ represents the proportion of the currently susceptible population in age group i that becomes infected during the time step starting at t_k. This quantity can be expressed in terms of the $\lambda_{r,t_{k-1},i}$ parameter of the model as

$$p^{\lambda}_{r,t_k,i} = \left(1 - \prod_{j=1}^{n_A} \left[(1 - b^{t_k}_{r,i,j})^{I^1_{r,t_k,j}+I^2_{r,t_k,j}}\right]\right)\delta t = \lambda_{r,t_{k-1},i}\delta t.$$

The second type of methods are aimed at generating trajectories of the daily number of deaths registered in the region we simulate. We include methods for sampling and computing the expected number of deaths per age group at a given time point, as well as a method for computing the vector of the log-likelihood of the observed number of deaths at that time point for the different age groups. The age-structured number of deaths is assumed to follow a negative-binomial distribution with the following mean and variance:

$$\begin{aligned} \mathbb{E}(X_{r,t_k,i}) &= \mu_{r,t_k,i} = p_i \sum_{l=0}^{k} f_{k-l}\Delta^{\text{infec}}_{r,t_l,i}, \\ \mathbb{V}(X_{r,t_k,i}) &= (\eta + 1)\mu_{r,t_k,i}, \end{aligned} \tag{C.4}$$

where $\Delta^{\text{infec}}_{r,t_l,i}$ is the number of new infections in age group i in region r at timepoint t_l, p_i is the age-specific fatality ratio of age group i, f_{k-l} is the probability of the individual dying $k-l$ days after contracting the infection, and η is the dispersion factor of the negative binomial distribution.

These need to be prescribed as parameters for the methods. In the original paper [Birrell et al., 2021], the prior distribution for the dispersion parameter η is given by a gamma$(1, 0.2)$-distribution; hence, we select its prior mean, 5, to be the value of η parameter in the model implementation. The age-specific fatality-ratio is an adapted version of the infection fatality ratio (IFR) vector of estimates computed by [Verity et al., 2020]. For the vector of probabilities of death f, we assume a gamma-distribution with mean time-to-death from symptom onset 15 and standard deviation 12.

Similarly to the case of the number of deaths, the final set of methods implemented for the PHE model focuses on the number of identified positive test results in a given region on a specific day. These methods take as input not only the age-structured number of tests conducted per day and the output of a simulation of the model, but also specific information

about the infection survey test accuracy, i.e. the sensitivity (true-positive rate) and specificity (true-negative rate) of a COVID-19 test. For these last two parameters, we assume a sensitivity of 70% and a 95% specificity, as in [Watson et al., 2020]. The distribution of the number of positive tests is considered to be binomial with parameters:

$$Y_{r,t_k,i} \sim \text{Bin}\left(n_{r,t_k,i}, k_{\text{sens}}\left(1 - \frac{I^1_{r,t_k,i} + I^2_{r,t_k,i}}{N_{r,i}}\right) + (1 - k_{\text{spec}})\frac{I^1_{r,t_k,i} + I^2_{r,t_k,i}}{N_{r,i}}\right), \tag{C.5}$$

where $n_{r,t_k,i}$ is the number of tests collected from individuals of age i, in region r on day t_k, k_{sens} and k_{spec} are the test sensitivity and specificity respectively. This differs slightly from the definition used in [Birrell et al., 2021] – which looks at the number of people not previously exposed by the virus i.e. antibody-testing; we use REACT1 data [Nicholson et al., 2021] which represents antigen-testing and therefore a positive test result shows an individual is currently infected and not yet recovered.

Having established all these elements for constructing the infrastructure of the PHE model, we can now analyse some example implementations of the model.

C2 Roche Model

The second model we reimplemented with the use of our 'epimodels' Python library is one developed by our industrial partner at F. Hoffmann-La Roche Ltd which we will refer to as the Roche model throughout this thesis. We reconstruct the Roche model based on the preprint published in 2020 [Lemenuel-Diot et al., 2020]. However, the model presented in the paper lacks age- or region-structure. Therefore, we extend our implementation of the model to include these features, to allow for a comprehensive model comparison analysis in Chapter 4.

C2.1 Roche Model Structure

Similarly to the previous PHE model [Birrell et al., 2021], the Roche model is a deterministic compartmental model which accounts for both infectious and exposed individuals. The exposed compartment E is introduced to reflect the observed delay between the time an individual becomes infected and the moment they can start spreading the disease to other individuals – an important feature of COVID-19. However, unlike the PHE model, the

Roche model includes a compartment D for people that die. Therefore, the Roche model is an extension of a basic SEIRD compartmental model.

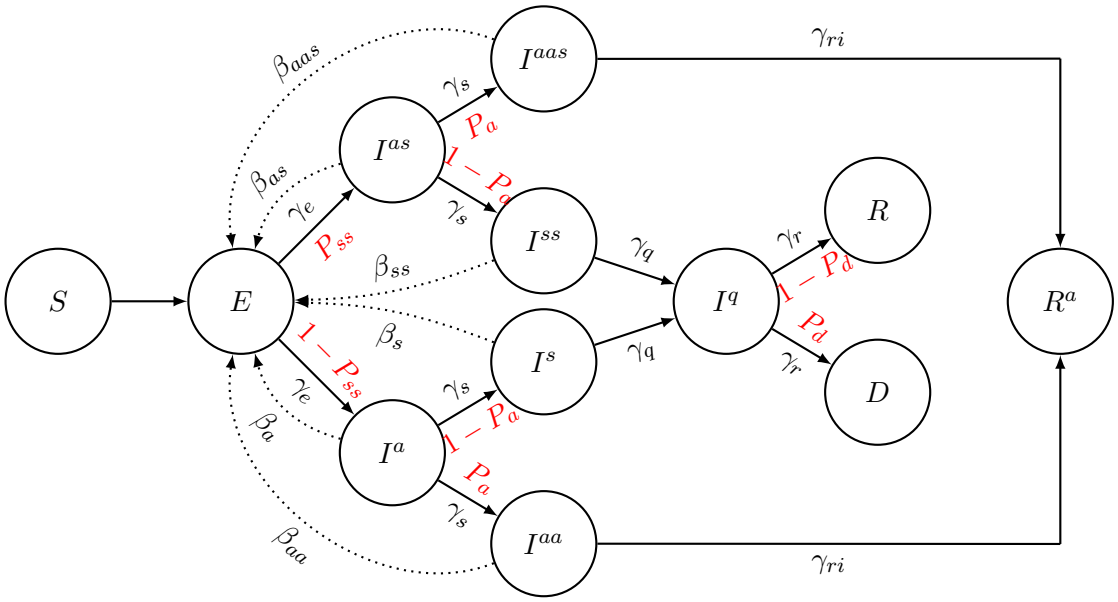


Figure C5: Roche Model Structure

SARS-CoV-2 is both an airborne and mildly contagious virus, which means that a non-negligible number of infections are caused by individuals who either have many close contacts, or do not manifest any symptoms. Therefore, differentiating between asymptomatic, super-spreader and normal transmissions improves the accuracy of the epidemic modelling. The Roche model assumes both asymptomatics and super-spreaders, with an additional stage of presymptomatic infection during which the individual can spread the virus but is not manifesting any symptoms. Passing the presymptomatic stage they either develop symptoms (i.e. moves to the symptomatic infectious compartment) or continue in the same state until they recover (i.e. moves to the asymptomatic infectious compartment). Moreover, if an individual becomes symptomatic, they proceed to the quarantined compartment; at this stage they can no longer infect any other susceptibles. From the quarantined compartment, the individual goes on to either recover or die. Worth mentioning is the inclusion of two recovered compartments – one for asymptomatic and one for symptomatic. This is because in practice the observed number of recovered individuals is mostly composed of symptomatic cases, since they are more likely to have tested and be recorded as having contracted the virus.

In Figure C5 we indicate how individuals move across the different compartments. Similar to the PHE model, we do not take into account population ageing or inter-regional migration. The total number of individuals in the model is conserved because we choose not to model births or deaths that are not due to COVID-19. For each region, the ODE system describing the Roche model is summarised by the following 12 equations:

$$
\begin{aligned}
\frac{dS_i}{dt} &= \sum_j C_{ij}\Big(-\frac{\beta_a}{N} S_i {I^a}_j - \frac{\beta_{aa}}{N} S_i {I^{aa}}_j - \frac{\beta_s}{N} S_i {I^s}_j - \frac{\beta_{as}}{N} S_i {I^{as}}_j - \frac{\beta_{aas}}{N} S_i {I^{aas}}_j \\
&\quad - \frac{\beta_{ss}}{N} S_i {I^{ss}}_j \Big) \\
\frac{dE_i}{dt} &= -\gamma_e E_i + \sum_j C_{ij}\Big(\frac{\beta_a}{N} S_i {I^a}_j + \frac{\beta_{aa}}{N} S_i {I^{aa}}_j + \frac{\beta_s}{N} S_i {I^s}_j + \frac{\beta_{as}}{N} S_i {I^{as}}_j \\
&\quad + \frac{\beta_{aas}}{N} S_i {I^{aas}}_j + \frac{\beta_{ss}}{N} S_i {I^{ss}}_j \Big) \\
\frac{d{I^a}_i}{dt} &= (1 - P_{ss})\gamma_e E_i - \gamma_s {I^a}_i \\
\frac{d{I^{aa}}_i}{dt} &= P_{ai}\gamma_s {I^a}_i - \gamma_{rai} {I^{aa}}_i \\
\frac{d{I^s}_i}{dt} &= (1 - P_{ai})\gamma_s {I^a}_i - \gamma_q {I^s}_i \\
\frac{d{I^{as}}_i}{dt} &= P_{ss}\gamma_e E_i - \gamma_s {I^{as}}_i \\
\frac{d{I^{aas}}_i}{dt} &= P_{ai}\gamma_s {I^{as}}_i - \gamma_{rai} {I^{aas}}_i \\
\frac{d{I^{ss}}_i}{dt} &= (1 - P_{ai})\gamma_s {I^{as}}_i - \gamma_q {I^{ss}}_i \\
\frac{d{I^q}_i}{dt} &= \gamma_q {I^{ss}}_i + \gamma_q {I^s}_i - \gamma_{ri} {I^q}_i \\
\frac{dR_i}{dt} &= (1 - P_{di})\gamma_{ri} {I^q}_i \\
\frac{d{R^a}_i}{dt} &= \gamma_{rai} {I^{aas}}_i + \gamma_{rai} {I^{aa}}_i \\
\frac{dD_i}{dt} &= P_{di}\gamma_{ri} {I^q}_i,
\end{aligned}
\tag{C.6}
$$

where i is the age group of the individual, C_{ij} is the $(i,j)^{\text{th}}$ element of the regional contact matrix defined as in the PHE model (see section C1.1.2) – which is equivalent to the expected number of new infections in age group i caused by an infection in age group j, and N is the total population size. The transmission rate of the virus can also vary based on the type of infector [Anderson et al., 2020; Wu et al., 2020; Lloyd-Smith et al., 2005] – whether they present as an asymptomatic, or symptomatic infection, or if they act as a super-spreader, but

are assumed equal across all age groups.

Parameter Name	Infective Compartment Label	Type of Infective
β_a	I^a	presymptomatic
β_{aa}	I^{aa}	asymptomatic
β_s	I^s	symptomatic
β_{as}	I^{as}	presymptomatic super-spreader
β_{aas}	I^{aas}	asymptomatic super-spreader
β_{ss}	I^{ss}	symptomatic super-spreader

Table C1: Transmission Rates of the Roche Model

In the Roche model each infectious compartment apart from the quarantined I^q, which we assume cannot transmit the virus, has its own transmission rate, as defined in Table C1. Furthermore, presymptomatic and asymptomatic transmission is assumed to be identical, i.e. individuals in either of the I^a or I^{aa} (and similarly for I^{as} and I^{aas}) pass on the virus to a susceptible with the same rate, which is taken to be half that of the symptomatic transmission. Moreover, as expected, super-spreaders transmit the virus more than the usual infectious individual by a constant factor for all groups. Below, we express mathematically these model assumptions:

$$\begin{aligned}
\beta_a &= \beta_{aa} = \frac{\beta_s}{2} \\
\beta_{as} &= \beta_{aas} = \frac{\beta_{ss}}{2} \\
\beta_s &= \beta_{\text{max}} - (\beta_{\text{max}} - \beta_{\text{min}}) \frac{\Theta^\gamma}{\Theta^\gamma + \Theta_{50}^\gamma} \\
\beta_{as} &= (1 + b_{ss})\beta_a \\
\beta_{aas} &= (1 + b_{ss})\beta_{aa} \\
\beta_{ss} &= (1 + b_{ss})\beta_s,
\end{aligned} \tag{C.7}$$

where b_{ss}, β_{max} and β_{min} represent the relative increase in transmission of a super-spreader case, the maximum possible transmission rate (i.e. when no non-pharmaceutical interventions are in place), and the minimum possible transmission rate, respectively (when all non-pharmaceutical interventions considered by the Roche model are put in place in their strictest form). The time-dependent stringency index Θ models the effect of non-pharmaceutical interventions currently in place on the virus transmission, with Θ_{50} representing the value of

the stringency index for which the transmission is reduced by 50%. The parameter γ is used for tuning the adaptability to changes in the interventions profile of the intervention-specific term in the expression for β_s; this is used for function continuity purposes. Larger values of this parameter cause the curve to more closely approach the step function.

The P_a, P_{ss} and P_d parameters are defined as the proportions of people that go on to become asymptomatic, super-spreaders or dead, respectively. Because we expect older people to be more likely to die [Verity et al., 2020] and younger people to be more likely to be asymptomatic [Nishiura et al., 2020; Mizumoto et al., 2020; Verity et al., 2020; Koelle et al., 2022], we consider P_a and P_d to be age-dependent in the context of the Roche model.

The final set of parameters to be defined in eq. (C.6) are the rates of progression through the different stages of the illness:

- γ_e: exposed to presymptomatic infectious status;
- γ_s: presymptomatic to (a)symptomatic infectious status;
- γ_q: symptomatic to quarantined infectious status;
- γ_r: quarantined infectious to recovered (or dead) status;
- γ_{ra}: asymptomatic to recovered (or dead) status.

Once infected, older and younger people may not recover from the virus in the same way; therefore, we assume age-dependency for both the γ_r and γ_{ra} parameters. These progression rates are computed according to the following rules:

$$\gamma_e = \frac{1}{k}, \gamma_s = \frac{1}{k_s}, \gamma_q = \frac{1}{k_q}, \gamma_{ri} = \frac{1}{k_{ri}}, \gamma_{rai} = \frac{1}{k_{rii}},$$

where k refers to mean incubation period until disease onset (i.e. from exposure to presymptomatic infection), k_s is the average time to developing symptoms since infectiousness onset, k_q is the average time until the case is quarantined once the symptoms appear, k_r is the average time until recovery since the start of the quarantining period, and k_{ri} is the average time to recovery since the end of the presymptomatic stage for an asymptomatic case. Moreover, in the Roche model the k_r, k_q and k_{ri} parameters are related through:

$$k_{rii} = k_{q_i} + k_{ri}, \forall i.$$

C2.1.1 Stringency Index

We mentioned previously that a particularity of the Roche model is the use of time-dependent transmission rates, which are sensitive to changes in the severity of the non-pharmaceutical interventions imposed by the government. The stringency index Θ, which is used to fine-tune the value of the symptomatic transmission parameter between its minimum and maximum possible values β_{min} and β_{max}, has been developed by the OxCGRT group [Hale et al., 2021] and acts as a measure of the severity of the NPIs. For ease of reference, in the following we present the methodology behind the computation of the real-time stringency index, as it was first introduced in the OxCGRT group study [Hale et al., 2021].

According to [Hale et al., 2021], we can incorporate in our computation of the stringency index Θ up to 20 distinct non-pharmaceutical interventions. Each intervention is characterised by an intervention indicator I_j which takes values between 0 and 100. The I_j indicators depend on a number of different factors:

- **Associated Level Scale with Maximum Level Value** N_j: e.g. for 'School closing' we consider a scale $0-3$, with the following level meanings:
  - *0*: in-person teaching occurs as normal;
  - *1* and *2*: partial closure with hybrid teaching;
  - *3*: no in-person teaching;
- **Targeting Binary Flag** F_j: indicator of whether the j^{th} intervention is tracked ('True') or not ('False') by the model throughout the epidemic;
- **Time-dependent General Value** G_j: this value indicates whether the j^{th} intervention is applied specifically (local intervention) of generally (at a national level) on any given day. We indicate through:
  - $G_j = 0$: the intervention is local;
  - $G_j = 1$: the intervention is general.

  There are two regimes for the possible values of G_j:
  - *just 0*: if the corresponding targeting flag $F_j =$ 'False';
  - *0 or 1*: if the corresponding targeting flag $F_j =$ 'True';

- **Current Intervention Level** C_j: this value must be ≥ 0 and $\leq N_j$.

In Table C2 we present the nine non-pharmaceutical interventions for the Roche model, together with their maximum levels N_j, targeted and time-dependent general values. Figure C6 shows a snapshot of a typical NPIs package structure at a given time point for the Roche model. The 'NA' in the $C8$ entry indicates a 'False' targeting flag F_8, while the 'No data' $C1$ entry means that for both current and general value we have $C_1 = G_1 = 0$.

Intervention	Maximum Level N_j	Targeted Flag F_j	General Value G_j
School closing	3 (0, 1, 2, 3)	Yes	0 or 1
Workplace closing	3 (0, 1, 2, 3)	Yes	0 or 1
Cancel public events	2 (0, 1, 2)	Yes	0 or 1
Restrictions on gatherings	4 (0, 1, 2, 3, 4)	Yes	0 or 1
Close public transport	2 (0, 1, 2)	Yes	0 or 1
Stay at home requirements	3 (0, 1, 2, 3)	Yes	0 or 1
Restrictions on internal movement	2 (0, 1, 2)	Yes	0 or 1
International travel controls	4 (0, 1, 2, 3, 4)	No	0
Public information campaigns	2 (0, 1, 2)	Yes	0 or 1

Table C2: Non-pharmaceutical Interventions included in the OxCGRT dataset

Indicator	Value	General?	Max value	Sub Index
Variable:	Cj	Gj	Nj	Ij
C1	No data	No data	3	0
C2	1	1	3	54
C3	2	0	2	69
C4	2	0	4	34.5
C5	1	1	2	65.5
C6	2	1	3	77
C7	2	1	2	100
C8	2	NA	4	50
H1	2	1	2	100
			Overall	61.11

C1:School closing
C2: Workplace closing
C3: Cancel public events
C4: Restrictions on gatherings
C5: Close public transport
C6: Stay at home requirements
C7: Restrictions on internal movement
C8: International travel controls
H1: Public information campaigns

Figure C6: *Example OxCGRT NPIs profile.*

In the following we present how we compute in practice the time-dependent intervention indicators I_j, and how these are then used to determine the value of the Θ stringency index of a set of non-pharmaceutical interventions – which is used to determine the new symptomatic transmission rate β_s in the Roche model, as defined in eq. (C.7). We use the nine NPIs currently tracked by the Roche model. Each intervention that has a targeting flag $F_j =$ 'True'

(therefore excluding 'International travel controls') has a different number of ordinal points on their severity scale (given by its maximum level N_j). In order to allow changes in interventional policies with large N_js and those with fewer ordinal points to contribute similarly in the computation of I_j, we will proceed to weight our interventions. This additional weight w is calculated with respect to the eight targeted interventions; N_j:

$$w = \frac{1}{8}\sum_{\substack{j=1 \\ j\neq 8}}^{9} \frac{1}{N_j + 1} \approx 0.29.$$

The time-dependent general value G_j with the current level C_j are weighted against each other in the computation of I_j as follows:

$$I_j = 100\left(C_j\frac{1-w}{N_j} + wG_j\right), \forall j = 1\ldots 9, j \neq 8.$$

For the remaining C_8 intervention which targets international travelling, we do not have a general value G_j (targeting flag is F_j = 'False') and therefore there is no need to use the weighting factor w – meaning the intervention index is given by the following formula:

$$I_8 = 100\frac{C_8}{N_8}.$$

We can tackle both F_j values by updating our formula for I_j into:

$$I_j = 100\left(C_j\frac{1-wF_j}{N_j} + wF_jG_j\right), \forall j = 1\ldots 9,$$

where F_j = 'False' = 0 and F_j = 'True' = 1. The stringency index is equal to the average of all the NPI indices. However, since 'Cancelling public events' is a subset of the 'Restricting gatherings' measure (similarly to how the 'Staying at home' requirement is included in 'Restrictions on internal movement'), we do not over-count those two pairs of measures. Instead we consider the contribution of the most effective one i.e. with higher I_j. The formula for Θ becomes:

$$\Theta = \frac{1}{7}(I_1 + I_2 + max(I_3, I_4) + I_5 + max(I_6, I_7) + I_8 + I_9).$$

In Figure C7, we plot in red the stringency index trajectory assuming all nine NPIs are

applied for the UK between 15th of February 2020 and 25th of June 2020. We profiled this trajectory using data we extracted and formatted in the *data* folder of our package. The raw data, which consist of:

- the current severity levels of the interventions covered by the [Hale et al., 2021] survey and
- the times of change of the severity levels of these interventions

, are available in the OxCGRT group repository [Hale et al., 2021]. An increase in the value of Θ is indicative of an increase in the severity of the restrictive measures meant to reduce the transmission of the virus in the community.

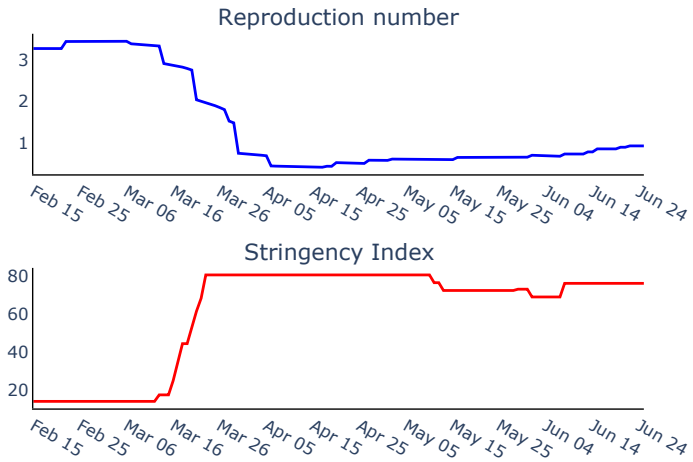


Figure C7: *Comparison plot of the Stringency Index Θ (red lines) and the instantaneous reproduction number R_t computed using PHE model implementation (blue lines).*

The effects of the reduction in transmission can be quantified through the instantaneous reproduction number R_t. In blue, we plot the average trajectory of R_t in England, calculated using our PHE model implementation. Comparing the Θ and R_t profiles we observe that as the stringency index increases, the reproduction number slowly decays, which proves the modelling validity of the former. Noteworthy is that the rate of decrease in R_t is slightly delayed compared to that of increase of Θ as the non-pharmaceutical interventions require some time before their effects become visible.

In the following section we present how we implemented the Roche model and integrated all the features described previously.

C2.1.2 Roche Model Class

The implementation of the Roche model (*RocheSEIRModel* class in the *roche_model.py* file) is done similarly to that of the PHE model, following the PINTS [Clerx et al., 2019] coding standards and structure. Unlike the PHE model – for which we included the numerical Euler ODE solver method because it appeared in the original paper [Birrell et al., 2021], we only include the *SciPy* [Virtanen et al., 2020] ODE solver method (*scipy_solver*) for the Roche model. The default setting remains the explicit Runge-Kutta method of order 5(4) [Dormand and Prince, 1980].

