IMPERIAL

Imperial College London
Department of Infectious Disease Epidemiology
School of Public Health

Imperial College PhD Thesis

Exploiting epidemiological surveillance to characterise the transmissibility and severity of emerging pandemic pathogens, and optimise response strategies

---

Pablo N Perez-Guzman

March 7, 2025

Submitted in part fulfilment of the requirements for the degree of Doctor of Philosophy in Clinical Medicine Research (Public Health) of Imperial College London and the Diploma of Imperial College London.

Abstract

In this thesis, I build on a validated SARS-CoV-2 transmission model and particle Markov chain Monte Carlo inference framework to develop a novel parametric method to quantify the effect of intrinsic and effective factors of transmission and severity dynamics of outbreak pathogens. I defined intrinsic factors as those determined by direct pathogen's interaction with a fully susceptible population assuming baseline social contact patterns and healthcare capacity, and effective factors as the time-varying pathogen transmissibility and severity given non-pharmaceutical interventions (NPIs) and changes in healthcare capacity, and the profile of population immunity from prior infection, vaccination or both.

To demonstrate the utility of my method, I conduct two COVID-19 case studies, one with rich surveillance data from England, and another with limited data from Kabwe, Zambia, a peri-urban district economically and demographically representative of low-income settings in the Southern Africa region. In England, I present estimations of the basic and effective reproduction number, infection fatality, infection hospitalisation and hospital fatality ratio the Wildtype, Alpha, Delta and Omicron BA.1 variants, systematically analysing how such key metrics responded to changing NPIs, healthcare, and population immunity (infection, vaccination or both) between March 2020 and February 2022. In Kabwe, I show the all-cause cumulative excess mortality rate between March 2020 and September 2021 was among the highest known for any low-income setting. I further demonstrate that an improved access to vaccines and other basic healthcare commodities could have averted most of this pandemic burden, and conduct an in-depth cost-effectiveness and extended dominance evaluation of a comprehensive range of counterfactual strategies, with earlier access to vaccination and improved healthcare, from the health system's perspective.

Deriving granular epidemiological insights across all income settings, regardless of data limitation, is critical to build an evidence base for a more equitable global preparedness to outbreak emergencies.

Executive summary

Reliable and streamlined surveillance systems are the cornerstone of advanced epidemiological modelling, which can play a vital role in informing response interventions to an outbreak emergency. The COVID-19 pandemic was a surveillance and modelling challenge like no other before it. Developing mathematical models that could perform robust inference on the many unobserved processes given available surveillance data was crucial to inform interventions in real-time. In settings where surveillance capacity was limited, modelling analyses relying solely on officially confirmed cases, hospitalisations and/or deaths yielded vast underestimations of the impact of COVID-19 and likely misinformed health policy decisions.

In this thesis, I describe how I built on a validated real-time mathematical modelling framework of SARS-CoV-2 to develop a novel parametric approach to quantifying the effect of intrinsic and effective factors driving the transmission and severity of the virus. My models were stochastic and used particle Markov chain Monte Carlo for conducting Bayesian inference on the epidemiological parameters of interest for the analyses I conducted. I defined *intrinsic drivers* as those explicitly determined by the virus interacting with a fully susceptible population given its baseline social contact patterns and healthcare capacity. On the other hand, I defined *effective drivers* as the time-varying transmissibility and severity of the virus in response to public health interventions, such as non-pharmaceutical interventions and changes in healthcare capacity, and the profile of population immunity from prior infection, vaccination or both.

In Chapters 2 and 3, I discuss how I used rich surveillance data from England to develop my framework and apply it to re-analysing the transmission and severity dynamics of the pandemic between March 2020 and February 2021. I thus fitted my model to representative data from symptomatic PCR cases positivity, variant genotype frequency (from among positive PCR cases), daily vaccinations by age and dose (e.g., first, second or booster), infection prevalence from two surveys, blood donors seroprevalence, hospital admissions, general hospital and ICU bed occupancy, deaths in hospital and the community. In doing so, I was able to describe how the basic reproduction number R_0 of SARS-CoV-2 increased from 2.6 (95% credible interval [CrI] 2.4-28) with the emergence of the Wildtype variant, to 8.4 (95%CrI 7.8-9.1) with Omicron BA.1. I also describe the evolution of the basic infection fatality ratio IFR_0 and other severity metrics across virus variant successions. I find the IFR_0 was highest for the Alpha variant, at 3.0% (95%CrI 2.8-3.2) and that the Omicron BA.1 variant had only a marginally lower IFR_0 than that of the Wildtype variant, at 0.7% (95%CrI 0.5-1.2) and 1.2% (95%CrI 1.1-1.2), respectively. I further systematically analyse how the effective transmission and severity of the pandemic changed in response to population-level changes in contact

rates (i.e., as a result of non-pharmaceutical interventions [NPIs]), the healthcare ecosystem, and the evolving profile of immunity from prior infection, vaccination or both.

To then demonstrate the applicability of my methods to settings with limited surveillance capacity, in Chapter 4 I present a case study of their adaptation to studying sparse SARS-CoV-2 data typical of low-income settings. I specifically conducted an in-depth analysis of the COVID-19 pandemic in Kabwe, Zambia, a peri-urban district economically and demographically representative of low-income settings in the Southern Africa region. For this analysis I relied on seroprevalence data from one representative survey across the local second and third waves, and all-cause deaths data. With the limitations of having limited data to inform my model, I discuss how Kabwe experienced severe constraints to conduct pandemic surveillance and a disproportionately delayed access to COVID-19 vaccines and other basic healthcare commodities to mount an effective mitigation response. As a result, I find that the pandemic-related all-cause excess mortality rate in Kabwe likely was of 402 deaths (95%CrI 277-473) per-100,000 people by the end of September 2021, among the highest known for any low-income setting.

To further quantify the impact of the limitations and delays to roll-out vaccination and implement healthcare improvements in Kabwe on driving this excess mortality burden, in Chapter 5 I conduct a cost-effectiveness evaluation of counterfactual pandemic response interventions. For this I assessed the disability adjusted life-years lost (DALY) averted and incremental costs to the health system in each counterfactual scenario against the factual outcomes quantified in Chapter 4. To assess cost-effectiveness, I further quantified the plausible range of the opportunity cost-based willingness-to-pay (ocWTP) threshold, the monetary threshold above which less health benefits are achieved by the health system given health forgone in other priority activities. I demonstrate that a local roll-out of vaccines in December 2020 with the same weekly vaccination rate achieved by high-income settings would have likely averted two-thirds of the excess mortality in Kabwe. However, I also find that such intervention might not have been cost-effective, assuming the lowest plausible value for the ocWTP threshold of the local health system, at $57.69 2020-US dollars per DALY averted.

Lastly, to better understand the determinants of cost-effectiveness of pandemic mitigation strategies in the context of low-income settings, in Chapter 6 I perform an extended dominance analysis of 250 mutually-exclusive plausible (hypothetical) adoptions of the Coalition for Epidemic Preparedness Initiative (CEPI) 100-day mission, had it been in place and achieved in Kabwe when SARS-CoV-2 emerged. I find that having had access to advanced modelling analytics, such as the ones produced with my methods, in near real-time would have been essential to understand how to implement CEPI's 100-day mission optimally, within the context of the local pandemic and healthcare capabilities. An ambitious roll-out of COVID-19 vaccination targeting 12+ year-olds at an average

weekly rate of 5% of the population, paired with improved healthcare for severe COVID-19 cases was a strategy that was never dominated against other available options. Implementing such an intervention in early May 2020, without delays from the 100-day mission, would have been highly cost-effective even when assuming the lowest plausible ocWTP threshold. Alas delays (particularly of 9 months or longer) to its implementation would have not only gradually undermined the health benefits of the mission, but also likely made it non cost-effective for the health system in Kabwe.

Health systems globally will likely be faced again with decision problems where there will be limited resources and competing interventions to tackle new outbreak emergencies. Robust modelling analytics, such as the ones made possible through the methods I developed during my PhD, open up new avenues to explore the drivers of disease transmission and severity in detail, either with rich or limited surveillance data. In doing so, I demonstrate that deriving granular epidemiological insights across all income settings, regardless of data limitation, is critical to build an evidence base for a more equitable global preparedness to outbreak emergencies.

Statement of originality

I, Pablo N Perez-Guzman, declare that the work in this thesis is my own and was completed during my PhD under the supervision of Prof Katharina Hauck, Dr Anne Cori and Dr Robert Verity. The work of others has been appropriately referenced and a full list of references is given in the bibliography.

For the work presented in Chapter 2, I acknowledge the contributions of the Imperial College COVID-19 real-time modelling team for co-developing the modelling framework upon which I built to create my parametric method. Also, in particular, I acknowledge Dr Marc Baguelin, Dr Edward Knock and Prof Neil Ferguson, for co-supervising me in the development of the parametric methods to quantify SARS-CoV-2 intrinsic and effective transmission and severity.

I have drafted this document in its entirety without the use of any AI tools.

Copyright statement

The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International Licence (CC BY NC-SA).

Under this licence, you may copy and redistribute the material in any medium or format. You may also create and distribute modified versions of the work. This is on the condition that; you credit the author, do not use it for commercial purposes and share any derivative works under the same licence.

When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text. Where a work has been adapted, you should indicate that the work has been changed and describe those changes.

Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law.

Acknowledgements

The COVID-19 pandemic inflicted huge, irreparable losses on so many people and families worldwide, nonetheless my own. Foremost, this work is for all of those left behind to pick up the broken pieces, with a sincere hope that by developing better analytic tools we can do much, much better in responding to future outbreaks as a global community. Mamá, esto es para ti.

To my wonderful supervisors, Katharina Hauck, Anne Cori and Robert Verity, your work ethics and brilliance as scientists have been my biggest source of inspiration before, during and, no doubt will be, after my PhD. I have been honoured to have you all as supervisors. Also, beyond your exceptional supervision, thank you for supporting me through the (more than a few) personal life bumps I have had in the last four years. I am forever in debt.

To my brilliant collaborators in Zambia, thank you enough for your generosity, kindness and the many unforgettable experiences. Jerry, Shadreck, the Kabwe DHO team, the valiant staff at Kabwe Central and Mine hospitals, Prof Chilengui, Dr Kapina, the ZNPHI team, the awe-inspiring TREATS COVID team, Helen, Kwame, Jeanette, Sarah, the Zambarat house team, the work you do truly creates ripples and impacts so the lives of so many people for the better. To Ab and Stephen, you are forces of nature, thank you for making me privy to your visions and aspirations.

To the Imperial College COVID-19 response team, thank you for being such outstanding scientists and colleagues. I learnt so much from you in a very short period of time. What a joy and honour that was, let's do it again! In particular, thank you Ed, Marc and Neil for your extraordinary genius that has inspired me so much. Also, to the wider DIDE team, what a privilege to work with you all these last few (almost eight... wow!) years. Hands down, working in this department makes me feel I am in the company of greatness.

A very, very special shout out to Jose Lourenco and Mikaela Smit. It is thanks to you that I found myself going down this rabbit hole. After several years of fever dreams of code and stakeholders meetings, I think I'm finally getting the hang of this modelling business. Thank you for showing me 'the way'.

Lastly, to my wonderful family. Sarah and Elisa, you are my core and drive. Through all of our successes and stumbles, we are always together and we are so much the better because of it. Thank you for not only putting up with my middle-aged whim of doing a PhD, but positively encouraging me through it all. Papá, Karen y Joakim, gracias por todo su apoyo y cariño también... como que ya va siendo hora de decir, ¿no?

Contents

1 Introduction **24**

1.1 SARS-CoV-2 pandemic in England . . . 25

1.1.1 Surveillance . . . 25

1.1.2 Analytics and scientific advice . . . 26

1.1.3 SARS-CoV-2 waves and pandemic response . . . 27

1.2 SARS-CoV-2 pandemic in Kabwe district, Zambia . . . 30

1.2.1 Surveillance . . . 31

1.2.2 SARS-CoV-2 waves and pandemic response . . . 31

1.3 Transmissibility and severity analytics . . . 34

1.3.1 Transmissibility . . . 34

1.3.2 Severity . . . 39

1.3.3 Age-structured, multi-variant compartmental models of SARS-CoV-2 . . . 40

1.4 Intrinsic vs effective transmissibility and severity . . . 42

1.5 Pandemic inequality . . . 44

1.6 Thesis aims and objectives . . . 45

1.7 Thesis overview . . . 47

1.8 Publications related to this thesis . . . 48

2 Novel parametric approach to quantify the intrinsic and effective drivers of transmissibility and severity **49**

2.1 Background . . . 49

2.2 Methods . . . 50

2.2.1 A foreword on pMCMC . . . 50

2.2.2 Model overview . . . 51

2.2.3 Modelling transmissibility . . . 56

2.2.4 Modelling severity . . . 63

2.3 Discussion . . . 72

3 Drivers of SARS-CoV-2 transmissibility and severity in England 74
3.1 Background . . . 74
3.2 Methods . . . 75
3.2.1 COVID-19 surveillance data . . . 75
3.2.2 Observation process . . . 77
3.2.3 Effective population-level immunity against infection . . . 84
3.2.4 Model parameters . . . 85
3.2.5 Running the model . . . 95
3.3 Results . . . 96
3.3.1 Transmissibility . . . 96
3.3.2 Severity . . . 97
3.4 Discussion . . . 99
3.5 Conclusion . . . 101

4 Drivers of SARS-CoV-2 transmissibility and severity in Kabwe, Zambia 102
4.1 Background . . . 102
4.2 Methods . . . 103
4.2.1 COVID-19 surveillance data . . . 103
4.2.2 Model overview . . . 105
4.2.3 Observation process . . . 107
4.2.4 Model inference . . . 110
4.2.5 Running the model . . . 114
4.3 Results . . . 115
4.3.1 Transmissibility . . . 115
4.3.2 Severity . . . 116
4.4 Discussion . . . 118
4.5 Conclusion . . . 122

5 The impact and cost-effectiveness of COVID-19 pandemic inequality in Kabwe, Zambia 123
5.1 Background . . . 123
5.2 Methods . . . 124
5.2.1 Modelling healthcare utilisation . . . 125
5.2.2 Vaccine efficacy . . . 130
5.2.3 Counterfactual scenarios . . . 131
5.2.4 Economic model . . . 134

5.3 Results . . . 138
5.3.1 Impact of limited and delayed pandemic interventions . . . 138
5.3.2 Cost-effectiveness of counterfactual interventions . . . 139
5.4 Discussion . . . 143
5.5 Conclusion . . . 146

6 Future pandemic preparedness in Kabwe, Zambia anchored on learnings from the COVID-19 pandemic: modelling the 100-day mission 147
6.1 Background . . . 147
6.2 Methods . . . 148
6.2.1 Mutually exclusive intervention options . . . 148
6.2.2 Extended dominance analysis . . . 151
6.2.3 Economic analysis uncertainty . . . 152
6.3 Results . . . 153
6.3.1 Extended dominance analysis . . . 153
6.3.2 Uncertainty analysis of non-dominated interventions . . . 155
6.4 Discussion . . . 157
6.5 Conclusion . . . 160

7 Discussion 162
7.1 Summary of key findings & limitations . . . 162
7.2 On the need for a holistic strengthening of surveillance systems . . . 164
7.2.1 Strengthening surveillance-rich systems . . . 166
7.2.2 Strengthening surveillance-limited systems . . . 168
7.3 On the need for collaborative approaches to the development of outbreak analytics . . . 169
7.4 On the host-pathogen-environment relationship and real-time outbreak analytics . . . 171
7.5 On the optimisation of outbreak response interventions . . . 174
7.6 Concluding remarks . . . 177
Appendices . . . 212
A England model technical appendix . . . 212
A.1 Model equations . . . 213
A.2 Vaccination . . . 226
A.3 Sensitivity analyses . . . 236
A.4 Fitted data trajectories . . . 246
A.5 Fitted parameters . . . 269
A.6 Inferred trajectories . . . 280

A.7 Comparison to non-fitted trajectories 286
B Kabwe model technical appendix 292
B.1 Model demography and infections seed 292
B.2 Model parameters 293
B.3 Model equations 297
B.4 Additional results: chapter 4 301
B.5 Additional results: chapter 5 311
B.6 Additional results: chapter 6 338

List of Figures

1.1 SARS-CoV-2 waves in England . . . 28
1.2 SARS-CoV-2 waves in Kabwe . . . 32
1.3 Contact matrix from socialmixr . . . 37
1.4 Intrinsic *vs* effective outbreak drivers . . . 43

2.1 State-space model inference . . . 51
2.2 SARS-CoV-2 England model diagram . . . 53

3.1 SARS-CoV-2 transmissibility in England between March 2020 and February 2022 . . . 97
3.2 SARS-CoV-2 severity in England across the Wildtype, Alpha, Delta and Omicron (BA.1) waves . . . 99

4.1 SARS-CoV-2 model for Kabwe, Zambia . . . 105
4.2 Deaths timeseries data . . . 109
4.3 SARS-CoV-2 epidemic in Kabwe, Zambia between March 2020 and September 2021 . . . 117

5.1 Modified SARS-CoV-2 model for Kabwe, Zambia CEA . . . 127
5.2 Vaccine type by age . . . 131
5.3 Vaccination schedules . . . 133
5.4 Counterfactual excess deaths in Kabwe . . . 139
5.5 Cost-effectiveness of counterfactual scenarios in Kabwe . . . 141
5.6 Quarterly net monetary benefit of counterfactual scenarios in Kabwe . . . 143

6.1 Factual vaccination capacity by income . . . 149
6.2 Extended dominance analysis in Kabwe, Zambia . . . 155
6.3 CEP and CEAC of non-dominated interventions in Kabwe, Zambia . . . 156
6.4 Incremental NMB of non-dominated interventions in Kabwe, Zambia . . . 158

7.1 ONS, REACT-1 and Pillar 2 comparison . . . 167
7.2 Disease triangle & intrinsic *vs* effective drivers . . . 172
A1 SARS-CoV-2 England model diagram . . . 212

A2 PCR and seropositivity structure . . . 225
A3 Vaccination strata . . . 226
A4 Vaccine uptake by type and age . . . 230
A5 VEalpha . . . 231
A6 VEdelta . . . 232
A7 VEomicron . . . 233
A8 Sensitivity analysis of vaccine efficacy . . . 241
A9 Sensitivity analysis of cross-immunity . . . 242
A10 Sensitivity analysis of serial interval and $mu_D(t)$ . . . 245
A11 Model fit to hospitals data . . . 246
A12 Model fit to Pillar 2 data . . . 246
A13 Model fit to ONS data . . . 247
A14 Model fit to REACT data . . . 247
A15 Model fit to Euroimmun serology . . . 248
A16 Model fit to Roche serology . . . 248
A17 Model fit to Alpha/Wildtype VAM data . . . 249
A18 Model fit to Delta/Alpha VAM data . . . 249
A19 Model fit to Omicron/Delta VAM data . . . 250
A20 Model fit admissions 0 to 9 year-olds . . . 250
A21 Model fit admissions 10 to 19 year-olds . . . 251
A22 Model fit admissions 20 to 29 year-olds . . . 251
A23 Model fit admissions 30 to 39 year-olds . . . 252
A24 Model fit admissions 40 to 49 year-olds . . . 252
A25 Model fit admissions 50 to 59 year-olds . . . 253
A26 Model fit admissions 60 to 69 year-olds . . . 253
A27 Model fit admissions 70 to 79 year-olds . . . 254
A28 Model fit admissions 80+ year-olds . . . 254
A29 Model fit community deaths 0 to 49 year-olds . . . 255
A30 Model fit community deaths 50 to 54 year-olds . . . 255
A31 Model fit community deaths 55 to 59 year-olds . . . 256
A32 Model fit community deaths 60 to 64 year-olds . . . 256
A33 Model fit community deaths 65 to 69 year-olds . . . 257
A34 Model fit community deaths 70 to 74 year-olds . . . 257
A35 Model fit community deaths 75 to 79 year-olds . . . 258
A36 Model fit community deaths 0 to 49 year-olds . . . 258

A37 Model fit hospital deaths 0 to 49 year-olds . . . 259
A38 Model fit hospital deaths 50 to 54 year-olds . . . 259
A39 Model fit hospital deaths 55 to 59 year-olds . . . 260
A40 Model fit hospital deaths 60 to 64 year-olds . . . 260
A41 Model fit hospital deaths 65 to 69 year-olds . . . 261
A42 Model fit hospital deaths 70 to 74 year-olds . . . 261
A43 Model fit hospital deaths 75 to 79 year-olds . . . 262
A44 Model fit hospital deaths 0 to 49 year-olds . . . 262
A45 Model fit Pillar 2 cases 0 to 14 year-olds . . . 263
A46 Model fit Pillar 2 cases 15 to 24 year-olds . . . 263
A47 Model fit Pillar 2 cases 25 to 49 year-olds . . . 264
A48 Model fit Pillar 2 cases 50 to 64 year-olds . . . 264
A49 Model fit Pillar 2 cases 65 to 79 year-olds . . . 265
A50 Model fit Pillar 2 cases 80+ year-olds . . . 265
A51 Model fit REACT-1 5 to 24 year-olds . . . 266
A52 Model fit REACT-1 25 to 34 year-olds . . . 266
A53 Model fit REACT-1 35 to 44 year-olds . . . 267
A54 Model fit REACT-1 45 to 54 year-olds . . . 267
A55 Model fit REACT-1 55 to 64 year-olds . . . 268
A56 Model fit REACT-1 65+ year-olds . . . 268
A57 $\beta(t)$ posterior distributions . . . 269
A58 Posterior distributions of variant transmission and severity parameters . . . 270
A59 Posterior distributions of time-varying severity parameters . . . 271
A60 Posterior distributions of all other parameters . . . 272
A61 MCMC traceplots for East of England . . . 272
A62 MCMC traceplots for London . . . 273
A63 MCMC traceplots for Midlands . . . 273
A64 MCMC traceplots for North East . . . 274
A65 MCMC traceplots for North West . . . 274
A66 MCMC traceplots for South East . . . 275
A67 MCMC traceplots for South West . . . 275
A68 pMCMC traceplots for East of England . . . 276
A69 pMCMC traceplots for London . . . 276
A70 pMCMC traceplots for Midlands . . . 277
A71 pMCMC traceplots for North East . . . 277

A72 pMCMC traceplots for North West . . . 278
A73 pMCMC traceplots for South East . . . 278
A74 pMCMC traceplots for South West . . . 279
A75 $\beta(t)$ trajectory . . . 280
A76 Cumulative attack rate . . . 280
A77 Daily incidence rate . . . 281
A78 μ_D trajectory . . . 281
A79 R_t trajectory . . . 282
A80 R_t^{eff} trajectory . . . 282
A81 Effective susceptible trajectory . . . 283
A82 Infection status trajectory . . . 283
A83 Vaccination status trajectory . . . 284
A84 Effective severity by age trajectory . . . 285
A85 Hospital admission by vaccination status . . . 286
A86 Hospital deaths by vaccination status . . . 287
A87 Hospitalisations and deaths age distribution . . . 288
A88 Hospital fatality ratio by NHS region . . . 289
A89 Hospital admissions by age, national level . . . 290
A90 Hospital deaths by age, national level . . . 290
A91 Community deaths by age, national level . . . 291
B1 $\beta(t)$ posterior distributions . . . 302
B2 Posterior distributions of all other fitted parameters . . . 303
B3 MCMC traceplots of fitted parameters . . . 304
B4 pMCMC traceplots of fitted parameters . . . 305
B5 Sensitivity analysis of fit to seroprevalence . . . 306
B6 Sensitivity analysis of fit to all-cause mortality . . . 307
B7 Sensitivity analysis of inferred R_t . . . 308
B8 Sensitivity analysis of inferred IFR_t . . . 309
B9 Inferred Rt and mobility data . . . 310
B10 Daily infections in counterfactual vaccination scenarios . . . 312
B11 Daily infections in counterfactual healthcare scenarios . . . 313
B12 Daily deaths in counterfactual vaccination scenarios . . . 314
B13 Daily deaths in counterfactual healthcare scenarios . . . 315
B14 R_t in counterfactual vaccination scenarios . . . 316
B15 R_t in counterfactual healthcare scenarios . . . 317

B16 IFR_t in counterfactual vaccination scenarios 318
B17 IFR_t in counterfactual healthcare scenarios 319
B18 Vaccination coverage in counterfactual scenarios 320
B19 Immunity profile with counterfactual vaccination 321
B20 Immunity profile with counterfactual healthcare 322
B21 Infections averted in counterfactual scenarios 323
B22 Deaths averted high uptake of vaccines 324
B23 Immunity profile high uptake of vaccines 325
B24 Deaths averted low uptake of vaccines 326
B25 Immunity profile low uptake of vaccines 327
B26 ICERs and NMB across counterfactual scenarios 328
B27 DALYs by age in vaccination counterfactual 329
B28 DALYs by age in healthcare counterfactuals 330
B29 Tornado plot of vaccination counterfactuals with low NMB 331
B30 Tornado plot of vaccination counterfactuals with high NMB 331
B31 Tornado plot of healthcare counterfactuals with low NMB 332
B32 Tornado plot of healthcare counterfactuals with high NMB 332
B33 DALYs averted in extended dominance analysis 338
B34 Total costs in extended dominance analysis 339
B35 Cost-effectiveness plane by dominance status 340

List of Tables

2.1 Definitions of compartments . . . 54
2.2 Definitions of parameters . . . 55
2.3 Variants of concern serial interval . . . 62
2.4 Piecewise severity change-points in England . . . 64
2.5 Age multipliers . . . 70

3.1 England data sources and definitions . . . 76
3.2 Fixed model parameters . . . 86
3.3 Inferred England model parameters . . . 88

4.1 Data sources for analysis in Kabwe, Zambia . . . 104
4.2 Definitions of compartments . . . 106
4.3 Definitions of parameters . . . 106
4.4 Definitions of parallel flow compartments . . . 107
4.5 Age-specific severity . . . 114

5.1 Definitions of hospital parallel flow compartments . . . 126
5.2 Kabwe baseline probability of severe disease by age . . . 129
5.3 Fixed healthcare parameters for Kabwe . . . 129
5.4 Vaccine efficacy parameters for Kabwe. . . . 130
5.5 Counterfactual Kabwe, Zambia scenarios . . . 132
5.6 Cost parameters for Kabwe . . . 134
5.7 Disability weights . . . 136
5.8 Age-specific parameters of cost-effectiveness analysis . . . 136
5.9 Opportunity cost-based willingness to pay threshold parameters . . . 137

A1 Definitions of parallel flow compartments . . . 224
A2 Vaccination strata . . . 227
A3 VE for AZ, PF and Mod by vaccine dose. . . . 229
A4 Conditional vaccine effectiveness values . . . 235

A5 Summary of sensitivity analyses explored. Source: Perez-Guzman et al. 2023 [1]. . . . 237
A6 Sensitivity analysis for vaccine efficacy *vs* the Alpha variant 238
A7 Sensitivity analysis for vaccine efficacy *vs* the Delta variant 239
A8 Sensitivity analysis parameters for vaccine efficacy of boosters *vs* the Omicron BA.1 variant 240
A9 Sensitivity analysis of $\mu_D(t)$ change points 244
B1 Population of Kabwe 292
B2 Inferred model parameter notations and prior distributions 294
B3 Counterfactual DALYs and DALYs averted 333
B4 Counterfactual incremental cost-effectiveness ratio 334
B5 Counterfactual incremental and total costs 336
B6 Counterfactual net monetary benefit 337
B7 League table of extended dominance 341
B7 League table of extended dominance 342
B8 Net-monetary benefit of extended dominance scenarios with low WTP 343
B8 Net-monetary benefit of extended dominance scenarios with low WTP 344
B9 Net-monetary benefit of extended dominance scenarios with high WTP 344
B9 Net-monetary benefit of extended dominance scenarios with high WTP 345
B9 Net-monetary benefit of extended dominance scenarios with high WTP 346

List of Acronyms

AZ Astra-Zeneca

CEA Cost-effectiveness analysis

CEAC Cost-effectiveness acceptability curve

CEP Cost-effectiveness plane

CEPI Coalition for Epidemic Preparedness Initiative

CERP Centre for Excellence in Respiratory Pathogens

CFR Case fatality ratio

CHESS COVID-19 Hospitalisation in England Surveillance System

CI Confidence interval

CO-CIN COVID-19 Clinical Information Network

COG-UK UK COVID-19 Genomics Consortium

CPU Central processing unit

CI Confidence interval

CrI Credible interval

DALY Disability-adjusted life-year

EPI Effective population-level immunity

ESS Effective sample size

FF100 First few hundred cases

GDPpc Gross domestic product per capita

HE Health-economic evaluation

HFR Hospital fatality ratio

HIC High-income country

HMM Hidden Markov models

ICER Incremental cost-effectiveness ratio

ICL-RTM Imperial College real-time modelling team

ICU Intensive Care Unit

IFR Infection fatality ratio

IHR Infection hospitalisation ratio

IQR Interquartile range

JBC Joint Biosecurity Centre

LIC Low-income country

L-MIC Lower middle-income country

LMICs Low- and middle-income countries

MCMC Markov chain Monte Carlo

mRNA Messenger ribonucleic acid

NGM Next generation matrix

NGMs Next generation matrices

NHS National Health Service

NMB Net monetary benefit

NPIs Non-pharmaceutical interventions

ocWTP Opportunity cost-based willingness to pay

ONS Office for National Statistics

PATH Program for Appropriate Technology in Health

PCR Polymerase chain reaction

Pf Pfizer

PHE Public Health England

pMCMC Particle Markov chain Monte Carlo

QALY Quality-adjusted life-year

REACT-1 Real-time Assessment of Community Transmission

SAGE Scientific Advisory Group for Emergencies

SARS Severe Acute Respiratory Syndrome

SEIR Susceptible - Exposed - Infected - Recovered

SIREN SARS-CoV-2 immunity and reinfection evaluation

SMC Sequential Monte Carlo

SPI-M-O Scientific Pandemic Influenza group on Modelling, Operational sub-group

TREATS Tuberculosis Reduction through Expanded Antiretroviral and Screening for Active TB

UEFA Union of European Football Associations

UK United Kingdom

USD United States dollar

U-MIC Upper middle-income country

VAM Variant and Mutation data

VCV Variance covariance

VE Vaccine effectiveness

VOC Variant of concern

VOCs Variants of concern

WFH Working from home

WHO World Health Organization

WTP Willingness to pay

YLD Years lived with disability

YLL Yeasts of life lost

ZGSC Zambia Genomic Sequencing Consortium

ZNPHI Zambia National Public Health Institute

Chapter 1

Introduction

The overarching aim of my PhD project was to assess how to robustly exploit available outbreak data to: a) unpick and quantify the factors that drive its transmissibility and severity dynamics over time, and b) use the latter knowledge to optimise public health response interventions. Three central themes underpinned my work. Firstly, that the transmissibility and severity dynamics of an outbreak are not solely determined by the pathogen itself. Rather, there are two levels of epidemiological factors that drive these dynamics, ***intrinsic***, or pathogen-driven, and ***effective*** or population-level driven (Section 1.4). Secondly, that both the intrinsic and effective drivers of outbreak dynamics are constantly changing over time. And, thirdly, that a robust understanding of the effect of these drivers on outbreak dynamics is of vital importance to devising and optimising public health response interventions.

During real-world outbreak responses, however, the necessary data to robustly study these drivers and the resulting dynamics is imperfect, delayed and limited to the surveillance capacity of health systems [2]. The work I conducted during my PhD introduced several technical innovations for analysing the transmission and severity dynamics of outbreak pathogens in granular detail, either with data from health systems with advanced or parsimonious surveillance capacity.

In this thesis, I summarise my research proposing a novel parametric approach to quantify the effect of *intrinsic* and *effective* drivers on the transmissibility and severity dynamics of an outbreak. I then present two case studies of its applicability to retrospectively studying the time-varying drivers and dynamics of SARS-CoV-2 using rich surveillance data from England and limited data from Kabwe district, Zambia, as contrasting examples of factual outbreak dynamics in high-income (HIC) and low-income country (LIC) settings, respectively. Although these two case studies were retrospective, I will argue the potential utility and limitations of my proposed approach in real-time analytics. I then elaborate on my case study of Kabwe to demonstrate how a granular analysis of said drivers and dynamics allowed me to quantify the dire effects of pandemic response inequity in LICs. Lastly, I further elaborate on this to estimate the cost-effectiveness of having had optimal capacity to mitigate

the effects of the COVID-19 pandemic in Kabwe or, more importantly, how to assess the optimisation of future responses in the context of limited healthcare resources, towards equitable future outbreak preparedness.

In this chapter, I summarise the relevant context that informed the rational and need for my research. I begin by giving an overview of the timeline of events during the COVID-19 pandemic study periods in England and Kabwe, providing key details of the waves by different SARS-CoV-2 variants, response interventions and breath of surveillance data available. I then describe a selected range of real-time model-based analytics that were employed to assess the transmissibility and severity dynamics of SARS-CoV-2. As relevant, I will highlight a) how they were used to inform mitigation interventions, and b) their critical limitations that I sought to address with my research. I then define intrinsic *vs* effective drivers of transmissibility and severity, for the purpose of my analyses. Lastly, I provide a brief comment on pandemic inequality and on the role that prioritising real-time analytics to inform the response of higher income settings played in them. I conclude this chapter by outlining my thesis aims and objectives and providing an overview of the structure of the following chapters.

1.1 SARS-CoV-2 pandemic in England

1.1.1 Surveillance

The first case of COVID-19 was detected in England on January 31, 2020 [3]. Local preparedness actions and mechanisms to undertake epidemiological intelligence pre-dated the arrival of SARS-CoV-2. In particular, England benefited from pre-existing strengths and capacity to implement surveillance and conduct clinical trials of novel therapeutics [3].

In many regards, England was a case study of extraordinarily rich pandemic surveillance data [4]. The First Few Hundred Cases Study (FF100) was set-up in January 2020 to provide early insights into the clinical profile of COVID-19 cases [3]. The methodology for the FF100 [5] fed directly into the World Health Organization's (WHO) first surveillance action plan, issued in February 2020, prescribing the undertaking of similar studies across member states [6]. For the case of England, however, the FF100 study was rapidly superseded by larger, nationally representative studies and optimised frontline healthcare data collection systems.

An early coordination between national funding, implementation and healthcare institutions allowed for the rapid deployment of the Government's flagship polymerase chain reaction (PCR) testing programme on March 23, 2020 [3]. This programme became fully operational from May 2020 [7], allowing universal testing and detection of symptomatic cases across several settings. Pillar

1 of the programme covered healthcare workers and patients accessing healthcare settings, and Pillar 2 the general public [3]. For the latter, the government recommended testing be based on the onset of symptoms like fever, a persistent cough and/or anosmia [3]. However, population test-seeking behaviours changed over the course of the pandemic, biasing Pillar 2 surveillance given dynamic sociodemographic aspects, including proximity to test centres, perceived individual-level risk and digital literacy [7].

Concurrently, England established the Joint Biosecurity Centre (JBC) in May 2020, which coordinated surveillance and analytics capacity across the NHS, Pillar 2, and a flurry of emerging national studies [3]. The latter included two landmark representative infection prevalence surveys, the Office for National Statistics (ONS) Coronavirus Infection Survey [8] and the REal-time Assessment of Community Transmission (REACT-1) study [9], and what can be considered as the best example of genomic sequencing surveillance for any pathogen anywhere in the world to date, the COVID-19 Genomics Consortium (COG-UK) [3]. These studies were instrumental in complementing Pillar 2 to track community-level transmission of the virus and the emergence of its different variants [3]. On the one hand, the ONS and REACT-1 studies were prospective, stratified random sample-based household surveys that tested with PCR individuals across England, regardless of the presence of symptoms [8, 9]. On the other, COG-UK conducted whole genome sequencing of a representative, daily random sample of all positive cases detected through Pillar 2 [3].

Lastly, the JBC also coordinated a range of clinical trials and routine, near real-time collection of healthcare data. This included the largest multi-national clinical trials of the pandemic, the International Severe Acute Respiratory and emerging Infection Consortium's (ISARIC), and the COVID-19 Clinical Information Network (CO-CIN) studies [10]. Daily linelist data bases of hospitalised cases linked to outcomes, such as the COVID-19 Hospitalisation in England Surveillance System (CHESS) [11], and streamlined reporting of situational reports by National Health Service (NHS) trusts were also set-up in early March 2020, and conducted prospectively throughout the pandemic.

1.1.2 Analytics and scientific advice

The rich landscape of pandemic surveillance data in England was an enabling factor for a range of granular analytics that were instrumental to inform the UK Government's scientific advisory groups. The Imperial College real-time modelling (ICL-RTM) team (of which I was a member of prior to my PhD) and several other academic groups used those data to produce weekly modelling analytics between May 2020 and April 2022. These were routinely reported to the Scientific Pandemic Influenza Group on Modelling (SPI-M-O), which in turn aggregated them and summarised into academic consensus statements for the Scientific Advisory Group for Emergencies (SAGE), co-chaired

by UK Government officials that directly provided scientific advice to the Prime Minster and Cabinet Office [3].

There were many great scientific breakthroughs in the field of model-based inference and analytics pioneered by the UK-based academic groups conforming SPI-M-O (see section 1.3). These not only served to provide the timely analytics that informed the UK's response, but were also widely adopted by academic groups and public health agencies world wide. Yet, it was not always the case that sound analytics underpinned pandemic response decision-making. With the benefit of hindsight, I discuss instances where analytic evidence and pandemic policies were misaligned in the following section and in Chapter 3.

1.1.3 SARS-CoV-2 waves and pandemic response

My analysis of the COVID-19 pandemic in England (Chapter 3) spanned from March 16, 2020 to February 24, 2022. During this period, there were four distinct epidemic waves, with documented dominance by the initial Wildtype variant of SARS-CoV-2, followed by the Alpha, Delta and Omicron BA.1 variants of concern (VOCs), respectively (Figure 1.1). The landscape of non-pharmaceutical interventions (NPI) policies across these waves was highly dynamic, at times informed by the UK Government's perception of risks to the health system and, at others, by the consensus analytics from SPI-M-O submitted to SAGE [3].

Analyses from COG-UK [3] allowed for the complete documentation of several hundreds of the international importation events of the Wildtype variant in England that occurred from as early January 2020 [12]. The first official NPIs came into force on March 16, 2020 [13], following the report of the first confirmed death by COVID-19 in England [14]. The aim of these NPIs was to mitigate transmission (whereby the rate of social contacts is decreased to limit, but not necessarily stop, new infections) and prevent hospitalisations from overwhelming the health system [3]. On the same day these NPIs came into force, the landmark Report 9 from the ICL-RTM team was published [15]. It highlighted that, in the UK and US context, a suppression strategy (whereby contact rates are decreased to a sufficient level to cause epidemic decline) through the implementation of stringent NPIs would be needed in order to avoid the hundreds of thousands of deaths that were likely to occur with a mitigation-only strategy [15].

As it happened, a decision to implement lockdowns, and thus a suppression strategy, had been taken by several European countries from as early as March 11, 2020 [16]. Ultimately, the UK also adopted this strategy when the first national lockdown began on March 23, 2020 [13]. Modelling analyses at the time [3, 16–18] indeed revealed that less stringent NPIs were unlikely to achieve the goal of protecting the health system from being overrun.

The first national lockdown in England ended on May 11, 2020, with an official recommendation

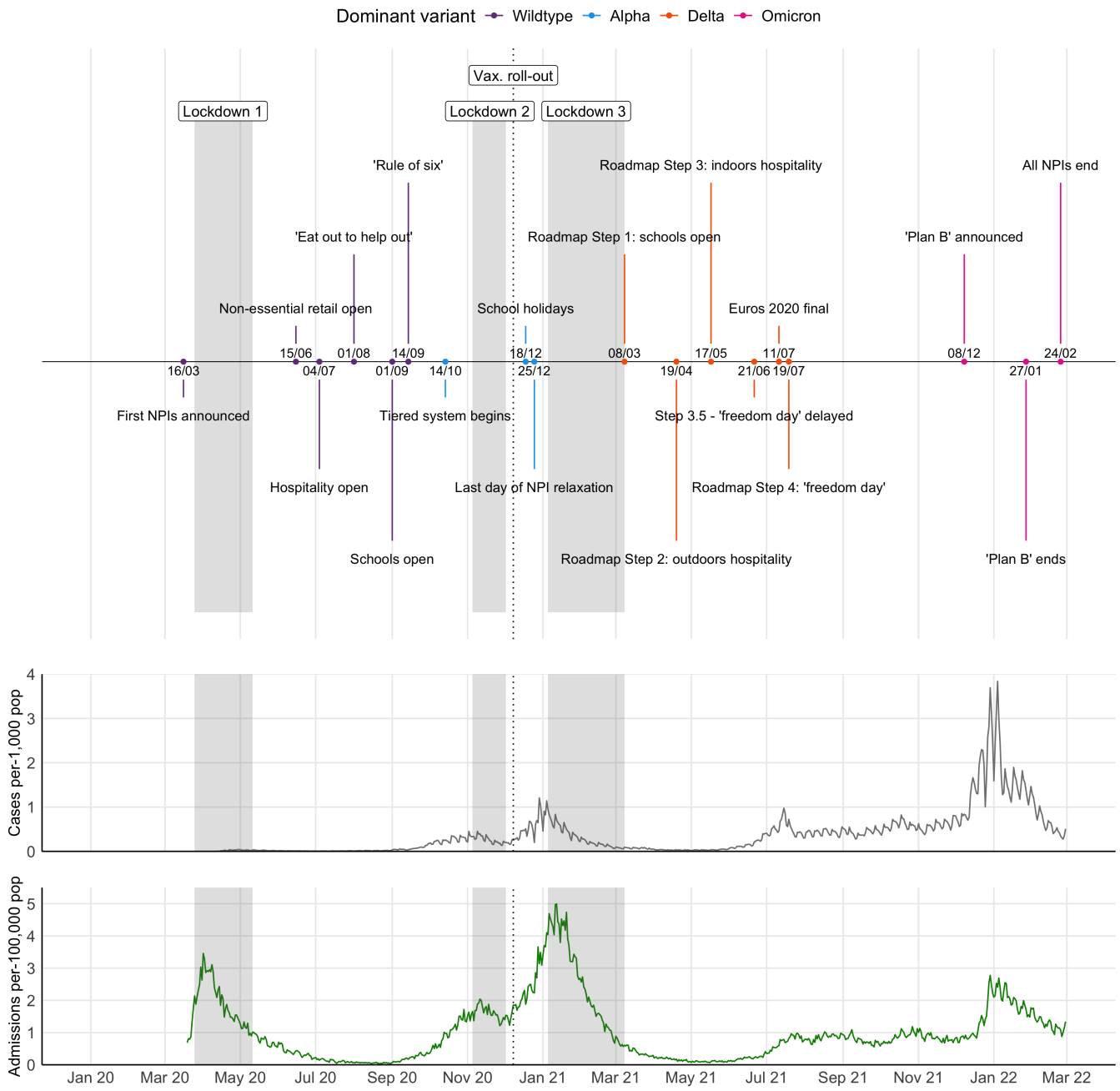


Figure 1.1: SARS-CoV-2 waves in England from March 16, 2020 to February 24, 2022. The landscape of NPI policies varied across waves, which were dominated, in succession, by the Wildtype, Alpha, Delta and Omicron BA.1 variants. For contextual reference, bottom panels shows the number of cases per-1,000 population detected by the Pillar 2 programme (community based, universal testing) and the number daily hospital admissions per-100,000 population across England.

to continue working from home where possible and avoid indoor social contacts [3]. Restrictions were gradually eased and, following the UK Government Treasury's "Eat Out to Help Out" scheme, non-essential shops and the hospitality sector re-opened in August 2020 [19] (Figure 1.1). The R_t of the pandemic increased above 1 in response to this sequential relaxation of NPIs [3, 18, 20], and a second wave thus ensued (Figure 1.1). Although new NPIs were introduced, including the "rule of six" (banning of social gatherings of more than six individuals) on September 14, 2020 [21], followed by a tiered regional system for the introduction and lifting of NPIs [22], pandemic transmission continued

to rise across all England regions [18, 23].

Amidst this new wave of infections, the Alpha VOC was first detected in late September 2020 as a rise in Pillar 2 cases with a negative S-gene target in qualitative PCR [24]. Early genomic analysis from cases with linked hospitalisation and death outcomes showed that, compared to the Wildtype variant, the Alpha VOC was more transmissible and severe [25, 26]. The progressive replacement of the Wildtype by the Alpha VOC led to a further increase in the number of COVID-19 cases and hospitalisations (Figure 1.1). A second national lockdown, therefore, came into force between November 5 and December 1, 2020 [27].

By mid-December 2020, the Alpha VOC dominated transmission and, following the ending of the second lockdown, caused the winter 2020/21 wave in England which culminated in the highest peak in hospitalisations seen at any point in the national pandemic (Figure 1.1). During the early stages of this winter wave, on December 8, 2020 England became the first country in the world to deliver a COVID-19 vaccine outside of a clinical trial [28]. Alas, it became evident that the second national lockdown had not been as effective at controlling transmission compared to the first one [18]. This was partly explained by the fact that schools and some workplaces were allowed to remain open during this period [27], but also because of the emergence and dominance of the Alpha VOC [29]. The third (and final) lockdown of the pandemic thus began on January 4, 2021 [29].

During the third lockdown, as the national vaccination programme was well underway, the SPI-M-O and SAGE groups commissioned modelling analyses by several academic groups looking to define a strategy for the gradual relaxation of NPIs [3]. This became known as the UK Government's *'Roadmap out of lockdown'* (Figure 1.1). It established a series of four steps to be undertaken gradually, as follows [3]:

1. Step 1: reopening of schools,
2. Step 2: limited reopening of non-essential business and outdoor hospitality,
3. Step 3: reopening of indoor hospitality, and
4. Step 4: *'freedom day'*, where all NPI restrictions would be lifted and night clubs and other mass-entertainment venues would be allowed to also reopen.

On March 8, 2021, almost a year after the first NPIs were issued in England, the first step of the roadmap took place (Figure 1.1). This coincided with early reports of the emergence of three novel VOCs, Beta, Gamma and Delta in South Africa, Brazil and India, respectively, all of which exhibited an increased transmissibility profile over Alpha, as well as a partial capacity to evade prior infection-derived immunity [30, 31]. Although all these VOCs were eventually detected in England [30], it was Delta which replaced Alpha as the dominant variant by May 2021 [32, 33]. At this point, step

3 of the roadmap had been taken (Figure 1.1) and the European UEFA football tournament was taking place. As England advanced through the knockout rounds of the competition new increases in cases and hospitalisations were seen (Figure 1.1) [32, 33]. This led the UK Government to take a decision to delay step 4 from its originally planed date on June 21 [3]. Despite the dominance of the more-transmissible Delta VOC, it the vaccination programme had been successful at effectively suppressing its transmissibility and severity [32,33] and, thus, 'freedom day' did eventually take place on July 19, 2021.

A cumulative 79% of the population had received two doses of an approved COVID-19 vaccine by early October 2021 [34]. So, in the lead-up to the winter of 2021/22, this meant that the accrued levels of immunity in the population, toped-up with the roll-out of booster (third) doses was expected to maintain a suppressed trajectory of the epidemic [34]. My PhD began at this stage of the pandemic, in October 2021, just before the genomic analysis of four sequences from Southern Africa allowed for the characterisation of a novel VOC, Omicron BA.1, which exhibited the highest number of mutations to date of any VOC [35]. What became quickly evident at the time was that, as a result of such genetic divergence from previous VOCs, Omicron BA.1 had escape properties to both prior infection- and vaccine-induced immunity [36]. This led to the final reinstatement of NPIs in England, known as *'Plan B'*, which included working from home, wearing face masks and/or showing proof of a negative COVID-19 test in public venues, and a ramp-up in the roll-out of booster (3rd) vaccine doses [37]. On February 24, 2022 the UK Government's plan for living with COVID-19 came into force, marking the end to all official NPI restrictions, including mandatory testing and self-isolation, and ushered the gradual disbandment of Pillar 2 and many other flagship pandemic surveillance efforts [38]. Over 70% and 50% of the population in England had received a second and booster (third) dose, respectively, by this point [28].

1.2 SARS-CoV-2 pandemic in Kabwe district, Zambia

Kabwe is a peri-urban district of the central province of Zambia, located about 140 km north of the nation's capital, Lusaka [39]. It is home to an estimated 299,206 people [40], with a median age of 17 years [41], comparable to that of other sub-Saharan African countries (excluding South Africa) at 18 years, but considerably lower than that of Western European and North American HICs, at around 40 years [42].

Before the emergence of SARS-CoV-2, the health system of Zambia, as it is the case of other LICs in the region, was already constrained by the demands of the rising double burden of infections and non-communicable diseases [43]. The profile of population health and available healthcare in Kabwe lagged behind the national average for Zambia, as evidenced by the 2017 Kabwe Household

Socioeconomic Survey [41]. This survey identified that the proportion of the population self-reported as ill or injured in Kabwe stood at 19.1%, compared to a 9.1% national average for urban settings [41]. Although the participation in the mining industry in the population of Kabwe had declined, with only 1.7% of those economically active being employed by it in 2018 [41], groundwater contamination [44] and high levels of lead in blood across the population of Kabwe [45] had been reported as significant drivers of a more deteriorated health of the local population compared to the rest of Zambia.

1.2.1 Surveillance

The first confirmed case of COVID-19 in Zambia was reported on March 18, 2020 [43]. The Zambia Government's COVID-19 surveillance strategy, implemented in May 2020, was anchored on the continuous surveillance of confirmed and suspected case and death counts [43]. Initially, official guidance indicated that PCR testing was to be targeted to individuals classified as "high-risk" [43], given the presence of fever or a persistent cough or international travel history [43, 46]. These criteria were subsequently revised to include asymptomatic individuals and those without a history of travel or prior contact with a confirmed case [43]. Additionally, the national government established the Zambia Genomic Sequencing Consortium (ZGSC) in May 2020, in collaboration between the University of Zambia, the international implementation health research organisation, PATH, and the Zambia National Public Health Institute (ZNPHI) [47]. The implementation of these official testing guidelines, however, was forestalled by global shortages in basic laboratory commodities and operational constraints, as I will discuss in the section below.

The WHO's recommendation for implementing an FF100 surveillance system [6], was also carried out in Zambia [48]. The local FF100 study, however, only recorded patients with a COVID-19 infection confirmed by PCR in the capital, Lusaka and did not include a complete clinical characterisation of cases or their outcomes [48]. A series of subsequent clinical case cohort analyses from tertiary care hospitals in the country revealed that the overall in-hospital mortality of COVID-19 patients was high, at 42.6% [49], that triaging protocols prioritised the more severe cases for admission [50], and that most pandemic-related deaths actually occurred in the community, and were thus not being recorded by the national COVID-19 surveillance programme [51–53].

1.2.2 SARS-CoV-2 waves and pandemic response

The time period of the study I conducted in Kabwe, Zambia spanned from March 1, 2020 to September 30, 2021 (Chapter 4). This period encompassed three distinct pandemic waves, known to have been caused in succession by the Wildtype, Beta and Delta variants (Figure 1.2) [54]. Although the ZGSC only conducted whole genome sequencing of 300 samples in its first year of operations [55],

a pooled analysis of their samples and those from other African countries revealed the first wave in Zambia resulted from repeated importation events from Europe and South Africa, notably between January and May 2020 [56].

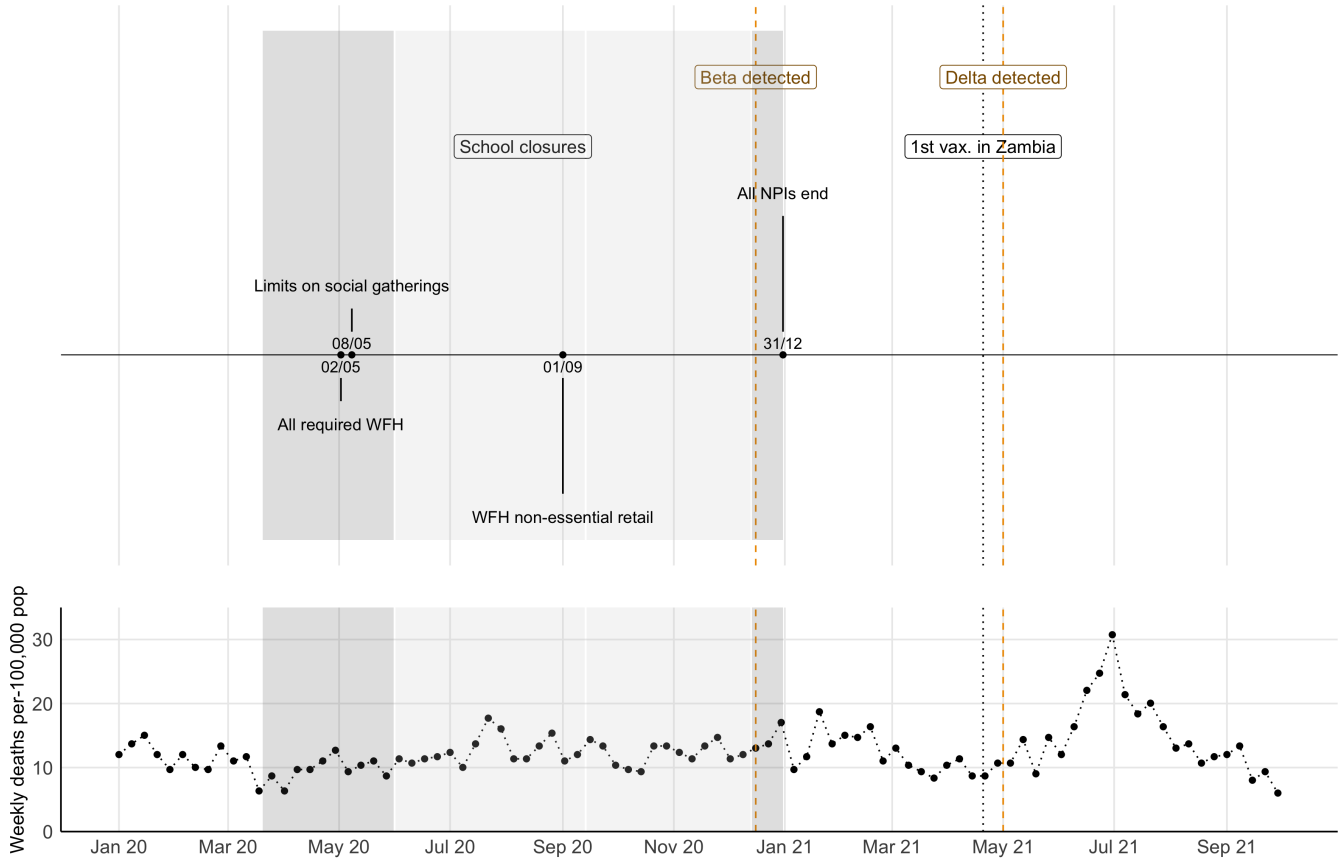


Figure 1.2: SARS-CoV-2 waves in Kabwe from March 1, 2020 to September 30, 2021. The landscape of NPI policies varied across waves, which were dominated, in succession, by the Wildtype, Beta and Delta variants. Official policies on school closures included either all educational sectors (dark grey areas) or basic education only (light grey areas). The precise dates of Beta and Delta VOC introduction are unknown; genomic analyses with a first detection of Beta was done on pooled samples from December 16-23, 2021 [57], and for the case of Delta from early May 2021 [54] (yellow lines and labels), at which point several samples tested positive for the respective VOCs. Officially confirmed cases and deaths amounted to less than a 100 across the study period, and hospital surveillance was not routinely conducted in Kabwe district (no official source - from personal communication whilst conducting my study in Kabwe, Zambia). For contextual reference, bottom panel shows the number of all-cause deaths per-100,000 population, as reported from paper-based mortuary records to the ZNPHI [52].

The first set of NPI policies implemented nationally in Zambia came into effect on March 20, 2020, only two days after the detection of the first case in the country (Figure 1.2) [48]. These included the closing of schools and workplaces, except essential businesses, of international ports of entry to all but essential travel, the limitation of gatherings to less than 100 people, and a strong recommendation to stay at home [58]. These NPIs were estimated to decrease social contact in Zambia by 52% between March and May 2020 [58]. However, they were soon lifted and, by May 31, schools and workplaces had been reopened, with social contacts rapidly returning to pre-pandemic

levels by June 2020 [58], ahead of the peak in official cases and deaths during the first local pandemic wave [51, 52].

A lack of materials to conduct diagnostic testing throughout the African continent made it operationally unfeasible to implement population-wide surveillance in Zambia [46, 59–62]. Amidst the first local wave, testing was limited to international points of entry, contact tracing of confirmed cases with a history of international travel and (sparsely) confirmatory testing in hospital settings in the capital, Lusaka [59–62]. This lack of surveillance data in Zambia and the region gave rise to the false notion that the severity of COVID-19 in the African continent was somehow paradoxically low compared to what was being reported in HICs [60–62].

It was only in the aftermath of the first wave that a single cross-sectional study emerged to shed some light into the likely true severity of the local COVID-19 pandemic in the setting of LICs in the southern Africa region [60]. This study, conducted at the University Teaching Hospital of Lusaka, highlighted that contrary to what the myth of the 'African COVID-19 paradox' suggested, most deaths that occurred during the local first wave were indeed linked to COVID-19 and, furthermore, the majority of deaths were happening outside of health facilities and, thus, not recorded by official surveillance systems [60].

This report was met with great discontent from national public health officials in Zambia [63]. However, a cross sectional seroprevalence study conducted in six districts in Zambia, Kabwe included, in July 2020 (before the peak of the first wave) found the ratio of official cases to estimated infection prevalence was of 1:92 [46]. Additional cross-section studies soon followed, all pointing to the same conclusion: that official figures gave a vast underestimation of the true burden of the local pandemic [46, 51, 59, 61, 64].

By the time I began my PhD, there had only been one other representative seroprevalence study in the country. This was the TREATS COVID study, a nested study within the larger Tuberculosis Reduction through Expanded Antiretroviral and Screening for Active TB (TREATS) study in the Bwacha-Ngungu community of Kabwe, Zambia between November 2020 and February 2021 [59]. In this study, it was found that local SARS-CoV-2 seroprevalence increased from 2.8% in December 2020 to 33.7% by February 2021, roughly coinciding with the start and end, respectively, of the second pandemic wave in Zambia [59]. The TREATS COVID study included an additional survey round in the same population between June and July 2021. This had been unpublished until my collaboration and subsequent publication with Zambart and ZNPHI during my PhD [65], so I will reserve the discussion of its findings to Chapter 4.

The space of time over which the above three pandemic waves unfolded in Kabwe (Figure 1.2), was characterised by a low enforcement and transient adherence to NPIs [58,66]. Less stringent NPIs were put in place in response to the second wave (Figure 1.2), alas by December 31, 2020 all official

NPIs had been lifted [43,58]. This was further met by severe constraints in healthcare capacity [50,67], with in-hospital mortality later being found to increase in Kabwe (and elsewhere in Zambia) during the second and third waves [50, 51, 67] given increasing pressures on an already-constrained health system [50, 67].

Additionally, the roll-out of the COVID-19 vaccination programme in Zambia suffered from severe delays, a limited supply of doses and low initial uptake [28, 68, 69]. The first doses of COVID-19 vaccines arrived in Kabwe in April 2021, and were used to target frontline healthcare and related support workers [70]. The coverage with a first dose of an approved COVID-19 vaccine at a national level, thus, stood at a meagre 1% of the general population by the end of May 2021, as the third wave by the Delta variant started [50]. An efficient roll-out of vaccination at the population-level did not commence until December 2021 [68], by which point only 4.5% of the national population had received a first dose of an approved vaccine [50], the Omicron BA.1 variant had been detected in Zambia and a fourth wave was ensuing [54]. It took until November 2022 for the national vaccination programme to achieve a 70% coverage with two doses among the eligible population of Zambia [68], with Kabwe being the first district to achieve this two months earlier.

1.3 Transmissibility and severity analytics

The use of analytic methods to infer transmission and severity dynamics of a pathogen is of critical public health relevance during an outbreak. Many methodological developments pioneered by academic groups during the pandemic allowed for a detailed understanding of the public health implications of the SARS-CoV-2 pandemic [3, 4]. As I highlight above (section 1.1 and section 1.2), where available, these analytics were an angular component of the epidemiological intelligence informing the COVID-19 response of governments and public health agencies [3, 42, 71].

In this section, I summarise key innovations made in the field of methods to infer transmissibility and severity dynamics, in context with what they meant for assessing COVID-19 control interventions. Yet, as I will argue in Section 1.6, no approach by the start of my PhD had been capable to systematically account for the specific and combined effect of changing contact rates, variants, immunity and healthcare landscape in driving the transmission and severity dynamics of SARS-CoV-2.

1.3.1 Transmissibility

The transmissibility of infectious disease outbreaks has long been studied with estimations of their basic R_0 and time-varying reproduction numbers R_t. From basic mathematical ecological principles R_0 can be expressed as

$$R_0 = exp(rG) \tag{1.1}$$

where r is the growth rate of a free-living organisms and G its generation time distribution, or time from the birth of one organism to the birth of its offspring [72]. When adapted to the field of infectious diseases, R_0 takes on a meaning of public health relevance. It represents the average number of individuals that will acquire an infection from a single infected case in a large fully susceptible population [73]. Thus, a pathogen with $R_0 > 1$ will be likely to establish in a population at which point an outbreak could ensue, whereas one with $R_0 < 1$ is unlikely to cause an outbreak [73].

The R_t on the other hand is commonly used to express pathogen transmissibility over the course of an outbreak [74,75]. It represents the average number of secondary infections each primary case generates at time t [74], which can vary in response to the emergence of pathogen variants, changes in contact patterns (e.g. NPIs) or population immunity. Similar to the above, an $R_t > 1$ implies ongoing transmissibility and epidemic growth, and $R_t < 1$ decreased transmissibility and possible epidemic control [18]. Several methods for estimating R_t in near real-time rely on the use of a branching process renewal equation relating present time new cases I_t to past ones I_{t-s} as the result of a Poisson process [75]

$$I_t \sim Pois(R_t \sum_{s=1}^{t} I_{t-s}\omega_s) \tag{1.2}$$

were the sum of past cases up to the immediately previous time point is weighted by the serial interval distribution ω_s [75].

Such methods offer the convenience of minimal data and parameter needs, ease of implementation through several open-source packages, and robustness to constant case underreporting [74, 75]. Given the many unobserved social and biological nuances that drove the transmissibility of SARS-CoV-2, however (Section 1.4), real-time analytics to inform public health responses demanded innovations to infer how R_t varied given

- Changes in contact rates, such as in response to NPIs
- Population immunity as conditioned from prior infection, vaccination or both
- Emergence of variants

Some expansions to renewal equation methods were undertaken but, predominantly, transmission model-based analytics were heavily employed for this. The latter require considerably higher technical capacity to develop and are data- and computationally-intensive. Yet, they offered unique possibilities to explicitly account for the aforementioned aspects on R_t and, thus, assess the impact of control interventions in near real-time.

Contact rates and NPIs

The study by Flaxman et al. (initially not peer-reviewed) in March 2020 provided one of the earliest real-world indications of the effect of lockdowns and other NPIs on R_t [16]. Their estimates were obtained by applying a renewal equation model to reverse engineer the expected epidemic curve of infections in eight European countries, given officially reported deaths and accounting for the NPIs that had been in place at time t when infectious should have occurred. Their approach introduced several assumptions, raising the possibility of miscalculating the infections curve and the effects of NPIs. Notably, they assumed representativeness of deaths data, which had known systematic reporting delays and varying COVID-19 death definitions across countries; they did not account for differences in country-specific population age distributions and assumed homogeneous mixing; and assumed that the sequential effect of NPIs on R_t was multiplicative and instantaneous (piece-wise constant parametric form) rather than gradual. Despite these limitations, their study strongly demonstrated that lockdowns had been the only early-pandemic interventions to consistently and decisively drive $R_t < 1$ [16].

The age structure and social mixing patterns of a population are critical determinant factors of R_t [76] and, thus, heavily weight on infection attack rate dynamics [77] and the likely impact of control interventions [78]. Robustly capturing these dynamics required the parameterisation of age-stratified models with social mixing matrices (Figure 1.3), to allow non-homogenous interactions across age groups [79]. Social mixing matrices, however, are notably hard to estimate. They require extensive population surveys to accurately record daily contacts across a representative period of time [80]. Pre-pandemic, the POLYMOD social contact survey, conducted across eight European countries [76], had been extrapolated across 152 countries to provide the largest global repository of mixing matrices [81]. During the pandemic, the CoMix study in the UK and an additional 21 countries was the largest social contact survey of the COVID-19 pandemic [82].

Both the POLYMOD and CoMix social mixing matrices were extensively employed, either alone to estimate how contact rates changed in response to NPIs [20, 82–87], or within transmission modelling frameworks to infer R_t as a function of time-varying next generation matrices (NGMs). I will reserve a technical description of this approach, which underpinned my own PhD analyses, to Section 2.2.3. Briefly, an NGM defines the non-linear dynamics between current infections at time t and expected number of secondary infections across age classes, given contact rates (Figure 1.3) between them [88]. Mathematical models that followed this approach were shown to far outperform those without age-disaggregation or without an explicit accounting for non-homogeneous social contacts when now-casting and forecasting expected SARS-CoV-2 trajectories in real-world settings [79]. The use of social contact data and NGMs in pandemic analytics was thus extremely valuable, not only from an academic perspective, but also to inform NPI strategies by governments and health

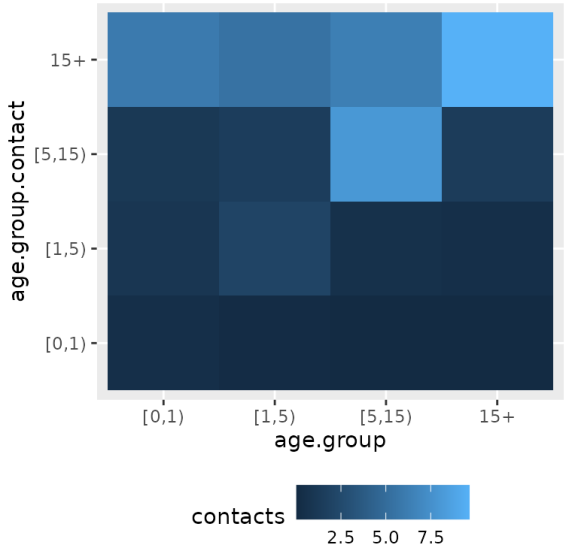


Figure 1.3: Example of a symmetric contact matrix with four age ij (rows and columns, respectively) classes, obtained from the R package `socialmixr` [80]. In the context of compartmental models, the force of infection acting on a susceptible individual of age i is determined by the product of the ij relative infectivity (per-contact probability of infection) and the sum of the ij contact rate-weighted prevalence of infection [72].

agencies across may settings globally [18, 32, 33, 79, 86, 87, 89–98].

Population immunity

The SIREN study from the UK, one of the largest front-line health worker COVID-19 infection cohorts, provided early evidence that protection against re-infection was imperfect (i.e. break-through infections amongst individuals with positive antibodies) and relatively short-lived, with re-infection episodes seen from as early as 95 days after a first infection episode [99]. Concurrently, preliminary results from phase 2/3 vaccine trials showed that break-through infections also occurred after vaccination and that protection seemed to peak and then wane with time [100–104].

The use of age-stratified transmission models, again, proved invaluable to account for the imperfect and waning nature of both types of immunity and their effect on SARS-CoV-2 transmissibility. An innovative approach we proposed at the ICL-RTM team during the first pandemic wave was the estimation of R_t with and without the effect of immunity [18]. I will describe this technical process in Chapter 2, as it was also central to my PhD analyses. In brief, the aim was to quantify the effect of changing contact rates (from NPIs) and that of accruing population-level immunity on R_t independently. This allowed us to estimate that by December 2020 immunity levels in the population were low and, therefore stringent NPIs were still needed in the absence of mass vaccination [18].

Similar findings were echoed by other modelling studies, within [89, 105–107] and outside [108] the UK. The capacity of these modelling frameworks to account for the effects of infection- and vaccine-derived immunity was critical to estimate the impact of the lifting of NPI restrictions during

England's roadmap out of lockdown, and in context with the simultaneous roll-out of the national vaccination programme [32,33]. Elsewhere, this also allowed the simulation of impact across several vaccination scenarios [108], which were ultimately pivotal to inform the WHO's global vaccination strategy [109].

Variant emergence

The emergence and co-circulation of variants, with different intrinsic mutations altering their transmission and severity profiles, became a hallmark of the pandemic [25,110]. Genomic sequencing data was critical to delve into what these lineage selection mechanism entailed in terms population-level dynamics of the pandemic [110].

Several approaches were used to analyse genomic data, including generalised (multinomial) models [30, 96], molecular clock analyses [25, 31, 110], expanded renewal equations [111] and multi-strain transmission mathematical models [32, 33, 96]. Powered by an early access to population-representative data from COG-UK [3], UK modelling groups again spearheaded in the development of the latter approach.

This posed several technical challenges for which simplifying assumptions had to be made. Namely, around the efficacy and duration of cross-protection (i.e. protection against infection from other variants recovery from infection by one), the efficacy of vaccine-induced protection, hybrid immunity (interaction, or lack thereof, of infection and vaccine derived immunity) and, for all types of immunity, accounting for efficacy against explicitly simulated clinical endpoints (e.g. infection, symptomatic disease, severe disease, hospitalisation, ICU admission, death, onward transmission). Accounting for the above was critical, as the aim of these models was to allow the monitoring of transmission of two or more co-circulating variants, in near real-time, amidst an evolving landscape of NPIs and population immunity [32, 33, 89, 96].

Regardless of the approach to modelling transmission of the variants themselves or set of assumptions made, all models found that variants that dominant at any one point outcompeted others given a transmission advantage [32, 33, 89, 96]. In principle, this meant that the R_t with and without immunity specific to a dominating variant had to be higher than those of the outcompeted one(s). Furthermore, the latter implied their R_0 would had to be intrinsically higher in sequential order or replacement/dominance. Yet, these hypotheses could not be proven at the time, given model-structure and mathematical inference limitations in the frameworks used [32, 33, 96]. I will circle back to these aspects in Chapters 2 to 4, as addressing them was a central aspect of my research.

1.3.2 Severity

A commonly estimated metric of severity of an emerging pandemic pathogen is the case fatality ratio (CFR), calculated as [112]

$$\mathbb{E}(CFR) = \frac{\sum_{i=0}^{t} d_i}{\sum_{i=0}^{t} c_i} \tag{1.3}$$

where t represents the last day with observations and d_i and c_i represent the incident number of deaths and cases, respectively, at time i. However, real-time surveillance of cases and deaths is particularly affected by underascertainment, overdispersion (e.g. weekend effect, relatively small population) and, crucially, right censoring (ignoring unknown outcomes of active cases at time t) [112, 113]. These biases can be, at least partly, corrected for using parametric statistical models, such as generalised mixture (cure) [114–116] or competing hazard models [112].

However, neither of these traditional statistical methods had been able to address a further layer of estimation error, which comes from the inherent underascertainment of the true size of an outbreak. That is, case and death count-based surveillance reflects only one, of potentially many, conditional severity pathways whereby death is conditional on symptomatic case detection, which is conditional on infection itself [2,117]. This is particularly problematic during outbreaks by pathogen that cause a high (or unknown) proportion of asymptomatic or mild infections, when case ascertainment changes over time (e.g. new, more virulent variants emerge, changes in testing and surveillance policies), or when comparing severity of the same pathogen across settings with different healthcare and/or surveillance provisions [2].