The *simulate* method in the Roche model class mimics the PHE model's implementation by taking as arguments the parameters of the model organised by the *RocheParameterController*. The *RocheParameterController* parameters are classified into:

- **the initial conditions**: the initial values of the compartments;
- **the compartment times parameters**: the average times individuals spend in each state between exposure and their subsequent recovery/death – k, k_s, k_q, k_r and k_{ri};
- **the proportion parameters**: the age-structured fractions of individuals that go on to become asymptomatic (P_a), and the age-invariant fraction of infections that go on to become super-spreaders (P_{ss}), or die respectively (P_d);
- **the transmission parameters**: the minimum and maximum transmission rates β_{min} and β_{max}, the relative increase in transmission of a super-spreader case b_{ss}, the parameter governing the flexibility of the intervention function γ, and the stringency index needed to reach 50% of the maximum effect on the infection rate Θ_{50};
- **the simulation-specific parameters**: the region for which we choose to simulate, the method of the solver used (the *SciPy* ODE solver), and the times for which we wish to run the model.

The time-dependent regional and contact matrices together with the time points of their changes are implemented beforehand. We use dedicated methods to add the age groups and the regions targeted to the model structure. The non-pharmaceutical intervention data, which encompass the time-dependent general and current level values of intervention severity as described in the previous section, are also implemented into the Roche model structure before running the simulation through the *read_npis_data* method.

The other methods implemented into the PHE model which include computation of daily new infections from the simulation output, calculating the log-likelihoods associated with the death and infection survey data, and sampling trajectories of the daily number of deaths and positive tests are also included for the Roche model. However, the way in which these quantities are computed differs slightly between the two models.

If the number of new infections $\Delta^{\text{infec}}_{r,t_k,i}$ is dependent in the context of the PHE model on the remaining number of susceptibles $S_{r,t_k,i}$ and the probability of infection $p^{\lambda}_{r,t_k,i} = \lambda_{r,t_{k-1},i}\delta_t$, i.e. the right hand-side term of the susceptible equation in the PHE model, for the Roche model we replace the $\lambda_{r,t_{k-1},i}\delta_t$ term with the corresponding term in its system of ODEs. The quantity can thus be expressed as follows:

$$\Delta^{\text{infec}}_{r,t_k,i} = S_{r,t_k,i} \sum_j C_{ij} \Big(\frac{\beta_a}{N} {I^a}_{r,t_k,j} + \frac{\beta_{aa}}{N} {I^{aa}}_{r,t_k,j} + \frac{\beta_s}{N} {I^s}_{r,t_k,j} + \frac{\beta_{as}}{N} {I^{as}}_{r,t_k,j} + \frac{\beta_{aas}}{N} {I^{aas}}_{r,t_k,j} + \\ + \frac{\beta_{ss}}{N} {I^{ss}}_{r,t_k,j} \Big). \quad \text{(C.8)}$$

For the Roche model we also compute the daily number of new deaths, which is calculated as the difference between the death compartment counts, D, on consecutive days. The k^{th} row entry of this output is then taken to be the mean number of deaths μ_{r,t_k}. We assume the age-structured number of deaths follows a negative-binomial distribution just as in the case of the PHE model, with the mean $\mathbb{E}(X_{r,t_k,i}) = \mu_{r,t_k,i}$ and variance, $\mathbb{V}(X_{r,t_k,i}) = (\eta+1)\mu_{r,t_k,i}$ where η is the dispersion factor of the negative binomial distribution – we use $\eta = 5$, similarly to what we did for the PHE model:

$$X_{r,t_k,i} \sim \text{NegativeBinomial}(\mu_{r,t_k,i}, \eta). \quad \text{(C.9)}$$

The Roche model's methods for sampling and computing the log-likelihood function associated with the number of identified positive test results in a given region on a specific day is done identically to our reimplementation of the PHE model. The number of positive tests is distributed according to a binomial distribution with parameters depending on the age-structured number of tests conducted per day (we use REACT1 data [Nicholson et al., 2021]), the simulation output of the model, and a COVID-19 infection survey antigen test

sensitivity (true-positive rate) and specificity (true-negative rate):

$$Y_{r,t_k,i} \sim \text{Bin}\left(n_{r,t_k,i}, k_{\text{sens}}\left(1 - \frac{I_{r,t_k,i}}{N_{r,i}}\right) + (1 - k_{\text{spec}})\frac{I_{r,t_k,i}}{N_{r,i}}\right) \tag{C.10}$$
$$I_{r,t_k,i} = I^{a}_{r,t_k,i} + I^{aa}_{r,t_k,i} + I^{s}_{r,t_k,i} + I^{as}_{r,t_k,i} + I^{aas}_{r,t_k,i} + I^{ss}_{r,t_k,i} + I^{q}_{r,t_k,i},$$

where $n_{r,t_k,i}$ is the number of tests collected from individuals of age i, in region r on day t_k, k_{sens} and k_{spec} are the test sensitivity and specificity respectively with assumed default values of 70% and 95% [Watson et al., 2020].

The final method included in the Roche model class is a method for computing the instantaneous reproduction number at any given time t using the next generation matrix approach [Diekmann et al., 2010]. This method uses as input the simulation output of a run of the Roche model in order to compute both the transition (Σ) and transmission matrices (T) of the process at the indicated time point and returns the reproduction number as the spectral radius of the matrix $-T\Sigma^{-1}$, where the spectral radius of a matrix refers to the maximum eigenvalue of said matrix. We provide detailed forms of the transition and transmission matrix in the following section.

C3 Warwick-Household Model

The third and last model included in the 'epimodels' Python library is the Warwick-Household model. This model has been developed by the University of Warwick and has been used as part of the SPI-M consortium to support their advice to the UK government on the appropriate measures to mitigate the growth of the COVID-19 epidemic. In the following, we use the original paper and its more technical sequel [Keeling et al., 2021, 2022] to faithfully reconstruct the Warwick-Household model in the same software framework as the PHE and the Roche models.

C3.1 Warwick-Household Model Structure

Like both of the two models previously discussed, the Warwick-Household model extends the basic SEIR model to more accurately reflect the transmission dynamics of the SARS-CoV-2 virus. We assume the modelled population to be both age-structured and split into

the same seven PHE regions discussed in section C1.1.1. Similar to the PHE model, the Warwick-household model considers multiple connected exposed compartments (three compared to the PHE model's two) so that the overall waiting time from exposure to becoming infectious of new cases is Gamma-distributed. As previously discussed in C1.1 and [van Leeuwen et al., 2017; Anderson et al., 2020], this assumption ensures that the observed latency period of any particular occurring case is less likely to be significantly smaller than the average expected value – as would be the case with an exponential waiting time between the exposed and infectious states, as it would be dictated by using only one E compartment.

Additionally, the Warwick model differentiates between symptomatic (or detectable I) and asymptomatic (or undetectable A) infections which is similar to the Roche model. However, it foregoes the presymptomatic infectious stage. Super-spreader events are also not accounted for in the context of this model. However, the main feature setting aside the Warwick-Household model from both the PHE and Roche models is that it takes into account within-household dynamics. To this effect, all the exposed E^1, E^2 and E^3, symptomatic (referred to as detectable infection in the original [Keeling et al., 2021] paper) infectious I and asymptomatic (or undetectable) infectious A compartments are split further based on the order of their becoming exposed in the household:

- Those individuals who contracted the infection first, that is outside of the household, are flagged with the F superscript. These compartments include: $E^{1,F}$, $E^{2,F}$, $E^{3,F}$, I^F, A^F.

- Those who are infected by someone in the household, also known as subsequent infections are flagged with the S superscript. We can differentiate on the infection type of the infector, whether they are an asymptomatic – undetectable (these include the $E^{1,SU}$, $E^{2,SU}$, $E^{3,SU}$ and I^{SU} compartments) or a symptomatic – detectable infected individual (for the $E^{1,SD}$, $E^{2,SD}$, $E^{3,SD}$ and I^{SD} compartments) or not as is the case with subsequent asymptomatic infections represented by the A^S compartment.

- And lastly those individuals who are quarantined are flagged with the Q superscript. These include the individuals in the $E^{1,Q}$, $E^{2,Q}$, $E^{3,Q}$ and A^Q. For the symptomatic infections occurring among the quarantined individuals we differentiate based on the provenance of these infections within the household: whether they were first infected in the household (I^{QF}) or they were subsequently infected by another household mem-

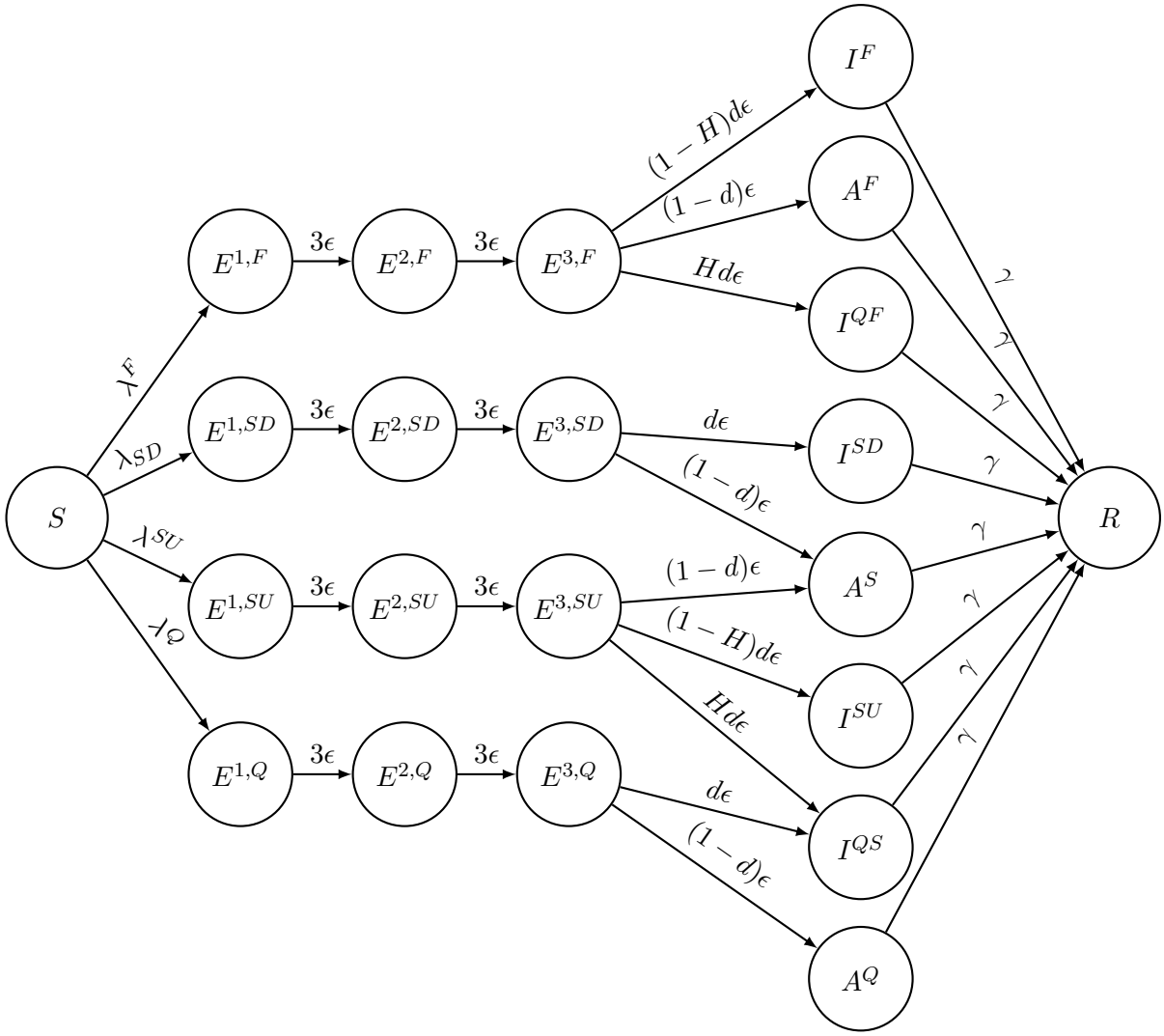


Figure C8: Warwick-Household Model Structure

ber (I^{QS}).

We summarise the transmission mechanisms across the different compartments in Figure C8. Just as in the case of both the PHE and Roche models, no ageing, migration between regions, natural deaths or births are considered. For each region, the ODE system underlying the Warwick-Household model is given by:

$$\begin{aligned}\frac{dS_i}{dt} =& -\sigma_i \sum_j \Big(C^N_{ij} \big(\frac{S_i}{N} I^F_j + \frac{S_i}{N} I^{SD}_j + \frac{S_i}{N} I^{SU}_j + \tau(\frac{S_i}{N} A^F_j + \frac{S_i}{N} A^S_j)\big) + \\ & + C^H_{ij} \big(\frac{S_i}{N} I^F_j + \frac{S_i}{N} A^F_j + \frac{S_i}{N} I^{QF}_j\big)\Big) \\ \frac{dE^{1,F}_i}{dt} =& \sigma_i \sum_j C^N_{ij} \big(\frac{S_i}{N} I^F_j + \frac{S_i}{N} I^{SD}_j + \frac{S_i}{N} I^{SU}_j + \tau(\frac{S_i}{N} A^F_j + \frac{S_i}{N} A^S_j)\big) - 3\epsilon E^{1,F}_i\end{aligned}$$

$$
\begin{aligned}
\frac{dE_i^{1,SD}}{dt} &= \sigma_i \sum_j C_{ij}^H \frac{S_i}{N} I_j^F - 3\epsilon E_i^{1,SD} \\
\frac{dE_i^{1,SU}}{dt} &= \sigma_i \sum_j C_{ij}^H \frac{S_i}{N} A_j^F - 3\epsilon E_i^{1,SU} \\
\frac{dE_i^{1,Q}}{dt} &= \sigma_i \sum_j C_{ij}^H \frac{S_i}{N} I_j^{QF} - 3\epsilon E_i^{1,Q} \\
\frac{dE_i^{2,F}}{dt} &= 3\epsilon E_i^{1,F} - 3\epsilon E_i^{2,F} \\
\frac{dE_i^{2,SD}}{dt} &= 3\epsilon E_i^{1,SD} - 3\epsilon E_i^{2,SD} \\
\frac{dE_i^{2,SU}}{dt} &= 3\epsilon E_i^{1,SU} - 3\epsilon E_i^{2,SU} \\
\frac{dE_i^{2,Q}}{dt} &= 3\epsilon E_i^{1,Q} - 3\epsilon E_i^{2,Q} \\
\frac{dE_i^{3,F}}{dt} &= 3\epsilon E_i^{2,F} - 3\epsilon E_i^{3,F} \\
\frac{dE_i^{3,SD}}{dt} &= 3\epsilon E_i^{2,SD} - 3\epsilon E_i^{3,SD} \\
\frac{dE_i^{3,SU}}{dt} &= 3\epsilon E_i^{2,SU} - 3\epsilon E_i^{3,SU} \\
\frac{dE_i^{3,Q}}{dt} &= 3\epsilon E_i^{2,Q} - 3\epsilon E_i^{3,Q} \\
\frac{dI_i^F}{dt} &= 3(1-H)\epsilon d_i E_i^{3,F} - \gamma I_i^F \\
\frac{dI_i^{SD}}{dt} &= 3\epsilon d_i E_i^{3,SD} - \gamma I_i^{SD} \\
\frac{dI_i^{SU}}{dt} &= 3(1-H)\epsilon d_i E_i^{3,SU} - \gamma I_i^{SU} \\
\frac{dI_i^{QF}}{dt} &= 3H\epsilon d_i E_i^{3,F} - \gamma I_i^{QF} \\
\frac{dI_i^{QS}}{dt} &= 3H\epsilon d_i E_i^{3,SU} + 3\epsilon d_i E_i^{3,Q} - \gamma I_i^{QS} \\
\frac{dA_i^F}{dt} &= 3\epsilon(1-d_i) E_i^{3,F} - \gamma A_i^F \\
\frac{dA^S}{dt} &= 3\epsilon(1-d_i)(E_i^{3,SD} + E_i^{3,SU}) - \gamma A_i^S \\
\frac{dA_i^Q}{dt} &= 3\epsilon(1-d_i) E_i^{3,Q} - \gamma A_i^Q \\
\frac{dR_i}{dt} &= \gamma(I_i^F + I_i^{QF} + A_i^F + I_i^{SD} + A_i^S + I_i^{SU} + I_i^{QS} + A_i^Q)
\end{aligned}
\tag{C.11}
$$

where i is the age group of the individual, C^H_{ij} is the $(i,j)^{\text{th}}$ element of the regional household contact matrix, and represents the expected number of contacts in age group i within the household made by an individual in age group j on a given day. Similarly, C^N_{ij} is the $(i,j)^{\text{th}}$ the element of the regional non-household contact matrix and represents the expected number of contacts in age group i outside the household made by an individual in age group j on a given day. The Warwick-Household model differentiates between the contacts occurring in the household (summarised by the household contact matrix C^H), at school (C^S), in the workplace (C^W) and in all other contexts (C^O) – the latter three of which sum to the non-household contact matrix $C^N = C^S + C^W + C^O$. The total population size is given by the parameter N.

C3.1.1 Social distancing modelling

Non-pharmaceutical interventions affect the number of contacts individuals will have per day. Moreover, these interventions can target different types of human interactions at different rates as seen from the OxCGRT study [Hale et al., 2021]. Therefore, to compute the correct household and non-household contact matrices C^H, C^S, C^W and C^O respectively on each day, we calculate the weighted average between the baseline matrices for contacts within household ($C^{H,base}$), at school ($C^{S,base}$), in the workplace ($C^{W,base}$) and in all other contexts ($C^{O,base}$) and their full-lockdown equivalents which are given by:

$$\begin{aligned} C^{H,lock} &= q_H C^{H,base} \\ C^{S,lock} &= q_S C^{S,base} \\ C^{W,lock} &= q_W C^{W,base} \\ C^{O,lock} &= q_O C^{O,base}, \end{aligned}$$

where q_H is the coefficient of increase in contacts within the household and q_S, q_W and q_O are the decrease in contacts at school, in the workplace, and, respectively, in all other contexts when a full-lockdown is implemented. In the original paper [Keeling et al., 2021] the authors multiply the within household contact matrix C^H by 1.3 in order to correctly match the model to a model without the household structure.

We further define ϕ as the coefficient indicating the strength of the implemented interventions and θ as the scaling factor of any public-facing interactions. The θ parameter is used

to separate interaction between the public-facing and the non-public facing industries cases; in the former the effect of interventions is two-fold as NPIs affect both the worker as well as the general population with whom they come into contact.

Hence, the daily household and non-household contact matrices C^H and C^N are given by:

$$
\begin{aligned}
C^H &= 1.3\big((1-\phi)C^{H,base} + \phi C^{H,lock}\big) \\
C^S &= \big((1-\phi)C^{S,base} + \phi C^{S,lock}\big) \\
C^W &= \big(1-\theta+\theta(1-\phi+\phi q_O)\big)\big((1-\phi)C^{W,base} + \phi C^{W,lock}\big) \\
C^O &= (1-\phi+\phi q_O)\big((1-\phi)C^{O,base} + \phi C^{O,lock}\big),
\end{aligned}
$$

where $C^N = C^S + C^W + C^O$. According to [Keeling et al., 2021, 2022], the coefficients of relative change of the full-lockdown contact matrices and the scaling factor of public-facing interactions are fixed at the values prescribed in Table C3.

Definition	Parameter	Value
Increase in household contacts	q_H	1.25
Reduction in school contacts	q_S	0.05
Reduction in workplace contacts	q_W	0.2
Reduction in all-other contexts contacts	q_O	0.05
Scaling factor for all public-facing interactions	θ	0.3

Table C3: Fixed Intervention values for the Warwick-Household model

The transmission parameters are the rates with which different types of infectious individual infect susceptible ones. Asymptomatic infections are assumed to have a reduced transmission compared to their symptomatic counterpart [Anderson et al., 2020; Wu et al., 2020; Lloyd-Smith et al., 2005]. This reduction in transmission is indicated through the parameter τ in the force of infection. The vector of susceptibilities to the disease σ which goes into the computation of the age-dependent force of infection for types of infections in the household:

$$
\begin{aligned}
\lambda_i^F &= \sigma_i \sum_j C_{ij}^N \big(I_j^F + I_j^{SD} + I_j^{SU} + \tau(A_j^F + A_j^S)\big) \\
\lambda_i^{SD} &= \sigma_i \sum_j C_{ij}^H I_j^F \\
\lambda_i^{SU} &= \sigma_i \sum_j C_{ij}^H A_j^F
\end{aligned}
$$

$$\lambda_i^Q = \sigma_i \sum_j C_{ij}^H I_j^{QF}$$

is age-dependent, unlike the Roche model. This is because we expect older people to be more likely to be susceptible to infection. Similarly, younger people are more likely to be asymptomatic; hence we also take the vector d of proportions of people that go on to develop symptomatic infections to be age-dependent. Therefore, the Warwick-Household model offers more flexibility and accounts for differences in transmission and symptom development risks across different demographics.

The rates of progression through the different stages of the illness are:

- ϵ: exposed to (a)symptomatic infectious status;
- γ: (a)symptomatic to recovered status.

The last parameter to account for, H, is defined as the household quarantine compliance and acts as a measure of the social adherence to the interventions of the population. According to the additional technical material of the model [Keeling et al., 2022], this parameter is related to the strength of interventions parameter ϕ via the following relationship: $H = h\phi$, where $h = 0.8$ by default, but we allow users to change this value – along with allowing temporal variations in the values of the other intervention parameters q_H, q_S, q_W, q_O and ϕ. There are two main reasons for this choice: (1) in Chapter 4, using the default parameter values lead to poor parameter fitting to available death data, and (2) in the private repository code of this project, these parameters are varied over the course of the considered time period. These discrepancies give rise to essentially two different reported versions of the Warwick-Household model; this makes our task of model reimplementation significantly more difficult, because we cannot determine which features belong to the true expression of the model with any certainty. Moreover, this ambiguity regarding the true dynamics of the model also impacts its value as a reliable tool for modelling a future epidemic with different initial epidemic conditions and applied interventions.

C3.1.2 Warwick-Household Model Class

The Warwick-Household model follows a similar integration approach into the 'epimodels' Python package as the PHE and Roche models. The forward-model is included in a separate class called *WarwickSEIRModel* (this can be found in *warwick_model.py* file) and follows a

similar structure to the two previously discussed models; we use the same coding standards as those implemented in the PINTS [Clerx et al., 2019] package, which we hold to as a standard of good software engineering practices. We use a *SciPy* [Virtanen et al., 2020] ODE solver with and explicit Runge-Kutta method of order 5(4) [Dormand and Prince, 1980] to run forward-simulations of the *WarwickSEIRModel* class (as observed in the *_scipy_solver* class method).

Moreover, just as we have proceeded for both the Roche and PHE models, the *simulate* method of the Warwick-Household model takes the parameters of the model in the form of a *WarwickParameterController* type object. The *WarwickParameterController* parameters are classified into:

- **the initial conditions**: the initial values of the compartments;
- **the regional parameters**: the region for which we choose to simulate and the regional quarantine proportions multiplier h;
- **the disease-specific parameters**: the reduction in transmission for an asymptomatic infection compared to the symptomatic case τ and the age-dependent probabilities of displaying symptoms d;
- **the transmission parameters**: the rate of progression to infection from exposed ϵ the recovery rate γ and the age-dependent susceptibility to infection σ;
- **the simulation-specific parameters**: the method of the solver used (the *SciPy* ODE solver) and the times for which we wish to run the model;
- **the social distancing parameters**: the proportion of work interactions in public-facing 'industries' θ, the time-dependent trajectories of the scaling factor between pre- and full-lockdown contact matrices ϕ, the factor of increase in the amount of household interactions during lockdown q_H, and the factors of reduction in attendance at school, in the workplace and engagement with shopping and leisure activities during lockdown – q_S, q_W and q_O respectively, and the times when these changes in these social distancing parameters are implemented.