All the above factors were relevant to the SARS-CoV-2 pandemic and, thus, to monitoring its severity over time. During the earliest stages of the pandemic, In early March 2020, Verity et al. [2] used infection prevalence data among repatriated expat residents in Wuhan and time-matched local cases and deaths data to provide the earliest estimates of the CFR and IFR of COVID-19. They found a strong indication that the severity of the pandemic was age-dependent, with older individuals being at a substantially higher risk of death [2]. Despite their data limitations and assumptions of its representativeness of the underlying transmission and disease severity pathway dynamics, their estimates were further corroborated by several representative infection prevalence studies that followed.

Of note, O'Driscoll et al. [118] estimated the COVID-19 cumulative infection attack rates across 45 countries (HICs and U-MICs predominantly, with Kenya and South Africa being the only two sub-Saharan African countries represented). They found that, whilst there were significant difference in the overall IFR by country, there was less variation in age-specific IFR estimates [118]. Limited

by likely discrepancies in the underlying datasets and parametric assumptions around the delay distribution of time from infection to death and seroconversion [118], they were unable to corroborate the driving factors behind the trends they observed. It was ultimately clarified by Brazeau et al. [119] that accounting for seroreversion (i.e. negativisation of a serological test due to diminishing antibody titres) and variations in the infection attack rates across geographies, which impact the specificity of serological tests, was crucial to derive robust IFR estimates in the context of COVID-19 [119]. They thus found that differences in the overall IFR across countries were explained by differences in the age distribution of the population [119]. Although it should be noted that, in contrast to O'Driscoll et al. [118], this conclusion was drawn by analysing data from the first wave of the pandemic in 10 countries only, all but one (Brazil) of which were HICs [119].

The age-specific COVID-19 IFR estimates from Brazeau et al. [119] were, nevertheless, the most robust and nuanced obtained during the course of the pandemic up to that point. They were directly employed to drive the mechanistic processes of age-specific severity in several SARS-CoV-2 mathematical modelling studies across hundreds of countries [90–93, 120, 121]. What became evident from these and other transmission modelling studies that tracked the pandemic, along with its severity, over prolonged periods of time was that the IFR was highly dynamic. It varied in response to intrinsic and effective drivers of immunity and healthcare factors [32, 33, 89, 106]. Given the complexity of these factors, and the sheer number of them at play, the development of advanced age-structured, multi-variant mechanistic models by UK-based academic groups was particularly relevant during the COVID-19 pandemic.

1.3.3 Age-structured, multi-variant compartmental models of SARS-CoV-2

A key scientific contribution of my research during my PhD was developing a novel parametric framework to systematically unpick and quantify the specific and joint effect of variants, immunity from vaccines and prior infection, NPIs and healthcare changes in driving the transmission and severity of the pandemic over time. This section focuses on describing the high-level approach employed by other academic groups in the UK to examine these factors within compartmental SARS-CoV-2 modelling frameworks.

Among the models employed by academic groups contributing to inform the UK Government's SPI-M and SAGE during the pandemic, only a handful were age-structured, multi-variant compartmental models [33, 106]. Previous to these and my model, it had not been possible to quantify the specific effects of different health interventions on the transmission and severity of the virus over time, in context with the emergence of different variants. That is, the decisions to increase the complexity

of our models were sequentially informed by the real-time surveillance and analytics challenges of the pandemic *ad hoc*. Moreover, the type and nature of the surveillance data required to perform inference on was not widely available across many settings (e.g., representative variant genotype from infections in the community to inform transmission advantage inference).

I specifically focus on the models described by Barnard et. al. 2022 [106] and Keeling et. al. 2022 [33], as they focus on a similar time period of the pandemic in England and employ the same or comparable surveillance data sources as I did for my analysis in Chapter 3. An additional noteworthy example was the model by Birrell et al. [122], which structure and complexity evolved as the need to account for vaccination and waning immunity, among other nuances, demanded it. This model, however, did not include the explicit modelling of different co-circulating variants, and so I will not comment on it further in this section.

For their 2022 analysis of the COVID-19 pandemic in England, Barnard et. al. [106] expanded on a deterministic, compartmental SARS-CoV-2 model and differential evolution Markov chain Monte Carlo (MCMC) inference framework to apply continuous-time multipliers of disease transmission and severity. In particular, they implemented a 3-variant model where given time-varying contact matrices and fitted introduction dates and relative transmission advantage of the variants (over the preceding variant), they then fitted a stepwise transmission multiplier with 5-day step changes. Similarly, they estimated baseline (given model fit to data) IFR, IHR and infection critical-illness rate (ICR) parameters, which were then scaled over the course of the pandemic given three factors: a) fitted stepwise change multiplier functions, b) variant-specific relative hospitalisation, ICU and death ratios (compared to the previous variant) and c) vaccine protection.

The model by Keeling et. al. [33], on the other hand, evolved in complexity along the crucial steps of England's roadmap out of lockdown (Section 1.1). Specifically, it evolved from an age-stratified model, to a multi-variant and multi-vaccine class one with accounting for waning immunity and vaccine boosting. The framework was also deterministic and employed MCMC to fit several key parameters, including of increased transmissibility of the variants (i.e., in order to recover data on variant frequency from S-gene target failure time series [33, 105]. Importantly, whilst their model accounted for age-specific probabilities of symptomatic and severe disease, it did not implement mechanisms to infer (or assume) changes in severity with the emergence of the simulated variants, and maintained an assumption of no waning immunity after infection [33].

As my discussion of methods in Chapters 2 and 3 and in Appendix A will show, there are similarities in the mechanisms to capture variant transmission and severity in the work I did during my PhD and the above two approaches [33,106]. There are, however, several differences to note too. Namely

- Barnard et. al. [106] fitted stepwise multiplier functions and variant-specific transmission and

severity characteristics, which I also did. They, however, did this to improve the recovery of (fit to) epidemiological time-series [106], as opposed to a hypothesis driven exercise to identify the individual and combined contribution of variants and other factors to model-observable transmission and severity.

- Relatedly, I implemented parametric calculations to infer the intrinsic and effective (see next section) transmission and severity of SARS-CoV-2 variants, which leveraged the above fitted stepwise multiplier functions, variant-specific severity multipliers and time-varying contact matrices. These were not performed in either of the above two studies.

- My model was stochastic, as opposed to deterministic as in the case of both of the aforedescribed ones [33, 106]. Deterministic equations are particularly suitable to describe epidemic dynamics with large population numbers and during periods of high levels of infections [123]. The stochastic, state-space framework I employed however offered several advantages, including a wider exploration of model uncertainty during episodes of low infection rates and/or population numbers in compartments relating to observable data streams. I will discuss these aspects in context in the following chapters.

1.4 Intrinsic vs effective transmissibility and severity

As I have alluded to, many epidemiological drivers influence the transmissibility and severity dynamics of pathogens over the course of an outbreak. All these factors will ultimately act in combination to determine how an outbreak pathogen transmits in a population, how much of a threat it poses to population health and health systems, and ultimately how it becomes 'observable' through epidemiological surveillance. Surveillance data, in turn, informs epidemiological analytics which can be used to provide insights into the drivers of outbreak dynamics and assess the likely effects of control interventions (Figure 1.4).

For the purpose of my PhD work, I defined the intrinsic drivers as those that directly relate to the transmissibility and pathogenicity mechanisms of an infectious organism. Evolutionary mechanisms of single-stranded mRNA viruses, as is the case for SARS-CoV-2, are known to naturally select lineages with a transmission advantage [124]. In addition, mRNA coding errors are also common, which can lead to alterations in proteins targeted by the host's immune response and/or alter its pathogenicity (disease severity) on the host [35, 36]. When an outbreak occurs, for example with the emergence of SARS-CoV-2 in 2019 [125], the intrinsic transmissibility and pathogenicity mechanisms of the microorganism are modified by population-level factors, which I defined as effective drivers.

I specifically considered three effective drivers of transmissibility and severity of critical impor-

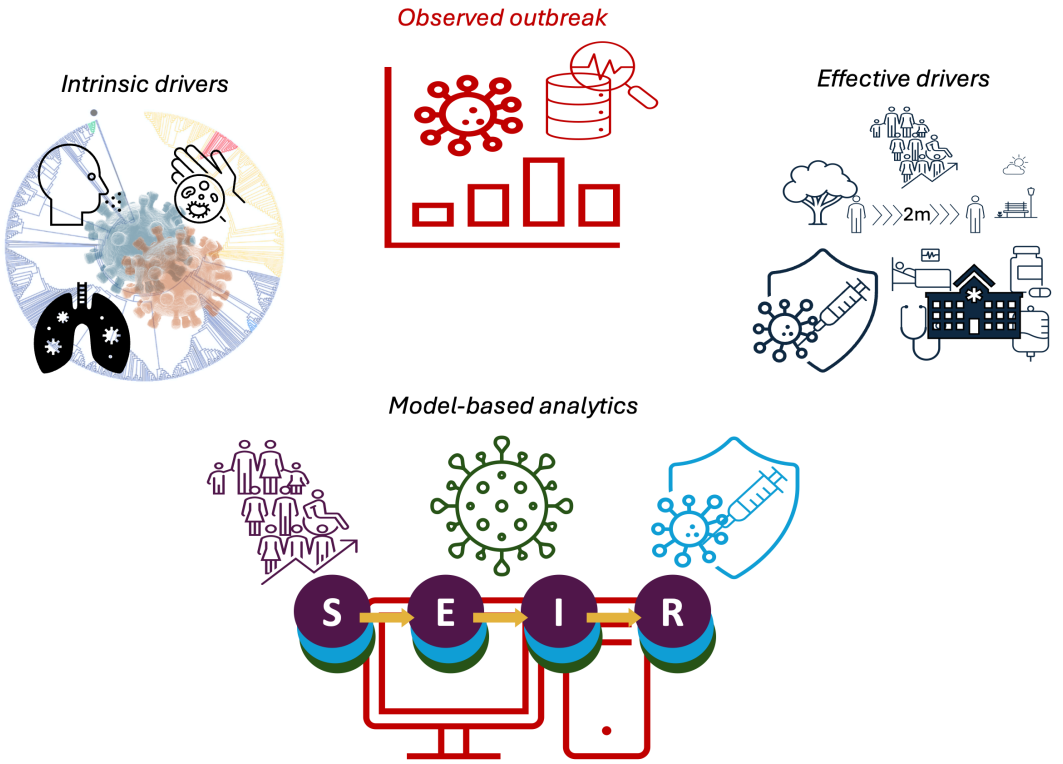


Figure 1.4: Conceptual representation of the intrinsic *vs* effective drivers of the transmission and severity of an outbreak pathogen. Compartmental transmission models are one type of analytics that can be used to assess outbreak dynamics in a population, along with key characteristics of the latter to infer the effects of intrinsic and effective drivers.

tance for outbreak response strategies. Namely, immunity, contact rates and healthcare. Over the course of the COVID-19 pandemic, immunity against SARS-CoV-2 was conditioned by the efficacy and duration of protection against its transmissibility and severity characteristics conferred by either prior-infection, vaccination or both [113]. Transmissibility further changed in response to modifications in social contacts, either as a result of the adoption of precautionary behaviours or (social-distancing) NPIs [15, 16, 78]. The capacity and quality of available healthcare was context-specific and highly dynamic, for example as a result of surges in demand and/or creation of excess capacity during outbreak peaks [92, 126, 127] or with the adoption of improved treatment protocols [128].

Understanding the specific and combined effect of the intrinsic and effective drivers is a critical step towards assessing how they can be leveraged to modify the transmissibility and severity of an outbreak [71]. In turn, timely having and acting on this knowledge is a cornerstone of the planning and implementation of public health response interventions [71, 117]. Model-based epidemiological analytics have been, and will remain, central to triangulating surveillance data, infer how unobservable social and biological process drive the transmissibility and severity dynamics of an outbreak and, ultimately, assess the likely efficacy of control interventions [117].

1.5 Pandemic inequality

The need for robust epidemiological intelligence, anchored on timely and reliable surveillance and analytics, to inform response interventions during an outbreak emergency is universal [71]. The SARS-CoV-2 pandemic drove unprecedented surges in excess healthcare demand and mortality worldwide [129]. It was, however, in the most resource constrained settings that its impact was felt the hardest, fuelled by the many inequities and inadequacies of the global pandemic response mechanisms [42].

Studying inequalities is a nuanced endeavour, and one that requires, at minimum, clarity on the study focus, metrics and comparator groups, all of which were outside the scope of my PhD. Yet, it is perhaps all this complexity that made it very difficult to gauge, particularly in the early stages of the COVID-19 pandemic, the disproportionately higher burden borne by resource constrained settings. My comment in this section centres around highlighting how inequality in a capacity to conduct surveillance and analytics ultimately impacted other areas of the pandemic response, in the context of African LMICs.

Global pandemic response financing mechanisms imposed grave delays and limitations to implementing local surveillance, analytics and mitigation interventions in low- and middle-income countries (LMICs) [42]. A global shortage of molecular diagnostic and other essential laboratory commodities at the start of the pandemic meant that available tests had to be triaged to confirm suspected diagnoses or screen higher-risk individuals. Although great efforts were directed to establish global COVID-19 surveillance data repositories [28, 130, 131], which soon became a trusted source of information, these collated official figures of cases, deaths and, in some countries, hospitalisations and vaccinations. Whilst most HICs had the capabilities to establish large-scale universal screening programmes, most LMICs did not [42]. Indeed, by September 2020, before the known widespread circulation of more transmissible and severe VOCs, countries like Luxembourg had been able to implement testing at a staggering capacity, reaching a cumulative 1,023 COVID-19 test conducted per-1,000 population [28]. At the same time, South Africa, the country in the region with the highest testing capacity, had conducted a cumulative 62 tests per-1,000 population and Zambia 6 tests per-1,000 population [28].

As I have alluded to in the sections above, data is paramount to inform modelling analytics. A systematic review of applied epidemiological modelling in the African continent found that, as of April 2021, there had only been 74 SARS-CoV-2 modelling studies in Africa, of which the majority had not been peer-reviewed [132]. Critically, only nine of the available studies had used local data for model parameterisation, whilst most have relied solely on confirmed cases data (from official sources) [132]. The lack of timely surveillance data and analytics, particularly during the first wave, thus gave

way to the stipulation of the 'African Paradox' (Section 1.2) [60, 61]. In turn, the dependency of many LMICs, in Africa and elsewhere, on global pandemic response financing conditioned a lack of coordination of local and international response mechanisms [42].

The median age of the sub-Saharan Africa region is the lowest for any region in the world [42,62] and importation of the virus into the region was mainly via Europe, thus the start of the local pandemic lagged behind that of HICs [56]. Both factors could and should have been plausible explanations for a markedly lower impact of the pandemic in Africa, but it was not. It is now known that the 'paradox' was merely a lack of timely data and analytics and that, instead, the COVID-19 pandemic drove a disproportionately higher burden [42, 61, 92, 129]. Rather soberingly, pandemic inequalities continued to be the norm even after the paradox was debunked and, as HICs were rolling out booster vaccines in late 2021 (section 1.1), the likely true impact of the pandemic remained largely unknown across many LMICs [42, 93, 133] and, in the most constrained of settings, like Kabwe, the coverage with first doses was still negligible (section 1.2).

It was at this point that I began my PhD. The focus of my research was informed not only by what I had learnt about the limitations of current-day analytics to infer the effect of intrinsic and effective drivers of transmission and severity dynamics. As part of the ICL-RTM team and a national of an LMIC, I was poignantly aware that the many successes we had in using real-time analytics to inform the pandemic response of the UK had been facilitated by the timely availability of a wealth of surveillance data. This was, in itself, a limitation for the utility of our methods in settings with lesser surveillance capacity. The WHO's Pandemic Preparedness and Resilience for Emerging Threats initiative prescribed member states should formulate their preparedness plans and priority actions informed by their learnings from the COVID-19 pandemic [134]. Achieving these preparedness goals towards a more effective and equitable global response to future pandemics, however, requires strategies underpinned by epidemiological learnings from across all settings, regardless of income or surveillance data constrains [65].

1.6 Thesis aims and objectives

1. To develop a novel parametric method to quantify the **intrinsic** and **effective** drivers of transmissibility and severity of an outbreak pathogen, generalisable to age-disaggregated compartmental transmission modelling frameworks (Chapter 2).

2. To conduct a granular case study of the SARS-CoV-2 pandemic in England, with multiple competing intrinsic and effective drivers of its dynamics over time (Chapter 3).

   (a) Quantify and compare the intrinsic transmissibility and severity of the Wildtype, Alpha,

Delta and Omicron BA.1 variants

(b) Assess the specific contributing factor on the effective transmissibility and severity, overall and by said variants, of the backdrop of changes in

i. NPI policies between March 2020 and February 2022

ii. Prior infection-, vaccine- or hybrid-derived population immunity profile

iii. Healthcare ecosystem conditions by changes in treatment protocols and/or surge critical care demand

3. To conduct a case study of a parsimonious adaptation of my modelling framework to limited surveillance data from an LIC setting: Kabwe district, Zambia

(a) Quantify the likely true burden of COVID-19 in a setting with a near-unmitigated pandemic

(b) Test the utility of my approach to infer effective transmissibility and severity dynamics with limited surveillance data

(c) Assess its capacity to provide an indication of intrinsic severity, in the absence of variant data

4. To assess the effects of a limited and delayed access to improved healthcare and vaccines on pandemic burden in Kabwe

(a) Simulate counterfactual pandemic trajectories, had Kabwe had implemented vaccination and healthcare improvements in line with what was achieved by higher in come settings

(b) Assess the cost-effectiveness and NMB of such interventions, against plausible values of the ocWTP threshold of the local health system

5. To assess pragmatic nuances of policy decisions around the adoption of CEPI's 100-day mission in the context of Kabwe

(a) Perform an extended dominance analysis of a comprehensive range of mutually exclusive, plausible local implementations of the mission, either without delay or at quarterly delay windows up to a year later

(b) Assess the impact of model-driven uncertainty on choosing the optimal implementation strategy by delay window

1.7 Thesis overview

In Chapter 2, I present the rationale and methods the novel parametric approach I proposed during my PhD. I specifically propose a parametric method to estimate the intrinsic and effective transmission and severity of a pathogen within a transmission modelling framework. For this, I built on the validated mathematical model of SARS-CoV-2 transmission developed by the ICL-RTM team. The innovations I performed during my PhD were, in turn, absorbed back into the ICL-RTM model whilst it was still in active use to provide analytics for SPI-M-O and SAGE. Although this model was a bespoke stochastic, compartmental model of SARS-CoV-2 transmission, I will argue my approach is generalisable to any age- and pathogen- class disaggregated compartmental model of an outbreak pathogen. An additional requisite, should the aim of the model be to fit to and estimate epidemiological parameters from data, is the capacity to perform multi-dimensional parameter state-space inference. For this, my framework was paired with a particle Markov chain Monte Carlo (pMCMC) inference framework, also developed in-house at ICL [135].

Then, in Chapter 3, I present the implementation of my method to the pandemic surveillance data of England between March 2020 and February 2022 (Section 1.1). Given the wealth of data I had available for this analysis, I used it as a case study of the robustness of my novel approach with two objectives. Namely, to quantify and compare the intrinsic transmissibility and severity properties of four SARS-CoV-2 variants; and to assess the specific contributing factor of the backdrop of a) NPI policies, b) prior infection-, vaccination- or hybrid-derived population immunity, and c) healthcare landscape on driving the effective transmissibility and severity of the pandemic, overall and by variant.

In Chapter 4 I present the adaptation of my method to a case study of the near-unmitigated burden imposed by the COVID-19 pandemic in Kabwe, Zambia between March 2020 and September 2021 (Section 1.2). This resulted from a research collaboration I instigated and led with two Zambian organisations, Zambart, an implementation research institution, and the ZNPHI. My objectives here were to test the utility of my method in evaluating effective transmissibility and severity dynamics with limited data, and to quantify the likely true burden of the pandemic in a setting with a near-unmitigated pandemic.

In Chapter 5, I expand on the work from the previous chapter to evaluate to what extent the restricted access to vaccines and improved healthcare drove the burden of the pandemic in an LIC setting. For this, I conducted a cost-effectiveness analysis of counterfactual strategies of the implementation of vaccination and healthcare as per what was achieved by higher income settings during the pandemic.

Lastly, in Chapter 6, I further expand on my analysis of the pandemic in Kabwe with the

objective of assessing the pragmatic nuances of adopting CEPI's 100-day mission in the context of an LIC setting with severe resource constraints. For this, I performed an extended dominance analysis of a comprehensive range of mutually exclusive, plausible implementations of the mission in Kabwe, either without delay or at quarterly delays up to a year later.

1.8 Publications related to this thesis

The work presented in Chapter 2 and Chapter 3 was published in its entirety in altered form in Perez-Guzman PN, Knock E, Imai N, Rawson T, Elmaci Y, Alcada J, et al. Epidemiological drivers of transmissibility and severity of SARS-CoV-2 in England. *Nature Communications*, 2023, 14(1):4279 [1].

The work presented in Chapter 4 and Chapter 5 was partly published in altered form in Perez-Guzman PN, Chanda SL, Schaap A, Shanaube K, Baguelin M, et al. Pandemic burden in low-income settings and impact of limited and delayed interventions: A granular modelling analysis of COVID-19 in Kabwe, Zambia. *Int J Infect Dis*, 2024, 147:107182 [65].

Chapter 2

Novel parametric approach to quantify the intrinsic and effective drivers of transmissibility and severity

2.1 Background

In Chapter 1, I presented an overview of some of the many analytic methods that were used to assess the transmissibility and severity dynamics of the SARS-CoV-2 pandemic. Whilst the epidemiological insights gained with these methods informed response interventions (section 1.3), none had the inferential capacity to disentangle and quantify the effects of intrinsic*vs* effective drivers (section 1.4) of transmissibility and severity across multiple emerging SARS-CoV-2 VOCs and epidemic waves.

Herein, I proposed a parametric approach to quantifying the effects of intrinsic and effective drivers on the transmissibility and severity of SARS-CoV-2 across several pandemic waves caused by different variants. This chapter does not contain analytical results, but is a technical discussion of the rationale and methodological aspects to develop my method. For this, I built on the stochastic, compartmental model of SARS-CoV-2 transmission developed in-house by the ICL-RTM team, of which I was a member of before my PhD.

Beyond my familiarity with this model, there were considerable advantage it had over other prior or contemporary similar models which facilitated the development of my methods. Foremost, it was stochastic and was equipped with a state-space particle Markov chain Monte Carlo (pMCMC) inference framework. It was thus suitable to absorbing multiple surveillance data streams, and performing multi-dimensional parameter space inference. For convenience, I give a brief overview of the transmission model structure and parameters, and then focus on detailing the technical expansions I undertook during my PhD. Additional technical aspects of the model are summarised in Appendix A.

2.2 Methods

For reading convenience, this methods section and Appendix A provide the full details of the model's structure, parameters and inference framework, as I employed it for my own analysis of England data (Chapter 3). As I have already noted, however, I developed my methods and analysis by building on the SARS-CoV-2 transmission model developed by the ICL-RTM team, of which I was a member. During my PhD, the specific expansions I performed to this model for my analysis were:

- explicitly simulating the emergence of VOCs $j = Alpha, Delta, Omicron$, accounting for differences in their serial interval, their transmission advantage σ_j, their effect on the relative susceptibility to infection $\chi^{i,j,k}$ and relative infectivity $\xi^{i,j,k}$ (Section 2.2.3);
- implementing parametric calculations of the variant specific R_0^j, $R_t^{j,eff}$ and R_t^j (Section 2.2.3);
- overhauling the $h(t)$ functions to infer the effect of new treatments and other modifications in healthcare provision on disease severity over time (Section 2.2.4);
- implementing parametric equations to calculate the variant-specific basic infection hospitalisation ratio IHR_0^j, hospital fatality ratio HFR_0^j and infection fatality ratio IFR_0^j (Section 2.2.4);
- and implementing parametric equations to calculate the respective effective severity metrics overall, $IHR(t)$, $HFR(t)$ and $IFR(t)$, and given age, variant and vaccination classes $IHR^{i,j,k}(t)$, $HFR^{i,j,k}(t)$ and $IFR^{i,j,k}(t)$ (Section 2.2.4).

2.2.1 A foreword on pMCMC

Conceptually, a time t dynamic system can be simplistically represented as a stochastic model θ with x_t hidden states and f_θ state transition functions that relate to y_t time series observations (e.g. data) via a g_θ observation function (Figure 2.1). Such models often suffer from intractability, whereby for the sake of computational efficiency, obtaining posterior distribution and marginal-likelihood calculations of the parameters of θ is ignored, especially when dealing with non-liner, non-Gaussian models [136].

An approach proposed to overcome the computational inference limitation of dealing with multi-dimensionality in θ, known as hidden Markov models (HMM), is the use of particle filtering with marginal likelihood calculation [136]. To this end, pMCMC is the incorporation of sequential Monte Carlo (SMC), to estimate time-varying hidden state variables, and traditional Markov Chain Monte Carlo (MCMC), to sample from a high-dimensional parameter space [137]. This approach is particularly suitable to address inference problems where it is necessary to apply knowledge from a component of the system to infer the rest. For example, when knowledge is needed on the parameters

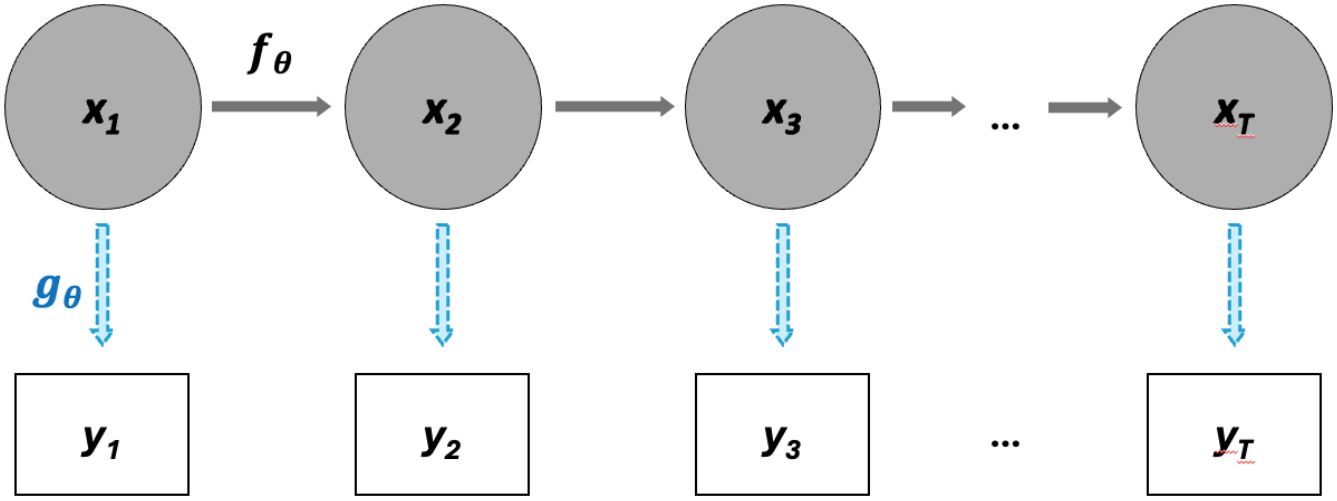


Figure 2.1: Conceptual representation of state-space models that relate to data observations via an observation function. Modified with permission from an internal ICL-RTM team training session by Marc Bageulin.

and marginal likelihoods of θ, given the HMM fit to y_t time series observations; or when the time series observations y_t are not a direct outcome of x_t states, but rather an observation process given g_θ with unknown parameters that must be estimated; or when the future state of the HMM is determined by the distribution of its states and transition parameters at current time t, without relation with previous history [138].

The in-house ICL-RTM model successfully implemented pMCMC for COVID-19 real-time analytics [135]. In my research to define a parametric quantifications of intrinsic and effective drivers of transmissibility and severity and their impact on control interventions, I made an extensive use of the above qualities of pMCMC inference. In section 2.2.3 and section 2.2.4, I present the parametric equations I developed to quantify the specific effect of intrinsic and effective drivers of transmissibility and severity. Throughout, a core capability I built into the model was inferring temporal changes in the effective transmission and severity dynamics, which further required me to define piece-wise linear parameters, whose values I inferred through pMCMC. Then in Section 3.2.2 and Section 4.2.3, I describe how I defined the observation functions that related the available data to my models, in order to reproduce the SARS-CoV-2 pandemic trajectories of England and Kabwe, respectively. Lastly, in Chapter 5 and Chapter 6, I make extensive use of the inferred hidden states and parameters at specific time points, which I then use to project a wide range of counterfactual simulations of pandemic trajectories given the implementation of public health interventions.

2.2.2 Model overview

The model is a discrete-time (1/4 day time step) stochastic, compartmental, extended SEIR-type model that is stratified by age, strain and vaccine classes, henceforth indexed as i, j, k, respectively

(Figure 2.2). Regarding age stratification, the model includes 17 five-year age bands (0-4, 5-9, ..., 75-79, 80+).

Upon infection with SARS-CoV-2, individuals enter an exposed compartment, before becoming infectious (Figure 2.2A). The model then tracks infected individuals through a series of conditional severity pathways in a sequential fashion. Infected individuals can either undergo asymptomatic/mild disease leading to recovery or symptomatic disease, which can in turn condition severe disease requiring hospitalisation or recovery without hospitalisation. For the former, individuals can either enter hospital compartments or enter a state G_D^{ijk} of severe disease leading to death in the community. Conditional on entering hospital flows (Figure 2.2B) individuals can be triaged for intensive care unit (ICU) admission or remaining in general beds throughout from where they can either die or recover and be discharged. Those admitted to ICU, can either die in ICU or be transferred for stepdown care in general wards, where they can either die or recover and be discharged. Hospitalised cases are either confirmed as SARS-CoV-2 cases upon admission or may be tested and confirmed later during their stay.

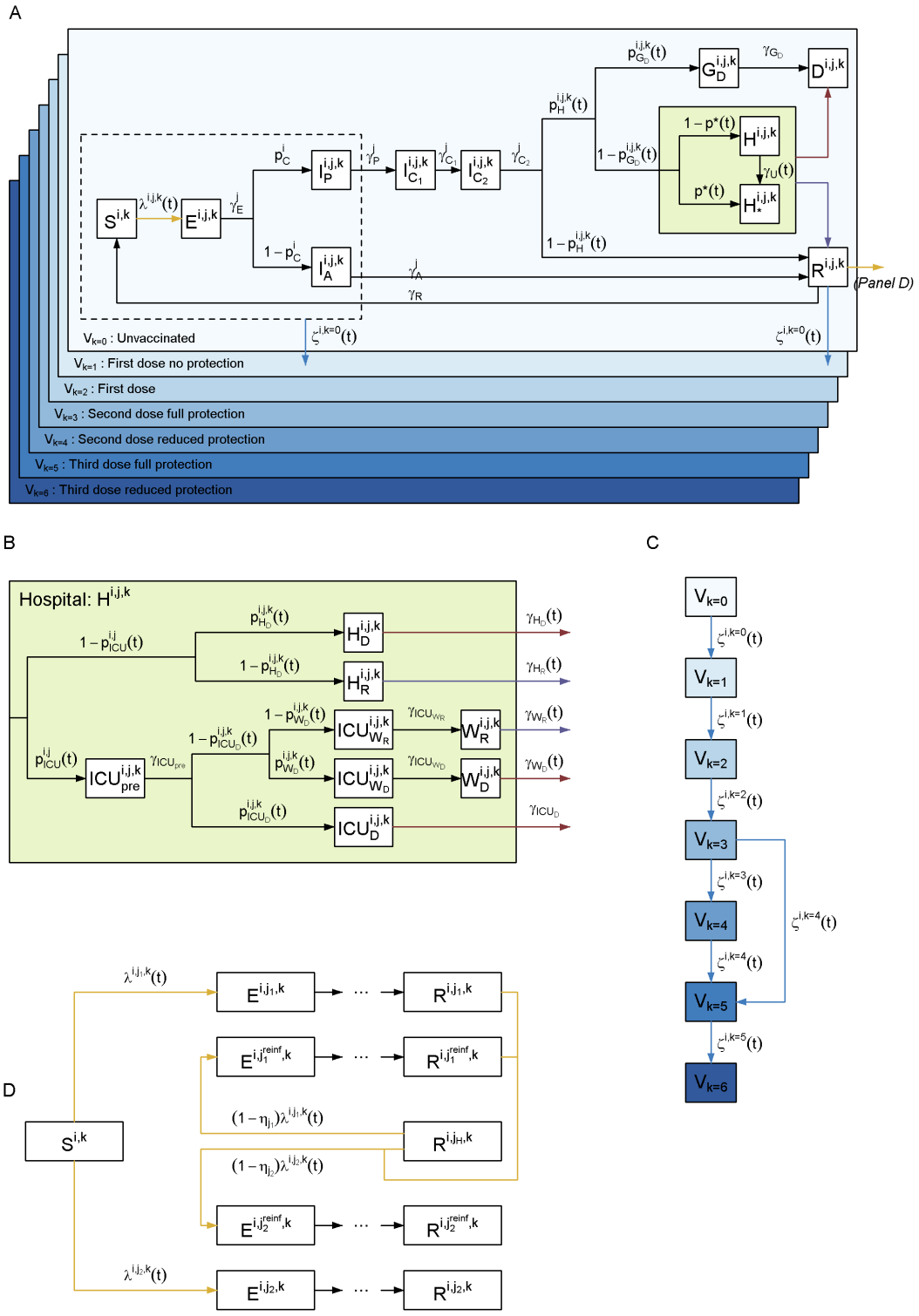


Figure 2.2: Model structure flow diagram with rates of transition between states. (A) Extended SEIR transmission model flow diagram overview. (B) Hospital flow diagram. (C) Vaccination flow diagram. (D) Multi-variant flow diagram. Definitions of state variables are presented in Table 2.1 and transition parameters in Table 2.2. Source: Perez-Guzman et. al. 2023 [1].

Table 2.1: Definitions of model compartments shown in Figure 2.2, with indices: i for age group ($i \in \{[0,5),[5,10),\ldots,[75-80),[80+)\}$), j for variant ($j \in \{$Wildtype,Alpha,Delta,Omicron$\}$ and k for vaccination strata ($k \in \{V_0, V_1, \ldots V_6\}$. Durations have Erlang (k_X, γ_X) distributions, with mean k_X/γ_X. Source: Perez-Guzman et al. 2023 [1].

		Erlang duration parameters			
Compartment	Definition	k_X	γ_X	Mean	
$S^{i,k}(t)$	Susceptible	Determined by infection dynamics			
$E^{i,j,k}(t)$	Exposed (latent infection)	2	γ_E^j	$2/\gamma_E^j$	(see Table 2.3)
$I_A^{i,j,k}(t)$	Asymptomatic infected	1	γ_A^j	$1/\gamma_A^j$	(see Table 2.3)
$I_P^{i,j,k}(t)$	Presymptomatic infected (infectious)	1	γ_P^j	$1/\gamma_P^j$	(see Table 2.3)
$I_{C_1}^{i,j,k}(t)$	Symptomatic infected (infectious)	1	$\gamma_{C_1}^j$	$1/\gamma_{C_1}^j$	(see Table 2.3)
$I_{C_2}^{i,j,k}(t)$	Symptomatic infected (not infectious)	1	$\gamma_{C_2}^j$	$1/\gamma_{C_2}^j$	(see Table 2.3)
$G_D^{i,j,k}(t)$	Severely diseased, leading to death (in the community)	2	0.67	3.00	
$D^{i,j,k}(t)$	Deceased (as a result of COVID-19)	-	-	-	
$H_D^{i,j,k}(t)$	Hospitalised on general ward leading to death	2	$0.19h_\gamma(t)$	$10.33/h_\gamma(t)$	(eq. (2.49))
$H_R^{i,j,k}(t)$	Hospitalised on general ward leading to recovery	1	$0.09h_\gamma(t)$	$10.67/h_\gamma(t)$	(2.49)
$ICU_{pre}^{i,j,k}(t)$	Awaiting admission to ICU	1	0.40	2.50	
$ICU_D^{i,j,k}(t)$	Hospitalised in ICU, leading to death	2	0.17	11.79	
$ICU_{W_R}^{i,j,k}(t)$	Hospitalised in ICU, leading to recovery	1	0.06	15.61	
$ICU_{W_D}^{i,j,k}(t)$	Hospitalised in ICU, leading to death following step-down from ICU	1	0.14	6.97	
$W_R^{i,j,k}(t)$	Step-down recovery period	2	$0.16h_\gamma(t)$	$12.22/h_\gamma(t)$	(eq. (2.49))
$W_D^{i,j,k}(t)$	Step-down post-ICU period, leading to death	1	$0.12h_\gamma(t)$	$8.06/h_\gamma(t)$	(eq. (2.49))
$R^{i,j,k}(t)$	Recovered	1	1/1095	1095	

Table 2.2: Definitions of model parameters shown in Figure 2.2. These parameters define the routes of progression through model compartments defined in Table 2.1.

Parameter	Definition
$\lambda^{i,j,k}(t)$	Force of infection.
γ_x	Rate of progression from compartment x.
$\gamma_U(t)$	Rate at which unconfirmed hospital patients are confirmed as infected.
p_C^i	Probability of being symptomatic if infected.
$p_H^{i,j,k}(t)$	Probability of admission to hospital, conditional on symptomatic infection.
$p_{G_D}^{i,j,k}(t)$	Probability of death for severe symptomatic cases outside of hospital.
$p^*(t)$	Probability of COVID-19 diagnosis confirmed prior to admission to hospital.
$p_{ICU}^{i,j}(t)$	Probability of admission to ICU, conditional on hospitalisation.
$p_{H_D}^{i,j,k}(t)$	Probability of death for hospitalised cases not in ICU.
$p_{ICU_D}^{i,j,k}(t)$	Probability of death for cases in ICU.
$p_{W_D}^{i,j,k}(t)$	Probability of death for cases after discharge from ICU.
$\zeta^{i,k}(t)$	Rate of movement from vaccine strata k to $k+1$.
η_j	Probability of being protected against infection with variant j for those recovered from earlier variants relative to those in the susceptible class.

To model heterogeneous mixing and time-varying transmission therein, e.g. given NPIs or other relevant events (e.g. school holidays), the model uses the POLYMOD contact matrix for England [76] which is scaled over time given a fitted piecewise-linear multiplier function, $\beta(t)$. The emergence and replacement of SARS-CoV-2 VOCs is modelled in a 2-strain framework whereby an established variant j_1 is replaced by an emerging one j_2 given fitted parameters of its seeding date t_j and transmission advantage σ_j over j_1. Upon recovery from infection with variant j, cross protection is not assumed to be 100% and there is waning of natural immunity (see Table 3.2), but individuals can be either (Figure 2.2D)

- Recovered from infection by variant $j = 1$, with no history of prior infection
- Recovered from infection by variant $j = 1$, with a history of prior infection by a non-circulating, historic variant (e.g. Wildtype, during Delta and Omicron BA.1 co-circulation)
- Recovered from infection by variant $j = 2$, with a history of prior infection by variant $j = 1$

- Recovered from infection by variant $j = 2$, with no history of prior infection

- Recovered from infection by variant $j = 2$, with a history of prior infection by a non-circulating 'historic' variant $j = H$

The model further explicitly accounts for waning of infection-induced immunity, assuming protection against reinfection with the same variant for an exponentially distributed duration with mean 3 years, with asymmetrical cross-immunity to a new SARS-CoV-2 variant for individuals who have recovered from previous variants (see in Table 3.2). Hybrid immunity is assumed to be a combination of the two independent effects of infection-induced and vaccine-induced protection (Appendix A.2) and waning.

Regarding vaccination classes k, the model stratifies individuals to describe a primary (two-dose) and booster regimen, and delays between receiving a dose and the onset of dose-specific vaccine effectiveness (VE), as well as waning of vaccine-induced protection following second and booster doses (Figure 2.2C). Of note, although the modelling of booster vaccinations and waning vaccine protection had not been published before my first PhD lead-author publication, this was a technical expansion done by other members of the ICL-RTM team at the time, which I incorporated into my model. I describe this process and parameters in detail in Appendix A.2.

2.2.3 Modelling transmissibility

A central objective of my work was to estimate and compare the intrinsic transmissibility of different SARS-CoV-2 variants, which I defined as their basic reproduction number R_0^j, and their effective reproduction number accounting for and void of the effect of immunity, $R_t^{j,eff}$ and R_t^j, respectively. Whilst the model had been previously equipped by the ICL-RTM team to infer the latter two effective reproduction numbers [32, 89], its focus before my PhD had been to provide rapid real-time analytics. So, simplifying assumptions had been made, including ignoring other VOCs and fitting to the emergence of Delta only and that the serial interval remained the same for Delta and non-Delta variants.

To address my PhD objective, as I discussed above, I implemented an explicit simulation of the emergence of VOCs $j = Alpha, Delta, Omicron$ accounting for differences in their serial interval, their transmission advantage σ_j, their effect on the relative susceptibility to infection $\chi^{i,j,k}$ and relative infectivity $\xi^{i,j,k}$. These nuances allowed me to, in turn, define parametric calculations of their R_0^j, $R_t^{j,eff}$ and R_t^j. This section describes these technical processes and assumptions.

Variant seeding and multi-strain model

Based on [32, 89], I seeded each variant j fixing their daily rate ϕ_j and period ν_j (Table 5.7), given a fitted emergence time t_j (Table 3.3). I assumed all seeded infections $S \rightarrow E$ occurred in the 15-19 year-old, unvaccinated group. For the Wildtype seed, I assumed a daily rate of 10 per million population-day over a 1-day period, so $\phi_{Wildtype} = \frac{1}{100{,}000} \sum_i N^i$ and $\nu_{Wildtype} = 1$. For the Alpha, Delta and Omicron BA.1 variants, I assumed a fixed rate of 2 per million population-day over a 7-day period, so for $j \in \{Alpha, Delta, Omicron\}$, $\phi_j = \frac{1}{500{,}000} \sum_i N^i$ and $\nu_j = 7$.

I initialised the model on March 1, 2020 as a single variant model, assuming all infections are due to the Wildtype variant. Subsequently, most model compartments are inflated to four classes in the j dimension: $j = j_1, j_2, j_1^{reinf}, j_2^{reinf}$ to simulate the emergence and co-circulation of VOCs. The sole exception is the R compartment where there is an additional layer, j_H accounting for individuals recovered from historic variants. For the analyses I performed in Chapter 3, I accounted for three two-variant phases:

1. From September 17, 2020, $j_1 = Wildtype$, $j_2 = Alpha$, $j_H = \{\}$
2. From March 8, 2021, $j_1 = Alpha$, $j_2 = Delta$, $j_H = \{Wildtype\}$
3. From November 1, 2021, $j_1 = Delta$, $j_2 = Omicron$, $j_H = \{Wildtype, Alpha\}$.

In addition to a fitted seeding date t_j, I fitted parameters of variant transmission advantage σ_j over the immediately preceding one (Table 3.3), as detailed in the section below.

Force of infection

Based on [32, 89], I defined $\chi^{i,j,k}$ as the susceptibility to variant j of individuals in S given age group i and vaccine stratum k, relative to a non vaccinated individual (so that $\chi^{i,j,0} = 1$ for all i and j), given by

$$\chi^{i,j,k} = (1 - e_{inf}^{i,j,k}), \tag{2.1}$$

where $e_{inf}^{i,j,k}$ was the vaccine effectiveness against infection of variant j in vaccine strata k (given values in Table A3 weighted by vaccine type distribution by age, according to Figure A4).

In turn, I defined $\xi^{i,j,k}$ as the infectivity of an individual in group i and vaccine stratum k infected with variant j relative to a non vaccinated individual infected with the Wildtype variant (so that $\xi^{i,Wildtype,0} = 1$), given by

$$\xi^{i,j,k} = (1 - e_{ins}^{i,j,k})\sigma_j, \tag{2.2}$$

where $e_{ins}^{i,j,k}$ was the vaccine effectiveness against onward transmission (Table A3) and σ_j captured the effect of variant-specific transmissibility advantage on the force of infection, defined by

$$\sigma_j = \begin{cases} 1 & \text{if } j = Wildtype \\ \sigma_{Wildtype}\sigma_{Alpha} & \text{if } j = Alpha \\ \sigma_{Alpha}\sigma_{Delta} & \text{if } j = Delta \\ \sigma_{Delta}\sigma_{Omicron} & \text{if } j = Omicron \end{cases} \tag{2.3}$$

In turn, I defined $\Theta_{i,j,k}(t)$ as the number of infectious individuals with variant j in group i and vaccination stratum k, weighted by infectivity, given by

$$\Theta_{i,j,k}(t) = \xi^{i,j,k}\left(\theta_{I_A} I_A^{i,j,k}(t) + I_P^{i,j,k}(t) + I_{C_1}^{i,j,k}(t)\right). \tag{2.4}$$

where θ_{I_A} was the infectivity of an asymptomatic infected individual relative to a symptomatic one infected with the same variant (Table 5.7), and in the same vaccination strata.

Given the above, I defined the force of infection, $\lambda^{i,j,k}(t)$, of variant $j \in \{j_1, j_2\}$ on a susceptible individual in group $i \in \{[0,5), \ldots, [75,80), [80+)\}$ and vaccination stratum $k = 0, 1, \ldots, 6$ as

$$\lambda^{i,j,k}(t) = \chi^{i,j,k} \sum_{i'} m_{i,i'}(t) \sum_{k'} \left(\Theta_{i',j,k'}(t) + \Theta_{i',j^{reinf},k'}(t)\right) \tag{2.5}$$

where $m_{i,i'}(t)$ corresponded to the time-varying contact matrix (section 2.2.3). This in turn, allowed me to define $\Lambda^{i,k}(t)$ as the total force of infection acting on susceptible individuals in group i and vaccination stratum k given by

$$\Lambda^{i,k}(t) = \lambda^{i,j_1,k}(t) + \lambda^{i,j_2,k}(t). \tag{2.6}$$

Time-varying next generation matrices

Based on [18, 32, 89], I simulated transmission via a time-varying contact matrix

$$m_{i,i'}(t) = c_{i,i'}\beta(t) \tag{2.7}$$

where $c_{i,i'}$ was the baseline social contact matrix for the UK [76], derived using the `socialmixr` package [80], and $\beta(t)$ a fitted piecewise-linear multiplier function given by

$$\beta(t) = \begin{cases} \beta_i, & \text{if } t \leq t_i,\ i = 1 \\ \dfrac{t_i - t}{t_i - t_{i-1}}\beta_{i-1} + \dfrac{t - t_{i-1}}{t_i - t_{i-1}}\beta_i, & \text{if } t_{i-1} < t \leq t_i,\ i \in \{2, \ldots, 36\} \\ \beta_i & \text{if } t > t_i,\ i = 36 \end{cases} \tag{2.8}$$

with 36 change points t_i corresponding to major NPI policy implementations or lifting and other relevant changes in contact rates (Table 3.3).

This in turn allowed me to determine two time-varying next generation matrices (NGM)

$$NGM_{i,i'}(t) = m_{i,i'}(t)\Delta_I^i N^j(t) \tag{2.9}$$

$$NGM_{i,i'}^{eff}(t) = m_{i,i'}(t)\Delta_I^i S^j(t) \tag{2.10}$$

where $N_j(t)$ and $S_j(t)$ were the total and susceptible population, respectively, in the infection-receiving age group i', and Δ_I^i represented the mean duration of infectiousness weighted by the infectivity θ_{I_A} of an individual in group i and vaccine stratum k as

$$\Delta_I^{i,k} = \theta_{I_A}\left(1 - p_C^i\right)\mathbb{E}\left[\tau_{I_A}\right] + p_C^i\left(\mathbb{E}\left[\tau_{I_P}\right] + \mathbb{E}\left[\tau_{I_{C_1}}\right]\right). \tag{2.11}$$

where p_C^i was the probability of symptomatic disease conditional on infection and age i (see section 2.2.4).

Effective reproduction number

The matrix $NGM_{i,i'}(t)$ thus represented the expected number of new infections given contact rates at time t by age classes in the entire population, void of the effect of any layer of immunity. The matrix $NGM_{i,i'}^{eff}(t)$, on the other hand, represented the corresponding number of infections but further accounting for immunity in the population.

Given that $\Delta_I^{i,k}$ did not depend on j, I calculated the NGM for variant j as

$$NGM_{i,i'}^{j}(t) = m_{i,i'}(t)\xi^{i,j,0}\Delta_I^{i,0}N^{i'} \tag{2.12}$$

Similarly, I calculated the effective NGMs for co-circulating variants j_1 and j_2 as

$$NGM_{D(i,k),D(i',k')}^{j_1,eff}(t) = M_{i,i'}(t)\chi^{i',j_1,k'}\xi^{i,j_1,k}\left(S^{i',k'}(t) + \left(1 - \eta_{inf}^{j_1}\right)R^{i',j_H,k'}(t)\right), \tag{2.13}$$

$$NGM^{j_2,eff}_{D(i,k),D(i',k')}(t) = M_{i,i'}(t)\chi^{i',j_2,k'}\xi^{i,j_2,k}\left(S^{i',k'}(t) + \left(1-\eta^{j_2}_{inf}\right)\left(R^{i',j_1,k'}(t) + R^{i',j_1^{reinf},k'}(t) + R^{i',j_H,k'}(t)\right)\right) \tag{2.14}$$

where $D : \{[0-4],[5-9],\ldots,[75-79],80+\} \times \{0,1,\ldots,6\} \rightarrow \{1,2,\ldots,119\}$ was a one-to-one mapping, $M_{i,i'}(t)$ was shorthand for $m_{i,i'}(t)\Delta^{i,k}_I$, and η^j_{inf} the parameter for cross-immunity against infection with variant j (Table 5.7). That is, the force of infection of variant j_1 on an individual recovered from historic variants j_H in group i and vaccine stratum k was $(1-\eta^{j_1}_{inf})\lambda^{i,j_1,k}(t)$; the force of infection of variant j_2 on an individual recovered from variant j_1 or historic variants j_H in group i and vaccine stratum k was $(1-\eta^{j_2}_{inf})\lambda^{i,j_2,k}(t)$; and there was no force of infection of variant j_1 on individuals recovered from variants j_1 or j_2, or of variant j_2 on individuals recovered from variant j_2.

Thus, I defined R^j_t as the dominant eigenvalue of the 17 by 17 matrix $NGM^j(t)$, and $R^{j,eff}_t$ as the dominant eigenvalue of the 119 by 119 matrix $NGM^{j,eff}(t)$. In turn,

$$R_t = \frac{w_{j_1}(t)R^{j_1}_t + w_{j_2}(t)R^{j_2}_t}{w_{j_1}(t) + w_{j_2}(t)} \tag{2.15}$$

$$R^{eff}_t = \frac{w_{j_1}(t)R^{j_1,eff}_t + w_{j_2}(t)R^{j_2,eff}_t}{w_{j_1}(t) + w_{j_2}(t)}, \tag{2.16}$$

where $w_j(t)$ were weightings based on the prevalence of each variant, accounting for the relative infectivity of each compartment, such that for $j = j_1, j_2$

$$w_j(t) = \sum_i \sum_k \left(\theta_{I_A}\left(I^{i,j,k}_A(t) + I^{i,j^{reinf},k}_A(t)\right) + I^{i,j,k}_P(t) + I^{i,j^{reinf},k}_P(t) + I^{i,j,k}_{C_1}(t) + I^{i,j^{reinf},k}_{C_1}(t)\right). \tag{2.17}$$

Basic reproduction number

By definition, the basic reproduction number R_0 is equal to R_t at time 0 in a fully susceptible population [88]. Given I assumed all infections at model initialisation were due to the Wildtype variant, I defined its R_0 as

$$R^{Wildtype}_0 \equiv R_{t=0} \tag{2.18}$$

For the VOCs, I then defined the values of R^j_0 as

$$R^{Alpha}_0 = R^{Wildtype}_0 * \sigma_{Alpha} \tag{2.19}$$

$$R^{Delta}_0 = R^{Wildtype}_0 * \sigma_{Alpha} * \sigma_{Delta} \tag{2.20}$$

$$R^{Omicron}_0 = R^{Wildtype}_0 * \sigma_{Alpha} * \sigma_{Delta} * \sigma_{Omicron} \tag{2.21}$$

That is, the instantaneous reproduction number at the start of the pandemic, $R_0^{j=Wildtype}$, by definition void of the effect of any immunity in the population, factored by the σ_j parameter of the corresponding VOCs in succession and, given the NGM determined by social mixing patterns at $t = 0$.

Serial interval of variants

Emerging evidence at that time I developed the above parametric calculations of transmissibility strongly indicated that the relative rate of replacement of a new variant was sensitive to assumptions of the mean serial interval [139]. I thus modelled proportional reductions in the mean serial interval of the VOCs compared to the Wildtype variant (Table 2.3), using the formula for mean backwards serial interval [140]

$$\mathbb{E}[SI] = \mathbb{E}\left[\tau_E\right] + \frac{\mathbb{E}\left[\tau_{I_P}\right]^2 + \mathbb{E}\left[\tau_{I_P}\right]\mathbb{E}\left[\tau_{I_{C_1}}\right] + \mathbb{E}\left[\tau_{I_{C_1}}\right]^2}{\mathbb{E}\left[\tau_{I_P}\right] + \mathbb{E}\left[\tau_{I_{C_1}}\right]} \tag{2.22}$$

For each variant in turn, I applied a proportional reduction in mean duration τ of the E, I_P, I_{C_1} and I_A compartments. I assumed the same total mean time from symptom onset to (possible) hospitalisation $\mathbb{E}\left[\tau_{I_{C_1}}\right] + \mathbb{E}\left[\tau_{I_{C_2}}\right]$ across variants remained unchanged, at 4 days.

Variant	Mean (days)	SI	% reduction	Source	Compartment	Erlang duration parameters k_X	γ_X^j	Mean (τ)
Wildtype	5.2		-	[141]	E	2	0.87	2.31
					I_A	1	0.35	2.88
					I_P	1	0.60	1.68
					I_{C_1}	1	0.47	2.14
					I_{C_2}	1		1.86
Alpha	4.9		6%	[142]	E	2	0.92	2.17
					I_A	1	0.37	2.71
					I_P	1	0.63	1.58
					I_{C_1}	1	0.50	2.01
					I_{C_2}	1	0.50	1.99
Delta	4.5		13%	[142]	E	2	0.99	2.01
					I_A	1	0.40	2.51
					I_P	1	0.68	1.46
					I_{C_1}	1	0.54	1.86
					I_{C_2}	1	0.47	2.14
Omicron	3.9		25%	[142]	E	2	1.15	1.73
					I_A	1	0.46	2.16
					I_P	1	0.79	1.26
					I_{C_1}	1	0.62	1.61
					I_{C_2}	1	0.42	2.39

Table 2.3: Serial interval (SI) assumptions for different variants of concern. All mean SI percent reductions are relative to the Wildtype variant. Mean (τ) is expressed in days and γ_x^j corresponds to respective k_x divided by mean duration. Please note that k dimension here refers to the number of successive sub-compartments used to generate a flexible Erlang delay distribution, rather than assuming exponential durations of all transitions. Source: Perez-Guzman et. al. 2023 [1].

2.2.4 Modelling severity

In its early stages [18], the model's severity pathways operated with a series of fixed (over time) age-specific parameters (henceforth *baseline*) $p_x^{i,j,k}$, which defined the probability of an individual of age i progressing to compartment x. These had been originally inferred from linelist data on hospitalisations and deaths in England during the first pandemic wave [18]. Before starting my PhD, I had expanded the model to incorporate a series of fitted piece-wise linear functions $h(t)$ (Table 2.4) to scale these baseline severity parameters [32, 89]. When I first implemented these $h(t)$ function, however, they where agnostic to the effect of specific drivers of effective severity. Rather, their main purpose was mechanistic, to improve the model's fit to hospital admission and occupancy data streams to provide rapid real-time analytics.

During my PhD, I performed several model expansions to: overhaul the $h(t)$ functions to infer the effect of new treatments and other modifications in healthcare provision on disease severity over time (Table 2.4); calculate the variant-specific basic infection hospitalisation ratio IHR_0^j, hospital fatality ratio HFR_0^j and infection fatality ratio IFR_0^j; calculate the respective effective severity metrics overall, $IHR(t)$, $HFR(t)$ and $IFR(t)$, and given age, variant and vaccination classes $IHR^{i,j,k}(t)$, $HFR^{i,j,k}(t)$ and $IFR^{i,j,k}(t)$.

Severity pathway probabilities and rates

As mentioned, the movement of individuals between disease severity pathways is determined by the parameters $p_x^{i,j,k}$, defining the probability of progressing to compartment x, as well as the respective rate parameters γ_x. The former parameters may vary between age groups i, variant of infection (j), and/or vaccine strata (k), as defined in the respective equations below.

The probability that an infected individual will have a symptomatic infection depends only on their age group i, as in [18, 32, 89], and is given by

$$p_C^i = \begin{cases} 0.25 & \text{for i = [0,5)} \\ 0.26875 & \text{for i = [5,10)} \\ 0.325 & \text{for i = [10,15)} \\ 0.41875 & \text{for i = [15,20)} \\ 0.55 & \text{for i = [20,25, ...[75,80), [80+).} \end{cases} \tag{2.23}$$

Table 2.4: Change points for fitted time-varying severity multiplier parameters with piecewise linear form. Where two dates are specified, the model assumes the corresponding fitted severity multiplier value to remain fixed in between. Source: [1].

Parameter	Dates of change	Rationale	Reference
$h_H(t)$	04-11-2021 31-12-2021	Novel outpatient treatments	[143–146]
$h_{ICU}(t)$	01-04-2020 01-06-2020	Dexamethasone indicated	[128, 147]
$h_{G_D}(t)$	01-05-2020 01-07-2020	Change healthcare after 1st wave	Assumed
$h_D(t)$	01-04-2020	First wave	
	01-07-2020 to 15-09-2020	Trough after first wave	Assumed
	15-10-2020 to 01-12-2020	Winter 2020/21 wave	
	04-02-2021	Trough after winter 2020/21 wave	
	01-04-2021 to 04-11-2021	Approval and roll-out of novel	[143–146]
	31-12-2021	oral treatments for COVID-19	

The probability that a symptomatic individual has severe disease requiring hospitalisation for $j \in \{j_1, j_2\}$ is given by

$$p_H^{i,j,k}(t) = \min\left\{h_H(t)\psi_H^i\left(1 - e_{SD|sympt}^{i,j,k}\right)\pi_H^j,\ 1\right\}, \tag{2.24}$$

$$p_H^{i,j^{reinf},k}(t) = \min\left\{h_H(t)\psi_H^i\left(1 - e_{SD|sympt}^{i,j,k}\right)\pi_H^j\left(1 - \eta_H^j\right),\ 1\right\}, \tag{2.25}$$

where ψ_H^i represents the age scaling (such that $\psi_H^i = 1$ for the group with the maximum probability (Table 2.5), $e_{SD|sympt}^{i,j,k}$ the vaccine effectiveness against symptomatic disease (Table A3), η_H^j the probability of cross-immunity against hospitalisation for an individual infected with variant j (Table 5.7),

and $h_H(t)$ was a fitted scaling function with piece-wise linear form, given change points (Table 2.4)

$$h_H(t) = \begin{cases} p_{H,1}^{max} & \text{on (and before) 04/11/21,} \\ p_{H,2}^{max} & \text{on (and after) 31/12/21,} \end{cases} \tag{2.26}$$

and π_H^j was a fitted (see Table 3.3) multiplier accounting for the changing severity of the VOCs with respect to the Wildtype variant:

$$\pi_H^j = \begin{cases} 1 & \text{if } j = Wildtype, \\ \pi_H^{Wildtype}\pi_H^{Alpha/Wildtype} & \text{if } j = Alpha, \\ \pi_H^{Alpha}\pi_H^{Delta/Alpha} & \text{if } j = Delta, \\ \pi_H^{Delta}\pi_H^{Omicron/Delta} & \text{if } j = Omicron, \end{cases} \tag{2.27}$$

where $\pi_H^{Alpha/Wildtype}$, $\pi_H^{Delta/Alpha}$ and $\pi_H^{Omicron/Delta}$ were the relative probabilities of severe disease requiring hospitalisation given infection for Alpha over Wildtype, Delta over Alpha and Omicron over Delta, respectively.

The probability that an individual with severe disease died in the community was given by, for $j \in \{j_1, j_2\}$,

$$p_{G_D}^{i,j,k}(t) = \min\left\{h_{G_D}(t)\psi_{G_D}^i\left(1 - e_{death|SD}^{i,j,k}\right)\pi_D^j,\ 1\right\}, \tag{2.28}$$

$$p_{G_D}^{i,j^{reinf},k}(t) = \min\left\{h_{G_D}(t)\psi_{G_D}^i\left(1 - e_{death|SD}^{i,j,k}\right)\pi_D^j\left(1 - \eta_D^j\right),\ 1\right\}, \tag{2.29}$$

where h_{G_D} also had a piecewise linear form with change points (Table 2.4)

$$h_{G_D}(t) = \begin{cases} p_{G_D,1}^{max} & \text{on (and before) 01/05/20,} \\ p_{G_D,2}^{max} & \text{on (and after) 01/07/20,} \end{cases} \tag{2.30}$$

ψ_{G_D} was an age multiplier (Table 2.5), and π_D^j a multiplier accounting for the changing severity of the VOCs with respect to the Wildtype variant:

$$\pi_D^j = \begin{cases} 1 & \text{if } j = Wildtype, \\ \pi_D^{Wildtype}\pi_D^{Alpha/Wildtype} & \text{if } j = Alpha, \\ \pi_D^{Alpha}\pi_D^{Delta/Alpha} & \text{if } j = Delta, \\ \pi_D^{Delta}\pi_D^{Omicron/Delta} & \text{if } j = Omicron, \end{cases} \tag{2.31}$$

where $\pi_D^{Alpha/Wildtype}$, $\pi_D^{Delta/Alpha}$ and $\pi_D^{Omicron/Delta}$ were fitted parameters of the relative probability

of death given severe disease for Alpha over Wildtype, Delta over Alpha and Omicron over Delta, respectively (Table 3.3).

The probability that an individual was admitted to ICU given hospitalisation with infection by $j \in \{j_1, j_2\}$ was given by

$$p_{ICU}^{i,j}(t) = \min\left\{h_{ICU}(t)\psi_{ICU}^{i}\pi_{ICU}^{j},\ 1\right\}, \tag{2.32}$$

$$p_{ICU}^{i,j^{reinf}}(t) = p_{ICU}^{i,j}(t) \tag{2.33}$$

where ψ_{ICU}^{i} was the age scaling (Table 2.5), $h_{ICU}(t)$ also had a piece-wise linear form with change points (Table 2.4)

$$h_{ICU}(t) = \begin{cases} p_{ICU,1}^{max} & \text{on (and before) } 01/04/20, \\ p_{ICU,2}^{max} & \text{on } 01/06/20, \end{cases} \tag{2.34}$$

and π_{ICU}^{j} was a multiplier accounting for the changing severity of the variants of concern with respect to the Wildtype variant

$$\pi_{ICU}^{j} = \begin{cases} 1 & \text{if } j = Wildtype, \\ \pi_{ICU}^{Wildtype}\pi_{ICU}^{Alpha/Wildtype} & \text{if } j = Alpha, \\ \pi_{ICU}^{Alpha}\pi_{ICU}^{Delta/Alpha} & \text{if } j = Delta, \\ \pi_{ICU}^{Delta}\pi_{ICU}^{Omicron/Delta} & \text{if } j = Omicron, \end{cases} \tag{2.35}$$

where $\pi_{ICU}^{Alpha/Wildtype}$, $\pi_{ICU}^{Delta/Alpha}$ and $\pi_{ICU}^{Omicron/Delta}$ were fitted parameters of the relative probability of admission to ICU given hospitalisation for Alpha over Wildtype, Delta over Alpha and Omicron over Delta, respectively (Table 3.3).

The probability that an individual died in general beds given they were not admitted to ICU was, for $j \in \{j_1, j_2\}$,

$$p_{H_D}^{i,j,k}(t) = \min\left\{p_{H_D}^{max}h_D(t)\psi_{H_D}^{i}\left(1 - e_{death|SD}^{i,j,k}\right)\pi_D^{j}, 1\right\}, \tag{2.36}$$

$$p_{H_D}^{i,j^{reinf},k}(t) = \min\left\{p_{H_D}^{max}h_D(t)\psi_{H_D}^{i}\left(1 - e_{death|SD}^{i,j,k}\right)\pi_D^{j}\left(1 - \eta_D^{j}\right), 1\right\}, \tag{2.37}$$

where $\psi_{H_D}^{i}$ represented the age scaling (Table 2.5) and $h_D(t)$ also had a piece-wise linear form with

change points (see Table 2.4)

$$h_D(t) = \begin{cases} 1 & \text{on (and before) } 01/04/20, \\ \mu_{D,1}^{max} & \text{on } 01/07/20, \\ \mu_{D,1}^{max} & \text{on } 15/09/20, \\ \mu_{D,2}^{max} & \text{on } 15/10/20, \\ \mu_{D,2}^{max} & \text{on } 01/12/20, \\ \mu_{D,3}^{max} & \text{on } 04/02/21, \\ \mu_{D,4}^{max} & \text{on } 01/04/21, \\ \mu_{D,4}^{max} & \text{on } 04/11/21, \\ \mu_{D,5}^{max} & \text{on (and after) } 31/12/21. \end{cases} \tag{2.38}$$

The probability that an individual died in ICU given admission into it with infection by variant $j \in \{j_1, j_2\}$:

$$p_{ICU_D}^{i,j,k}(t) = \min\left\{p_{ICU_D}^{max} h_D(t) \psi_{ICU_D}^i \left(1 - e_{death|SD}^{i,j,k}\right) \pi_D^j,\ 1\right\} \tag{2.39}$$

$$p_{ICU_D}^{i,j^{reinf},k}(t) = \min\left\{p_{ICU_D}^{max} h_D(t) \psi_{ICU_D}^i \left(1 - e_{death|SD}^{i,j,k}\right) \pi_D^j \left(1 - \eta_D^j\right),\ 1\right\} \tag{2.40}$$

where $\psi_{ICU_D}^i$ represented the age scaling (Table 2.5).

The probability that an individual who had been in ICU dying in a stepdown bed, given that they did not die in ICU, for $j \in \{j_1, j_2\}$ was given by

$$p_{W_D}^{i,j,k}(t) = \min\left\{p_{W_D}^{max} h_D(t) \psi_{W_D}^i \left(1 - e_{death|SD}^{i,j,k}\right) \pi_D^j,\ 1\right\}, \tag{2.41}$$

$$p_{W_D}^{i,j^{reinf},k}(t) = \min\left\{p_{W_D}^{max} h_D(t) \psi_{W_D}^i \left(1 - e_{death|SD}^{i,j,k}\right) \pi_D^j \left(1 - \eta_D^j\right),\ 1\right\}, \tag{2.42}$$

where $\psi_{W_D}^i$ represented the age scaling (Table 2.5).

Finally, the probability of individuals having had a COVID-19 diagnosis confirmed prior to

admission to hospital, $p^*(t)$ had a piecewise linear form with change points

$$p^*(t) = \begin{cases} 0.1 & \text{on (and before) } 15/03/20, \\ 0.42 & \text{on } 01/07/20, \\ 0.2 & \text{on } 20/09/20, \\ 0.45 & \text{on } 27/06/21, \\ 0.45 & \text{on } 01/12/21, \\ 0.33 & \text{on (and after) } 01/01/22. \end{cases} \tag{2.43}$$

These were informed by data on COVID-19 admissions and inpatient diagnoses from NHS England [148]. Meanwhile, the rate at which unconfirmed COVID-19 patients in hospital became confirmed, $\gamma_U(t)$, had a piece-wise linear form with change points

$$\gamma_U(t) = \begin{cases} 1/2.1 & \text{on (and before) } 15/03/20, \\ 1/1.2 & \text{on } 05/04/20, \\ 1/1.2 & \text{on } 15/09/20, \\ 1/0.7 & \text{on } 15/11/20, \\ 1/1.3 & \text{on } 10/12/20, \\ 1/0.7 & \text{on } 20/01/21, \\ 1/0.4 & \text{on } 01/08/21, \\ 1/0.9 & \text{on } 01/03/22. \end{cases} \tag{2.44}$$

These were informed by data on times to diagnosis from CHESS [149].

In addition, I allowed for some of the duration rates for hospital compartments to also be time-varying to account for changes in length of stay over time, given

$$\gamma_{H_R}(t) = h_\gamma(t)\gamma_{H_R} \tag{2.45}$$

$$\gamma_{H_D}(t) = h_\gamma(t)\gamma_{H_D} \tag{2.46}$$

$$\gamma_{W_R}(t) = h_\gamma(t)\gamma_{W_R} \tag{2.47}$$

$$\gamma_{W_D}(t) = h_\gamma(t)\gamma_{W_D} \tag{2.48}$$

where $h_\gamma(t)$ had a piece-wise linear form with change points

$$h_\gamma(t) = \begin{cases} 1 & \text{on (and before) 01/12/20,} \\ 1/\mu_{\gamma_H,1} & \text{on 01/01/21,} \\ 1/\mu_{\gamma_H,2} & \text{on 01/03/21,} \\ 1/\mu_{\gamma_H,3} & \text{on 01/06/21,} \\ 1/\mu_{\gamma_H,4} & \text{on (and after) 01/09/21.} \end{cases} \tag{2.49}$$

Table 2.5: Baseline age-specific multipliers for severity pathway probabilities. These had been previously estimated [18] using a progression model fitted with MCMC to age-stratified linelist data from CHESS [149]. Source: Perez-Guzman et al. 2023 [1].

Age group i	ψ^i_H	$\psi^i_{G_D}$	ψ^i_{ICU}	$\psi^i_{H_D}$	$\psi^i_{ICU_D}$	$\psi^i_{W_D}$
$[0, 5)$	0.0149	0.0000	0.2428	0.0386	0.2823	0.0911
$[5, 10)$	0.0027	0.0078	0.2891	0.0365	0.2861	0.0830
$[10, 15)$	0.0033	0.0033	0.3377	0.0353	0.2913	0.0775
$[15, 20)$	0.0065	0.0123	0.3894	0.0351	0.2991	0.0744
$[20, 25)$	0.0129	0.0139	0.4426	0.0362	0.3103	0.0739
$[25, 30)$	0.0191	0.0129	0.5027	0.0391	0.3276	0.0760
$[30, 35)$	0.0269	0.0242	0.5697	0.0447	0.3526	0.0802
$[35, 40)$	0.0255	0.0420	0.6530	0.0552	0.3909	0.0860
$[40, 45)$	0.0287	0.0604	0.7559	0.0743	0.4465	0.0927
$[45, 50)$	0.0336	0.0801	0.8659	0.1067	0.5196	0.1016
$[50, 55)$	0.0568	0.1054	0.9541	0.1568	0.6044	0.1167
$[55, 60)$	0.0769	0.1317	1.0000	0.2385	0.7047	0.1482
$[60, 65)$	0.1075	0.1679	0.9720	0.3528	0.8057	0.2113
$[65, 70)$	0.1364	0.2165	0.8544	0.5020	0.8988	0.3315
$[70, 75)$	0.2452	0.2992	0.6454	0.6750	0.9692	0.5263
$[75, 80)$	0.4218	0.4327	0.4024	0.8319	1.0000	0.7531
$[80+)$	1.0000	1.0000	0.1074	1.0000	0.9178	1.0000

Variant effective severity

Given the implementation of aforementioned time-varying conditional severity probabilities and rates, I defined the variant-specific effective severity for individuals in age group i and vaccine class k infected with variant j as

$$IHR^{i,j,k}(t) = p_C^i p_H^{i,j,k}(t)(1 - p_{G_D}^{i,j,k}(t)) \tag{2.50}$$

$$HFR^{i,j,k}(t) = \left(1 - p_{ICU}^{i,j,k}(t)\right) p_{H_D}^{i,j,k}(t) + p_{ICU}^{i,j,k}(t) p_{ICU_D}^{i,j,k}(t) + p_{ICU}^{i,j,k}(t) \left(1 - p_{ICU_D}^{i,j,k}(t)\right) p_{W_D}^{i,j,k}(t) \tag{2.51}$$

$$IFR^{i,j,k}(t) = IHR^{i,j,k}(t) HFR^{i,j,k}(t) + p_C^i p_H^{i,j,k}(t) p_{G_D}^{i,j,k}(t). \tag{2.52}$$

To then obtain overall estimates of effective severity, I aggregated the above measures weighting by new infections $n_I^{i,j,k}(t)$ for IFR and IHR, or new hospitalisations $n_H^{i,j,k}(t)$ for HFR across age groups i, variants j and vaccine classes k as follows:

$$IHR(t) = \frac{\sum_i \sum_j \sum_k n_I^{i,j,k}(t) IHR^{i,j,k}(t)}{\sum_i \sum_j \sum_k n_I^{i,j,k}(t)} \tag{2.53}$$

$$HFR(t) = \frac{\sum_i \sum_j \sum_k n_H^{i,j,k}(t) HFR^{i,j,k}(t)}{\sum_i \sum_j \sum_k n_H^{i,j,k}(t)} \tag{2.54}$$

$$IFR(t) = \frac{\sum_i \sum_j \sum_k n_I^{i,j,k}(t) IFR^{i,j,k}(t)}{\sum_i \sum_j \sum_k n_I^{i,j,k}(t)}, \tag{2.55}$$

Additionally, I obtained estimates of effective severity over time for age group i, given

$$IHR^i(t) = \frac{\sum_j \sum_k n_I^{i,j,k}(t) IHR^{i,j,k}(t)}{\sum_j \sum_k n_I^{i,j,k}(t)} \tag{2.56}$$

$$HFR^i(t) = \frac{\sum_j \sum_k n_H^{i,j,k}(t) HFR^{i,j,k}(t)}{\sum_j \sum_k n_H^{i,j,k}(t)} \tag{2.57}$$

$$IFR^i(t) = \frac{\sum_j \sum_k n_I^{i,j,k}(t) IFR^{i,j,k}(t)}{\sum_j \sum_k n_I^{i,j,k}(t)}. \tag{2.58}$$

For the case of variant j, I further calculated the effective severity

$$IHR^j(t) = \frac{\sum_i \sum_k \left(n_I^{i,j,k}(t) IHR^{i,j,k}(t) + n_I^{i,j^{reinf},k}(t) IHR^{i,j^{reinf},k}(t)\right)}{\sum_i \sum_k \left(n_I^{i,j,k}(t) + n_I^{i,j^{reinf},k}(t)\right)} \tag{2.59}$$

$$HFR^j(t) = \frac{\sum_i \sum_k \left(n_H^{i,j,k}(t) HFR^{i,j,k}(t) + n_H^{i,j^{reinf},k}(t) HFR^{i,j^{reinf},k}(t)\right)}{\sum_i \sum_k \left(n_H^{i,j,k}(t) + n_H^{i,j^{reinf},k}(t)\right)} \tag{2.60}$$

$$IFR^j(t) = \frac{\sum_i \sum_k \left(n_I^{i,j,k}(t) IFR^{i,j,k}(t) + n_I^{i,j^{reinf},k}(t) IFR^{i,j^{reinf},k}(t)\right)}{\sum_i \sum_k \left(n_I^{i,j,k}(t) + n_I^{i,j^{reinf},k}(t)\right)}. \tag{2.61}$$

Variant basic severity

In a similar fashion to R_0^j (section 2.2.3), I defined the variant-specific IHR_0^j, HFR_0^j and IFR_0^j as those that would be observable in an entirely immunologically naïve population, given its age-distribution and baseline social contact rates and, critically, assuming the healthcare ecosystem at $t \sim 0$. For the comparison of variants I performed in Chapter 3, I set time t to the 16th of May 2020 to capture the basic severity at a point in time where there were enough individuals across compartments to calculate weights (as above), but no significant population-level immunity and prior to the introduction of the first targeted treatments for severe COVID-19 in early June 2020 [150].

Thus, formally

$$IHR_0^j(t) = \frac{\sum_i w_I^i(t) IHR^{i,j,0}(t)}{\sum_i w_I^i(t)}, \tag{2.62}$$

$$HFR_0^j(t) = \frac{\sum_i w_I^i(t) IHR^{i,j,0}(t) HFR^{i,j,0}(t)}{\sum_i w_I^i(t) IHR^{i,j,0}(t)}, \tag{2.63}$$

$$IFR_0^j(t) = \frac{\sum_i w_I^i(t) IFR^{i,j,0}(t)}{\sum_i w_I^i(t)}, \tag{2.64}$$

where $w_I^i(t)$ were the weightings from the eigenvector corresponding to leading eigenvalue of $NGM_{i,i'}(t)$ and, in the case of HFR_0^j, further weighted by $IHR^{i,j,0}(t)$ to obtain the distribution of hospitalised cases by age corresponding to the eigenvector at time t.

2.3 Discussion

The methods I proposed in this chapter overcame several critical limitation in transmission models employed during the COVID-19 pandemic. Namely, I proposed a formal parametric method to infer the effect of intrinsic *vs* effective drivers of transmission and severity. This approach explicitly unpicked and quantified the specific effect of the following factors on the 'observable' population-level basic and effective transmission and severity of SARS-CoV-2:

- Population age distribution and social contact patters
- Immunity profile
- Age distribution in the case mix of infections and, for HFR, hospitalisations
- Emerging VOCs

Although the equations, parameters and assumptions I defined in this chapter pertained to the structure and mechanisms of a specific modelling framework of SARS-CoV-2 transmission in England, the ICL-RTM team's model, their underlying mathematical principles are applicable to other models.

The representation of age-, vaccine- and strain-classes and healthcare compartments in transmission modelling frameworks was common to the study of COVID-19 across many settings [33, 90, 95, 96]. Two aspects made my modelling framework specific to England. Namely the the specification of change-point dates in the piece-wise linear function scaling transmission and severity and, as I will discuss in the following chapter, the coding of the marginal-likelihood functions comparing the model's hidden state variables to the data.

These are aspected that can be easily adapted to the other settings. Shortly after the publication of my methods [1], another study adopted its parametric calculations of transmissibility and severity and the use of piece-wise liner functions to an age-, strain- and vaccine-class stratified model of SARS-CoV-2 transmission in French Polynesia [94]. Two additional SARS-CoV-2 modelling studies have made use of parameters estimated by my model, one in the UK [151] and another in China [152], and six others have contextualised their own analyses of the COVID-19 pandemic in relation to the novel insights my methods brought [153–158].

For the case of preparedness, the parametric approach I developed is in principle applicable to the study of other outbreak pathogens, provided the aim is to study them with the use of a discrete-time transmission model with age- and immune- and/or strain-class disaggregation equipped with a multi-dimensional parameter space inference framework, such as pMCMC. Indeed, during an international workshop on the use of modelling for pandemic preparedness, held at the Centre for Excellence in Respiratory Pathogens (CERP) in Lyon, France, a central theme that was discussed was the critical need for multi-source utilisation of data [159]. In this regard, my work was cited as an example of a holistic approach to the use of data for the understanding of outbreak disease spread and severity, amidst a landscape of rapidly evolving epidemiological drivers [159].

Chapter 3

Drivers of SARS-CoV-2 transmissibility and severity in England

3.1 Background

The emergence of SARS-CoV-2 in late 2019 signalled the start of a global health emergency that would inflict unprecedented pressures on healthcare systems and economies globally, and cause surges in excess mortality [42]. Even the more resilient public health surveillance and healthcare systems of HICs were put under severe pressure and, thus, stringent NPIs that included lockdowns (i.e. education, social and economic closures of all but essential services) and a re-structuring of health services to deal with pandemic-driven demand were imposed [3, 71].

The work conducted by the ICL-RTM team and other academic groups providing analytics to SPI-M-O and SAGE [3, 160] was seminal in driving an understanding of population-level dynamics of the COVID-19 pandemic (Figure 1.1). Yet, prior to the work I conducted during my PhD, it had not been possible to systematically quantified the individual and combined effect of the intrinsic and severity factors of the 'observable' population-level transmission and severity dynamics of SARS-CoV-2.

In this chapter I discuss how I tailored the model I presented in Chapter 2 to COVID-19 surveillance data from England between March 2020 and February 2022, and thus estimate the time-varying transmission and severity dynamics of the pandemic in the most granular detail for any modelling study to date. Specifically, I quantified the impact on these dynamics given changes in contact rates due to NPIs, the introduction of novel treatments for severe COVID-19, the emergence of VOCs, and the profile of infection-, vaccine- or hybrid-induced immunity in the population.

3.2 Methods

I illustrated the model's structure in Figure 2.2, detailed its compartments in table 2.1 and its parameters in table 2.2. In this methods section, I describe the technical details of the data employed, model parameterisation and fitting approach. Additional technical details, including the model's equations and approach to simulating vaccination are presented in Appendix A.

3.2.1 COVID-19 surveillance data

I fitted the model independently to COVID-19 surveillance data for each of the seven National Health Service (NHS) regions in England (Table 3.1). These data included the number of positive and negative PCR tests in the community, population-representative infection prevalence surveys, genetic characterisation of a sample of the PCR positive cases, seroprevalence from blood donor residual sera, hospital admissions and deaths, and community deaths.

Table 3.1: England data sources and definitions. All data are reported by NHS region or processed to match these regions. Source: Perez-Guzman et al. 2023 [1].

Data	Description	Stratification	Source	Reference
Hospital deaths	Daily deaths with COVID-19 in hospital	Age: 0-49, 50-54, 55-59, 60-64, 65-69, 70-74, 75-79 and 80+	ONS	[161]
Community deaths	Daily deaths with COVID-19 in any other place outside of hospital	Age: 0-49, 50-54, 55-59, 60-64, 65-69, 70-74, 75-79 and 80+	ONS	[161]
ICU occupancy	Daily number of confirmed COVID-19 patients in ICU	None	Gov.uk Dashboard	[161]
General bed occupancy	Daily number of confirmed COVID-19 patients in non-ICU beds	None	Gov.uk Dashboard	[161]
Admissions	Daily confirmed COVID-19 patients admitted to hospital	Age: 0-9, 10-19, 20-29, 30-39, 40-49, 50-59, 60-69, 70-79 and 80+	PHE	[161]
Pillar 2 testing	Daily number of positive (cases) and negative PCR test results	Age: 15-24, 25-49, 50-64, 65-79 and 80+	PHE	[161]
REACT-1 testing	Daily number of positive and negative PCR test results	Age: 5-24, 25-34, 35-44, 45-54, 55-64 and 65+	REACT	[9]
ONS testing	Weekly number of positive and negative PCR test results	None	ONS	[8]
Serology	Serology survey conducted on blood donors aged 15-65.	EuroImmun pre-Jan 14, 2021; Roche N assays thereafter	PHE	[162].
Vaccinations	Daily number of first, second and booster vaccine doses	Age: reported in 5-year bands	PHE	[161]
Variant and Mutation	Daily number of variant tests of symptomatic individuals from Pillar 2	None	VAM	[163]

3.2.2 Observation process

I fitted the model to the above data using a bespoke pMCMC algorithm developed for the ICL-RTM model [18]. This process involved a series of marginal-likelihood functions (i.e. of probabilities of observing the data given model parameterisation). This section summarises the notation and technical details of this process.

Notations

If $X \sim \text{Binom}(n, p)$, then X follows a binomial distribution with mean np and variance $np(1-p)$, such that

$$P(X = x) = P_{\text{Binom}}(x|n,\ p) = \binom{n}{x} p^x (1-p)^{(n-x)}. \tag{3.1}$$

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

$$P(Y = y) = P_{\text{NegBinom}}(y|m, \kappa) = \frac{\Gamma(\kappa + y)}{y!\Gamma(\kappa)} \left(\frac{\kappa}{\kappa + m}\right)^{\kappa} \left(\frac{m}{\kappa + m}\right)^{y}, \tag{3.2}$$

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

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

$$P(Z = z) = P_{\text{BetaBinom}}(z|n, \omega, \rho) = \binom{n}{z} \frac{B(z + a, n - z + b)}{B(a, b)}, \tag{3.3}$$

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

In the following subsections, $d_{x,y}^{i,j,k}(t)$ represents the model inferred distribution of the number of individuals moving from compartment x to y at a time t (day), given their age i, strain j and vaccination k class. It should be noted that different data streams had different sets of days with observations.

Hospital admissions and new diagnoses in hospital

I represented the daily number of confirmed COVID-19 hospital admissions and new diagnoses for existing hospitalised patients, $Y_{adm}^{z}(t)$, in age band $z \in Z_{adm}$ ($Z_{adm} = \{[0, 10), [10, 20), \ldots, [70, 80), [80+)\}$), as the observed realisations of an underlying hidden Markov process. This corresponded to the model-

predicted number of admissions to hospital and diagnoses in hospital $X^z_{adm}(t)$, calculated as

$$\begin{aligned} X^z_{adm}\,(t) \coloneqq \sum_{i\in z}\sum_j\sum_k \Big(& d^{i,j,k}_{I_C,H_{R^*}}(t) + d^{i,j,k}_{I_C,H_{D^*}}(t) + d^{i,j,k}_{I_C,ICU_{pre^*}}(t) + d^{i,j,k}_{H_R,H_{R^*}}(t) + d^{i,j,k}_{ICU_{pre},ICU_{pre^*}}(t) \\ & + d^{i,j,k}_{ICU_{W_R},ICU_{W_{R^*}}}(t) + d^{i,j,k}_{ICU_{W_D},ICU^*_{W_D}}(t) + d^{i,j,k}_{W_D,W_{D^*}}(t) + d^{i,j,k,1,1}_{H_D H_{D^*}}(t) + d^{i,j,k,1,2}_{H_D,H_{D^*}}(t) \\ & + d^{i,j,k,2,2}_{H_D,H_{D^*}}(t) + d^{i,j,k,1,1}_{ICU_D,ICU_{D^*}}(t) + d^{i,j,k,1,2}_{ICU_D,ICU_{D^*}}(t) + d^{i,j,k,2,2}_{ICU_D,ICU_{D^*}}(t) + d^{i,j,k,1,1}_{W_R,W_{R^*}} \\ & + d^{i,j,k,1,2}_{W_R,W_{R^*}}(t) + d^{i,j,k,2,2}_{W_R,W_{R^*}}(t) + d^{i,j,k}_{ICU_{pre},ICU_{W_{R^*}}}(t) + d^{i,j,k}_{ICU_{pre},ICU_{W_{D^*}}}(t) \\ & + d^{i,j,k}_{ICU_{pre},ICU_{D^*}}(t) + d_{ICU^{i,j,k}_{W_D,W_{D^*}}(t)+d_{ICU^{i,j,k}_{W_R,W_{R^*}}(t)}} \Big) \end{aligned} \tag{3.4}$$

This was related to the surveillance data stream of new hospitalisations via a negative-binomial reporting distribution:

$$Y^z_{adm}(t) \sim \text{NegBinom}\left(X^z_{adm}(t), \kappa_A\right), \tag{3.5}$$

with a fitted overdispersion parameter, $\alpha_A = \frac{1}{\kappa_A}$, which accounted for "noise" in the underlying data streams (e.g. day-of-week effects).

The contribution to the likelihood of the data on hospital admissions and new diagnoses in hospital in age band z was therefore:

$$\mathcal{L}_{adm} = \prod_t \prod_{z\in Z_{adm}} P_{\text{NegBinom}}\left(Y^z_{adm}(t)\,|X^z_{adm}\,(t),\ \kappa_A\right) \tag{3.6}$$

Hospital bed occupancy by confirmed COVID-19 cases

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

$$\begin{aligned} X_{hosp}(t) \coloneqq \sum_i\sum_j\sum_k \big(& H^{i,j,k}_{R^*}(t) + H^{i,j,k,1}_{D^*}(t) + H^{i,j,k,2}_{D^*}(t) + \\ & ICU^{i,j,k}_{pre^*}(t) + W^{i,j,k}_{D^*}(t) + W^{i,j,k,1}_{R^*}(t) + W^{i,j,k,2}_{R^*}(t)\Big), \end{aligned} \tag{3.7}$$

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

$$Y_{hosp}\,(t) \sim \text{NegBinom}\left(X_{hosp}(t), \kappa_H\right). \tag{3.8}$$

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

$$X_{ICU}(t) \coloneqq \sum_i\sum_j\sum_k \Big(ICU^{i,j,k}_{W_{R^*}}(t) + ICU^{i,j,k}_{W_{D^*}}(t) + ICU^{i,j,k,1}_{D^*}(t) + ICU^{i,j,k,2}_{D^*}(t)\Big), \tag{3.9}$$

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

$$Y_{ICU}(t) \sim \text{NegBinom}\left(X_{ICU}(t), \kappa_H\right). \tag{3.10}$$

As above, this allowed the fitting of an overdispersion parameter $\alpha_H = \frac{1}{\kappa_H}$, used for both general hospital bed and ICU bed occupancy.

The overall contribution to the likelihood of the data on general bed and ICU bed occupancy was:

$$\begin{aligned}\mathcal{L}_{beds} = &\prod_t P_{\text{NegBinom}}\left(Y_{hosp}(t)\,|X_{hosp}(t), \kappa_H\right) \\ &\times \prod_t P_{\text{NegBinom}}\left(Y_{ICU}(t)\,|X_{ICU}(t), \kappa_H\right).\end{aligned} \tag{3.11}$$

Hospital and community COVID-19 deaths

The model considered $Y^z_{D_{hosp}}(t)$, the reported number of daily COVID-19 deaths in hospitals in age band $z \in Z_D$ ($Z_D = \{[0,50),\ \ [50,55), [55,60)\ldots, [75,80), [80+)\}$), as the observed realisation of an underlying hidden Markov process, $X^z_{D_{hosp}}(t)$, defined as:

$$X^z_{D_{hosp}}(t) \coloneqq \sum_{i\in z}\sum_j\sum_k \left(d^{i,j,k}_{H_D,D}(t) + d^{i,j,k}_{H_{D^*},D}(t) + d^{i,j,k}_{ICU_D,D}(t) + d^{i,j,k}_{ICU_{D^*},D}(t) + d^{i,j,k}_{W_D,D}(t) + d^{i,j,k}_{W_{D^*},D}(t)\right), \tag{3.12}$$

which was related to the data via a reporting distribution:

$$Y^z_{D_{hosp}}(t) \sim \text{NegBinom}\left(X^z_{D_{hosp}}(t), \kappa_{D_{hosp}}\right). \tag{3.13}$$

Similarly, the model represented the reported number of daily COVID-19 deaths in the community in age band $z \in Z_D$, $Y^z_{D_{comm}}(t)$, as the observed realisations of an underlying hidden Markov process, $X^z_{D_{comm}}(t)$, defined as:

$$X^z_{D_{comm}}(t) \coloneqq \sum_{i\in z}\sum_j\sum_k d^{i,j,k}_{G_D,D}(t), \tag{3.14}$$

which was related to the data via a reporting distribution:

$$Y^z_{D_{comm}}(t) \sim \text{NegBinom}\left(X^z_{D_{comm}}(t), \kappa_{D_{comm}}\right). \tag{3.15}$$

This process also involved the fitting of overdispersion parameters $\alpha_{D_{hosp}} = \frac{1}{\kappa_{D_{hosp}}}$ for hospital deaths and $\alpha_{D_{comm}} = \frac{1}{\kappa_{D_{comm}}}$ for community deaths.

The overall contribution to the likelihood of the data on COVID-19 deaths in hospitals and

the community was:

$$\begin{aligned}\mathcal{L}_{deaths} = &\prod_t \prod_{z \in Z_D} P_{\text{NegBinom}}\left(Y_{D^z_{hosp}}(t) \,\middle|\, X_{D^z_{hosp}}(t), \kappa_{D^z_{hosp}}\right) \\ &\times \prod_t \prod_{z \in Z_D} P_{\text{NegBinom}}\left(Y_{D^z_{comm}}(t) \,\middle|\, X_{D^z_{comm}}(t), \kappa_{D^z_{comm}}\right)\end{aligned} \tag{3.16}$$

Serosurveys

There were two different serological tests employed in England during the study period, EuroImmun and Roche N (Table 3.1). The model explicitly simulated serological testing of all individuals aged 15-64, inclusive, via a parallel flow approach (see Appendix A.1). That is, assuming all individuals aged 15-65 were to be tested, the resulting number of seropositive and seronegative individuals with test j (where $j = 1$ corresponds to EuroImmun and $j = 2$ to Roche N), as:

$$X_{sero^j_{pos}}(t) \coloneqq \sum_{i=[15,20)}^{[60,65)} T^i_{sero^j_{pos}}(t) \tag{3.17}$$

$$X_{sero^j_{neg}}(t) \coloneqq \left(\sum_{i=[15,20)}^{[60,65)} N^i\right) - X_{sero^j_{pos}}(t). \tag{3.18}$$

were compared to the observed number of seropositive results, $Y_{sero^j_{pos}}(t)$, allowing for i) the sample size of each serological survey, $Y_{sero^j_{test}}(t)$ and ii) imperfect sensitivity ($p_{sero_{sens}}$) and specificity ($p_{sero_{spec}}$) of the serological assay given a binomial likelihood:

$$Y_{sero^j_{pos}}(t) \sim \text{Binom}\left(Y_{sero^j_{test}}(t), \omega_{sero_{pos}}(t)\right) \tag{3.19}$$

where:

$$\omega_{sero^j_{pos}}(t) \coloneqq \frac{p_{sero_{sens}} X_{sero^j_{pos}}(t) + \left(1 - p_{sero_{spec}}\right) X_{sero^j_{neg}}(t)}{X_{sero^j_{pos}}(t) + X_{sero^j_{neg}}(t)}. \tag{3.20}$$

The contribution to the likelihood of the serosurvey data was:

$$\mathcal{L}_{sero} = \prod_t \prod_{j=1,2} P_{\text{Binom}}\left(Y_{sero^j_{pos}}(t) \,\middle|\, X_{sero^j_{test}}(t), \omega_{sero^j_{pos}}(t)\right) \tag{3.21}$$

PCR testing

I had access to three different data sources of SARS-CoV-2 PCR testing to fit the model to (Table 3.1). This included the population-level government programme for symptomatic case detection, Pillar 2 [3], and two population-level longitudinal infection surveys [8, 9].

Using a parallel flow approach (Appendix A.1), the model assumes that results from individuals

tested through Pillar 2 PCR for each age band $z \in Z_{P2}$ ($Z_{P2} = \{[15, 25), [25, 50), [50, 65), [65, 80), [80+)\}$) were representative of the true population-level prevalence of symptomatic cases (i.e. positive and negative results that would have been observed if everyone in the corresponding age group had been tested). Thus, newly symptomatic SARS-CoV-2 cases were defined as

$$X^z_{P2_{pos}}(t) \coloneqq \sum_{i \in z} \sum_j \sum_k d^{i,j,k}_{I_P, I_{C_1}}(t) \tag{3.22}$$

and non-SARS-CoV-2 cases who had symptoms consistent with COVID-19 as

$$X^z_{P2_{neg}}(t) \coloneqq g^z(t) \left(\left(\sum_{i \in z} N^i \right) - X^z_{P2_{pos}}(t) \right), \tag{3.23}$$

where

$$g^z(t) = \begin{cases} p^z_{NC} & \text{if } t \text{ is a weekday} \\ p^z_{NCW} & \text{if } t \text{ is a weekend day} \end{cases} \tag{3.24}$$

was the probability of non SARS-CoV-2 cases in age band z having symptoms consistent with COVID-19 that might lead them to get a PCR test.

These were then compared to the observed number of positive PCR tests, $Y^z_{P2_{pos}}(t)$, accounting for the number of PCR tests conducted each day under Pillar 2, $Y^z_{P2_{test}}(t)$ and by calculating the probability of a positive PCR result (assuming perfect sensitivity and specificity of the PCR test):

$$\omega^z_{P2_{pos}}(t) \coloneqq \frac{X^z_{P2_{pos}}(t)}{X^z_{P2_{pos}}(t) + X^z_{P2_{neg}}(t)} \tag{3.25}$$

To account for the fact that people may have sought PCR testing through Pillar 2 for many other reasons which would have introduced competing biases (e.g. flu-like symptoms caused by illness other than COVID-19, before visiting acquaintances, caring for vulnerable person), the model included an over-dispersion parameter $\rho_{P2_{test}}$, fitted through a beta-binomial likelihood process:

$$Y^z_{P2_{pos}}(t) \sim \text{BetaBinom}\left(Y^z_{P2_{test}}(t), \omega^z_{P2_{pos}}(t), \rho_{P2_{test}} \right). \tag{3.26}$$

To incorporate PCR data from the two infection surveys in England at the time [8, 9], the model accounted for the included age-eligibility of each separately.

For each age band $z \in Z_{R1}$ ($Z_{R1} = \{[5, 24), [25, 34), [35, 44), [45, 54), [55, 64), \quad [65+)\}$) the model incorporated the REACT-1 PCR data into the likelihood analogously to the serology data, by considering the model-predicted number of PCR-positives, $X^z_{R1_{pos}}(t)$, and PCR-negatives, $X^z_{R1_{neg}}(t)$:

$$X^z_{R1_{pos}}(t) \coloneqq \sum_{i \in z} T^i_{PCR_{pos}}(t), \tag{3.27}$$

$$X^z_{R1_{neg}}(t) \coloneqq \left(\sum_{i \in z} N^i \right) - X^z_{R1_{pos}}(t). \tag{3.28}$$

and comparing them to the daily number of positive results, $Y^z_{R1_{pos}}(t)$, given the number of people tested on that day, $Y^z_{R1_{test}}(t)$, and by calculating the probability of a positive result, assuming perfect sensitivity and specificity of the REACT-1 assay:

$$\omega^z_{R1_{pos}}(t) \coloneqq \frac{X^z_{R1_{pos}}(t)}{X^z_{R1_{pos}}(t) + X^z_{R1_{neg}}(t)} \tag{3.29}$$

so

$$Y^z_{R1_{pos}}(t) \sim \text{Binom}\left(Y^z_{R1_{test}}(t), \omega^z_{R1_{pos}}(t)\right). \tag{3.30}$$

For ONS PCR testing data the pool of eligible individuals was those aged 2 years and over. I thus assumed homogenous distribution in the 0-4 year age group and that the fraction from this group and all other age classes in the model represented the entire population of England. The model thus considered the predicted number of PCR-positives, $X_{ONS_{pos}}(t)$, and PCR-negatives, $X_{ONS_{neg}}(t)$, in the population eligible for ONS PCR testing:

$$X_{ONS_{pos}}(t) \coloneqq \frac{3}{5} T^{[0,5)}_{PCR_{pos}}(t) + \sum_{i=[5,10)}^{[80+)} T^i_{PCR_{pos}}(t), \tag{3.31}$$

$$X_{ONS_{neg}}(t) \coloneqq \frac{3}{5} N^{[0,5)} + \left(\sum_{i=[5,10)}^{[80+)} N^i \right) - X_{ONS_{pos}}(t), \tag{3.32}$$

and compared them to the daily number of positive results observed in the ONS infection survey, $Y_{ONS_{pos}}(t)$, given the number of people tested on that day, $Y_{ONS_{test}}(t)$ and by calculating the probability of a positive result, assuming perfect sensitivity and specificity of the assay:

$$\omega_{ONS_{pos}}(t) \coloneqq \frac{X_{ONS_{pos}}(t)}{X_{ONS_{pos}}(t) + X_{ONS_{neg}}(t)} \tag{3.33}$$

so

$$Y_{ONS_{pos}}(t) \sim \text{Binom}\left(Y_{ONS_{test}}(t), \omega_{ONS_{pos}}(t)\right). \tag{3.34}$$

The contribution to the likelihood of the PCR testing data was:

$$\begin{aligned}\mathcal{L}_{PCR} = &\prod_t \prod_{z \in Z_{P2}} P_{\text{BetaBinom}}\left(Y^{\text{-}}P2_{pos}(t) \,\middle|\, Y^z_{P2_{test}}(t), \omega^z_{P2_{pos}}(t), \rho_{P2_{test}}\right) \\ &\times \prod_t \prod_{z \in Z_{R1}} P_{\text{Binom}}\left(Y^z_{R1_{pos}}(t) \,\middle|\, Y^z_{R1_{test}}(t), \omega^z_{R1_{pos}}(t)\right) \\ &\times \prod_t P_{\text{Binom}}\left(Y_{ONS_{pos}}(t) \,\middle|\, Y_{ONS_{test}}(t), \omega_{ONS_{pos}}(t)\right)\end{aligned} \tag{3.35}$$

Variant and Mutation data

A crucial aspect to this analysis was to capture the replacement of the Wildtype by the Alpha variant, of Alpha by Delta, and of Delta by Omicron BA.1. The SARS-CoV-2 genomic sequencing capacity of England was second-to-none globally [4]. In particular, the variant and mutation (VAM) data stream (Table 3.1) was a SARS-CoV-2 genotype survey collected from daily random sampling of new daily cases detected through the government's Pillar 2 programme [4]. I thus aggregated these linelist data of infecting genotypes to daily VAM frequency time series (e.g. Alpha variant *vs* non-Alpha variant in late 2020 to early 2021) and assumed it to be representative the true VAM infection frequency in the population.

During each of these variant replacement events, I was able to account for a variant pair (j_1, j_2), where j_2 was the emerging variant and assumed that genomic samples tested were newly symptomatic cases (across all age groups). The model thus simulated the number of infections for variants j_1 and j_2 given

$$X_{VAM_1}(t) \coloneqq \sum_i \sum_k d^{i,j_1,k}_{I_P,I_{C_1}} + d^{i,j_1^{reinf},k}_{I_P,I_{C_1}} \tag{3.36}$$

$$X_{VAM_2}(t) \coloneqq \sum_i \sum_k d^{i,j_2,k}_{I_P,I_{C_1}} + d^{i,j_2^{reinf},k}_{I_P,I_{C_1}} \tag{3.37}$$

which were then compared to the observed number of variant j_2 VAM test results, $Y_{VAM_2}(t)$, accounting for the combined number of variant j_1 and j_2 VAM test results each day, $Y_{VAM_{test}}(t)$, and the probability of a variant j_2 VAM test result:

$$\omega_{VAM_2}(t) \coloneqq \frac{X_{VAM_2}(t)}{X_{VAM_1}(t) + X_{VAM_2}(t)} \tag{3.38}$$

so

$$Y_{VAM_2}(t) \sim \text{Binom}\left(Y_{VAM_{test}}(t), \omega_{VAM_2}(t)\right). \tag{3.39}$$

The contribution to the likelihood of the VAM data was:

$$\mathcal{L}_{VAM} = \prod_t P_{\text{Binom}}\left(Y_{VAM_2}(t)\,|Y_{VAM_{test}}(t), \omega_{VAM_2}(t)\right) \tag{3.40}$$

Full likelihood

Given available data to fit the model to, the overall likelihood was calculated as the product of the likelihoods of the individual observations:

$$\mathcal{L} = \mathcal{L}_{adm} \times \mathcal{L}_{beds} \times \mathcal{L}_{deaths} \times \mathcal{L}_{sero} \times \mathcal{L}_{PCR} \times \mathcal{L}_{VAM}. \tag{3.41}$$

3.2.3 Effective population-level immunity against infection

Multiple drivers affect immunity against infection, from prior infection with different variants and vaccination. In order to quantify the make-up of this immunity profile at a population level, I defined a parametric calculation of the effective population-level immunity (EPI) by variant j, given vaccination, prior infection or both. It should be noted that, whilst the implementation of this calculation was done as part of my PhD project, I did not develop it independently, but rather co-developed with the ICL-RTM team.

For the susceptible S and recovered R compartments, in terms of susceptibility to infection $\chi^{i,j,k}(t)$ (see section 2.2.3), individuals could either be

1. susceptible and unvaccinated (fully susceptible)
2. susceptible and vaccinated (vaccination-derived immunity only)
3. recovered and unvaccinated (infection-derived immunity only)
4. recovered and vaccinated (both vaccination- and infection-derived immunity, i.e. "hybrid immunity").

For the latter two, note that individuals in the R compartment recovered from variants j_1 or j_2 are fully protected against infection with variant j_1, and those recovered from variant j_2 are additionally fully protected against infection with variant j_2.

Thus, we weighted individuals by $1 - \chi^{i,j,k}(t)$ relative to an unvaccinated individual in the S compartment, which weighted 0, and defined the EPI due to vaccination-derived immunity only as

$$EPI_V^j(t) = \sum_i \sum_{k>0} \left(1 - \chi^{i,j,k}(t)\right) S^{i,k}(t), \tag{3.42}$$

the EPI due to infection-derived immunity only as

$$EPI_I^j(t) = \begin{cases} \sum_i \eta_{j_1} R^{i,j_H,0}(t) + R^{i,j_1,0}(t) + R^{i,j_1^{reinf},0}(t) + R^{i,j_2,0}(t) + R^{i,j_2^{reinf},0}(t) & \text{if } j = j_1, \\ \sum_i \eta_{j_2} \left(R^{i,j_H,0}(t) + R^{i,j_1,0}(t) + R^{i,j_1^{reinf},0}(t) \right) + R^{i,j_2,0}(t) + R^{i,j_2^{reinf},0}(t) & \text{if } j = j_2, \end{cases} \tag{3.43}$$

and the EPI due to hybrid immunity as

$$EPI_H^j(t) = \begin{cases} \sum_i \sum_{k>0} \left(1 - (1-\eta_{j_1})\,\chi^{i,j_1,k}\right) R^{i,j_H,k}(t) + R^{i,j_1,k}(t) + \\ \qquad R^{i,j_1^{reinf},k}(t) + R^{i,j_2,k}(t) + R^{i,j_2^{reinf},k}(t) & \text{if } j = j_1, \\ \sum_i \sum_{k>0} \left(1 - (1-\eta_{j_2})\,\chi^{i,j_2,k}\right) \left(R^{i,j_H,k}(t) + R^{i,j_1,k}(t) + \right. \\ \qquad \left. R^{i,j_1^{reinf},k}(t) \right) + R^{i,j_2,k}(t) + R^{i,j_2^{reinf},k}(t) & \text{if } j = j_2. \end{cases} \tag{3.44}$$

Lastly, the effective population-level immunity against infection with variant j was thus given by

$$EPI^j(t) = EPI_V^j(t) + EPI_I^j(t) + EPI_H^j(t), \tag{3.45}$$

and the overall effective population-level immunity against infection, $EPI(t)$ given variants j_1 and j_2 was weighted by the model-inferred number of new infections by variant, as I presented previously for the IFR calculation in eq. (2.61).

3.2.4 Model parameters

Vaccination

The integration of vaccination into the ICL-RTM model had been led by others before my PhD [32, 89]. At the time, I collaborated to this work by leading on fitting the model to available vaccination data, alongside all other data streams.

The model's mechanisms for simulating vaccination are presented in detail in Appendix A.2. Briefly, it considers the original (i.e. non-multivalent) AZ, PF and Moderna vaccines, which were approved for use in England at the time. Given the analysis was capped to February 24, 2022, before the roll-out of fourth doses (second "boosters") or multi-valent vaccines, only six distinct vaccination strata were simulated (V_k, for $k \in \{1, 2, 3, ..., 6\}$), representing the stages of VE detailed in Table A2 and illustrated in Figure A3.

Fixed parameters

Table 3.2 presents the fixed parameters in the model.

Table 3.2: Fixed model parameter notations, values, and evidence-base. Source: Perez-Guzman et al. 2023 [1].

Parameter	Definition	Value	Source
$1/\gamma_R$	Mean duration of natural immunity following infection.	3 years	Assumed
$p_{sero_{pos}}$	Probability of seroconversion following infection.	0.85	[164]
$1/\gamma_{sero_{pre}}$	Mean time to seroconversion from onset of infectiousness.	13 days	[165]
$1/\gamma_{sero^1_{pos}}$	Mean duration of seropositivity (Euroimmun assay).	400 days	[164, 166, 167]
$1/\gamma_{sero^2_{pos}}$	Mean duration of seropositivity (Roche N).	1000 days	[164, 166, 167]
$p_{sero_{spec}}$	Specificity of serology test.	0.99	[164]
$p_{sero_{sens}}$	Sensitivity of serology test.	1	Assumed
η_{Alpha}	Relative probability of infection by *Alpha* following infection with $Wildtype$.	0.95	Assumed as in [89]
η_{Delta}	Relative probability of infection by *Delta* following infection with $Wildype$ or *Alpha*.	0.85	[168, 169]
$\eta_{Omicron}$	Relative probability of infection by *Omicron* following infection with $Wildtype$, *Delta* or *Omicron*.	0.25	Assumed based on [170, 171]
η_H^j	Relative probability (1 - eta_x^j) of hospitalisation given infection with variant j, following prior infection with a historic variant.	0.85 for *Alpha* and *Delta*, 0.55 for *Omicron*.	Assumed as in [32, 89]
η_D^j	Relative probability (1 - eta_x^j) of death given severe disease with variant j, following prior infection with a historic variant.	0.18 for *Omicron*, 0 otherwise	Assumed
θ_{I_A}	Infectivity of an asymptomatic individual, relative to a symptomatic individual.	0.223	[18]

Fitted parameters

Prior distributions are described in table 3.3. As I discussed in Section 2.2.2, the model is initialised as a single variant model and, subsequently to simulate the emergence of VOCs, expanded to a multi-strain model. Of note, for this analysis, I used uninformative (or weakly informative) prior distributions for the parameters related to variant replacements, including their transmissibility and severity cross multipliers (see Section 2.2.3 and Section 2.2.4), given an absence of literature at the time to indicate otherwise.

Table 3.3: Inferred model parameter notations and prior distributions for England analysis. *Time-varying severity parameters with piece-wise linear form, with structure and dates for timepoints as described in Section 2.2.4. Source: Perez-Guzman et al. 2023 [1]

Parameter	Description	Prior distribution	95% probability interval	Rationale
$t_{Wildtype}$	Start date of regional outbreak (dd/mm/2020)	$U[01/01, 15/03]$	(01/02 - 13/03)	Wide range of dates in early 2020 before our first data point
t_{Alpha}	Regional Alpha seeding date (dd/mm/2020)	$U[17/09/2020, 03/01/2021]$	(17/09/2020 - 03/01/2021)	16-week window around first Alpha case detection in England
t_{Delta}	Regional Delta seeding date (dd/mm/2021)	$U[08/03, 24/07]$	(08/03 - 24/07)	20-week window around first Delta case detection in England
$t_{Omicron}$	Regional Omicron seeding date (dd/mm/2021)	$U[14/09, 01/01]$	(14/09/2021 - 01/01/2022)	20-week window around first Omicron case detection in England
$\sigma_{Alpha/Wildtype}$	Transmission advantage of Alpha over Wildtype	$U(0, 3)$	(0.075 - 2.925)	Wide range capturing epidemiologically plausible transmission advantages
$\sigma_{Delta/Alpha}$	Tansmission advantage of Delta over Alpha	$U(0, 3)$	(0.075 - 2.925)	As above
$\sigma_{Omicron/Delta}$	Transmission advantage of Omicron over Delta	$U(0, 3)$	(0.075 - 2.925)	As above
$\beta(t)$	Transmission rate (pp) at $t = dd/mm/yy$			
β_1	16/03/20: PM advises WFH and essential travel only	$\Gamma(136, 0.0008)$	(0.0918,0.0128)	Range corresponding to a basic reproduction numbers between 2.5 and 3.5 consistent with [16, 172]

Continued on next page

Table 3.3 – continued from previous page

	Description	Prior distribution	95% probability interval	Rationale
β_2	23/03/20: PM announces lockdown 1	$\Gamma(3.73, 0.0154)$	(0.0147,0.128)	Corresponding to R0 between 0.9 and 3.5, consistent with a 0% to 75% relative decrease from R0
β_3	25/03/20: Lockdown 1 into full effect	$\Gamma(4.25, 0.0120)$	(0.0147,0.110)	Corresponding to R0 between 0.4 and 3, allowing a further decrease in contact rates due to NPIs
β_4	11/05/20: Initial easing of lockdown 1	$\Gamma(4.25, 0.0120)$	(0.0147,0.110)	As above
β_5	15/06/20: Non-essential shops re-open	$\Gamma(4.25, 0.0120)$	(0.0147,0.110)	As above
β_6	04/07/20: Hospitality re-opens	$\Gamma(4.25, 0.0120)$	(0.0147,0.110)	As above
β_7	01/08/20: "Eat out to help out" scheme starts	$\Gamma(4.25, 0.0120)$	(0.0147,0.110)	As above
β_8	01/09/20: Schools and universities re-open	$\Gamma(4.25, 0.0120)$	(0.0147,0.110)	As above
β_9	14/09/20: "Rule of six" introduced	$\Gamma(4.25, 0.0120)$	(0.0147,0.110)	As above
β_{10}	14/10/20: Tiered system introduced	$\Gamma(4.25, 0.0120)$	(0.0147,0.110)	As above
β_{11}	31/10/20: Lockdown 2 announced	$\Gamma(4.25, 0.0120)$	(0.0147,0.110)	As above
β_{12}	05/11/20: Lockdown 2 starts	$\Gamma(4.25, 0.0120)$	(0.0147,0.110)	As above
β_{13}	02/12/20: Lockdown 2 ends	$\Gamma(4.25, 0.0120)$	(0.0147,0.110)	As above
β_{14}	18/12/20: School holidays start	$\Gamma(4.25, 0.0120)$	(0.0147,0.110)	As above
β_{15}	25/12/20: Last day of holiday season relaxation	$\Gamma(4.25, 0.0120)$	(0.0147,0.110)	As above
β_{16}	05/01/21: Lockdown 3 starts	$\Gamma(4.25, 0.0120)$	(0.0147,0.110)	As above
β_{17}	08/03/21: Roadmap step one - schools reopen	$\Gamma(4.25, 0.0120)$	(0.0147,0.110)	As above
β_{18}	01/04/21: School holidays	$\Gamma(4.25, 0.0120)$	(0.0147,0.110)	As above

Continued on next page

Table 3.3 – continued from previous page

	Description	Prior distribution	95% probability interval	Rationale
β_{19}	19/04/21: Roadmap step two - outdoor rule of 6 (12/04) and schools re-open (19/04)	$\Gamma(4.25, 0.0120)$	(0.0147,0.110)	As above
β_{20}	17/05/21: Roadmap step three - Indoor hospitality opens	$\Gamma(4.25, 0.0120)$	(0.0147,0.110)	As above
β_{21}	21/06/21: Wedding and care home restrictions eased	$\Gamma(4.25, 0.0120)$	(0.0147,0.110)	As above
β_{22}	03/07/21: Euro 2020 quarter finals (cited as significant influence [173])	$\Gamma(4.25, 0.0120)$	(0.0147,0.110)	As above
β_{23}	11/07/21: End of Euros football tournament	$\Gamma(4.25, 0.0120)$	(0.0147,0.110)	As above
β_{24}	19/07/21: Full lift of NPIs	$\Gamma(4.25, 0.0120)$	(0.0147,0.110)	As above
β_{25}	15/08/21: Summer festivals/holidays	$\Gamma(4.25, 0.0120)$	(0.0147,0.110)	As above
β_{26}	01/09/21: Schools return	$\Gamma(4.25, 0.0120)$	(0.0147,0.110)	As above
β_{27}	22/09/21: Mid-point between school start and half term	$\Gamma(4.25, 0.0120)$	(0.0147,0.110)	As above
β_{28}	01/10/21: Point before sharp increase in epidemic wave	$\Gamma(4.25, 0.0120)$	(0.0147,0.110)	As above
β_{29}	22/10/21: School holidays (half term)	$\Gamma(4.25, 0.0120)$	(0.0147,0.110)	As above
β_{30}	01/11/21: School return	$\Gamma(4.25, 0.0120)$	(0.0147,0.110)	As above
β_{31}	08/12/21: Plan B announced	$\Gamma(4.25, 0.0120)$	(0.0147,0.110)	As above
β_{32}	23/12/21: School holidays start	$\Gamma(4.25, 0.0120)$	(0.0147,0.110)	As above
β_{33}	04/01/22: Schools return	$\Gamma(4.25, 0.0120)$	(0.0147,0.110)	As above
β_{34}	19/01/22: Announcement of end of Plan B	$\Gamma(4.25, 0.0120)$	(0.0147,0.110)	As above

Continued on next page

Table 3.3 – continued from previous page

	Description	Prior distribution	95% probability interval	Rationale
β_{35}	27/01/22: End of Plan B	$\Gamma(4.25, 0.0120)$	(0.0147,0.110)	As above
β_{36}	24/02/22: End of self-isolation policy in England, end of fits	$\Gamma(4.25, 0.0120)$	(0.0147,0.110)	As above
$p_{H,1}^{max}$, $p_{H,2}^{max}$	The probability of symptomatic individuals developing serious disease requiring hospitalisation, for the group with the largest probability at first and second timepoint, respectively*	$U(0,1)$	(0.025,0.975)	Uninformative
$p_{G_D,1}^{max}$, $p_{G_D,2}^{max}$	Probability of death in the community given disease severe enough for hospitalisation for the group with the largest probability at first and second timepoint, respectively*	$U(0,1)$	(0.025,0.975)	Uninformative
$p_{ICU,1}^{max}$	Probability of triage to ICU for new hospital admissions, for the group with the largest probability at first timepoints*	$B(13.9, 43.9)$	(0.140,0.357)	Informed by previous work [18, 32] based on [149]
$p_{ICU,2}^{max}$	Probability of triage to ICU for new hospital admissions, for the group with the largest probability at second timepoint*	$U(0,1)$	(0.025,0.975)	Uninformative
$p_{H_D}^{max}$	Initial probability of death for general inpatients	$B(42.1, 50.1)$	(0.356,0.558)	Informed by previous work [18, 32] based on [149]

Continued on next page

Table 3.3 – continued from previous page

	Description	Prior distribution	95% probability interval	Rationale
$p^{max}_{ICU_D}$	Initial probability of death for ICU inpatients	$B(60.2, 29.3)$	(0.573,0.766)	Informed by previous work [18, 32] based on [149]
$p^{max}_{W_D}$	Initial probability of death for step-down inpatients	$B(28.7, 52.1)$	(0.255,0.462)	Informed by previous work [18, 32] based on [149]
$\mu_{D,1}, \mu_{D,2}$, $\mu_{D,3}, \mu_{D,4}$, $\mu_{D,5}$	Hospital mortality multipliers due to changes in clinical care at different timepoints*	$U(0, 1)$	(0.025,0.975)	Uninformative
$\pi_H^{Alpha/Wildtype}$	Multiplier of the probability of hospitalisation with Alpha relative to the Wildtype variant*	$U(0, 3)$	(0.075,2.925)	Wide range capturing epidemiologically plausible intrinsic changes in the severity of SARS-CoV-2
$\pi_H^{Delta/Alpha}$	Multiplier of the probability of hospitalisation with Delta relative to the Alpha variant*	$U(0, 3)$	(0.075,2.925)	As Above
$\pi_H^{Omicron/Delta}$	Multiplier of the probability of hospitalisation with Omicron relative to the Delta variant*	$U(0, 3)$	(0.075,2.925)	As Above
$\pi_{ICU}^{Alpha/Wildtype}$	Multiplier of the probability of ICU admission with Alpha relative to the Wildtype variant*	$U(0, 3)$	(0.075,2.925)	Wide range capturing epidemiologically plausible intrinsic changes in the severity of SARS-CoV-2

Continued on next page

Table 3.3 – continued from previous page

	Description	Prior distribution	95% probability interval	Rationale
$\pi_{ICU}^{Delta/Alpha}$	Multiplier of the probability of ICU admission with Delta relative to the Alpha variant*	$U(0,3)$	(0.075,2.925)	As Above
$\pi_{ICU}^{Omicron/Delta}$	Multiplier of the probability of ICU admission with Omicron relative to the Delta variant*	$U(0,3)$	(0.075,2.925)	As Above
$\pi_{D}^{Alpha/Wildtype}$	Multiplier of the probability of death with Alpha relative to the Wildtype variant*	$U(0,3)$	(0.075,2.925)	Wide range capturing epidemiologically plausible intrinsic changes in the severity of SARS-CoV-2
$\pi_{D}^{Delta/Alpha}$	Multiplier of the probability of death with Delta relative to the Alpha variant*	$U(0,3)$	(0.075,2.925)	As Above
$\pi_{D}^{Omicron/Delta}$	Multiplier of the probability of death with Omicron relative to the Delta variant*	$U(0,3)$	(0.075,2.925)	As Above
$\mu_{\gamma_H,1}$, $\mu_{\gamma_H,2}$, $\mu_{\gamma_H,3}$, $\mu_{\gamma_H,4}$	Mean duration multipliers for non-ICU hospital compartments at different timepoints*	$U(0,2)$	$(0.050, 1.950)$	Wide range capturing epidemiologically plausible relative changes in durations
$p_{NC}^{[15,25)}$, $p_{NC}^{[25,50)}$, $p_{NC}^{[50,65)}$, $p_{NC}^{[65,80)}$, $p_{NC}^{[80+)}$	Prevalence of non-COVID symptomatic illness that could lead to getting a PCR test for different age bands	$U(0,1)$	(0.025,0.975)	Uninformative

Continued on next page

Table 3.3 – continued from previous page

	Description	Prior distribution	95% probability interval	Rationale
$p_{NCW}^{[15,25)}$, $p_{NCW}^{[24,50)}$, $p_{NCW}^{[50,65)}$, $p_{NCW}^{[65,80)}$, $p_{NCW}^{[80+)}$	Prevalence of non-COVID symptomatic illness that could lead to getting a PCR test on a weekend for different age bands	$U(0,1)$	(0.025,0.975)	Uninformative
$\rho_{P2_{test}}$	Overdispersion of PCR positivity	$U(0,1)$	(0.025,0.975)	Uninformative
α_A	Overdispersion for hospital admission data streams	$U(0,1)$	(0.025,0.975)	Uninformative
α_H	Overdispersion for hospital bed streams	$U(0,1)$	(0.025,0.975)	Uninformative
$\alpha_{D_{hosp}}$	Overdispersion for hospital death data streams	$U(0,1)$	(0.025,0.975)	Uninformative
$\alpha_{D_{comm}}$	Overdispersion for community death data streams	$U(0,1)$	(0.025,0.975)	Uninformative

3.2.5 Running the model

I ran and fitted the model independently for each NHS England region by fitting it to data (Table 3.1) between March 16, 2020 and February 24, 2022. Given the multi-dimensionality of the parameter space over which I conducted inference (see Table 3.3), I relied on an innovative, multi-step model fitting approach pioneered by our group. This involved a first step pre-tuning of the model's parameters with a deterministic version of the model and traditional MCMC, followed by the full stochastic model implementation and pMCMC inference.

Briefly, the idea was to run a computationally efficient inference (with the use of adaptive MCMC) on the equivalent deterministic model, in order to obtain a good initial parameter conditions and proposal distributions for the pMCMC. I describe the equations of the deterministic version of the model in Appendix A.1, where the mean of the proposed distribution for a given parameter is used instead of a random draw.

For the present analysis, I ran the deterministic model with eight adaptive MCMC chains of 30,000 steps, each initialised from a different place of the 85-dimensional (table 3.3) parameter space. I assessed convergence of the parameter samples produced with

- Visual inspection of the MCMC chains (Appendix A.5), looking for a caterpillar appearance of the chains and an asymptotic appearance of the log-likelihood curve with coverage of an area of the parameter space
- Gelman-Rubin statistic associated with the full parameter chains with $\hat{R}$ below the recommended 1.1 threshold
- Effective Sample Size (ESS) of the full parameter chains above 1,000

I then used the pre-tuned parameter set as the initial values and proposal kernel, respectively, to run the stochastic model and pMCMC inference detailed in Appendix A.1. I ran the latter with four chains of 3,000 steps each.

Of note, the pMCMC uses a re-scaled version of the posterior VCV matrix generated by the MCMC as its VCV proposal matrix, in order to capture the main correlation structure and the "scale" of the different parameters, where diagonal elements of the VCV matrix represent the variances of the different parameters.

The final pMCMC model fit was produced through approximately 23,000 CPU hours, spread across 7 nodes of 32-core Xeons (dual 16-core 2.6 GHz). All necessary code and scripts to reproduce my analyses were made public at the time of publication, and can be found in GitHub.

3.3 Results

3.3.1 Transmissibility

I found that the R_0 of SARS-CoV-2 increased with the emergence of each variant in succession, from 2.6 (95% credible interval (CrI) 2.4 – 2.8) for the initial Wildtype virus, to 4.2 (95% CrI 4.0 – 4.6), 7.2 (95% CrI 6.7 – 7.8) and 8.4 (95% CrI 7.8 – 9.1), respectively, for the Alpha, Delta and Omicron (BA.1) variants (Figure 3.1d). The R_t, on the other hand, varied as contact rates decreased or increased in response to NPIs (Figure 3.1f). Before the COVID-19 vaccination programme started, R_t^{eff} was only slightly lower than R_t, indicating very low levels of infection-induced immunity in the population. Accounting for waning of infection-induced immunity, only 7.9% (95% CrI 7.4 – 8.3) of the population were immune by early September 2020, before the emergence of the Alpha variant (Figure 3.1g).

With the roll-out and scale-up of the national vaccination programme from December 2020 onwards, immunity against infection in the population rapidly rose from 12.6% (95% CrI 11.9 – 13.2), to 33.9% (95% CrI 32.9 – 35.0) in March 2021. Notably, this period of time coincided with the emergence and dominance of the more transmissible Alpha and Delta variants. From this point onward, R_t increased gradually as NPIs were lifted in the first half of 2021, leading to a sustained high level of Delta infections for between June and October 2021 (Figure 3.1a). R_t^{eff} estimates, however, remained close to 1 (Figure 3.1f) suggesting increases in contact rates were balanced out by increasing immunity, primarily vaccine-derived, for an extended period (Figure 3.1g).

In contrast, the emergence and dominance of the initial Omicron BA.1 variant marked a decline in effective population immunity (Figure 3.1g). The relative increase in Omicron's R_0 compared to Delta was not as high compared to previous variant replacements (Figure 3.1d), but it was paired with Omicron's antigenic divergence and, thus, immune escape properties. Population-level immunity against infection decreased by a third relative to its late November peak, to 47.0% (95% CrI 45.4 – 48.6) in late December 2021 (Figure 3.1g). As a result, infection levels rose to an all-time high, with re-infections accounting for up to a third of daily new infections (Figure 3.1c).

The re-implementation of some NPIs (work from home and wearing face masks and/or showing proof of negative COVID-19 tests in public venues) between December 12, 2021 and January 27, 2022 achieved a decrease in R_t (Figure 3.1f). However, it was the combined effect of rapidly rolling out booster vaccinations and infection-induced immunity, that yielded an increased population-level immunity and, consequently, a sustained decrease in R_t^{eff} to around 1 (Figure 3.1f-g) from January 2021 onward.

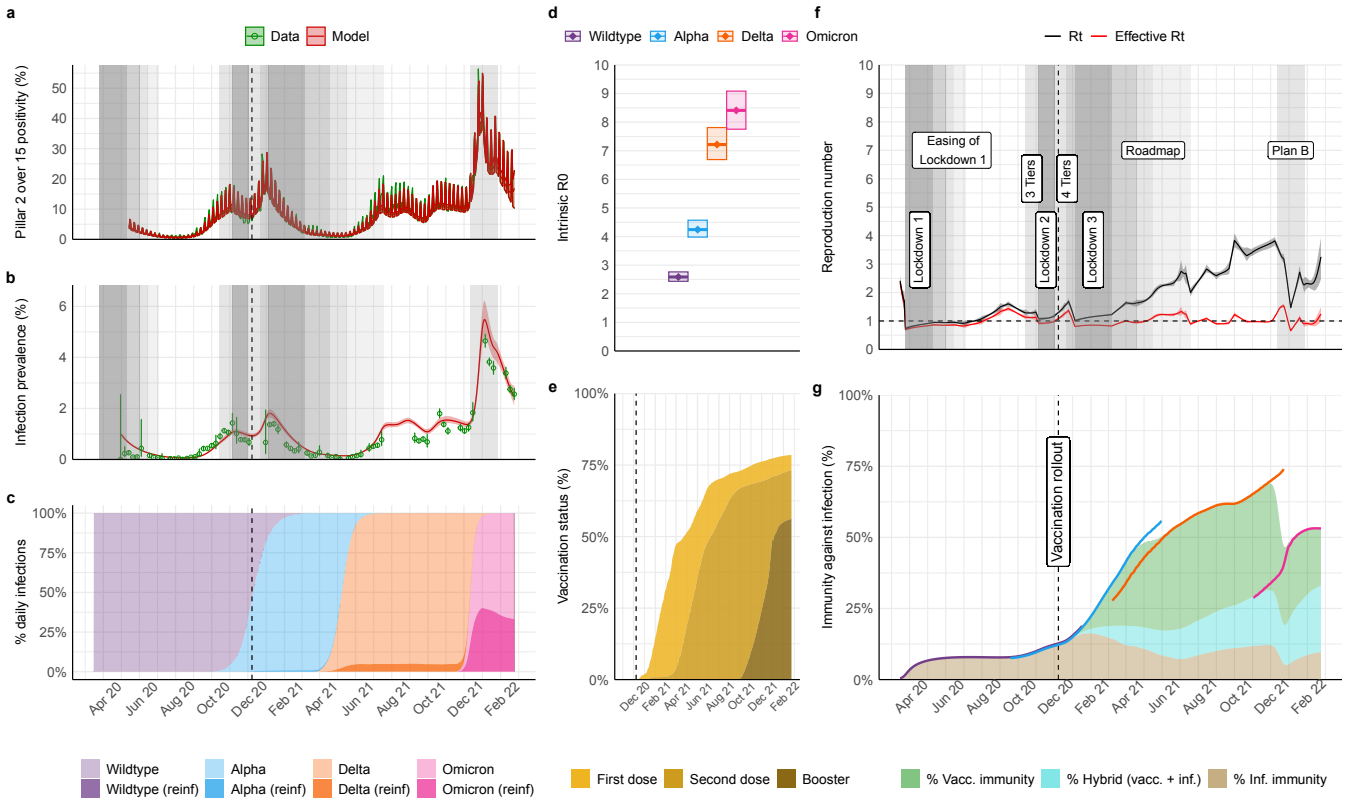


Figure 3.1: Population-level transmission of SARS-CoV-2 between March 2020 and February 2022 in England. a) Infection positivity amongst individuals 15 years and older in the community through the national PCR testing programme (Pillar 2). b) Infection prevalence (PCR) in representative samples of households from the REACT-1 study. Panels a and b show data in green (lines and point, with binomial 95% confidence intervals in error bars) and model trajectory in red (average and 95% credible interval (95% CrI)). Grey shading indicates periods of non-pharmaceutical interventions (NPIs) of interest for the analysis; see in f. c) Model-inferred average frequency of daily infections by variant and type of infection (either primary or re-infection following any prior infection, "reinf"). d) Intrinsic basic reproduction number (R_0) estimates by variant (mean and 95% CrI). e) Model trajectory of vaccine status of the national population (all ages as denominator), as informed by official data of daily doses administered; transition between vaccination classes was modelled stochastically to capture smooth changes in population-level immunity over time. f) Model trajectories of the instantaneous reproduction number in the absence of the effect of immunity (R_t) or accounting for immunity (R_t^{eff}). Legends and grey areas specify date and duration of official NPIs in England over the study period. g) Inferred effective levels of immunity in the population (all ages as denominator). Lines correspond to the effective immunity against specific variants, with colour scheme as in c and d, whereas areas indicate overall effective immunity by type of immunity (vacc: vaccine-induced, inf: from prior infection, or hybrid). Note that the coloured areas corresponding to the different types of immunity cover different periods of variant dominance and should be interpreted in the context of the circulating variants. During periods of variant replacement (e.g. Alpha to Delta) the effective immunity transitions from the levels associated with the variant being replaced to the level of the variant that becomes dominant. Source: Perez-Guzman et al. 2023 [1].

3.3.2 Severity

The basic IHR of SARS-CoV-2 increased from 2.2% (95% CrI 2.1 – 2.3) for Wildtype, to 3.4% (95% CrI 3.1 – 3.6) for Alpha, 3.4% (95% CrI 3.1 – 34.64) for Delta, but subsequently decreased to 4.2% (95% CrI 3.8 – 4.4) for Omicron (Figure 3.2a). The overall risk of dying for hospitalised patients,

however, was highest for the Alpha variant, with a basic HFR of 50.3% (95% CrI 46.7 – 53.6), lower for Wildtype and Delta, at 33.0% (95% CrI 30.7 – 35.3) and 32.3% (95% CrI 27.4 – 35.8), respectively, and lowest for Omicron (BA.1), at 15.5% (95% CrI 13.1 – 18.8) (Figure 3.2b). Lastly, the basic IFR was highest for the Alpha variant at 3.0% (95% CrI 2.8 – 3.2), followed by Delta at 2.1% (95% CrI 1.9 – 2.4), Wildtype at 1.2% (95% CrI 1.1 – 1.2), and Omicron at 0.7% (95% CrI 0.5 – 1.2) (Figure 3.2c).

Despite the above, the effective observable severity (given healthcare provision, immunity, and age-dependent mixing patterns) of the pandemic generally decreased over time. Substantial decreases were seen in response to improvements in healthcare practices (Figure 3.2a-c) and, more so, with the introduction and roll-out of effective vaccines (Figure 3.2d-f). Notably, improved healthcare at and after the peak of the first wave (e.g. remdesivir and dexamethasone, major re-adaptation of hospital capacity to manage severe COVID-19 cases) [128] signalled a halving in the effective HFR and IFR of the pandemic to much lower levels than the intrinsic severity of the then-dominant Wildtype variant (Figure 3.2b-c).

On the other hand, pressures on the healthcare system had a detrimental impact on the severity of the pandemic. Between October 2020 and January 2021 infection levels increased gradually in relation to the ineffectiveness of the multi-tiered regional NPIs system introduced by the UK Government at the time. In turn, all effective severity metrics gradually increased, particularly the HFR (Figure 3.2a-c). Critically, I found there was an overall higher risk of death in hospitals, independently of the basic severity properties of the Alpha variant. Given there was no representative data available on hospital deaths by variant during this period to fit our model to, however, I could not parameterise the model to fully differentiate the specific contribution of the effects of the increase in Alpha variant frequency and healthcare effects.

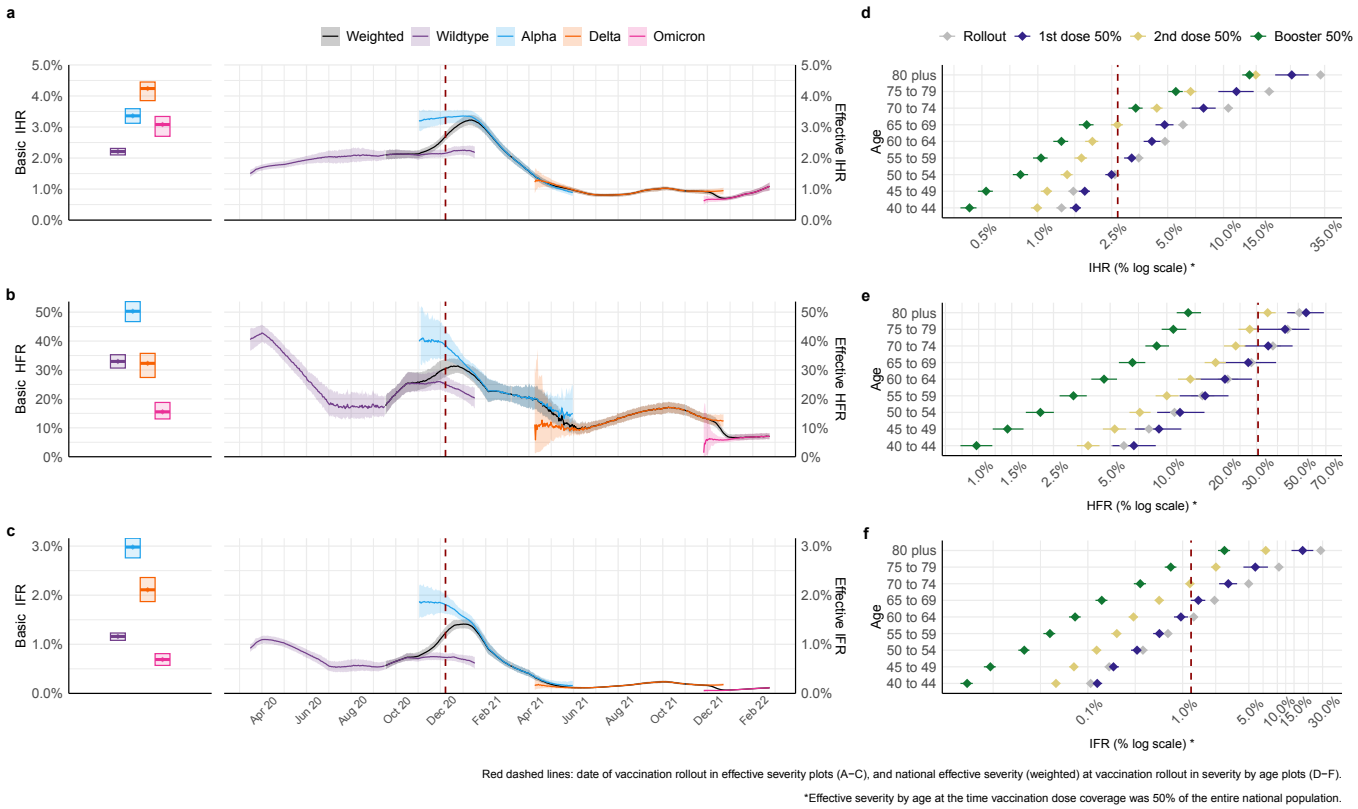


Figure 3.2: Inferred severity of SARS-CoV-2 variants in England between March 2020 and February 2022. a–c Inferred basic and effective infection hospitalisation ratio (IHR, a), hospital fatality ratio (HFR, b) and infection fatality ratio (IFR, c). Boxplots to the left show the basic severity measure and trajectories to the right show severity measure over time (mean and 95% credible interval, 95% CrI) for each variant, with black showing the weighted average across co-circulating variants at any time. Basic severity is measured for each variant assuming healthcare characteristics of the early epidemic, to allow a like-for-like comparison. Effective severity trajectories account for changing vaccine- and infection-induced protection against severe disease, as well as underlying healthcare variations. d–f Age-specific (selected age groups) effective IHR (d), HFR (e) and IFR (f), assessed at the date of key milestones of the national vaccination programme, "roll-out" refers to the start of the vaccination programme on 8th December 2020. Source: Perez-Guzman et al. 2023 [1].

3.4 Discussion

The study period for my analysis captured the SARS-CoV-2 epidemic, up to the official end of the national policy for self-isolation, capturing the emergence phase of all main variants in England. Estimating the relative and combined role of multiple competing factors shaping pandemic transmissibility and severity was challenging. Multiple dynamic interacting processes were involved, many of which are not directly measurable through traditional surveillance methods [113].

The intrinsic transmission advantage of Alpha and Delta allowed them to become dominant, while Omicron's additional immune escape properties reduced effective population-level immunity against infection, enabling it to become dominant despite high vaccine-derived immunity in the population at the time. More importantly, I found strong evidence that, whilst the tiered regional approach to the implementation and lifting of NPIs in England between October 2020 and January 2021 was

ineffective at controlling virus transmission, a national level approach to their re-introduction and gradual lifting during early to mid 2021 and partial reintroduction in late 2021 were successful at this. These findings underscore that spells of gradual implementation and lifting of NPIs must be carefully assessed in context with concurrent population-level immunity levels against an emerging pandemic pathogen.

On the other hand, I found that the Alpha variant had the highest basic HFR and IFR. This contrasted with evidence from observational studies at the time, which suggested a higher risk of hospitalisation and death in individuals infected with Delta compared to Alpha [174–177]. It should be noted, however, observational studies can be biased when estimating variant severity metrics over time, as they are inherently restricted to periods when variants co-circulate, and given population healthcare seeking behaviours during such period. In this sense, estimates of severity during the early phase of Delta emergence, when Alpha co-circulated, would not have characterised the effect of partial replacement of the original Delta variant by the AY.4.2 sub-variant, which showed a decreased severity [178]. A critical limitation of my estimates, though, was the lack of representative healthcare data to fit the model to hospital admissions and deaths by variant infection status. The model could have, thus, partly weighted higher the severity of Alpha if significant healthcare factors I did not account for occurred between the time of Alpha emergence and before it became dominant.

For settings with high surveillance capacity, such as England, representative genomic surveillance across the spectrum of pandemic disease severity will be crucial going forward in order to better characterise such factors. Importantly, for the case of a public health response to a pandemic emergency, my results strongly highlighted that improvements in clinical healthcare can be nullified if not paired with effective and timely interventions to control pathogen transmission rates in the population. As I have now alluded to, NPIs in England were not always effective or timely. For the case of the winter of 2020/21, this led to a peak in hospitalisations and, in turn, in the $HFR^{eff}(t)$ which could not be explained solely by the severity characteristics of the Wildtype and Alpha variants, nor changes in the age distribution of hospital admissions.

My analysis also demonstrated the critical role of risk-prioritised vaccination programmes to reduce pandemic severity alongside transmission. Purely infection-driven immunity in the population remained low from the start of the pandemic, through to February 2022. At the point of Omicron emergence, nevertheless, the accrued profile of combined infection- and vaccine-induced immunity at the population level were crucial to mitigate the burden on the health system in the absence of any new lockdowns or other stringent NPIs. Analyses looking forward are needed to better characterise the role of booster vaccines, optimised healthcare performance and NPIs to achieve sustained pandemic control as emerging pandemic pathogens become endemic. Comprehensive quantitative frameworks, such as the one I developed here, will be crucial to integrate evidence from multiple data

streams and derive robust inference on the transmissibility and severity of pandemic pathogens.

3.5 Conclusion

Such rich surveillance as the one I employed for this analysis, however, was not the norm for other settings, particularly LICs. In the following chapter, I thus developed a parsimonious adaptation of my SARS-CoV-2 inference framework and applied it to study the transmissibility and severity dynamics of the COVID-19 pandemic in Kabwe Zambia, an LIC setting with limited COVID-19 surveillance capacity and where the access to and capacity to implement vaccination- and healthcare improvement-based strategies was severely impaired.

Chapter 4

Drivers of SARS-CoV-2 transmissibility and severity in Kabwe, Zambia

4.1 Background

In Chapter 3, I presented how the parametric framework I developed allowed me to assess the time-varying transmission and severity dynamics of SARS-CoV-2 in detail. Many of the innovations I implemented for this analysis were absorbed back into the ICL-RTM model, which fed into the real-time analytics submitted to SPI-M-O and SAGE to inform the UK Government's COVID-19 response. Such granular analytic innovations were made possible largely through the availability of a wide range of reliable surveillance data streams. Reliable surveillance and real-time analytics to inform decision-making during the pandemic, however, was not available in most resource limited settings [42, 71].