The time-dependent regional and contact matrices together with the time points of their changes are implemented beforehand.

The other methods implemented into the Warwick-Household model which include computation of daily new infections from the simulation output, calculating the log-likelihoods associated with the death and infection survey data, and sampling the number of deaths and positive tests are also included as for the Roche model. However, the way in which these quantities are computed differs slightly between the two models.

To determine the number of new symptomatic infections in the Warwick-Household model we sum over all rates of entry into the symptomatic infected compartment coming from the exposed compartment, which gives us:

$$\Delta^{\text{infec}}_{r,t_k,i} = 3\epsilon d_i (E^{3,F}_{r,t_k,i} + E^{3,SD}_{r,t_k,i} + E^{3,SU}_{r,t_k,i} + E^{3,Q}_{r,t_k,i}). \tag{C.12}$$

The incidence of infections can then be used to compute the number of new daily hospitalisations ($H_{r,t,i}$) and ICU admissions ($ICU_{r,t,i}$). For both outputs of interest, their daily number of new occurrences will depend on a previous interval of new infections according to:

$$\begin{aligned} H_{r,t,i} &= P^{I\to H}_{r,i} \sum_q \text{Delay}^{I\to H}_{r,q} \Delta^{\text{infec}}_{r,t-q,i} \\ ICU_{r,t,i} &= P^{I\to ICU}_{r,i} \sum_q \text{Delay}^{I\to ICU}_{r,q} \Delta^{\text{infec}}_{r,t-q,i}, \end{aligned} \tag{C.13}$$

where $\text{Delay}^{I\to H}_{r,q}$ and $\text{Delay}^{I\to ICU}_{r,q}$ represent the regional probability that an individual who becomes hospitalised and respectively admitted to the ICU does so q days after entering a symptomatic infectious class. Additionally, we denote by $P^{I\to H}_{r,i}$ and $P^{I\to ICU}_{r,i}$ the regional probability that a symptomatic infectious individual aged i goes on to be hospitalised or respectively admitted to the ICU. The Warwick-Household model also includes methods for computing the daily incidence of deaths ($D_{r,t,i}$) and the hospital ($H^{\text{bed}}_{r,t,i}$) and ICU beds occupancy ($ICU^{\text{bed}}_{r,t,i}$) each day using the number of hospitalisation and ICU admissions described in Eq. (C.13):

$$\begin{aligned} D_{r,t,i} &= P^{H\to D}_{r,i} \sum_q \text{Delay}^{H\to D}_{r,q} H_{r,t-q,i} \\ H^{\text{bed}}_{r,t,i} &= \sum_q T^{H}_{r,q} H_{r,t-q,i} + \sum_q T^{ICU\to H}_{r,q} ICU_{r,t-q,i} \end{aligned}$$

$$ICU^{\text{bed}}_{r,t,i} = \sum_q T^I_{r,q} ICU_{r,t-q,i}, \tag{C.14}$$

where $\text{Delay}^{H\to D}_{r,q}$ is the regional probability that an individual dies q days after becoming hospitalised, $P^{H\to D}_{r,i}$ the regional probability that a hospitalised individual aged i will die, $T^H_{r,q}$, and $T^{ICU\to H}_{r,q}$ and $T^I_{r,q}$ represent the regional probability that someone admitted to hospital is still occupying a non-ICU hospital bed or respectively if an ICU admission is still occupying a bed in non-ICU or ICU q days later.

We conclude the listing of the methods included in the Warwick-Household model's class with methods for sampling and computing the log-likelihood function associated with the number of identified positive test results and deaths in a given region on a specific day. These closely resemble the approaches we employed previously for both the Roche and PHE models, that is we assume a negative-binomial noise on the number of death. For the number of positive tests we consider a binomial distributed noise; the computation of the mean of the prescribed distribution relies on REACT1 study data [Nicholson et al., 2021] – which provides the age-structured number of tests conducted per day and the simulation output of the model and a COVID-19 infection survey antigen test sensitivity (true-positive rate) of 70% and specificity (true-negative rate) of 95% , as prescribed in [Watson et al., 2020]:

$$\begin{aligned} Y_{r,t_k,i} &\sim \text{Bin}\left(n_{r,t_k,i}, k_{\text{sens}}\left(1 - \frac{I_{r,t_k,i}}{N_{r,i}}\right) + (1 - k_{\text{spec}})\frac{I_{r,t_k,i}}{N_{r,i}}\right) \\ I_{r,t_k,i} &= I^a_{r,t_k,i} + I^{aa}_{r,t_k,i} + I^s_{r,t_k,i} + I^{as}_{r,t_k,i} + I^{aas}_{r,t_k,i} + I^{ss}_{r,t_k,i} + I^q_{r,t_k,i} \end{aligned} \tag{C.15}$$

where $n_{r,t_k,i}$ is the number of tests collected from individuals of age i, in region r on day t_k, k_{sens} and k_{spec} are the test sensitivity and specificity.

C4 Example model implementations

In this section we present example implementation discuss and compare some of the results produced by running the PHE, the Roche and the Warwick-Household models with UK epidemic data between 15$^{\text{th}}$ of February 2020 and 25$^{\text{th}}$ of June 2021. We wish to investigate the

effects of implementing policy decisions similar to those made in response to the COVID-19 pandemic on the daily number of infections and deaths and on the time-dependent reproduction number.

C4.1 PHE Model Implementation

For the PHE model, these results are contained in two separate notebooks in the 'epimodels' repository, each named for the type of contacts we assume: fixed – no interventions; variable – apply Google mobility data [Google] to model the effect of the various government interventions which restrict the number of contacts between individuals. Then, in a separate notebook in the *examples/phe_examples* folder of the 'epimodels' we include our results for a comparative scenario analysis performed using the PHE model: we compare (1) the scenario in which no interventions have been implemented, i.e. keeping the contact matrices to their baseline value, with (2) the case when interventions are accounted for, i.e. we apply the Google mobility data to the contact matrices.

To initialise the model, we use the information in the section about prior parameters from [Birrell et al., 2021]. In the original paper, both the mean latent period d_L and mean infectious period d_I are assumed to be equal to 4 (days) – in our instantiation of the model, we will also use $d_L = 4$ and $d_I = 4$. The sensitivity and specificity are taken as in [Watson et al., 2020] to be 70% and 95% respectively. The initial reproduction number $R_{0,r}$ is region specific and is dependent on the physiology of the disease. Hence, it makes sense that these quantities depend in some capacity on d_L and d_I, as seen in [Birrell et al., 2021]:

$$R_{0,r} = \psi_r d_I \frac{(\frac{\psi_r d_L}{2} + 1)^2}{1 - \frac{1}{(\frac{\psi_r d_I}{2}+1)^2}},$$

where the region-specific exponential growth parameters ψ_r are drawn from a gamma$(31.36, 224)$-distribution. This equation is directly referenced in [Wearing et al., 2005]. A longer infectious period implies more time for individuals to be infected by an already infected case; a longer latent period has the opposite effect since exposed individuals do not present symptoms and the model structure assumes that only symptomatic cases can transmit the virus. For our purposes we will use the value for ψ_r across all the seven PHE regions, equal to the mean of the gamma-distribution, i.e. $\psi_r = 0.14$. Therefore, the basic reproduction number is the same across all regions, $R_{0,r} = 2.355$; this is significantly lower than the value

for R_0 we inferred from UK death data [GOV.UK] back in section 4.3.7.1, which serves as a first indication that the parameter estimates reported in [Birrell et al., 2021] do not correctly fit the data. Additionally, in this instantiation of the PHE model, we forgo any regional variations in the contact matrix, that is, we fix the weekly β_t parameters defined in eq. (C.2) to be equal to one throughout the simulation window.

For the initial values in each compartment for each region, we assume no individuals in any of the E^1, E^2, I^2, or R compartments. For the first type of infected individuals, i.e. those in the I^1 age-separated compartments, we assume in each age-compartment 50 individuals; we do not use the realised counts of cases at that time (total of 8 in the whole of England, as indicated by ONS case data [GOV.UK]), as these lead to overall poor fitting of the model to the ONS death data, as we see in Chapter 4 – this could be explained by both the mass under-reporting of the cases, common in the beginning of an epidemic, as well as the change of the definition of a case later on in the epidemic. For the susceptibles S compartments, we make use of the census data for England from the Office of National Statistics [ONS], where we have the numbers of individuals of each age in each of the historical regions of England, collapsing those columns and rows to obtain the subtotals for each of the age groups and regions as mentioned before.

C4.2 Roche Model Implementation

For the Roche model, we include a Python notebook in the 'epimodels' repository, where we compare two possible scenarios: (1) when no interventions are modelled, keeping the contact matrix at its baseline levels versus (2) the case when all interventions applied for the first wave of the UK COVID-19 epidemic are accounted for, that is we apply the OxCGRT NPIs data.

We instantiate the Roche model using the same parameter values as those referenced in section 4.3.5.2 – with the exception of the β_{min} and b_{ss} parameters for which we use their mean estimates, as inferred in section 4.3.7.2. It is important to note that we do not use the parameter values introduced in the original paper [Lemenuel-Diot et al., 2020] – we see in section 4.3.7.2 that by using the values stated in the paper.

We also use the same values of 70% and 95% for the sensitivity and specificity parameters used in sampling the number of infections and deaths identified in [Watson et al., 2020].

For each PHE region and age group we assume initially no individuals in any of the E, I^a, I^{aa}, I^{as}, I^{ss}, I^{aas}, I^q, R, R^a or D compartments. The susceptibles are initialised at the values indicated in the age-structured regional 2011 census data for England collected by the Office of National Statistics [ONS]. All that remains is to instantiate the symptomatic compartments I^s. When we used the same initial number of infections as in the PHE model (see section C4.1), the resulting estimated trajectory of the number of deaths produced by the Roche model after parameter optimisation (see Chapter 4 for more details) did not fit well to the observed death data. Instead, we assume 40 individuals in each age group. Of note is that this number differs from the value of initial number of infections we use to parameterise the PHE model. This indeed serves as a first clue of the intrinsic differences between the PHE and the Roche models and why they are not substitutable for each another.

Additionally, in the original Roche model paper, no definition of the time-dependent reproduction number is provided. To compute its trajectory we use the next generation matrix approach as described in [Diekmann et al., 2010]. To do so, we first identify the infection system – that is all infection states that create or transition infections. For the Roche model, this system is given by the following vector of compartments $V = (E, I^a, I^{aa}, I^s, I^{as}, I^{aas}, I^{ss}, I^q)$:

$$
\begin{aligned}
\frac{dE_i}{dt} &= -\gamma_e E_i + \sum_j C_{ij}\Big(\frac{\beta_a}{N} S_i {I^a}_j + \frac{\beta_{aa}}{N} S_i {I^{aa}}_j + \frac{\beta_s}{N} S_i {I^s}_j + \frac{\beta_{as}}{N} S_i {I^{as}}_j \\
&\quad + \frac{\beta_{aas}}{N} S_i {I^{aas}}_j + \frac{\beta_{ss}}{N} S_i {I^{ss}}_j\Big) \\
\frac{d{I^a}_i}{dt} &= (1 - P_{ss})\gamma_e E_i - \gamma_s {I^a}_i \\
\frac{d{I^{aa}}_i}{dt} &= P_{ai}\gamma_s {I^a}_i - \gamma_{rai} {I^{aa}}_i \\
\frac{d{I^s}_i}{dt} &= (1 - P_{ai})\gamma_s {I^a}_i - \gamma_q {I^s}_i \\
\frac{d{I^{as}}_i}{dt} &= P_{ss}\gamma_e E_i - \gamma_s {I^{as}}_i \\
\frac{d{I^{aas}}_i}{dt} &= P_{ai}\gamma_s {I^{as}}_i - \gamma_{rai} {I^{aas}}_i \\
\frac{d{I^{ss}}_i}{dt} &= (1 - P_{ai})\gamma_s {I^{as}}_i - \gamma_q {I^{ss}}_i \\
\frac{d{I^q}_i}{dt} &= \gamma_q {I^{ss}}_i + \gamma_q {I^s}_i - \gamma_{ri} {I^q}_i,
\end{aligned}
$$

where i is the age group of the individual. We group compartments together based on age

group, so that the infectious system vector is more accurately given by

$$V = (E_1, \dots E_M, I_1^a, \dots I_M^a, \dots, I_1^q \dots, I_M^q),$$

where we indicate by $M = 8$ the total number of age groups modelled. The $(i,j)^{\text{th}}$ element of the transition matrix Σ comprises of the rate at which the individuals in the infectious states j move to the infectious state i. Similarly, for the transmission matrix, its $(i,j)^{\text{th}}$ element is the rate at which individuals in the infectious states j generate new infectious into the infectious state i. Therefore, we determine the following block structures for the transmission (T) and transition (Σ) matrices respectively:

$$T = \begin{pmatrix} 0 & S^T\frac{\beta_a}{N}C & S^T\frac{\beta_{aa}}{N}C & S^T\frac{\beta_s}{N}C & S^T\frac{\beta_{as}}{N}C & S^T\frac{\beta_{aas}}{N}C & S^T\frac{\beta_{ss}}{N}C & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ & & & \ddots & & & & \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

$$\Sigma = \begin{pmatrix} -\gamma_e\mathbb{I}_M & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ (1-P_{ss})\gamma_e\mathbb{I}_M & -\gamma_s\mathbb{I}_M & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \gamma_s(P_a)_M & -\gamma_{ra}\mathbb{I}_M & 0 & 0 & 0 & 0 & 0 \\ 0 & \gamma_s(1-P_a)_M & 0 & -\gamma_q\mathbb{I}_M & 0 & 0 & 0 & 0 \\ P_{ss}\gamma_e\mathbb{I}_M & 0 & 0 & 0 & -\gamma_s\mathbb{I}_M & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \gamma_s(P_a)_M & -\gamma_{ra}\mathbb{I}_M & 0 & 0 \\ 0 & 0 & 0 & 0 & \gamma_s(1-P_a)_M & 0 & -\gamma_q\mathbb{I}_M & 0 \\ 0 & 0 & 0 & \gamma_q\mathbb{I}_M & 0 & 0 & \gamma_q\mathbb{I}_M & -\gamma_r\mathbb{I}_M \end{pmatrix},$$

where 0 refers to the $M \times M$ matrix with all elements equal to zero and $\mathbb{I}_M$ is the $M \times M$ identity matrix. $(P_a)_M$ is a diagonal matrix with the i^{th} element on the diagonal equal to P_{ai} (and similarly $(1-P_a)_M$ is a diagonal matrix with the i^{th} element on the diagonal equal to $1 - P_{ai}$). Additionally S^T is the row vector or remaining susceptibles at the time of estimation of the reproduction number R_t and C is the contact matrix. The reproduction number is then equal to the maximal eigenvalue (also known as the *spectral radius*) of the matrix $-T\Sigma^{-1}$, that is

$$R_t = \rho(-T\Sigma^{-1}).$$

C4.3 Warwick Model Implementation

To study the impact of non-pharmaceutical interventions on the epidemic dynamics in the context of the Warwick-Household model, we proceed similarly to both the PHE and the Roche model: in a separate Python notebook in the 'epimodels' repository we model the daily number of deaths, cases and the trajectories of the time-variant reproduction number in two scenarios: (1) when no NPIs have been applied, that is the level of contacts are kept at their baseline values versus (2) the case when interventions have been applied as during the first wave of the UK COVID-19 epidemic. For both scenarios, we model the population in the seven different PHE regions of England between the 15$^{\text{th}}$ of February 2020 and 25$^{\text{th}}$ of June 2021.

To instantiate the Warwick-Household model we use the same parameter values as those inferred for the first wave of the UK COVID-19 epidemic (see Chapter 4 and the original and technical papers [Keeling et al., 2021, 2022]). To sample the number of infections and deaths using large-scale infection survey REACT1 data we assume 70% test sensitivity and 95% specificity, as seen in [Watson et al., 2020]. Initially, we consider individuals only in the susceptible S, and the symptomatic infectious that occur first within the household I^F: for those in the age-structured S compartments, we consider the number of individuals recorded in each age bracket as recorded in the 2011 census for England [ONS]; for the remaining I^F compartments we assume a total of 20 individuals, spread across the different age groups according to a previously prescribed vector of fractions of initial infections within the modelled population. In our case, we adapt the vector of initial infection fraction from eq. (4.6).

Similar to the Roche model, the original code describing Warwick-Household model has no precise methodology for the computation of the time-dependent reproduction number detailed. Instead, it only presents a way to compute the basic reproduction number R_0 using a next-generation matrix approach; in that sense, R_0 is equal to the dominant eigenvalue of the non-household contact matrix C^N weighted by the susceptibility vector σ, the symptom development probability vector d and the reduction in transmission effect of the asymptomatic infections τ:

$$R_0 = \rho\Big((C_{ij} d_j \sigma_j \frac{1+\tau(1-d_i)}{d_i})_{ij}\Big),$$

where $\rho(\cdot)$ indicates the spectral radius of a matrix and we indicate by $(\cdot)_{ij}$ the matrix whose $(i,j)^{\text{th}}$ element is written between the brackets.

In the following, to compute the time-dependent R_t for the Warwick-Household model, we proceed instead from first principles using the methodology described in [Diekmann et al., 2010]. The infectious system consists of the following infectious states: $V = (E^{1,F}, E^{2,F}, E^{3,F}, E^{1,SD}, E^{2,SD}, E^{3,SD}, E^{1,SU}, E^{2,SU}, E^{3,SU}, E^{1,Q}, E^{2,Q}, E^{3,Q}, I^F, A^F, I^{QF}, I^{SD}, A^S, I^{SU}, I^{QS}, A^Q)$; that is the infectious system of the Warwick-Household model is determined by the following ODEs:

$$
\begin{aligned}
\frac{dE_i^{1,F}}{dt} &= \sigma_i \sum_j C_{ij}^N \big(\frac{S_i}{N} I_j^F + \frac{S_i}{N} I_j^{SD} + \frac{S_i}{N} I_j^{SU} + \tau(\frac{S_i}{N} A_j^F + \frac{S_i}{N} A_j^S)\big) - 3\epsilon E_i^{1,F} \\
\frac{dE_i^{2,F}}{dt} &= 3\epsilon E_i^{1,F} - 3\epsilon E_i^{2,F} \\
\frac{dE_i^{3,F}}{dt} &= 3\epsilon E_i^{2,F} - 3\epsilon E_i^{3,F} \\
\frac{dE_i^{1,SD}}{dt} &= \sigma_i \sum_j C_{ij}^H \frac{S_i}{N} I_j^F - 3\epsilon E_i^{1,SD} \\
\frac{dE_i^{2,SD}}{dt} &= 3\epsilon E_i^{1,SD} - 3\epsilon E_i^{2,SD} \\
\frac{dE_i^{3,SD}}{dt} &= 3\epsilon E_i^{2,SD} - 3\epsilon E_i^{3,SD} \\
\frac{dE_i^{1,SU}}{dt} &= \sigma_i \sum_j C_{ij}^H \frac{S_i}{N} A_j^F - 3\epsilon E_i^{1,SU} \\
\frac{dE_i^{2,SU}}{dt} &= 3\epsilon E_i^{1,SU} - 3\epsilon E_i^{2,SU} \\
\frac{dE_i^{3,SU}}{dt} &= 3\epsilon E_i^{2,SU} - 3\epsilon E_i^{3,SU} \\
\frac{dE_i^{1,Q}}{dt} &= \sigma_i \sum_j C_{ij}^H \frac{S_i}{N} I_j^{QF} - 3\epsilon E_i^{1,Q} \\
\frac{dE_i^{2,Q}}{dt} &= 3\epsilon E_i^{1,Q} - 3\epsilon E_i^{2,Q} \\
\frac{dE_i^{3,Q}}{dt} &= 3\epsilon E_i^{2,Q} - 3\epsilon E_i^{3,Q} \\
\frac{dI_i^F}{dt} &= 3(1-H)\epsilon d_i E_i^{3,F} - \gamma I_i^F \\
\frac{dA_i^F}{dt} &= 3\epsilon(1-d_i) E_i^{3,F} - \gamma A_i^F
\end{aligned}
$$

$$
\begin{aligned}
\frac{dI_i^{QF}}{dt} &= 3H\epsilon d_i E_i^{3,F} - \gamma I_i^{QF} \\
\frac{dI_i^{SD}}{dt} &= 3\epsilon d_i E_i^{3,SD} - \gamma I_i^{SD} \\
\frac{dA^S}{dt} &= 3\epsilon(1-d_i)(E_i^{3,SD} + E_i^{3,SU}) - \gamma A_i^S \\
\frac{dI_i^{SU}}{dt} &= 3(1-H)\epsilon d_i E_i^{3,SU} - \gamma I_i^{SU} \\
\frac{dI_i^{QS}}{dt} &= 3H\epsilon d_i E_i^{3,SU} + 3\epsilon d_i E_i^{3,Q} - \gamma I_i^{QS} \\
\frac{dA_i^Q}{dt} &= 3\epsilon(1-d_i) E_i^{3,Q} - \gamma A_i^Q,
\end{aligned}
$$

where i is the age group of the individual. Accounting for the different age-structured compartments, the infectious system vector becomes:

$$V = (E_1^{1,F}, \dots E_M^{1,F}, E_1^{2,F}, \dots E_M^{2,F}, \dots, A_1^Q \dots, A_M^Q),$$

where we indicate by $M = 8$ the total number of age groups modelled. It follows then that the transmission (T) and transition (Σ) matrices, which we defined previously in section C4.2 are equal to:

$$
T = \left(\begin{array}{c|c} \mathbf{0} & \mathcal{T} \\ \hline \mathbf{0} & \mathbf{0} \end{array}\right) \text{ and } \Sigma = \left(\begin{array}{c|c} -3\epsilon\mathbb{I}_{12M} + 3\epsilon\mathcal{J} & \mathbf{0} \\ \hline 3\epsilon\mathcal{S} & -\gamma\mathbb{I}_{8M} \end{array}\right),
$$

where $\mathbb{I}_{8M}$ is the $8M \times 8M$ identity matrix and $\mathbb{I}_{12M}$ is the $12M \times 12M$ identity matrix and the $\mathbf{0}$s represent blocks matrices with all elements equal to zero. We used the following $\mathcal{T}$

and $\mathcal{S}$ as shorthand for the following block matrices:

$$\mathcal{T} = \begin{pmatrix} S^T\sigma\frac{1}{N}C^N & S^T\sigma\frac{\tau}{N}C^N & 0 & S^T\sigma\frac{1}{N}C^N & S^T\sigma\frac{\tau}{N}C^N & S^T\sigma\frac{1}{N}C^N & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ S^T\sigma\frac{1}{N}C^H & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & S^T\sigma\frac{1}{N}C^H & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & S^T\sigma\frac{1}{N}C^H & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

and, respectively,

$$\mathcal{S} = \begin{pmatrix} 0 & 0 & (1-H)(d)_M & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & (1-d)_M & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & H(d)_M & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & (d)_M & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & (1-d)_M & 0 & 0 & (1-d)_M & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & (1-H)(d)_M & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & H(d)_M & 0 & 0 & (d)_M \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & (1-d)_M \end{pmatrix},$$

where the 0 in these block matrices refers to the $M \times M$ matrix with all elements equal to zero. $(d)_M$ is a diagonal matrix with the i^{th} element on the diagonal equal to d_i (and similarly $(1-d)_M$ is a diagonal matrix with the i^{th} element on the diagonal equal to $1-d_i$). Additionally S^T is the row vector or remaining susceptibles at the time of estimation of the reproduction number R_t and C^H and C^N are the household respectively the non-household contact matrices.