At the time I began the work summarised in this chapter, epidemiological studies in most African LMICs had invariably relied on official test-positive cases, hospitalisations and/or deaths counts to assess the local dynamics and burden of the COVID-19 pandemic [179–183]. Such an approach had been then demonstrated to yield underestimates of pandemic burden, mainly due to systematic underascertainment issues [91, 93]. In the aftermath of the first two years of the pandemic, a global comparison of official COVID-19 case counts to time- and geography-linked serosurveys indicated that up to 99% of infections were not being detected by official surveillance systems in African LICs [133]. Similarly, a global excess mortality study reported that official COVID-19 death counts in sub-Saharan Africa only accounted for around 7% of the likely true death toll of the pandemic [129]. Limited, delayed and, potentially, erroneous analytics to inform response strategies was indeed a widespread challenge of many LMICs that depended on global pandemic financing mechanisms, rooted in inequalities in the sharing of basic healthcare and surveillance testing commodities [42].

Such an incomplete picture of the dynamics and burden of the pandemic COVID-19 in LMICs will further have grave implications for informing their preparedness strategies to future outbreak emergencies. In this chapter, I thus present a case study of the adaptation of my modelling framework to exploit the sparse surveillance data typical of an LIC setting. With the aim of assessing the drivers of transmissibility, severity and likely true burden of SARS-CoV-2 in a setting with severe healthcare constraints, I analysed COVID-19 seroprevalence and all-cause mortality data from Kabwe district between March 2020 and September 2021. This period captured the first three local pandemic waves, known to have been dominated by the Wildtype, Beta and Delta variants, respectively [54]. This work was the result of a research collaboration I instigated and led with Zambart and the ZNPHI during the second and third years of my PhD.

4.2 Methods

Based on the approach I described in Chapters 2 and 3, I developed a flexible, parsimonious compartmental SARS-CoV-2 transmission model to infer the transmissibility and severity dynamics of the pandemic, under limited surveillance data constraints. In this methods sections, I provide details of the model's structure, its parameters and fitting approach.

4.2.1 COVID-19 surveillance data

For the first local pandemic wave, I relied on a single data point of IgG prevalence from Kabwe from a random representative sample of individuals 5-years and older collected on July 17, 2020 [46]. For the second wave, I used age-disaggregated seroprevalence data from a previously published household cluster weekly seroprevalence survey conducted in Kabwe between November 2020 and February 2021 [59]. Lastly, for the third wave, I had access to previously unpublished age-disaggregated seroprevalence data from the same surveyed population between June and July 2021.

Throughout, to inform mortality in the model, I relied on an aggregated time series of all-cause deaths, which was purposely collected from paper-based mortuary records in Kabwe as part of a multi-district pandemic excess mortality study [52]. To triangulate these data and calculate excess mortality (see Section 4.2.4), I also accessed monthly aggregated all-cause death counts from Kabwe between January 2017 and December 2020. These came from routine reports to the ZNPHI, collected as part of the same excess mortality study [52].

Table 4.1: SARS-CoV-2 pandemic surveillance data available for analysis from Kabwe, Zambia between March 2020 and September 2021. Source: Perez-Guzman et al. [65].

Data type	Description	Source
Seroprevalence	Aggregated data (all-ages) IgG prevalence in randor sample of individuals from Kabwe on 2020-07-17	[46]
Seroprevalence	Age-disaggregated (15-19, 20-29, 30-39, 40-49, 50+) weekly IgG prevalence in Ngungu-Bwacha, Kabwe across two cross-sectional samples:	
	From 2020-11-18 to 2021-02-24	[59]
	From 2021-06-23 to 2021-07-28	[65]
Historic deaths	Monthly all-cause death counts in health facilities from 2017 to 2020.	[52]
Pandemic period deaths	Daily all-cause death counts by place of death (community vs health facility) from 2020 to 2022.	[52]

4.2.2 Model overview

Based on the modelling framework described in Chapters 2 and 3, I developed a stochastic, compartmental model of SARS-CoV-2 transmission paired with a pMCMC inference framework. The full model equations are presented in Appendix B.3. Briefly, the population was disaggregated into 16 age groups ($\{[0-4],[5-9],\ldots,[65-69],70+\}$) and five vaccination classes (unvaccinated, 1st dose - no effect, 1st dose - full effect, 2nd dose - full effect, 2nd dose - waned) (Figure 4.1).

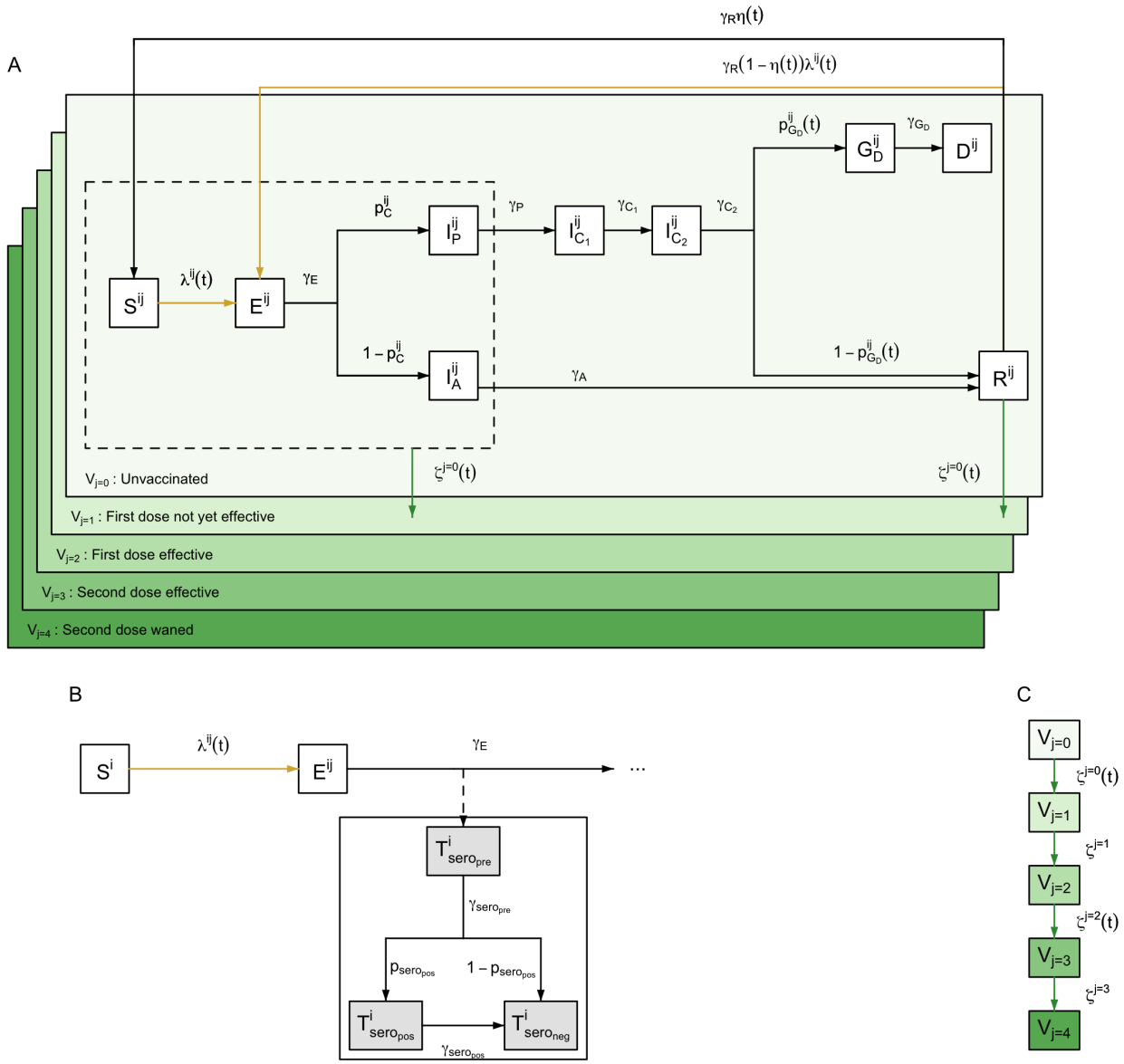


Figure 4.1: Model diagram of SARS-CoV-2 transmission applied to surveillance data from Kabwe, Zambia between March 2020 and September 2021. A) Overview of model compartments and transition rates; B) parallel flows developed to flexibly accomodate for serology data; C) detail of vaccine progression classess: $j = 0$ for unvaccinated, $j = 1$ for 1st dose - no effect, $j = 2$ for 1st dose - full effect, $j = 3$ for 2nd dose - full effect, and $j = 4$ for 2nd dose - waned. i denotes age groups ($\{[0-4),[5-10),\ldots,[70-75),75+\}$). Source: Perez-Guzman et al. 2024 [65].

Table 4.2: Definitions of model compartments shown in Figure 4.1, with indices: i for age group ($i \in \{[0,5),[5,10),\ldots,[70-75),[75+)\}$), and j for vaccination strata ($j \in \{0,1,\ldots 4\}$. Durations have Erlang (k_X, γ_X) distributions, with mean k_X/γ_X. Parallel flows compartments are described below in Section 4.2.3. Source: Perez-Guzman et al. [65].

		Erlang duration parameters		
Compartment	Definition	k_X	γ_X (per day)	Mean (days)
$S^{i,j}(t)$	Susceptible	Determined by infection dynamics		
$E^{i,j}(t)$	Exposed (latent infection, not infectious)	2	0.87	2.31
$I_A^{i,j}(t)$	Asymptomatic infected (infectious)	1	0.35	2.88
$I_P^{i,j}(t)$	Presymptomatic infected (infectious)	1	0.60	1.68
$I_{C_1}^{i,j}(t)$	Symptomatic infected (infectious)	1	0.47	2.14
$I_{C_2}^{i,j}(t)$	Symptomatic infected (not infectious)	1	0.54	1.86
$G_D^{i,j}(t)$	Severely diseased, leading to death (infectious)	2	0.67	3.00
$D^{i,j}(t)$	Deceased (as a result of COVID-19)	-	-	-
$R^{i,j}(t)$	Recovered	1	1/1095	1095

Table 4.3: Definitions of model parameters shown in Figure 4.1. These parameters define the routes of progression through model compartments defined in Table 4.2 with indices: i for age group ($i \in \{[0,5),[5,10),\ldots,[70-75),[75+)\}$), and j for vaccination strata ($j \in \{0,1,\ldots 4\}$. Source: Perez-Guzman et al. [65].

Parameter	Definition
$\lambda^{i,j}(t)$	Force of infection.
γ_x	Rate of progression from compartment x.
$p_C^{i,j}$	Probability of being symptomatic if infected.
$p_{G_D}^{i,j}(t)$	Probability of death conditional on severe symptomatic disease.
$\zeta^{i,j}(t)$	Rate of movement from vaccine strata j to $j+1$.
$\eta(t)$	Cross-protection from prior infection.

4.2.3 Observation process

Seroprevalence

To model seroprevalence, I implemented a parallel flow (Figure 4.1B), as proposed by Baguelin et al. [184]. This flow simulates IgG seroconversion as a result of infection whereby, after a pre-conversion period $T_{sero_{pre}}$ individuals can enter a seropositive compartment $T_{sero_{pos}}$ or not $T_{sero_{neg}}$. For those that enter the seropositive compartment, they eventually serorevert to $T_{sero_{neg}}$.

Table 4.4: Definitions of model parallel flow compartments shown in Figure 4.1B, with index i for age group ($i \in \{[0,5),[5,10),\ldots,[70-75),[75+)\}$). Durations (in days) have Erlang (k_X, γ_X) distributions, with mean k_X/γ_X. *Assay used for serosurvey data (see Table 4.1) was Abbott Architect IgG, for which I assumed parameters from the literature for its sensitivity and specificity, at 0.927 and 0.999, respectively [185], time to seropositivity, fixed at 13 days [186], and mean (exponential) time to sero-reversion at 242.5 [187]. Source: Perez-Guzman et al. [65].

		Erlang duration parameters		
Compartment	Definition	k_X	γ_X (per day)	Mean (days)
$T^i_{sero_{pre}}(t)$	Pre-seropositive for Abbott Architect (AA) assay	1	0.077	13
$T^i_{sero_{pos}}(t)$	True* seropositive for AA assay	1	0.0041	242.5
$T^i_{sero_{neg}}(t)$	True* seronegative (after infection) for AA assay	-	-	-

The quality and completeness of the seroprevalence data was reportedly good (see description in Table 4.1) [59]. I thus assumed it to be representative of the true population-level seroprevalence of SARS-CoV-2 IgG in Kabwe.

Using t to represent a week with observations within each of the two rounds of serosurvey data (see Table 4.1), I defined the number of seropositive and seronegative individuals for age band $b \in B = \{[0+),[15,20),[20,30),\ [30,40),[40,50),[50+)\}$ as the sum over all age groups within the age band, given

$$X^b_{sero_{pos}}(t) := \sum_{i \in b} T^i_{sero_{pos}}(t) \tag{4.1}$$

$$X^b_{sero_{neg}}(t) := \left(\sum_{i \in b} N^i\right) - X^b_{sero_{pos}}(t). \tag{4.2}$$

I then compared these to the observed number of weekly seropositive individuals from the data,

$Y^b_{sero_{pos}}(t)$, accounting for the size of the sample $Y^b_{sero_{test}}(t)$, and imperfect sensitivity ($p_{sero_{sens}}$) and specificity ($p_{sero_{spec}}$) of the assay from the literature [185] given

$$Y^b_{sero_{pos}}(t) \sim \text{Binom}\left(Y^b_{sero_{test}}(t), \omega^b_{sero_{pos}}(t)\right) \tag{4.3}$$

where

$$\omega^b_{sero_{pos}}(t) := \frac{p_{sero_{sens}} X^b_{sero_{pos}}(t) + \left(1 - p_{sero_{spec}}\right) X^b_{sero_{neg}}(t)}{X^b_{sero_{pos}}(t) + X^b_{sero_{neg}}(t)}. \tag{4.4}$$

The contribution to the likelihood of the seroprevalence data was then given as the independent product over all times and age bands, given

$$\mathcal{L}_{sero} = \prod_t \prod_{b \in B} P_{\text{Binom}}\left(Y^b_{sero_{pos}}(t) \,\middle|\, T^b_{sero_{pos}}(t), \omega^b_{sero_{pos}}(t)\right). \tag{4.5}$$

All-cause mortality

Triangulating a leading indicator representative of population-level infection prevalence, such as seroprevalence, and all-cause mortality has been demonstrated to allow accurate estimations of the burden of pandemic emergencies in the absence or reliable cause-specific mortality data [91,93]. The implementation of this analytic approach requires mortality data sets both from before and during a pandemic period, pertaining to the same population and with no systematic reporting differences (e.g. varying underreporting levels). This then allows to estimate the expected mortality during the pandemic period, in its hypothetical absence [91].

The all-cause mortality pandemic period data I had access to from Kabwe, however, suffered from an unknown level of under-ascertainment [52]. Furthermore, historic monthly aggregate counts only accounted for deaths occurring at health facilities, as opposed to both health facilities and the community, as was the case for the pandemic period data. Thus, I firstly estimated the potential size of historic underreporting, to allow a more direct comparison between the historic and pandemic period data streams.

I aggregated deaths from the daily timeseries in January and February 2020 (i.e. before the local start of the pandemic and the introduction of any NPIs that could have affected healthcare seeking / reporting behaviours in the population) and calculated the ratio between these and the routine monthly counts, assuming perfect reporting in the former (i.e. known to include deaths from the community). Routine monthly counts in early 2020, thus, recorded only 57% of the deaths included in the timeseries for the same period. I used this calculated ratio as a scaling factor to inflate monthly pre-pandemic deaths data, and conducted sensitivity analyses adjusting the assumed absolute historic deaths reporting by ±15% (see Appendix B).

Following a previously proposed approach [91,93], I then parametrised a linear model of baseline

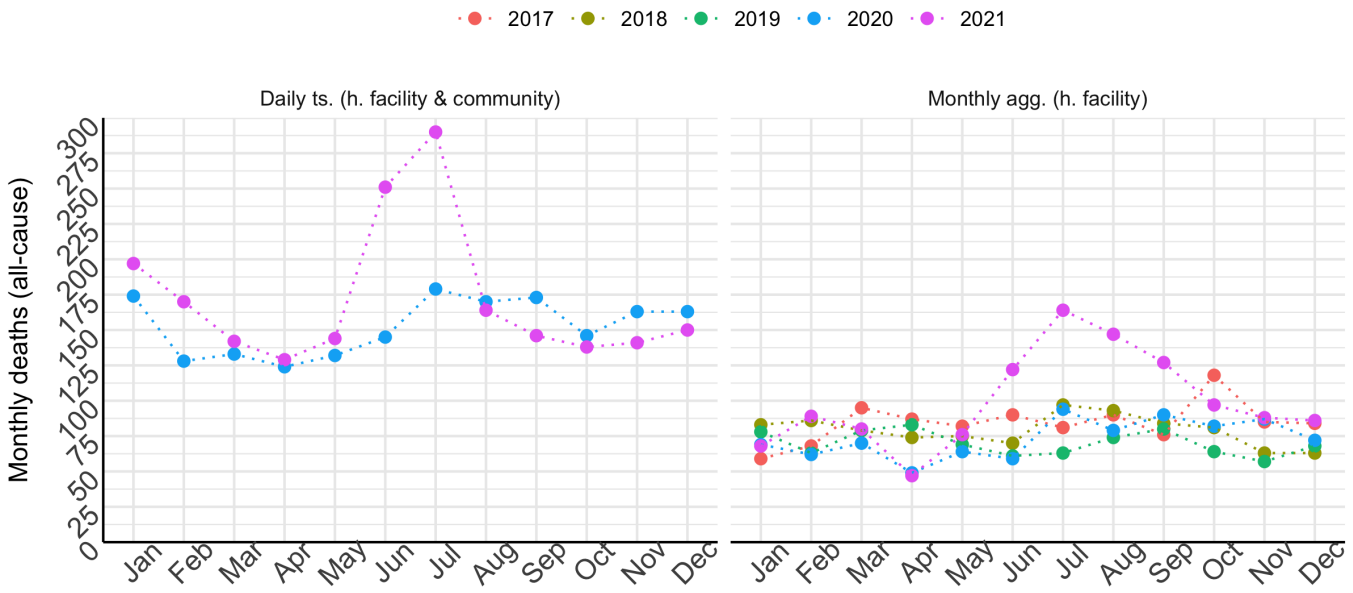


Figure 4.2: Monthly deaths (all-cause) recorded in Kabwe between 2017 and 2021 by type of data stream. Daily timeseries all-cause deaths data (left) was available for the pandemic period between 2020 and 2021 and included both deaths occurring in health facilities or the community in Kabwe. Routinely reported monthly aggregated all-cause deaths (right), on the other hand, were available between 2017 and 2021 and accounted only for deaths registered in health facilities (Table 4.1). Note raw time series data from 2020 and 2021 were in the form of daily timeseries, which I aggregate here for ease of visualisation.

deaths using the adjusted monthly counts from January 2017 to December 2019, with year and month as independent variables. From this model, I projected the expected monthly baseline deaths (i.e. in the absence of the pandemic) between March 2020 and September 2021, which I used as a fixed input parameter of the underlying, non-COVID-19 death rate $\rho_{base}(t)$ with a piece-wise constant form, assuming a uniform split across the days of the corresponding month.

Subsequently, using t to represent a day with observations, I considered the reported daily death counts, $Y_{D_{all}}$, as the observed realisation of a hidden Markov process, $X_{D_{all}}$, in the entire population by age (i.e. $i \in \{[0,5), \ldots, [70,75), [75+)\}$) with:

$$X_{D_{all}}(t) := \sum_i \left(X^i_{D_{covid}}(t)\right) + \rho_{base}(t) \tag{4.6}$$

where $X^i_{D_{covid}}$ represents the age specific deaths related to the COVID-19 pandemic and ρ_{base} baseline deaths that would have occurred in the absence of the pandemic. This related to the data via a reporting distribution:

$$Y_{D_{all}}(t) \sim \text{NegBinom}\left(X_{D_{all}}(t), \kappa_{D_{all}}\right), \tag{4.7}$$

where the overdispersion parameter $\kappa_{D_{all}}$ corresponded to the inverse of the fitted model parameter

i.e. $\kappa_{D_{all}} = 1/\alpha_D$.

The contribution to the likelihood of the daily deaths data was:

$$\mathcal{L}_{deaths} = \prod_t P_{\text{NegBinom}}\left(Y_{D_{all}}(t) \,|X_{D_{all}}(t), \kappa_{D_{all}}(t).\right) \tag{4.8}$$

The overall likelihood of the observed data given the model parameter was then:

$$\mathcal{L} = \mathcal{L}_{sero} \times \mathcal{L}_{deaths} \tag{4.9}$$

4.2.4 Model inference

Force of infection

Following the approach I proposed in Chapter 2, I modelled the force of infection as a function of transmission patterns given baseline contact rates by age $m_{i,i'}$ and a fitted piece-wise linear scaling parameter $\beta(t)$ to infer changes in contact rates over time.

I, thus, let $\Theta_{i,j}(t)$ be the number of infectious individuals in group i and vaccination stratum j, weighted by infectivity, given by:

$$\Theta_{i,j}(t) = \xi^{i,j}\left(\theta_{I_A} I_A^{i,j}(t) + I_P^{i,j}(t) + I_{C_1}^{i,j}(t)\right). \tag{4.10}$$

where $\xi^{i,j}$ represented the infectivity of an individual in strata i, j relative to an unvaccinated individual of the same age and, in counterfactual analysis presented in the next chapter, further accounting for vaccine efficacy against onward transmission, $e_{ins}^{i,j}$:

$$\xi^{i,j} = (1 - e_{ins}^{i,j}) \tag{4.11}$$

and θ_{I_A} is the infectivity of an asymptomatic infected individual relative to a symptomatic individual in the same age and vaccination strata.

The force of infection, $\lambda^{i,j}(t)$ on a susceptible individual in age group $i \in \{[0,5), \ldots, [75,80), [80+)\}$ and vaccination stratum $j = 0, 1, \ldots, 4$ was thus modelled as

$$\lambda^{i,j}(t) = \chi^{i,j} \sum_{i'} m_{i,i'}(t) \sum_{j'} \Theta_{i',j'}(t) \tag{4.12}$$

where $\chi^{i,j}$ is the relative susceptibility to infection of a susceptible individual in strata i, j relative to a non-vaccinated individual of the same age,

$$\chi^{i,j} = (1 - e_{inf}^{i,j}) \tag{4.13}$$

and the time-varying person-to-person transmission rate from age group i' to $i \in \{[0,5), \ldots, [70,75), 75+\}^2$, is given by

$$m_{i,i'}(t) = \beta(t) c_{i,i'}, \tag{4.14}$$

where $c_{i,i'}$ is the (symmetric) person-to-person contact rate between age group i and i'.

In the absence of contact surveys specific to Kabwe, Zambia, I assumed the same contact matrix as in an analysis of the COVID-19 pandemic in Lusaka, Zambia [121]. Of note, this matrix was derived pre-pandemic from Manicaland, Zimbabwe [188].

The time-varying transmission rate $\beta(t)$ encompassed both changes in transmission efficiency (e.g. emergence of variants, temperature and other seasonality factors) and temporal changes in the overall level of contacts in the population (due to changes in policy and behaviours). As mentioned, I assumed $\beta(t)$ to be a piecewise linear function:

$$\beta(t) = \begin{cases} \beta_i, & \text{if } t \leq t_i,\ i = 1 \\ \dfrac{t_i - t}{t_i - t_{i-1}}\beta_{i-1} + \dfrac{t - t_{i-1}}{t_i - t_{i-1}}\beta_i, & \text{if } t_{i-1} < t \leq t_i,\ i \in \{2, \ldots, 42\} \\ \beta_i & \text{if } t > t_i,\ i = 42 \end{cases} \tag{4.15}$$

Given unclear local non-pharmaceutical intervention (NPI) policies and known low adherence to social distancing measures during the period of interest [58, 66], I took an agnostic approach to parametrising $\beta(t)$, and fitted a total of 42 change points at regular fortnightly intervals, with the first change point on 2020-03-01.

Reproduction number

As proposed in section 2.2.3, I calculated the R_t and R_t^{eff} using the next generation matrix (NGM) method [88], assuming (for this calculation only) that individuals were not able to change vaccination class between the start and end of their infectiousness period.

For an individual in age group i and vaccine class j, the mean duration of infectiousness $\Delta_I^{i,j}$ accounted for infectivity (the infectivity of asymptomatic infected individuals is θ_{I_A} relative to symptomatic ones) as:

$$\Delta_I^{i,j} = \theta_{I_A}\left(1 - p_C^i\right)\mathbb{E}\left[\tau_{I_A}\right] + p_C^i\left(\mathbb{E}\left[\tau_{I_P}\right] + \mathbb{E}\left[\tau_{I_{C_1}}\right]\right). \tag{4.16}$$

The next generation matrix was thus:

$$NGM_{i,i'}(t) = m_{i,i'}(t)\xi^{i,j=0}\Delta_I^{i,j=0}N^{i'} \tag{4.17}$$

with N^i defined as the total population of age group i.

I then defined R_t as the dominant eigenvalue of the 16 by 16 matrix $NGM(t)$. The effective next generation matrix, on the other hand, was:

$$NGM^{eff}_{D(i,j),D(i',j')}(t) = m_{i,i'}(t)\chi^{i',j'}\xi^{gi,j}\Delta_I^{i,j}\left(S^{i',j'}(t) + R^{i',j'}(t)\right) \tag{4.18}$$

where $D : \{[0-5), [5-10), \ldots, [70-75), 75+\} \times \{0, 1, \ldots, 4\} \rightarrow \{1, 2, \ldots, 80\}$ is a one-to-one mapping. The R_t^{eff}, thus, corresponded to the dominant eigenvalue of the 80 by 80 matrix $NGM^{eff}(t)$.

Infection-induced immunity

I assumed waning infection-induced immunity γ_R with mean 3 years, as in my analysis of the pandemic in England. I further allowed for breakthrough infections determined by a fixed partial cross-protection parameter $\eta(t)$ with piecewise linear form with change-points and values

$$\eta(t) = \begin{cases} 1 & \text{for (t)} \leq \text{2020-12-01} \\ 0.95 & \text{if 2020-12-01} < \text{(t)} < \text{2021-06-30} \\ 0.85 & \text{for (t)} \geq \text{2021-06-30} \end{cases} \tag{4.19}$$

I also assumed values for this function as in my England analysis, whereas change-points corresponded to the 15^{th} day prior to the first known detection of the Beta [57] and Delta [189] VOC in Zambia, as a proxy of VOC emergence given the single-variant structure of the model.

Baseline and time-varying severity

As I proposed in Chapter 2, I modelled disease severity as a sequence of conditional transition probabilities. Namely, of clinical disease given infection p_C^i, and death given severe disease $p_{G_D}^{i,j}$.

I assumed the p_C^i values from section 2.2.4 by age group i

$$p_C^i = \begin{cases} 0.25 & \text{for i = [0,5)} \\ 0.26875 & \text{for i = [5,10)} \\ 0.325 & \text{for i = [10,15)} \\ 0.41875 & \text{for i = [15,20)} \\ 0.55 & \text{for i = [20,25), ...[70,75), [75+).} \end{cases} \tag{4.20}$$

and defined $IFR^{i,j}(t)$ as:

$$IFR^{i,j}(t) = p_C^i p_{G_D}^{i,j}(t) \tag{4.21}$$

where the probability of death given symptomatic disease was defined as:

$$p_{G_D}^{i,j}(t) = \psi_{G_D}^i p_{G_D}^{max} \left(1 - e_{death|SD}^{i,j}\right) h_D(t) \tag{4.22}$$

where $\psi_{G_D}^i$ was given by:

$$\psi_{G_D}^i = IFR_{lit}^i / p_C^i, \tag{4.23}$$

and both $p_{G_D}^{max}$ and the time-varying scaling function $h_D(t)$ were fitted parameters.

For the case of $p_{G_D}^{max}$, this corresponded to a multiplier of the resulting baseline spline $\psi_{G_D}^i$ from the literature (Table 4.5) [119], with respect to the age-group with the maximum value $i = 75+$ and $j = 0$.

For, $h_D(t)$, I assumed a piece-wise linear form, with change-points:

$$h_D(t) = \begin{cases} 1 & \text{on (and before) 15/08/20 - first wave,} \\ \mu_{D,1} & \text{from 01/12/20 to 15/03/21 - second wave,} \\ \mu_{D,2} & \text{on (and after) 01/05/21 - third wave.} \end{cases} \tag{4.24}$$

Lastly, $e_{death|SD}^{i,j}$ was the age- and vaccine-class specific vaccine efficacy against death given severe disease (see in Section 5.2.2).

Table 4.5: Age-specific probabilities of clinical disease given infection p_C^i and death given severe disease $p_{G_D}^{i,j}$. Parameters of the age-specific infection fatality ratio from the literature IFR_{lit}^i [119] were used to derive values of $\psi_{G_D}^i$, given fixed p_C^i values as per [1, 18, 32, 89] and eq. (4.22). Source: Perez-Guzman et al. 2024 [65].

Age group i	p_C^i	$\psi_{G_D}^i$	IFR_{lit}^i [119]
$[0, 5)$	0.25	0.000(	0.0000
$[5, 10)$	0.268750.00037		0.00010
$[10, 15)$	0.325	0.0003	0.00010
$[15, 20)$	0.418750.00048		0.00020
$[20, 25)$	0.55	0.0005	0.00030
$[25, 30)$	0.55	0.00073	0.00040
$[30, 35)$	0.55	0.0010	0.00060
$[35, 40)$	0.55	0.00182	0.00100
$[40, 45)$	0.55	0.0029	0.00160
$[45, 50)$	0.55	0.00436	0.00240
$[50, 55)$	0.55	0.0069	0.00380
$[55, 60)$	0.55	0.01073	0.00590
$[60, 65)$	0.55	0.0162	0.00920
$[65, 70)$	0.55	0.02600	0.01430
$[70, 75)$	0.55	0.0405	0.02230
$[75+)$	0.55	0.09163	0.05040

4.2.5 Running the model

Following my approach in Chapter 3, I pre-tuned the model's parameters by running an adaptive MCMC on the equivalent "expectation model", defined as the same model as the full stochastic model (Appendix B) but using the mean of the corresponding distribution wherever a random draw was required. I then used the pre-tuned parameter set from the "expectation" model with the highest posterior density and VCV matrix of the posterior distribution parameters as the initial values and proposal kernel, respectively, for subsequent pMCMC runs. The final pMCMC model fit was produced through approximately 3,000 CPU hours, in a single node of 32-core Xeon (dual 16-core 2.6 GHz). All necessary code and scripts to reproduce this analysis were made public at the time of publication, and can be found open-source in GitHub.

4.3 Results

4.3.1 Transmissibility

The model was able to recover the available data well (Figure 4.3A & B), which allowed me to infer the transmission and severity dynamics of the COVID-19 pandemic in Kabwe, Zambia between March 2020 and September 2021. Virus transmission was high across pandemic waves, which yielded an increasing profile of seroprevalence and infection-derived immunity in the local population (Figure 4.3A).

Local factors driving transmissibility varied across waves. Although the enforcement of and adherence to NPIs in Zambia was known to be low [58, 66], I found there was an initial gradual decrease in the R_t of the pandemic (fig. 4.3C). In fact this decline, which is most easily explained by decreases in social contacts given no initial immunity in the population in early 2020, was enough to drive R_t below 1 (i.e. the critical value for controlling transmission) between July and September 2020. As a result, infection attack rates in Kabwe peaked at a relatively low 2.6 (95%CrI 1.4-4.3) per-1000 people during the first wave, in early June 2020, compared to subsequent waves (Figure 4.3E).

The inferred reduction in contact rates was, however, reversed shortly after the peak of the first epidemic wave. Importantly, I found that levels or R_t in Kabwe increased again above 1 from early October 2020. This increase in transmission was, again, most likely driven by increasing contact rates, as the more-transmissible Beta VOC was not detected in Zambia until December that year [54, 57]. The combined effect of increasing contact rates and (likely) Beta VOC emergence drove much higher infection attack rates during this wave, which peaked at 5.7 (95%CrI 4.0-8.2) per-1000 people in early January 2021.

After this peak, the model estimated a significant 38% (95%CrI 32-43) of the population of Kabwe accrued infection-induced immunity as a result of the high attack rates seen during the second wave. Despite this, the R_t steadily rose from March 2021 onward, reaching levels above 1 from late April to mid-June 2020 (Figure 4.3C). It was at this point that the Delta VOC was detected in Zambia [54, 57]. Likely driven once more by a synergistic effect of higher intrinsic virus transmission, and high contact rates in the population, the infection attack rates of the pandemic were highest during the third wave, peaking at 7.2 (95%CrI 4.5-10.5) per-1000 people. Moreover, reinfections accounted for 14.9% (95%CrI 14.1–15.5) of daily infections, compared to only 0% (95%CrI 0-0) and 1.1% (95%CrI 0.8–1.2) during the first and second waves, respectively (Figure 4.3E).

4.3.2 Severity

To explain the high inferred SARS-CoV-2 infection attack rates in context with the observed all-cause mortality data from Kabwe between March 2020 and September 2021, my model estimated a total of 1,124 (95%CrI 774-1,323) excess deaths over the study period (Figure 4.3B & F). This translated into a cumulative excess mortality rate of 402 (95%CrI 277-473) per-100,000 people, a result that was robust to my assumptions of underascertainment in the input data.

The inferred trajectory of the IFR_t of SARS-CoV-2 provided important insights into the drivers of such high excess mortality rates of Kabwe. At the peak of transmission of the first wave, I estimated the IFR_t was of 0.34% (95%CrI 0.19-0.51) (Figure 4.3D). This was, on average, 2.8-times higher than the IFR of 0.12% that would have been expected in Kabwe (i.e. had similar healthcare provisions for COVID-19 cases as those of the settings from which my baseline age-specific values were derived [119] been in place locally), given the age distribution of the population.

Furthermore, I estimated the IFR_t of the pandemic in Kabwe increased from 0.39% (95%CrI 0.23-0.63) to 0.61% (95%CrI 0.37-0.87) at the peaks of transmission of the second and third pandemic waves, respectively (Figure 4.3D).

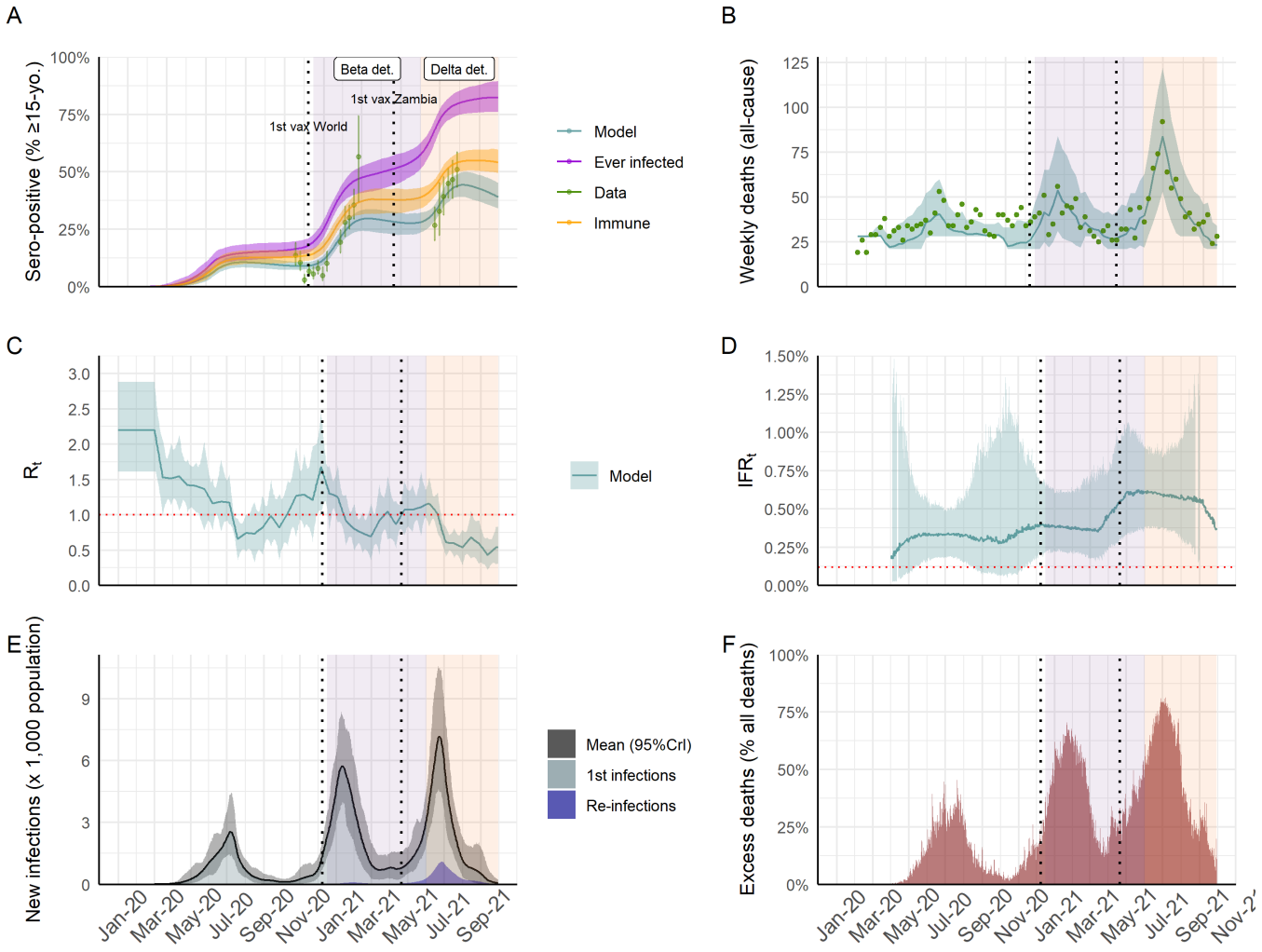


Figure 4.3: SARS-CoV-2 pandemic transmissibility and severity in Kabwe, Zambia from March 2020 to September 2021. Across all panels, shared rectangle areas illustrate the periods of time where the Beta (light purple) and Delta (light orange) variants were detected and known to dominate transmission in Zambia [54, 57], and black dashed vertical lines the date when the first COVID-19 vaccine was delivered in the world (2020-12-08) and in Zambia (2021-04-19) [28]. (a and b) Model fit (blue line mean and shaded area 95% CrI) to data (green dots; in [a], the dot represents the mean, and error bars the 95% binomial confidence interval) to seroprevalence and all-cause excess deaths, respectively. In (a), the purple line and shaded area show the inferred mean and 95% CrI for the cumulative proportion of the population ever infected, and the orange line and shaded area the mean and 95% CrI of population immune against infection (as a result of prior infection). (c) Inferred time-varying reproduction number, Rt (line mean and 95% CrI shaded area). The red horizontal dotted line at a y-axis value of 1.0 shows the critical threshold for pathogen transmission. (d) Inferred time-varying infection fatality ratio, IFR_t (line mean and 95% CrI shaded area). The red horizontal dotted line at a y-axis value of 0.12% shows the naïve estimate drawn by adjusting age-specific IFR estimates from Brazeau et al. [119], which came from a range of countries with known reliable hospital reporting systems, by the age distribution of the population in Kabwe. (e) Inferred new daily infection rate per 10 00 population (black line mean and 95% CrI shaded area), with breakdown of area under the curve into incident first infections (blue area) or reinfections (orange). (f) Inferred likely proportion of daily deaths that correspond to the excess mortality driven by the local COVID-19 pandemic. CrI, credible interval; det., detected; vax., vaccine; yo, year-old. Source: Perez-Guzman et al. 2024 [65].

4.4 Discussion

The study period for this analysis captured the first three SARS-CoV-2 pandemic waves in Kabwe, during which the majority of the population was infected at least once, amidst a complete lack of COVID-19 vaccines for the general population and a severely limited healthcare capacity. Adapting my modelling framework to infer the local drivers of transmission and severity from limited surveillance data was extremely testing. It required a careful conceptualisation of the observation process equations informed by the data, and its relation with the underlying hidden state variables. Despite these technical challenges, my analysis of the drivers of SARS-CoV-2 transmissibility and severity yielded several new insights of public health relevance for the response to outbreak emergencies in Kabwe and other similar LIC settings with a limited capacity to conduct surveillance and implement response interventions.

For LIC settings like Kabwe, relying on stringent and/or prolonged NPIs as a key pillar of a pandemic response is unsustainable, given their markedly higher economic burden on such settings compared to HICs [190]. At the start of the local pandemic, the transient effect of the NPIs implemented in late March 2020 (Section 1.2) [58, 66] effectively achieved a gradual reduction in contact rates between then and mid-July 2020. As a result, the R_t of SARS-CoV-2 decreased below 1, leading to a relatively mild first pandemic wave in Kabwe. The effect of local NPIs, however, was short-lived and the subsequent steady rise in contact rates synergised with a low level of infection-derived immunity at the after the first wave, and the known emergence of the more-transmissible Beta and Delta variants. Together, this myriad of intrinsic and effective drivers of transmissibility led to the very high incidence rates I inferred for subsequent waves. Given NPIs cannot be a sustainable outbreak response intervention in LICs, a rapid and sustained access to healthcare-based mitigation strategies, such as vaccines and other treatments for severe-case management, should be ensured.

Alas, I found strong indications that the severity of the pandemic in Kabwe was higher than expected and steadily increased across waves. Prior to my work, it had been demonstrated that differences in the IFR of COVID-19 across countries with comparable health care capacity were explained by their population age distribution [119]. By this token, it had been further demonstrated that adjusting the age-specific IFR estimates from the latter countries by the age distribution of the population in Lusaka explained the local excess mortality of the first wave [121]. In sharp contrast, I estimated the IFR_t during the first wave peaked at a value almost 3-times higher than that expected if adjusting those same age-specific IFR estimates by the age distribution of the population of Kabwe. Moreover, accounting for the how the characteristics of imperfect and waning immunity from prior infection likely changed over time (i.e. during the Wildtype, Beta and Delta waves), I estimated the IFR_t continued to increase across waves. My findings were thus consistent with early-pandemic

warning that indicated that the 'protective' effect of having young populations against the burden of the pandemic in LICs could be nullified in the presence of severe healthcare constraints [92].

Overall, my results underscore that the burden of SARS-CoV-2 in Kabwe was gravely accentuated by local limitations to implement the healthcare interventions that higher-income settings did. Between March 2020 and September 2021 the accrued excess mortality rate in Kabwe was of 402 (95%CrI 277-473) per-100,000 people, well exceeding that estimated by an influential (contemporary) study for the whole of Zambia and Western Europe, at 228.2 (95%CI 165.9-322.8) and 140.0 (95%CI 133.5-146.3) per 100,000-people, respectively [129]. Several factors could have conditioned such disproportionately high mortality burden during the pandemic in Kabwe, including a higher vulnerability to SARS-CoV-2 in the population of Kabwe compared to the national average of Zambia [41, 44, 45]. However, it remains true that, where possible to implement, an early and rapid implementation of improved healthcare protocols targeted to patients with severe COVID-19 was paramount to mitigate the severity of the pandemic [128, 150]. Similarly, COVID-19 vaccines offered lasting protection against entire clinical spectrum of COVID-19 disease, more so against the most severe of outcomes of hospitalisation and death [102, 104, 191]. It is thus not surprising that, accounting for the different age profiles of either setting, the excess mortality of Zambia at a national level was almost double that of Western Europe.

My model, for the most part, was able to recover the epidemiological data within the 95%CrI despite the structural and parametric assumptions I made. Moreover, its design was informed by best-practices identified in the studies where issues around limited data uncertainty were explored previously [91, 93, 121]. In particular, I

- Relied on available population-representative data streams
- Made structural assumptions informed by my validated framework for England (Chapter 2), including of pandemic seeding in early 2020 ahead of known reported cases and R_0 values of the Wildtype variant in-line with literature to inform early-pandemic dynamics
- Where appropriate, assumed fixed key parameters, including waning immunity, seroconversion and seroreversion rates, from the literature whilst varying them in sensitivity analysis (figures B5 to B8)
- Was agnostic when parametrising the fitted piece-wise linear contact rate scaling function $\beta(t)$, with fortnightly change points (as done in analyses across LMIC settings with biased case-based surveillance [121, 192]), more so given existing evidence of poor correlation of implementation/lifting of NPIs and mobility data across Zambia [58]

However, my analysis had several limitations as a result of the sparse quantity and uncertain

quality of the available surveillance data. For seroprevalence, I had access to age-specific data that covered the second and third local waves, with only one all-age aggregated seroprevalence estimate for the first-wave, before its peak. It is possible these data constraints limited the ability of my framework to correctly assess some key pandemic dynamics in Kabwe. The paucity of seroprevalence or other data stream reflecting community level transmission during the first wave introduced identifiability issues for the respective fitted values of the transmission scaling function $\beta(t)$. Despite high model precision to capture the one data point of aggregated seroprevalence, model-inferred dynamics during this wave were, as a result, only informed by the fit to mortality data (see respective limitations below) and assumptions on seroconversion-seroreversion rates (from the literature). Critically, this would have impacted $\beta(t)$ values and, in turn, the inferred declining trajectory of R_t during this wave. This could have partly further impacted the fit to deaths during the initial wave, where the model's mean trajectory fell short of the data, and had a knock-on impact on severity inference, whereby with a potentially higher number of infections during the first wave, the model's inferred IFR_t could have been lower. It should be, however, noted that whilst the model largely captured most of the deaths data within its 95%CrI (area of uncertainty), particularly during the peak in deaths of the first wave in July 2020, the respective uncertainty around the inferred IFR_t value at the same point was higher than the 0.12% naïve IFR value that would have been expected in Kabwe, as previously discussed.

Arguably, there could have been additional data sources, like mobility data and PCR cases to further inform my model's inference during the first wave. However, for the case of the latter, Kabwe district authorities reported a cumulative total of 3 cases per-1,000 population by March 2021 (i.e., 500 total PCR-positive cases detected by May 2021) [59]. Moreover, national policies in place during the first wave indicated the use of PCR testing for symptomatic individuals and/or contact tracing activities, and there were known shortages, delays and limitations in the availability of tests [59]. It is such biases and limitations in PCR case-based surveillance that lad to severe underestimations of the burden of COVID-19 where modelling analyses were informed by these data [93]. For the case of mobility data, comparing these with my model's inferred R_t during 2020 does suggest that reductions in mobility in the early stages of the first wave correlated to decreases in SARS-CoV-2 transmission (see Figure B9). It should be noted, however, that mobility data from Zambia was representative of Lusaka only [193]. Furthermore, the dedicated analysis by Loisante et. al. [58] demonstrated that: a) inter- and intra-province mobility across Zambia returned to pre-pandemic levels swiftly after official NPI policies on May 8, 2021, ahead of the peak of the first wave; b) mobility remained stable during subsequent waves, regardless of any NPI policies in place; and c) there were no statistically significant changes in mobility driven by school closures or holidays. I considered that analysing these data in my framework to inform transmission levels at face value would have introduced additional uncertainties at little gain.

Regarding deaths data I had access to, this suffered from an unknown level of underreporting, more so the historic deaths data I used to estimate the expected baseline mortality (in the absence of the pandemic). Further, whilst the daily timeseries data for the pandemic period was disaggregated by age, the historic timeseries was not. I was thus not able to assess the proportion of excess mortality in Kabwe that was not related to COVID-19 or fit my model to age-specific deaths data. The former could have implied that, for example, with an increase in non-COVID excess mortality during pandemic waves my model could have overestimated the true value of IFR_t. By the time I published my analysis in this chapter, others had sought to assess the implications of using different deaths data streams with on informing SARS-CoV-2 models in resource-constrained settings with limited surveillance capacity [91, 93, 121]. These analyses demonstrated that in settings with limited surveillance capacity and, thus, high risk of biased and underreported confirmed COVID-19 deaths, fitting transmission models to excess mortality yielded more robust estimations of the true burden of the pandemic [91, 93], and that adjusting global estimates of age-specific IFR by the age distribution of the population to be modelled yielded a good fit to excess mortality [121]. Here, I followed the validated methodology presented by these studies and fitted my model to available excess mortality data and informed my model's severity with the same global age-specific IFR estimates [119]. Future analyses will greatly benefit from reliable age- and cause-specific mortality.

Despite the above limitations, a key aim of my PhD was to assess how advanced analytics can and should be adapted to conduct robust inference, in light of such data gaps and uncertainties. That is, I sought to develop a parsimonious model that could have been employed for near real-time scientific advise in a resource and surveillance constrained like Kabwe setting during the pandemic. Focusing at a district level to me was important, as a key theme that emerged early on during my collaboration with the ZNPHI and Zambart was that countries like Zambia suffered from different sub-national pandemics in lieu of systematic surveillance and response capacity differences within the country. Indeed, the evidence of pandemic dynamics I found in Kabwe stands in contrast with that inferred for other areas of Zambia [46, 52, 121].

Surveillance data limitations were not unique to Kabwe, but rather prevalent across resource-constrained settings during the pandemic [71]. My work demonstrated that purposeful research collaborations between academic and public health institutions to develop outbreak advanced analytics are critical in LIC settings at the subnational level. A judicious approach to developing my modelling framework allowed me to conduct robust inference of the transmission and severity dynamics and burden of the pandemic, even with very sparse and uncertain data. Future analyses in Kabwe and similar settings will greatly benefit from exploring in depth the implications of incorporating other data streams, like mobility and (very limited) cases data. Importantly, my analysis further shows key surveillance gaps that could be prioritised for strengthening in resource-constrained set-

tings, including reliable all-cause death registration, cause- and age-specific mortality surveillance, and genomic surveillance.

4.5 Conclusion

During the course of my collaboration in Zambia, the WHO and CEPI have issued their calls to formulate pandemic preparedness strategies underpinned by learnings from the COVID-19 pandemic [134, 194]. These calls are leading to binding agreements for member states, despite of geography or income group. Equitable global pandemic preparedness will require robust analytics from all income settings and regions of the world. In highlighting key learnings from the COVID-19 pandemic in Kabwe, my work timely began to shed light into the implications that abiding such agreements will entail for LIC settings, as I will present in the following two chapters.

Chapter 5

The impact and cost-effectiveness of COVID-19 pandemic inequality in Kabwe, Zambia

Results around the impact of delayed access to vaccines and optimised healthcare on deaths in Kabwe, Zambia were included in my *International Journal of Infectious Diseases* publication on July 25, 2024 [65]. Given stochastic variability in the model, the epidemiological results summarised herein may show marginal variations from those in the latter publication, as I reran the model to perform the cost-effectiveness analyses of this and the following chapter, which remain unpublished.

5.1 Background

In Chapter 4, I demonstrated that the burden of the COVID-19 pandemic in Kabwe, Zambia was amongst the highest estimated for any low-income country (LIC) setting. Regardless of the many limitations in my analysis, I demonstrated that the local transmissibility and severity of SARS-CoV-2 was high, which conditions high infection attack and excess mortality rates locally. As I discussed, beyond likely constraints in local healthcare capacity, a higher burden of clinical vulnerability in the population of Kabwe [41,44,45] could have conditioned the disproportionately higher excess mortality seen in Kabwe.

However, during the same time period of my analysis in Kabwe, higher income settings successfully rolled-out COVID-19 vaccination programmes and implemented novel life-saving therapeutic protocols for patients with severe infections [1, 28, 94]. Global pandemic cooperation failures meant such an early adoption of effective pharmaceutical countermeasures was not possible across LIC settings [42]. In Kabwe, it was not until late 2022 that 70% of the population had received two doses of an approved vaccine [50, 68], a delay of well over a year compared to HICs and upper middle-income

countries (U-MICs) [28]. Moreover, it is also known that in-hospital mortality in Kabwe and Zambia at large increased during the second and third waves [50, 51, 67] as hospital pressures rose [50, 67].

A global modelling analysis of vaccine inequality estimated that over 50% of the COVID-19 deaths that occurred by October 2021 in LMICs could have been averted, if these countries had vaccinated at the same rate as HICs [183]. Critically, this analysis was informed by official COVID-19 case and death counts from the included LMICs which, as I have argued, yields vast underestimations of pandemic burden [91, 93]. Additionally, it did not account for the capacity of local health systems to actually implement the counterfactual vaccination strategies simulated. At the time I began the work I summarise in this chapter, there were only three additional studies of vaccine inequality, one data-driven multi-country modelling analysis [195] and two hypothetical meta-population models [196, 197], all suffering from the same aforementioned limitations.

Despite their limitations, 'umbrella' analyses such as these provided early insights into the impact of inequities in the global response to the COVID-19 pandemic, relevant for advocacy purposes. However, employing modelling analytics to inform public health interventions requires a more nuanced approach. Namely, it requires analytics carefully tailored to the context of available epidemiological surveillance (Chapter 4) and the capacities of local health systems. Failures to do so risk inflicting an overall negative impact on population health and healthcare affordability, given resources are diverted from current-time health system interventions [198, 199]. To address my aim of assessing the health and economic effects of a limited and delayed access to mitigation interventions in Kabwe, herein I present a counterfactual cost-effectiveness analysis of having implemented improved healthcare and COVID-19 vaccination as achieved by higher-income settings by September 2021. My analysis takes on the perspective of the health system, critically accounting for plausible values of its opportunity cost-based willingness to pay (ocWTP) threshold, the monetary threshold at which an intervention would result in one less health-benefit unit as a result of diverting funds from other current-day health system interventions [200].

5.2 Methods

I expanded the age- and vaccine-class explicit compartmental transmission model of SARS-CoV-2 I developed for Kabwe, Zambia to simulate a series of mutually-exclusive counterfactual scenarios, which I compared against the factual, model-fitted trajectories (henceforth *factual scenario*), as presented in Chapter 4. The scenarios I considered simulated the implementation of COVID-19 vaccination at the population level or of improved health care for severe cases, as achieved in higher income settings.

I have described the structure, parameters, data sources and equations of my epidemiological

model in Chapter 4 and Appendix B. In this methods section, I provide a technical description of the adaptations I made to my model, in order to explicitly track healthcare utilisation (Section 5.2.1), simulate the counterfactual implementation of interventions (Section 5.2.3), and the integration of an economic model to assess their cost-effectiveness (Section 5.2.4).

5.2.1 Modelling healthcare utilisation

Beyond intrinsic pathogen severity, factors such as time- and space-varying availability of healthcare resources and triaging practices can drive differences in hospital admission rates [1, 128, 150]. As I discussed in Chapter 3, identifying the effect of such factors on hospital admission rates, however, requires a wealth of healthcare data. To overcome the limitations of not having access to such data from Kabwe, I implemented a parallel flow to probabilistically simulate hospitalisations for individuals with a symptomatic infection leading to either recovery or death (Figure 5.1D).

Parallel flow equations

Based on the methods I developed to estimate the effective severity in Kabwe (Section 4.2.4), the model simulated the transition of infections through one of three possible clinical severity pathways (Figure 5.1A)

- Asymptomatic/mild infection leading to recovery ($I_A^{ij} \rightarrow R^{ij}$)
- Symptomatic infection leading to recovery ($I_{C_2}^{ij} \rightarrow R^{ij}$)
- Symptomatic infection leading to death ($I_{C_2}^{ij} \rightarrow \dots D^{ij}$)

I expanded this section of the model with a parallel flow (Figure 5.1D and Table 5.1) to simulate the probability of individuals with a symptomatic infection receiving hospital care given by

$$\frac{H_R^{i,j}}{dt} = I_{C_2}^{i,j} * (1 - p_{G_D}^{i,j}) * \gamma_{C_2} * p_{sev}^{i,j}(t) - H_R^{i,j} * \gamma_H \tag{5.1}$$

$$\frac{H_D^{i,j}}{dt} = I_{C_2}^{i,j} * p_{G_D}^{i,j} * \gamma_{C_2} * p_{D_{hosp}} - H_D^{i,j} * \gamma_{H_D}. \tag{5.2}$$

Table 5.1: Definitions of hospital parallel flow compartments shown in Figure 4.1D, with index i for age group ($i \in \{[0,5),[5,10),\ldots,[70-75),[75+)\}$). Durations (in days) have Erlang (k_X, γ_X) distributions, with mean k_X/γ_X.

		Erlang duration parameters			
Compartment	Definition	k_X	γ_X (per day)	Mean (days)	Reference
$H_R^{i,j}(t)$	Severe case in hospital leading to recovery	2	0.0299	6	[49]
$H_D^{i,j}(t)$	Severe case in hospital leading to death	2	0.8	2.5	[49]

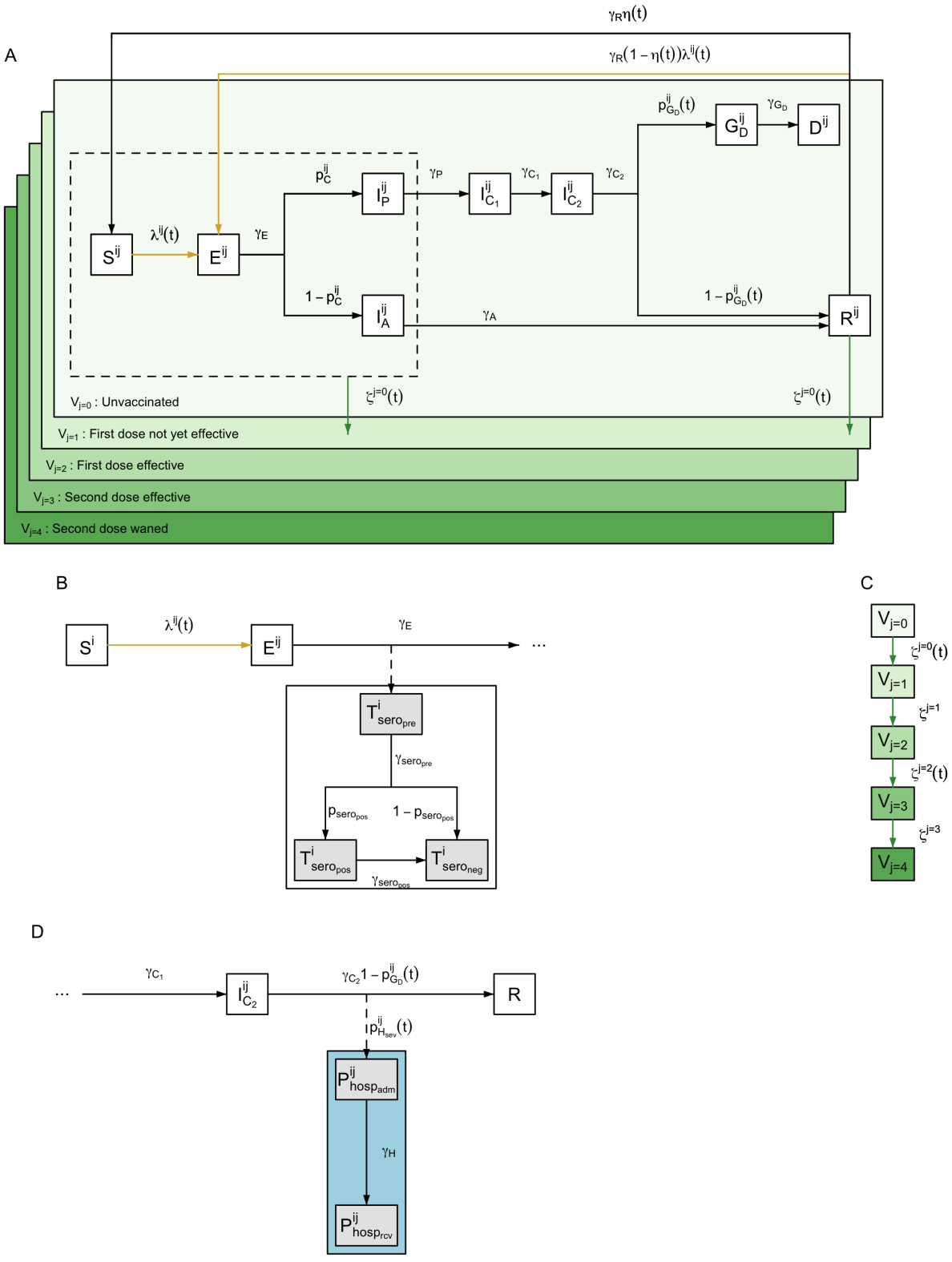


Figure 5.1: Adaptation of the model of SARS-CoV-2 transmission applied to surveillance data from Kabwe, Zambia between March 2020 and September 2021 to conduct cost-effectiveness evaluation. As in [65], A) presents the overview of model compartments and transition rates; B) the parallel flows developed simulate seroprevalence; and C) details vaccine progression classes: $j = 0$ for unvaccinated, $j = 1$ for 1st dose - no effect, $j = 2$ for 1st dose - full effect, $j = 3$ for 2nd dose - full effect, and $j = 4$ for 2nd dose - waned. i denotes age groups ($\{[0-4), [5-10), \ldots, [70-75), 75+\}$). Here, D) was added to simulate the progression from symptomatic infection to hospitalisation prior to recovery.

Parameters and assumptions

The implementation of the above parallel flow required the following parameters and simplifying assumptions.

Firstly, I assumed that a proportion of the individuals with a symptomatic infection en route to recovery ($I_{C_2}^{ij} \rightarrow R^{ij}$) would have needed hospital care to indeed recover, defined by

$$p_{sev}^{i,j}(t) = p_{sev}^{max}\left(1 - e_{hosp|SD}^{i,j}\right) h_D(t) \tag{5.3}$$

where p_{sev}^{max} was the multiplier of the baseline (i.e. at model onset and, thus, in the absence of immunity) age-specific probability of hospitalisation given symptomatic infection p_{sev}^{i}, with respect to the age-group with the maximum value $i = 75+$ and $j = 0$. For $h_D(t)$, this was the fitted piece-wise linear severity modifier (as defined in Section 4.2.4) with change points

$$h_D\left(t\right) = \begin{cases} 1 & \text{on (and before) 15/08/20 - first wave,} \\ \mu_{D,1} & \text{from 01/12/20 to 15/03/21 - second wave,} \\ \mu_{D,2} & \text{on (and after) 01/05/21 - third wave.} \end{cases} \tag{5.4}$$

and, $e_{hosp|SD}^{i,j}$ was the age- and vaccine-class specific vaccine efficacy against hospitalisation given symptomatic disease (Section 5.2.2).

Again, constrained by the lack of local healthcare data, I was not able to derive p_{sev}^{i} values specific to the pandemic in Kabwe. I thus assumed these to be the same as the ones I used for my analysis in England, at the start of the pandemic (Table 5.2).

It should be noted that I further assumed $p_{sev}^{i,j}(t)$ to vary over time given the fitted piece-wise linear function $h_D(t)$, whose values were inferred from my model's fit to local excess mortality data (Section 4.2.4). Therefore, despite the lack of locally-inferred values for p_{sev}^{i}, the overall effect of the parameter $p_{sev}^{i,j}(t)$ governing the entry into the hospital parallel flow still captured changes in severity specific to Kabwe (i.e. driven by the likely epidemiological factors underpinning local excess mortality, such as the local emergence of VOCs and surges in hospital demand).

Secondly, for symptomatic infections leading to death ($I_{C_2}^{i,j} \rightarrow \ldots D^{i,j}$), I assumed a fixed proportion $p_{D_{hosp}}$ (see Equation (5.1)) received in-hospital care before death, at 20% informed by a COVID-19 healthcare surveillance study from Lusaka between April 2020 and August 2021 [64].

Lastly, to model healthcare utilisations of individuals with a symptomatic infection leading to recovery, who would not have been captured by the above hospital parallel flow, I assumed a proportion $1 - p_{sev}^{i,j}(t)p_{outpx}$ received outpatient care at a primary care facility, where p_{outpx} was fixed at 66.5%, as informed by the same healthcare surveillance study from Lusaka [64]. This and the

assumed value of $p_{D_{hosp}}$ were varied in sensitivity analysis (Appendix B).

Table 5.2: Pathogen-driven age-specific probability of hospitalisation to lead to recovery given symptomatic infection $p_{sev}^{i,j=0}$, assumed as in [1].

Age group i	$p_{sev}^{i,j=0}$
$[0, 5)$	0.004
$[5, 10)$	0.001
$[10, 15)$	0.001
$[15, 20)$	0.002
$[20, 25)$	0.004
$[25, 30)$	0.006
$[30, 35)$	0.008
$[35, 40)$	0.008
$[40, 45)$	0.009
$[45, 50)$	0.010
$[50, 55)$	0.017
$[55, 60)$	0.023
$[60, 65)$	0.032
$[65, 70)$	0.041
$[70, 75)$	0.073
$[75+)$	0.126

Table 5.3: Fixed healthcare utilisation parameters for economic evaluation of COVID-19 counterfactual scenarios in Kabwe, Zambia between March 2020 and September 2021.

Parameter	Description	Value (range)	Source
p_{outpx_R}	Probability of receiving outpatient care leading to recovery (if not hospitalised)	0.665 (0.3 - 0.7)	[64]
p_{hosp_D}	Probability of receiving hospital care leading to death	0.2 (0.15 - 0.25)	[64]
γ_{H_D}	Mean length of hospital stay (days) leading to death (gamma distributed)	2.5	[49]

5.2.2 Vaccine efficacy

I assumed the values for VE against the Wildtype/Alpha variants of SARS-CoV-2 from my analysis in England (Table A3), across the entire study period. These values were derived from UK-based studies, so my implicit assumption was that there was no difference in VE between the UK and Kabwe, supported by results from a recently published systematic review and meta-analysis of available VE studies in Africa [201]. To assess the robustness of my results to my VE assumptions, I varied the absolute VE (all endpoints, Table A3) by ± 10% in sensitivity analysis (see Appendix B).

Table 5.4: Vaccine effectiveness parameters for the AstraZeneca and Pfizer vaccines by dose. "Infection" refers to vaccine effectiveness protecting an individual from being infected with SARS-CoV-2, whilst "infectiousness" refers to the vaccine effectiveness at preventing onward transmission by an infected individual. VE against death is modelled conditional on clinical disease (i.e. transition to I_P compartment), whilst all other endpoints are modelled conditional on infection. Source: Perez-Guzman et al. 2024 [65]

End point	Dose	AstraZeneca	Pfizer
Death	1	88%	89%
	2 (Full protection)	99%	99%
	2 (Waned protection)	83%	90%
Disease or infection	1	64%	79%
	2 (Full protection)	92%	99%
	2 (Waned protection)	29%	77%
Infectiousness	1	45%	45%
	2 (Full protection)	45%	45%
	2 (Waned protection)	40%	40%

Given an absence of age-specific data of vaccines administered in Kabwe, I assumed a fixed proportion of AZ and Pf vaccines by age, targeting those under the age of 18 with Pf and varying proportions of those older than 18 with AZ, in-line with global recommendations (Figure A4).

Lastly, whilst I assumed VE did not vary by age, the weighted VE did vary given the proportion of each vaccine (PF or AZ) assumed to be administered to each age group (Figure A4) and VE values by vaccine (Table A3).

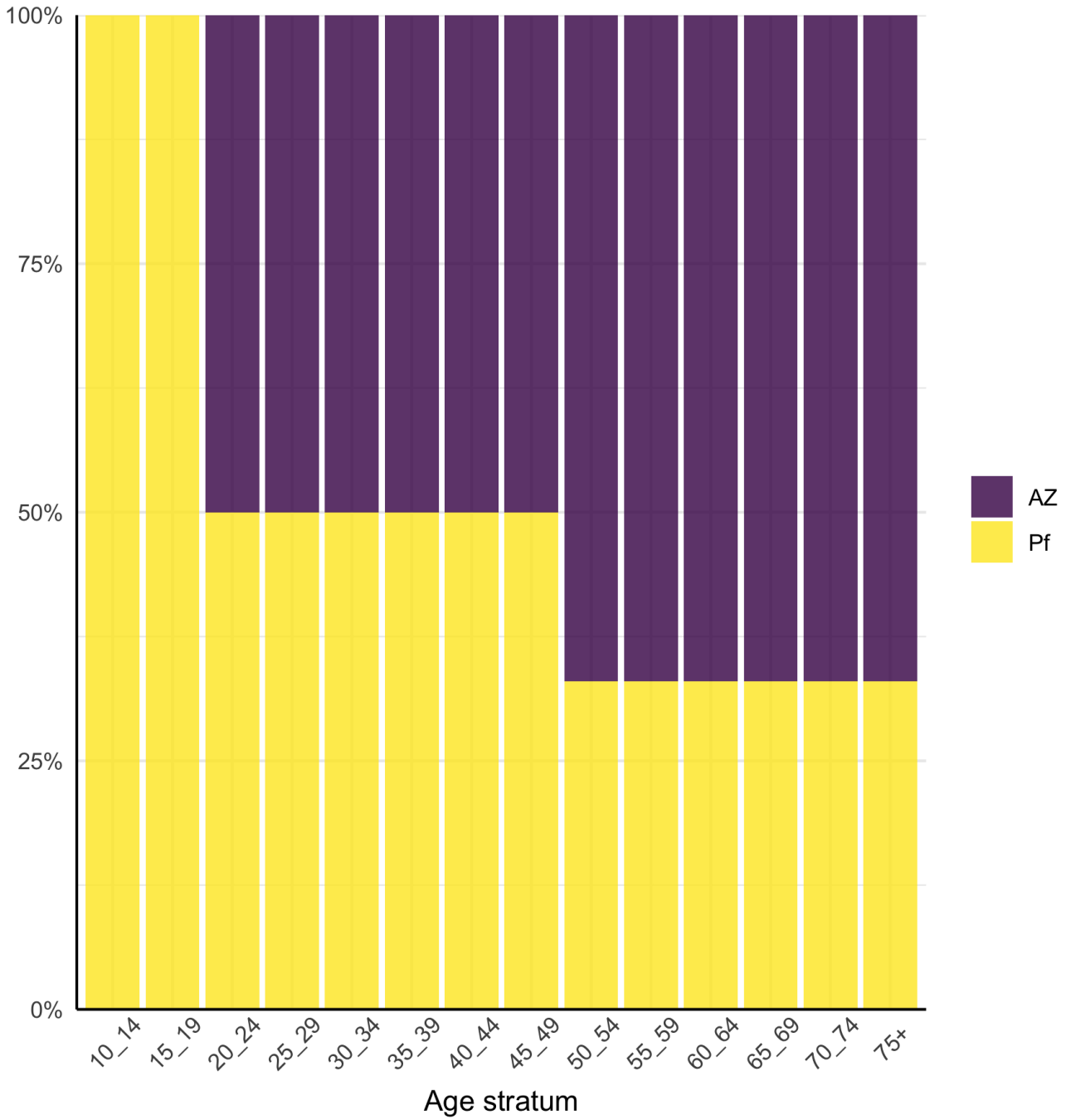


Figure 5.2: Assumed proportion of vaccine type (Pf = Pfizer, AZ = AstraZeneca) administered by age group. Source: Perez-Guzman et al. 2024 [65]

5.2.3 Counterfactual scenarios

Compared to the factual scenario (Chapter 4), I considered a total of 30 counterfactual intervention scenarios (Table 5.5) whereby I simulated the roll-out of a 2-dose primary course of COVID-19 vaccines (21 scenarios) and improvement of healthcare for the treatment of severe COVID-19 cases (9 scenarios).

Table 5.5: Counterfactual scenarios of vaccination and severe-case healthcare implementation in Kabwe, Zambia as achieved by higher income settings.