Using a similar notation to above, that is we write the matrix $\mathbb{I}_M$ to represent the $M \times M$ identity matrix, we define the final block to be defined, $\mathcal{J}$ as having the following matrix structure:

$$\mathcal{J} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \mathbb{I}_M & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \mathbb{I}_M & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \mathbb{I}_M & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \mathbb{I}_M & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \mathbb{I}_M & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbb{I}_M & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbb{I}_M & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbb{I}_M & 0 \end{pmatrix}.$$

The time-dependent reproduction number R_t is then defined as the spectral radius of the matrix $-T\Sigma^{-1}$:

$$R_t = \rho(-T\Sigma^{-1}).$$

C4.4 Comparison of the trajectories of daily infection incidence

C4.4.1 The PHE Model

First, we compare the total number of infectious people at a specific time point classified by region and age group. For each region and age group we plot the trajectory of the number of infectious individuals for each of the two regimes in which we run the model: non-changing contact matrices in blue, and in red the scenario in which we account for interventions. However, it is important to keep in mind that our choice to not use the average estimated values for the parameters we infer for the PHE model implies that the simulated trajectories of the outcomes of interest will look different from the ones from the original paper – see Figure C9 for reference, which has a more jagged outline for the total number of infections for the regions of London and North West, compared to the corresponding subplots of Figure C10 where this effect is not present, as in this case the β_t regional variations are assumed constant.

In Figure C9 we reproduced some of the results of the original PHE paper [Birrell et al.,

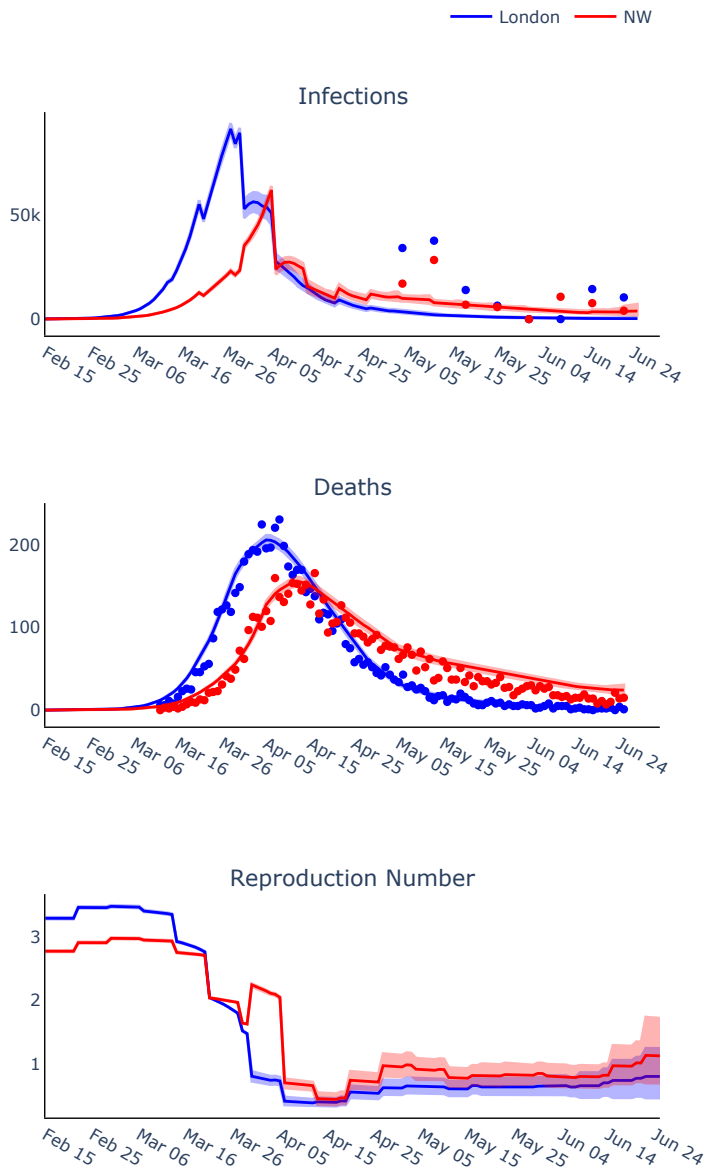


Figure C9: *Mean and* 95% *confidence regions of trajectories of the daily total number of new infections, deaths and instantaneous reproduction number* R_t *for London (blue lines) and North West (red lines) regions. These trajectories are produced using the PHE model. The dots represent the observed number of positive tests scaled to the population (first panel) and number of deaths (second panel).*

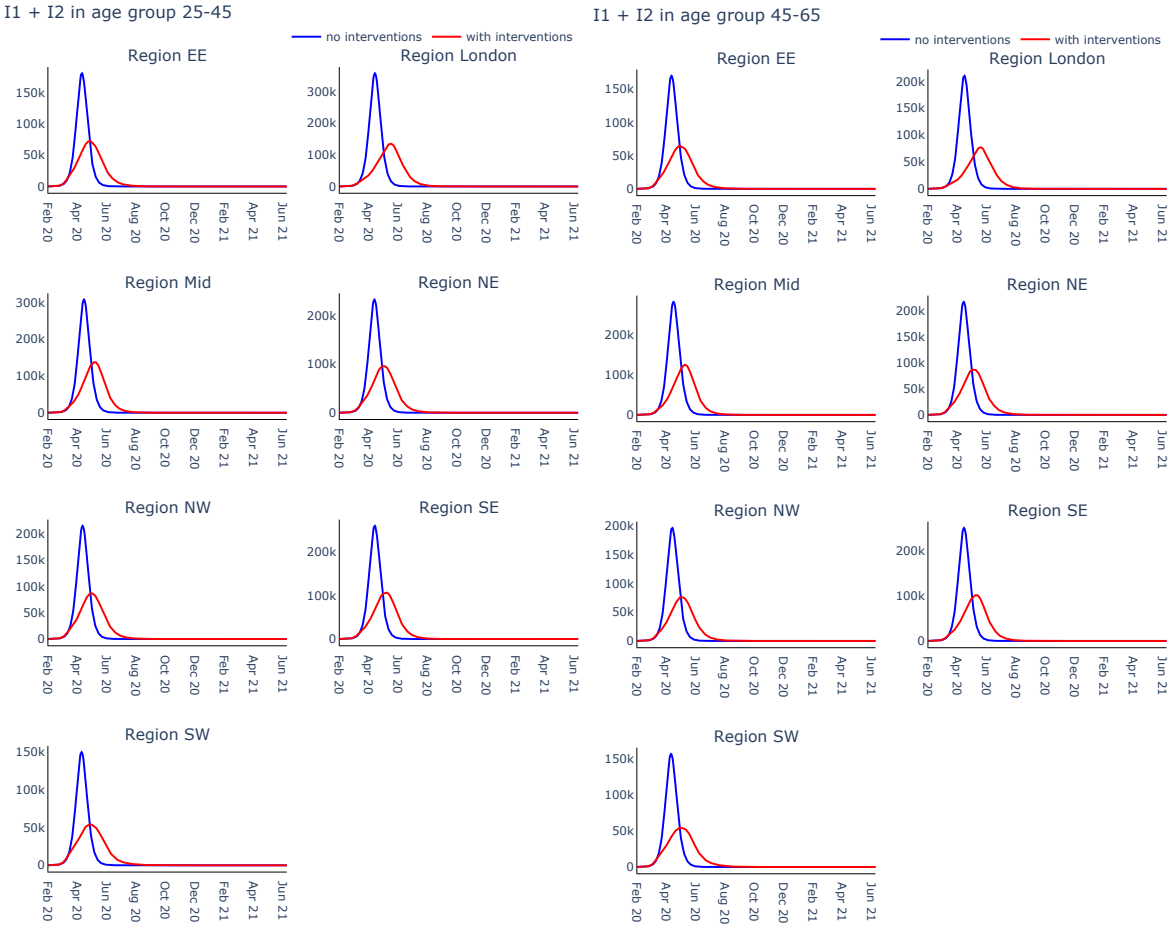
2021], mainly those of Figure 3. We run two inference algorithms, one for each of the PHE regions considered – London (plotted in blue) and North West (plotted in red). The initial number of infected individuals in the North West is fixed at a lower level compared to London due to differences in demography (30 individuals in each age group compared to 50 respectively). The shapes of both the number of infections and the number of deaths matches those in the paper, which validates our reimplementation of the model. However,

the width of the uncertainty measure is smaller in our figure, compared to the original; this is because we aim to infer a much smaller number of parameters compared to the Cambridge group, in the interest of both time and computation resources.

Also worth mentioning is the difference in the instantaneous reproduction number profiles R_t for the two regions: for the North West, R_0 is smaller than in London, mainly an effect of the population counts. However, the subsequent drop in the R_t occurs much later than for London; in fact it appears to be stuck around $R_t = 2$ for a couple of weeks after the 16$^{\text{th}}$ of March mark. The first observation could explain why the peak for both the number of deaths and new infections is lower and occurs later in the North West compared to London, while the second motivates the slower down-slope of the number of deaths in the North West region.

In Figure C10 we observe the trajectories for the two most populated age groups, which are coincidentally the age groups with the largest average number of contacts, i.e. the $25-45$ and $45-65$ years old. For both age groups, we observe for all the regions that in the case of no interventions, the highest count for the total number of infectious individuals occurs not only earlier in the epidemic compared to the other scenario – usually by around 2 months, but the height of the peak is also larger, usually by a factor of 1.5. This effect is what has been colloquially called 'flattening of the curve' through interventions and was one of the main arguments made by the British Government for the guidelines issued early in the COVID-19 epidemic in early 2020 [Ferguson et al., 2020; Adam, 2020]. Witnessing this effect for these particular age groups, shows that the measures taken would have had an effect and so a health crisis has been averted. Similarly shaped graphs occur for the other age groups as well.

There is always only one peak present in either of the scenarios, which does not follow the shape of the curve of the numbers reported [GOV.UK]. This is due to the fact that in this version of the PHE model no waning immunity is considered, i.e. once recovered, an individual cannot be infected again. However, such an assumption is not true for COVID-19; multiple studies [Leidi et al., 2021; Abu-Raddad et al., 2021; Gallais et al., 2021] indicate that the protection against the virus can last between four months and up to a year. This means that in its current form, the PHE model can only reliably model the short-term behaviour of the epidemic; for longer-term predictions, some changes need to be made to the model structure.



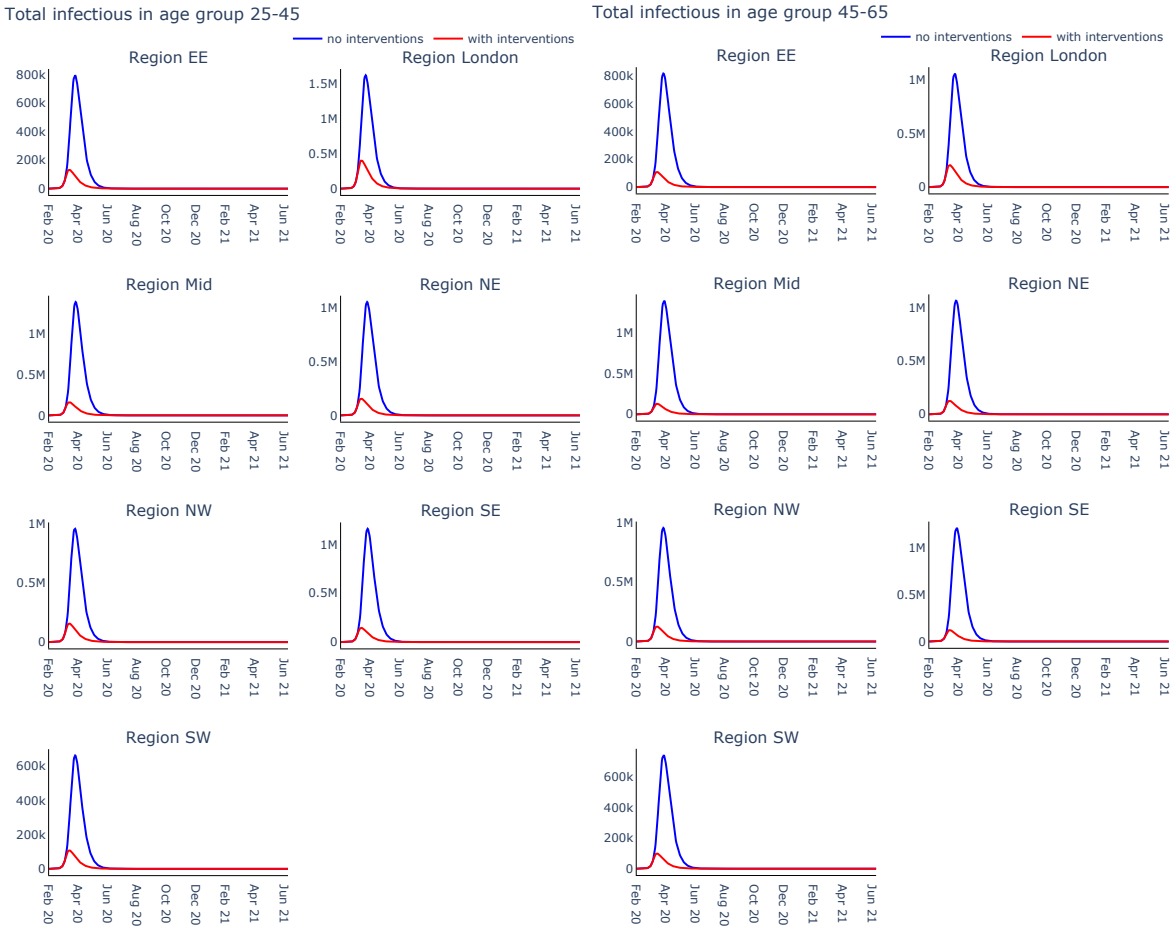
(a) 25 to 45 years old (b) 45 to 65 years old

Figure C10: *Trajectories of the total number of infected individuals split by region for different age groups: (a) 25 to 45 years old vs (b) 45 to 65 years old for both the interventions (red lines) and baseline scenarios (blue lines). These trajectories were produced by running forward simulations of the PHE model.*

C4.4.2 The Roche Model

Similar to the PHE model, we now compare in Figure C11 the time-dependent trajectories of the total number of infectious individuals in each PHE region for the two largest age groups ($25-45$ and $45-65$ years old) for two regimes: when interventions are applied as seen during the first wave of the UK COVID-19 epidemic (in red) and when no interventions are imposed and the contact matrix of the model is kept at its baseline level (in blue). We observe that for both age groups and in all considered regions, the peak of the number infections when the population retains their baseline behaviour pre-pandemic is on average four time higher than in the case when NPIs are applied. Moreover, we observe that until mid-March when lockdowns were imposed the trajectories for both scenarios considered

line up, which suggests that interventions have a significant impact in lowering the number of infections from the moment of their introduction. Similar conclusions follow for the other age groups considered as well.



(a) 25 to 45 years old (b) 45 to 65 years old

Figure C11: *Trajectories of the total number of infected individuals split by region for different age groups: (a) 25 to 45 years old vs (b) 45 to 65 years old for both the interventions (red lines) and baseline scenarios (blue lines). These trajectories were produced by running forward simulations of the Roche model.*

Comparing these trajectories of infections with their equivalents generated from the PHE model in Figure C10, we observe that number of infected predicted by the Roche model is more then three orders of magnitude larger than in the case of the PHE model when contacts are maintained at their baseline levels; additionally, the peak of infections in the scenario when interventions were implemented occurred before April 2020 in the Roche model, compared to after April 2020 as it did in the case of the PHE model. These results suggest that the instantiations of the PHE and the Roche models do not capture the same

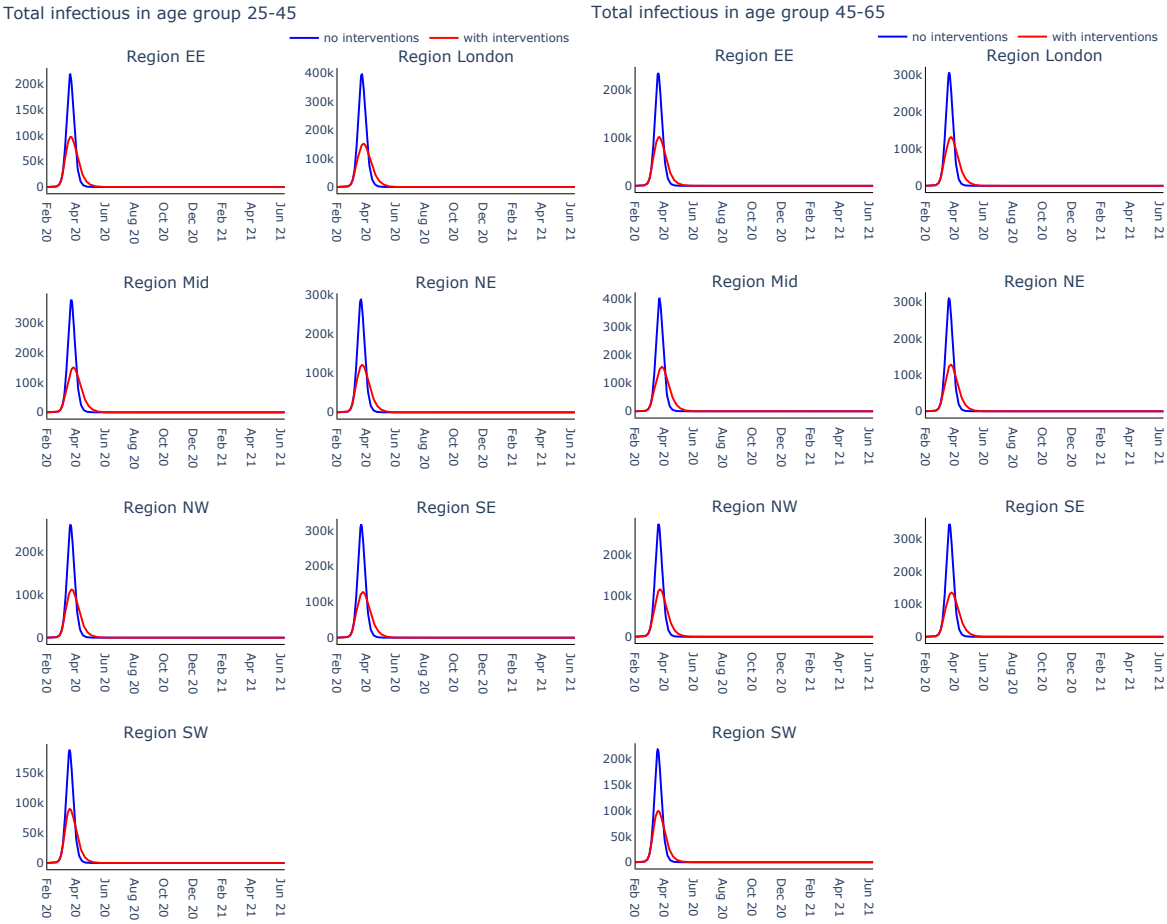
epidemic dynamics, and hence they imply that we cannot substitute the results of one for the model for the others when quantify the effects of interventions.

It is important to note that only one peak is observed in either of the scenarios for all regional and age-group trajectories of infection incidence. This contrasts with the shape of the curve of the numbers reported by ONS [GOV.UK] and is due to the fact that, just like the PHE model, the Roche model does not take into account waning immunity in the model structure – that is, once recovered, an individual cannot be infected again. As discussed in multiple studies such as [Leidi et al., 2021; Abu-Raddad et al., 2021; Gallais et al., 2021] immunity can last between four months and up to a year, which means that disregarding this feature in setting up the model will have a significant impact in the long-term prediction of epidemic dynamics – primarily by the fact that no secondary wave can be ever be predicted. Therefore, in its current form, we can only use the Roche model for short-term, less than three months, predictions.

C4.4.3 The Warwick-Household Model

We now proceed with analysing how interventions affect the trajectories of the daily number of new infections generated when using the Warwick-Household model instead. Comparing the trajectories of the number of infections in each of the seven PHE regions introduced in subsection C1.1.1 for the two largest age groups (the $25-45$ and the $45-65$ years old) we observe similar to the Roche model implementation (described in section C4.2) that the peak of the number of infections obtained when interventions have been applied as they have been during the first wave of the UK COVID-19 epidemic (the red curve) is approximately two-fold lower than the peak of the number of infections that would have been observed if the contact matrix remains unchanged from its initial value and no other interventions are applied (the blue curve). Additionally, until mid-March when NPIs begin to be applied during the first COVID-19 epidemic wave in the UK, the trajectories of the number of infections for both simulation scenarios align, which further suggest that the NPIs implemented are responsible for the significant decrease in the predicted number of infections.

Similar to the Roche and the PHE models, the Warwick-Household model structure does not account for any renewal of susceptibles – that is, there is no waning immunity incorporated. This prevents the modelling of multi-peak epidemics, which was the case with the number



(a) 25 to 45 years old (b) 45 to 65 years old

Figure C12: *Trajectories of the total number of infected individuals split by region for different age groups: (a) 25 to 45 years old vs (b) 45 to 65 years old for both the interventions (red lines) and baseline scenarios (blue lines). These trajectories were produced by running forward simulations of the Warwick-Household model.*

of cases observed in the first six months of the COVID-19 epidemic in the UK [GOV.UK]. Hence, due to the lack of waning immunity in the model structure, the Warwick-Household model is only suitable for modelling the early part of an epidemic or short-time scale epidemics, less in duration than the average waning immunity cycle for COVID-19, i.e. less then an year, according to [Leidi et al., 2021; Abu-Raddad et al., 2021; Gallais et al., 2021].

C4.5 Comparison of the trajectories of daily number of deaths

C4.5.1 The PHE Model

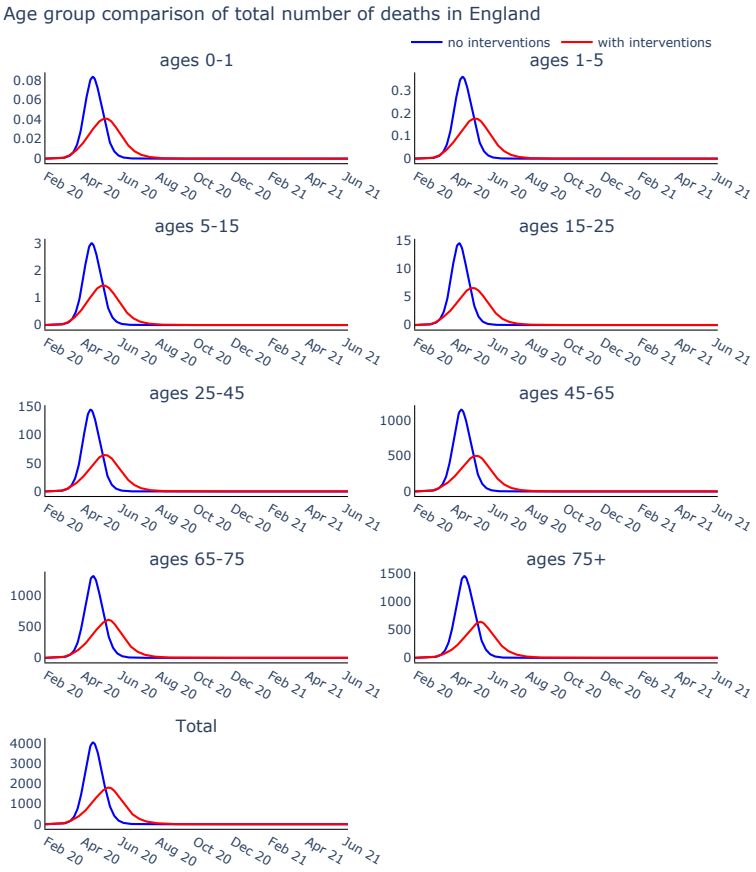


Figure C13: *Trajectories of the cumulative number of deaths for the whole of England split by different age groups for both the interventions (red lines) and baseline scenarios (blue lines). These trajectories were produced by running forward simulations of the PHE model. NB: Since this is an ODE system we can obtain values of deaths that are less than* 1 *in the younger age group.*

In this section, we study the trajectory of the daily number of deaths throughout the epidemic. Using the corresponding sampling methods in the PHE model class and assuming we know the levels in each compartment of the model at all times, e.g. by using the output of the *simulate* method, we sample possible number of deaths for each age category and region. We plot the cumulative number of deaths for each age group considered, across all regions, as well as the total number of deaths in England as in Figure C13. These trajectories were produced using the original initialisation considered in this section i.e. 10 individuals in each age and region category for the I^1 compartments, with all other non-susceptible compartments initialised at 0. However, we observe very high daily death tolls for the whole of

England, at over 2000 even in the intervention case scenario. This is much larger than the number of deaths registered at the the peak of the first UK epidemic wave, which leads us to believe there might be a problem with the setup of the model – circling back to the challenge of finding a suitable instantiation the model.

C4.5.2 The Roche Model

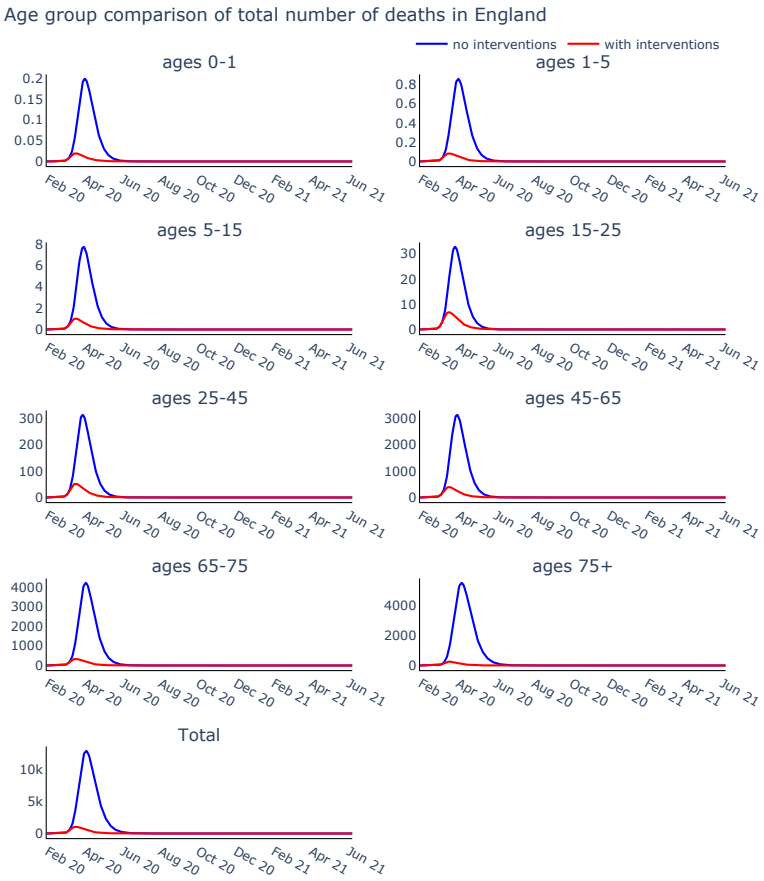


Figure C14: *Trajectories of the cumulative number of deaths for the whole of England split by different age groups for both the interventions (red lines) and baseline scenarios (blue lines). These trajectories were produced by running forward simulations of the Roche model. NB: Since this is an ODE system we can obtain values of deaths that are less than* 1 *in the younger age group.*

Similar conclusions can drawn in the case of the Roche model. We consider the trajectories of the daily number of deaths for each age group across all PHE region as well as for the total number of deaths in England, the latter of which we plot in Figure C14. These numbers were isolated from the outputs produced by running the *simulate* method of the Roche model class using the initialisation of the model described above, that is with 40 individuals in each age

group and region for all symptomatic infectious I^s compartments and with no individuals present in any of the other non-susceptible compartments.

The total death toll in England at the peak of the first epidemic wave when contacts remain at their baseline pre-pandemic levels and no interventions are applied exceeds the total number of fatalities recorded (represented by the red trajectories in the Figure C14) more than twelve-fold, which imply a significant impact of NPIs in containing the number of severe disease outcomes under this model. This effect is much more pronounced than in the case of the PHE model predictions (see Figure C13): in the case of that model, the peak in the number of deaths when interventions were applied occurs around June 2020 (as opposed to March 2020 in the Roche model); additionally, there is a more than three-fold increase in the peak number of deaths predicted by the Roche model compared to the PHE, if contacts remain unchanged throughout. This implies we are unable to reproduce the same epidemic dynamics of the number of deaths described by the PHE model using the Roche model.

C4.5.3 The Warwick-Household Model

In the following, we compare the trajectories of the daily number of deaths for each age group across all PHE regions as well as for the total number of deaths in England (the latter of which we plot in Figure C15) obtained by running forward-simulation of the Warwick-Household model, parametrised as specified in section 4.3.5.3. We observe that, unlike the incidence of infection trajectories, there is a substantial diminution in the effect of the implemented NPIs in terms of severe outcomes of infection – indeed for most age-group, including the $75+$ in which the largest incidence of deaths is recorded, the trajectories described by the two scenarios are almost equal.

Similar to the daily number of infection curves, a singular peak is observed in either of the scenarios considered, which again differs significantly from with the shape of the curve of the numbers reported by the ONS [GOV.UK]; this is because just as the PHE and the Roche models, the Warwick-Household model does not account for waning immunity. However, the magnitude difference between the peak in the total number of deaths in England in the case when interventions are applied is only approximately 1.3 times less compared to the no-interventions scenario, when the contact matrix equals their baseline pre-pandemic one.

These trajectories were obtained from the outputs produced by running the *simulate* method

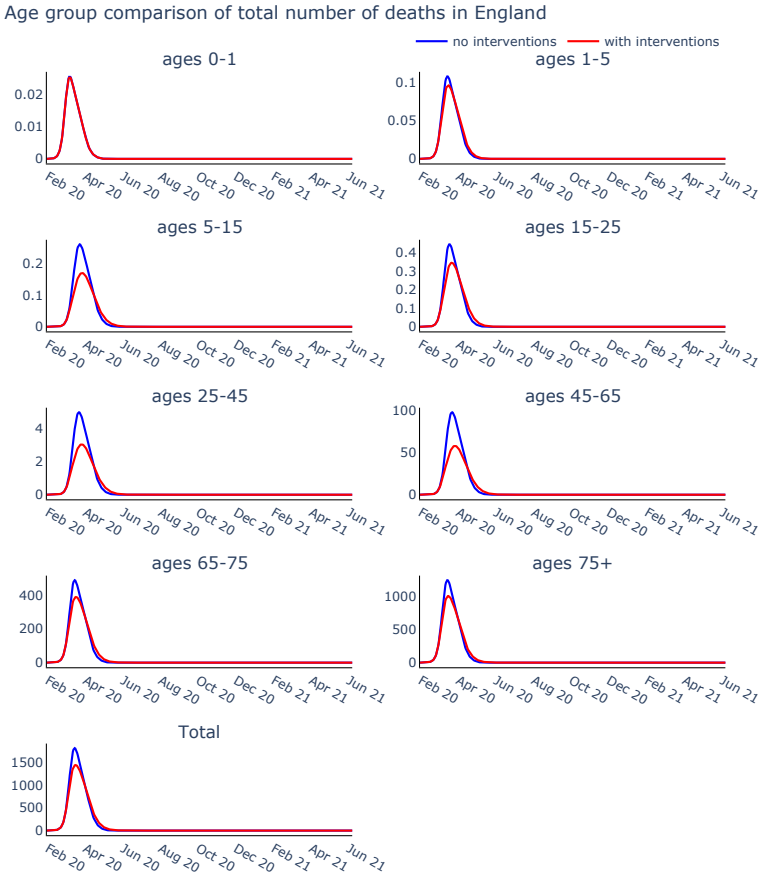


Figure C15: *Trajectories of the cumulative number of deaths for the whole of England split by different age groups for both the interventions (red lines) and baseline scenarios (blue lines). These trajectories were produced by running forward simulations of the Warwick-Household model. NB: Since this is an ODE system we can obtain values of deaths that are less than* 1 *in the younger age group.*

of the Warwick-Household model class using the initialisation of the model described above, that is with 200 total number of individuals symptomatically infected (I^F) in the initial population in each region, distributed across the different age groups according to the initial infection ratio from eq. (4.6); no other individuals are assumed present in any of the other non-susceptible compartments.

C4.6 Comparison of the trajectories of time-dependent reproduction number

C4.6.1 The PHE Model

In this final set of analyses for the PHE model, we plot the trajectories of the instantaneous reproduction number of the virus over the course of the epidemic period considered, as seen in Figure C16. All the way back in the Introduction to this thesis, we defined the

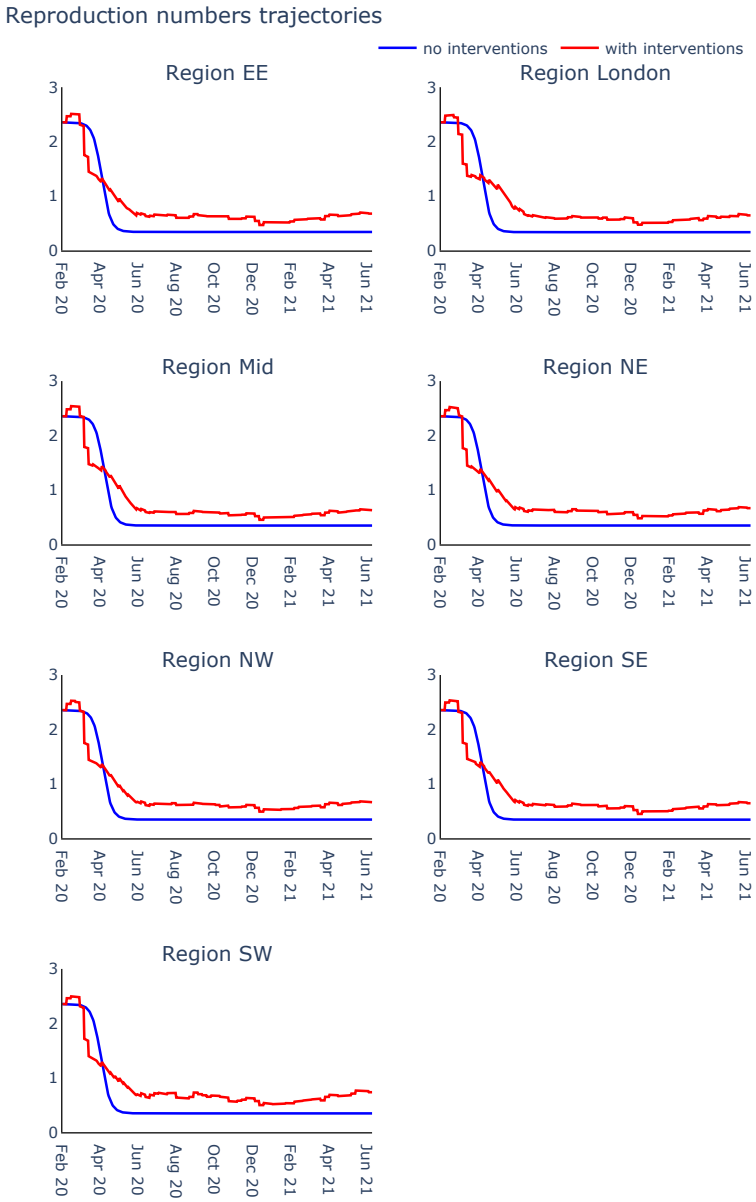


Figure C16: *Trajectories of the instantaneous reproduction numbers split by different regions for both the interventions (red lines) and baseline scenarios (blue lines). These trajectories were produced by running forward simulations of the PHE model.*

instantaneous reproduction number as the expected number of secondary infections caused by one infective host at time t assuming that future transmission remains the same as at time t. From the way the number of total infectious individuals form peaks in the two scenarios considered, they replicate the delay seen in Figure C10. The more stochastic looking behaviour of the 'with interventions' curve is mainly due to the changes in the profile of contacts with respect to time. The slight differences of the R_t curves for 'no interventions' between the regional subplots are most likely due to the differences in the number of individuals and age-structure between each region.

The fact that the instantaneous reproduction number does not go up after the first plummeting, in either of the scenarios, means that no second wave can be predicted by the PHE model. This suggests that we lack some information in the model: either the Google mobility data is not accurate enough, or we disregard some parameter that captures the seasonality of the disease, e.g. the effect of temperature on transmissibility and the waning immunity of those recovered.

C4.6.2 The Roche Model

Here we discuss the behaviour of the instantaneous reproduction number for each of the two scenarios discussed in section C4.2, as seen in Figure C17. In both cases we observe a very high initial reproduction number (above 10) which quickly drops to below one once interventions are put in place – this could lead us to believe that the Roche model mechanism for predicting R_t produces unrealistic values. However, when using the formula for computing the final epidemic size given the initial reproduction number R_0 in the case of a simple SIR model [Tildesley and Keeling, 2009; Kermack et al., 1927]:

$$R_\infty = S(0)\big(1 - e^{R_0 R_\infty}\big), \tag{C.16}$$

the attack rate AT $= 1 - R_\infty$ of an epidemic with initial $R_0 = 10.46$ (as was the case for the London region) is equal to 99.98% – which is very close to the percentage of the population not susceptible at the end of the simulation time interval of 456 days for the Roche example implementation, i.e. 99.6%. We can explain this high R_0 value by the fact that the Roche model assumes all symptomatic individuals move very quickly to the quarantined state (2.5 day); this requires a high initial reproduction number, in order to sustain the number of

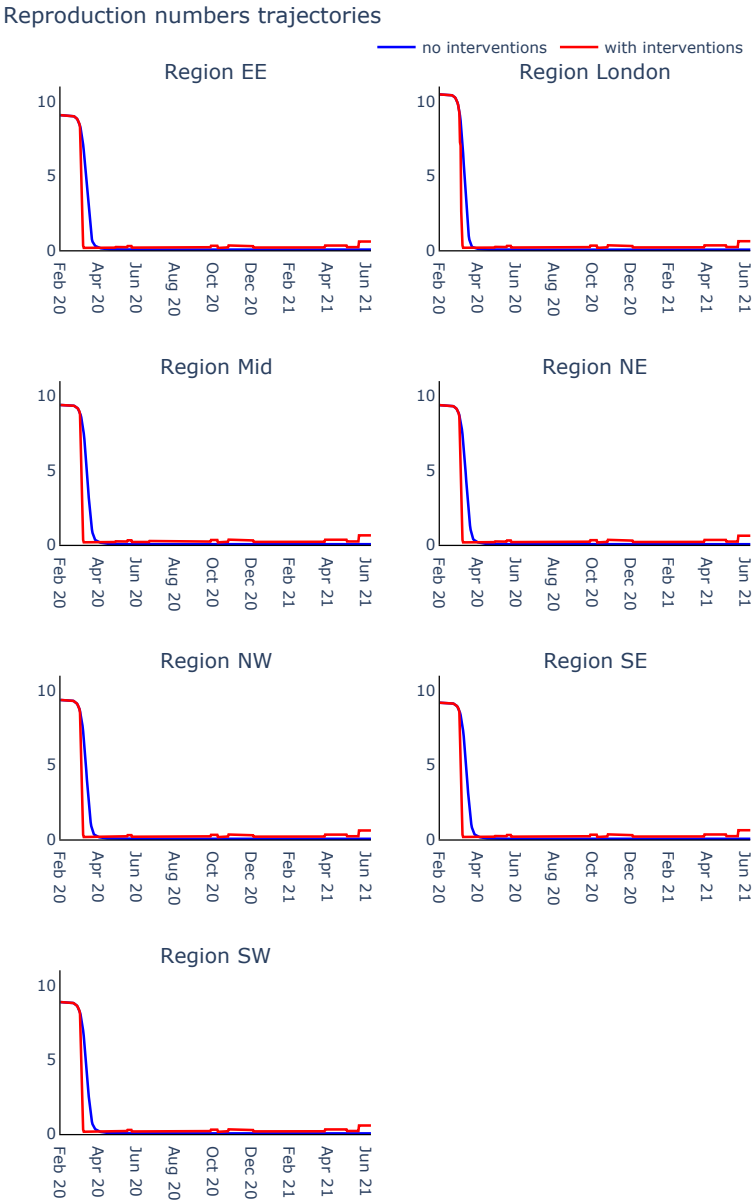


Figure C17: *Trajectories of the instantaneous reproduction numbers split by different regions for both the interventions (red lines) and baseline scenarios (blue lines). These trajectories were produced by running forward simulations of the Roche model.*

infections and deaths observed in the real data.

However, a similar if slightly delayed behaviour is observed when the contacts remain at their pre-pandemic levels and no NPIs are implemented. The sharp drop in the instantaneous R_t trajectory in the case when contacts remain the same throughout the simulation interval can be explained by the quick uptake in infection predicted by the Roche model in this scenario (see Figure C11) – in this case all available susceptible individuals go on to become

exposed and then infected very quickly and we are left with very few other individuals on pass the diseases onto. In contrast, the steep decline in R_t for the original scenario with interventions could be explained by the strong effect of interventions in suppressing transmission. This behaviour contrasts what we have observed in the case of the PHE model, where a longer but shallower decline in R_t can be observed over the 465 days modelled.

C4.6.3 The Warwick-Household Model

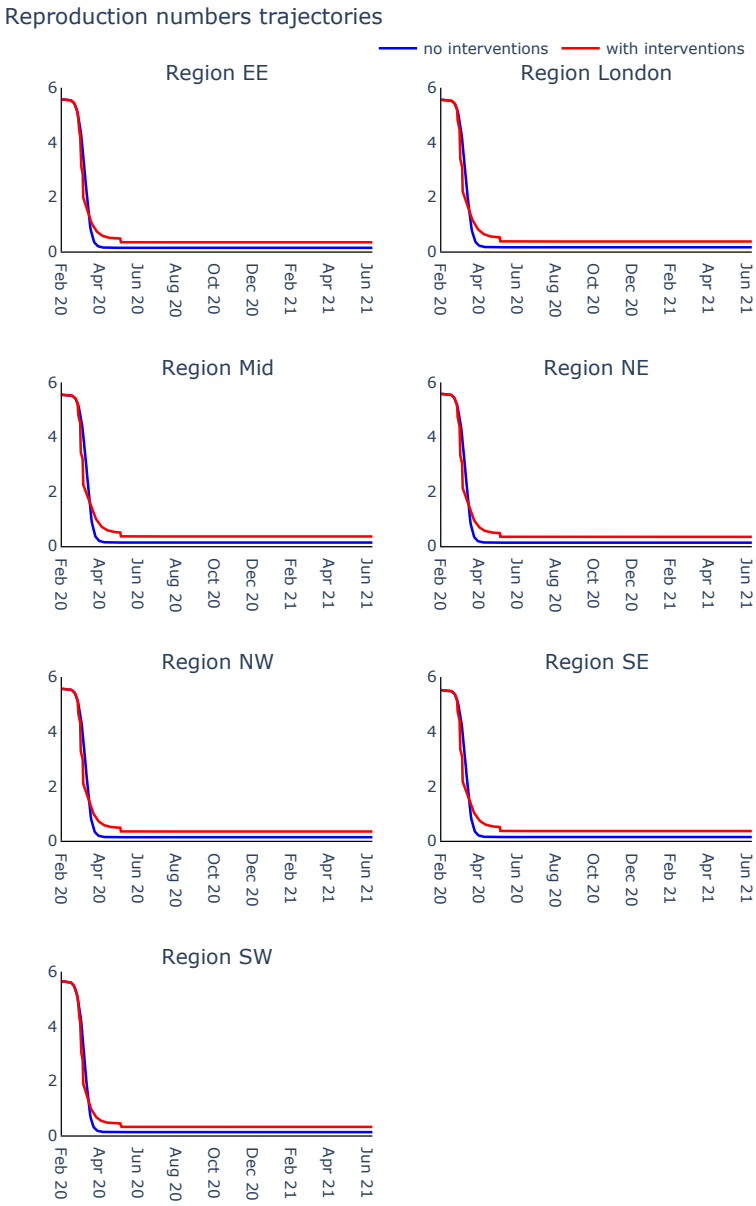


Figure C18: *Trajectories of the instantaneous reproduction numbers split by different regions for both the interventions (red lines) and baseline scenarios (blue lines). These trajectories were produced by running forward simulations of the Warwick-Household model.*

Finally, in Figure C18 we illustrate the changes in the instantaneous reproduction number over time for each of the two scenarios discussed above. In both cases we observe a moderate initial reproduction number (around 6 in all PHE regions) which quickly drops to below one once interventions are put in place (around the April 2020 mark). However, a similar if slightly delayed behaviour is observed when the contacts remain at their pre-pandemic levels and no NPIs are implemented; this behaviour contrasts what we have observed in the case of the PHE Model, where a longer but shallower decline in R_t can be observed over the 465 days modelled, but it more closely resembles the one observed in the Roche model.

Now, in order to assess how realistic are the Warwick-Household model's predictions of the trajectory of R_t, considering its high values for the initial reproduction number R_0, we proceed in a similar manner to the Roche model case. The total percentage of the population that has been infected at the end of the 456 day simulation window of the Warwick-Household model is equal to 96.7% in the region of London, which is in agreement with the 99.6% value of the epidemic attack rate predicted for an initial $R_0 = 5.5$ using the SIR model.

C5 Shortcomings of the existing form of the models

In the previous section, we compared the outputs of the PHE, the Roche and the Warwick-Household models for two scenarios for a specified parameterisation of the model: (1) the baseline behaviour and (2) with non-pharmaceutical intervention in place to affect the value of the contact matrix. We concluded that the produced results do not fully resemble the values from observed data for any of the three models considered. This is due to our lack of knowledge of certain parameters such as the initial number of individuals in each compartment at the start of the simulation time window.

Moreover, by foregoing modelling of waning immunity in the structure of the PHE, the Roche and the Warwick-Household models, the outputs of the model cannot reproduce the multi-peak shape of the long-term observed data on the number of deaths or cases.

The second issue implies that we should restrict the use of these models to only short-term forecasting, less in length than the average waning immunity cycle for COVID-19 [Leidi et al., 2021; Abu-Raddad et al., 2021; Gallais et al., 2021]. However, the first challenge can be addressed by letting the data inform on the suitable parameterisation of the model, especially the parameters for which no clear, study-led defined values have been identified.

In Chapter 4, we illustrate the methodology through which we tackle this issue: we first complete a parameter sensitivity analysis to identify those model parameters which produce the biggest shift in the model output when letting their values vary – big fluctuations in the output indicate that the true value of the parameter can be recovered from the data, making it an ideal candidate for the next steps of parameter optimisation or inference.

Additionally, in its current form the Roche model considers the effect of only a small number of state-imposed interventions, while completely disregarding the impact of region- and age-specific population response: whether they adhere to the implemented restrictions or they practice self-isolation. Similarly, the Roche model assumes all symptomatic individuals move eventually to the quarantined state, which is not a realistic assumption – in reality, a not insignificant percentage of those infected may not self-isolate at all.