Intervention	Start date - rationale	Weekly coverage / IFR_t reduction
Vaccination	• April 20, 2020 - roll-out as per a hypothetical implementation of CEPI's 100-day mission [194] • December 14, 2020 - roll-out as factually achieved in HICs [28, 202] • April 19, 2021 - roll-out as factually achieved in L-MICs and LICs [28]	At each date, vaccination weekly coverage simulates the average doses per-100 population delivered in HICs, U-MICs, L-MICs and LICs [28] (Figure 5.3).
Improved healthcare	• June 15, 2020 - decrease in pandemic mortality due to early-pandemic improvements in healthcare (e.g. dexamethasone, supportive and respiratory care) [128, 150] • October 5, 2020 - before the onset of local second wave • April 5, 2021 - before the onset of local third wave	At each date, the inferred (from model fit) IFR_t decreases by 15% (minimal), 30% (moderate) and 45% (optimal) as in [92].

Vaccination roll-out as achieved by income setting

Local policies for vaccine eligibility followed WHO guidance at the time [203] with an old-to-young prioritisation of doses, a minimum age for vaccines eligibility of 12 years and an interval of 12 weeks between the first and second dose of a 2-dose primary immunisation course [40]. All vaccination scenarios I considered simulated these policies, assuming a 70% uptake amongst the eligible population. Given reports of high vaccination hesitancy in Zambia initially followed by a high uptake [43], I varied this assumption in sensitivity analysis, at 95%, 90% and 80% amongst 50+, 40-49-, and 12–39-year-olds, respectively, and a lower uptake at 50% across all age groups (Appendix B).

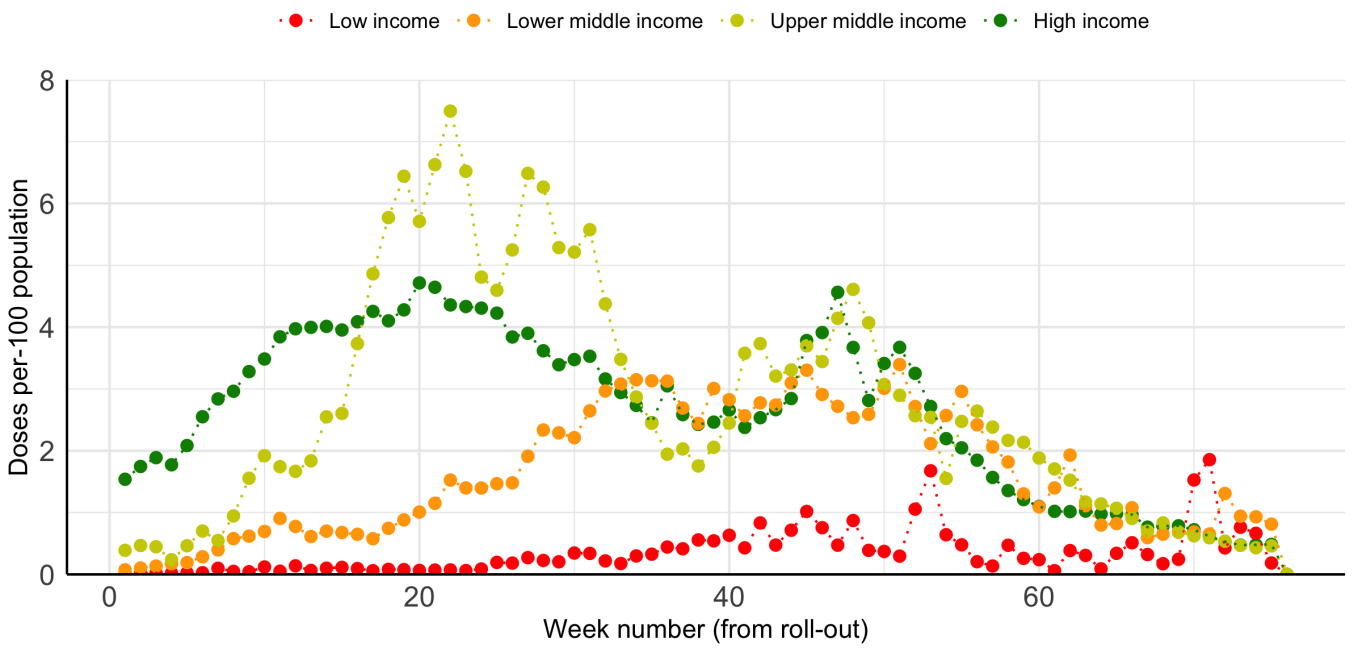


Figure 5.3: Modelled weekly vaccination schedules matching officially reported average weekly coverage (doses per-100 population) of LICs (red), L-MICs (orange), U-MICs (yellow) or LICs (green), [28] as presented in [65]. Schedules are plotted as weekly doses by week, from the moment of vaccination roll-out, which varied across scenarios (Table 5.5. Of note, during its first year of operation, the Zambian COVID-19 vaccination programme closely resembled the average trend of LICs (red), [28].

Healthcare improvement scenarios

I considered an increase in the Kabwe health system's capacity to effectively treat severe COVID-19 cases in hospital (Table 5.5). An early-pandemic analysis of the hypothetical absence of constraints in the availability of healthcare resources (i.e. health workers, beds, medicinal oxygen, other basic medical commodities, intensive care/mechanical ventilation capabilities) to manage severe COVID-19 cases would have resulted in a 45% relative reduction in the IFR, compared to a an absence thereof [92]. I considered this as a scenario of an *optimal* improvements in healthcare capacity and further simulated minimal and moderate improvements, assuming a reduction in IFR by 15% and 30%, respectively.

To implement these scenarios, I modified Equation (4.22) given

$$p^{i,j}_{G_D}(t) = \psi^i_{G_D} p^{max}_{G_D} \left(1 - e^{i,j}_{death|SD}\right) h_D(t)(1 - IFR_{reduction}) \tag{5.5}$$

where $IFR_{reduction} \in (0.15, 0.3, 0.45)$ was given by a minimal, moderate or optimal healthcare improvement, respectively.

For all scenarios, I assumed the factor $IFR_{reduction}$ applied linearly over a period of 3 weeks (piece-wise linear), thus simulating a gradual rather than sudden restructuring/upscaling of available healthcare [126]. Given the assumed severity reduction factor was applied on $h_D(t)$, I was able to

retain the model-inferred uncertainty in $IFR^{i,j}(t)$ across all simulated counterfactual trajectories. This was important, as the latter was not only the result of my model's stochastic behaviour. That is, I retained the uncertainty given underlying intrinsic and effective drivers of severity I could not explicitly model but were implicitly assumed in the data (e.g. time-varying healthcare constraints, basic severity properties of emerging VOCs).

5.2.4 Economic model

I projected the cost-effectiveness of the simulated counterfactual scenarios with respect to the factual scenario from March 2020 to September 2021, using the incremental costs-effectiveness ratio (ICERs) and net-monetary benefit (NMB). I quantified health outcomes both in terms of total accrued DALYs and DALYs averted, and obtained total costs and incremental costs of the interventions, given parameters from the literature (Table 5.6). I accounted for the cost of per-person-day of treatment in hospital or outpatient clinics (see Section 5.2.1) across scenarios and, for the case of vaccination counterfactuals, per-dose administered costs. Given stochastic behaviour in the model, all simulations were based a sample of 200 posterior pMCMC particles and presented results as the median and 95% credible interval (95%CrI).

Table 5.6: Cost parameters for economic evaluation of COVID-19 counterfactual scenarios in Kabwe, Zambia between March 2020 and September 2021. All cost parameters have been adjusted to 2020 US$ given year-on-year inflation rates for Zambia or an annual discounting rate of 3% [204], as appropriate.
*The methodology for the estimation of outpatient consultation and hospital day costs by WHO-CHOICE does not account for the cost of any consumables or overheads related to the treatment of patients at the respective level [205]. Here, I estimated the costs of consumables as the ratio of per-consultation or per-day hospitalisation, as appropriate, to all-other costs as reported in an ingredient's-based analysis of COVID-19 treatments costs from Kenya [206,207]. These encompassed pharmaceutical (e.g. drugs, IV fluids, oxygen) and non-pharmaceutical consumable items (e.g. personal protective equipment, devices wear-and-tear), diagnostics (i.e. laboratory, radiology and other) and facility overheads [206, 207].

Parameter	Value (range)	Source
Outpatient consultation	2.73 (0.56 - 2.73)	[205]
Per consultation consumables*	17.19 (3.54 - 55.86)	[206]
Hospital day	13.65 (5.22 - 29.10)	[205]
Per hospital day consumables*	113.30 (43.29 - 241.26)	[206, 207]
Per-vaccine dose administered cost	7.30	[208]

Cost data and assumptions

I gathered cost estimates from the literature, (Table 5.6), using a top-down approach for both the vaccination and improved healthcare interventions. All monetary input parameters and model outcomes are summarised in 2020 US Dollars, adjusting cost inputs from previous or later years by the average inflation or deflation rate, as appropriate, of Zambia of 3% per-year. I did not include the costs of NPIs, which I rather assumed to remain as inferred by the model (i.e. estimated changes in $\beta(t)$ contact rates) across all scenarios.

Tracking disability adjusted life-years

The model tracked years of life lost (YLL) to COVID-19 given the age-specific life-expectancy $life^{i}_{exp}$ at age of death in Zambia, using the WHO's 2019 life expectancy estimates. [109] To track years lived with disability (YLD), I assumed the mean disability weights for severe and mild respiratory infection from the 2021 Global Burden of Disease study (Table 5.7) [209], given no COVID-19 specific disability estimates exist.

$$\frac{YLL^{i}}{dt} = \sum_{j} D^{i,j}(t) life^{i}_{exp}, \tag{5.6}$$

$$\frac{YLD^{i}}{dt} = \sum_{j} \left(\Delta^{i,j}_{I_{C_2} \to H}(t) Disab_{sev} + \Delta^{i,j}_{I_{C_2} \to R}(t) Disab_{mild} \right), \tag{5.7}$$

$$\frac{DALY^{i}}{dt} = YLL^{i} + YLD^{i}, \tag{5.8}$$

where

$$\Delta^{i,j}_{I_{C_2} \to H}(t) = I^{i,j}_{C_2} * (1 - p^{i,j}_{G_D}) * \gamma_{C_2} * p^{i,j}_{sev}(t), \tag{5.9}$$

$$\Delta^{i,j}_{I_{C_2} \to R}(t) = I^{i,j}_{C_2} * (1 - p^{i,j}_{G_D}) * \gamma_{C_2} * (1 - p^{i,j}_{sev}(t)), \tag{5.10}$$

and $Disab_{sev}$ and $Disab_{mild}$ where the product of the disability weights and durations of severe and mild disease, respectively (Table 5.7).

Table 5.7: Disability weights for mild and severe COVID-19 disease in Kabwe, Zambia. All cases requiring hospitalisation to lead to recovery are assumed to be severe, whilst all other cases leading to recovery were mild (see Section 5.2.1).

Parameter	Value	Source
Disability weight mild	0.051	[209]
Disability duration mild	14	[206, 210]
Disability weight severe	0.133	[209]
Disability duration severe	21	[206, 210]

Table 5.8: Pathogen-drive age-specific probability of requiring hospitalisation to lead to recovery given symptomatic infection p^i_{sev}, as in [1], and life-expectancy at age of death in Zambia from [109].

Age group i	p^i_{sev}	$life^i_{exp}$
$[0, 5)$	0.004	63.40
$[5, 10)$	0.001	61.59
$[10, 15)$	0.001	56.99
$[15, 20)$	0.002	52.22
$[20, 25)$	0.004	47.62
$[25, 30)$	0.006	43.14
$[30, 35)$	0.008	38.73
$[35, 40)$	0.008	34.50
$[40, 45)$	0.009	30.46
$[45, 50)$	0.010	26.66
$[50, 55)$	0.017	23.03
$[55, 60)$	0.023	19.62
$[60, 65)$	0.032	16.39
$[65, 70)$	0.041	13.41
$[70, 75)$	0.073	10.67
$[75+)$	0.126	6.39

Willingness to pay threshold, incremental cost-effectiveness and net monetary benefit estimation

I calculated the ICERs from the total sum of DALYS and costs for each scenario compared to the baseline outcomes, given

$$ICERs = \frac{Costs_{scen} - Costs_{fact}}{DALYs_{fact} - DALYs_{scen}} \tag{5.11}$$

To determine the cost-effectiveness of each scenario, I then compared its ICERs to the ocWTP of Zambia. The long-standing 2001 WHO Commission on Macroeconomics and Health recommendation to set WTPs as 1x to 3x the gross-domestic product per-capita (GDPpc) per health gain (e.g. DALY averted or QALY gained) was recently revised [198]. Updated recommendations highlight that economic thresholds for the implementation of health interventions should contextualise potential health gains in light of an affordability roadmap to achieving universal health coverage [198]. To account for this, I calculated the ocWTP for Zambia WTP_{Zam} as

$$WTP_{Zam} = WTP_{UK} \left(\frac{GDP_{Zam}}{GDP_{UK}} \right)^{\epsilon} \tag{5.12}$$

The comparator value for the UK, at WTP_{UK} 0.39-times the GDPpc, accounts for purchasing power parity and value of statistical life [200,211]. The scaling factor ϵ was, thus, the assumed income elasticity for the purchasing power parity and value of statistical life in Zambia, which I assumed constant at between 0.7 and 1.5 that of the UK (Table 5.9), as per Woods et al [200].

Table 5.9: Parameters to estimate the range for the opportunity cost-based willingness to pay threshold in Kabwe, Zambia for the simulated COVID-19 interventions estimated as in Woods et all. [200]. GDPpc values are presented in 2020-USD. [204]

Parameter	Value	Range	Source
WTP_{UK}	0.39	-	[200, 211]
GDP_{Zam}	$958	-	[204]
GDP_{UK}	$40,217	-	[204]
ϵ	-	0.7 to 1.5	Assumed as in [200]

Lastly, I calculated the range in the NMB of each simulated scenario as

$$NMB_{min} = DALYs_{averted} * WTP_{min} - Costs_{incremental} \tag{5.13}$$

$$NMB_{max} = DALYs_{averted} * WTP_{max} - Costs_{incremental} \tag{5.14}$$

Given the range in ϵ (Table 5.9), the ocWTP used to estimate the NMB ranged from \$57.69 to \$1,146.64 2020-USD per DALY-averted at NMB_{min} and NMB_{max}, respectively, or between 6% and 120% the GDPpc of Zambia per DALY averted, respectively.

One-way sensitivity analysis

I performed a one-way sensitivity analysis varying parameters across their uncertainty range from source literature (Table 5.6). It should be noted that, given the analysis was done from the health system's perspective, I assumed that individuals in the model with an asymptomatic infection or those with a symptomatic infection that did not enter the aforedescribed healthcare parallel flow (Section 5.2.1) did not contribute to the estimated costs of the main or sensitivity analysis scenarios.

5.3 Results

5.3.1 Impact of limited and delayed pandemic interventions

The roll-out of COVID-19 vaccines on December 2020 or earlier at an average capacity of L-MICs or higher would have averted more excess deaths in Kabwe, Zambia than even the most optimistic of improvements in severe-case healthcare (Figure 5.4).

Vaccination roll-out in December 2020 in Kabwe, as some HICs did [28], could have averted 48% (95%CrI 45-53), 60% (95%CrI 56-64) and 66% (95%CrI 62-70) of the excess deaths accrued between March 2020 and September 2021 given the average capacity (Figure 5.3) of L-MIC, U-MIC and HICs, respectively, compared to only 23% (95%CrI 20-28) with the average capacity of LICs. Whilst my model inferred a higher average excess deaths averted with higher weekly coverage across all roll-out dates, there was a greater overlap in the estimated uncertainty across scenarios at earlier roll-out dates. Notably, this suggests that a hypothetical roll-out in April 2020, in line with CEPI's 100-day mission [194], with a weekly coverage of U-MICs or higher could have almost fully mitigated the COVID-19 pandemic in Kabwe, Zambia (Figure 5.4A).

On the other hand, a decrease in the IFR_t by 45%, 30% and 15% from as early as June 2020 as a result of healthcare improvements during the first wave, as implemented in higher income settings [128], could have averted 46% (95%CrI 41-54), 31% (95%CrI 27-44) or 17% (95%CrI 12-31) of excess deaths, respectively. Of note, the model inferred that a moderate improvement in healthcare (resulting in a 30% decrease in IFR_t) during the first wave would have sufficed to drive the mean IFR_t down to 0.12%, in-line with what would have been expected in the absence of healthcare constraints in Kabwe (Figure B17). Lastly, given the severity of the pandemic in Kabwe sequentially increased across pandemic waves, even a delay in implementing improved healthcare until the third

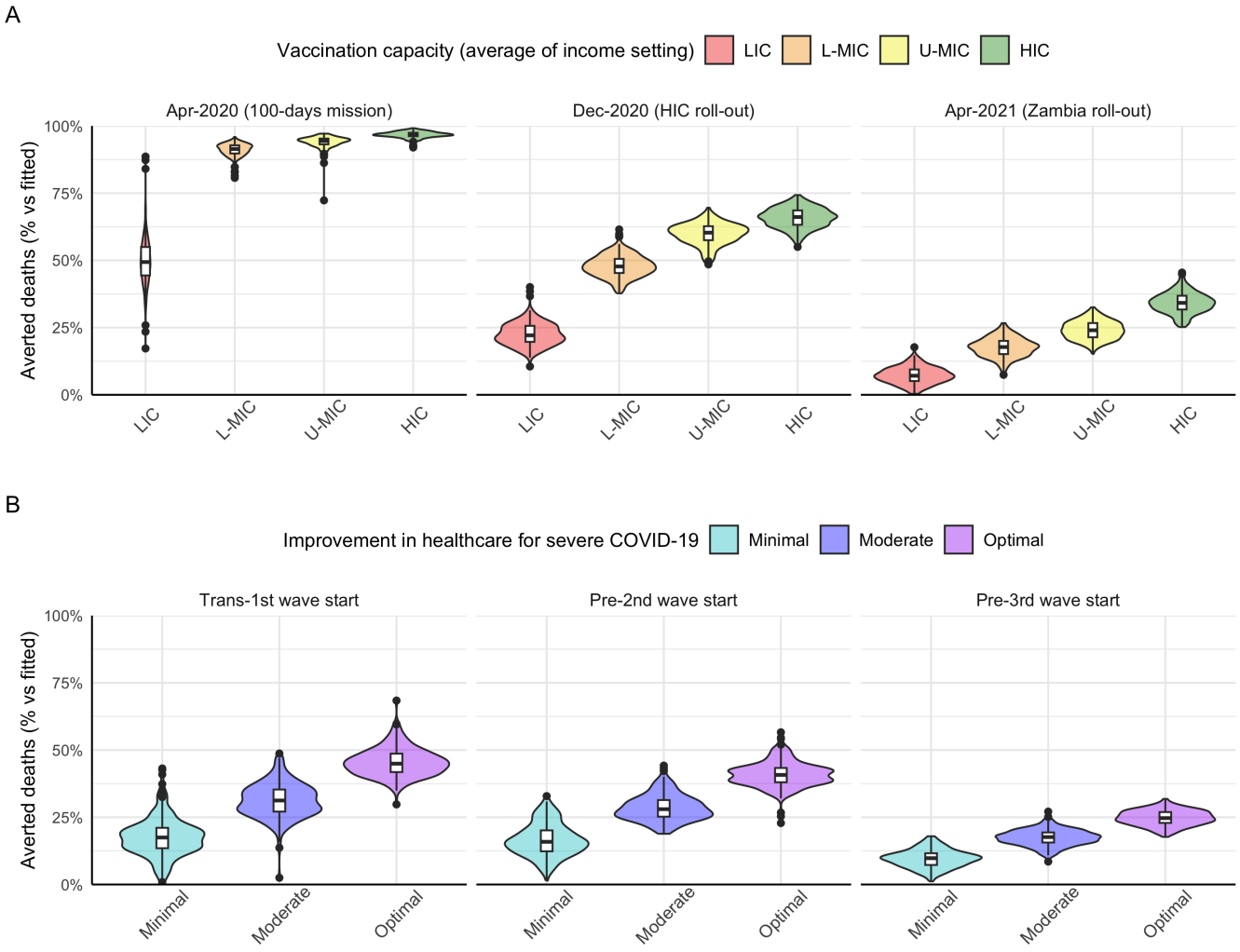


Figure 5.4: Estimated excess deaths averted in counterfactual simulated scenarios (Table 5.5), compared with the model fit to data, as reported in [65]. Timing of scenario intervention implementation (columns), weekly coverage (colour legend in A) and healthcare improvement level (colour legend in B) as per Table 5.5.
Abbreviations: HIC, high-income countries; LIC, low-income countries; L-MIC, lower-middle-income countries; U-MIC, upper-middle-income countries.

wave would have still averted a significant burden, at 10% (95%CrI 9-13), 17% (95%CrI 15-20) or 25% (95%CrI 24-26) of excess deaths averted, respectively (Figure 5.4B).

5.3.2 Cost-effectiveness of counterfactual interventions

Under the factual scenario (Chapter 4), I estimated that between March 2020 and September 2021 the total burden of DALYs amounted to 22,536 (95%CrI 19,679 to 25,394), at a total cost to the health system of $1.79m (95%CrI 1.61-1.91) 2020-USD in Kabwe, Zambia as a result of the COVID-19 pandemic. Given the size of the population of Kabwe, this translated to a burden of 8,055 (95%CrI 7,034-9,076) DALYs and $0.64m (95%CrI 0.57-0.68) 2020-USD per-100,000 population. Assuming a high ocWTP, at $1,146.64 2020-USD or 120% the GDPpc of Zambia per-DALY averted, all simulated

counterfactual interventions would have been cost-effective for the local health system (Figure 5.5A & D), but not assuming the lowest plausible ocWTP threshold, at $57.69 2020-USD or 6% of the GDPpc of Zambia.

Incremental cost-effectiveness ratio of interventions

During the study period, the second and third local waves were responsible for the majority of infections and deaths in Kabwe (Chapter 4), thus most DALYs were accrued during these waves across all age groups (Figures B27 and B28). An earlier roll-out of vaccination at a higher weekly rate would have led to an earlier achievement of the 70% target coverage with 2 doses (Figure B18), at which point the costs of the intervention would have peaked (Figure 5.5B). Thus, accounting for the efficacy and duration of vaccine-induced immunity, more DALYs would have been averted with an earlier roll-out as a result of the downstream prevention of infections, cases in need of healthcare and deaths in subsequent waves (Figure 5.5C).

Such dynamic relation would have, in turn, conditioned a higher incremental cost and lower health benefit of any vaccination strategy with delays to roll-out. For example, the roll-out of vaccination in April 2020 at HICs capacity would have been highly cost-effective, even if assuming the lowest ocWTP threshold, but not with a later roll-out. That is, a vaccination intervention with the capacity of HICs would have averted a total of 21,729 (95%CrI 18,725-24,623), 14,256 (95%CrI 12,064-17,058) and 6,981 (95%CrI 5,113-8,947) DALYs if rolled-out in April 2020, December 2020 and April 2021, respectively, at an ICER of $10 (95%CrI 4-21), $66 (95%CrI 52-89) and $159 (95%CrI 111-231) 2020-USD per DALY averted, respectively.

Furthermore, at the earliest roll-out of April 2020, the model showed no significant difference in the health or economic outcomes of the vaccination interventions with either an HIC, U-MICs or even L-MICs capacity (Figure 5.5A). Conversely, with a roll-out a year later in April 2021, rolling-out vaccination would have had an ICER of $159 (95%CrI 111-231), $211 (95%CrI 138-409) and $61 (95%CrI 21-263) 2020-USD per-DALY averted at the capacity of HICs, U-MICs and L-MICs, respectively, thus making higher capacity roll-out unlikely cost-effective at a low ocWTP value.

Only vaccination at the capacity of LICs would have remained below the health system's lowest plausible ocWTP threshold across all dates (Table B4), at an ICER of $13 (95%CrI -18 to 99), $-10 (95%CrI -36 to 61) and $-16 (95%CrI -190 to 250) with a roll-out in April 2020, December 2020 and April 2021, respectively. Beyond the wide uncertainty in estimated ICERs, it should be further noted that the evaluation horizon was curtailed to September 2021. Such vaccination strategies would, therefore, had a limited efficacy within this horizon, at 8,556 (95%CrI 3,230 to 14,380), 3,391 (95%CrI 1,006 to 6,046) and 898 (95%CrI -778 to 2,839) DALYs averted, respectively.

All counterfactual healthcare improvement interventions simulated, on the other hand, were

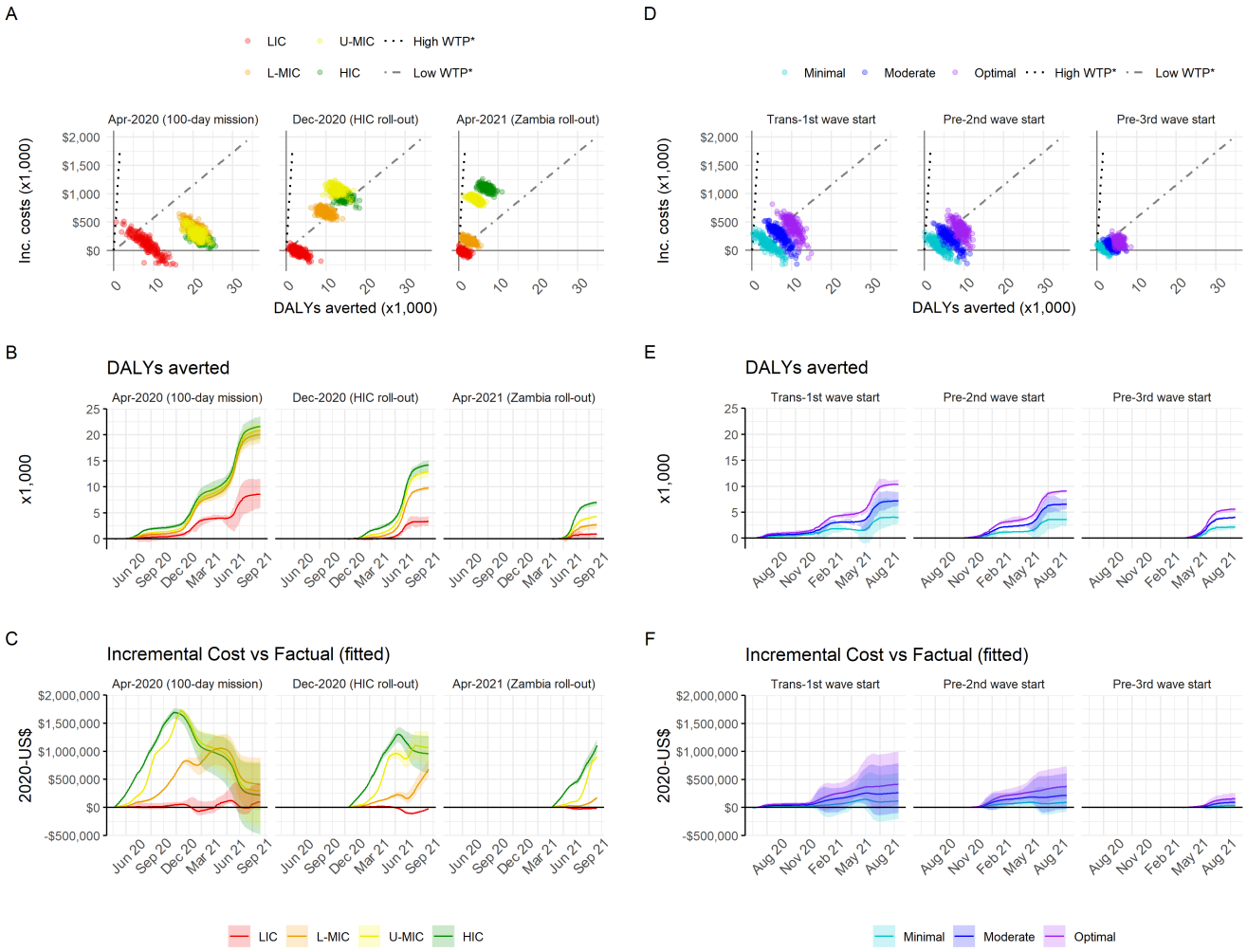


Figure 5.5: Estimated cost-effectiveness of counterfactual COVID-19 vaccination and healthcare improvement scenarios in Kabwe, Zambia (Table 5.5), compared with factual (model fitted) outcomes (see Chapter 4). Timing of scenario intervention implementation (columns), vaccination rates (colour legend in A-C) and healthcare improvement levels (colour legend in D-F) as per Table 5.5.
*Lowest (dash-dotted line) and highest (dotted line) opportunity-cost based WTP estimated at $57.69 ant $1,146.64 2020-USD, respectively, given assumed range in income elasticity for the purchasing power parity and value of statistical life of Zambia (Table 5.9), as per Woods et al. [200].
Abbreviations: DALYs, disability-adjusted life years; HIC, high-income countries; Inc., incremental; LIC, low-income countries; L-MIC, lower-middle-income countries; U-MIC, upper-middle-income countries; WTP, opportunity cost-based willingness to pay threshold.

cost-effective even when assuming the lowest ocWTP threshold (Figure 5.5D), again with considerable model uncertainty. Given such interventions would have not have had an effect on immunity (Figure B20), there were no time-varying dynamics between the timing of implementation and incremental costs (Figure 5.5F). Considering, again, that the 2nd and 3rd waves caused the most DALYs in the population (Figure B28), even with delays to their implementation a significant burden of DALYS would have still been averted.

Thus, if implemented during the first wave, as higher income settings did, a total of 3,771 (95%CrI 490-8,387), 7,046 (95%CrI 4,194-11,068) and 10,275 (95%CrI 7,547-13,767) DALYs would have been averted with minimal, moderate and optimal healthcare improvements, respectively, at

an ICER of $28 (95%CrI -32 to 353), $41 (95%CrI -8 to 111) and $41 (95%CrI 5 to 80) 2022-USD per-DALY averted (Table B3). Conversely, if implemented by the 3rd wave, a total of 2,193 (95%CrI 412-3,833), 3,899 (95%CrI 2,429-5,791) and 5,492 (95%CrI 4,288-7,144) would have still been averted, respectively, at an ICER of $17 (95%CrI -23 to 205), $24 (95%CrI -4 to 70) and $29 (95%CrI 6 to 55), respectively (Table B4).

Net monetary benefit of interventions

When assessing the NMB by fiscal quarters, the model showed that earlier implementations and at ideal capacity (i.e. HIC capacity for vaccination and optimal improvement of healthcare) would have maximised the monetary value of the interventions by the end of the evaluation horizon (Figure 5.6, Q3-2021). Importantly, however, assuming the lowest plausible ocWTP threshold yielded high uncertainty in the model estimates, which meant that interventions with lesser than ideal capacities might not have guaranteed a positive NMB.

For the case of vaccination, assuming the highest ocWTP threshold, a roll-out in April 2020 at HIC capacity would have resulted in a NMB of $24.8m (95%CrI 21.1 to 28.2) 2020-USD, compared to $22.8m (95%CrI 19.5 to 26.1) with an L-MIC capacity. Conversely, assuming the lowest ocWTP threshold, the model estimated such vaccination interventions would have resulted in a NMB of $1.1m (95%CrI 0.7 to 1.4) and 0.8m (95%CrI 0.4 to 1.1). Furthermore, with a roll-out in December 2020 at HIC and L-MIC capacity, the expected NMB would have been of $15.5m (95%CrI 12.9 to 18.6) and $10.5m (95%CrI 8.0 to 13.5), respectively, if assuming the highest ocWTP threshold, alas of $-0.7m (95%CrI -0.9 to -0.5) and $-0.1m (95%CrI 0.0 to 0.2) 2020-USD, if assuming the lowest ocWTP threshold, respectively (Figure 5.6A).

Similarly, a healthcare improvement strategy implemented in the first wave would have resulted in a NMB of $11.3m (95%CrI 8.1 to 15.4) if optimal and $4.2m (95%CrI 0.3 to 9.8) if minimal, assuming the highest ocWTP threshold. Conversely, assuming the lowest ocWTP threshold, the NMB of the same healthcare improvement interventions would have been of $0.2m (95%CrI -0.2 to 0.7) and $0.1m (95%CrI -0.3 to 0.6), respectively (Figure 5.6B).

The above results were sensitive to variations in the input parameters of vaccination, outpatient and hospital costs (Figures B29 to B32). For the case of vaccination scenarios, the model indicated a higher sensitivity in the estimated ICERs to variations in vaccination costs, particularly with earlier roll-out dates. On the other hand, whilst the ICERs of healthcare improvement scenarios were relatively robust to variations in costs of outpatient and hospital care, their NMB estimates were highly sensitive to such changes if assuming the lowest ocWTP threshold, but not at its highest value.

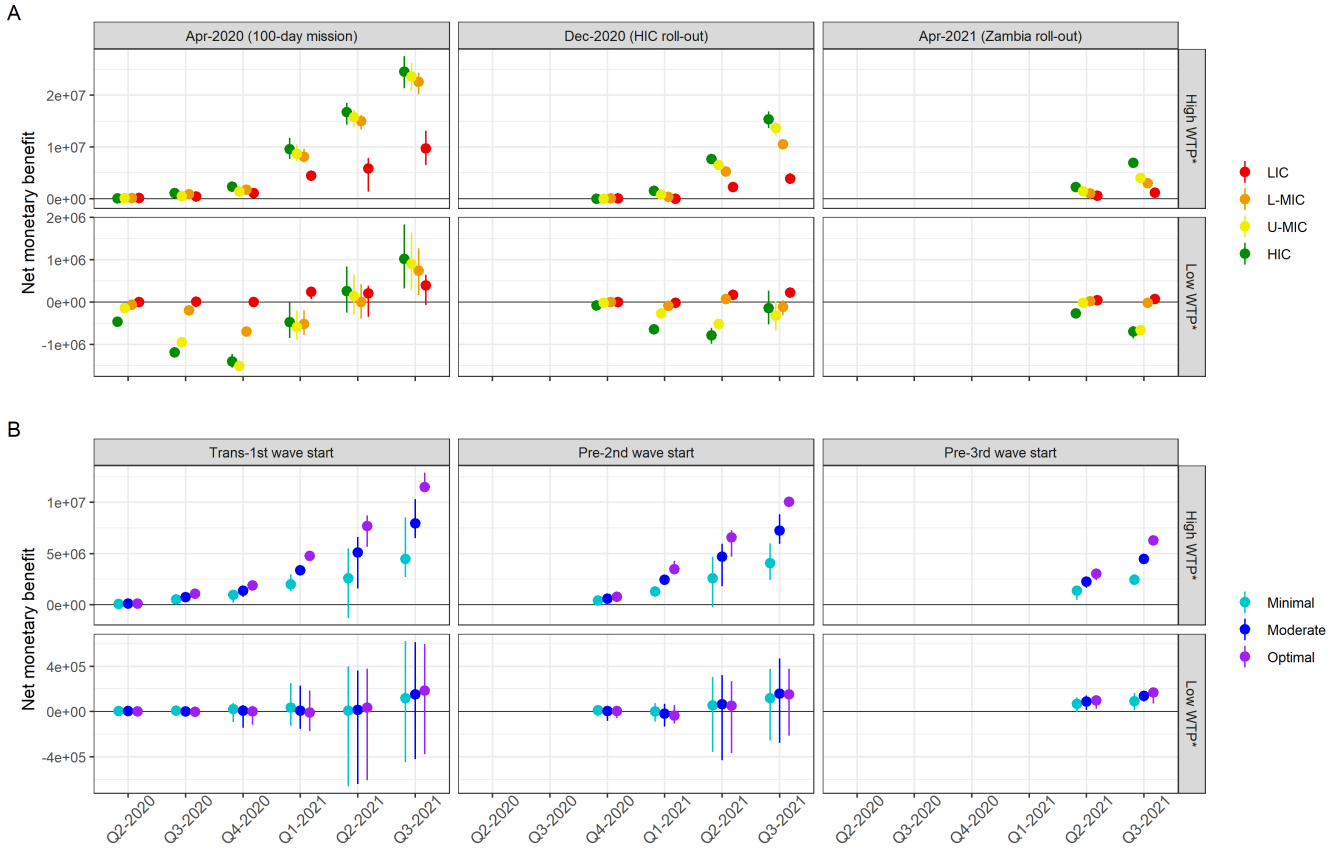


Figure 5.6: Estimated excess deaths averted in counterfactual simulated scenarios (Table 5.5), compared with the model fit to data, as reported in [65].
Abbreviations: HIC, high-income countries; LIC, low-income countries; L-MIC, lower-middle-income countries; U-MIC, upper-middle-income countries.

5.4 Discussion

In the previous chapter, I described how the health burden of the COVID-19 pandemic in Kabwe, Zambia was among the highest recorded for the African continent. My counterfactual health economic analysis herein further showed that an early roll-out of COVID-19 vaccines and improvement in healthcare capacity could have averted a large proportion of this burden. However, for all counterfactual interventions I considered, delays to their implementation would have been associated with a decrease in their cost-effectiveness, especially if assuming the lowest plausible ocWTP value of the health system in Kabwe. My findings not only provided several novel insights into the public health implications of pandemic inequality, but also have implications for informing response interventions contextualised in the capacities of LIC settings similar to Kabwe.

A roll-out of vaccines at the same time and with the same capacity as HICs would have averted two-thirds of the excess deaths of the pandemic. This findings was in keeping with previous evidence from two multi-country analyses that estimated that over 50% of COVID-19 deaths could have been averted across LMICs with a vaccination roll-out in-line with that of HICs [90, 183]. However, I further found that the cost-effectiveness of such intervention would have been subject to the ocWTP

threshold of the health system in Kabwe.

Given the many constraints of health systems, the introduction of new health interventions should be considered in light of its opportunity costs with respect to other priority services [198]. With this in mind and following a validated methodology [200], I estimated an ocWTP threshold at between 6% and 120% the GDPpc of Zambia. This range was in line with results from a recent multi-country WTP estimation exercise which found (also given the above constraint considerations) that ocWTPs for LICs lie within a plausible range of between 5% and 102% of their GDPpc [199]. In this context, what my analysis highlighted was that the cost of rolling-out vaccination during the pandemic under a more equitable global sharing of vaccines could have exceeded the lowest ocWTP threshold of the health system in Kabwe. Other health economic analyses of COVID-19 vaccination strategies in LMICs have also found uncertainty of the cost-effectiveness of ambitious roll-out strategies [206, 212].

Importantly, I found that progressively longer delays to the implementation of any vaccination intervention would have not only deemed them less efficacious at averting DALYs, but also more costly. When looking at more hypothetical scenarios, whereby vaccines had been rolled out in April 2020 (CEPI's 100-day mission [194]), even the most ambitious roll-out of vaccines at the capacity of HICs would have been well below the lowest ocWTP threshold. Moreover, my model showed there would have been no significant differences in terms of DALYs averted or incremental costs with a roll-out at L-MIC, U-MIC or HIC capacity. Conversely, the incremental costs of a population-level vaccination programme roll-out in April 2021 at either of said capacities would have likely deemed it not cost-effective, well exceeding the lowest ocWTP threshold. Similarly, an earlier roll-out at higher capacity would have maximised the likelihood of attaining a high NMB of any vaccination intervention. That is, a positive NMB for any vaccination strategy was not materialised immediately given upfront costs of their roll-out. As time accrued, such costs were substantially offset by their lasting effect on population-level immunity and, thus, to continue to avert DALYs over prolonged periods of time. Assuming a low ocWTP threshold, the model suggested it would have taken over a year for vaccination strategies rolled-out in April 2020 to attain a positive NMB.

Additionally, I found that improvements in available healthcare in Kabwe would have likely levelled the severity of the pandemic to what would have been expected with similar healthcare provisions as in higher income settings. In Chapter 4 I underscored that the IFR_t of the first pandemic wave was almost 3-times higher than that estimated by adjusting global age-specific IFR estimates [119] by the age distribution of the population of Kabwe. My model showed that a moderate improvement in healthcare, driving the severity down by 30%, would have yielded an IFR_t of around 0.12% during the first wave, in line with my comparison with said global estimates. Moreover, I found that this and any other healthcare improvement strategy at any point of the local pandemic would

have likely be cost-effective, although with high model uncertainty. Indeed, the model indicated no significant differences in the burden of DALYS averted by an optimal healthcare improvement intervention by the 3rd local wave and a roll-out of vaccination at HIC capacity in April 2021 roll-out, albeit the ICER of the latter would have been over 5-times that of the healthcare intervention. Whilst this might suggest implementing healthcare improvements in Kabwe at such late stages of the pandemic would have been preferable to rolling-out vaccination, it should be noted again that the model did not account for health and economic benefits beyond the evaluation horizon.

My analysis herein had several limitations. Foremost, the evaluation horizon was curtailed at September 30, 2021. It was, thus, possible I underestimated the incremental costs and benefits of all interventions, particularly the vaccination ones with a late roll-out (i.e. not enough time to reach sufficient coverage and/or account for DALYs averted in subsequent waves). The absence of data on seroprevalence or other leading indicators from Kabwe to inform the model during subsequent waves was a limitation to extend my analysis. Alas, as my analysis in Chapter 4 demonstrated, by September 2021 79% (95%CrI 74 to 84) of the population had already been infected, and 55% (95%CrI 51 to 58) had immunity against infection. It was, therefore, also possible that the effectiveness of any vaccination strategy beyond this date, even at high roll-out capacities, would have been limited. Moreover, the study period of my analysis in Kabwe encompassed the waves caused by the Beta and Delta VOCs, which were the most severe to circulate in Zambia and, as a result, the ones over which most of the excess mortality of the pandemic accrued [52].

Secondly, the limitations I listed in my previous chapter, as a result of a lack of cause- and age-specific mortality or COVID-19 healthcare data in Kabwe also applied to this health economic analysis. This implied a limited capacity of my model to reduce uncertainty around the estimated health and economic outcomes of the simulated interventions. As I discussed, though, my key aim was to asses how advanced analytics can make the most of the available data to the end of informing decision-making in resource constrained settings. With considerable milestones now achieved to improve surveillance in Africa [213], future analyses will greatly benefit from improved data.

Thirdly, I did not explicitly factor in healthcare capacity scale-up or consumables' volatility costs (e.g. personal-protective equipment procurement under market disruptions). However, I estimated case-management costs as the ratio of per-patient outpatient and per-patient-day inpatient case management to all other costs reported in a granular, ingredients' based COVID-19 analysis in Kenya [207]. The latter accounted for the overhead costs of staffing, facility equipment and consumables, which may have indirectly absorbed some scale-up and volatility costs. Moreover, I conducted a cost-sensitivity analysis across a wide range of variability in my model's cost parameters, and assessed the cost-effectiveness of intervention under the highest and lowest plausible range for the ocWTP using a validated approach [200].

Lastly, I did not consider other sources of potentially considerable costs to any setting's pandemic response. Namely, I did not account for costs related to investment into research and development of therapeutics, incremental programmatic and logistics costs when delivering doses at shorter timelines (e.g. if the 100-day mission had been achieved) or even the logistics and costs related to the factual procurement of the doses that arrived in Zambia. However, such costs would have been largely borne by international and multi-lateral organisations, rather than by the local health system.

5.5 Conclusion

The use of advanced analytics for decision-making support was seldomly available for LMIC settings, particularly for the most resource-constrained, as it was the case for Kabwe. Some notable exceptions included the use of global data dashboards [28, 131] and modelling projections of pandemic trajectories [192,214]. Of the latter, it was only the projections by Watson et al. [192] which were continuously updated during the pandemic and, as vaccines became available, offered the possibility to decision-makers around the world to conduct scenario-based projections. Yet, public health decision-making during a pandemic emergency is even more nuanced than this. It requires a careful consideration of the available capacity and competing priorities within a health system. Moreover, the consideration of potential health and economic outcomes of potential intervention options requires analytics underpinned by a granular understanding of local pandemic transmission and severity dynamics.

By this token, my work opened up new possibilities to explore the epidemiological drivers of the burden of outbreak emergencies leveraging on limited surveillance data. Beyond this, as of the submission date of my PhD dissertation, my analysis herein was, to my knowledge, the most detailed assessment of the real-world health and economic implications of pandemic inequality in an LIC setting. My analysis in this chapter, however, only included a range of counterfactual interventions informed by what was achieved in higher income settings during the COVID-19 pandemic. In the next chapter, I present an extended dominance analysis of a comprehensive range of mutually exclusive strategies to the implementation of CEPI's 100-day mission in Kabwe.

Chapter 6

Future pandemic preparedness in Kabwe, Zambia anchored on learnings from the COVID-19 pandemic: modelling the 100-day mission

The work presented in this chapter remain unpublished to date.

6.1 Background

My analysis in Chapter 5 showed that the majority of the burden of deaths as a result of the pandemic in Kabwe could have been averted with greater global pandemic response equality, in line with prior evidence from across LMICs [90, 183, 195]. Alas, a roll-out of vaccines in December 2020 at the same capacity as HICs may have not been cost-effective in Kabwe, depending on the opportunity cost-based willingness to pay (ocWTP) threshold of the local health system. In turn, a hypothetical roll-out of vaccines as per CEPI's 100-day mission [194] with the same capacity of HICs could have achieved an almost complete mitigation of the local pandemic in Kabwe and be highly cost-effective. In addition, I showed that the implementation of healthcare improvements were likely to be cost-effective throughout the pandemic, with earlier implementations achieving the highest benefit.

Going forward, implementing CEPI's 100-day mission and/or early healthcare improvements in a setting with severe healthcare constrains like Kabwe, however, will likely be more complicated than what my analysis so far has shown. Several operational decisions will need to be considered and weighted against each other and against the capacity of the health system. Namely, given delays to the availability of vaccines, decision-makers will need to consider issues around which sectors of the population (e.g. age groups) to target, given a quota of vaccine doses available. In addition,

directing resources to healthcare improvements might detract them from a vaccination programme and vice versa. Such pragmatic nuances will, in turn, need to be weighted against a health system's ocWTP threshold and uncertainty in the estimated health and economic outcomes of the possible interventions under consideration.

My final PhD aim was to assess how, by accounting for such pragmatic nuances, outbreak response interventions can be optimised in the context of resource constrained health systems. In this final results chapter, I thus expanded on my previous analysis to perform an extended dominance of a comprehensive set of mutually exclusive intervention options for implementing CEPI's 100-day mission in Kabwe, Zambia. I assessed these in the context of COVID-19, had vaccines been available in May 2020, at virtually no delay from the mission, or at quarterly delays up to a year thereafter. At each delay window, the simulated interventions varied: the minimum age for COVID-19 vaccines eligibility; doses delivered per-week, per-100 population (henceforth referred to as *weekly coverage rate*); and whether to implemented a *vaccination only* interventions, or a *combined* vaccination and healthcare improvement one.

6.2 Methods

In Chapter 4 I described the structure, parameters, data sources and equations of my epidemiological model for Kabwe, Zambia, and in Chapter 5 I provided a technical description of the adaptations I made to the epidemiological model to simulate counterfactual vaccination and healthcare improvement scenarios, and the integration of the economic model to assess the cost-effectiveness of such interventions.

In this methods section, I focus on describing the mutually exclusive scenarios of the implementation of the 100-day mission I simulated in Kabwe, and the technical approach I followed to perform my extended dominance analysis. The evaluation horizon was from March 2020 and September 2021. For each scenario, the economic output assessed was the total cost of implementing the intervention, from the health system's perspective. The health outcome, on the other hand, was the number of averted disability-adjusted life years (DALYs), compared to the factual (model-fitted) scenario as in Chapter 5, at 22,536 (95%CrI 19,679 to 25,394).

6.2.1 Mutually exclusive intervention options

I considered a total of 180 mutually exclusive intervention option scenarios. I assumed the same cost (Table 5.6) and vaccine efficacy (Table A3) parameters, the same probabilities of receiving outpatient or hospital care and length of hospital stay parameters (Table 5.3), and calculated a lower and upper plausible range for the ocWTP as in Chapter 5.

Vaccination

During 2021, the first year of COVID-19 vaccines availability, the vaccination capacity of LICs severely lagged behind that of higher income settings, many of which even began the roll-out of booster doses by the autumn of 2021 (Figure 6.1) [28]. Despite all countries following an older-to-young prioritisation of doses, policies on the minimum age for vaccines eligibility and the weekly vaccination rate varied over time and geography (see Figure 5.3).

Many HICs, for example, initially targeted individuals aged 50 and older, and subsequently included younger age groups as coverage targets were gradually achieved [28, 215]. Their weekly coverage rate capacity increased from a median of 9.5% (doses per-100 population-week, interquartile range [IQR] 4.1 to 16.6) during the first quarter of 2021 to 143.2% (IQR 135.1 to 153.9) by the last quarter of that year (Figure 6.1) [28]. In contrast to this, given the severe delays to the local availability of vaccines LICs, the median weekly coverage rate was of only 0.1% (IQR 0 to 0.1) in early 2021, increasing to 6% (IQR 4.4 to 7.8) by the end of the year (Figure 6.1) [28]. At the time the population-level vaccination programme was rolled out in Zambia in late 2021, WHO guidelines stipulated a minimum age eligibility of 12 years, which was followed locally [68, 203].

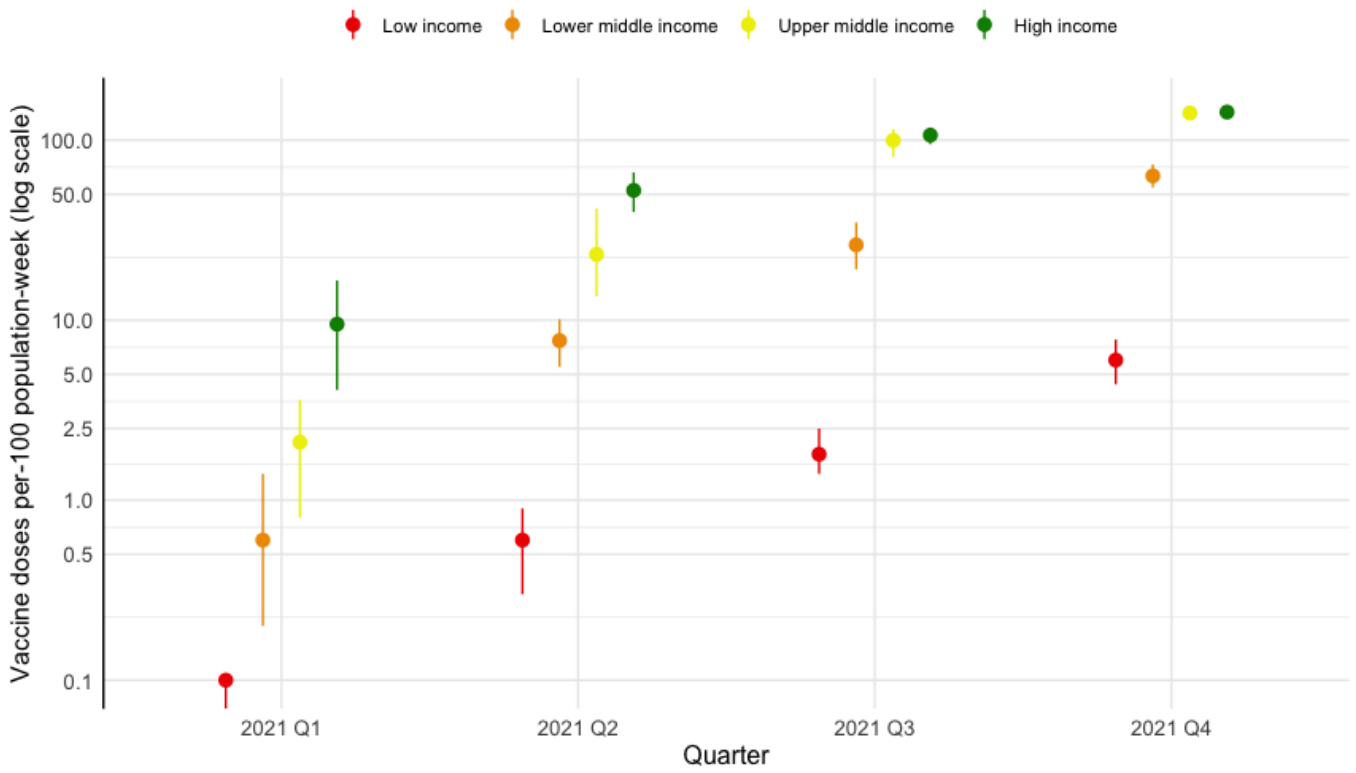


Figure 6.1: Factual vaccination capacity by quarter and income setting during 2021, the first full year of COVID-19 vaccines availability. Points and lines correspond to the median and interquartile range (IQR) of factual data aggregated by income setting, weighting by population size. The median (IQR) vaccination rate (doses delivered per-100 population-week) during 2021 increased from 0.1 (0-0.1), 0.6 (0.2-1.4), 2.1 (0.8-3.6) and 9.5 (4.1-16.6), in Q1 in LICs, L-MICs, U-MICs and HICs, respectively, to 6 (4.4-7.8) 63.3 (54.6-73.3), 141.4 (130.6-152.4) and 143.2 (135.1-153.9), respectively, by Q4. Source of data [28].

The disparities in weekly vaccination rates attained by different size economies were driven by several factors. The COVAX initiative, launched before the approval of any vaccine candidate in 2020, was a global multilateral collaboration led by the WHO and CEPI to catalyse the equitable access to COVID-19 vaccines [216]. However, advanced purchasing agreements (i.e. before vaccine production began) put in place by HICs, the centralisation of vaccines manufacturing capacity, and logistic constraints for the transportation of doses (e.g. cold chain requirements) were determinant factors in delaying the availability of vaccines to LICs at the start of 2021 [42]. Furthermore, once doses were available in LICs, limits in the operational capacities of local health systems to put "jabs to arms" forestalled the delivery of doses [42, 68, 208].

In light of the above real-world nuances of vaccination policies and capacities of LICs, I simulated the following mutually exclusive vaccination interventions:

- Five roll-out scenarios, either without delay from CEPI's 100-day mission, in early May 2020, or at quarterly delays thereafter up to a year later
- Three weekly coverage rate scenarios, at 0.5%, 2.5% and 5% (doses per-100 population-week, in line with the factual capacity of LICs attained by quarters two to four of 2021)
- Six scenarios of minimum age for vaccines eligibility, at 12+, 20+, 30+, 40+, 50+ and 65+ year-olds

As I presented in Chapter 5 (see Table 5.6), my assumed cost per-dose delivered represented an average cost to the health system in Zambia across the first year of operations of its vaccination programme [208]. Notably this estimate accounted for personnel salaries, investments in transportation and cold chain and community mobilisation campaigns, which were found to account for a large proportion of the programme's operational costs [208]. I thus did not assume any difference in costs given time to roll-out or coverage rate. All the above 90 scenarios were simulated either in isolation, as a vaccination only intervention, or as a combined vaccination and healthcare improvement intervention, as I describe in the section below. It should be noted, however, that it is unlikely the latter estimate accounted for upstream costs, such as vaccines research and development and manufacturing investments.

Improved healthcare

In Chapter 4, I found that, the IFR_t of the first wave in Kabwe stood at a staggering 0.34% (95%CrI 0.19-0.51), almost three-times higher than what would have been expected given the very young age distribution of the local population. In turn, my counterfactual analysis in Chapter 5, showed that a decrease in the inferred levels of IFR_t by 30%, given improvements in healthcare provisions for

COVID-19 cases, would have been needed to observe the expected level of pandemic severity, if local healthcare capacity had been comparable to other settings [92, 119].

Here, the combined vaccination and healthcare scenarios I simulated accounted for this reduction in IFR_t, driven by a corresponding 30% increase in outpatient consultations and inpatient admissions, as I described in Section 5.2.3. I maintained the assumptions of a linear decrease in IFR_t and increase in case-management over a period of 3 weeks and of fixed per-patient costs, regardless of the time of implementation, which I further assumed to be the same as the time of vaccination roll-out (see above).

6.2.2 Extended dominance analysis

A decision problem with multiple mutually exclusive interventions requires a comparative assessment of the health and economic benefits of each, given all possible intervention pair combinations [217]. Extended dominance provides a methodology to perform such a comparison given the estimated health benefits of the considered interventions and the cost of implementing them [218].

In this methodology, strongly dominated alternatives are those that are more costly and yield less benefit than the compared alternative [218]. Such intervention should not be chosen, irrespective of any WTP threshold, and are excluded from any subsequent analyses [217, 218]. The remainder of strategies can then be ranked according to their net cost and compared sequentially given their incremental cost-effectiveness ratio (ICER) [217]

$$ICER = (Cost_2 - Cost_1)/(Outcome_2 - Outcome_1) \tag{6.1}$$

where option 1 corresponds to the one with the lowest cost and option 2 the next more costly one. Extensively dominated strategies, also known as weakly dominated, will be those with an ICER (i.e. cost per unit of health benefit) greater than that of a more effective alternative [217, 218]. The linear combination of the remainder dominant strategies forms an efficient frontier, which thus represent the set of preferable, more effective options at a given budget [217].

To perform an extended dominance analysis, I calculated the median DALYs averted compared to the factual baseline, and its total costs in 2022-USD, from the health system perspective, across 250 stochastic realisations (simulations) for each of the afore-described 180 interventions. Then, at each possible implementation of CEPI's 100-day mission for Kabwe (i.e. no delay to its implementation or a 3-, 6-,9- or 12-month delay), I evaluated the presence of strong and weak (extended) dominance across available intervention options. Lastly, I calculated the ICERs between the remaining, non-dominated interventions, as above, to determine an effective frontier by delay window.

6.2.3 Economic analysis uncertainty

When presented with sets of mutually exclusive interventions, decision makers will often aim to maximise the net-monetary benefit (NMB) given their WTP threshold for each additional unit of health benefit [219]. Choosing a strategy merely through the assessment of cost-effectiveness planes (CEP), however, can lead to incorrect identifications of the intervention with the highest expected NMB in the presence of high variance in modelled estimates (e.g. given parameter uncertainty or stochastic variance) and/or similar cost-effectiveness ratios across strategies [219]. The calculation of cost-effectiveness acceptability curve (CEAC) and, in particular, incremental NMB curves have been proposed as means to clearly present and assess the impact of uncertainty in determining the optimal intervention option [219].

My extended dominance analysis, as described above, was done on the median effect and costs of the considered mutually exclusive interventions, as per accepted methods for it [217, 218]. Thus, I conducted an uncertainty analysis on the non-dominated interventions by assessing the median and 95%CrI of their CEP, CEAC and incremental NMB curves across stochastic realisations ($n = 250$) by scenarios of delay from CEPI's 100-day mission.

Cost-effectiveness planes

The CEP of mutually exclusive strategies i is used to present the incremental cost $\overline{\Delta C_i}$ and effect $\overline{\Delta E_i}$, compared to status quo, under uncertainty [217, 219]. For each stochastic realisation of a non-dominated intervention i and a given delay scenario, I calculated these as

$$\overline{\Delta C_i} = C_i - C_{factual} \tag{6.2}$$

$$\overline{\Delta E_i} = E_i - E_{factual} \tag{6.3}$$

where C_i was the sum of total hospitalisation, outpatient consultation and vaccination costs (assuming parameters from Table 5.6), and E_i the sum of DALYs averted, accrued across the entire simulation window.

I contextualised the estimated CEP of the interventions against the ocWTP. As I noted in Chapter 5, this represents the monetary threshold at which an intervention would result in one less health-benefit unit as a result of diverting funds from other current-day health system interventions [200]. Assuming the same parameters as in Table 5.9, I calculated a plausible range for the ocWTP from \$57.69 to \$1,146.64 2020-USD per DALY-averted, or between 6% and 120% the GDPpc of Zambia per DALY averted, respectively.

Cost-effectiveness acceptability curves

The CEAC compares the probability of each mutually exclusive strategy i being the one with the highest NMB, given a range of WTP per-benefit values ω as [217, 219]

$$CEAC(\omega) = \frac{I_i}{N} \tag{6.4}$$

where I_i is the number of simulated iterations where intervention i has the highest NMB and N the total number of iterations [219]. Here, I considered a wider range of possible values of the WTP ω, including the aforementioned ocWTP thresholds, at between \$0 and \$2,294 2022-USD, 2-times the GDPpc of Zambia. For example, if considering two interventions, x and y, the CEACs at a WTP of 0.5-times the GDPpc per DALY averted might show that x maximises the NMB in 30% of iterations N and y the remaining 70%, a ratio that might then be either maintained or dynamically change at different WTP thresholds.

Incremental net-monetary benefit

In the absence of further information, risk-neutral decision-makers might opt for options with the highest CEAC (i.e. probability of maximising the NMB) at a chosen WTP value. However, high model variance can bias the interpretability of the latter and, therefore, assessing uncertainty-augmented incremental NMB lines should be considered instead [219].

Interventions that attain the highest incremental NMB with the lowest uncertainty at a given WTP represent the optimal choice [219]. I thus estimated the incremental NMB $NMB(\omega)$ as [219]

$$NMB(\omega) = \omega\overline{E_i} - \overline{C_i}. \tag{6.5}$$

To present model uncertainty around incremental values of $NMB(\omega)$, I calculated the median and 95% credible interval across stochastic realisations ($n = 250$) of each mutually exclusive intervention i. As above, I considered a range of ω between 0 and 2-times the GDPpc of Zambia per DALY averted.

6.3 Results

6.3.1 Extended dominance analysis

At a roll-out without delay from CEPI's 100-day mission, all intervention scenarios targeting individuals aged 40 or older with a weekly coverage of 2.5% of less were strongly dominated and, thus, excluded from subsequent analyses. Conversely, a combined healthcare improvement and vaccination

intervention targeting individuals aged 12+ with a 5% weekly coverage rate was never dominated (Figure 6.2). Alas at delay windows of 9 months or greater, its incremental cost (with respect to baseline, model-fitted outcomes) would have exceeded the lowest plausible ocWTP threshold of the Kabwe health system. Such delays would, thus, have conditioned the consideration of more pragmatic interventions within the non-dominated options in the efficient frontier. That is, interventions with either a higher minimum age for vaccines eligibility and/or a vaccination-only strategy, instead of a combined one (Figure 6.2).

If rolled-out in May 2020, without delay from the 100-day mission, interventions targeting individuals aged 12+, 20+ or 30+ with a 5% coverage, either as vaccination-only or combined, formed the efficient frontier (Figure 6.2). For the case of the former two age-eligibility options, the model inferred there would not have been a significant difference in DALYs averted or total costs if rolled-out as a combined or vaccination-only intervention. For instance, a combined intervention targeting 12+ year-olds would have averted 21,983 (95%CrI 18,829 to 24,732) DALYs, compared to 21,794 (95%CrI 18,504 to 24,775) with a vaccination only one. The total costs to the health system of these intervention would have been of $1.99m (95%CrI 1.92 to 2.09) and $1.97m (95%CrI 1.92 to 2.06) 2020-USD, respectively (Figure 6.2).

Progressive delays windows, however, would have reduced the efficacy or any intervention at averting DALYs and increased its total costs. At a 6-, 9- and 12-month delay from CEPI's 100-day mission, the health impact of the above combined intervention targeting 12+ year-olds at 5% weekly coverage would have progressively decreased, averting approximately 94% (20,637, 95%CrI 17,460 to 23,415), 90% (19,707, 95%CrI 16,481 to 22,258), 72% (15,736, 95%CrI 12,112 to 18,971) and 56% (12,223, 95%CrI 7,434 to 16,058) of the DALYs burden, respectively, compared to the no-delay scenario (Figure B33). In turn, its total costs would have increased by 1.1-times ($2.22m, 95%CrI 2.03 to 2.42), 1.2-times ($2.41, 95%CrI 2.29 to 2.61), 1.5-times ($3.06, 95%CrI 2.69 to 3.33) and 1.7-times ($3.42, 95%CrI 3.19 to 3.84), respectively (Figure B34).

Whilst at a 6-month or shorter delay window this combined intervention would have been cost-effective, delays of 9 months or longer would have pushed its ICER with respect to factual (model-fitted) outcomes beyond the lowest ocWTP threshold of the health system in Kabwe (Figure B35 and Table B7). The next non-dominated options with a median incremental cost below the lowest ocWTP threshold at 9-month and 12-month delay windows would have been a combined healthcare improvement and vaccination intervention with 5% weekly coverage targeting 20+ and 40+ year-olds, respectively (Figure B35). These interventions would have still averted 15,546 (95%CrI 11,877 to 17,695) and 10,918 (95%CrI 5,538 to 15,549) DALYs, respectively, at a total cost of $2.51m (95%CrI 2.32 to 2.74) and $2.19m (95%CrI 1.89 to 2.57) 2020-USD, respectively (Figure 6.2).

The health impact and total cost of the remaining options in the efficient frontier, particularly

at the longer delay windows, was highly uncertain (Table B7). For example in the 12-month delay window, a combined and a vaccination-only intervention targeting 50+ year-olds at a 5% weekly coverage - which formed the lower tail of the efficient frontier - would have averted 9,855 (95%CrI 5,562 to 15,065) and 4,917 (95%CrI -3,031 to 11,926) DALYs, respectively, at a total cost of $2.06m (95%CrI 1.68 to 2.44) and $1.74m (95%CrI 1.41 to 2.11) 2020-USD, respectively.



Figure 6.2: Efficient frontier and dominance planes of modelled interventions by delay to roll-out from CEPI's 100-day mission. The interventions considered the implementation of COVID-19 vaccination only vs. vaccination and healthcare improvement in May 2020 in Kabwe, Zambia as per CEPI's 100-day mission as the (*No delay)* scenario, or at quarterly delays up to a year thereafter. The Y-axis reflects the median total cost of implementing the interventions. **Strongly dominated** interventions were those with yielding a lower benefit at a higher cost than compared other alternatives and were thus excluded from further analyses. Remaining interventions were ranked by total cost and ICERs were sequentially calculated for the one with the lowest cost and its immediate neighbour. **Weakly dominated** interventions were, thus, those with an ICER greater than that of a more effective alternative, with the interventions that were **not dominated** forming the efficient frontier.

6.3.2 Uncertainty analysis of non-dominated interventions

The calculation of extended dominance and the efficient frontier by delay window, as above, only accounted for the median effect and cost of the interventions. Greater delay windows and the choice of more pragmatic options drove higher uncertainty in modelled outcomes. High variance in model estimates and/or similar cost-effectiveness ratios across interventions can lead to incorrect identifications of the options with the highest NMB [219]. Thus, to further account for the effects of such uncertainty in assessing optimal (non-dominated) intervention options by delay windows, I estimated the CEP, CEAC and incremental NMB of each intervention.

In Chapter 5 (section 5.3.2), I estimated the factual (model-fitted) total burden of DALYs and costs to the health system at 22,536 (95%CrI 19,679 to 25,394) and $1.79m (95%CrI 1.61-1.91) 2020-USD, respectively. Compared to this baseline, the CEPs (Figure 6.3A) indeed showed a high variance in outcomes, particularly with regards to DALYs averted at longer delay windows. It should

be noted all economic model parameters were fixed, as described in the methods section, and so this uncertainty solely reflects the epidemiological uncertainty (variance given posterior distribution of fitted parameters) across stochastic model realisations ($n = 250$ per intervention).

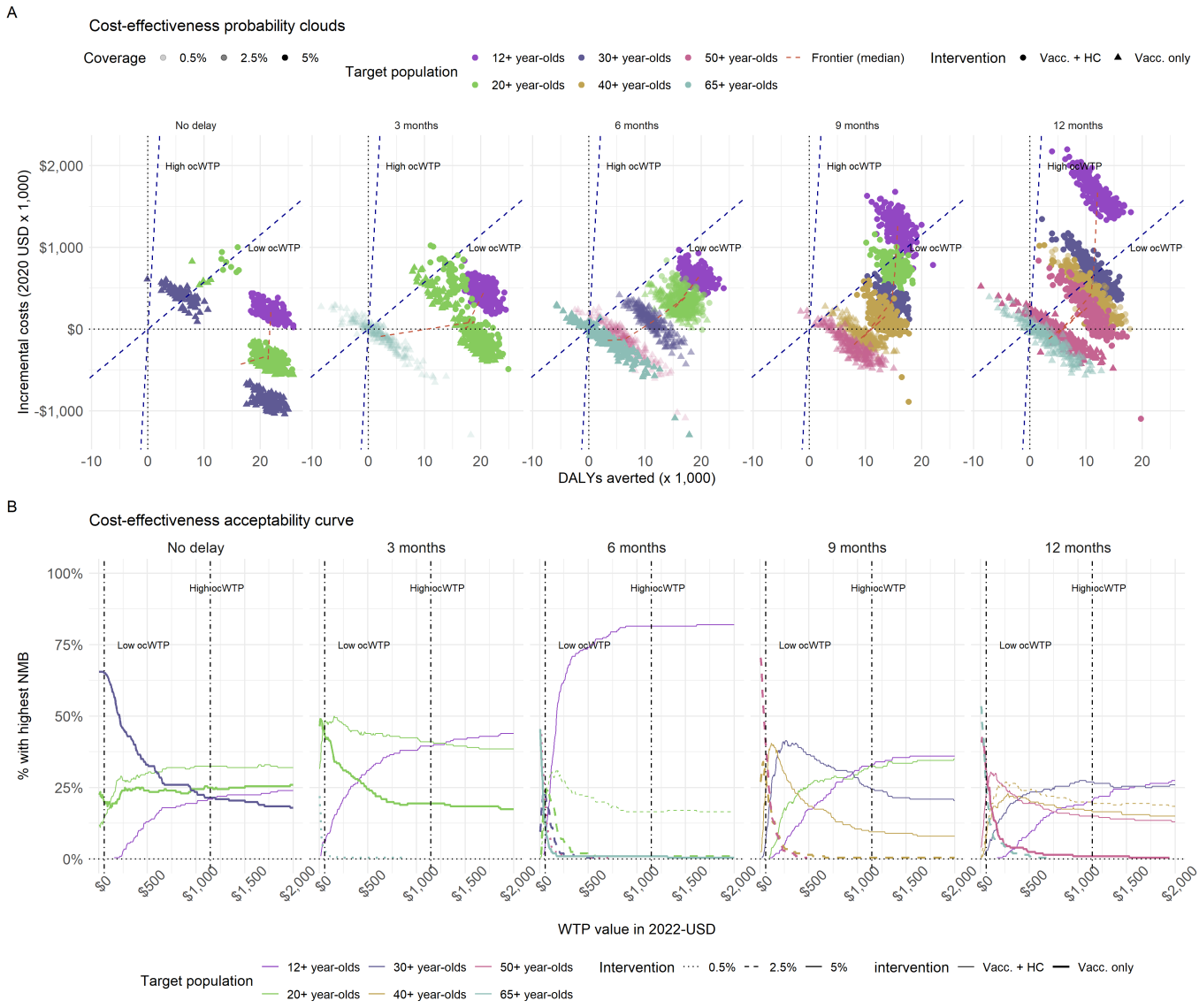


Figure 6.3: A) Cost-effectiveness uncertainty clouds and B) cost-effectiveness acceptability curves of non-dominated COVID-19 mitigation interventions in Kabwe, Zambia (from Figure 6.2) by delay scenarios, at either *No delay* from CEPI's 100-day mission (May 2020) or quarterly delays up to a year thereafter. In A), axes reflect DALYs averted (in thousands) and incremental costs (2020-USD in thousands) from the health system's perspective, with respect to the baseline factual (model-fitted) DALYs and costs, at 26,417.98 (95%CrI 23,789.45-29,908.77) and $1.84m (95%CrI 1.73-1.97) 2020-USD, respectively, as presented in Chapter 5, section 5.3.2. In B), risk-neutral decision-makers might opt for the (mutually exclusive) intervention with the highest probability of attaining the highest NMB (Y-axis) given their WTP (X-axis), in the absence of further information (e.g. incremental NMB, as in Figure 6.4) [219].