D | Appendix II Chapter 4

D1 The PHE model

The one region, one age group PHE model consists of the following system of ODEs:

$$\begin{aligned}
\frac{dS(t)}{dt} &= -\lambda_t S(t) \\
\frac{dE^1(t)}{dt} &= \lambda_t S(t) - \frac{2}{d_L} E^1(t) \\
\frac{dE^2(t)}{dt} &= \frac{2}{d_L} E^1(t) - \frac{2}{d_L} E^2(t) \\
\frac{dI^1(t)}{dt} &= \frac{2}{d_L} E^2(t) - \frac{2}{d_I} I^1(t) \\
\frac{dI^2(t)}{dt} &= \frac{2}{d_I} I^1(t) - \frac{2}{d_I} I^2(t) \\
\frac{dR(t)}{dt} &= \frac{2}{d_I} I^2(t).
\end{aligned} \tag{D.1}$$

Our aim is to determine which parameter combinations of the SEIR model – satisfying the system of equations (4.1), allows us to retain as many features of the original PHE model as possible.

The PHE model assumes no asymptomatic infections and has two compartments for the exposed and two compartments for the infectious individuals, compared to the SEIR model which has only one for each. As mentioned in the Appendix C, this allows for more realistic Erlang-distributed waiting times in the exposed and infectious infection states. In the PHE model the total number of exposed and infectious individuals is given respectively by:

$$E(t) = E^1(t) + E^2(t) \text{ and } I(t) = I^1(t) + I^2(t),$$

therefore, we will collapse the PHE model structure along these two infection states.

To ensure equivalence of the first equation in both of the ODEs systems, the PHE force of infection parameter $\lambda_t = 1 - (1 - b_t)^{I^1(t)+I^2(t)}$ has to be similar in shape to the one associated with the SEIR model, i.e. $\beta I(t)$. For this to happen the conditions of the *binomial*

approximation need to be met, that is:

$$\lambda_t = 1 - (1 - b_t)^{I^1(t)+I^2(t)} \approx 1 - (1 - b_t(I^1(t) + I^2(t))) = 1 - (1 - b_t I(t)) = b_t I(t)$$

if $|b_t| < 1$ and $|b_t I(t)| \ll 1$. In the following, we investigate if these conditions are satisfied in the context of the PHE model.

The PHE parameter for the probability of transmission, b_t, is computed according to the following formula:

$$b_t = \frac{\beta_t R_0}{R_0^*} \tilde{C}_t,$$

where β_t accounts for any misspecification in the time-dependent fluctuations in transmission, R_0 is the basic reproduction number, $\tilde{C}_t$ is the current contact matrix and R_0^* is the dominant eigenvalue of the initial transmission matrix $S_0 \tilde{C}_0 d_I$, with S_0 being the initial number of susceptibles and d_I the average time an individual is infectious. Since we have only one age group, the contact matrix is one-dimensional and therefore $R_0^* = S_0 \tilde{C}_0 d_I$. As the parameters of the SEIR model do not exhibit any temporal variation, we further assume that the contact matrix for the PHE model is constant throughout; hence $\beta_t = 1, \forall t$. The probability of transmission parameter thus becomes

$$b_t = \frac{R_0}{S_0 d_I},$$

which satisfies the *binomial approximation* condition $\left| \frac{R_0 I(t)}{S_0 d_I} \right| \ll 1$, as the number of infectives is much smaller than that of the initial susceptible population. Moreover, since the rates of transmission between the compartments are defined as the inverse of the average time spent in the compartment, the system of equations of the PHE model indicate that the average time an individual spends in each of the exposed and infectious compartments are:

- E^1 and E^2: $\frac{d_L}{2}$ each, so a total average of d_L as exposed;
- I^1 and I^2: $\frac{d_I}{2}$ each, so a total average of d_I as infectious.

Therefore, the PHE model reduces to an SEIR framework with the model parameters given by $\beta = \frac{R_0}{S_0 d_I}$, $\kappa = \frac{1}{d_L}$ and $\gamma = \frac{1}{d_I}$. In Figure D1 we plot the trajectories of the number of individuals in each compartment for an example instantiation of the PHE model (the blue curve) and its SEIR model counterpart with matched parameters (in red). We choose the

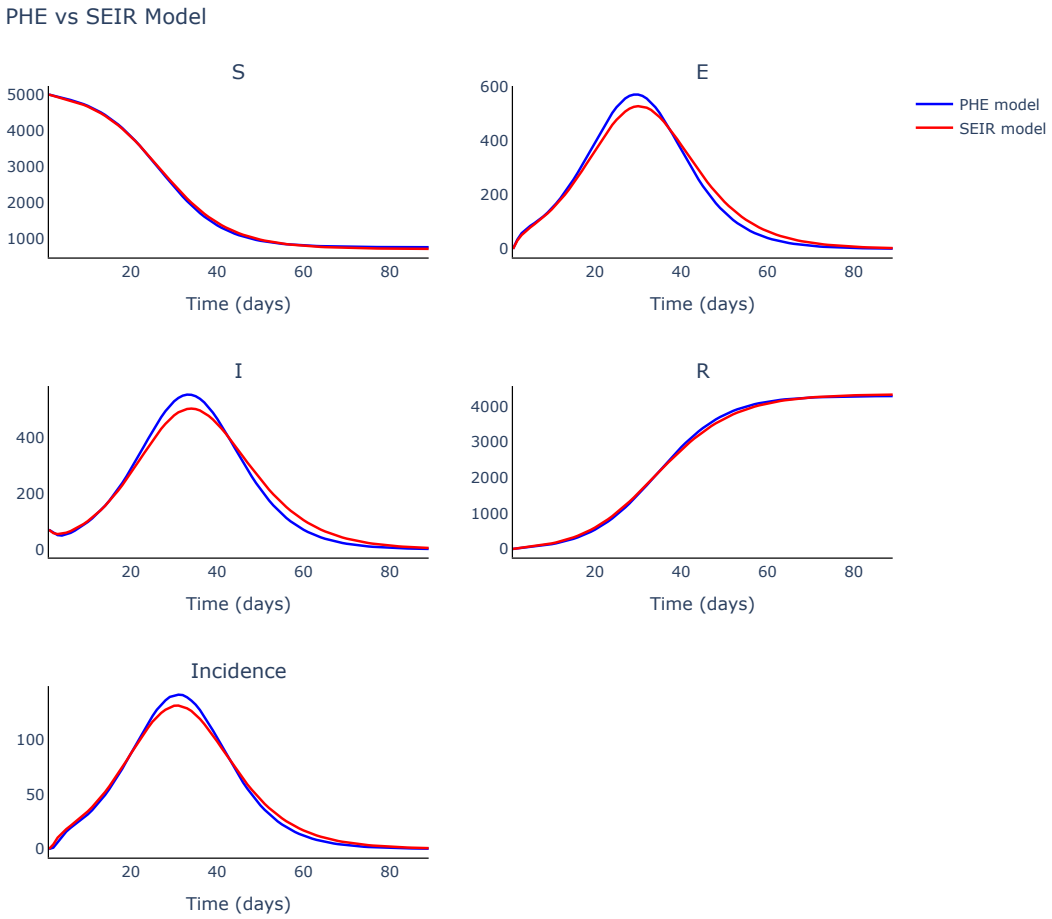


Figure D1: *Comparison of trajectories of compartment counts between an SEIR model (RED) – eq.* (4.1)*, and the simplified version of the PHE model (BLUE) – eq.* (D.1)*. Both models were run over* 90 *days.*

values of the PHE model so that the equivalent parameters of the SEIR are $\beta = 1.102 \times 10^{-4}$, $\kappa = 0.25$ and $\gamma = 0.25$ – hence, we set the free parameters of the PHE model as such: $R_0 = 2.2164$, $d_L = 4$, $d_I = 4$, the one-group one-region level of contacts at $C = 0.516$ (average of the UK POLYMOD contact matrix [Prem et al., 2021]) respectively. We simulate the population dynamics for both models over 90 days. The initial number of susceptibles is fixed at 5000 individuals, while the initial number of infections is set to be 70 individuals. These parameter values areComparing these trajectories, we observe that except for small initial perturbations in the exposed, infectious and incidence counts, which result from the difference in the distribution of the waiting time of individuals in the infectious and exposed compartments (Erlang-distributed for the PHE model versus exponential-distributed for the SEIR model), the temporal variations in the individual counts produced by the PHE model closely resemble those of the SEIR model for each population compartment.

Hence, under the conditions indicated above, the PHE model is equivalent to and can be substituted with the SEIR model for modelling the transmission dynamics of an ongoing COVID-19-like epidemic.

D2 The Roche model

We proceed now to identify under which parameter conditions the SEIR model is able to reproduce the original Roche model dynamics. We proceed similarly to the PHE model case in reducing the Roche model's ODE system:

$$
\begin{aligned}
\frac{dS(t)}{dt} &= -C\frac{S(t)}{N}(\beta_a(t)I^a(t) + \beta_{aa}(t)I^{aa}(t) + \beta_s(t)I^s(t) + \beta_{as}(t)I^{as}(t) + \beta_{aas}(t)I^{aas}(t) + \\
&+ \beta_{ss}(t)I^{ss}(t)) \\
\frac{dE(t)}{dt} &= C\frac{S(t)}{N}(\beta_a(t)I^a(t) + \beta_{aa}(t)I^{aa}(t) + \beta_s(t)I^s(t) + \beta_{as}(t)I^{as}(t) + \beta_{aas}(t)I^{aas}(t) + \\
&+ \beta_{ss}(t)I^{ss}(t)) - \gamma_e E(t) \\
\frac{dI^a(t)}{dt} &= (1-P_{ss})\gamma_e E(t) - \gamma_s I^a(t) \\
\frac{dI^{aa}(t)}{dt} &= P_a\gamma_s I^a(t) - \gamma_{ra} I^{aa}(t) \\
\frac{dI^s(t)}{dt} &= (1-P_a)\gamma_s I^a(t) - \gamma_q I^s(t) \\
\frac{dI^{as}(t)}{dt} &= P_{ss}\gamma_e E(t) - \gamma_s I^{as}(t) \\
\frac{dI^{aas}(t)}{dt} &= P_a\gamma_s I^{as}(t) - \gamma_{ra} I^{aas}(t) \\
\frac{dI^{ss}(t)}{dt} &= (1-P_a)\gamma_s I^{as}(t) - \gamma_q I^{ss}(t) \\
\frac{dI^q(t)}{dt} &= \gamma_q I^{ss}(t) + \gamma_q I^s(t) - \gamma_r I^q(t) \\
\frac{dR(t)}{dt} &= (1-P_d)\gamma_r I^q(t) \\
\frac{dR_a(t)}{dt} &= \gamma_{ra} I^{aas}(t) + \gamma_{ra} I^{aa}(t) \\
\frac{dD(t)}{dt} &= P_d\gamma_r I^q(t)
\end{aligned}
\tag{D.2}
$$

to the four equations of the SEIR model in eq. (4.1).

Unlike the SEIR model (which has only one compartment) and the PHE model (which has

two), the Roche model differentiates between seven different infectious categories: asymptomatic, symptomatic and presymptomatic, for both normal or super-spreader individuals, as well as a quarantined infectious group. Each category has their own designated compartment in the model structure as seen in eq. (C.6). Additionally, the Roche model introduces a compartment for the dead – not present in the SEIR model, and distinguishes between recovered individuals based on whether or not they manifested symptoms during their infectious period, as it assumes only symptomatic infections can go on to either recover or die; those who had asymptomatic infections will always recover.

As stated in the previous section, the parameters of the simple SEIR model are time-invariant; therefore, we impose a similar restriction on the transmission rates used in the Roche model, that is, we assume the levels of the non-pharmaceutical interventions (NPIs), as by OxCGRT (see section C2.1.1 for details), remain unchanged throughout the simulation (or the stringency index $\Theta = \text{const.}$ as described in section C2).

The flow of individuals into each of the infectious, recovered and dead compartments of the Roche model is controlled through the proportion parameters P_a, P_{ss} and P_d of asymptomatic, super-spreader transmission and death respectively. The force of infection parameter of the Roche model encompasses the effects from each of the seven infectious compartments, each governed by its own (different) transmission rates. However, these rates are dependent on the symptomatic transmission rate β_s and hence cannot individually be set to zero. We cannot keep both of the asymptomatic and symptomatic transmission mechanisms; we choose to retain only the asymptomatic transmission as according to the Roche model structure, this infection type are neither quarantined, nor are they able to die – similar to the SEIR model. We also turn off all super-spreader transmission. The proportionality parameters thus become $P_{ss} = P_d = 0$ and $P_a = 1$ and the Roche model ODE system is updated to:

$$
\begin{aligned}
\frac{dS(t)}{dt} &= -C\frac{S(t)}{N}(\beta_a I^a(t) + \beta_{aa} I^{aa}(t) + \beta_s I^s(t) + \beta_{as} I^{as}(t) + \beta_{aas} I^{aas}(t) + \beta_{ss} I^{ss}(t)) \\
\frac{dE(t)}{dt} &= C\frac{S(t)}{N}(\beta_a I^a(t) + \beta_{aa} I^{aa}(t) + \beta_s I^s(t) + \beta_{as} I^{as}(t) + \beta_{aas} I^{aas}(t) + \beta_{ss} I^{ss}(t)) \\
&\quad - \gamma_e E(t) \\
\frac{dI^a(t)}{dt} &= \gamma_e E(t) - \gamma_s I^a(t)
\end{aligned}
$$

$$
\begin{aligned}
\frac{dI^{aa}(t)}{dt} &= \gamma_s I^a(t) - \gamma_{ra} I^{aa}(t) \\
\frac{dI^{s}(t)}{dt} &= -\gamma_q I^s(t) \\
\frac{dI^{as}(t)}{dt} &= -\gamma_s I^{as}(t) \\
\frac{dI^{aas}(t)}{dt} &= \gamma_s I^{as}(t) - \gamma_{ra} I^{aas}(t) \\
\frac{dI^{ss}(t)}{dt} &= -\gamma_q I^{ss}(t) \\
\frac{dI^{q}(t)}{dt} &= \gamma_q I^{ss}(t) + \gamma_q I^s(t) - \gamma_r I^q(t) \\
\frac{dR(t)}{dt} &= \gamma_r I^q(t) \\
\frac{dR_a(t)}{dt} &= \gamma_{ra} I^{aas}(t) + \gamma_{ra} I^{aa}(t) \\
\frac{dD(t)}{dt} &= 0.
\end{aligned}
$$

The last equation in the current system of ODEs indicates that we have a constant number of dead. To ensure there are no individuals in the dead compartment at any time point (and similarly for all the other symptomatic recovered and symptomatic, quarantined and all super-spreader infections), we impose the following set of constrained initial conditions:

$$I^s(0) = I^{as}(0) = I^{aas}(0) = I^{ss}(0) = I^q(0) = R(0) = D(0) = 0.$$

These assumptions paired together with their corresponding ODE equations ensure that the compartments remain empty throughout the simulation, i.e. $I^s(t) = I^{as}(t) = I^{aas}(t) = I^{ss}(t) = I^q(t) = R(t) = D(t) = 0, \forall t$. Dropping the now redundant equations, the ODE system of the Roche model further simplifies to:

$$
\begin{aligned}
\frac{dS(t)}{dt} &= C(-\frac{\beta_a}{N} S(t) I^a(t) - \frac{\beta_{aa}}{N} S(t) I^{aa}(t)) \\
\frac{dE(t)}{dt} &= -\gamma_e E(t) + C(\frac{\beta_a}{N} S(t) I^a(t) + \frac{\beta_{aa}}{N} S(t) I^{aa}(t)) \\
\frac{dI^a(t)}{dt} &= \gamma_e E(t) - \gamma_s I^a(t) \\
\frac{dI^{aa}(t)}{dt} &= \gamma_s I^a(t) - \gamma_{ra} I^{aa}(t) \\
\frac{dR_a(t)}{dt} &= \gamma_{ra} I^{aa}(t).
\end{aligned}
$$

We still retain two infectious compartments: one for presymptomatic (I^a) and one for asymptomatic (I^{aa}) infections. The Roche model assumes no differences in the virus transmission between asymptomatic and presymptomatic cases, that is $\beta_a = \beta_{aa}$. To collapse the current infectious structure into a single compartment, we add together the infected compartments according to the identity:

$$I(t) = I^a(t) + I^{aa}(t).$$

Dropping the subscript for the R_a compartment, i.e. $R(t) = R_a(t)$, the compressed Roche model becomes:

$$\begin{aligned}
\frac{dS(t)}{dt} &= -C\frac{\beta_a}{N}S(t)(I^a(t) + I^{aa}(t)) = -C\frac{\beta_a}{N}S(t)I(t) = -C\frac{\beta_s}{2N}S(t)I(t) \\
\frac{dE(t)}{dt} &= -\gamma_e E(t) + C\frac{\beta_s}{2N}S(t)I(t) \\
\frac{dI(t)}{dt} &= \gamma_e E(t) - \gamma_{ra} I^{aa}(t) \\
\frac{dR(t)}{dt} &= \gamma_{ra} I^{aa}(t).
\end{aligned}$$

We now assess the transmission parameters into and out of the I and R compartments: the incubation rate $\gamma_e = \frac{1}{k}$, the symptomatic infection rate $\gamma_s = \frac{1}{k_s}$, the quarantined infection rate $\gamma_q = \frac{1}{k_q}$, the symptomatic recovery rate $\gamma_r = \frac{1}{k_r}$ and the asymptomatic recovery rate $\gamma_{ra} = \frac{1}{k_{ri}}$. The rate of transmission between the compartments is defined as the inverse of the average time spent in the compartment. Therefore, in the Roche model individuals wait in the I^a and I^{aa} compartments an average time of k_s and k_{ri} respectively – hence they spend a total average of $k_s + k_{ri}$ as infectious.

A final necessary assumption we need to make about the Roche model in order to simplify it to the SEIR model is that the average time people are presymptomatic is very small, i.e. $k_s \ll 1$. In this case, the total number of infections observed at time t is roughly equal to all asymptomatic infections at that time point, or $I(t) \approx I^{aa}(t), \forall t$. Therefore, the Roche model reduces to an SEIR with the model parameters given by:

$$\begin{aligned}
\beta &= C\frac{\beta_s}{2N} = C\frac{\beta_{\text{max}} - (\beta_{\text{max}} - \beta_{\text{min}})\frac{\Theta^\gamma}{\Theta^\gamma + \Theta_{50}^\gamma}}{2N} \\
\kappa &= \gamma_e = \frac{1}{k}
\end{aligned}$$

$$\gamma = \frac{1}{k_s + k_{ri}} \cong \frac{1}{k_{ri}}.$$

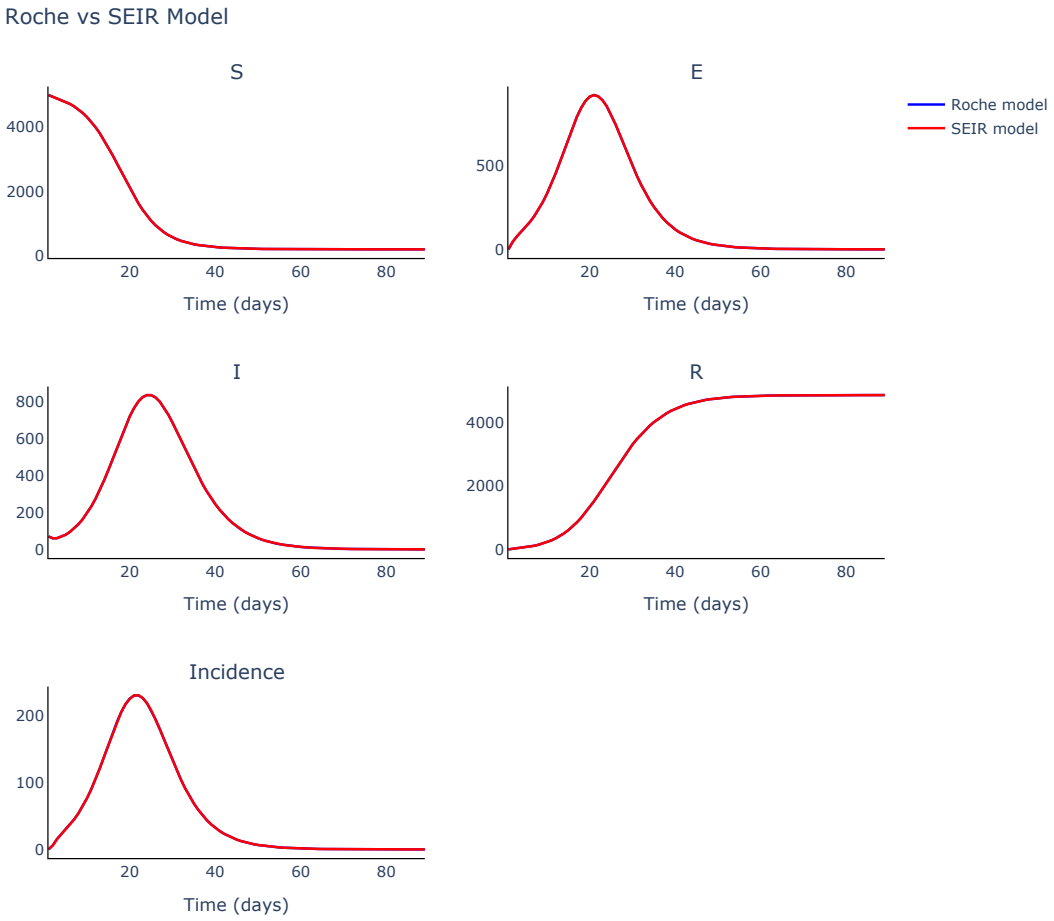


Figure D2: *Comparison of trajectories of compartment counts between an SEIR model (RED) – eq.* (4.1)*, and the simplified version of the Roche model (BLUE) – eq.* (D.2)*. Both models were run over* 90 *days. The trajectories described by the two models are superimposed.*

In Figure D2 we plot the trajectories of the compartment counts for an example instantiation of the Roche model (the blue curve) and an SEIR model with matched parameters (in red), similarly to the PHE model case. We choose the values of the Roche model so that the equivalent parameters of the SEIR are $\beta = 1.102 \times 10^{-4}$, $\kappa = 0.25$ and $\gamma = 0.25$ – hence, we set the free parameters of the PHE model as such: $\beta_{\text{max}} = 2.176$, $k = 4$, $k_s = 0.0001$, $k_q = 1$, $k_r = 3$, $k_{ri} = 4$ and the one-group one-region level of contacts at $C = 0.516$ respectively – all other parameter values are inconsequential for driving the dynamics of the Roche model in this scenario and are set at the same values as in section 4.3.5.2. We simulate the population dynamics for both models over 90 days. The initial number of susceptibles

is fixed at 5000 individuals, while the initial number of infections is set to be 70 individuals. We observe that the trajectories of the population counts for all the compartments described by the two models line up perfectly throughout the simulation. Therefore, we conclude that the Roche model and SEIR model can be used interchangeably when modelling a COVID-19 like epidemic when the model parameters are chosen as described in this section.

D3 The Warwick-Household model

We aim to collapse the Warwick-Household model into a simpler SEIR version. Similar to the PHE and Roche models, we look for specific parameter combinations such that the following ODE system:

$$
\begin{aligned}
\frac{dS(t)}{dt} =& -\sigma\Big(C^N\big(\frac{S(t)}{N}I^F(t)+\frac{S(t)}{N}I^{SD}(t)+\frac{S(t)}{N}I^{SU}(t)+\tau(\frac{S(t)}{N}A^F(t)+\\
&+\frac{S(t)}{N}A^S(t))\big)+C^H(\frac{S(t)}{N}I^F(t)+\frac{S(t)}{N}A^F(t)+\frac{S(t)}{N}I^{QF}(t))\Big)\\
\frac{dE^{1,F}(t)}{dt} =& \sigma C^N\big(\frac{S(t)}{N}I^F(t)+\frac{S(t)}{N}I^{SD}(t)+\frac{S(t)}{N}I^{SU}(t)+\tau(\frac{S(t)}{N}A^F(t)+\\
&+\frac{S(t)}{N}A^S)\big)-3\epsilon E^{1,F}(t)\\
\frac{dE^{1,SD}(t)}{dt} =& \sigma C^H\frac{S(t)}{N}I^F(t)-3\epsilon E^{1,SD}(t)\\
\frac{dE^{1,SU}(t)}{dt} =& \sigma C^H\frac{S(t)}{N}A^F(t)-3\epsilon E^{1,SU}(t)\\
\frac{dE^{1,Q}(t)}{dt} =& \sigma C^H S(t)I^{QF}(t)-3\epsilon E^{1,Q}(t)\\
\frac{dE^{2,F}(t)}{dt} =& 3\epsilon E^{1,F}(t)-3\epsilon E^{2,F}(t)\\
\frac{dE^{2,SD}(t)}{dt} =& 3\epsilon E^{1,SD}(t)-3\epsilon E^{2,SD}(t)\\
\frac{dE^{2,SU}(t)}{dt} =& 3\epsilon E^{1,SU}(t)-3\epsilon E^{2,SU}(t)\\
\frac{dE^{2,Q}(t)}{dt} =& 3\epsilon E^{1,Q}(t)-3\epsilon E^{2,Q}(t)\\
\frac{dE^{3,F}(t)}{dt} =& 3\epsilon E^{2,F}(t)-3\epsilon E^{3,F}(t)\\
\frac{dE^{3,SD}(t)}{dt} =& 3\epsilon E^{2,SD}(t)-3\epsilon E^{3,SD}(t)\\
\frac{dE^{3,SU}(t)}{dt} =& 3\epsilon E^{2,SU}(t)-3\epsilon E^{3,SU}(t)
\end{aligned}
$$

$$
\begin{aligned}
\frac{dE^{1,Q}(t)}{dt} &= 3\epsilon E^{2,Q}(t) - 3\epsilon E^{3,Q}(t) \\
\frac{dI^{F}(t)}{dt} &= 3(1-H)\epsilon d E^{3,F}(t) - \gamma I^{F}(t) \\
\frac{dI^{SD}(t)}{dt} &= 3\epsilon d E^{3,SD}(t) - \gamma I^{SD}(t) \\
\frac{dI^{SU}(t)}{dt} &= 3(1-H)\epsilon d E^{3,SU}(t) - \gamma I^{SU}(t) \\
\frac{dI^{QF}(t)}{dt} &= 3H\epsilon d E^{3,F}(t) - \gamma I^{QF}(t) \\
\frac{dI^{QS}(t)}{dt} &= 3H\epsilon d E^{3,SU}(t) + 3\epsilon d E^{3,Q}(t) - \gamma I^{QS}(t) \\
\frac{dA^{F}(t)}{dt} &= 3\epsilon(1-d) E^{3,F}(t) - \gamma A^{F}(t) \\
\frac{dA^{S}(t)}{dt} &= 3\epsilon(1-d)(E^{3,SD}(t) + E^{3,SU}(t)) - \gamma A^{S}(t) \\
\frac{dA^{Q}(t)}{dt} &= 3\epsilon(1-d) E^{3,Q}(t) - \gamma A^{Q}(t) \\
\frac{dR(t)}{dt} &= \gamma(I^{F}(t) + I^{QF}(t) + A^{F}(t) + I^{SD}(t) + A^{S}(t) + I^{SU}(t) + I^{QS}(t) + A^{Q}(t))
\end{aligned}
\tag{D.3}
$$

simplifies to the following set of four equations (4.1) corresponding to the basic SEIR model.

We proceed similarly to section D1 and fix the contact matrices associated with the Warwick-Household model, as all the parameters of the SEIR model are time-invariant.

Just like the Roche model, the Warwick-Household model differentiates between asymptomatic and symptomatic infections. However, it also considers within-household dynamics with separate compartments for those infections first to occur within the household and subsequent asymptomatic, subsequent symptomatic or quarantined infections – contrasting with the SEIR model which only has one infection compartment. Additionally, the Warwick-Household model has three compartments for the exposed individuals, similar to the PHE model.

The flow of individuals into each of the infectious compartments is regulated through the parameters H and d – the probabilities of household quarantine and symptom development respectively. We cannot keep both the asymptomatic and symptomatic infections as that distinction does not appear in the SEIR model. Hence, we turn off the asymptomatic transmission by choosing all the parameters related to asymptomatic or quarantined symptomatic

infections to be equal to 0 – that is $H = 0$ and $1 - d = 0$, or $d = 1$.

The Warwick-Household model ODE system thus becomes:

$$
\begin{aligned}
\frac{dS(t)}{dt} =& -\sigma\Big(C^N\big(\frac{S(t)}{N}I^F(t) + \frac{S(t)}{N}I^{SD}(t) + \frac{S(t)}{N}I^{SU}(t) + \tau(\frac{S(t)}{N}A^F(t) + \\
&+ \frac{S(t)}{N}A^S(t))\big) + C^H(\frac{S(t)}{N}I^F(t) + \frac{S(t)}{N}A^F(t) + \frac{S(t)}{N}I^{QF}(t))\Big) \\
\frac{dE^{1,F}(t)}{dt} =& \sigma C^N\big(\frac{S(t)}{N}I^F(t) + \frac{S(t)}{N}I^{SD}(t) + \frac{S(t)}{N}I^{SU}(t) + \tau(\frac{S(t)}{N}A^F(t) + \\
&+ \frac{S(t)}{N}A^S(t))\big) - 3\epsilon E^{1,F}(t) \\
\frac{dE^{1,SD}(t)}{dt} =& \sigma C^H\frac{S(t)}{N}I^F(t) - 3\epsilon E^{1,SD}(t) \\
\frac{dE^{1,SU}(t)}{dt} =& \sigma C^H\frac{S(t)}{N}A^F(t) - 3\epsilon E^{1,SU}(t) \\
\frac{dE^{1,Q}(t)}{dt} =& \sigma C^H\frac{S(t)}{N}I^{QF}(t) - 3\epsilon E^{1,Q}(t) \\
\frac{dE^{2,F}(t)}{dt} =& 3\epsilon E^{1,F}(t) - 3\epsilon E^{2,F}(t) \\
\frac{dE^{2,SD}(t)}{dt} =& 3\epsilon E^{1,SD}(t) - 3\epsilon E^{2,SD}(t) \\
\frac{dE^{2,SU}(t)}{dt} =& 3\epsilon E^{1,SU}(t) - 3\epsilon E^{2,SU}(t) \\
\frac{dE^{2,Q}(t)}{dt} =& 3\epsilon E^{1,Q}(t) - 3\epsilon E^{2,Q}(t) \\
\frac{dE^{3,F}(t)}{dt} =& 3\epsilon E^{2,F}(t) - 3\epsilon E^{3,F}(t) \\
\frac{dE^{3,SD}(t)}{dt} =& 3\epsilon E^{2,SD}(t) - 3\epsilon E^{3,SD}(t) \\
\frac{dE^{3,SU}(t)}{dt} =& 3\epsilon E^{2,SU}(t) - 3\epsilon E^{3,SU}(t) \\
\frac{dE^{1,Q}(t)}{dt} =& 3\epsilon E^{2,Q}(t) - 3\epsilon E^{3,Q}(t) \\
\frac{dI^F(t)}{dt} =& 3\epsilon E^{3,F}(t) - \gamma I^F(t) \\
\frac{dI^{SD}(t)}{dt} =& 3\epsilon E^{3,SD}(t) - \gamma I^{SD}(t) \\
\frac{dI^{SU}(t)}{dt} =& 3\epsilon E^{3,SU}(t) - \gamma I^{SU}(t) \\
\frac{dI^{QF}(t)}{dt} =& -\gamma I^{QF}(t) \\
\frac{dI^{QS}(t)}{dt} =& 3\epsilon E^{3,Q}(t) - \gamma I^{QS}(t)
\end{aligned}
$$

$$
\begin{aligned}
\frac{dA^F(t)}{dt} &= -\gamma A^F(t) \\
\frac{dA^S(t)}{dt} &= -\gamma A^S(t) \\
\frac{dA^Q(t)}{dt} &= -\gamma A^Q(t) \\
\frac{dR(t)}{dt} &= \gamma(I^F(t) + I^{QF}(t) + A^F(t) + I^{SD}(t) + A^S(t) + I^{SU}(t) + I^{QS}(t) + A^Q(t)).
\end{aligned}
$$

To turn off completely the asymptomatic, quarantined exposed and infections compartments as we have done in section D2, we impose the following set of constrained initial conditions:

$$
\begin{aligned}
E^{k,SU}(0) = E^{k,Q}(0) = 0, \forall k = 1, \dots 3 \text{ and} \\
A^F(0) = I^{QF}(0) = A^S(0) = I^{QS}(0) = A^Q(0) = 0.
\end{aligned}
$$

These assumption ensure that these model compartments remain empty for the entire duration of the simulation, that is $E^{k,SU}(t) = E^{k,Q}(t) = A^F(t) = I^{QF}(t) = A^S(t) = I^{QS}(t) = A^Q(t) = 0, \forall t, \forall k = 1, \dots 3$. Therefore, we can drop the equations governing these compartments altogether, slimming down the ODE system to:

$$
\begin{aligned}
\frac{dS(t)}{dt} &= -\sigma\Big(C^N\big(\frac{S(t)}{N}I^F(t) + \frac{S(t)}{N}I^{SD}(t) + \frac{S(t)}{N}I^{SU}(t)\big) + C^H\frac{S(t)}{N}I^F(t)\Big) \\
\frac{dE^{1,F}(t)}{dt} &= \sigma C^N\big(\frac{S(t)}{N}I^F(t) + \frac{S(t)}{N}I^{SD}(t) + \frac{S(t)}{N}I^{SU}(t)\big) - 3\epsilon E^{1,F}(t) \\
\frac{dE^{1,SD}(t)}{dt} &= \sigma C^H\frac{S(t)}{N}I^F(t) - 3\epsilon E^{1,SD}(t) \\
\frac{dE^{2,F}(t)}{dt} &= 3\epsilon E^{1,F}(t) - 3\epsilon E^{2,F}(t) \\
\frac{dE^{2,SD}(t)}{dt} &= 3\epsilon E^{1,SD}(t) - 3\epsilon E^{2,SD}(t) \\
\frac{dE^{3,F}(t)}{dt} &= 3\epsilon E^{2,F}(t) - 3\epsilon E^{3,F}(t) \\
\frac{dE^{3,SD}(t)}{dt} &= 3\epsilon E^{2,SD}(t) - 3\epsilon E^{3,SD}(t) \\
\frac{dI^F(t)}{dt} &= 3\epsilon E^{3,F}(t) - \gamma I^F(t) \\
\frac{dI^{SD}(t)}{dt} &= 3\epsilon E^{3,SD}(t) - \gamma I^{SD}(t) \\
\frac{dI^{SU}(t)}{dt} &= -\gamma I^{SU}(t) \\
\frac{dR(t)}{dt} &= \gamma(I^F(t) + I^{SD}(t) + I^{SU}(t)).
\end{aligned}
$$

Now, according to this new set of differential equations, the subsequent symptomatic infection compartment caused by an asymptomatic infection I^{SU} has only individuals leaving the compartment. Hence, to turn off this compartment completely, i.e. $I^{SU}(t) = 0, \forall t$ it is enough to set the initial number of individuals in this compartment to 0, or $I^{SU}(0) = 0$.

We are still left with two types of exposed and two symptomatic infected compartments: those individuals first exposed or infected in the household ($E^{k,F}$ and I^F respectively), and those individuals subsequently exposed and symptomatically infected due to a symptomatic infection ($E^{k,SD}$ and I^{SD}). As there are no within-household dynamics embedded in the SEIR model structure, we also assume no household contacts ($C^H = 0$) for the Warwick-Household model; hence all contacts that individuals make are outside the household, or $C^N = C$. Additionally, to insure that no individuals become subsequently infected throughout the simulation, we impose $E^{k,SD}(0) = I^{SD}(0) = 0$ which is a necessary and sufficient condition for $E^{k,SD}(0) = I^{SD}(t) = 0, \forall t, \forall k = 1, \dots 3$. Therefore, the ODE systems associated with the Warwick-Household model becomes:

$$\begin{aligned}
\frac{dS(t)}{dt} &= -\sigma C \frac{S(t)}{N} I(t) \\
\frac{dE^1(t)}{dt} &= \sigma C \frac{S(t)}{N} I(t) - 3\epsilon E^1(t) \\
\frac{dE^2(t)}{dt} &= 3\epsilon E^1(t) - 3\epsilon E^2(t) \\
\frac{dE^3(t)}{dt} &= 3\epsilon E^2(t) - 3\epsilon E^3(t) \\
\frac{dI(t)}{dt} &= 3\epsilon E^3(t) - \gamma I(t) \\
\frac{dR(t)}{dt} &= \gamma I(t),
\end{aligned}$$

once we drop the superscript for the first infected, i.e. $I^F(t) = I(t), \forall t$.

In this form, the Warwick-Household model is still left with three exposed compartments, compared to only one in the SEIR model, which allows for an Erlang-distributed waiting time in the exposed state. The total number of exposed individuals in the Warwick-household model is given by summing over these three exposed compartments: $E(t) = E^1(t) + E^2(t) + E^3(t)$. Therefore, we will collapse this model along this infection state. Because the rate of transmission between the compartments is defined as the inverse of the average time spent in the compartment and that individuals spend on average $\frac{1}{3\epsilon}$ time in

each of the E^1, E^2 and E^3 in the Warwick-Household model, we arrive at the equivalent incubation rate of $\frac{1}{\epsilon}$ for the SEIR model counterpart.

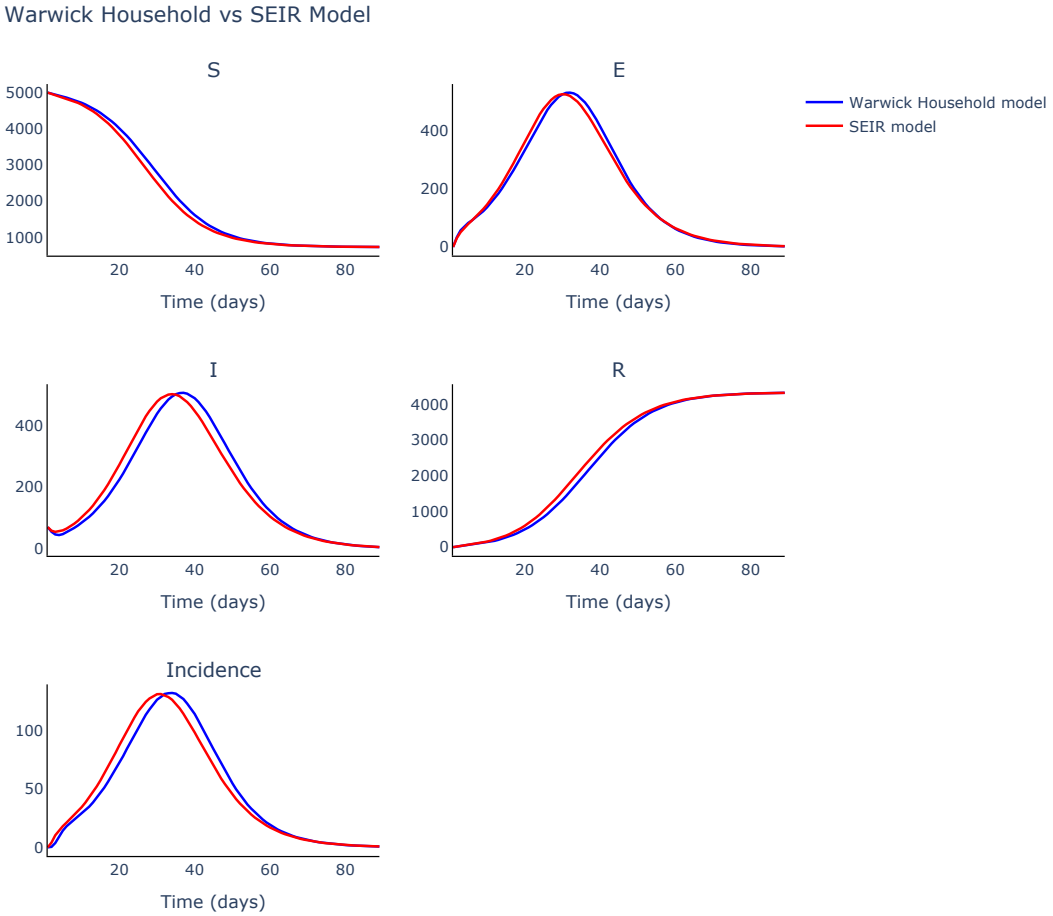


Figure D3: *Comparison of trajectories of compartment counts between an SEIR model (RED) – eq.* (4.1)*, and the simplified version of the Warwick-Household model (BLUE) – eq.* (D.3)*. Both models were run over* 90 *days.*

Therefore, under the parameter restrictions introduced above, the Warwick-Household model becomes equivalent to a SEIR model with its parameters given by $\beta = C\frac{\sigma}{N}$, $\kappa = \epsilon$ and $\gamma = \gamma$. In Figure D3 we plot the trajectories of the number of individuals present in different compartments for both the Warwick-Household model and its SEIR counterpart. We choose the values of the Roche model so that the equivalent parameters of the SEIR are $\beta = 1.102 \times 10^{-4}$, $\kappa = 0.25$ and $\gamma = 0.25$ – hence, we set the free parameters of the PHE model as such: $\sigma = 1.088$, $\epsilon = 0.25$, $\gamma = 0.25$ and the one-group one-region level of contacts at $C = 0.516$ respectively. We simulate the population dynamics for both models over 90 days. The initial number of susceptibles is fixed at 5000 individuals, while the initial number of infections is set to be 70 individuals. Similar to the PHE case discussed in sec-

tion D1, we observe that apart from small initial differences in the individual counts of the exposed, infected and incidence compartments, the rendered trajectories by the two models match up throughout the simulation. Hence, we can substitute the Warwick-model for its simpler SEIR model counterpart when the parameters of the former are chosen as described in this section.

D4 Parameter relationships across the models

In this section we discuss how the parameters of the three models for COVID-19 introduced in Appendix C relate to one another. We begin by analysing the connection between the parameters of these more complex models for epidemic transmission and the parameters governing the SEIR model, which we have identified in the previous sections of this appendix and summarised in Table D1.

Parameter Name	SEIR model	PHE model	Roche model	Warwick-Household model
Transmission rate	β	$\frac{R_0}{S_0 d_I}$	$C\frac{\beta_s}{2N}$	$C\frac{\sigma}{N}$
Incubation rate	κ	$\frac{1}{d_L}$	$\frac{1}{k}$	ϵ
Recovery rate	γ	$\frac{1}{d_I}$	$\frac{1}{k_{ri}}$	γ

Table D1: *The compartment progression parameters of the SEIR model in terms of the parameters of the PHE, the Roche and the Warwick-Household models.*

While we have considered only one age group and one model region in order to equate the SEIR framework to either of the PHE, Roche and Warwick-Household models, these equivalence relationships can still offer important insights into how the models resemble and contrast with one another – both in terms of disease assumptions, and notation.

For example, all three models use a matrix of contacts. However, the Warwick-Household model differentiates between within-household contacts (C_t^H) and outside-of-household ones (C_t^N), which when summed recover the matrix of total contacts used by both the PHE and the Roche model. Additionally, the total population size uses two distinct notations: S_0 for the PHE model and N for the Roche and Warwick-Household models. For all three models the initial number of susceptibles is equal to the total population size, while all initial infected individuals are present in the first symptomatic infectious compartment for each of the models: I^1 for the PHE model, I^s for the Roche and I^F for the Warwick-Household

one. Below we summarise the relationships between the population-specific parameters for the three models.

Parameter Name	Value	PHE model	Roche model	Warwick-Household model
Total population size	5000	S_0	N	N
Total contact matrix	0.516	$\tilde{C}_t$	C_t	$C_t^N + C_t^H$

Table D2: *Equivalence relation between the population and contact-specific parameters of the PHE, the Roche and the Warwick-Household models. We also include for reference the values of the model parameters used to produce Figures D1, D2 and D3.*

Additionally, from Appendix C, we remember that non-pharmaceutical interventions are implemented in each of the three models in a different manner.

- In the PHE model, the effect of NPIs acts as a multiplier of the contact matrix $\tilde{C}_t$.
- The Roche model assumes a constant contact matrix C, with the impact of interventions being applied directly onto the transmission parameters (via a multiplier in the definition of the β_s which is used to determine the transmission rate values for all other infection types).
- Finally, in the Warwick-Household model, similar to the PHE model, the effect on interventions is modelled through the intervention strength parameter ϕ. These parameters weight the baseline contact matrices within the household (C_t^H), at school (C_t^S), in the workplace (C_t^W) and in all other locations respectively (C_t^O) – when no interventions are in place, against the full lock-down reduced contact matrices in each of these settings (reduced by q_H, q_S, q_W and q_O respectively).

Parameter Name	PHE model	Roche model	Warwick-Household model
NPIs effect	Google Mobility Data multiplier	$1-(1-\frac{\beta_{\min}}{\beta_{\max}})\frac{\Theta^\gamma}{\Theta^\gamma+\Theta_{50}^\gamma}$	a combination of ϕ, q_H, q_S, q_W and q_O

Table D3: *Equivalence relation between the non-pharmaceutical intervention effects of the PHE, the Roche and the Warwick-Household models.*

For the progression of individuals across their different serological stages, we use the equivalences to the SEIR model summarised in Tables D1 and D3 to get the susceptibility vector and the incubation and infection periods. Additionally, we are able to retrieve a form for

the basic reproduction number for both the Roche and Warwick-Household model for a one-region, one-age group modelled population, as seen in the Table D4:

Parameter Name	Value	PHE model	Roche model	Warwick-Household model
Susceptibility vector	1.088	$\frac{R_0}{\tilde{C}_0 d_I}$	$\frac{\beta_{\max}}{2}$	σ
Incubation period	4 (days)	d_L	k	$1/\epsilon$
Infection period	4 (days)	d_I	k_{ri}	$1/\gamma$
Basic reproduction number	2.2164	R_0	$C\beta_{\max}\frac{k_{ri}}{2}$	$C\frac{\sigma}{\gamma}$

Table D4: *Equivalence relation between the transmission parameters of the PHE, the Roche and the Warwick-Household models. We also include for reference the values of the model parameters used to produce Figures D1, D2 and D3.*

Unlike the PHE model which does not include asymptomatic transmission at all, the Roche and the Warwick-Household models have very different mechanisms for monitoring the different proportions of infections: in the Roche model, only symptomatic infected individuals can become quarantined infectious individuals and proceed directly to the I^q compartment after $\frac{1}{\gamma_q}$ days; meanwhile, in the Warwick-Household model, both symptomatic and asymptomatic individuals can be quarantined with a given probability H, i.e. there are unconstrained symptomatic infections. Additionally, in the Roche model, super-spreader individuals can transmit the infection at a higher rate compared to their normal infection counterparts – unlike the Warwick-Household model where there is no such enhancement. However, they both assume that asymptomatic infected individuals have a smaller transmission effect than their symptomatic counterparts. We summarise these observations in Table D5.