I find that the high variance in modelled outputs indeed biased the interpretability of the CEACs, with varying implications given the assumed delay window (Figure 6.3B). At no delay (Figure 6.2, left-most panel), for instance, the four intervention options that were never dominated, and therefore considered here, were: two combined ones targeting either 12+ or 20+, and two

vaccination-only ones targeting 20+ or 30+ year-olds, all at 5% weekly coverage.

The CEACs show that the vaccination-only intervention targeting 30+ year-olds attained the highest probability of maximising the NMB within a range of the ocWTP threshold between $0 and $580 2020-USD per-DALY averted (Figure 6.3B, left-most panel). Beyond this point, the combined intervention targeting 20+ year-olds emerged as the option with the highest probability of maximising the NMB, plateauing at around a 31% probability. These estimated CEACs, however, were contrasted by the respective incremental NMB lines (Figure 6.4), which showed that the combined intervention targeting 12+ with 5% weekly coverage was, in fact, yielded a similar NMB, albeit with the lowest (narrowest) uncertainty. For instance, at no delay, the model estimated that the incremental NMB of the combined intervention targeting 12+ or 20+ year-olds would have been of $1.11m (0.79 to 1.42) and $1.61m (95%CrI 0.12 to 1.96) 2020-USD at the lowest ocWTP value, and of $24.96m (21.34 to 28.33) and $25.0m (95%CrI 17.3 to 28.6) at the highest ocWTP value, respectively (Figure 6.4, left-most column).

Progressively longer delays were associated with greater uncertainty in modelled outcomes across considered, non-dominated interventions (Figure 6.3A two right-most panels, ans Figure 6.4 right-most columns). This yielded, in turn, a greater overlap in the CEACs rendering their interpretability more difficult, which was partly resolved by the assessment of the respective incremental NMB lines. At a 9-month delay window, for example, it was evident that the more pragmatic intervention options had very high uncertainty and, in turn, combined interventions targeting 12+, 20+ or 30+ at 5% weekly coverage would have been preferable across the entire range of WTP values (Figure 6.4, penultimate panels column from left to right).

6.4 Discussion

In this chapter I have performed a robust accounting of the uncertainty behind a comprehensive range of plausible (hypothetical) implementations of CEPI's 100-day mission during the COVID-19 pandemic in an LIC setting by assessing their CEP, CEAC and incremental NMB, which had been missing from the literature in this field of research. In doing so, I have unpicked key nuances and drivers of uncertainty in the formulation of optimal pandemic mitigation strategies in Kabwe, had CEPI's 100-day mission [194] been in place and achieved with the emergence of SARS-CoV-2. My results thus provide critical novel insights and actionable evidence of the plausible implications of such global pandemic preparedness agreements for LIC settings going forward.

During the evaluation horizon of March 2020 to September 2021 in Kabwe, the COVID-19 pandemic imposed as disproportionately higher local burden compared to higher income settings. This was likely accentuated by the constrained and delayed healthcare and vaccination capacity

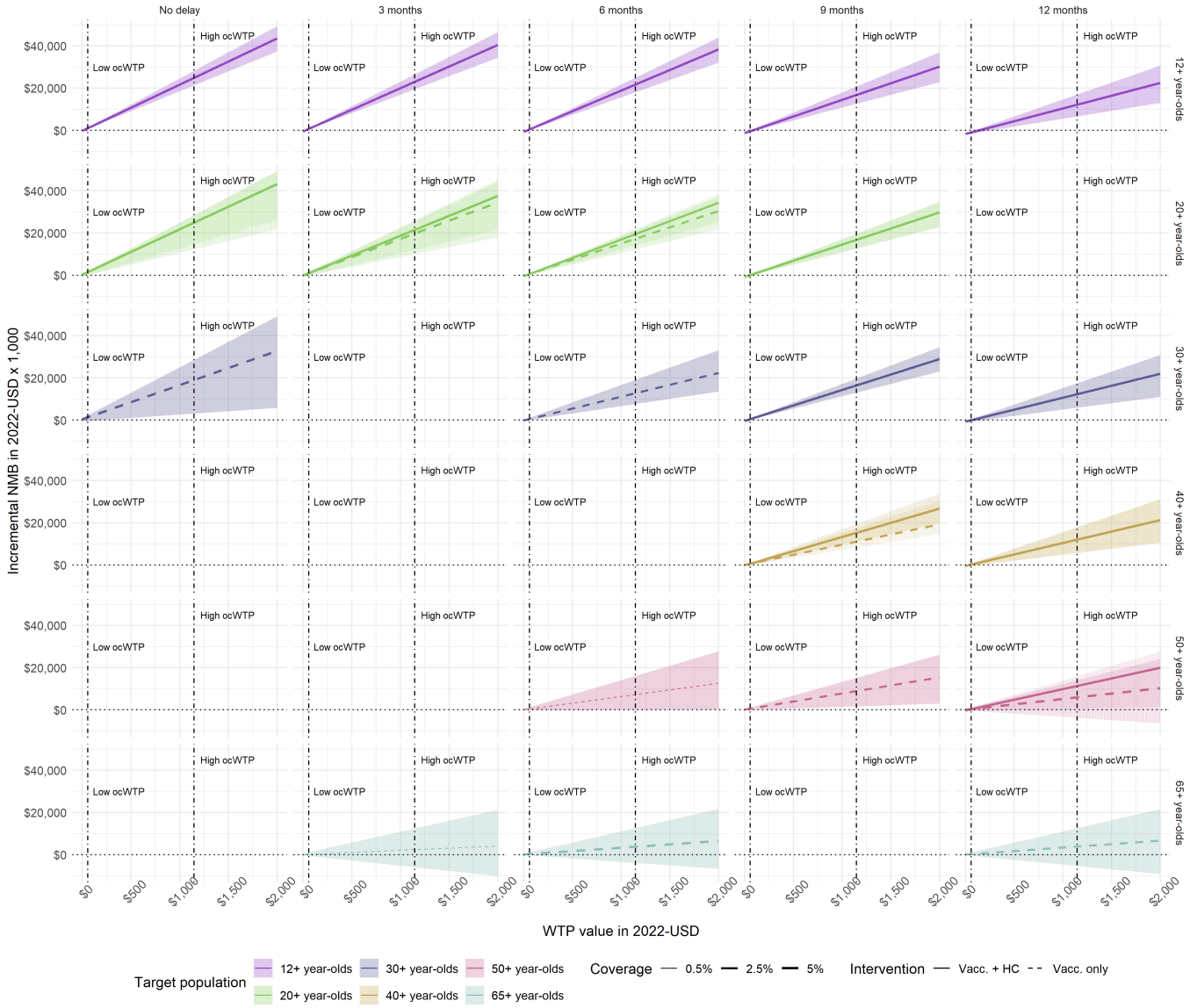


Figure 6.4: Incremental net-monetary benefit lines of non-dominated interventions implemented in May 2020 in Kabwe, Zambia as per CEPI's 100-day mission (*No delay)*, or at quarterly delays up to a year thereafter, and by minimum age for vaccine eligibility. Solid and dotted lines present the estimated median NMB and shaded areas its 95%CrI given the variance across stochastic model realisations (n = 250 per-intervention). Under high uncertainty, risk-neutral decision-makers will favour the (mutually exclusive) intervention that attains the highest incremental NMB (Y-axis) with the lowest (narrowest) uncertainty at a given WTP (X-axis) [219]. This analysis is complementary to the CEP and CEAC in Figure 6.3.

locally (Chapter 4 and Chapter 5). I found that delays to the local implementation of CEPI's 100-day mission in Kabwe, had it been in place when SARS-CoV-2 emerged, would have severely undermined its potential health and economic benefits. Rolling out the most ambitious intervention option I considered here (combined improved healthcare leading to a 30% decrease in the pandemic IFR_t, and vaccination targeting 12+ year-olds at a 5% weekly coverage) by May 2020, without delay from CEPI's 100-day mission, would have averted approximately 93% (24,545, 95%CrI 20,068 to 27,646) of the estimated baseline (model-fitted) burden of accrued DALYs. With delays to its roll-out, however, the burden of DALYs averted by this intervention would have progressively decreased

to approximately 44% (11,539, 95%CrI 5,953 to 15,738) only at a 12-month delay. This adverse effect on estimated health impact was more pronounced for more pragmatic intervention options with a higher minimum age for vaccine eligibility, lower coverage or vaccination-only without improved healthcare. Such delays would have had an even more deleterious effect on the economic outcomes of the interventions. Compared to baseline (model-fitted) accrued costs, implementing the same comprehensive intervention as above without delay would have come at an incremental cost of $0.25m 2020-USD only (95%CrI 0.11 to 0.44), but soared to an over 6-times greater cost, at $1.67m (95%CrI 1.42 to 2.17), given a 12-month delay.

My analysis further highlighted nuanced non-linearities when it comes to assessing optimal pandemic mitigation interventions under uncertainty, more so at the lowest WTP values which health systems of LICs are confined to [198–200]. On the one hand, at no delay from the 100-day mission, I estimated the efficient frontier was, rather intuitively, formed by interventions targeting individuals aged 12+ and 20+, either as vaccination-only or combined with improved healthcare, and a vaccination-only one targeting either 20+ or 30+ year-olds, all with a 5% weekly coverage. All of these interventions would have had an incremental cost below the lowest ocWTP threshold of the health system in Kabwe and, thus, would have been affordable. On the other hand, at a 9-month and 12-month delay window, the non-dominated options remaining below the lowest ocWTP threshold would have been a combined healthcare improvement and vaccination intervention with 5% weekly coverage targeting 20+ and 40+ year-olds, respectively. When considering solely their CEP and CEAC, the (arguably more pragmatic) intervention options targeting older individuals at each of these delays window would have appeared to be optimal choices under the severe economic constraints typical of LICs. Namely, they would have been cheaper than the more comprehensive counterparts (i.e. with wider age-eligibility) and, within the lower spectrum of WTP values considered, would have had the highest probability of maximising the expected NMB. However, the higher uncertainty associated with of longer delay windows and the modelling of pragmatic options drove a significantly higher variance in estimated health and economic outcomes, thus biasing the interpretability of the CEPs and CEACs. My further assessment of incremental NMB lines, in such instances, clearly showed that the (most comprehensive) combined intervention targeting 12+ with 5% weekly coverage was not only the most likely one to maximise the NMB across the entire range of WTP values considered, but was also the one with the lowest uncertainty.

In my analysis herein, I intended to comprehensively account for many real-world nuances around what would have been the decision problem of an optimal implementation of CEPI's 100-day mission in an LIC setting. However, my analysis still has several limitations. Firstly, the lack of access to cause- and age-specific mortality and COVID-19 healthcare data (e.g. admissions, occupancy), as well as granular cost data for Kabwe, as discussed in Chapter 5, are also relevant

to my extended dominance analysis herein. Such data constraints limited my model's performance to estimate the health and economic benefit of the simulated interventions. Informed by up-to-date methodological approaches in the analysis of extended dominance under uncertainty [198–200, 217, 219], however, I minimised the impact of such limitations on my results by estimating of CEP, CEAC and incremental NMB curves of all 180 interventions considered across a comprehensive range of WTP values. Surveillance data availability limitations were generalised across resource-constrained settings during the COVID-19 pandemic [71]. My work demonstrates how advanced analytics can be judiciously used to assess optimal pandemic mitigation strategies even under high uncertainty from surveillance data limitations.

Secondly, across all simulated scenarios I assumed no changes in the fitted contact rate change scaling function $\beta(t)$. It is possible that, under an earlier vaccination and healthcare improvement situation, the contact rates in Kabwe would have differed given official social distancing policies and population behaviour. Therefore, the simulated interventions can only represent what could have been assuming only changes in vaccination and healthcare for COVID-19 cases. Assessing how behavioural and social distancing policies affect the latter, however, lied outside the scope of my analysis. For the case of future pandemic preparedness, the transmissibility and severity characteristics of emerging pathogens will need to be carefully considered in context with evidence on the efficacy of pharmacological interventions, including vaccines, to weight in the role of social distancing measures.

Lastly, my results are relevant to the health system in Kabwe and other LIC settings with similar age distribution and healthcare provision. Similar analyses in other resource constrained settings with different population composition and healthcare characteristics are urgently needed to ensure future pandemic preparedness plans are adequately tailored to achieve positive health and economic outcomes.

6.5 Conclusion

A hallmark of the pandemic response in LIC settings was a limited and delays access to COVID-19 vaccines and basic healthcare commodities [42, 71], which forced decision-makers and frontline health workers to making pragmatic decisions when triaging available resources [42]. My results provide a strong indication that long delays to the implementation of emerging pandemic mitigation interventions must be avoided, more so for the case of LIC settings with severely constrained health systems. Nuances around the implications of vaccination roll-out policies, like the minimum age eligibility and vaccination rate, need to be carefully considered given delays to their availability. The methodology I have presented herein opens up avenues for exploring how to weight such considerations and opti-

mise response strategies in LICs, regardless of surveillance data and health system response capacity constraints. Going forward, it will be paramount that such nuances are accounted for by decision makers when considering local and global pandemic preparedness plans.

Chapter 7

Discussion

7.1 Summary of key findings & limitations

During my PhD, I sought to develop a parametric method to quantify the intrinsic and effective drivers of transmissibility and severity of outbreak dynamics. My further aims were to:

1. Conduct two case studies of the applied use of my approach with SARS-CoV-2 surveillance data: a) one with the rich data from England; and b) another with parsimonious data from Kabwe, Zambia

2. Demonstrate how a more systematic approach to assessing the transmissibility and severity of SARS-CoV-2 could reveal novel insights into:

    - For the case of England, i) the intrinsic transmissibility and severity of different variants; and ii) the specific contributing factor of NPIs, population immunity and healthcare ecosystem on variant-specific and overall pandemic transmissibility and severity.
    - For the case of Kabwe: i) the likely true burden of COVID-19 in a setting with a near-unmitigated pandemic; ii) an indication of the effect of limited and delayed vaccination and healthcare on the burden of the pandemic.

3. Demonstrate how accounting for the epidemiological nuances of transmission and severity can aid in optimising public-health decisions during an outbreak response under severe resource constraints.

In Chapters 2 and 3, I presented how I developed my parametric method based off the advanced mathematical model of SARS-CoV-2 transmission developed by the ICL-RTM team to inform the UK's COVID-19 response. I found that the R_0 of each variant in succession reflected its intrinsic transmissibility properties, and quantified the effect of NPIs and population immunity on the latter

to shape the dynamics of COVID-19 transmission across waves. I further showed that the IFR_0 (and all other severity metrics) of the variants was extremely sensitive to changes in healthcare provisions and population immunity. I highlighted that the breadth and depth of epidemiological surveillance in England were an enabling factor to perform this analysis, whilst acknowledging the strengths and limitations of the data throughout. Furthermore, I emphasised that it was this granularity of data that became a limitation for the applicability of many advanced analytics to settings with lesser surveillance capacity.

In Chapter 4, I studied how those limitations can be addressed through collaborative approaches to the development of outbreak analytics. I thus demonstrated that, after a short-lived reduction of contact rates as a result of NPIs in early to mid-2020, SARS-CoV-2 transmission was largely unmitigated in Kabwe, as the efficacy of subsequent NPIs was limited and vaccination at the population-level unavailable until late 2021. I further found that the severity of the local pandemic was higher than expected, likely due to constraints in healthcare provisions which, in turn, led to an excess mortality rate that was among the highest in the region.

In Chapter 5 I expanded on my analysis of the pandemic in Kabwe to assess how the counterfactual trajectory of the pandemic would have been under vaccination and healthcare provisions similar to those achieved by higher income settings. I demonstrated that the majority of the burden of the pandemic could have been avoided with the access to vaccines at the same time and with the same capacity as in HICs. Yet, given the length of the evaluation horizon considered by policy makers and their ocWTP threshold, implementing such an ambitious vaccination intervention might not have been cost-effective for the health system in Kabwe, but optimal healthcare improvements might have been.

Lastly, in Chapter 6 I proved how, if CEPI's 100-day mission had it been in place and achieved when SARS-CoV-2 emerged, the use of the analytic method I developed would have been of critical value to optimise its adoption in settings like Kabwe. My analysis gave a strong indication that an optimal implementation of the mission in Kabwe would have almost completely mitigated the local COVID-19 pandemic, and would have been highly cost-effective to the local health system. I further showed that delays to the implementation of the mission would have not only severely undermined its health benefits but likely make its implementation not cost-effective. Across the comprehensive range of (mutually exclusive) implementations of the mission I simulated, I unpicked the key drivers of epidemiological and economic uncertainty of their implementation, making a strong case of the critical need of advanced analytics from resource limited settings towards formulating more equitable global preparedness plans going forward.

Model-based epidemiological studies have been, and will remain, central to inferring how the unobserved social and biological processes behind the above factors ultimately drive the 'observable'

transmissibility and severity dynamics of an outbreak, and the likely effect of control measures [117]. However, as I found along the course of my research, three overarching limitations emerged across the different analyses I undertook, which will be relevant to development of future model-based outbreak analytic techniques.

Firstly, developing robust transmission modelling frameworks is a highly involved process. Beyond considerable technical expertise, it requires a holistic knowledge of the available surveillance data and the context under which it is collected. Whilst I strived to account for the many data limitations in my analyses at the time I conducted them, there were critical issues that my modelling approach could not overcome when the data to inform them was sub-optimal and/or insufficient.

Secondly, beyond the specific drivers I looked into, there is a myriad of other interrelated and/or complex biological and social drivers of transmissibility and severity. Of these, for example, my modelling approach did not account for detailed correlates of immune cross-protection, including individual-level neutralising antibodies and different waning rates in protection to specific variants and/or clinical severity end-points. Accounting for these and many other factors was outside of the scope of my work and, presently, lies outside of the technical capacity of available model-based inference techniques.

Thirdly, and related to the last point, the cross-discipline integration of modelling techniques adds another layer of technical challenges, with which come many additional modelling uncertainties. In the case of my analyses, this involved the integration of epidemiological and economic modelling methods, a process that demanded a thoughtful consideration of the uncertainties in each method separately (e.g. simplifying assumptions, data limitations) and uncertainty feedbacks across them.

In the following sections, I contextualise the above key findings and limitations of my work within relevant current literature, and signal avenues for future research into the development of modelling analytics to exploit surveillance data and inform response strategies to present and outbreak emergencies.

7.2 On the need for a holistic strengthening of surveillance systems

"... random sampling will cost, but bad information is expensive too." - Dean, N. 2022 [4]

Recurrent policy questions over the course of an outbreak include whether it is being caused by a novel/re-emergent pathogen or a new variant of an existing pathogen, what are its transmissibility and severity characteristics, how to monitor it and what control interventions are needed [71]. Answering all these questions, and others of both scientific and public health value, requires the use

of epidemiological intelligence (EI) systems, the cyclic process of conducting surveillance and analytics to monitor and respond to outbreak emergencies [71]. At the very foundation of this EI cycle, thus, lies the availability of locally contextualised, representative, timely and reliable epidemiological surveillance systems [71].

I have made reference to the several surveillance shortcomings that gave way to the, now debunked, 'African paradox' [61] and, conversely, to how England was an example of robust pandemic surveillance [3, 4]. However, it is also true that the COVID-19 pandemic made evident the many limitations in the surveillance systems of even the most data- and expertise-rich settings [220].

In the US, Europe and the UK, EI frameworks were proposed and widely adopted in the 2000's as means to strengthen preparedness to future epidemic threats [221–223]. In short, what these frameworks stipulated was the establishment and continuous monitoring of indicator-based and event-based surveillance systems, whereby the former included monitoring local routine data streams (e.g. healthcare, agricultural and animal health data) and the latter monitoring international and unconventional (e.g. media reports) data streams [221]. It is now known, however, that these EI systems had several critical failures. Recent analyses and independent inquiries have found that EI practitioners in both public health agencies and academic groups experienced significant, systematic challenges in accessing reliable data in a timely manner to conduct analytics throughout the COVID-19 pandemic, and that policy-makers failed to act upon independent analytics and scientific advice in a timely manner [222–224]. Specific constraints included structural inadequacies in surveillance systems and data-access restrictions of political, legal and/or economic nature [222–224].

Prior to my PhD, I conducted one of the earliest clinical cohort studies looking into the drivers of COVID-19 disease severity [225]. Prevalent pandemic misinformation myths in the UK and other HICs at the time included that individuals dying from COVID-19 were only those with extreme clinical vulnerability and who already had a short life expectancy, that the surges in clinical demand were akin to those of a 'bad flu season', or that everyone was at equal risk of severe disease or death from COVID-19 [226]. My study and similar contemporary ones helped to (gradually) debunk these myths. Yet, from my own experience and that of other clinical researchers in the UK, Europe and US I was in touch with at the time, automated health information systems at the start of the pandemic were inadequate for such analyses. Conducting them amidst the first pandemic wave instead was a time-consuming endeavour, where teams of frontline clinical staff manually collected anonymised linelist data into secure, closed-sever databases within hospitals. Alas, it was during this time that many of the above myths emerged and gained mainstream media traction [226]. Ultimately, when our analyses were completed and made public, we demonstrated that several biologic and socio-demographic factors in fact determined a heterogeneous risks to severe COVID-19 outcomes, that ethnic minorities were disproportionately affected despite a younger age at hospital presentation

and that hospital pressures had in fact increased and reasons for attendance shifted compared to pre-pandemic levels [225, 227–231].

The noxious cycle of paucity of timely EI and emergence of misinformation was, therefore, not restricted to LMIC settings. Further to my own experience with the shortfalls of epidemiological surveillance in HICs, there was a widespread notion that the COVID-19 pandemic was not disproportionately affecting resource-limited settings [232]. This came about as a result of basic analyses done with the use of case-based official surveillance systems across several countries which, at face value, gave the false impression that HICs had a higher per-capita mortality rate compared to LMICs [232]. Regrettably, this erroneous belief, along with the 'African paradox' myth, prevailed until well into 2022, when more nuanced analyses began to shed light into the devastating extent of the pandemic burden in LMICs [90, 91, 93, 129, 233].

The work I summarised in this thesis, by virtue of being retrospective, had the enormous benefit of hindsight. By the time I began my PhD, the England surveillance data streams I had access to had been streamlined. Similarly, Zambart and the ZNPHI in Zambia had already collected the data that I would later rely on. Moreover, I had accumulated a wealth of relevant experience, from how to collect and clean surveillance data to extensive secondary use of it in basic statistical analyses and the development of advanced modelling analytics. Still, I faced several limitations given the type of surveillance data I had access to, or lack thereof, which I could not overcome within the scope of my PhD project. For the case of better preparedness to epidemic threats, it is issues like this that beg the question of *where does the balance lie in terms of strengthening surveillance and EI systems?*

7.2.1 Strengthening surveillance-rich systems

In England, the ONS and REACT-1 studies were praised as examples of random sample-based, representative infection prevalence surveys [4]. As I summarised in Section 1.1, these co-existed with the UK Government's flagship community-based, universal PCR testing Pillar 2 programme which, in turn, benefited from daily random genomic sequencing of test-positive cases, and linked hospital, deaths and vaccination data streams [3]. I, however, found systematic discrepancies among data streams which were hard to reconcile, even with the use of the state-of-the-art transmission modelling framework at my disposal.

For one, there were critical periods of the pandemic in England where a fast succession of NPI changes (and likely of evolving test-seeking behaviours in the population) occurred over short periods of time, notably during the winter of 2020/21, summer of 2021 and winter of 2021/22. Although the model captured overall trends very well, it was evident it struggled to reconcile these data streams during such critical periods (Figure 7.1), albeit it should be noted all were weighted equally in the marginal log-likelihood calculation (Section 3.2.2). Additionally, I performed a series of comparisons

of key modelled trajectories to data that I did not fit the model to, which brought additional discrepancies to the foreground (Appendix A.7). In particular, whilst the model performed reasonably well at recovering hospital admissions and deaths by vaccination status, and the age distribution of the case-mix of hospitalisations and deaths (Figures A85 to A87), there was considerable model uncertainty in the trajectories of the HFR_t by NHS England region (Figure A88), most notably during the period of Alpha and Delta circulation (circa September 2020 to November 2021).

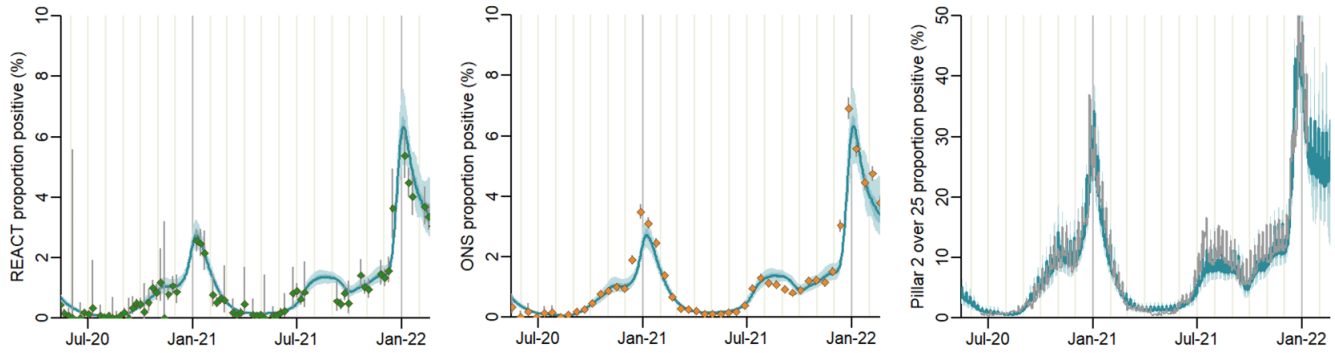


Figure 7.1: Comparison of my model's fit to data from REACT-1, ONS and Pillar 2 during the study period of my analysis in England (Chapter 3). The model was fitted independently for each NHS England region; for demonstrative purposes, panels only show a comparison of the modelled trajectory for London. It should further be noted that the model was fitted to age-disaggregated REACT-1 and Pillar 2 data, and all-ages aggregated data for ONS. Only aggregated trajectories are shown. For the full fitted trajectories, see Appendix A.4.

Exploring systematic data biases and drivers of model uncertainty around hospital trajectories, however, lied outside of the scope of my PhD. Such issues have been partially addressed by other studies. Infection-survey data, when well conducted, constitutes a gold standard for the study of community level transmission, albeit their lagged nature limits their utility during real-time analyses [117]. To this end, symptomatic case data has been traditionally employed [74]. For the case of Pillar 2 data, though, an analysis that pre-dated my PhD work demonstrated that, in isolation, it led to biased estimates of transmissibility [234]. Less-resource intensive approaches have been proposed as potential alternatives to overcome the above limitations. In particular, digital contact tracing apps in England have been proposed as non-lagged leading indicator of disease transmission with reliability (spearman statistic correlation) comparable to the gold-standard of random sample-based infection surveys [235, 236]. Unpublished work by an MRes student at Imperial College, who was supervised by leading researchers of the ICL-RTM team, looked at reproducing my analysis in Chapter 3 whilst systematically exploring the information loss (Kullback-Leibler divergence) when parametrically estimating the basic and effective variant-specific severity metrics as defined by myself (IHR, HFR and IFR), with and without specific data streams. It was thus found that, whilst VAM, hospital admission, Pillar 2 and general bed occupancy data were most informative for severity indicators, fitting the model to data from the ONS, REACT-1, seroprevalence (blood donors) and

ICU occupancy did not contribute to information loss and thus might have been less relevant for this purpose (unpublished MRes thesis by Huang, Yilei, 2024). It should be, once again, noted that during my analysis, I found that VAM data was not representative of the complete spectrum of disease severity. When linking VAM to hospitalisations and deaths linelists I, instead, found systematic discrepancies in total missing records by NHS regions (i.e. thus not allowing a linkage of databases) and a bias towards more severe cases being represented in records with full-linkage (unpublished analysis, data not publicly available).

Over its lifetime, the ONS infection survey collected approximately 11 million swabs and 3 million blood samples [237]. Yet, to the best of my knowledge, the accrued costs of this or other COVID-19 surveillance surveillance activities in England is not known. Official estimates by the UK Government have calculated that, as of March 2022, a total of £311 billion (or £4,631 per-person) were spent in 'pandemic-related support measures' across the UK, of which £96.7 billion were spent in the Health and Social Care sector [238]. There is, thus, an urgent need to direct further research into the strengthening of surveillance data in resource-rich settings. Specific areas of research must include holistic approaches for a fast-track integration of multiple data streams into ready-to-use pipelines; the assessment of potentially less-costly, scalable surveillance options as leading indicators of transmission for real-time analytics; and value of information analyses of the relative contribution of transmission, impact (severity) and genomic surveillance data streams for said analytics of value in public health decision-making.

7.2.2 Strengthening surveillance-limited systems

For the case of my analysis in Kabwe, the landscape of available surveillance data was markedly different. As I have made reference to in this thesis, much like the case of other LIC settings [42], Kabwe suffered from severe constraints to conduct real-time surveillance during the COVID-19 pandemic.

Beyond the limitations this imposed on my ability to perform inference, as detailed above, deriving meaningful inference from the available seroprevalence and all-cause mortality data required adopting fixed parameters and assumptions from my England analysis and/or relevant literature (see Figures B5 to B8). Whilst this is a common approach in modelling-based analytics, in all likelihood, this would been a critical bottleneck to performing my analysis or similar in near real-time in a setting like Kabwe. Conducting an in-depth analysis of the impact of alternative assumptions on information loss or comparing different analytic inference methods was completely outside of the scope of my work. Again, what other analyses have shown on this topic, though, is that for settings where traditional surveillance data streams are limited, triangulating multiple non-traditional data sources such as burial site worker reports, satellite imagery of cemeteries and social media-conducted

infection surveys can be an alternative to inform real-time epidemiological assessments [93, 239].

The costs of conducting representative, random sample-based surveillance during the COVID-19 pandemic in settings with severe resource constraints were prohibitively high [240, 241]. Relatedly, with landmark initiatives now well underway to strengthen vital registration systems and surveillance systems across LMICs [213, 242], calls have been made to trial the integration of unsupervised machine-learning methods into emerging surveillance data streams [159, 243]. All this underscores the case for thoroughly considering where the balance lies into what is sufficient *vs* too little surveillance. In particular, further research into the strengthening of surveillance data in resource-limited settings should focus on holistic approaches for integrating traditional and non-traditional epidemiological surveillance; fast-track integration of these data into ready-to-use pipelines; and value of information analyses of the relative contribution of both traditional and non-traditional data sources, with a robust accounting for their uncertainties and biases, when employing them to address policy-relevant questions during outbreak emergencies.

7.3 On the need for collaborative approaches to the development of outbreak analytics

The work I conducted during my PhD demonstrates that exploiting surveillance data to derive granular epidemiological insights during an outbreak response and robustly estimate the impact and optimisation of response strategies requires more holistic analytic approaches to surveillance data generation and usage. Moreover, such a holistic approach requires embedding the design and development of the analytic tools that make use of these data within efforts to strengthen surveillance.

Prior to my PhD, the experience of working alongside frontline clinicians to analyse the clinical drivers of severe COVID-19 outcomes informed the need to adapt the ICL-RTM modelling framework. When I joined the real-time modelling team in September 2020 (a year prior to beginning my PhD project) one of the tasks I was responsible for was to integrate such knowledge in order to better account for the drivers of population-level severity. Key adaptations that I, thus, performed then included a capacity to recover (fit to) age-specific infection, hospitalisation and deaths data (see Figures A20 to A56). Furthermore, whilst performing my PhD analysis in England, I found that having access to a wealth of surveillance data enabled me to explore different possibilities for fitting my model to and develop my parametric calculations. Some of such technical experimentations led to innovations that trickled back into the real-time model in use by the ICL-RTM team and, thus, into analytic reports submitted to SPI-M-O and SAGE. Increasing analytic skills at this level, whether individually and/or collectively, in order to foster innovation is time-demanding [224] and, thus, cannot be considered a sustainable approach to advancing outbreak sciences [159, 224].

Relatedly, to conduct my analyses of the pandemic in Kabwe, I spent a total of six weeks in Zambia between September 2022 and March 2024. This began with a 10-day workshop on outbreak analytics in Lusaka, which I led in collaboration with Zambart and the London School of Hygiene and Tropical Medicine [244]. Participants included academic and public health researchers from the TREATS COVID study team, Zambart, the University of Zambia, Kabwe District Health Office and the ZNPHI. Beyond strengthening capacity in the use of model-based analytics, the workshop organically led to the research collaboration that defined the latter two thirds of my PhD. Over the course of several months, we followed a systematic approach with collaborators to identify available data, analysing its quality and completeness and set research priority questions that we could answer with the modelling framework I was developing in parallel. Along the way, I had the opportunity to speak to several stakeholders, from national public health officials to frontline healthcare personnel, visit the community, health facilities and mortuaries in Kabwe from which the data was collected, and was invited to attend policy meetings where strategic plans to strengthen epidemiological surveillance in Zambia and the Southern Africa region were discussed.

These experiences were invaluable to allow myself and my collaborators to contextualise the many strengths and limitations of the available surveillance data. Furthermore, it was essential to allow me to develop my modelling framework (Chapter 4) in context with those data and the research and policy questions formulated by the team. This type of collaborative experience to conduct retrospective COVID-19 pandemic analyses was not unique to my PhD. Along the course of my own work, I got to speak to several other researchers who followed similarly collaborative approaches to develop or adapt model-based analytics to LMICs, all with a similar common objective: to retrospectively infer the likely true dynamics and impact of the pandemic in resource-limited settings [91, 93, 95, 121, 132, 206, 245, 246]. As my work has further demonstrated, however, looking at the pandemic in hindsight is not nearly enough. Rather, key epidemiological learnings from it must be put in the context of striving for improved, more equitable response interventions.

There is thus a need for systematic efforts to harness and innovate on the technical learnings made in the field of real-time outbreak analytics during the COVID-19 pandemic, in light of:

- the common policy questions they need to answer,
- the many strengths and limitations of surveillance systems,
- and the implications of future preparedness accords for settings in all geographic and income classifications.

.

Indeed, fostering cross-institution and/or multi-national EI collaborations, strengthening traditional and non-traditional surveillance, adopting scalable data collection approaches and increas-

ing analytic capacity to innovate in holistic approaches are all recurring themes of recent analyses [132, 247], lessons learnt workshops and surveys [159, 224], public inquiries [222, 223] and expert opinion pieces [4, 42, 71, 117, 220, 243]. Beyond this, though, international collaborative networks of academic and public health institutions in HIC and LMIC settings must be fostered in order to allow the co-development of future outbreak analytic innovations [242], which is urgently needed towards global preparedness equality.

7.4 On the host-pathogen-environment relationship and real-time outbreak analytics

'A model should be suited to its purpose... [it] should concentrate on the characteristics that are of interest while simplifying all others... [it] should provide a comprehensive picture of the full dynamics, and include all the relevant features of the disease and host, although determining which factors are relevant and which may be safely ignored is a complex and skilled process.' - Keeling, M 2008 [248]

The interplay of host-pathogen-environment interactions has been long considered in the development of analytics to study infectious disease dynamics [248, 249]. Almost as a parallel process, modelling analytics in the field of mathematical ecology have considered this interplay with a focus on studying the risk of pathogen spill-over across animal species, humans included, and pathogen evolutionary dynamics at the molecular level [250–252]. It feels safe to say that one positive outcome of the COVID-19 pandemic was the unprecedented innovation it ushered in the development of real-time modelling analytics (Section 1.3).

Models were able to put ever more nuanced attention to detailing non-linear dynamics of disease transmission given shifting patters in heterogeneous contact rates and disease severity given age- and clinical-vulnerability characteristics [18, 32, 33, 78, 79, 82, 89–93, 96, 106, 107, 120, 121, 195], the profile of population immunity [18, 105], and the co-transmission of SARS-CoV-2 variants [32, 33, 89, 96, 107]. Again, almost as a parallel process, genetic epidemiological analytics excelled at providing insights into the dynamics of variant emergence and dominance [25, 31, 110].

It is that much more astonishing to think about all these breakthroughs in context with how they were developed. Under relentless, unprecedented pressure on the scientists to deliver near real-time analytics to support vital public health policy decision-making by governments and public health agencies around the world. As I alluded to in the section above, though, this was far from an ideal, sustainable environment to fostering innovation in outbreak sciences [159, 224]. When I undertook my PhD studies in October 2021, the ICL-RTM team and many other academic groups continued

to be attached to this process of outbreak emergency response for several more months. Giving the intimate relation of the latter and the initial aims and objectives of my thesis, fully detaching myself from supporting the work of the ICL-RTM team (and vice versa) was not straightforward. Nevertheless, it provided the necessary pause to take a step back and re-imagine our modelling framework within the classic disease triangle (Figure 7.2).

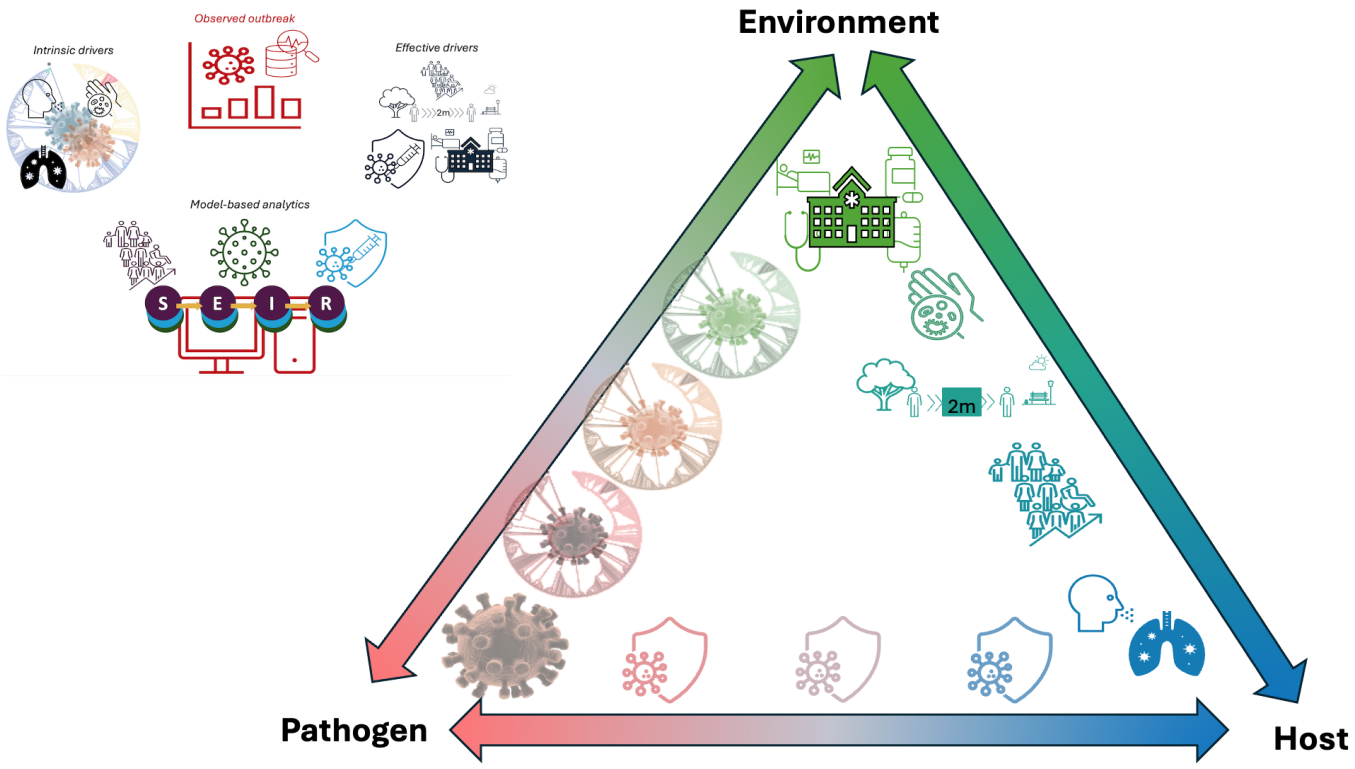


Figure 7.2: From Chapter 1, my conceptual representation of the intrinsic *vs* effective drivers of the transmission and severity of an outbreak pathogen, and how they relate to the disease triangle.

The technical innovation of incorporating a pMCMC inference framework into a real-time SARS-CoV-2 transmission model (Section 2.2.1) had been unprecedented until it was done for the ICL-RTM model [135]. With this level of inference capacity, I reconsidered the picture of dynamics characteristics that, thus far, had been simplified or ignored and began to incorporate them, informed by my prior knowledge and experience of what would be desirable in real-time modelling analytics to optimise response strategies.

When I assessed and compared the intrinsic transmissibility and severity of the Wildtype, Alpha, Delta and Omicron BA.1 variants with respects to one another in England, I demonstrated that each emerging variant in succession had a higher R_0. Although this was not surprising, what this comparison hinted at was the intrinsic characteristics of SARS-CoV-2's evolutionary processes (pathogen) acting against the backdrop of competing variants (pathogen/environment) and the dynamics of population contact rates and immunity (host/environment). I also found that the IFR_0 of the Omicron BA.1 variant was the lowest among all variants, at 59% (95%CrI 0.49 to 0.71) that of Wildtype, whilst the one of Alpha was the highest, at 2.59-times (95% 2.38 - 2.75) that of Wildtype.

Despite the caveats in my analysis I have previously discussed, this presented a systematic approach to disentangling and inferring the intrinsic severity characteristics of a pathogen and how within a transmission modelling framework, to the best of my knowledge. That is, I quantified the intrinsic (pathogen) severity of SARS-CoV-2, given the healthcare provisions and population social mixing (environment/host) characteristics of particular points of interest in the pandemic.

One can further consider response interventions to an outbreak as response modifiers of the host-pathogen-environment interactions (Figure 7.2). In this regard, pharmaceutical interventions and NPIs modify the dynamics of outbreak transmission and severity by acting on (mitigating or suppressing) its intrinsic and/or effective drivers. Whilst the intrinsic transmission advantage of Alpha and Delta over their predecessors allowed them to become dominant, the Omicron BA.1 VOC processed immune breakthrough mechanisms that gave it a decisive advantage over all preceding variant lineages [36, 167, 170]. Whilst this is now common knowledge, what my analysis categorically demonstrated was that the common misconception at the time that Omicron BA.1 was no longer a public health concern was erroneous. In fact, the isolated effect of the optimisations in available healthcare during the time of Omicron BA.1 circulation (i.e. even in the total absence of population immunity) decreased its IFR_0 by 28.6% (95%CrI 28.6 to 42.9). Now, given this effect and the existing immunity in the population from vaccination, prior infection or both, its effective IFR peaked at a mere 16.3% (95%CrI 13.4 to 19.1) of its IFR_0. Indeed, for the case of countries that had previously followed containment strategies and subsequently abandoned them during the emergence of Omicron BA.1 [42], the profile of population immunity against the virus was markedly lower [253], which led to unforeseen surges in healthcare demand [254, 255].

Despite the considerably parsimonious approach I took to adapting my modelling framework to Kabwe, I was able to demonstrate the dire effects of near-unmitigated transmissibility and constraints in host-pathogen-environment response modifiers (interventions), as I will further discuss in the next section of this chapter. My findings in Kabwe corroborated the early-pandemic notion that the protective effect of a young population profile against the IFR of SARS-CoV-2 would be nullified in the presence of healthcare constraints [92]. Elsewhere, my approach for the assessment of intrinsic and effective pandemic dynamics has been further adapted. In an analysis of the COVID-19 pandemic in French Polynesia, a similarly structured model was built to estimate the impact of vaccination and NPIs across the Wildtype, Delta and Omicron BA.1 waves [94].

I strongly believe my approach has applications beyond COVID-19 and viral pandemic pathogens, although this remains to be demonstrated. The general principle of considering mathematical models given the host-pathogen-environment disease triangle is universal to the study of animal and human infectious disease dynamics [248]. Quantifying the intrinsic and severity drivers of disease dynamics, and their response modifiers, is thus applicable to any compartmental transmission model with

disaggregation by age and pathogen (variant or clade) classes, provided replacement events do occur (i.e. one pathogen class becomes dominant over other[s]). The considerations I have given above on the need for holistic approaches to surveillance and collaborative analytics will be paramount to allow the further development of this type of transmission modelling frameworks. I consider key areas of future research in this regard should include assessing the capacity to conduct inference of correlates of pathogen transmission and severity with finer detail. In particular, identifying the types and granularity of data required, and parametric approaches to infer levels of infection-induced cross-immunity between pathogen variants/clades, vaccination-induced immunity and hybrid immunity, considering various endpoints of clinical severity (e.g. symptomatic disease, hospitalisation, death) and rates of immunity waning.

7.5 On the optimisation of outbreak response interventions

As I discussed in the previous section, outbreaks can be conceptually understood as epidemiological phenomena that result from the myriad of host-pathogen-environment interactions. In my work, I defined the latter into intrinsic and effective drivers, that are modified by response interventions. Health systems, however, have finite resources and competing opportunity costs to balance in decision-making processes during an outbreak emergency. This is a highly contextual process (e.g. geographic setting, income classification) and one that, given the many unobservable interactions within drivers in the disease triangle, requires anchoring on robust analytics.

Designing an analytic tool, from the ground up, with all the above considerations was a central motivation behind my analyses in Kabwe. That is, the contrasting burden of the COVID-19 pandemic and blatant inequalities in the global response to it, particularly between HICs and LICs [42], to me begged several poignant questions. Namely,

- Foregoing the adoption of NPIs that would be economically unfeasible for LICs, what mitigation interventions could have been adopted that would have yielded time-varying severity dynamics more in line with those seen in higher income settings (i.e. equitable response)?

- How feasible would it have been to implement such interventions, given the opportunity cost-based willingness to pay (ocWTP) threshold of health systems with severe constraints and many competing priorities?

- Going forward, how can real-time analytics holistically contextualised on available surveillance and local host-pathogen-environment interactions aid in optimising response interventions?

During the course of my collaboration with Zambart and the ZNPHI, the WHO issued its call to formulate future pandemic preparedness strategies underpinned by learnings from the COVID-19

pandemic [134] and CEPI its 100-day mission to strive for the availability of effective vaccines within 100 days of the first sequencing of a novel pandemic pathogen [194]. These calls timely spoke to the relevance of my questions. Yet, to understand how future response strategies can adequately (and equitably) mitigate pandemic threats, it is not only essential to consider what could have been done better in hindsight, but also to understand what the implementation of 'global' preparedness accords will entail in the local context of the settings where they are to be implemented.

For the case of LICs in sub-Saharan Africa, key competing disease priority areas that severely suffered during the COVID-19 pandemic included child and maternal care, tuberculosis, HIV/AIDS and malaria services [108]. Indeed, it was argued that the relative burden of disability-adjusted life-years (DALYs) imposed by the COVID-19 pandemic in sub-Saharan African (excluding South Africa) was less than a tenth of that estimated for tuberculosis, HIV/AIDS and malaria [256]. This assessment, however, was the result of a study that estimated COVID-19 DALYS given officially reported cases and deaths between March 2020 and March 2021 [256], which implies the burden of COVID-19 could have been greatly underestimated. Nevertheless, this study did underscore the crucial need to account for local contexts and opportunity costs of implementing pandemic control interventions.

As I summarised at start of this chapter, my results were in agreement with those of other studies [90, 183, 195], highlighting that most of the local pandemic burden in Kabwe could have been averted with equitable response (i.e. as per achieved by higher income settings) interventions. What had not been previously considered, however, by many counterfactual evaluations was that, given the markedly lower ocWTP thresholds that health systems in LICs, implementing such ambitions interventions could have resulted in diverting resources from other (current-day) health system interventions. For one, I estimated that the cost-effectiveness of vaccination strategies with a similar roll-out as higher income settings exceeded the lowest plausible ocWTP threshold of the health system in Zambia and were likely to result in negative NMB (i.e. a net loss in the valuable health system's economic resources). More importantly, what I clearly demonstrated was that delays to the implementation of any intervention were the biggest driver of increasing costs and, thus, unaffordability to LIC health systems.

Responding to future pandemic threats effectively will likely require an early roll-out of effective vaccines [120, 257]. For the case of resource-limited settings, however, this will only be achieved with equitable improvements in vaccines manufacturing and health system capacities [120]. My results echoed these estimations, as I found that a) the affordability of having adopted CEPI's 100-day mission in Kabwe during the COVID-19 pandemic would have been subject to an early roll-out; and that b) combined healthcare improvement and vaccination strategies were most likely to maximise the NMB (i.e. a net gain in economic resources within the health system, given disease burden

averted). For the case of LMICs in sub-Saharan Africa, it will be of the outmost importance to ensure that non-pandemic life-saving services, such as those for child-maternal health, HIV/AIDS, tuberculosis, malaria and malnutrition, are strengthened in the lead to and are not de-prioritised during a pandemic response [206, 258, 259]. To this end, what my analysis further indicated was that a judicious integration of advanced transmission and economic modelling techniques is critical to correctly assess uncertainties in the possible health and economic outputs of interventions under consideration. Many factors, as it was the case in my analysis, will likely drive high modelling uncertainty in future outbreak emergencies in LICs, including limited surveillance data, simplifying assumptions/adoption of parameters from other settings, unknown cost parameters, among others. Under any of these circumstances, or combinations thereof, the interpretability of cost-effectiveness outcomes can be compromised, which will demand more involved economic analysis methodologies, such as the use of NMB curves.

My work has laid a foundation upon which countries of any income classification, regardless of rich or parsimonious surveillance capacities, can critically appraise their learnings from the COVID-19 pandemic towards the adoption of future preparedness agreements. Certainly, however, further research on the matter is urgently needed. My work has only begun to hint at how the intrinsic and effective drivers of outbreak dynamics ultimately condition the cost-effectiveness of response interventions. Case studies from countries in different income settings, with access to different granularity of surveillance data, are needed to better quantify and understand these non-linear interactions. Relatedly, it us urgently needed to conduct case studies for outbreak pathogens other that SARS-CoV-2, with intrinsically different transmission and severity characteristics and modes of transmission. Furthermore, as I hinted to in the previous sections, the effect of correlates of intrinsic transmissibility and severity at a population level are poorly understood and, beyond having a significant role in epidemiological dynamics, are likely to impact the potential health and economic outcomes of interventions. Additionally, whilst so far I have evaluated reactive response interventions, there is an urgent need to understand the expected return on investment of preparedness strategies. This should be assessed given outbreaks by pathogen with a range of intrinsic profiles (of transmissibility and severity) and modes of transmission, and by geographic and income-level setting.

7.6 Concluding remarks

The work I conducted during my PhD was greatly motivated by a drive to understand how judiciously developed analytics can robustly exploit many and most types of surveillance data to the end of providing a granular understanding of the drivers of outbreak dynamics and plan response interventions. What my work has underscored is that considering host-pathogen-environment interactions in light of the current landscape of public health questions that modelling analytics are required to answer during outbreak emergencies requires innovation. The scope of the analytic breakthroughs I sought to address in this thesis had a specific focus, given the many biologic, social and health economic factors behind observable dynamics and public health policy decision-making during an outbreak response. I have, nonetheless, laid a cornerstone for systematically exploring the transmission and severity dynamics of pathogens in granular detail, with contrasting levels of data availability. As I have further proven, it is the systematic appraisal of those granular details that, when assessing response interventions against competing priorities, can ultimately uncover fine nuances around the optimisation of past and future response interventions.

I have made a case that, as systematic efforts are put in place to strengthen surveillance [159,213,242], the interdependence of the latter and analytics must not be forgotten, nor the relevance of cross-sectorial and cross-border collaborations that are needed to contextualise both within the geographic settings they intend to inform. A holistic approach to strengthen one component of EI systems in context with the other is what will ultimately yield fit-for-purpose preparedness plans and response strategies, equitably.

Indeed, systematic efforts to strengthen EI systems [159, 213, 242] and preparedness plans [134, 194] are underway, in the wake of the global COVID-19 pandemic. Alas, the sustainability of these efforts has been recently put into question, amidst seismic changes in the landscape of investment in global health and development [260–263]. The global community, however, must remember now more than ever before that *whether to prioritise surveillance, healthcare and/or vaccination capacity locally and/or globally* are not valid questions any more. The sobering message of my work is that, given the many known priority pathogens of outbreak potential [264], building the necessary capacity to adopt early vaccination strategies across all geographic and economic sectors cannot be de-prioritised, neither can be building resilience within healthcare systems. Limiting the devastating health and economic effects of outbreak emergencies equitably will require efforts and resources into both, as holistic surveillance and analytics capacities are also developed and strengthened.

Bibliography

[1] Perez-Guzman PN, Knock E, Imai N, Rawson T, Elmaci Y, Alcada J, et al. Epidemiological drivers of transmissibility and severity of SARS-CoV-2 in England. Nature Communications. 2023 Jul;14(1):4279. Available from: https://www.nature.com/articles/s41467-023-39661-5.

[2] Verity R, Okell LC, Dorigatti I, Winskill P, Whittaker C, Imai N, et al. Estimates of the severity of coronavirus disease 2019: a model-based analysis. The Lancet Infectious Diseases. 2020 Jun;20(6):669-77. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1473309920302437.

[3] UK Government. Technical report on the COPVID-19 pandemic in the UK; 2022. Available from: https://www.gov.uk/government/publications/technical-report-on-the-covid-19-pandemic-in-the-uk/chapter-5-modelling.

[4] Dean N. Tracking COVID-19 infections: time for change. Nature. 2022. Available from: https://www.nature.com/articles/d41586-022-00336-8.

[5] Public Health England. The First Few Hundred (FF100)" Enhanced Case and Contact Protocol v12; 2013.

[6] World Health Organization. The First Few X cases and contacts (FFX) investigation protocol for coronavirus disease 2019 (COVID-19); 2020. Available from: https://iris.who.int/bitstream/handle/10665/332023/WHO-2019-nCoV-FFXprotocol-2020.3-eng.pdf?sequence=1.

[7] Bajaj S, Chen S, Creswell R, Naidoo R, Tsui JLH, Kolade O, et al. COVID-19 testing and reporting behaviours in England across different sociodemographic groups: a population-based study using testing data and data from community prevalence surveillance surveys. The Lancet Digital Health. 2024 Nov;6(11):e778-90. Publisher: Elsevier. Available from: https://www.thelancet.com/journals/landig/article/PIIS2589-7500(24)00169-9/fulltext.

[8] Office for National Statistics. Coronavirus (COVID-19) latest insights: Infections; 2022. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/articles/coronaviruscovid19latestinsights/infections.

[9] Riley S, Walters CE, Wang H, Eales O, Ainslie KE, Atchinson C, et al. REACT-1 round 7 updated report: regional heterogeneity in changes in prevalence of SARS-CoV-2 infection during the second national COVID-19 lockdown in England. medRxiv. 2020. Publisher: Cold Spring Harbor Laboratory Press.

[10] ISARIC. ISARIC4C Comprehensive Clinical Characterisation Collaboration; 2012. Available from: https://isaric4c.net.

[11] Public Health England. COVID-19 Hospitalisation in England Surveillance System (CHESS) – daily reporting; 2020. Available from: https://www.england.nhs.uk/coronavirus/wp-content/uploads/sites/52/2020/03/phe-letter-to-trusts-re-daily-covid-19-hospital-surveillance-11-march-2020.pdf.

[12] du Plessis L, McCrone JT, Zarebski AE, Hill V, Ruis C, Gutierrez B, et al. Establishment and lineage dynamics of the SARS-CoV-2 epidemic in the UK. Science. 2021 Feb;371(6530):708-12. Publisher: American Association for the Advancement of Science. Available from: https://www.science.org/doi/10.1126/science.abf2946.

[13] UK Government. Prime Minister's statement on coronavirus (COVID-19): 20 March 2020; 2020. Available from: https://www.gov.uk/government/speeches/pm-statement-on-coronavirus-20-march-2020.

[14] Mahase E. Covid-19: UK records first death, as world's cases exceed 100 000. BMJ. 2020 Mar;368:m943. Publisher: British Medical Journal Publishing Group Section: News. Available from: https://www.bmj.com/content/368/bmj.m943.

[15] Ferguson N, Laydon D, Nedjati Gilani G, Imai N, Ainslie K, Baguelin M, et al. Report 9: Impact of non-pharmaceutical interventions (NPIs) to reduce COVID19 mortality and healthcare demand. Imperial College London; 2020. Available from: http://spiral.imperial.ac.uk/handle/10044/1/77482.

[16] Flaxman S, Mishra S, Gandy A, Unwin HJT, Mellan TA, Coupland H, et al. Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. Nature. 2020

Aug;584(7820):257-61. Publisher: Nature Research. Available from: https://doi.org/10.1038/s41586-020-2405-7.

[17] Kucharski AJ, Russell TW, Diamond C, Liu Y, Edmunds J, Funk S, et al. Early dynamics of transmission and control of COVID-19: a mathematical modelling study. The Lancet Infectious Diseases. 2020 May;20(5):553-8. Publisher: Elsevier. Available from: https://www.thelancet.com/journals/laninf/article/PIIS1473-3099(20)30144-4/fulltext.

[18] Knock ES, Whittles LK, Lees JA, Perez-Guzman PN, Verity R, FitzJohn RG, et al. Key epidemiological drivers and impact of interventions in the 2020 SARS-CoV-2 epidemic in England. Science Translational Medicine. 2021 Jul;13(602):eabg4262. Available from: https://stm.sciencemag.org/lookup/doi/10.1126/scitranslmed.abg4262.

[19] UK Government. [Withdrawn] Get a discount with the Eat Out to Help Out Scheme; 2020. Available from: https://www.gov.uk/guidance/get-a-discount-with-the-eat-out-to-help-out-scheme.

[20] Gimma A, Munday JD, Wong KLM, Coletti P, Zandvoort Kv, Prem K, et al. Changes in social contacts in England during the COVID-19 pandemic between March 2020 and March 2021 as measured by the CoMix survey: A repeated cross-sectional study. PLOS Medicine. 2022 Mar;19(3):e1003907. Publisher: Public Library of Science. Available from: https://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.1003907.

[21] UK Government. Rule of six comes into effect to tackle coronavirus; 2020. Available from: https://www.gov.uk/government/news/rule-of-six-comes-into-effect-to-tackle-coronavirus.

[22] UK Government. Prime Minister announces new local COVID Alert Levels; 2020. Available from: https://www.gov.uk/government/news/prime-minister-announces-new-local-covid-alert-levels.

[23] Davies NG, Barnard RC, Jarvis CI, Russell TW, Semple MG, Jit M, et al. Association of tiered restrictions and a second lockdown with COVID-19 deaths and hospital admissions in England: a modelling study. The Lancet Infectious Diseases. 2021 Apr;21(4):482-92.

[24] Rambaut A, Loman N, Pybus O, Barclay WS, Barrett J, Carabelli A, et al. Preliminary genomic characterisation of an emergent SARS-CoV-2 lineage in the UK defined by a novel set of spike mutations - SARS-CoV-2 coronavirus / nCoV-2019 Genomic Epidemiology. Virological. 2020 Dec. Section: SARS-CoV-2 coronavirus. Available from: https://virological.org/t/

preliminary-genomic-characterisation-of-an-emergent-sars-cov-2-lineage-in-the-uk-def
563.

[25] The COVID-19 Genomics UK (COG-UK) consortium, Volz E, Mishra S, Chand M, Barrett JC, Johnson R, et al. Assessing transmissibility of SARS-CoV-2 lineage B.1.1.7 in England. Nature. 2021 May;593(7858):266-9. Available from: http://www.nature.com/articles/s41586-021-03470-x.

[26] Davies NG, Jarvis CI, Edmunds WJ, Jewell NP, Diaz-Ordaz K, Keogh RH. Increased mortality in community-tested cases of SARS-CoV-2 lineage B.1.1.7. Nature. 2021 May;593(7858):270-4. Publisher: Nature Publishing Group. Available from: https://www.nature.com/articles/s41586-021-03426-1.

[27] UK Government. Prime Minister announces new national restrictions; 2020. Available from: https://www.gov.uk/government/news/prime-minister-announces-new-national-restrictions.

[28] Mathieu E, Ritchie H, Rodés-Guirao L, Appel C, Giattino C, Hasell J, et al. Coronavirus Pandemic (COVID-19). Our World in Data. 2020 Mar. Available from: https://ourworldindata.org/coronavirus/country/zambia.

[29] UK Government. Prime Minister announces national lockdown; 2021. Available from: https://www.gov.uk/government/news/prime-minister-announces-national-lockdown.

[30] Campbell F, Archer B, Laurenson-Schafer H, Jinnai Y, Konings F, Batra N, et al. Increased transmissibility and global spread of SARS-CoV-2 variants of concern as at June 2021. Eurosurveillance. 2021 Jun;26(24):2100509. Publisher: European Centre for Disease Prevention and Control. Available from: https://www.eurosurveillance.org/content/10.2807/1560-7917.ES.2021.26.24.2100509.

[31] Faria NR, Mellan TA, Whittaker C, Claro IM, Candido DdS, Mishra S, et al. Genomics and epidemiology of the P.1 SARS-CoV-2 lineage in Manaus, Brazil. Science. 2021 May;372(6544):815-21. Publisher: American Association for the Advancement of Science. Available from: https://www.science.org/doi/10.1126/science.abh2644.

[32] Sonabend R, Whittles LK, Imai N, Perez-Guzman PN, Knock ES, Rawson T, et al. Non-pharmaceutical interventions, vaccination, and the SARS-CoV-2 delta variant in England: a mathematical modelling study. The Lancet. 2021 Oct:S0140673621022765. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0140673621022765.

[33] Keeling MJ, Dyson L, Tildesley MJ, Hill EM, Moore S. Comparison of the 2021 COVID-19 roadmap projections against public health data in England. Nature Communications. 2022 Aug;13(1):4924. Publisher: Nature Publishing Group. Available from: https://www.nature.com/articles/s41467-022-31991-0.

[34] Perez-Guzman PN, Imai N, Knock ES, Rawson T, Sonabend R, Kanapram DT, et al. Autumn and Winter 2021-2022: potential COVID-19 epidemic trajectories. London: Imperial College London; 2021. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1027920/S1386_SPI-M_potential_winter_trajectories_Imperial_College.pdf.

[35] Viana R, Moyo S, Amoako DG, Tegally H, Scheepers C, Althaus CL, et al. Rapid epidemic expansion of the SARS-CoV-2 Omicron variant in southern Africa. Nature. 2022 Mar;603(7902):679-86. Available from: https://www.nature.com/articles/s41586-022-04411-y.

[36] Cao Y, Wang J, Jian F, Xiao T, Song W, Yisimayi A, et al. Omicron escapes the majority of existing SARS-CoV-2 neutralizing antibodies. Nature. 2022 Feb;602(7898):657-63.

[37] UK Government. COVID-19 Response: Autumn and Winter Plan 2021; 2021. Available from: https://www.gov.uk/government/publications/covid-19-response-autumn-and-winter-plan-2021/covid-19-response-autumn-and-winter-plan-2021.

[38] UK Government. COVID-19 Greenbook chapter 14a; 2022. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/984310/Greenbook_chapter_14a_7May2021.pdf.

[39] Zambia Statistics Agency. 2022 Census of Population and Housing - Preliminary Report; 2022. Available from: https://www.zamstats.gov.zm/2022-census/reports/.

[40] Zambia National Public Health Institute. Revised COVID-19 guidelines for strengthening surveillance at points of entry for the Republic of Zambia; 2022. Available from: https://w2.znphi.co.zm/wp-content/uploads/2023/04/Revised_Travel_Guidelines_Zambia-1.pdf.

[41] Hiwatari M, Yamada D, Hangoma P, Narita D, Mphuka C, Chitah B. The Kabwe Household Socioeconomic Survey (KHSS). Kabwe Mine Pollution Amelioration Initiative (KAMPAI); 2018.

[42] Sachs JD, Karim SSA, Aknin L, Allen J, Brosbøl K, Colombo F, et al. The Lancet Commission on lessons for the future from the COVID-19 pandemic. The Lancet. 2022 Sep:S0140673622015859. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0140673622015859.

[43] Mudenda S, Chileshe M, Mukosha M, Hikaambo CN, Banda M, Kampamba M, et al. Zambia's Response to the COVID-19 Pandemic: Exploring Lessons, Challenges and Implications for Future Policies and Strategies. Pharmacology & Pharmacy. 2022 Jan;13(1):11-33. Number: 1 Publisher: Scientific Research Publishing. Available from: https://www.scirp.org/journal/paperinformation.aspx?paperid=115052.

[44] Sorensen JPR, Lapworth DJ, Nkhuwa DCW, Stuart ME, Gooddy DC, Bell RA, et al. Emerging contaminants in urban groundwater sources in Africa. Water Research. 2015 Apr;72:51-63.

[45] Yabe J, Nakayama SM, Nakata H, Toyomaki H, Yohannes YB, Muzandu K, et al. Current trends of blood lead levels, distribution patterns and exposure variations among household members in Kabwe, Zambia. Chemosphere. 2020 Mar;243:125412.

[46] Mulenga LB, Hines JZ, Fwoloshi S, Chirwa L, Siwingwa M, Yingst S, et al. Prevalence of SARS-CoV-2 in six districts in Zambia in July, 2020: a cross-sectional cluster sample survey. The Lancet Global Health. 2021 Jun;9(6):e773-81. Available from: https://linkinghub.elsevier.com/retrieve/pii/S2214109X2100053X.

[47] PATH. In Zambia scientists use new tech to track SARS-CoV-2 virus; 2020. Available from: https://www.path.org/our-impact/articles/zambian-scientists-use-new-tech-track-sars-cov-2-virus/.

[48] Chipimo PJ, Barradas DT, Kayeyi N, Zulu PM, Muzala K, Mazaba ML, et al. First 100 Persons with COVID-19 - Zambia, March 18-April 28, 2020. MMWR Morbidity and mortality weekly report. 2020 Oct;69(42):1547-8.

[49] Banda J, Mweemba A, Bulaya T, Chenga N, Siziya S. Predictors of mortality in acute hospitalised COVID-19 pneumonia patients: A retrospective cohort study at two tertiary-level hospitals in Zambia. SAMJ: South African Medical Journal. 2022 Apr;112(4):273-8. Publisher: South African Medical Association. Available from: http://www.scielo.org.za/scielo.php?script=sci_abstract&pid=S0256-95742022000400008&lng=en&nrm=iso&tlng=en.

[50] Minchella PA, Chanda D, Hines JZ, Fwoloshi S, Itoh M, Kampamba D, et al. Clinical Characteristics and Outcomes of Patients Hospitalized With COVID-19 During the First

4 Waves in Zambia. JAMA Network Open. 2022 Dec;5(12):e2246152. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9856263/.

[51] Chanda SL, Tembo E, Sinyange N, Kayeyi N, Musonda K, Chewe O, et al. Characteristics of cases and deaths arising from SARS-CoV-2 infection in Zambia: March 2020 to April 2021. The Pan African Medical Journal. 2023 Aug;45:155. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10589408/.

[52] Chanda SL, Hines JZ, Malambo W, Hamukale A, Kapata N, Sinyange N, et al. An assessment of excess mortality during the COVID-19 pandemic, a retrospective post-mortem surveillance in 12 districts – Zambia, 2020–2022. BMC Public Health. 2024 Sep;24(1):2625. Available from: https://bmcpublichealth.biomedcentral.com/articles/10.1186/s12889-024-20045-3.

[53] Chanda SL, Imamura T, Malambo W, Bain R, Musuka C, Sinyange N, et al. Geographic factors associated with SARS-CoV-2 prevalence during the first wave 6 districts in Zambia, July 2020. BMC Public Health. 2025 Jan;25(1):123. Available from: https://bmcpublichealth.biomedcentral.com/articles/10.1186/s12889-025-21347-w.

[54] Katowa B, Kalonda A, Mubemba B, Matoba J, Shempela DM, Sikalima J, et al. Genomic Surveillance of SARS-CoV-2 in the Southern Province of Zambia: Detection and Characterization of Alpha, Beta, Delta, and Omicron Variants of Concern. Viruses. 2022 Sep;14(9):1865. Number: 9 Publisher: Multidisciplinary Digital Publishing Institute. Available from: https://www.mdpi.com/1999-4915/14/9/1865.

[55] PATH. Genomic sequencing explains variant spread in Zambia; 2021. Available from: https://www.path.org/our-impact/articles/genomic-sequencing-explains-variant-spread-zambia/.

[56] Wilkinson E, Giovanetti M, Tegally H, San JE, Lessells R, Cuadros D, et al. A year of genomic surveillance reveals how the SARS-CoV-2 pandemic unfolded in Africa. Science (New York, NY). 2021 Oct;374(6566):423-31.

[57] Mwenda M. Detection of B.1.351 SARS-CoV-2 Variant Strain — Zambia, December 2020. MMWR Morbidity and Mortality Weekly Report. 2021;70. Available from: https://www.cdc.gov/mmwr/volumes/70/wr/mm7008e2.htm.

[58] Loisate S, Mutembo S, Arambepola R, Makungo K, Kabalo EN, Sinyange NB, et al. Changes in mobility patterns during the COVID-19 pandemic in Zambia: Implications for the effectiveness of NPIs in Sub-Saharan Africa. PLOS Global Public Health. 2023 Oct;3(10):e0000892. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10617722/.

[59] Shanaube K, Schaap A, Klinkenberg E, Floyd S, Bwalya J, Cheeba M, et al. SARS-CoV-2 seroprevalence and associated risk factors in periurban Zambia: a population-based study. International Journal of Infectious Diseases. 2022 May;118:256-63. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1201971222001588.

[60] Mwananyanda L, Gill CJ, MacLeod W, Kwenda G, Pieciak R, Mupila Z, et al. Covid-19 deaths in Africa: prospective systematic postmortem surveillance study. BMJ. 2021 Feb;372:n334. Publisher: British Medical Journal Publishing Group Section: Research. Available from: https://www.bmj.com/content/372/bmj.n334.

[61] Gill CJ, Mwananyanda L, MacLeod WB, Kwenda G, Pieciak RC, Etter L, et al. What is the prevalence of COVID-19 detection by PCR among deceased individuals in Lusaka, Zambia? A postmortem surveillance study. BMJ open. 2022 Dec;12(12):e066763.

[62] Ghosh D, Bernstein JA, Mersha TB. COVID-19 pandemic: The African paradox. Journal of Global Health. 2020;10(2):020348. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7506193/.

[63] Kapata N, Mukonka V, Siyange N, Ngomah AM, Hamoonga R, Liwewe MM, et al.. Re: Covid-19 deaths in Africa: prospective systematic postmortem surveillance study; 2021. Available from: https://www.bmj.com/content/372/bmj.n334/rr-0.

[64] Hines JZ, Kapombe P, Mucheleng'anga A, Chanda SL, Hamukale A, Cheelo M, et al. COVID-19 mortality sentinel surveillance at a tertiary referral hospital in Lusaka, Zambia, 2020–2021. PLOS Global Public Health. 2024 Mar;4(3):e0003063. Publisher: Public Library of Science. Available from: https://journals.plos.org/globalpublichealth/article?id=10.1371/journal.pgph.0003063.

[65] Perez-Guzman PN, Chanda SL, Schaap A, Shanaube K, Baguelin M, Nyangu ST, et al. Pandemic burden in low-income settings and impact of limited and delayed interventions: A granular modelling analysis of COVID-19 in Kabwe, Zambia. International Journal of Infectious Diseases. 2024 Oct;147. Publisher: Elsevier. Available from: https://www.ijidonline.com/article/S1201-9712(24)00253-4/fulltext.