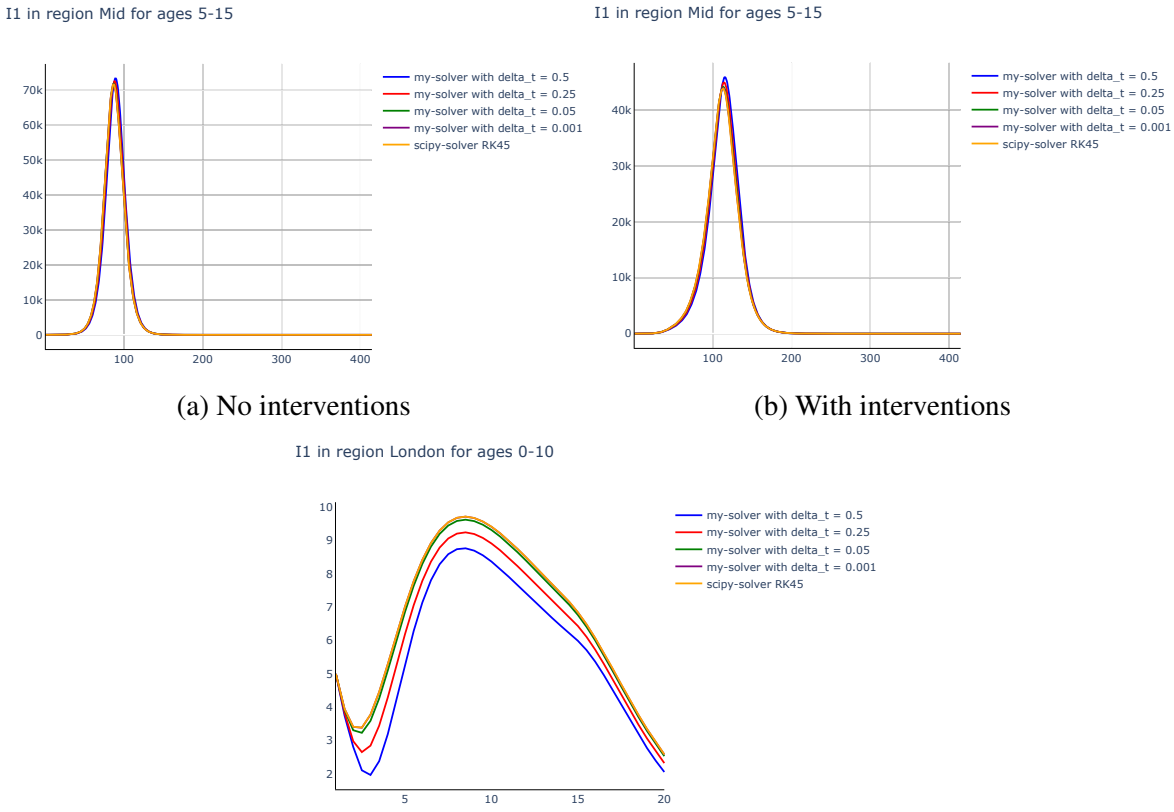
Parameter Name	Roche model	Warwick-Household model
Decrease in asymptomatic transmission	$\frac{1}{2}$	τ
Increase in super-spreader transmission	$1 + b_{ss}$	—
Probability of symptom detectability	$1 - P_a$	d
Probability of super-spreader infection	P_{ss}	—
Quarantine probability	1	H

Table D5: *Equivalence relation between the infection type-specific parameters of the PHE, the Roche and the Warwick-Household models.*

E | Appendix III Chapter 4

E1 Changing the time-step of the Euler method

The PHE model states a time-step δt of 1/2 a day for the discretised solver method. Choosing smaller step sizes results in more steps in the Euler method for computing compartment values; in this situation the solver method yields more accurate results, at the expense of higher computational times. In Figure E1, we observe the implementation of the PHE model in three different scenarios: (a) toy problem, (b) no interventions, (c) time-varying contacts.



(a) No interventions

(b) With interventions

(c) Toy problem

Figure E1: *Number of infectious individuals in the I^1 in given regions and age group, for multiple solver and scenarios: (a) no interventions, (b) time-varying contacts, (c) toy problem.*

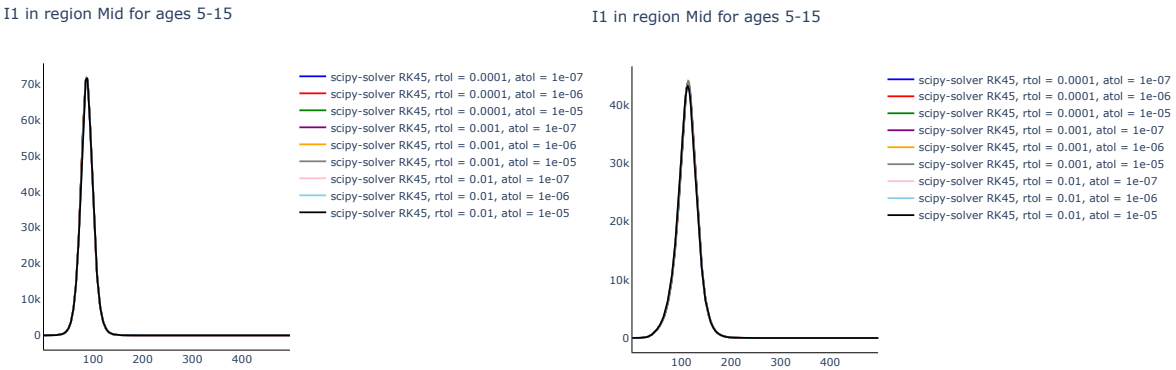
We can definitely see an effect in the toy problem: the smaller the δt time-step, the more the curve of the infectious obtained using the discretised solver resembles the result we obtain using the more accurate *scipy_solver*. When applying the same range of different solvers

to the realistic scenarios from Figures E.1(b) and E.1(c), the differences in the trajectories differ too little to seem significant.

E2 Changing the tolerance of the Scipy solver

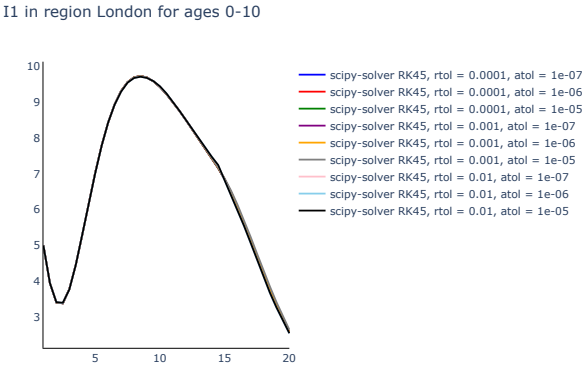
The *scipy_solver* method for solving the ODE system of the PHE model also has two key parameters through which we can fine-tune the accuracy of the predicted trajectory:

- the **relative tolerance** $rtol$: which is defined as the maximum relative error admissible in a single step of the solver – and thus controls the number of significant digits in the solution; and
- the **absolute tolerance** $atol$: which is defined as the maximum absolute error allowed in a solution of the ODE solver.



(a) No interventions

(b) With interventions



(c) Toy problem

Figure E2: *Number of infectious individuals in the I^1 in given regions and age group, for multiple relative and absolute tolerance values of the Scipy RK5(4) solver, for three different scenarios: (a) no interventions, (b) time-varying contacts, (c) toy problem.*

In the default implementation of the Scipy solver, $rtol = 1 \times 10^{-3}$ and $atol = 1 \times 10^{-6}$. In

Figure E2, we illustrate the effects of changing the values of the $rtol$ and $atol$ by one order of magnitude on either side of their default values; we observe that varying the relative and tolerance levels of the Scipy ODE solver makes no difference in the predicted trajectories of the number of daily infections. Therefore, in our analyses, we will maintain the values for $rtol$ and $atol$ to their default levels.

F | Appendix IV Chapter 4

F1 Methods for loading the data

Prior to running either of the inference methods described in Chapter 4, we run the following methods for loading data:

- For **the PHE model**:
  - *Read_deaths_data*: the daily number of deaths and when they occur, the probability distribution of dying of an individual in a given number of days after infection f and the age-specific infection fatality ratio (IFR) p (see Eq. (C.4));
  - *Read_serology_data*: the weekly number of tests conducted, the number of positive results and when they occur and the sensitivity and specificity of the test;
  - *Read_model_data* method: regional age-structured number of susceptibles and infectives.
- For **the Roche model**:
  - *Read_deaths_data*: the daily number of deaths and when they occur;
  - *Read_serology_data*: the weekly number of tests conducted, the number of positive results and when they occur, and the sensitivity and specificity of the test;
  - *Read_npis_data*: the maximum levels the non-pharmaceutical interventions can reach according to the OxCGRT study [Hale et al., 2021], a list of whether or not these interventions where targeted, a list of whether the non-pharmaceutical interventions are targeted locally or globally and when these changes are made, as well as the region-specific levels the non-pharmaceutical interventions and when they come into effect – for more details consult section C2.1.1;
  - *Read_model_data* method: regional age-structured number of susceptibles and infectives.
- For **the Warwick-Household model**:
  - *Read_deaths_data*: the daily number of deaths and when they occur;

- *Read_serology_data*: the weekly number of tests conducted, the number of positive results and when they occur, and the sensitivity and specificity of the test;
- *Read_model_data* method: regional age-structured number of susceptibles and infectives.

F2 Numerical and Computational Considerations

In this Appendix we summarise some of the numerical considerations used in running our analyses for the compartmental ODE models in Chapter 4 – for details about the assumptions of each of these models, see Appendix C.

For each of the three models, we use the *SciPy* [Virtanen et al., 2020] standard ODE solver (*solve_ivp*) from *SciPy* [Virtanen et al., 2020], with the explicit Runge-Kutta method of order 5(4) [Dormand and Prince, 1980] as the default solver. The relative and absolute tolerance parameters of the solver are kept at their default values of 1×10^{-3} and 1×10^{-6} respectively – see Appendix E2 for more details about the impact of changing the tolerance of the RK5(4) solver.

As stated in sections 4.3.7.1, 4.3.7.2 and 4.3.7.3, we use for each of the three models the CMA-ES algorithm [Hansen et al., 2003; Hansen, 2006] for running the optimisation procedure in our parameter inference; to further refine those estimates, we run three chains of the Haario-Bardenet sampling algorithm [Haario et al., 2001] of 60000 samples each, with each chain starting from the optimised parameter values. We cast-off half of the produced samples as burn-in, and further thinned to 9000 samples after concatenating the results from the three chains in order to reduce the autocorrelation between the remaining samples.

In Figures F1, F2 and F3, we present diagnostic plots for the inference procedure in the case of the PHE model. The traces described by the three Markov chains and the histograms of the resulting parameter samples (Figure F1) overlap over the three chains; this indicates that the MCMC algorithm properly converged and that the mean and 95% quantiles of the generated samples correctly estimate the uncertainty about the true parameter vales. In addition, from Figures F2 and F3, we observe that the samples produced by each of the three Markov chains do not exhibit high levels of autocorrelation, which implies that the MCMC algorithm does a good job of fully exploring the parameter surface and has low

rejection probability of proposed next steps.

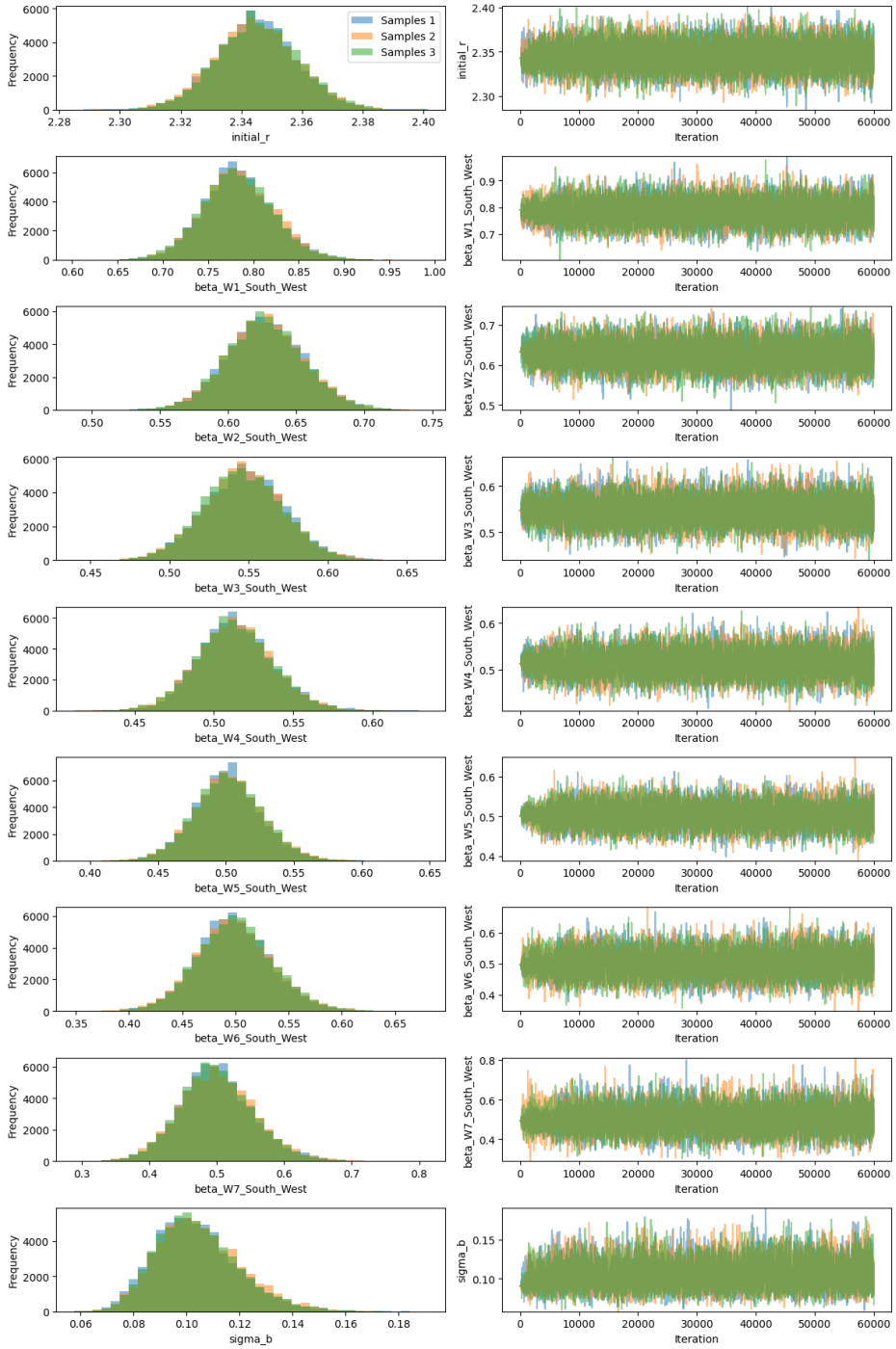


Figure F1: *Main diagnostic plot assessing the quality of convergence of the MCMC algorithm used to infer the parameters of the PHE model. The panels on the left show the histograms of the samples from the three chains for each of the inferred parameters; in the panels on the right we plot the traces described by each of the three MCMC chains for each of the estimated models.*

We finish with providing estimates of the expected time it takes to run each of the key analyses (forward simulation, optimisation and inference) for each of the three models discussed in Appendix C. Comparing the values recorded in Table F1, we conclude that it takes us roughly the same amount of time to run a forward simulation (or preform parameter inference) for each of the three models.

Model Name	Type of Analysis		
	Forward Simulation	Optimisation	Inference
PHE	0.1 seconds	8 minutes	134 minutes
Roche	0.1 seconds	1 minutes	100 minutes
Warwick-Household	0.1 seconds	8.5 minutes	216 minutes

Table F1: *Average duration of the three key analyses performed in Chapter 4 For the inference procedure, the times do not include the duration of the initial optimisation step.*

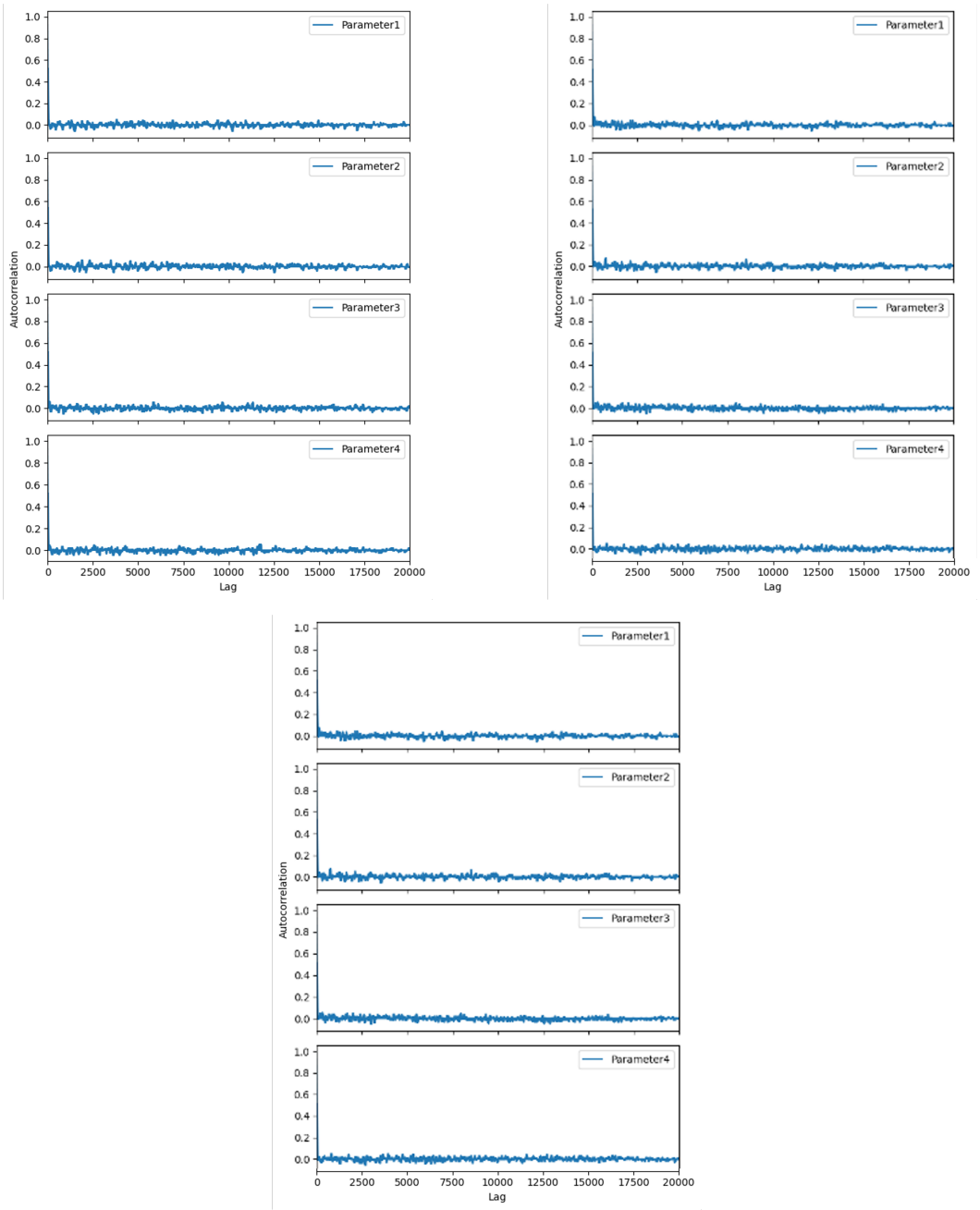


Figure F2: *Autocorrelation plots of the sampled parameter estimates for each of the first four inferred model parameter and each of the three Markov chains produced by running the Haario-Bardenet sampling algorithm [Haario et al., 2001; Clerx et al., 2019]. Each column shows the results for one of the chains of the algorithm.*

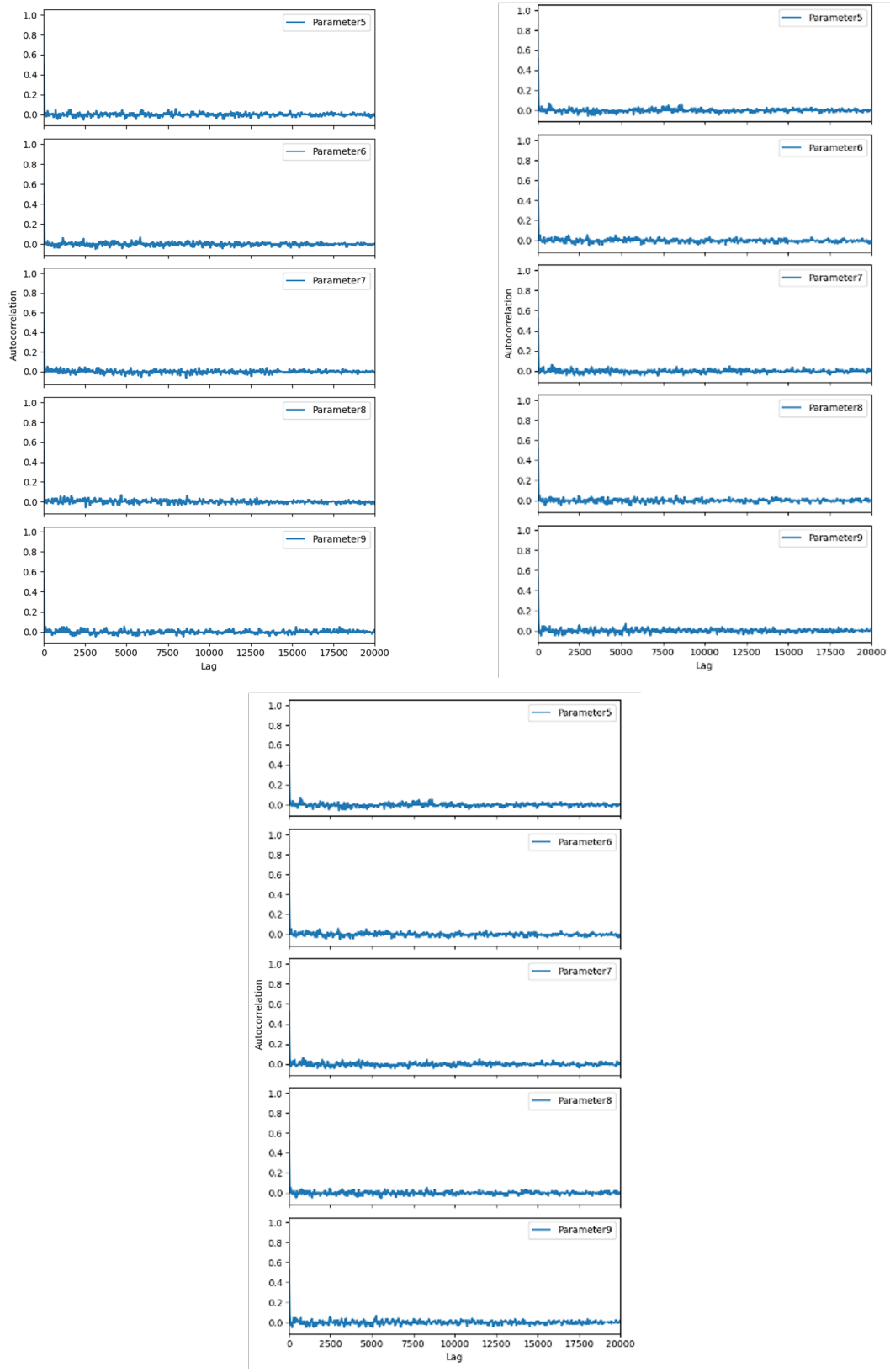


Figure F3: *Autocorrelation plots of the sampled parameter estimates for each of the last five inferred model parameter and each of the three Markov chains produced by running the Haario-Bardenet sampling algorithm [Haario et al., 2001; Clerx et al., 2019]. Each column shows the results for one of the chains of the algorithm.*