[66] Solomon H, Thea DM, Galea S, Sabin LL, Lucey DR, Hamer DH. Adherence to and enforcement of non-pharmaceutical interventions (NPIs) for COVID-19 prevention in Nigeria, Rwanda, and Zambia: A mixed-methods analysis; 2022. Available from: https://www.medrxiv.org/content/10.1101/2022.03.29.22273120v1.full.pdf.

[67] Mukwasa LL, Mutemwa R, Zambwe M, Nkhama E, Chipimo PJ. Prevalence and predictors of stress among COVID-19 health workers in Kabwe District - a cross sectional study; 2021. Available from: http://medrxiv.org/lookup/doi/10.1101/2021.12.20.21267832.

[68] Kfoury N. Zambia Reaches 70% Vaccinated; 2023. Available from: https://www.jsi.com/zambia-reaches-70-percent-vaccinated/.

[69] Kuhfeldt KJ, Kaiser JL, Morgan AJ, Ngoma T, Hamer DH, Fink G, et al. Motivation, Cues to Action, and Barriers to COVID-19 Vaccine Uptake: A Qualitative Application of the Health Belief Model among Women in Rural Zambia. The American Journal of Tropical Medicine and Hygiene. 2024 Nov;111(5):1118-26.

[70] Lusakatimes. Kabwe set for the launch of COVID 19 Vaccinations; 2021. Section: Headlines. Available from: https://www.lusakatimes.com/2021/04/20/kabwe-set-for-the-launch-of-covid-19-vaccinations/.

[71] Bhatia S, Imai N, Watson OJ, Abbood A, Abdelmalik P, Cornelissen T, et al. Lessons from COVID-19 for rescalable data collection. The Lancet Infectious Diseases. 2023 May:S1473309923001214. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1473309923001214.

[72] Bjørnstad ON. Epidemics: Models and Data using R. Use R!. Cham: Springer International Publishing; 2018. Available from: http://link.springer.com/10.1007/978-3-319-97487-3.

[73] Wallinga J, Teunis P. Different epidemic curves for severe acute respiratory syndrome reveal similar impacts of control measures. American Journal of Epidemiology. 2004 Sep;160(6):509-16.

[74] Cori A, Ferguson NM, Fraser C, Cauchemez S. A New Framework and Software to Estimate Time-Varying Reproduction Numbers During Epidemics. American Journal of Epidemiology. 2013 Nov;178(9):1505-12. Available from: https://academic.oup.com/aje/article-lookup/doi/10.1093/aje/kwt133.

[75] Nash RK, Bhatt S, Cori A, Nouvellet P. Estimating the epidemic reproduction number from temporally aggregated incidence data: A statistical modelling approach and software tool. PLOS Computational Biology. 2023 Aug;19(8):e1011439. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10491397/.

[76] Mossong J, Hens N, Jit M, Beutels P, Auranen K, Mikolajczyk R, et al. Social Contacts and Mixing Patterns Relevant to the Spread of Infectious Diseases. PLoS Medicine. 2008 Mar;5(3):e74. Available from: https://dx.plos.org/10.1371/journal.pmed.0050074.

[77] Kucharski AJ, Andreasen V, Gog JR. Capturing the dynamics of pathogens with many strains. Journal of Mathematical Biology. 2016 Jan;72(1):1-24. Available from: https://doi.org/10.1007/s00285-015-0873-4.

[78] Davies NG, Klepac P, Liu Y, Prem K, Jit M, Eggo RM. Age-dependent effects in the transmission and control of COVID-19 epidemics. Nature Medicine. 2020 Aug;26(8):1205-11. Publisher: Nature Publishing Group. Available from: https://www.nature.com/articles/s41591-020-0962-9.

[79] Munday JD, Abbott S, Meakin S, Funk S. Evaluating the use of social contact data to produce age-specific short-term forecasts of SARS-CoV-2 incidence in England. PLOS Computational Biology. 2023 Sep;19(9):e1011453. Publisher: Public Library of Science. Available from: https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1011453.

[80] Funk S, Willem L, Gruson H. Social Mixing Matrices for Infectious Disease Modelling; 2024. Available from: https://epiforecasts.io/socialmixr/.

[81] Prem K, Cook AR, Jit M. Projecting social contact matrices in 152 countries using contact surveys and demographic data. PLOS Computational Biology. 2017 Sep;13(9):e1005697. Available from: https://dx.plos.org/10.1371/journal.pcbi.1005697.

[82] Jarvis C, Gimma A, Wong K. Social contacts in the UK from the CoMix social contact survey Report for survey week 101 - Final report. Epidemics. 2024.

[83] Wong KLM, Gimma A, Coletti P, Faes C, Beutels P, Hens N, et al. Social contact patterns during the COVID-19 pandemic in 21 European countries – evidence from a two-year study. BMC Infectious Diseases. 2023 Apr;23:268. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10132446/.

[84] Loedy N, Coletti P, Wambua J, Hermans L, Willem L, Jarvis CI, et al. Longitudinal social contact data analysis: insights from 2 years of data collection in Belgium during the COVID-19 pandemic. BMC Public Health. 2023 Jul;23(1):1298. Available from: https://doi.org/10.1186/s12889-023-16193-7.

[85] Verelst F, Hermans L, Vercruysse S, Gimma A, Coletti P, Backer JA, et al. SOCRATES-CoMix: a platform for timely and open-source contact mixing data during and in between COVID-19

surges and interventions in over 20 European countries. BMC Medicine. 2021 Sep;19(1):254. Available from: https://doi.org/10.1186/s12916-021-02133-y.

[86] Jarvis CI, Gimma A, van Zandvoort K, Wong KLM, Abbas K, Villabona-Arenas CJ, et al. The impact of local and national restrictions in response to COVID-19 on social contacts in England: a longitudinal natural experiment. BMC Medicine. 2021 Feb;19(1):52. Available from: https://doi.org/10.1186/s12916-021-01924-7.

[87] Abbott S, Hellewell J, Thompson RN, Sherratt K, Gibbs HP, Bosse NI, et al. Estimating the time-varying reproduction number of SARS-CoV-2 using national and subnational case counts. Wellcome Open Research. 2020 Dec;5:112. Available from: https://wellcomeopenresearch.org/articles/5-112/v2.

[88] Diekmann O, Heesterbeek JAP, Metz JAJ. On the definition and the computation of the basic reproduction ratio R 0 in models for infectious diseases in heterogeneous populations. Journal of Mathematical Biology. 1990 Jun;28(4). Available from: http://link.springer.com/10.1007/BF00178324.

[89] Imai N, Rawson T, Knock ES, Sonabend R, Elmaci Y, Perez-Guzman PN, et al. Quantifying the impact of delaying the second COVID-19 vaccine dose in England: a mathematical modelling study. The Lancet Public Health. 2023 Feb. Https://doi.org/10.1016/S2468-2667(22)00337-1. Available from: https://www.thelancet.com/journals/lanpub/article/PIIS2468-2667(22)00337-1/fulltext.

[90] Watson OJ, Barnsley G, Toor J, Hogan AB, Winskill P, Ghani AC. Global impact of the first year of COVID-19 vaccination: a mathematical modelling study. The Lancet Infectious Diseases. 2022 Sep;22(9):1293-302. Publisher: Elsevier. Available from: https://www.thelancet.com/journals/laninf/article/PIIS1473-3099(22)00320-6/fulltext.

[91] Ghafari M, Watson OJ, Karlinsky A, Ferretti L, Katzourakis A. A framework for reconstructing SARS-CoV-2 transmission dynamics using excess mortality data. Nature Communications. 2022 May;13(1):3015. Available from: https://www.nature.com/articles/s41467-022-30711-y.

[92] Walker PGT, Whittaker C, Watson OJ, Baguelin M, Winskill P, Hamlet A, et al. The impact of COVID-19 and strategies for mitigation and suppression in low- and middle-income countries. Science. 2020 Jul;369(6502):413-22. Available from: https://www.science.org/doi/10.1126/science.abc0035.

[93] McCabe R, Whittaker C, Sheppard RJ, Abdelmagid N, Ahmed A, Alabdeen IZ, et al. Alternative epidemic indicators for COVID-19 in three settings with incomplete death registration systems. Science Advances. 2023 Jun;9(23):eadg7676. Available from: https://www.science.org/doi/10.1126/sciadv.adg7676.

[94] Chapman LAC, Aubry M, Maset N, Russell TW, Knock ES, Lees JA, et al. Impact of vaccinations, boosters and lockdowns on COVID-19 waves in French Polynesia. Nature Communications. 2023 Nov;14(1):7330. Available from: https://www.nature.com/articles/s41467-023-43002-x.

[95] Finch E, Nilles EJ, Paulino CT, Skewes-Ramm R, Lau CL, Lowe R, et al. Effects of mobility, immunity and vaccination on SARS-CoV-2 transmission in the Dominican Republic: a modelling study. The Lancet Regional Health - Americas. 2024 Sep;37:100860. Available from: https://linkinghub.elsevier.com/retrieve/pii/S2667193X2400187X.

[96] Davies NG, Abbott S, Barnard RC, Jarvis CI, Kucharski AJ, Munday JD, et al. Estimated transmissibility and impact of SARS-CoV-2 lineage B.1.1.7 in England. Science. 2021 Apr;372(6538):eabg3055. Publisher: American Association for the Advancement of Science. Available from: https://www.science.org/doi/10.1126/science.abg3055.

[97] Kucharski AJ, Klepac P, Conlan AJK, Kissler SM, Tang ML, Fry H, et al. Effectiveness of isolation, testing, contact tracing, and physical distancing on reducing transmission of SARS-CoV-2 in different settings: a mathematical modelling study. The Lancet Infectious Diseases. 2020 Oct;20(10):1151-60. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1473309920304576.

[98] Tildesley MJ, Vassall A, Riley S, Jit M, Sandmann F, Hill EM, et al. Optimal health and economic impact of non-pharmaceutical intervention measures prior and post vaccination in England: a mathematical modelling study. Royal Society Open Science. 2022 Aug;9(8):211746. Publisher: Royal Society. Available from: https://royalsocietypublishing.org/doi/10.1098/rsos.211746.

[99] Hall VJ, Foulkes S, Charlett A, Atti A, Monk EJM, Simmons R, et al. SARS-CoV-2 infection rates of antibody-positive compared with antibody-negative health-care workers in England: a large, multicentre, prospective cohort study (SIREN). The Lancet. 2021 Apr;397(10283):1459-69. Publisher: Elsevier. Available from: https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(21)00675-9/fulltext.

[100] Ramasamy MN, Minassian AM, Ewer KJ, Flaxman AL, Folegatti PM, Owens DR, et al. Safety and immunogenicity of ChAdOx1 nCoV-19 vaccine administered in a prime-boost regimen in young and old adults (COV002): a single-blind, randomised, controlled, phase 2/3 trial. The Lancet. 2020;396(10267):1979-93. Publisher: Elsevier.

[101] Mulligan MJ, Lyke KE, Kitchin N, Absalon J, Gurtman A, Lockhart S, et al. Phase I/II study of COVID-19 RNA vaccine BNT162b1 in adults. Nature. 2020;586(7830):589-93. Publisher: Nature Publishing Group.

[102] Lopez Bernal J, Andrews N, Gower C, Robertson C, Stowe J, Tessier E, et al. Effectiveness of the Pfizer-BioNTech and Oxford-AstraZeneca vaccines on covid-19 related symptoms, hospital admissions, and mortality in older adults in England: test negative case-control study. BMJ. 2021 May:n1088. Available from: https://www.bmj.com/lookup/doi/10.1136/bmj.n1088.

[103] Voysey M, Clemens SAC, Madhi SA, Weckx LY, Folegatti PM, Aley PK, et al. Safety and efficacy of the ChAdOx1 nCoV-19 vaccine (AZD1222) against SARS-CoV-2: an interim analysis of four randomised controlled trials in Brazil, South Africa, and the UK. The Lancet. 2021;397(10269):99-111. Publisher: Elsevier.

[104] Andrews N, Tessier E, Stowe J, Gower C, Kirsebom F, Simmons R, et al. Duration of Protection against Mild and Severe Disease by Covid-19 Vaccines. New England Journal of Medicine. 2022 Jan;386(4):340-50. Available from: http://www.nejm.org/doi/10.1056/NEJMoa2115481.

[105] Moore S, Hill EM, Tildesley MJ, Dyson L, Keeling MJ. Vaccination and non-pharmaceutical interventions for COVID-19: a mathematical modelling study. The Lancet Infectious Diseases. 2021 Jun;21(6):793-802. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7972312/.

[106] Barnard RC, Davies NG, Jit M, Edmunds WJ. Modelling the medium-term dynamics of SARS-CoV-2 transmission in England in the Omicron era. Nature Communications. 2022 Aug;13(1):4879. Publisher: Nature Publishing Group. Available from: https://www.nature.com/articles/s41467-022-32404-y.

[107] Keeling MJ, Moore S, Penman BS, Hill EM. The impacts of SARS-CoV-2 vaccine dose separation and targeting on the COVID-19 epidemic in England. Nature Communications. 2023 Feb;14(1):740. Publisher: Nature Publishing Group. Available from: https://www.nature.com/articles/s41467-023-35943-0.

[108] Hogan AB, Winskill P, Watson OJ, Walker PGT, Whittaker C, Baguelin M, et al. Within-country age-based prioritisation, global allocation, and public health impact of a vaccine against

SARS-CoV-2: A mathematical modelling analysis. Vaccine. 2021 May;39(22):2995-3006. Available from: https://www.sciencedirect.com/science/article/pii/S0264410X21004278.

[109] World Health Organization. Global Health Observatory - Life tables by country; 2021. Available from: https://apps.who.int/gho/data/node.main.LIFECOUNTRY.

[110] Hodcroft EB, Zuber M, Nadeau S, Vaughan TG, Crawford KHD, Althaus CL, et al. Spread of a SARS-CoV-2 variant through Europe in the summer of 2020. Nature. 2021 Jul;595(7869):707-12. Publisher: Nature Publishing Group. Available from: https://www.nature.com/articles/s41586-021-03677-y.

[111] Bhatia S, Wardle J, Nash RK, Nouvellet P, Cori A. Extending EpiEstim to estimate the transmission advantage of pathogen variants in real-time: SARS-CoV-2 as a case-study. Epidemics. 2023 Sep;44:100692. Available from: https://www.sciencedirect.com/science/article/pii/S1755436523000282.

[112] Ghani AC, Donnelly CA, Cox DR, Griffin JT, Fraser C, Lam TH, et al. Methods for estimating the case fatality ratio for a novel, emerging infectious disease. American Journal of Epidemiology. 2005 Sep;162(5):479-86.

[113] Bhattacharyya RP, Hanage WP. Challenges in Inferring Intrinsic Severity of the SARS-CoV-2 Omicron Variant. New England Journal of Medicine. 2022 Feb;386(7):e14. Available from: http://www.nejm.org/doi/10.1056/NEJMp2119682.

[114] Peng Y, Dear KBG, Denham JW. A generalized F mixture model for cure rate estimation. Statistics in Medicine. 1998;17(8):813-30. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/%28SICI%291097-0258%2819980430%2917%3A8%3C813%3A%3AAID-SIM775%3E3.0.CO%3B2-%23. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1002/%28SICI%291097-0258%2819980430%2917%3A8%3C813%3A%3AAID-SIM775%3E3.0.CO%3B2-%23.

[115] Nishiura H, Klinkenberg D, Roberts M, Heesterbeek JAP. Early Epidemiological Assessment of the Virulence of Emerging Infectious Diseases: A Case Study of an Influenza Pandemic. PLOS ONE. 2009 Aug;4(8):e6852. Publisher: Public Library of Science. Available from: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0006852.

[116] Reed C, Biggerstaff M, Finelli L, Koonin LM, Beauvais D, Uzicanin A, et al. Novel Framework for Assessing Epidemiologic Effects of Influenza Epidemics and Pandemics. Emerging Infectious Diseases. 2013 Jan;19(1):85-91. Available from: http://wwwnc.cdc.gov/eid/article/19/1/12-0124_article.htm.

[117] Cori A, Kucharski A. Inference of epidemic dynamics in the COVID-19 era and beyond. Epidemics. 2024 Sep;48:100784. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1755436524000458.

[118] O'Driscoll M, Ribeiro Dos Santos G, Wang L, Cummings DAT, Azman AS, Paireau J, et al. Age-specific mortality and immunity patterns of SARS-CoV-2. Nature. 2021 Feb;590(7844):140-5. Available from: http://www.nature.com/articles/s41586-020-2918-0.

[119] Brazeau NF, Verity R, Jenks S, Fu H, Whittaker C, Winskill P, et al. Estimating the COVID-19 infection fatality ratio accounting for seroreversion using statistical modelling. Communications Medicine. 2022 Dec;2(1):54. Available from: https://www.nature.com/articles/s43856-022-00106-7.

[120] Barnsley G, Mesa DO, Hogan AB, Winskill P, Torkelson AA, Walker DG, et al. Impact of the 100 days mission for vaccines on COVID-19: a mathematical modelling study. The Lancet Global Health. 2024 Nov;12(11):e1764-74. Publisher: Elsevier. Available from: https://www.thelancet.com/journals/langlo/article/PIIS2214-109X(24)00286-9/fulltext.

[121] Sheppard RJ, Watson OJ, Pieciak R, Lungu J, Kwenda G, Moyo C, et al. Using mortuary and burial data to place COVID-19 in Lusaka, Zambia within a global context. Nature Communications. 2023 Jun;14(1):3840. Available from: https://www.nature.com/articles/s41467-023-39288-6.

[122] Birrell PJ, Blake J, Kandiah J, Alexopoulos A, van Leeuwen E, Pouwels KB, et al. Real-time modelling of the SARS-CoV-2 pandemic in England 2020–2023: a challenging data integration. Journal of the Royal Statistical Society Series A: Statistics in Society. 2025 Apr:qnaf030. Available from: https://doi.org/10.1093/jrsssa/qnaf030.

[123] Keeling MJ, Dyson L, Moore S, Tildesley MJ, Hill E. Projections of SARS-CoV-2 transmission and COVID-19 disease until June 2022: the action of waning efficacy and boosters. Warwick University; 2021. Available from: https://assets.publishing.service.gov.uk/media/61729a3de90e07197a68c30c/S1384_Warwick_Autumn_and_Winter_scenarios.pdf.

[124] Moya A, Holmes EC, González-Candelas F. The population genetics and evolutionary epidemiology of RNA viruses. Nature Reviews Microbiology. 2004 Apr;2(4):279-88. Available from: http://www.nature.com/articles/nrmicro863.

[125] Zhou P, Yang XL, Wang XG, Hu B, Zhang L, Zhang W, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature. 2020 Mar;579(7798):270-3.

Publisher: Nature Publishing Group. Available from: https://www.nature.com/articles/s41586-020-2012-7.

[126] McCabe R, Schmit N, Christen P, D'Aeth JC, Løchen A, Rizmie D, et al. Adapting hospital capacity to meet changing demands during the COVID-19 pandemic. BMC Medicine. 2020 Dec;18(1):329. Available from: https://bmcmedicine.biomedcentral.com/articles/10.1186/s12916-020-01781-w.

[127] D'Aeth JC, Ghosal S, Grimm F, Haw D, Koca E, Lau K, et al. Optimal national prioritization policies for hospital care during the SARS-CoV-2 pandemic. Nature Computational Science. 2021 Aug;1(8):521-31. Available from: https://www.nature.com/articles/s43588-021-00111-1.

[128] Docherty AB, Mulholland RH, Lone NI, Cheyne CP, De Angelis D, Diaz-Ordaz K, et al. Changes in in-hospital mortality in the first wave of COVID-19: a multicentre prospective observational cohort study using the WHO Clinical Characterisation Protocol UK. The Lancet Respiratory Medicine. 2021 Jul;9(7):773-85. Available from: https://linkinghub.elsevier.com/retrieve/pii/S2213260021001752.

[129] Wang H, Paulson KR, Pease SA, Watson S, Comfort H, Zheng P, et al. Estimating excess mortality due to the COVID-19 pandemic: a systematic analysis of COVID-19-related mortality, 2020–21. The Lancet. 2022 Apr;399(10334):1513-36. Publisher: Elsevier. Available from: https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(21)02796-3/fulltext.

[130] World Health Organization. COVID-19 cases | WHO COVID-19 dashboard; 2020. Available from: https://data.who.int/dashboards/covid19/cases.

[131] Moss B, Watson C, Rutkow L, Garibaldi B, Blauer B, Gardner L. COVID-19 Map;. Available from: https://coronavirus.jhu.edu/map.html.

[132] Kimani TN, Nyamai M, Owino L, Makori A, Ombajo LA, Maritim M, et al. Infectious disease modelling for SARS-CoV-2 in Africa to guide policy: A systematic review. Epidemics. 2022 Sep;40:100610. Available from: https://www.sciencedirect.com/science/article/pii/S1755436522000536.

[133] Bergeri I, Whelan MG, Ware H, Subissi L, Nardone A, Lewis HC, et al. Global SARS-CoV-2 seroprevalence from January 2020 to April 2022: A systematic review and meta-analysis of standardized population-based studies. PLOS Medicine. 2022 Nov;19(11):e1004107. Publisher:

Public Library of Science. Available from: https://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.1004107.

[134] World Health Organization. WHO launches new initiative to improve pandemic preparedness; 2023. Available from: https://www.who.int/news/item/26-04-2023-who-launches-new-initiative-to-improve-pandemic-preparedness.

[135] FitzJohn RG, Knock ES, Whittles LK, Perez-Guzman PN, Bhatia S, Guntoro F, et al. Reproducible parallel inference and ... | Wellcome Open Research. Wellcome Open Research. 2021;5. Available from: https://wellcomeopenresearch.org/articles/5-288.

[136] Doucet A, Johansen A. A Tutorial on Particle Filtering and Smoothing: Fifteen years later; 2008. Available from: https://www.stats.ox.ac.uk/~doucet/doucet_johansen_tutorialPF2011.pdf.

[137] Endo A, van Leeuwen E, Baguelin M. Introduction to particle Markov-chain Monte Carlo for disease dynamics modellers. Epidemics. 2019 Dec;29:100363. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1755436519300301.

[138] Endo A, Centre for the Mathematical Modelling of Infectious Diseases COVID-19 Working Group, Abbott S, Kucharski AJ, Funk S. Estimating the overdispersion in COVID-19 transmission using outbreak sizes outside China. Wellcome Open Research. 2020 Jul;5:67. Available from: https://wellcomeopenresearch.org/articles/5-67/v3.

[139] Park SW, Bolker BM, Funk S, Metcalf CJE, Weitz JS, Grenfell BT, et al. The importance of the generation interval in investigating dynamics and control of new SARS-CoV-2 variants. Journal of The Royal Society Interface. 2022 Jun;19(191):20220173. Available from: https://royalsocietypublishing.org/doi/10.1098/rsif.2022.0173.

[140] Svensson A note on generation times in epidemic models. Mathematical Biosciences. 2007 Jul;208(1):300-11. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0025556406002094.

[141] Geismar C, Fragaszy E, Nguyen V, Fong WLE, Shrotri M, Beale S, et al. Household serial interval of COVID-19 and the effect of Variant B.1.1.7: analyses from prospective community cohort study (Virus Watch). Wellcome Open Research. 2021 Dec;6:224. Available from: https://wellcomeopenresearch.org/articles/6-224/v2.

[142] an der Heiden M, Buchholz U. Serial interval in households infected with SARS-CoV-2 variant B.1.1.529 (Omicron) is even shorter compared to Delta. Epidemiology and Infection.

2022;150:e165. Available from: https://www.cambridge.org/core/product/identifier/S0950268822001248/type/journal_article.

[143] UK Government. Regulatory approval of Ronapreve; 2021. Available from: https://www.gov.uk/government/publications/regulatory-approval-of-ronapreve.

[144] UK Government. Regulatory approval of Lagevrio (molnupiravir); 2022. Available from: https://www.gov.uk/government/publications/regulatory-approval-of-lagevrio-molnupiravir.

[145] UK Government. Regulatory approval of Paxlovid; 2022. Available from: https://www.gov.uk/government/publications/regulatory-approval-of-paxlovid.

[146] UK Government. Regulatory approval of Xevudy (sotrovimab); 2022. Available from: https://www.gov.uk/government/publications/regulatory-approval-of-xevudy-sotrovimab.

[147] Medicines and Healthcare products Regulatory Agency. Dexamethasone in the treatment of COVID-19: Implementation and management of supply for treatment in hospitals; 2020. Available from: https://www.cas.mhra.gov.uk/ViewandAcknowledgment/ViewAlert.aspx?AlertID=103054.

[148] NHS England. COVID-19 Hospital Activity;. Available from: https://www.england.nhs.uk/statistics/statistical-work-areas/covid-19-hospital-activity/.

[149] NHS England. SGSS and CHESS data;. Available from: https://digital.nhs.uk/about-nhs-digital/corporate-information-and-documents/directions-and-data-provision-notices/data-provision-notices-dpns/sgss-and-chess-data.

[150] The RECOVERY Collaborative Group. Dexamethasone in Hospitalized Patients with Covid-19. New England Journal of Medicine. 2021 Feb;384(8):693-704. Publisher: Massachusetts Medical Society _eprint: https://doi.org/10.1056/NEJMoa2021436. Available from: https://doi.org/10.1056/NEJMoa2021436.

[151] Mendes D, Machira Krishnan S, O'Brien E, Padgett T, Harrison C, Strain WD, et al. Modelling COVID-19 Vaccination in the UK: Impact of the Autumn 2022 and Spring 2023 Booster Campaigns. Infectious Diseases and Therapy. 2024 May;13(5):1127-46. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC11098993/.

[152] Tsang TK, Sullivan SG, Meng Y, Lai FTT, Fan M, Huang X, et al. Evaluating the impact of extended dosing intervals on mRNA COVID-19 vaccine effectiveness in adolescents. BMC medicine. 2024 Sep;22(1):384.

[153] Rawson T, Hinsley W, Sonabend R, Semenova E, Cori A, Ferguson NM. The impact of health inequity on spatial variation of COVID-19 transmission in England. PLOS Computational Biology. 2024 May;20(5):e1012141. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC11161116/.

[154] Liu H, Xu X, Deng X, Hu Z, Sun R, Zou J, et al. Counterfactual analysis of the 2023 Omicron XBB wave in China. Infectious Disease Modelling. 2024 Mar;9(1):195-203.

[155] Bosse NI, Abbott S, Bracher J, van Leeuwen E, Cori A, Funk S. Human judgement forecasting of COVID-19 in the UK. Wellcome Open Research. 2023;8:416.

[156] Sim JK, Lee HS, Yang J, Gwack J, Kim BI, Cha JO, et al. Comparative Analysis of Clinical Outcomes Using Propensity Score Matching: Coronavirus Disease 2019 vs. Seasonal Influenza in Korea. Journal of Korean Medical Science. 2024 Apr;39(14):e128.

[157] Faherty LJ, Nascimento de Lima P, Lim JZ, Roberts D, Karr S, Lawson E, et al. Effects of non-pharmaceutical interventions on COVID-19 transmission: rapid review of evidence from Italy, the United States, the United Kingdom, and China. Frontiers in Public Health. 2024;12:1426992.

[158] Espinosa O, White L, Bejarano V, Aguas R, Rincón D, Mora L, et al. Predictive modelling of the effectiveness of vaccines against COVID-19 in Bogotá: Methodological innovation involving different variants and computational optimisation efficiency. Heliyon. 2024 Nov;10(21):e39725.

[159] Nunes MC, Thommes E, Fröhlich H, Flahault A, Arino J, Baguelin M, et al. Redefining pandemic preparedness: Multidisciplinary insights from the CERP modelling workshop in infectious diseases, workshop report. Infectious Disease Modelling. 2024 Feb;9(2):501-18. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10912817/.

[160] Imperial College London. Imperial College COVID-19 Response Team 2020-2021 Report; 2021. Available from: https://www.imperial.ac.uk/media/imperial-college/medicine/mrc-gida/Imperial-College-COVID-19-Response-Team-2020-2021-Report-(22-03-2021).pdf.

[161] UK Government. Coronavirus (COVID-19) in the UK; 2022. Available from: https://coronavirus.data.gov.uk.

[162] UK Government. Sero-surveillance of COVID-19; 2021. Available from: https://www.gov.uk/government/publications/national-covid-19-surveillance-reports/sero-surveillance-of-covid-19.

[163] UK Health Security Agency. Variants: distribution of cases data, 20 May 2021 - GOV.UK; 2021. Available from: https://www.gov.uk/government/publications/covid-19-variants-genomically-confirmed-case-numbers/variants-distribution-of-cases-data.

[164] Brazeau N, Verity R, Jenks S, Fu H, Whittaker C, Winskill P, et al. Report 34: COVID-19 infection fatality ratio: estimates from seroprevalence. Imperial College London COVID-19 Reports. 2020.

[165] Borremans B, Gamble A, Prager K, Helman SK, McClain AM, Cox C, et al. Quantifying antibody kinetics and RNA detection during early-phase SARS-CoV-2 infection by time since symptom onset. Elife. 2020;9:e60122. Publisher: eLife Sciences Publications Limited.

[166] Harris RJ, Whitaker HJ, Andrews NJ, Aiano F, Amin-Chowdhury Z, Flood J, et al. Serological surveillance of SARS-CoV-2: Six-month trends and antibody response in a cohort of public health workers. Journal of Infection. 2021;82(5):162-9. Publisher: Elsevier.

[167] Chen S, Flegg JA, White LJ, Aguas R. Levels of SARS-CoV-2 population exposure are considerably higher than suggested by seroprevalence surveys. MedRxiv. 2021. Publisher: Cold Spring Harbor Laboratory Press.

[168] Planas D, Veyer D, Baidaliuk A, Staropoli I, Guivel-Benhassine F, Rajah MM, et al. Reduced sensitivity of SARS-CoV-2 variant Delta to antibody neutralization. Nature. 2021:1-7. Publisher: Nature Publishing Group.

[169] Kim P, Gordon SM, Sheehan MM, Rothberg MB. Duration of SARS-CoV-2 Natural Immunity and Protection against the Delta Variant: A Retrospective Cohort Study. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America. 2021 Dec.

[170] Planas D, Saunders N, Maes P, Guivel-Benhassine F, Planchais C, Buchrieser J, et al. Considerable escape of SARS-CoV-2 Omicron to antibody neutralization. Nature. 2022 Feb;602(7898):671-5. Available from: https://www.nature.com/articles/s41586-021-04389-z.

[171] Chen J, Wang R, Gilby NB, Wei GW. Omicron Variant (B.1.1.529): Infectivity, Vaccine Breakthrough, and Antibody Resistance. Journal of Chemical Information and Modeling. 2022 Jan;62(2):412-22. Available from: https://pubs.acs.org/doi/10.1021/acs.jcim.1c01451.

[172] Wu JT, Leung K, Leung GM. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. The Lancet. 2020;395(10225):689-97. Available from: https://www.sciencedirect.com/science/article/pii/S0140673620302609.

[173] Public Health Scotland. COVID-19 Statistical Report - 23 June 2021 [Statistical report]; 2021. Publisher: Public Health Scotland. Available from: https://publichealthscotland.scot/publications/covid-19-statistical-report/covid-19-statistical-report-23-june-2021/.

[174] Nyberg T, Ferguson NM, Nash SG, Webster HH, Flaxman S, Andrews N, et al. Comparative analysis of the risks of hospitalisation and death associated with SARS-CoV-2 omicron (B.1.1.529) and delta (B.1.617.2) variants in England: a cohort study. The Lancet. 2022 Apr;399(10332):1303-12. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0140673622004627.

[175] Lin L, Liu Y, Tang X, He D. The Disease Severity and Clinical Outcomes of the SARS-CoV-2 Variants of Concern. Frontiers in Public Health. 2021;9. Available from: https://www.frontiersin.org/journals/public-health/articles/10.3389/fpubh.2021.775224.

[176] Twohig KA, Nyberg T, Zaidi A, Thelwall S, Sinnathamby MA, Aliabadi S, et al. Hospital admission and emergency care attendance risk for SARS-CoV-2 delta (B.1.617.2) compared with alpha (B.1.1.7) variants of concern: a cohort study. The Lancet Infectious Diseases. 2022 Jan;22(1):35-42. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1473309921004758.

[177] Bast E, Tang F, Dahn J, Palacio A. Increased risk of hospitalisation and death with the delta variant in the USA. The Lancet Infectious Diseases. 2021 Dec;21(12):1629-30. Available from: https://linkinghub.elsevier.com/retrieve/pii/S147330992100685X.

[178] Nyberg T, Harman K, Zaidi A, Seaman SR, Andrews N, Nash SG, et al. Hospitalization and Mortality Risk for COVID-19 Cases With SARS-CoV-2 AY.4.2 (VUI-21OCT-01) Compared to Non-AY.4.2 Delta Variant Sublineages. The Journal of Infectious Diseases. 2022 Sep;226(5):808-11. Available from: https://academic.oup.com/jid/article/226/5/808/6532623.

[179] Cabore JW, Karamagi HC, Kipruto HK, Mungatu JK, Asamani JA, Droti B, et al. COVID-19 in the 47 countries of the WHO African region: a modelling analysis of past trends and future patterns. The Lancet Global Health. 2022 Aug;10(8):e1099-114. Publisher: Elsevier. Avail-

able from: https://www.thelancet.com/journals/langlo/article/PIIS2214-109X(22)00233-9/fulltext.

[180] Otshudiema JO, Folefack GLT, Nsio JM, Mbala-Kingebeni P, Kakema CH, Kosianza JB, et al. Epidemiological Comparison of Four COVID-19 Waves in the Democratic Republic of the Congo, March 2020–January 2022. Journal of Epidemiology and Global Health. 2022 Sep;12(3):316-27. Available from: https://doi.org/10.1007/s44197-022-00052-6.

[181] Semakula M, Niragire F, Nsanzimana S, Remera E, Faes C. Spatio-temporal dynamic of the COVID-19 epidemic and the impact of imported cases in Rwanda. BMC Public Health. 2023 May;23(1):930. Available from: https://doi.org/10.1186/s12889-023-15888-1.

[182] Mgbere O, Nwabuko OC, Olateju OA, Adepoju OE, Liaw W, Darkoh C, et al. Population-based assessment of the burden of COVID-19 infection in African countries: a first-year report card and public health implications. Therapeutic Advances in Infectious Disease. 2023 Jan;10:20499361231202116. Publisher: SAGE Publications. Available from: https://doi.org/10.1177/20499361231202116.

[183] Gozzi N, Chinazzi M, Dean NE, Longini Jr IM, Halloran ME, Perra N, et al. Estimating the impact of COVID-19 vaccine inequities: a modeling study. Nature Communications. 2023 Jun;14(1):3272. Available from: https://www.nature.com/articles/s41467-023-39098-w.

[184] Baguelin M, Flasche S, Camacho A, Demiris N, Miller E, Edmunds WJ. Assessing Optimal Target Populations for Influenza Vaccination Programmes: An Evidence Synthesis and Modelling Study. PLoS Medicine. 2013 Oct;10(10):e1001527. Available from: https://dx.plos.org/10.1371/journal.pmed.1001527.

[185] Public Health England, Nuffield Department of Medicine, University of Oxford, Oxford University Hospital NHS Foundation Trust. Evaluation of sensitivity and specificity of four commercially available SARS-CoV-2 antibody immunoassays; 2020. Available from: https://assets.publishing.service.gov.uk/media/5f05887ad3bf7f2be6e02161/Evaluation__of_sensitivity_and_specificity_of_4_commercially_available_SARS-CoV-2_antibody_immunoassays.pdf.

[186] Bryan A, Pepper G, Wener MH, Fink SL, Morishima C, Chaudhary A, et al. Performance Characteristics of the Abbott Architect SARS-CoV-2 IgG Assay and Seroprevalence in Boise, Idaho. Journal of Clinical Microbiology. 2020 Jul;58(8):e00941-20. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7383515/.

[187] Krutikov M, Palmer T, Tut G, Fuller C, Azmi B, Giddings R, et al. Prevalence and duration of detectable SARS-CoV-2 nucleocapsid antibodies in staff and residents of long-term care facilities over the first year of the pandemic (VIVALDI study): prospective cohort study in England. The Lancet Healthy Longevity. 2022 Jan;3(1):e13-21. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8676418/.

[188] Melegaro A, Fava ED, Poletti P, Merler S, Nyamukapa C, Williams J, et al. Social Contact Structures and Time Use Patterns in the Manicaland Province of Zimbabwe. PLOS ONE. 2017 Jan;12(1):e0170459. Publisher: Public Library of Science. Available from: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0170459.

[189] Gill CJ, Mwananyanda L, MacLeod W, Kwenda G, Pieciak R, Etter L, et al.. Sustained high prevalence of COVID-19 deaths from a systematic post-mortem study in Lusaka, Zambia: one year later. medRxiv; 2022. Pages: 2022.03.08.22272087. Available from: https://www.medrxiv.org/content/10.1101/2022.03.08.22272087v2.

[190] Chowdhury R, Luhar S, Khan N, Choudhury SR, Matin I, Franco OH. Long-term strategies to control COVID-19 in low and middle-income countries: an options overview of community-based, non-pharmacological interventions. European Journal of Epidemiology. 2020;35(8):743-8. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7354877/.

[191] Imai N, Hogan AB, Williams L, Cori A, Mangal TD, Winskill P, et al. Interpreting estimates of coronavirus disease 2019 (COVID-19) vaccine efficacy and effectiveness to inform simulation studies of vaccine impact: a systematic review. Wellcome Open Research. 2021 Jul;6:185. Available from: https://wellcomeopenresearch.org/articles/6-185/v1.

[192] Watson OJ, Walker P, Whittaker C, Winskill P, Charles G. squire: SEIR transmission model of COVID-19; 2020. Available from: https://mrc-ide.github.io/squire/index.html.

[193] Google. COVID-19 Community Mobility Report; 2022. Available from: https://www.google.com/covid19/mobility?hl=en.

[194] Coallition for Epidemic Preparedness Initiative. CEPI's 100 Days Mission; 2022. Available from: https://100days.cepi.net/.

[195] Moore S, Hill EM, Dyson L, Tildesley MJ, Keeling MJ. Retrospectively modeling the effects of increased global vaccine sharing on the COVID-19 pandemic. Nature Medicine. 2022;28(11):2416-23. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9671807/.

[196] Ye Y, Zhang Q, Wei X, Cao Z, Yuan HY, Zeng DD. Equitable access to COVID-19 vaccines makes a life-saving difference to all countries. Nature Human Behaviour. 2022;6(2):207-16. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8873023/.

[197] Li Q, Huang Y. Optimizing global COVID-19 vaccine allocation: An agent-based computational model of 148 countries. PLoS Computational Biology. 2022 Sep;18(9):e1010463. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9447912/.

[198] Bertram MY, Edejer TTT. Introduction to the Special Issue on "The World Health Organization Choosing Interventions That Are Cost-Effective (WHO-CHOICE) Update". International Journal of Health Policy and Management. 2021 Sep:1. Available from: http://www.ijhpm.com/article_4139.html.

[199] Pichon-Riviere A, Drummond M, Palacios A, Garcia-Marti S, Augustovski F. Determining the efficiency path to universal health coverage: cost-effectiveness thresholds for 174 countries based on growth in life expectancy and health expenditures. The Lancet Global Health. 2023 Jun;11(6):e833-42. Available from: https://linkinghub.elsevier.com/retrieve/pii/S2214109X23001626.

[200] Woods B, Revill P, Sculpher M, Claxton K. Country-Level Cost-Effectiveness Thresholds: Initial Estimates and the Need for Further Research. Value in Health. 2016 Dec;19(8):929-35. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1098301516000644.

[201] Mengstu S, Beyene Berha A. Safety and Efficacy of COVID-19 Vaccine in Africa: Systematic Review. Infection and Drug Resistance. 2023 May;16:3085-100. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10202204/.

[202] NHS England. NHS England Landmark moment as first NHS patient receives COVID-19 vaccination; 2020. Available from: https://www.england.nhs.uk/2020/12/landmark-moment-as-first-nhs-patient-receives-covid-19-vaccination/.

[203] World Health Organization. Strategy to achieve global COVID-19 vaccination by mid-2022; 2021. Available from: https://cdn.who.int/media/docs/default-source/immunization/covid-19/strategy-to-achieve-global-covid-19-vaccination-by-mid-2022.pdf?sfvrsn=5a68433c_5&download=true.

[204] World Bank. The World by Income and Region; 2023. Available from: https://datatopics.worldbank.org/world-development-indicators/the-world-by-income-and-region.html.

[205] WHO-CHOICE estimates of cost for inpatient and outpatient health service delivery; 2021. Available from: https://www.who.int/publications/m/item/who-choice-estimates-of-cost-for-inpatient-and-outpatient-health-service-delivery.

[206] Orangi S, Ojal J, Brand SP, Orlendo C, Kairu A, Aziza R, et al. Epidemiological impact and cost-effectiveness analysis of COVID-19 vaccination in Kenya. BMJ Global Health. 2022 Aug;7(8):e009430. Available from: https://gh.bmj.com/lookup/doi/10.1136/bmjgh-2022-009430.

[207] Barasa E, Kairu A, Ng'ang'a W, Maritim M, Were V, Akech S, et al. Examining unit costs for COVID-19 case management in Kenya. BMJ Global Health. 2021 Apr;6(4):e004159. Available from: https://gh.bmj.com/lookup/doi/10.1136/bmjgh-2020-004159.

[208] Schermbrucker K. Roadblocks at the last mile: What's. 2022.

[209] Institute for Health Metrics and Evaluations. Global Burden of Disease Study 2021 (GBD 2021) Disability Weights | GHDx; 2021. Available from: https://ghdx.healthdata.org/record/ihme-data/gbd-2021-disability-weights.

[210] Fan CY, Fann JCY, Yang MC, Lin TY, Chen HH, Liu JT, et al. Estimating global burden of COVID-19 with disability-adjusted life years and value of statistical life metrics. Journal of the Formosan Medical Association. 2021 Jun;120:S106-17. Available from: https://www.sciencedirect.com/science/article/pii/S0929664621002345.

[211] Sun L, Peng X, Li S, Huang Z. Cost-effectiveness thresholds or decision-making threshold: a novel perspective. Cost Effectiveness and Resource Allocation. 2023 Oct;21(1):72. Available from: https://doi.org/10.1186/s12962-023-00472-6.

[212] Pearson CAB, Bozzani F, Procter SR, Davies NG, Huda M, Jensen HT, et al. COVID-19 vaccination in Sindh Province, Pakistan: A modelling study of health impact and cost-effectiveness. PLOS Medicine. 2021 Oct;18(10):e1003815. Publisher: Public Library of Science. Available from: https://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.1003815.

[213] Africa CDC. Continental Framework on Strengthening Mortality Surveillance Systems in Africa; 2023.

[214] Pearson CAB, van Zandvoort K, Jarvis CI, Davies N, Checchi F, working group CC, et al.. Projections of COVID-19 epidemics in LMIC countries; 2020. Available from: https://cmmid.github.io/topics/covid19/LMIC-projection-reports.html.

[215] UK Government. UK COVID-19 vaccines delivery plan; 2021. Available from: https://www.gov.uk/government/publications/uk-covid-19-vaccines-delivery-plan/uk-covid-19-vaccines-delivery-plan.

[216] World Health Organization. COVAX: Working for global equitable access to COVID-19 vaccines; 2024. Available from: https://www.who.int/initiatives/act-accelerator/covax.

[217] Drummond MF, Sculpher M, Claxton K, Stoddard GL, Torrance GW. Methods for the Economic Evaluation of Health Care Programmes. 4th ed. Oxford: Oxford University Press; 2015. Available from: https://books.google.co.uk/books?id=lvWACgAAQBAJ.

[218] Postma MJ, Vries Rd, Welte R, Edmunds WJ. Health economic methodology illustrated with recent work on Chlamydia screening: the concept of extended dominance. Sexually Transmitted Infections. 2008 Apr;84(2):152-4. Publisher: The Medical Society for the Study of Venereal Disease Section: Health services research. Available from: https://sti.bmj.com/content/84/2/152.

[219] Yaesoubi R, Kunst N. Net Monetary Benefit Lines Augmented with Value-of-Information Measures to Present the Results of Economic Evaluations under Uncertainty. Medical Decision Making. 2024 Oct;44(7):770-86. Available from: https://journals.sagepub.com/doi/10.1177/0272989X241262343.

[220] Williams BA, Jones CH, Welch V, True JM. Outlook of pandemic preparedness in a post-COVID-19 world. npj Vaccines. 2023 Nov;8(1):1-12. Publisher: Nature Publishing Group. Available from: https://www.nature.com/articles/s41541-023-00773-0.

[221] Kaiser R, Coulombier D, Baldari M, Morgan D, Paquet C. What is epidemic intelligence, and how is it being improved in Europe? Weekly releases (1997–2007). 2006 Feb;11(5):2892. Publisher: European Centre for Disease Prevention and Control. Available from: https://www.eurosurveillance.org/content/10.2807/esw.11.05.02892-en.

[222] Module 1 Report 'In Brief' summary - The resilience and preparedness of the United Kingdom UK Covid-19 Inquiry Archives; 2024. Available from: https://covid19.public-inquiry.uk/documents/module-1-in-brief-report/.

[223] Sepulveda J, Faechem NS, Sanders K, Barker F. The United States' response to COVID-19: a case study of the first year. Institute for Global Health Sciences; 2021. Available from: https://globalhealthsciences.ucsf.edu/wp-content/uploads/2024/02/summary_us_case_study5.pdf.

[224] Bouyer F, Thiongane O, Hobeika A, Arsevska E, Binot A, Corrèges D, et al. Epidemic intelligence in Europe: a user needs perspective to foster innovation in digital health surveillance. BMC Public Health. 2024 Apr;24(1):973. Available from: https://doi.org/10.1186/s12889-024-18466-1.

[225] Perez-Guzman PN, Daunt A, Mukherjee S, Crook P, Forlano R, Kont MD, et al. Clinical Characteristics and Predictors of Outcomes of Hospitalized Patients With Coronavirus Disease 2019 in a Multiethnic London National Health Service Trust: A Retrospective Cohort Study. Clinical Infectious Diseases. 2020 Aug:ciaa1091. Available from: https://academic.oup.com/cid/advance-article/doi/10.1093/cid/ciaa1091/5885151.

[226] Krelle H, Tallack C. One year on: Three myths about COVID-19 that the data proved wrong; 2021. Available from: https://www.health.org.uk/reports-and-analysis/analysis/one-year-on-three-myths-about-covid-19-that-the-data-proved-wrong.

[227] Daunt A, Perez-Guzman PN, Liew F, Hauck K, Costelloe CE, Thursz MR, et al. Validity of the UK early access to medicines scheme criteria for Remdesivir use in patients with COVID-19 disease. Journal of Infection. 2020 Jun:S0163445320304308. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0163445320304308.

[228] Forlano R, Mullish BH, Mukherjee SK, Nathwani R, Harlow C, Crook P, et al. In-hospital mortality is associated with inflammatory response in NAFLD patients admitted for COVID-19. PloS One. 2020;15(10):e0240400.

[229] Vollmer MAC, Radhakrishnan S, Kont MD, Flaxman S, Bhatt S, Costelloe C, et al. The impact of the COVID-19 pandemic on patterns of attendance at emergency departments in two large London hospitals: an observational study. BMC health services research. 2021 Sep;21(1):1008.

[230] Middleton P, Perez-Guzman PN, Cheng A, Kumar N, Kont MD, Daunt A, et al. Characteristics and outcomes of clinically diagnosed RT-PCR swab negative COVID-19: a retrospective cohort study. Scientific Reports. 2021 Dec;11(1):2455. Available from: http://www.nature.com/articles/s41598-021-81930-0.

[231] Daunt A, Perez-Guzman PN, Cafferkey J, Manalan K, Cooke G, White PJ, et al. Factors associated with reattendance to emergency services following COVID-19 hospitalization. Journal of Medical Virology. 2021 Mar;93(3):1250-2.

[232] Ferreira FHG. Inequality and COVID-19. International Monetary Fund; 2021. Available from: https://www.imf.org/external/pubs/ft/fandd/2021/06/inequality-and-covid-19-ferreira.htm.

[233] Unwin HJT, Hillis S, Cluver L, Flaxman S, Goldman PS, Butchart A, et al. Global, regional, and national minimum estimates of children affected by COVID-19-associated orphanhood and caregiver death, by age and family circumstance up to Oct 31, 2021: an updated modelling study. The Lancet Child & Adolescent Health. 2022 Apr;6(4):249-59. Publisher: Elsevier. Available from: https://www.thelancet.com/journals/lanchi/article/PIIS2352-4642(22)00005-0/fulltext.

[234] Sherratt K, Abbott S, Meakin SR, Hellewell J, Munday JD, Bosse N, et al. Exploring surveillance data biases when estimating the reproduction number: with insights into subpopulation transmission of COVID-19 in England. Philosophical Transactions of the Royal Society of London Series B, Biological Sciences. 2021 Jul;376(1829):20200283.

[235] Brainard J, Lake IR, Morbey RA, Jones NR, Elliot AJ, Hunter PR. Comparison of surveillance systems for monitoring COVID-19 in England: a retrospective observational study. The Lancet Public Health. 2023 Nov;8(11):e850-8. Available from: https://linkinghub.elsevier.com/retrieve/pii/S2468266723002190.

[236] Ferretti L, Wymant C, Petrie J, Tsallis D, Kendall M, Ledda A, et al. Digital measurement of SARS-CoV-2 transmission risk from 7 million contacts. Nature. 2024 Feb;626(7997):145-50. Available from: https://www.nature.com/articles/s41586-023-06952-2.

[237] UK Health Security Agency. COVID-19 Infection Survey participants thanked for 'huge contribution' to pandemic response; 2023. Available from: https://www.gov.uk/government/news/covid-19-infection-survey-participants-thanked-for-huge-contribution-to-pandemic-res

[238] Brien P, Keep M. Public spending during the Covid-19 pandemic. UK Parliament House of Commons; 2023. Available from: https://researchbriefings.files.parliament.uk/documents/CBP-9309/CBP-9309.pdf.

[239] Dedhe AM, Chowkase AA, Gogate NV, Kshirsagar MM, Naphade R, Naphade A, et al. Conventional and frugal methods of estimating COVID-19-related excess deaths and undercount factors. Scientific Reports. 2024 May;14(1):10378.

[240] Torres-Rueda S, Sweeney S, Bozzani F, Naylor NR, Baker T, Pearson C, et al. Stark choices: exploring health sector costs of policy responses to COVID-19 in low-income and middle-income countries. BMJ Global Health. 2021 Dec;6(12):e005759. Publisher: BMJ Specialist Journals Section: Original research. Available from: https://gh.bmj.com/content/6/12/e005759.

[241] Yigezu A, Zewdie SA, Mirkuzie AH, Abera A, Hailu A, Agachew M, et al. Cost-analysis of COVID-19 sample collection, diagnosis, and contact tracing in low resource setting: The case of Addis Ababa, Ethiopia. PloS One. 2022;17(6):e0269458.

[242] World Health Organization. Pandemic Hub; 2025. Available from: https://pandemichub.who.int.

[243] Kraemer MUG, Tsui JLH, Chang SY, Lytras S, Khurana MP, Vanderslott S, et al. Artificial intelligence for modelling infectious disease epidemics. Nature. 2025 Feb;638(8051):623-35. Publisher: Nature Publishing Group. Available from: https://www.nature.com/articles/s41586-024-08564-w.

[244] Zambart. Zambart, Imperial College London and London School of Hygiene and Tropical Medicine Host 10 Day Workshop - Zambart; 2022. Available from: https://www.zambart.org.zm/2022/10/06/1797-2/,https://www.zambart.org.zm/2022/10/06/1797-2/.

[245] Han Q, Bragazzi N, Asgary A, Orbinski J, Wu J, Kong JD. Estimation of epidemiological parameters and ascertainment rate from early transmission of COVID-19 across Africa. Royal Society Open Science. 2023 Sep;10(9):230316. Publisher: Royal Society. Available from: https://royalsocietypublishing.org/doi/10.1098/rsos.230316.

[246] Johnson R, Carnalla M, Basto-Abreu A, Haw D, Morgenstern C, Doohan P, et al. Promoting healthy populations as a pandemic preparedness strategy: a simulation study from Mexico. Lancet Regional Health Americas. 2024 Feb;30:100682.

[247] Sherratt K, Gruson H, Grah R, Johnson H, Niehus R, Prasse B, et al. Predictive performance of multi-model ensemble forecasts of COVID-19 across European nations. eLife. 2023 Apr;12:e81916.

[248] Keeling MJ, Rohani P. Modeling Infectious Diseases in Humans and Animals. Princeton University Press; 2008. Google-Books-ID: G8enmS23c6YC.

[249] Lloyd-Smith JO, George D, Pepin KM, Pitzer VE, Pulliam JRC, Dobson AP, et al. Epidemic dynamics at the human-animal interface. Science (New York, NY). 2009 Dec;326(5958):1362-7. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3891603/.

[250] Engering A, Hogerwerf L, Slingenbergh J. Pathogen–host–environment interplay and disease emergence. Emerging Microbes & Infections. 2013 Feb;2(2):e5. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3630490/.

[251] Ewald J, Sieber P, Garde R, Lang SN, Schuster S, Ibrahim B. Trends in mathematical modeling of host–pathogen interactions. Cellular and Molecular Life Sciences: CMLS. 2019 Nov;77(3):467-80. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7010650/.

[252] Considering the host in host–pathogen interactions. Nature Microbiology. 2024 May;9(5):1149-9. Publisher: Nature Publishing Group. Available from: https://www.nature.com/articles/s41564-024-01708-8.

[253] Leung K, Leung GM, Wu JT. Modelling the adjustment of COVID-19 response and exit from dynamic zero-COVID in China. Epidemiology; 2022. Available from: http://medrxiv.org/lookup/doi/10.1101/2022.12.14.22283460.

[254] Yang B, Lin Y, Xiong W, Liu C, Gao H, Ho F, et al. Comparison of control and transmission of COVID-19 across epidemic waves in Hong Kong: an observational study. The Lancet Regional Health – Western Pacific. 2024 Feb;43. Publisher: Elsevier. Available from: https://www.thelancet.com/journals/lanwpc/article/PIIS2666-6065(23)00287-0/fulltext.

[255] Peng L, Huang X, Wang C, Xin H, Cowling BJ, Wu P, et al. Comparative epidemiology of outbreaks caused by SARS-CoV-2 Delta and Omicron variants in China. Epidemiology and Infection. 2024 Mar;152:e43.

[256] Bell D, Schultz Hansen K. Relative Burdens of the COVID-19, Malaria, Tuberculosis, and HIV/AIDS Epidemics in Sub-Saharan Africa. The American Journal of Tropical Medicine and Hygiene. 2021 Dec;105(6):1510-5. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8641365/.

[257] Bilgin GM, Munira SL, Lokuge K, Glass K. Mathematical modelling of the 100-day target for vaccine availability after the detection of a novel pathogen: A case study in Indonesia. Vaccine. 2024 Aug;42(21):126163. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0264410X24008260.

[258] Chanda-Kapata P, Ntoumi F, Kapata N, Lungu P, Mucheleng'anga LA, Chakaya J, et al. Tuberculosis, HIV/AIDS and Malaria Health Services in sub-Saharan Africa – A Situation Analysis of the Disruptions and Impact of the COVID-19 Pandemic. International Journal of Infectious Diseases. 2022 Nov;124:S41-6. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8949686/.

[259] Velavan TP, Meyer CG, Esen M, Kremsner PG, Ntoumi F. COVID-19 and syndemic challenges in 'Battling the Big Three': HIV, TB and malaria. International Journal of Infectious Dis-

eases. 2021 May;106:29-32. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7997707/.

[260] US Government. Withdrawing The United States From The World Health Organization. The White House; 2025. Available from: https://www.whitehouse.gov/presidential-actions/2025/01/withdrawing-the-united-states-from-the-worldhealth-organization/.

[261] Seddon S. What is USAID and why does Donald Trump want to end it? BBC News. 2025 Feb. Section: US & Canada. Available from: https://www.bbc.com/news/articles/clyezjwnx5ko.

[262] Buschschluter V, Hughes D. Argentina's Milei to pull out of WHO in move mirroring Trump's. BBC News. 2025 Feb. Section: Latin America. Available from: https://www.bbc.com/news/articles/c8975qp1n4qo.

[263] UK Parliament. International Development Committee Chair responds to aid cuts - Committees - UK Parliament; 2025. Available from: https://committees.parliament.uk/committee/98/international-development-committee/news/205430/international-development-committee-chair-responds-to-aid-cuts/.

[264] Coallition for Epidemic Preparedness Initiative. Priority pathogens | CEPI; 2024. Available from: https://cepi.net/priority-pathogens.

[265] Lauer SA, Grantz KH, Bi Q, Jones FK, Zheng Q, Meredith HR, et al. The Incubation Period of Coronavirus Disease 2019 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application. Annals of Internal Medicine. 2020 May;172(9):577-82. Available from: https://www.acpjournals.org/doi/10.7326/M20-0504.

[266] Gillespie DT. Approximate accelerated stochastic simulation of chemically reacting systems. The Journal of chemical physics. 2001;115(4):1716-33. Publisher: American Institute of Physics.

[267] UK Government. Regulatory approval of Pfizer/BioNTech vaccine for COVID-19; 2020. Available from: https://www.gov.uk/government/publications/regulatory-approval-of-pfizer-biontech-vaccine-for-covid-19.

[268] UK Government. Regulatory approval of COVID-19 Vaccine AstraZeneca; 2020. Available from: https://www.gov.uk/government/publications/regulatory-approval-of-covid-19-vaccine-astrazeneca.

[269] UK Government. Moderna vaccine becomes third COVID-19 vaccine approved by UK regulator; 2021. Available from: https://www.gov.uk/government/news/moderna-vaccine-becomes-third-covid-19-vaccine-approved-by-uk-regulator.

[270] Polack FP, Thomas SJ, Kitchin N, Absalon J, Gurtman A, Lockhart S, et al. Safety and efficacy of the BNT162b2 mRNA Covid-19 vaccine. New England Journal of Medicine. 2020;383(27):2603-15. Publisher: Mass Medical Soc.

[271] Harris RJ, Hall JA, Zaidi A, Andrews NJ, Dunbar JK, Dabrera G. Impact of vaccination on household transmission of SARS-COV-2 in England. medRxiv. 2021.

[272] Ismail S, Vilaplana T, Elgohari S, others. Effectiveness of BNT162b2 mRNA and ChAdOx1 adenovirus vector COVID-19 vaccines on risk of hospitalisation among older adults in England: an observational study using surveillance data; 2021.

[273] Hogan AB, Wu SL, Doohan P, Watson OJ, Winskill P, Charles G, et al. The value of vaccine booster doses to mitigate the global impact of the Omicron SARS-CoV-2 variant. medRxiv. 2022 Jan:2022.01.17.22269222. Publisher: Cold Spring Harbor Laboratory Press. Available from: https://www.medrxiv.org/content/10.1101/2022.01.17.22269222v1https://www.medrxiv.org/content/10.1101/2022.01.17.22269222v1.abstract.

[274] Khoury DS, Cromer D, Reynaldi A, Schlub TE, Wheatley AK, Juno JA, et al. Neutralizing antibody levels are highly predictive of immune protection from symptomatic SARS-CoV-2 infection. Nature Medicine 2021 27:7. 2021 May;27(7):1205-11. Publisher: Nature Publishing Group. Available from: https://www.nature.com/articles/s41591-021-01377-8.

[275] Andrews N, Tessier E, Stowe J, Gower C, Kirsebom F, Simmons R, et al. Vaccine effectiveness and duration of protection of Comirnaty, Vaxzevria and Spikevax against mild and severe COVID-19 in the UK. medRxiv. 2020. Publisher: Cold Spring Harbor Laboratory Press.

[276] Vasileiou E, Simpson CR, Shi T, Kerr S, Agrawal U, Akbari A, et al. Interim findings from first-dose mass COVID-19 vaccination roll-out and COVID-19 hospital admissions in Scotland: a national prospective cohort study. The Lancet. 2021. Publisher: Elsevier.

[277] Public Health England. Public Health England vaccine effectiveness report; 2021. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/971017/SP_PH__VE_report_20210317_CC_JLB.pdf.

[278] Hyams C, Marlow R, Maseko Z, King J, Ward L, Fox K, et al. Assessing the effectiveness of bnt162b2 and chadox1ncov-19 covid-19 vaccination in prevention of hospitalisations in elderly

and frail adults: A single centre test negative case-control study. SSRN Electronic Journal. 2021.

[279] Stowe J, Andrews N, Gower C, Gallagher E, Utsi L, Simmons R. Effectiveness of COVID-19 vaccines against hospital admission with the Delta (B. 1.617. 2) variant. Public Health England. 2021.

[280] Hall VJ, Foulkes S, Saei A, Andrews N, Oguti B, Charlett A, et al. Effectiveness of BNT162b2 mRNA vaccine against infection and COVID-19 vaccine coverage in healthcare workers in England, multicentre prospective cohort study (the SIREN Study). SSRN Electronic Journal. 2021.

[281] Public Health England. COVID-19 vaccine surveillance report: week 48; 2022. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1121345/vaccine-surveillance-report-week-48-2022.pdf.

[282] Sheikh A, McMenamin J, Taylor B, Robertson C. SARS-CoV-2 Delta VOC in Scotland: demographics, risk of hospital admission, and vaccine effectiveness. The Lancet. 2021. Publisher: Elsevier.

[283] Bernal JL, Andrews N, Gower C, Gallagher E, Simmons R, Thelwall S, et al. Effectiveness of Covid-19 vaccines against the B. 1.617. 2 (Delta) variant. New England Journal of Medicine. 2021;385(7):385-94. Publisher: Mass Medical Soc.

[284] Voysey M, Clemens SAC, Madhi SA, Weckx LY, Folegatti PM, Aley PK, et al. Single-dose administration and the influence of the timing of the booster dose on immunogenicity and efficacy of ChAdOx1 nCoV-19 (AZD1222) vaccine: a pooled analysis of four randomised trials. The Lancet. 2021;397(10277):881-91. Publisher: Elsevier.

[285] Nasreen S, Chung H, He S, Brown KA, Gubbay JB, Buchan SA, et al. Effectiveness of COVID-19 vaccines against symptomatic SARS-CoV-2 infection and severe outcomes with variants of concern in Ontario. Nature Microbiology 2022 7:3. 2022 Feb;7(3):379-85. Publisher: Nature Publishing Group. Available from: https://www.nature.com/articles/s41564-021-01053-0.

[286] Andrews N, Stowe J, Kirsebom F, Toffa S, Rickeard T, Gallagher E, et al. Covid-19 vaccine effectiveness against the Omicron (B. 1.1. 529) variant. New England Journal of Medicine. 2022;386(16):1532-46. Publisher: Mass Medical Soc.

[287] Institute for Government. Coronavirus vaccine rollout; 2021. Available from: https://www.instituteforgovernment.org.uk/explainers/coronavirus-vaccine-rollout.

[288] UK Government. COVID-19 Response - Living with COVID-19; 2022. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1056229/COVID-19_Response_-_Living_with_COVID-19.pdf.

[289] United Nations. World Population Prospects; 2023. Available from: https://population.un.org/wpp/Download/Standard/MostUsed/.

[290] Giandhari J, Pillay S, Wilkinson E, Tegally H, Sinayskiy I, Schuld M, et al. Early transmission of SARS-CoV-2 in South Africa: An epidemiological and phylogenetic report. International Journal of Infectious Diseases. 2021 Feb;103:234-41. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1201971220323225.

A England model technical appendix

This appendix provides further technical details of the model I used to study the SARS-CoV-2 pandemic in England, as previously published [1]. For convenience, I have included here the model diagram here but will not discuss the disaggregation of compartments and transitions through severity pathways, as these were presented in Section 2.2.2.

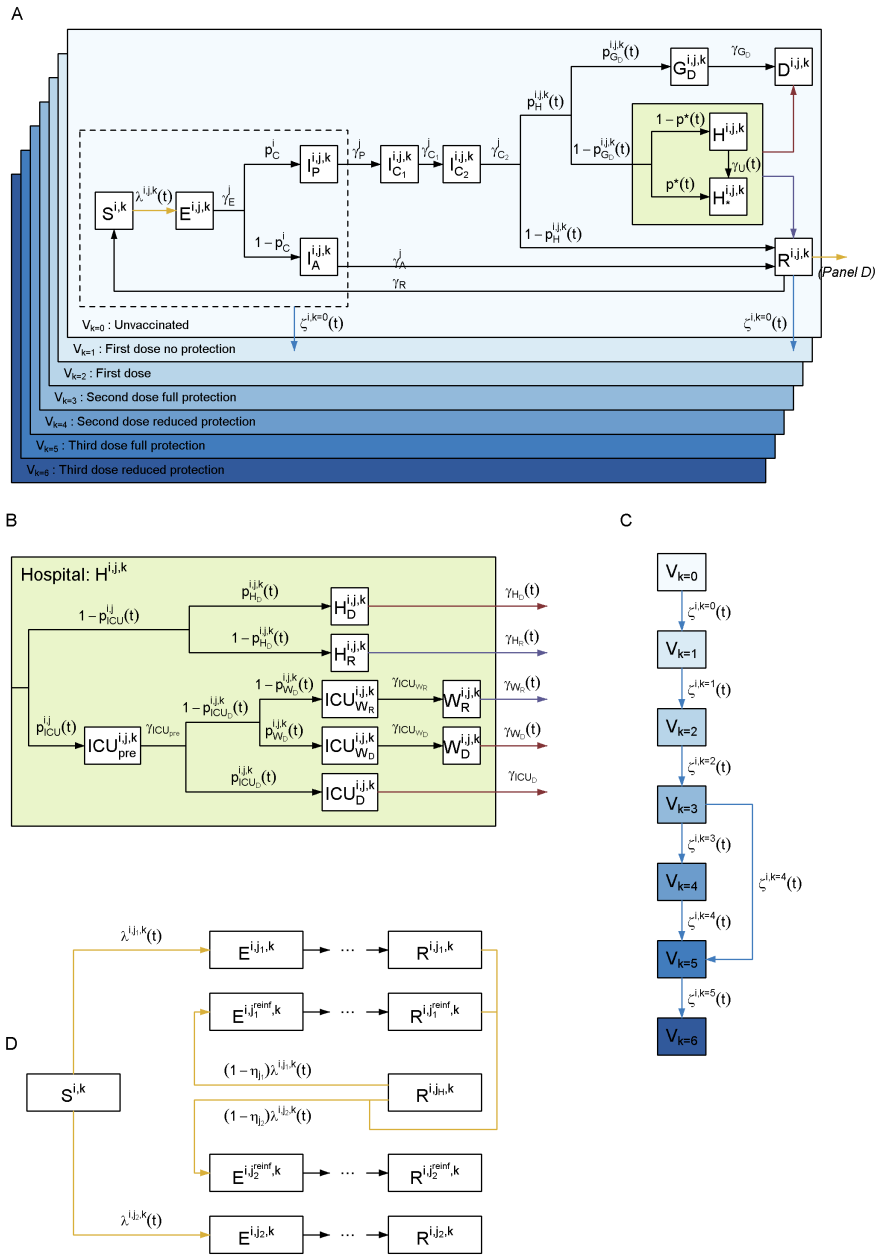


Figure A1: Model structure flow diagram with rates of transition between states. (A) Extended SEIR transmission model flow diagram overview. (B) Hospital flow diagram. (C) Vaccination flow diagram. (D) Multi-variant flow diagram. Definitions of state variables are presented in Table 2.1 and transition parameters in Table 2.2. Source: Perez-Guzman et. al. 2023 [1].

A.1 Model equations

Expectation model

To clearly illustrate the model dynamics, I describe a deterministic version of the model in differential equations (2)-(42), followed by the stochastic implementation used in the analysis. Full definitions of compartments and model parameters are set out in Tables 2.1 and 2.2.

Unless otherwise specified, $\sum_i$ refers to the sum across age groups (i.e. $i \in \{[0,5), \ldots, [75,80), [80+)\}$), $\sum_j$ refers to the sum across all combinations of co-circulating variants (j_1, j_2, j_1^{reinf}, j_2^{reinf}), and $\sum_k$ refers to the sum across all vaccination strata ($k \in \{0, 1, ..., 6\}$).

In the following model equations I use $\mathbb{1}_A$ as an indicator function, such that

$$\mathbb{1}_A(j) \coloneqq \begin{cases} 1 & \text{if } j \in A, \\ 0 & \text{if } j \notin A. \end{cases} \tag{1}$$

Also, to capture a non exponentially distributed duration of stay in certain compartments, some compartments are split in two distinct sub-compartments. For example, the exposed class, $E^{i,j,k}$, is modelled via two separate compartments, $E^{i,j,k,1}$ and $E^{i,j,k,2}$ (equations (3) and (4)). This is done to model durations of stay as an Erlang distribution instead (sum of independent exponential distributions) [265].

$$\begin{aligned} \frac{dS^{i,k}(t)}{dt} &= \zeta^{i,k-1}(t)S^{i,k-1}(t) + \mathbb{1}_{\{5\}}(k)\zeta^{i,4}(t)S^{i,3}(t) - \left(\zeta^{i,k}(t) + \mathbb{1}_{\{3\}}(k)\zeta^{i,4}(t) + \Lambda^{i,k}(t)\right) S^{i,k}(t) \\ &\quad - \sum_j \left(\delta^{i,j,k}(t) - \gamma_R R^{i,j,k}(t)\right) + \gamma_R R^{i,j_H,k} \end{aligned} \tag{2}$$

$$\begin{aligned} \frac{dE^{i,j,k,1}(t)}{dt} &= \mathbb{1}_{\{j_1,j_2\}}(j)\lambda^{i,j,k}(t)S^{i,k}(t) + \mathbb{1}_{\left\{j_1^{reinf},j_2^{reinf}\right\}}(j)\,(1-\eta_j)\,\lambda^{i,j,k}(t)R^{i,j_H,k}(t) \\ &\quad + \mathbb{1}_{\left\{j_2^{reinf}\right\}}(j)\,(1-\eta_j)\,\lambda^{i,j,k}(t)\left(R^{i,j_1,k}(t) + R^{i,j_1^{reinf},k}(t)\right) + \zeta^{i,k-1}(t)E^{i,j,k-1,1}(t) \\ &\quad + \mathbb{1}_{\{5\}}(k)\zeta^{i,4}(t)E^{i,j,3,1}(t) + \delta^{i,j,k}(t) - \left(\gamma_E^j + \zeta^{i,k}(t) + \mathbb{1}_{\{3\}}(k)\zeta^{i,4}(t)\right) E^{i,j,k,1}(t) \end{aligned} \tag{3}$$

$$\begin{aligned} \frac{dE^{i,j,k,2}(t)}{dt} &= \gamma_E^j E^{i,j,k,1}(t) + \zeta^{i,k-1}(t)E^{i,j,k-1,2}(t) + \mathbb{1}_{\{5\}}(k)\zeta^{i,4}(t)E^{i,j,3,2}(t) \\ &\quad - \left(\gamma_E^j + \zeta^{i,k}(t) + \mathbb{1}_{\{3\}}(k)\zeta^{i,4}(t)\right) E^{i,j,k,2}(t) \end{aligned} \tag{4}$$

$$\begin{aligned} \frac{dI_A^{i,j,k}(t)}{dt} &= \left(1 - p_C^i\right)\gamma_E^j E^{i,j,k,2}(t) + \zeta^{i,k-1}(t)I_A^{i,j,k-1}(t) + \mathbb{1}_{\{5\}}(k)\zeta^{i,4}(t)I_A^{i,j,3}(t) \\ &\quad - \left(\gamma_A^j + \zeta^{i,k}(t) + \mathbb{1}_{\{3\}}(k)\zeta^{i,4}(t)\right) I_A^{i,j,k}(t) \end{aligned} \tag{5}$$

$$\frac{dI_P^{i,j,k}(t)}{dt} = p_C^i \gamma_E^j E^{i,j,k,2}(t) + \zeta^{i,k-1}(t) I_P^{i,j,k-1}(t) + \mathbb{1}_{\{5\}}(k)\zeta^{i,4}(t) I_P^{i,j,3}(t) - \left(\gamma_P^j + \zeta^{i,k}(t) + \mathbb{1}_{\{3\}}(k)\zeta^{i,4}(t)\right) I_P^{i,j,k}(t) \tag{6}$$

$$\frac{dI_{C_1}^{i,j,k}(t)}{dt} = \gamma_P^j I_P^{i,j,k}(t) - \gamma_{C_1}^j I_{C_1}^{i,j,k}(t) \tag{7}$$

$$\frac{dI_{C_2}^{i,j,k}(t)}{dt} = \gamma_{C_1}^j I_{C_1}^{i,j,k}(t) - \gamma_{C_2}^j I_{C_2}^{i,j,k}(t) \tag{8}$$

$$\frac{dG_D^{i,j,k,1}(t)}{dt} = p_H^{i,j,k}(t) p_{G_D}^{i,j,k}(t) \gamma_{C_2}^j I_{C_2}^{i,j,k}(t) - \gamma_{G_D} G_D^{i,j,k,1}(t) \tag{9}$$

$$\frac{dG_D^{i,j,k,2}(t)}{dt} = \gamma_{G_D} G_D^{i,j,k,1}(t) - \gamma_{G_D} G_D^{i,j,k,2}(t) \tag{10}$$

$$\frac{dICU_{pre}^{i,j,k}(t)}{dt} = p_H^{i,j,k}(t)\left(1 - p_{G_D}^{i,j,k}(t)\right)(1 - p^*(t))\, p_{ICU}^{i,j}(t) \gamma_{C_2}^j I_{C_2}^{i,j,k}(t) - \left(\gamma_{ICU_{pre}} + \gamma_U(t)\right) ICU_{pre}^{i,j,k}(t) \tag{11}$$

$$\frac{dICU_{pre^*}^{i,j,k}(t)}{dt} = p_H^{i,j,k}(t)\left(1 - p_{G_D}^{i,j,k}(t)\right) p^*(t) p_{ICU}^{i,j}(t) \gamma_{C_2}^j I_{C_2}^{i,j,k}(t) - \gamma_{ICU_{pre}} ICU_{pre^*}^{i,j,k}(t) + \gamma_U(t) ICU_{pre}^{i,j,k}(t) \tag{12}$$

$$\frac{dICU_{W_R}^{i,j,k}(t)}{dt} = \left(1 - p_{ICU_D}^{i,j,k}(t)\right)\left(1 - p_{W_D}^{i,j,k}(t)\right) \gamma_{ICU_{pre}} ICU_{pre}^{i,j,k}(t) - \left(\gamma_{ICU_{W_R}} + \gamma_U(t)\right) ICU_{W_R}^{i,j,k}(t) \tag{13}$$

$$\frac{dICU_{W_{R^*}}^{i,j,k}(t)}{dt} = \left(1 - p_{ICU_D}^{i,j,k}(t)\right)\left(1 - p_{W_D}^{i,j,k}(t)\right) \gamma_{ICU_{pre}} ICU_{pre^*}^{i,j,k}(t) - \gamma_{ICU_{W_R}} ICU_{W_{R^*}}^{i,j,k}(t) + \gamma_U(t) ICU_{W_R}^{i,j,k}(t) \tag{14}$$

$$\frac{dICU_{W_D}^{i,j,k}(t)}{dt} = \left(1 - p_{ICU_D}^{i,j,k}(t)\right) p_{W_D}^{i,j,k}(t) \gamma_{ICU_{pre}} ICU_{pre}^{i,j,k}(t) - \left(\gamma_{ICU_{W_D}} + \gamma_U(t)\right) ICU_{W_D}^{i,j,k}(t) \tag{15}$$

$$\frac{dICU_{W_{D^*}}^{i,j,k}(t)}{dt} = \left(1 - p_{ICU_D}^{i,j,k}(t)\right) p_{W_D}^{i,j,k}(t) \gamma_{ICU_{pre}} ICU_{pre^*}^{i,j,k}(t) - \gamma_{ICU_{W_D}} ICU_{W_{D^*}}^{i,j,k}(t) + \gamma_U(t) ICU_{W_D}^{i,j,k}(t) \tag{16}$$

$$\frac{dICU_D^{i,j,k,1}(t)}{dt} = p_{ICU_D}^{i,j,k}(t) \gamma_{ICU_{pre}} ICU_{pre}^{i,j,k}(t) - \left(\gamma_{ICU_D} + \gamma_U(t)\right) ICU_D^{i,j,k,1}(t) \tag{17}$$

$$\frac{dICU_D^{i,j,k,2}(t)}{dt} = \gamma_{ICU_D} ICU_D^{i,j,k,1}(t) - \left(\gamma_{ICU_D} + \gamma_U(t)\right) ICU_D^{i,j,k,2}(t) \tag{18}$$

$$\frac{dICU_{D^*}^{i,j,k,1}(t)}{dt} = p_{ICU_D}^{i,j,k}(t) \gamma_{ICU_{pre}} ICU_{pre^*}^{i,j,k}(t) - \gamma_{ICU_D} ICU_{D^*}^{i,j,k,1}(t) + \gamma_U(t) ICU_D^{i,j,k,1}(t) \tag{19}$$

$$\frac{dICU_{D^*}^{i,j,k,2}(t)}{dt} = \gamma_{ICU_D} ICU_{D^*}^{i,j,k,1}(t) - \gamma_{ICU_D} ICU_{D^*}^{i,j,k,2}(t) + \gamma_U(t) ICU_D^{i,j,k,2}(t) \tag{20}$$

$$\frac{dW_R^{i,j,k,1}(t)}{dt} = \gamma_{ICU_{W_R}} ICU_{W_R}^{i,j,k}(t) - \left(\gamma_{W_R}(t) + \gamma_U(t)\right) W_R^{i,j,k,1}(t) \tag{21}$$

$$\frac{dW_R^{i,j,k,2}(t)}{dt} = \gamma_{W_R}(t) W_R^{i,j,k,1}(t) - \left(\gamma_{W_R}(t) + \gamma_U(t)\right) W_R^{i,j,k,2}(t) \tag{22}$$

$$\frac{dW_{R^*}^{i,j,k,1}(t)}{dt} = \gamma_{ICU_{W_R}} ICU_{W_{R^*}}^{i,j,k}(t) - \gamma_{W_R}(t) W_{R^*}^{i,j,k,1}(t) + \gamma_U(t) W_R^{i,j,k,1}(t) \tag{23}$$

$$\frac{dW_{R^*}^{i,j,k,2}(t)}{dt} = \gamma_{W_R}(t) W_{R^*}^{i,j,k,1}(t) - \gamma_{W_R}(t) W_{R^*}^{i,j,k,2}(t) + \gamma_U(t) W_R^{i,j,k,2}(t) \tag{24}$$

$$\frac{dW_D^{i,j,k}(t)}{dt} = \gamma_{ICU_{W_D}} ICU_{W_D}^{i,j,k}(t) - \left(\gamma_{W_D}(t) + \gamma_U(t)\right) W_D^{i,j,k}(t) \tag{25}$$

$$\frac{dW_{D^*}^{i,j,k}(t)}{dt} = \gamma_{ICU_{W_D}} ICU_{W_{D^*}}^{i,j,k}(t) - \gamma_{W_D}(t) W_{D^*}^{i,j,k}(t) + \gamma_U(t) W_D^{i,j,k}(t) \tag{26}$$

$$\begin{aligned}\frac{dH_R^{i,j,k}(t)}{dt} &= p_H^{i,j,k}(t)\left(1 - p_{G_D}^{i,j,k}(t)\right)\left(1 - p^*(t)\right)\left(1 - p_{ICU}^{i,j}(t)\right)\left(1 - p_{H_D}^{i,j,k}(t)\right)\gamma_{C_2}^j I_{C_2}^{i,j,k}(t) \\ &\quad - \left(\gamma_{H_R}(t) + \gamma_U(t)\right) H_R^{i,j,k}(t)\end{aligned} \tag{27}$$

$$\begin{aligned}\frac{dH_{R^*}^{i,j,k}(t)}{dt} &= p_H^{i,j,k}(t)\left(1 - p_{G_D}^{i,j,k}(t)\right) p^*(t)\left(1 - p_{ICU}^{i,j}(t)\right)\left(1 - p_{H_D}^{i,j,k}(t)\right)\gamma_{C_2}^j I_{C_2}^{i,j,k}(t) \\ &\quad + \gamma_U(t) H_R^{i,j,k}(t) - \gamma_{H_R}(t) H_{R^*}^{i,j,k}(t)\end{aligned} \tag{28}$$

$$\begin{aligned}\frac{dH_D^{i,j,k,1}(t)}{dt} &= p_H^{i,j,k}(t)\left(1 - p_{G_D}^{i,j,k}(t)\right)\left(1 - p^*(t)\right)\left(1 - p_{ICU}^{i,j}(t)\right) p_{H_D}^{i,j,k}(t)\gamma_{C_2}^j I_{C_2}^{i,j,k}(t) \\ &\quad - \left(\gamma_{H_D}(t) + \gamma_U(t)\right) H_D^{i,j,k,1}(t)\end{aligned} \tag{29}$$

$$\frac{dH_D^{i,j,k,2}(t)}{dt} = \gamma_{H_D}(t) H_D^{i,j,k,1}(t) - \left(\gamma_{H_D}(t) + \gamma_U(t)\right) H_D^{i,j,k,2}(t) \tag{30}$$

$$\begin{aligned}\frac{dH_{D^*}^{i,j,k,1}(t)}{dt} &= p_H^{i,j,k}(t)\left(1 - p_{G_D}^{i,j,k}(t)\right) p^*(t)\left(1 - p_{ICU}^{i,j}(t)\right)\ p_{H_D}^{i,j,k}(t)\gamma_{C_2}^j I_{C_2}^{i,j,k}(t) + \gamma_U(t) H_D^{i,j,k,1}(t) \\ &\quad - \gamma_{H_D}(t) H_{D^*}^{i,j,k,1}(t)\end{aligned} \tag{31}$$

$$\frac{dH_{D^*}^{i,j,k,2}(t)}{dt} = \gamma_{H_D}(t) H_{D^*}^{i,j,k,1}(t) - \gamma_{H_D}(t) H_{D^*}^{i,j,k,2}(t) + \gamma_U(t) H_D^{i,j,k,2}(t) \tag{32}$$

$$\begin{aligned}\frac{dR^{i,j,k}(t)}{dt} &= \mathbb{1}_{\left\{j_1, j_2, j_1^{reinf}, j_2^{reinf}\right\}}(j)\left(\gamma_A^j I_A^{i,j,k}(t) + \left(1 - p_H^{i,j,k}(t)\right)\gamma_{C_2}^j I_{C_2}^{i,j,k}(t) + \gamma_{H_R}(t)\left(H_R^{i,j,k}(t) + H_{R^*}^{i,j,k}(t)\right)\right. \\ &\qquad \left. + \gamma_{W_R}(t)\left(W_R^{i,j,k,2}(t) + W_{R^*}^{i,j,k,2}(t)\right)\right) \\ &\quad - \mathbb{1}_{\{j_H\}}(j)\left(1 - \eta_{j_1}\right)\lambda^{i,j_1,k}(t) R^{i,j,k}(t) - \mathbb{1}_{\left\{j_1, j_1^{reinf}, j_H\right\}}(j)\left(1 - \eta_{j_2}\right)\lambda^{i,j_2,k}(t) R^{i,j,k}(t) \\ &\quad + \zeta^{i,k-1}(t) R^{i,j,k-1}(t) + \mathbb{1}_{\{5\}}(k)\zeta^{i,4}(t) R^{i,j,3}(t) - \left(\gamma_R + \zeta^{i,k}(t) + \mathbb{1}_{\{3\}}(k)\zeta^{i,4}(t)\right) R^{i,j,k}(t)\end{aligned} \tag{33}$$

$$\frac{dT_{sero_{pre}^1}^i(t)}{dt} = -\gamma_{sero_{pre}} T_{sero_{pre}^1}^i(t) + \sum_j \sum_k \gamma_E^j E^{i,j,k,2}(t) \tag{34}$$

$$\frac{dT_{sero_{pos}^1}^i(t)}{dt} = p_{sero_{pos}} \gamma_{sero_{pre}} T_{sero_{pre}^1}^i(t) - \gamma_{sero_{pos}^1} T_{sero_{pos}^1}^i(t) \tag{35}$$

$$\frac{dT_{sero_{neg}^1}^i(t)}{dt} = \left(1 - p_{sero_{pos}}\right)\gamma_{sero_{pre}} T_{sero_{pre}^1}^i(t) + \gamma_{sero_{pos}^1} T_{sero_{pos}^1}^i(t) \tag{36}$$

$$\frac{dT_{sero_{pre}^2}^i(t)}{dt} = -\gamma_{sero_{pre}} T_{sero_{pre}^2}^i(t) + \sum_j \sum_k \gamma_E^j E^{i,j,k,2}(t) \tag{37}$$

$$\frac{dT_{sero_{pos}^2}^i(t)}{dt} = p_{sero_{pos}} \gamma_{sero_{pre}} T_{sero_{pre}^2}^i(t) - \gamma_{sero_{pos}^2} T_{sero_{pos}^2}^i(t) \tag{38}$$

$$\frac{dT_{sero_{neg}^2}^i(t)}{dt} = \left(1 - p_{sero_{pos}}\right)\gamma_{sero_{pre}} T_{sero_{pre}^2}^i(t) + \gamma_{sero_{pos}^2} T_{sero_{pos}^2}^i(t) \tag{39}$$

$$\frac{dT_{PCR_{pre}}^i(t)}{dt} = -\gamma_{PCR_{pre}} T_{PCR_{pre}}^i(t) + \sum_k \left(\lambda^{i,Alpha,k}(t) + \lambda^{i,Delta,k}(t)\right) S^{i,k}(t) \tag{40}$$

$$\frac{dT^i_{PCR_{pos}}(t)}{dt} = \gamma_{PCR_{pre}} T^i_{PCR_{pre}}(t) - \gamma_{PCR_{pos}} T^i_{PCR_{pos}}(t) \tag{41}$$

$$\frac{dT^i_{PCR_{neg}}(t)}{dt} = \gamma_{PCR_{pos}} T^i_{PCR_{pos}}(t). \tag{42}$$

Stochastic model

I used the tau-leap method [266] to create a stochastic, time-discretised version of the model described in equations (45) - (185), taking four update steps per day ($dt = 0.25$ days).

For each time step, the model iterated through the procedure described below. In the following, for transitions involving multiple onward compartments (e.g. transition from compartment E to compartments I_A or I_P or to the next vaccination strata within E), for conciseness, I omit the time dependency and use the following notation shorthand

$$\left(d^{i,j,k}_{E,I_A},\ d^{i,j,k}_{E,I_P},\ d^{i,j,k}_{E,v}\right) \sim \text{Multinom}\left(E^{i,j,k,2}(t),\ q^{i,j,k}_{E,I_A},\ q^{i,j,k}_{E,I_P},\ q^{i,j,k}_{E,v}\right) \tag{43}$$

instead of

$$\left(d^{i,j,k}_{E,I_A},\ d^{i,j,k}_{E,I_P},\ d^{i,j,k}_{E,v},\ d^{i,j,k}_{nomove}\right) \sim \text{Multinom}\left(E^{i,j,k,2}(t),\ q^{i,j,k}_{E,I_A},\ q^{i,j,k}_{E,I_P},\ q^{i,j,k}_{E,v},\ 1 - \sum_{x \in I_A, I_P, v} q^{i,j,k}_{E,x}\right) \tag{44}$$

where $d^{i,j,k}_{nomove}$ is a dummy variable counting the number of individuals remaining in compartment $E^{i,j,k,2}$. Using this convention, transition variables are drawn from the following distributions, with probabilities defined below:

$$q^{i,j_1,k}_{S,E} = \left(1 - e^{-\Lambda^{i,k}(t)dt}\right) \frac{\lambda^{i,j_1,k}(t)}{\Lambda^{i,k}(t)} \tag{45}$$

$$q^{i,j_2,k}_{S,E} = \left(1 - e^{-\Lambda^{i,k}(t)dt}\right) \frac{\lambda^{i,j_2,k}(t)}{\Lambda^{i,k}(t)} \tag{46}$$

$$\left(q^{i,k}_{S,v}, q^{i,k}_{S,v,2}\right) = \left(1 - e^{-\zeta^{i,k}(t)dt}, \mathbb{1}_{\{3\}}(k)\left(1 - e^{-\zeta^{i,4}(t)dt}\right) e^{-\zeta^{i,k}(t)dt}\right) \tag{47}$$

$$\left(d^{i,j_1,k}_{S,E},\ d^{i,j_2,k}_{S,E}\right) \sim \text{Multinom}\left(S^{i,k}(t),\ q^{i,j_1,k}_{S,E},\ q^{i,j_2,k}_{S,E}\right) \tag{48}$$

$$d^{i,j,k}_{seed} \sim \min\left(\text{Poisson}\left(\mathbb{1}_{\{j_1,j_2\}}(j)\hat{\delta}^{i,j,k}(t)dt\right),\ S^{i,k}(t) - d^{i,j_1,k}_{S,E} - d^{i,j_2,k}_{S,E}\right) \tag{49}$$

$$\left(d^{i,k}_{S,v}, d^{i,k}_{S,v,2}\right) \sim \text{Multinom}\left(S^{i,k}(t) - d^{i,j,k}_{seed} - d^{i,j_1,k}_{S,E} - d^{i,j_2,k}_{S,E},\ q^{i,k}_{S,v},\ q^{i,k}_{S,v,2}\right) \tag{50}$$

$$\left(q^{i,j,k}_{E,E},\ q^{i,j,k,1}_{E,v}\right) = \left(\ 1 - e^{-\gamma^j_E dt},\ e^{-\gamma^j_E dt}\left(1 - e^{-\zeta^{i,k}(t)dt}\right)\right) \tag{51}$$

$$q^{i,j,k,1}_{E,v,2} = \left(1 - q^{i,j,k,1}_{E,v}\right) \mathbb{1}_{\{3\}}(k)\left(1 - e^{-\zeta^{i,4}(t)dt}\right) \tag{52}$$

$$\left(d^{i,j,k}_{E,E},\ d^{i,j,k,1}_{E,v},\ d^{i,j,k,1}_{E,v,2}\right) \sim \text{Multinom}\left(E^{i,j,k,1}(t),\ q^{i,j,k}_{E,E},\ q^{i,j,k,1}_{E,v},\ q^{i,j,k,1}_{E,v,2}\right) \tag{53}$$

$$q^{i,j,k}_{E,I_A} = \left(1 - p^i_C\right)\left(1 - e^{-\gamma^j_E dt}\right) \tag{54}$$

$$q^{i,j,k}_{E,I_P} = p^i_C \left(1 - e^{-\gamma^j_E dt}\right) \tag{55}$$

$$q^{i,j,k,2}_{E,v} = e^{-\gamma^j_E dt} \left(1 - e^{-\zeta^{i,k}(t) dt}\right) \tag{56}$$

$$q^{i,j,k,2}_{E,v,2} = \left(1 - q^{i,j,k,2}_{E,v}\right) \mathbb{1}_{\{3\}}(k) \left(1 - e^{-\zeta^{i,4}(t) dt}\right) \tag{57}$$

$$\left(d^i_{E,I_A},\ d^i_{E,I_P},\ d^{i,j,k,2}_{E,v},\ d^{i,j,k,2}_{E,v,2}\right) \sim \text{Multinom}\left(E^{i,j,k,2}(t),\ q^{i,j,k}_{E,I_A},\ q^{i,j,k}_{E,I_P},\ q^{i,j,k,2}_{E,v},\ q^{i,j,k,2}_{E,v,2}\right) \tag{58}$$

$$\left(q^{i,j,k}_{I_A,R},\ q^{i,j,k}_{I_A,v}\right) = \left(1 - e^{-\gamma^j_A dt},\ e^{-\gamma^j_A dt} \left(1 - e^{-\zeta^{i,k}(t) dt}\right)\right) \tag{59}$$

$$q^{i,j,k}_{I_A,v,2} = \left(1 - q^{i,j,k}_{I_A,v}\right) \mathbb{1}_{\{3\}}(k) \left(1 - e^{-\zeta^{i,4}(t) dt}\right) \tag{60}$$

$$\left(d^{i,j,k}_{I_A,R},\ d^{i,j,k}_{I_A,v},\ d^{i,j,k}_{I_A,v,2}\right) \sim \text{Multinom}\left(I^i_A(t),\ q^{i,j,k}_{I_A,R},\ q^{i,j,k}_{I_A,v},\ q^{i,j,k}_{I_A,v,2}\right) \tag{61}$$

$$\left(q^{i,j,k}_{I_P,I_{C_1}},\ q^{i,j,k}_{I_P,v}\right) = \left(1 - e^{-\gamma^j_P dt},\ e^{-\gamma^j_P dt} \left(1 - e^{-\zeta^{i,k}(t) dt}\right)\right) \tag{62}$$

$$q^{i,j,k}_{I_P,v,2} = \left(1 - q^{i,j,k}_{I_P,v}\right) \mathbb{1}_{\{3\}}(k) \left(1 - e^{-\zeta^{i,4}(t) dt}\right) \tag{63}$$

$$\left(d^{i,j,k}_{I_P,I_{C_1}},\ d^{i,j,k}_{I_P,v},\ d^{i,j,k}_{I_P,v,2}\right) \sim \text{Multinom}\left(I^i_P(t),\ q^{i,j,k}_{I_P,I_{C_1}},\ q^{i,j,k}_{I_P,v},\ q^{i,j,k}_{I_P,v,2}\right) \tag{64}$$

$$d^{i,j,k}_{I_{C_1},I_{C_2}} \sim \text{Binom}\left(I^{i,j,k}_{C_1}(t),\ 1 - e^{-\gamma^j_{C_1} dt}\right) \tag{65}$$

$$q^{i,j,k}_{I_{C_2},G_D} = p^{i,j,k}_H(t) p^{i,j,k}_{G_D}(t) \left(1 - e^{-\gamma^j_{C_2} dt}\right) \tag{66}$$

$$q^{i,j,k}_{I_{C_2},R} = \left(1 - p^{i,j,k}_H(t)\right) \left(1 - e^{-\gamma^j_{C_2} dt}\right) \tag{67}$$

$$q^{i,j,k}_{I_{C_2},ICU_{pre}} = p^{i,j,k}_H(t) \left(1 - p^{i,j,k}_{G_D}(t)\right) \left(1 - p^*(t)\right) p^{i,j}_{ICU}(t) \left(1 - e^{-\gamma^j_{C_2} dt}\right) \tag{68}$$

$$q^{i,j,k}_{I_{C_2},ICU_{pre^*}} = p^{i,j,k}_H(t) \left(1 - p^{i,j,k}_{G_D}(t)\right) p^*(t) p^{i,j}_{ICU}(t) \left(1 - e^{-\gamma^j_{C_2} dt}\right) \tag{69}$$

$$q^{i,j,k}_{I_{C_2},H_R} = p^{i,j,k}_H(t) \left(1 - p^{i,j,k}_{G_D}(t)\right) \left(1 - p^*(t)\right) \left(1 - p^{i,j}_{ICU}(t)\right) \left(1 - p^{i,j,k}_{H_D}(t)\right) \left(1 - e^{-\gamma^j_{C_2} dt}\right) \tag{70}$$

$$q^{i,j,k}_{I_{C_2},H_{R^*}} = p^{i,j,k}_H(t) \left(1 - p^{i,j,k}_{G_D}(t)\right)\ p^*(t) \left(1 - p^{i,j}_{ICU}(t)\right) \left(1 - p^{i,j,k}_{H_D}(t)\right) \left(1 - e^{-\gamma^j_{C_2} dt}\right) \tag{71}$$

$$q^{i,j,k}_{I_{C_2},H_D} = p^{i,j,k}_H(t) \left(1 - p^{i,j,k}_{G_D}(t)\right) \left(1 - p^*(t)\right) \left(1 - p^{i,j}_{ICU}(t)\right) p^{i,j,k}_{H_D}(t) \left(1 - e^{-\gamma^j_{C_2} dt}\right) \tag{72}$$

$$q^{i,j,k}_{I_{C_2},H_{D^*}} = p^{i,j,k}_H(t) \left(1 - p^{i,j,k}_{G_D}(t)\right) p^*(t) \left(1 - p^{i,j}_{ICU}(t)\right) p^{i,j,k}_{H_D}(t) \left(1 - e^{-\gamma^j_{C_2} dt}\right) \tag{73}$$

$$\left(d^{i,j,k}_{I_{C_2},G_D}, \ldots, d^{i,j,k}_{I_{C_2},H_{D^*}}\right) \sim \text{Multinom}\left(I^{i,j,k}_{C_2}(t),\ q^{i,j,k}_{I_{C_2},G_D}, \ldots, q^{i,j,k}_{I_{C_2},H_{D^*}}\right) \tag{74}$$

$$d^{i,j,k}_{G_D,G_D} \sim \text{Binom}\left(G^{i,j,k,1}_D(t),\ 1 - e^{-\gamma_{G_D} dt}\right) \tag{75}$$

$$d^{i,j,k}_{G_D,D} \sim \text{Binom}\left(G^{i,j,k,2}_D(t),\ 1 - e^{-\gamma_{G_D} dt}\right) \tag{76}$$

$$q^{i,j,k}_{ICU_{pre},ICU_{W_R}} = \left(1 - p^{i,j,k}_{ICU_D}(t)\right) \left(1 - p^{i,j,k}_{W_D}(t)\right) \left(1 - e^{-\gamma_{ICU_{pre}} dt}\right) e^{-\gamma_U(t) dt} \tag{77}$$

$$q^{i,j,k}_{ICU_{pre},ICU_{W_{R^*}}} = \left(1 - p^{i,j,k}_{ICU_D}(t)\right) \left(1 - p^{i,j,k}_{W_D}(t)\right) \left(1 - e^{-\gamma_{ICU_{pre}} dt}\right) \left(1 - e^{-\gamma_U(t) dt}\right) \tag{78}$$

$$q^{i,j,k}_{ICU_{pre},ICU_{W_D}} = \left(1 - p^{i,j,k}_{ICU_D}(t)\right) p^{i,j,k}_{W_D}(t) \left(1 - e^{-\gamma_{ICU_{pre}} dt}\right) e^{-\gamma_U(t) dt} \tag{79}$$

$$q^{i,j,k}_{ICU_{pre},ICU_{W_{D^*}}} = \left(1 - p^{i,j,k}_{ICU_D}(t)\right) p^{i,j,k}_{W_D}(t) \left(1 - e^{-\gamma_{ICU_{pre}} dt}\right) \left(1 - e^{-\gamma_U(t) dt}\right) \tag{80}$$

$$q^{i,j,k}_{ICU_{pre},ICU_D} = p^{i,j,k}_{ICU_D}(t) \left(1 - e^{-\gamma_{ICU_{pre}} dt}\right) e^{-\gamma_U(t) dt} \tag{81}$$

$$q^{i,j,k}_{ICU_{pre},ICU_{D^*}} = p^{i,j,k}_{ICU_D}(t) \left(1 - e^{-\gamma_{ICU_{pre}} dt}\right) \left(1 - e^{-\gamma_U(t) dt}\right) \tag{82}$$

$$q^{i,j,k}_{ICU_{pre},ICU_{pre^*}} = e^{-\gamma_{ICU_{pre}} dt} \left(1 - e^{-\gamma_U(t) dt}\right) \tag{83}$$

$$\begin{aligned} &\left(d^{i,j,k}_{ICU_{pre},ICU_{W_R}}, \ \ldots, \ d^{i,j,k}_{ICU_{pre},ICU_{pre^*}}\right) \\ &\qquad \sim \text{Multinom}\left(ICU^{i,j,k}_{pre}(t), \ q^{i,j,k}_{ICU_{pre},ICU_{W_R}}, \ldots, q^{i,j,k}_{ICU_{pre},ICU_{pre^*}}\right) \end{aligned} \tag{84}$$

$$q^{i,j,k}_{ICU_{pre^*},ICU_{W_{R^*}}} = \left(1 - p^{i,j,k}_{ICU_D}(t)\right) \left(1 - p^{i,j,k}_{W_D}(t)\right) \left(1 - e^{-\gamma_{ICU_{pre}} dt}\right) \tag{85}$$

$$q^{i,j,k}_{ICU_{pre^*},ICU_{W_{D^*}}} = \left(1 - p^{i,j,k}_{ICU_D}(t)\right) p^{i,j,k}_{W_D}(t) \left(1 - e^{-\gamma_{ICU_{pre}} dt}\right) \tag{86}$$

$$q^{i,j,k}_{ICU_{pre^*},ICU_{D^*}} = p^{i,j,k}_{ICU_D}(t) \left(1 - e^{-\gamma_{ICU_{pre}} dt}\right) \tag{87}$$

$$\begin{aligned} &\left(d^{i,j,k}_{ICU_{pre^*},ICU_{W_{R^*}}}, \ \ldots, \ d^{i,j,k}_{ICU_{pre^*},ICU_{D^*}}\right) \\ &\qquad \sim \text{Multinom}\left(ICU^{i,j,k}_{pre^*}(t), \ q^{i,j,k}_{ICU_{pre^*},ICU_{W_{R^*}}}, \ldots, q^{i,j,k}_{ICU_{pre^*},ICU_{D^*}}\right) \end{aligned} \tag{88}$$

$$q^{i,j,k}_{H_D,H_D} = \left(1 - e^{-\gamma_{H_D}(t) dt}\right) e^{-\gamma_U(t) dt} \tag{89}$$

$$q^{i,j,k,1,1}_{H_D,H_{D^*}} = e^{-\gamma_{H_D}(t) dt} \left(1 - e^{-\gamma_U(t) dt}\right) \tag{90}$$

$$q^{i,j,k,1,2}_{H_D,H_{D^*}} = \left(1 - e^{-\gamma_{H_D}(t) dt}\right) \left(1 - e^{-\gamma_U(t) dt}\right) \tag{91}$$

$$\begin{aligned} &\left(d^{i,j,k}_{H_D,H_D}, \ \ d^{i,j,k,1,1}_{H_D,H_{D^*}}, \ d^{i,j,k,1,2}_{H_D,H_{D^*}}\right) \\ &\qquad \sim \text{Multinom}\left(H^{i,j,k,1}_D(t), \ q^{i,j,k}_{H_D,H_D}, \ q^{i,j,k,1,1}_{H_D,H_{D^*}}, \ q^{i,j,k,1,2}_{H_D,H_{D^*}}\right) \end{aligned} \tag{92}$$

$$d^{i,j,k}_{H_{D^*},H_{D^*}} \sim \text{Binom}\left(H^{i,j,k,1}_{D^*}(t), \ 1 - e^{-\gamma_{H_D}(t) dt}\right) \tag{93}$$

$$\left(d^{i,j,k}_{H_D,D}, \ d^{i,j,k,2,2}_{H_D,H_{D^*}}\right) \sim \text{Multinom}\left(H^{i,j,k,2}_D(t), \ 1 - e^{-\gamma_{H_D}(t) dt}, \ e^{-\gamma_{H_D}(t) dt} \left(1 - e^{-\gamma_U(t) dt}\right)\right) \tag{94}$$

$$d^{i,j,k}_{H_{D^*},D} \sim \text{Binom}\left(H^{i,j,k,2}_{D^*}(t), \ 1 - e^{-\gamma_{H_D}(t) dt}\right) \tag{95}$$

$$\left(d^{i,j,k}_{H_R,R}, \ d^{i,j,k}_{H_R,H_{R^*}}\right) \sim \text{Multinom}\left(H^{i,j,k}_R(t), \ 1 - e^{-\gamma_{H_R}(t) dt}, \ e^{-\gamma_{H_R}(t) dt} \left(1 - e^{-\gamma_U(t) dt}\right)\right) \tag{96}$$

$$d^{i,j,k}_{H_{R^*},R} \sim \text{Binom}\left(H^{i,j,k}_{R^*}(t), \ 1 - e^{-\gamma_{H_R}(t) dt}\right) \tag{97}$$

$$q^{i,j,k}_{ICU_{W_R},W_R} = \left(1 - e^{-\gamma_{ICU_{W_R}} dt}\right) e^{-\gamma_U(t) dt} \tag{98}$$

$$q^{i,j,k}_{ICU_{W_R},ICU_{W_{R^*}}} = e^{-\gamma_{ICU_{W_R}} dt} \left(1 - e^{-\gamma_U(t) dt}\right) \tag{99}$$

$$q^{i,j,k}_{ICU_{W_R},W_{R^*}} = \left(1 - e^{-\gamma_{ICU_{W_R}} dt}\right) \left(1 - e^{-\gamma_U(t) dt}\right) \tag{100}$$

$$\begin{aligned} &\left(d^{i,j,k}_{ICU_{W_R},W_R}, \ \ldots, d^{i,j,k}_{ICU_{W_R},W_{R^*}}\right) \\ &\qquad \sim \text{Multinom}\left(ICU^{i,j,k}_{W_R}(t), \ q^{i,j,k}_{ICU_{W_R},W_R}, \ldots, q^{i,j,k}_{ICU_{W_R},W_{R^*}}\right) \end{aligned} \tag{101}$$

$$d^{i,j,k}_{ICU_{W_{R^*}},W_{R^*}} \sim \text{Binom}\left(ICU^{i,j,k}_{W_{R^*}}(t), \ 1 - e^{-\gamma_{ICU_{W_R}} dt}\right) \tag{102}$$

$$q^{i,j,k}_{ICU_{W_D},W_D} = \left(1 - e^{-\gamma_{ICU_{W_D}}dt}\right) e^{-\gamma_U(t)dt} \tag{103}$$

$$q^{i,j,k}_{ICU_{W_D},ICU_{W_{D^*}}} = e^{-\gamma_{ICU_{W_D}}dt} \left(1 - e^{-\gamma_U(t)dt}\right) \tag{104}$$

$$q^{i,j,k}_{ICU_{W_D},W_{D^*}} = \left(1 - e^{-\gamma_{ICU_{W_D}}dt}\right) \left(1 - e^{-\gamma_U(t)dt}\right) \tag{105}$$

$$\begin{aligned}\left(d^{i,j,k}_{ICU_{W_D},W_D},\ \ldots, d^{i,j,k}_{ICU_{W_D},W_{D^*}}\right) \\ \sim \text{Multinom}\left(ICU^{i,j,k}_{W_D}(t),\ q^{i,j,k}_{ICU_{W_D},W_D}, \ldots, q^{i,j,k}_{ICU_{W_D},W_{D^*}}\right)\end{aligned} \tag{106}$$

$$d^{i,j,k}_{ICU_{W_{D^*}},W_{D^*}} \sim \text{Binom}\left(ICU^{i,j,k}_{W_{D^*}}(t),\ 1 - e^{-\gamma_{ICU_{W_D}}dt}\right) \tag{107}$$

$$q^{i,j,k}_{ICU_D,ICU_D} = \left(1 - e^{-\gamma_{ICU_D}dt}\right) e^{-\gamma_U(t)dt} \tag{108}$$

$$q^{i,j,k,1,1}_{ICU_D,ICU_{D^*}} = e^{-\gamma_{ICU_D}dt} \left(1 - e^{-\gamma_U(t)dt}\right) \tag{109}$$

$$q^{i,j,k,1,2}_{ICU_D,ICU_{D^*}} = \left(1 - e^{-\gamma_{ICU_D}dt}\right) \left(1 - e^{-\gamma_U(t)dt}\right) \tag{110}$$

$$\begin{aligned}\left(d^{i,j,k}_{ICU_D,ICU_D},\ \ d^{i,j,k,1,1}_{ICU_D,ICU_{D^*}},\ d^{i,j,k,1,2}_{ICU_D,ICU_{D^*}}\right) \\ \sim \text{Multinom}\left(ICU^{i,j,k,1}_{D}(t),\ q^{i,j,k}_{ICU_D,ICU_D},\ q^{i,j,k,1,1}_{ICU_D,ICU_{D^*}},\ q^{i,j,k,1,2}_{ICU_D,ICU_{D^*}}\right)\end{aligned} \tag{111}$$

$$d^{i,j,k}_{ICU_{D^*},ICU_{D^*}} \sim \text{Binom}\left(ICU^{i,j,k,1}_{D^*}(t),\ 1 - e^{-\gamma_{ICU_D}dt}\right) \tag{112}$$

$$\left(q^{i,j,k}_{ICU_D,D},\ q^{i,j,k,2,2}_{ICU_D,ICU_{D^*}}\right) = \left(\ 1 - e^{-\gamma_{ICU_D}dt},\ e^{-\gamma_{ICU_D}dt} \left(1 - e^{-\gamma_U(t)dt}\right)\right) \tag{113}$$

$$\left(d^{i,j,k}_{ICU_D,D},\ d^{i,j,k,2,2}_{ICU_D,ICU_{D^*}}\right) \sim \text{Multinom}\left(ICU^{i,j,k,2}_{D}(t),\ q^{i,j,k}_{ICU_D,D},\ q^{i,j,k,2,2}_{ICU_D,ICU_{D^*}}\right) \tag{114}$$

$$d^{i,j,k}_{ICU_{D^*},D} \sim \text{Binom}\left(ICU^{i,j,k,2}_{D^*}(t),\ 1 - e^{-\gamma_{ICU_D}dt}\right) \tag{115}$$

$$q^{i,j,k}_{W_R,W_R} = \left(1 - e^{-\gamma_{W_R}(t)dt}\right) e^{-\gamma_U(t)dt} \tag{116}$$

$$q^{i,j,k,1,1}_{W_R,W_{R^*}} = e^{-\gamma_{W_R}(t)dt} \left(1 - e^{-\gamma_U(t)dt}\right) \tag{117}$$

$$q^{i,j,k,1,2}_{W_R,W_{R^*}} = \left(1 - e^{-\gamma_{W_R}(t)dt}\right) \left(1 - e^{-\gamma_U(t)dt}\right) \tag{118}$$

$$\begin{aligned}\left(d^{i,j,k}_{W_R,W_R},\ \ d^{i,j,k,1,1}_{W_R,W_{R^*}},\ d^{i,j,k,1,2}_{W_R,W_{R^*}}\right) \\ \sim\ \text{Multinom}\left(W^{i,j,k,1}_{R}(t),\ q^{i,j,k}_{W_R,W_R},\ q^{i,j,k,1,1}_{W_R,W_{R^*}},\ q^{i,j,k,1,2}_{W_R,W_{R^*}}\right)\end{aligned} \tag{119}$$

$$d^{i,j,k}_{W_{R^*},W_{R^*}} \sim \text{Binom}\left(W^{i,j,k,1}_{R^*}(t),\ 1 - e^{-\gamma_{W_R}(t)dt}\right) \tag{120}$$

$$\left(q^{i,j,k}_{W_R,R},\ q^{i,j,k,2,2}_{W_R,W_{R^*}}\right) = \left(1 - e^{-\gamma_{W_R}(t)dt},\ e^{-\gamma_{W_R}(t)dt} \left(1 - e^{-\gamma_U(t)\ dt}\right)\right) \tag{121}$$

$$\left(d^{i,j,k}_{W_R,R},\ d^{i,j,k,2,2}_{W_R,W_{R^*}}\right) \sim \text{Multinom}\left(W^{i,j,k,2}_{R}(t),\ q^{i,j,k}_{W_R,R},\ q^{i,j,k,2,2}_{W_R,W_{R^*}}\right) \tag{122}$$

$$d^{i,j,k}_{W_{R^*},R} \sim \text{Binom}\left(W^{i,j,k,2}_{R^*}(t),\ 1 - e^{-\gamma_{W_R}(t)dt}\right) \tag{123}$$

$$\left(q^{i,j,k}_{W_D,D},\ q^{i,j,k}_{W_D,W_{D^*}}\right) = \left(1 - e^{-\gamma_{W_D}(t)dt},\ e^{-\gamma_{W_D}(t)dt} \left(1 - e^{-\gamma_U(t)dt}\right)\right) \tag{124}$$

$$\left(d^{i,j,k}_{W_D,D},\ d^{i,j,k}_{W_D,W_{D^*}}\right) \sim \text{Multinom}\left(W^{i,j,k}_{D}(t),\ q^{i,j,k}_{W_D,D},\ q^{i,j,k}_{W_D,W_{D^*}}\right) \tag{125}$$

$$d^{i,j,k}_{W_{D^*},D} \sim \text{Binom}\left(W^{i,j,k}_{D^*}(t),\ 1 - e^{-\gamma_{W_D}(t)dt}\right) \tag{126}$$

$$\gamma^{i,j,k,j_1}_{R,E} = \mathbb{1}_{\{j_H\}}(j)(1 - \eta_{j_1})\lambda^{i,j_1,k}(t) \tag{127}$$

$$\gamma_{R,E}^{i,j,k,j_2} = \mathbb{1}_{\left\{j_1, j_1^{reinf}, j_H\right\}}(j)(1-\eta_{j_2})\lambda^{i,j_2,k}(t) \tag{128}$$

$$q_{R,S}^{i,j,k} = \left(1 - e^{-\left(\gamma_R + \gamma_{R,E}^{i,j,k,j_1} + \gamma_{R,E}^{i,j,k,j_2}\right)dt}\right) \frac{\gamma_R}{\gamma_R + \gamma_{R,E}^{i,j,k,j_1} + \gamma_{R,E}^{i,j,k,j_2}} \tag{129}$$

$$q_{R,E}^{i,j,k,j_1} = \left(1 - e^{-\left(\gamma_R + \gamma_{R,E}^{i,j,k,j_1} + \gamma_{R,E}^{i,j,k,j_2}\right)dt}\right) \frac{\gamma_{R,E}^{i,j,k,j_1}}{\gamma_R + \gamma_{R,E}^{i,j,k,j_1} + \gamma_{R,E}^{i,j,k,j_2}} \tag{130}$$

$$q_{R,E}^{i,j,k,j_2} = \left(1 - e^{-\left(\gamma_R + \gamma_{R,E}^{i,j,k,j_1} + \gamma_{R,E}^{i,j,k,j_2}\right)dt}\right) \frac{\gamma_{R,E}^{i,j,k,j_2}}{\gamma_R + \gamma_{R,E}^{i,j,k,j_1} + \gamma_{R,E}^{i,j,k,j_2}} \tag{131}$$

$$q_{R,v}^{i,j,k} = e^{-\left(\gamma_R + \gamma_{R,E}^{i,j,k,j_1} + \gamma_{R,E}^{i,j,k,j_2}\right)dt}\left(1 - e^{-\zeta^{i,k}(t)dt}\right) \tag{132}$$

$$q_{R,v,2}^{i,j,k} = \left(1 - q_{R,v}^{i,j,k}\right)\mathbb{1}_{\{3\}}(k)\left(1 - e^{-\zeta^{i,4}(t)dt}\right) \tag{133}$$

$$\begin{aligned}\left(d_{R,S}^{i,j,k},\ d_{R,E}^{i,j,k,j_1},\ d_{R,E}^{i,j,k,j_2},\ \ d_{R,v}^{i,j,k},\ d_{R,v,2}^{i,j,k}\right)\\ \sim \text{Multinom}\left(R^{i,j,k}(t),\ q_{R,S}^{i,j,k},\ q_{R,E}^{i,j,k,j_1},\ q_{R,E}^{i,j,k,j_1},\ q_{R,v}^{i,j,k},\ q_{R,v,2}^{i,j,k}\right)\end{aligned} \tag{134}$$

$$q_{T_{sero_{pre}^1}, T_{sero_{pos}^1}}^{i} = p_{sero_{pos}}\left(1 - e^{-\gamma_{sero_{pre}}dt}\right) \tag{135}$$

$$q_{T_{sero_{pre}^1}, T_{sero_{neg}^1}}^{i} = \left(1 - p_{sero_{pos}}\right)\left(1 - e^{-\gamma_{sero_{pre}}dt}\right) \tag{136}$$

$$\begin{aligned}\left(d_{T_{sero_{pre}^1}, T_{sero_{pos}^1}}^{i},\ d_{T_{sero_{pre}^1}, T_{sero_{neg}^1}}^{i}\right)\\ \sim \text{Multinom}\left(T_{sero_{pre}^1}^{i}(t),\ q_{T_{sero_{pre}^1}, T_{sero_{pos}^1}}^{i},\ q_{T_{sero_{pre}^1}, T_{sero_{neg}^1}}^{i}\right)\end{aligned} \tag{137}$$

$$d_{T_{sero_{pos}^1}, T_{sero_{neg}^1}}^{i} \sim \text{Binom}\left(T_{sero_{pos}^1}^{i}(t),\ 1 - e^{-\gamma_{sero_{pos}^1}dt}\right) \tag{138}$$

$$q_{T_{sero_{pre}^2}, T_{sero_{pos}^2}}^{i} = p_{sero_{pos}}\left(1 - e^{-\gamma_{sero_{pre}}dt}\right) \tag{139}$$

$$q_{T_{sero_{pre}^2}, T_{sero_{neg}^2}}^{i} = \left(1 - p_{sero_{pos}}\right)\left(1 - e^{-\gamma_{sero_{pre}}dt}\right) \tag{140}$$

$$\begin{aligned}\left(d_{T_{sero_{pre}^2}, T_{sero_{pos}^2}}^{i},\ d_{T_{sero_{pre}^2}, T_{sero_{neg}^2}}^{i}\right)\\ \sim \text{Multinom}\left(T_{sero_{pre}^2}^{i}(t),\ q_{T_{sero_{pre}^2}, T_{sero_{pos}^2}}^{i},\ q_{T_{sero_{pre}^2}, T_{sero_{neg}^2}}^{i}\right)\end{aligned} \tag{141}$$

$$d_{T_{sero_{pos}^2}, T_{sero_{neg}^2}}^{i} \sim \text{Binom}\left(T_{sero_{pos}^2}^{i}(t),\ 1 - e^{-\gamma_{sero_{pos}^2}dt}\right) \tag{142}$$

$$d_{T_{PCR_{pre}}, T_{PCR_{pos}}}^{i} \sim \text{Binom}\left(T_{PCR_{pre}}^{i}(t),\ 1 - e^{-\gamma_{PCR_{pre}}dt}\right) \tag{143}$$

$$d_{T_{PCR_{pos}}, T_{PCR_{neg}}}^{i} \sim \text{Binom}\left(T_{PCR_{pos}}^{i}(t),\ 1 - e^{-\gamma_{PCR_{pos}}dt}\right) \tag{144}$$

Model compartments were then updated as follows (Note that $d_{S,E}^{i,j_1^{reinf},k} = d_{S,E}^{i,j_2^{reinf},k} = 0$):

$$\begin{aligned} S^{i,k}(t+dt) \coloneqq S^{i,k}(t) - d_{S,E}^{i,j_1,k} - d_{S,E}^{i,j_2,k} + d_{S,v}^{i,k-1} - d_{S,v}^{i,k} + \mathbb{1}_{\{5\}}(k) d_{S,v,2}^{i,3} - \mathbb{1}_{\{3\}} d_{S,v,2}^{i,3} - d_{seed}^{i,j,k} - d_{seed}^{i,j_2,k} \\ + \sum_j d_{R,S}^{i,j,k} + d_{R,S}^{i,j_H,k} \end{aligned} \tag{145}$$

$$\begin{aligned} E^{i,j,k,1}(t+dt) \coloneqq E^{i,j,k,1}(t) + d_{S,E}^{i,j,k} + \mathbb{1}_{\left\{j_1^{reinf}\right\}}(j) d_{R,E}^{i,j_H,k,j_1} + \mathbb{1}_{\left\{j_2^{reinf}\right\}}(j) \left(d_{R,E}^{i,j_H,k,j_1} + d_{R,E}^{i,j_1,k,j_1} + d_{R,E}^{i,j_1^{reinf},k,j_1} \right) \\ - d_{E,E}^{i,j,k} + d_{E,v}^{i,j,k-1,1} - d_{E,v}^{i,j,k,1} + \mathbb{1}_{\{5\}}(k) d_{E,v,2}^{i,j,3,1} - \mathbb{1}_{\{3\}} d_{E,v,2}^{i,j,3,1} + d_{seed}^{i,j,k} \end{aligned} \tag{146}$$

$$E^{i,j,k,2}(t+dt) \coloneqq E^{i,j,k,2}(t) + d_{E,E}^{i,j,k} - d_{E,I_A}^{i,j,k} - d_{E,I_P}^{i,j,k} + d_{E,v}^{i,j,k-1,2} - d_{E,v}^{i,j,k,2} + \mathbb{1}_{\{5\}}(k) d_{E,v,2}^{i,j,3,2} - \mathbb{1}_{\{3\}} d_{E,v,2}^{i,j,3,2} \tag{147}$$

$$I_A^{i,j,k}(t+dt) \coloneqq I_A^{i,j,k}(t) + d_{E,I_A}^{i,j,k} - d_{I_A,R}^{i,j,k} + d_{I_A,v}^{i,j,k-1} - d_{I_A,v}^{i,j,k} + \mathbb{1}_{\{5\}}(k) d_{I_A,v,2}^{i,j,3} - \mathbb{1}_{\{3\}} d_{I_A,v,2}^{i,j,3} \tag{148}$$

$$I_P^{i,j,k}(t+dt) \coloneqq I_P^{i,j,k}(t) + d_{E,I_P}^{i,j,k} - d_{I_P,I_{C_1}}^{i,j,k} + d_{I_P,v}^{i,j,k-1} - d_{I_P,v}^{i,j,k} + \mathbb{1}_{\{5\}}(k) d_{I_P,v,2}^{i,j,3} - \mathbb{1}_{\{3\}} d_{I_P,v,2}^{i,j,3} \tag{149}$$

$$I_{C_1}^{i,j,k}(t+dt) \coloneqq I_{C_1}^{i,j,k}(t) + d_{I_P,I_{C_1}}^{i,j,k} - d_{I_{C_1},I_{C_2}}^{i,j,k} \tag{150}$$

$$\begin{aligned} I_{C_2}^{i,j,k}(t+dt) \coloneqq I_{C_2}^{i,j,k}(t) + d_{I_{C_1},I_{C_2}}^{i,j,k} - d_{I_{C_2},G_D}^{i,j,k} - d_{I_{C_2},R}^{i,j,k} - d_{I_{C_2},ICU_{pre}}^{i,j,k} \\ - d_{I_{C_2},ICU_{pre^*}}^{i,j,k} - d_{I_{C_2},H_R}^{i,j,k} - d_{I_{C_2},H_{R^*}}^{i,j,k} - d_{I_{C_2},H_D}^{i,j,k} - d_{I_{C_2},H_{D^*}}^{i,j,k} \end{aligned} \tag{151}$$

$$G_D^{i,j,k,1}(t+dt) \coloneqq G_D^{i,j,k,1}(t) + d_{I_{C_2},G_D}^{i,j,k} - d_{G_D,G_D}^{i,j,k} \tag{152}$$

$$G_D^{i,j,k,2}(t+dt) \coloneqq G_D^{i,j,k,2}(t) + d_{G_D,G_D}^{i,j,k} - d_{G_D,D}^{i,j,k} \tag{153}$$

$$\begin{aligned} ICU_{pre}^{i,j,k}(t+dt) \coloneqq ICU_{pre}^{i,j,k}(t) + d_{I_{C_2},ICU_{pre}}^{i,j,k} - d_{ICU_{pre},ICU_{W_R}}^{i,j,k} - d_{ICU_{pre},ICU_{W_D}}^{i,j,k} \\ - d_{ICU_{pre},ICU_D}^{i,j,k} - d_{ICU_{pre},ICU_{pre^*}}^{i,j,k} - d_{ICU_{pre},ICU_{W_{R^*}}}^{i,j,k} - d_{ICU_{pre},ICU_{W_{D^*}}}^{i,j,k} \\ - d_{ICU_{pre},ICU_{D^*}}^{i,j,k} \end{aligned} \tag{154}$$

$$\begin{aligned} ICU_{pre^*}^{i,j,k}(t+dt) \coloneqq ICU_{pre^*}^{i,j,k}(t) + d_{I_{C_2},ICU_{pre^*}}^{i,j,k} - d_{ICU_{pre},ICU_{W_{D^*}}}^{i,j,k} - d_{ICU_{pre^*},ICU_{W_{R^*}}}^{i,j,k} \\ - d_{ICU_{pre^*},ICU_{D^*}}^{i,j,k} \end{aligned} \tag{155}$$

$$\begin{aligned} ICU_{W_R}^{i,j,k}(t+dt) \coloneqq ICU_{W_R}^{i,j,k}(t) + d_{ICU_{pre},ICU_{W_R}}^{i,j,k} - d_{ICU_{W_R},W_R}^{i,j,k} - d_{ICU_{W_R},ICU_{W_{R^*}}}^{i,j,k} \\ - d_{ICU_{W_R},W_{R^*}}^{i,j,k} \end{aligned} \tag{156}$$

$$\begin{aligned} ICU_{W_{R^*}}^{i,j,k}(t+dt) \coloneqq ICU_{W_{R^*}}^{i,j,k}(t) + d_{ICU_{pre^*},ICU_{W_{R^*}}}^{i,j,k} + d_{ICU_{W_R},ICU_{W_{R^*}}}^{i,j,k} + d_{ICU_{pre},ICU_{W_{R^*}}}^{i,j,k} \\ - d_{ICU_{W_{R^*}},W_{R^*}}^{i,j,k} \end{aligned} \tag{157}$$

$$\begin{aligned} ICU_{W_D}^{i,j,k}(t+dt) \coloneqq ICU_{W_D}^{i,j,k}(t) + d_{ICU_{pre},ICU_{W_D}}^{i,j,k} - d_{ICU_{W_D},W_D}^{i,j,k} - d_{ICU_{W_D},ICU_{W_{D^*}}}^{i,j,k} \\ - d_{ICU_{W_D},W_{D^*}}^{i,j,k} \end{aligned} \tag{158}$$

$$ICU_{W_{D^*}}^{i,j,k}(t+dt) := ICU_{W_{D^*}}^{i,j,k}(t) + d_{ICU_{pre^*},ICU_{W_{D^*}}}^{i,j,k} + d_{ICU_{W_D},ICU_{W_{D^*}}}^{i,j,k} + d_{ICU_{pre},ICU_{W_{D^*}}}^{i,j,k} - d_{ICU_{W_{D^*}},W_{D^*}}^{i,j,k} \tag{159}$$

$$ICU_{D}^{i,j,k,1}(t+dt) := ICU_{D}^{i,j,k,1}(t) + d_{ICU_{pre},ICU_D}^{i,j,k} - d_{ICU_D,ICU_D}^{i,j,k} - d_{ICU_D,ICU_{D^*}}^{i,j,k,1,1} - d_{ICU_D,ICU_{D^*}}^{i,j,k,1,2} \tag{160}$$

$$ICU_{D}^{i,j,k,2}(t+dt) := ICU_{D}^{i,j,k,2}(t) + d_{ICU_D,ICU_D}^{i,j,k} - d_{ICU_D,D}^{i,j,k} - d_{ICU_D,ICU_{D^*}}^{i,j,k,2,2} \tag{161}$$

$$ICU_{D^*}^{i,j,k,1}(t+dt) := ICU_{D^*}^{i,j,k,1}(t) + d_{ICU_{pre^*},ICU_{D^*}}^{i,j,k} + d_{ICU_D,ICU_{D^*}}^{i,j,k,1,1} + d_{ICU_{pre},ICU_{D^*}}^{i,j,k} - d_{ICU_{D^*},ICU_{D^*}}^{i,j,k} \tag{162}$$

$$ICU_{D^*}^{i,j,k,2}(t+dt) := ICU_{D^*}^{i,j,k,2}(t) + d_{ICU_{D^*},ICU_{D^*}}^{i,j,k} + d_{ICU_D,ICU_{D^*}}^{i,j,k,1,2} + d_{ICU_D,ICU_{D^*}}^{i,j,k,2,2} - d_{ICU_{D^*},D}^{i,j,k} \tag{163}$$

$$W_{R}^{i,j,k,1}(t+dt) := W_{R}^{i,j,k,1}(t) + d_{ICU_{W_R},W_R}^{i,j,k} - d_{W_R,W_R}^{i,j,k} - d_{W_R,W_{R^*}}^{i,j,k,1,1} - d_{W_R,W_{R^*}}^{i,j,k,1,2} \tag{164}$$

$$W_{R}^{i,j,k,2}(t+dt) := W_{R}^{i,j,k,2}(t) + d_{W_R,W_R}^{i,j,k} - d_{W_R,R}^{i,j,k} - d_{W_R,W_{R^*}}^{i,j,k,2,2} \tag{165}$$

$$W_{R^*}^{i,j,k,1}(t+dt) := W_{R^*}^{i,j,k,1}(t) + d_{ICU_{W_{R^*}},W_{R^*}}^{i,j,k} + d_{W_R,W_{R^*}}^{i,j,k,1,1} + d_{ICU_{W_R},W_{R^*}}^{i,j,k} - d_{W_{R^*},W_{R^*}}^{i,j,k} \tag{166}$$

$$W_{R^*}^{i,j,k,2}(t+dt) := W_{R^*}^{i,j,k,2}(t) + d_{W_{R^*},W_{R^*}}^{i,j,k} + d_{W_R,W_{R^*}}^{i,j,k,2,2} + d_{W_R,W_{R^*}}^{i,j,k,1,2} - d_{W_{R^*},R}^{i,j,k} \tag{167}$$

$$W_{D}^{i,j,k}(t+dt) := W_{D}^{i,j,k}(t) + d_{ICU_{W_D},W_D}^{i,j,k} - d_{W_D,D}^{i,j,k} - d_{W_D,W_{D^*}}^{i,j,k} \tag{168}$$

$$W_{D^*}^{i,j,k}(t+dt) := W_{D^*}^{i,j,k}(t) + d_{ICU_{W_{D^*}},W_{D^*}}^{i,j,k} + d_{W_D,W_{D^*}}^{i,j,k} + d_{ICU_{W_D},W_{D^*}}^{i,j,k} - d_{W_{D^*},D}^{i,j,k} \tag{169}$$

$$H_{D}^{i,j,k,1}(t+dt) := H_{D}^{i,j,k,1}(t) + d_{I_{C_2},H_D}^{i,j,k} - d_{H_D,H_D}^{i,j,k} - d_{H_D,H_{D^*}}^{i,j,k,1,1} - d_{H_D,H_{D^*}}^{i,j,k,1,2} \tag{170}$$

$$H_{D}^{i,j,k,2}(t+dt) := H_{D}^{i,j,k,2}(t) + d_{H_D,H_D}^{i,j,k} - d_{H_D,D}^{i,j,k} - d_{H_D,H_{D^*}}^{i,j,k,2,2} \tag{171}$$

$$H_{D^*}^{i,j,k,1}(t+dt) := H_{D^*}^{i,j,k,1}(t) + d_{I_{C_2},H_{D^*}}^{i,j,k} + d_{H_D,H_{D^*}}^{i,j,k,1,1} - d_{H_{D^*},H_{D^*}}^{i,j,k} \tag{172}$$

$$H_{D^*}^{i,j,k,2}(t+dt) := H_{D^*}^{i,j,k,2}(t) + d_{H_{D^*},H_{D^*}}^{i,j,k} + d_{H_D,H_{D^*}}^{i,j,k,2,2} + d_{H_D,H_{D^*}}^{i,j,k,1,2} - d_{H_{D^*},D}^{i,j,k} \tag{173}$$

$$H_{R}^{i,j,k}(t+dt) := H_{R}^{i,j,k}(t) + d_{I_{C_2},H_R}^{i,j,k} - d_{H_R,R}^{i,j,k} - d_{H_R,H_{R^*}}^{i,j,k} \tag{174}$$

$$H_{R^*}^{i}(t+dt) := H_{R^*}^{i,j,k}(t) + d_{I_{C_2},H_{R^*}}^{i,j,k} + d_{H_R,H_{R^*}}^{i,j,k} - d_{H_{R^*},R}^{i,j,k} \tag{175}$$

$$\begin{aligned} R^{i,j,k}(t+dt) := {} & R^{i,j,k}(t) \\ & + \mathbb{1}_{\left\{j_1,j_2,j_1^{reinf},j_2^{reinf}\right\}}(j)\left(d_{I_A,R}^{i,j,k} + d_{I_{C_2},R}^{i,j,k} + d_{H_R,R}^{i,j,k} + d_{H_{R^*},R}^{i,j,k} + d_{W_R,R}^{i,j,k} + d_{W_{R^*},R}^{i,j,k}\right) \\ & - d_{R,S}^{i,j,k} - \mathbb{1}_{\{j_H\}} d_{R,E}^{i,j,k,j_1} + \mathbb{1}_{\left\{j_H,j_1,j_1^{reinf}\right\}} d_{R,E}^{i,j,k,j_2} \\ & + d_{R,v}^{i,j,k-1} - d_{R,v}^{i,j,k} + \mathbb{1}_{\{5\}}(k) d_{R,v,2}^{i,j,3} - \mathbb{1}_{\{3\}} d_{R,v,2}^{i,j,3} \end{aligned} \tag{176}$$

$$T_{sero_{pre}^1}^{i}(t+dt) := T_{sero_{pre}^1}^{i}(t) - d_{T_{sero_{pre}^1},T_{sero_{pos}^1}}^{i} - d_{T_{sero_{pre}^1},T_{sero_{neg}^1}}^{i} + \sum_j \sum_k \left(d_{E,I_A}^{i,j,k} + d_{E,I_P}^{i,j,k}\right) \tag{177}$$

$$T_{sero_{pos}^1}^{i}(t+dt) := T_{sero_{pos}^1}^{i}(t) + d_{T_{sero_{pre}^1},T_{sero_{pos}^1}}^{i} - d_{T_{sero_{pos}^1},T_{sero_{neg}^1}}^{i} \tag{178}$$

$$T_{sero_{neg}^1}^{i}(t+dt) := T_{sero_{neg}^1}^{i}(t) + d_{T_{sero_{pre}^1},T_{sero_{neg}^1}}^{i} + d_{T_{sero_{pos}^1},T_{sero_{neg}^1}}^{i} \tag{179}$$

$$T_{sero_{pre}^2}^{i}(t+dt) := T_{sero_{pre}^2}^{i}(t) - d_{T_{sero_{pre}^2},T_{sero_{pos}^2}}^{i} - d_{T_{sero_{pre}^2},T_{sero_{neg}^2}}^{i} + \sum_j \sum_k \left(d_{E,I_A}^{i,j,k} + d_{E,I_P}^{i,j,k}\right) \tag{180}$$

$$T^i_{sero^2_{pos}}(t+dt) \coloneqq T^i_{sero^2_{pos}}(t) + d^i_{T_{sero^2_{pre}},T_{sero^2_{pos}}} - d^i_{T_{sero^2_{pos}},T_{sero^2_{neg}}} \tag{181}$$

$$T^i_{sero^2_{neg}}(t+dt) \coloneqq T^i_{sero^2_{neg}}(t) + d^i_{T_{sero^2_{pre}},T_{sero^2_{neg}}} + d^i_{T_{sero^2_{pos}},T_{sero^2_{neg}}} \tag{182}$$

$$\begin{aligned} T^i_{PCR_{pre}}(t+dt) \coloneqq T^i_{PCR_{pre}}(t) - d^i_{T_{PCR_{pre}},T_{PCR_{pos}}} + \\ \sum_j \sum_k \left(d^{i,j,k}_{S,E} + d^{i,j,k,j_1}_{R,E} + d^{i,j,k,j_2}_{R,E} \right) + \sum_k \left(d^{i,j_H,k,j_1}_{R,E} + d^{i,j_H,k,j_2}_{R,E} \right) \end{aligned} \tag{183}$$

$$T^i_{PCR_{pos}}(t+dt) \coloneqq T^i_{PCR_{pos}}(t) + d^i_{T_{PCR_{pre}},T_{PCR_{pos}}} - d^i_{T_{PCR_{pos}},T_{PCR_{neg}}} \tag{184}$$

$$T^i_{PCR_{neg}}(t+dt) \coloneqq T^i_{PCR_{neg}}(t) + d^i_{T_{PCR_{pos}},T_{PCR_{neg}}}. \tag{185}$$

$$\tag{186}$$

The seeding process (see Section 2.2.3) is handled within the discretisation to four update steps per day such that:

$$\hat{\delta}^{i,j,k}(t) = \begin{cases} \phi_j f_j(t) & \text{if } i = [15, 20),\ j \in \{Wildtype,\ Alpha,\ Delta,\ Omicron\},\ k = 0 \\ 0 & \text{otherwise,} \end{cases} \tag{187}$$

where

$$f_j(t) = \begin{cases} \left(\left\lceil \frac{t_j}{dt} \right\rceil - \frac{t_j}{dt} \right) & \text{if } t = dt \left\lfloor \frac{t_j}{dt} \right\rfloor \\ 1 & \text{if } dt \left\lfloor \frac{t_j}{dt} \right\rfloor < t < dt \left\lfloor \frac{t_j}{dt} \right\rfloor + \nu_j \\ \left(\frac{t_j}{dt} - \left\lfloor \frac{t_j}{dt} \right\rfloor \right) & \text{if } t = dt \left\lfloor \frac{t_j}{dt} \right\rfloor + \nu_j \\ 0 & \text{otherwise.} \end{cases} \tag{188}$$

where $\lfloor . \rfloor$ and $\lceil . \rceil$ denote the floor and ceiling functions respectively.

Parallel flows

In addition to compartments involved in the transmission dynamics and clinical progression, there are three parallel flows for fitting to testing data from surveys: (i) one for PCR testing and (ii) two for serology testing (Figure A2), with separate flows used for testing with the EuroImmun and Roche N assays.

Upon infection, an individual enters the PCR flow in a pre-positivity compartment ($T_{PCR_{pre}}$) before moving into the PCR positivity compartment ($T_{PCR_{pos}}$) and then ultimately into the PCR negativity compartment ($T_{PCR_{neg}}$). The duration of PCR positivity is modelled as an Erlang distribution, assuming a mean of 12 days (Table A1).

With regards to the serology parallel flows, the EuroImmun essay was used in England for testing NHS Blood and Transplant samples from the first wave onwards, while Roche N only started being used in November 2020. The latter distinguishes between seropositivity resulting from infection

or from vaccination, whereas EuroImmun does not. The serology flows was only designed to capture seroconversion resulting from infection and, therefore, I did not fit to samples using the EuroImmun assay from 15th January 2021 onwards, expecting that vaccination in the population would have had an impact on its results beyond this point.

After a pre-conversion period ($T_{sero^1_{pre}}$ for EuroImmun, $T_{sero^2_{pre}}$ for Roche N), individuals can seroconvert ($T_{sero^1_{pos}}$ for EuroImmun, $T_{sero^2_{pos}}$ for Roche N) or not ($T_{sero^1_{neg}}$ for EuroImmun, $T_{sero^2_{neg}}$ for Roche N). If they do seroconvert, they eventually serorevert to $T_{sero^1_{neg}}$ or $T_{sero^2_{neg}}$ accordingly.

Table A1: Definitions of model parallel flow compartments shown in Figure A2, with indices: i for age group ($i \in \{[0,5),[5,10),\ldots,[75-80),[80+)\}$). Durations (in days) have Erlang (k_X, γ_X) distributions, with mean k_X/γ_X. Source: Perez-Guzman et al. 2023 [1].

		Erlang duration parameters		
Compartment	Definition	k_X	γ_X	Mean
$T^i_{PCR_{pre}}(t)$	Pre-PCR positive	1	0.5	2
$T^i_{PCR_{pos}}(t)$	True PCR positive	1	0.083	12.05
$T^i_{PCR_{neg}}(t)$	True PCR negative (after infection)	-	-	-
$T^i_{sero^1_{pre}}(t)$	Pre-seropositive for EuroImmun assay	1	0.077	13
$T^i_{sero^1_{pos}}(t)$	True seropositive for EuroImmun assay	1	0.0025	400
$T^i_{sero^1_{neg}}(t)$	True seropositive (after infection) for EuroImmun assay	-	-	-
$T^i_{sero^2_{pre}}(t)$	Pre-seropositive for Roche N assay	1	0.077	13
$T^i_{sero^2_{pos}}(t)$	True seropositive for Roche N assay	1	0.001	1000
$T^i_{sero^2_{neg}}(t)$	True seropositive (after infection) for Roche N assay	-	-	-

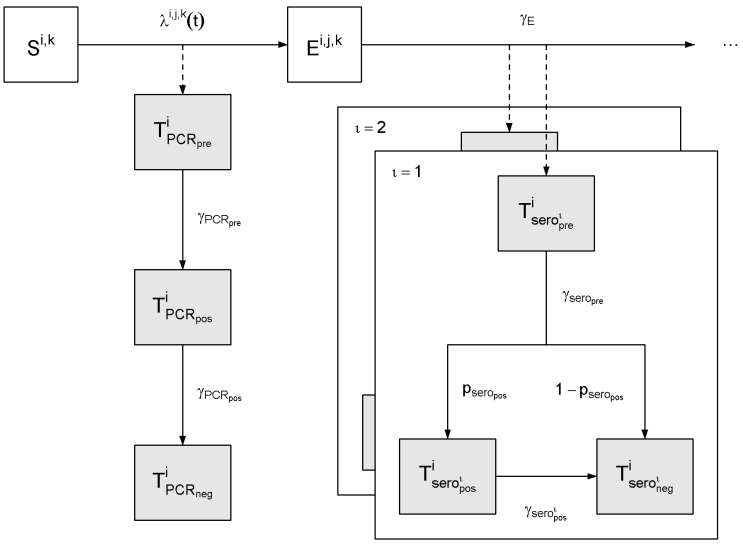


Figure A2: PCR positivity and seropositivity model structure flow diagram. Upon infection, an individual enters the pre-positive PCR compartment ($T_{PCR_{pre}}$) before moving into the PCR positivity compartment ($T_{PCR_{pos}}$) and then into the PCR negativity compartment ($T_{PCR_{neg}}$). After a pre-conversion period ($T_{sero_{pre}}$), individuals can seroconvert ($T_{sero_{pos}}$) or not ($T_{sero_{neg}}$); if they do seroconvert, they eventually serorevert to $T_{sero_{neg}}$. Source: Perez-Guzman et al. 2023 [1].

A.2 Vaccination

The model considers the original (i.e. non-multivalent) AZ, PF and Moderna vaccines, which were approved for use in England [267–269]. Given the analysis was capped to February 24, 2022, before the rollout of fourth doses (second "boosters") or multi-valent vaccines, only six distinct vaccination strata were simulated (V_k, for $k \in \{1, 2, 3, ..., 6\}$), representing the stages of VE detailed in Table A2 and illustrated in Figure A3.

Individuals move out of an unvaccinated (V_0) stratum at a rate determined by vaccine roll-out data and the prioritisation strategy adopted by the UK government. Only individuals who are not symptomatic and not hospitalised can receive a vaccine dose and/or transition across vaccination strata, i.e. only in compartments: susceptible (S), exposed (E), infected asymptomatic (I_A), infected pre-symptomatic (I_P) or recovered (R).

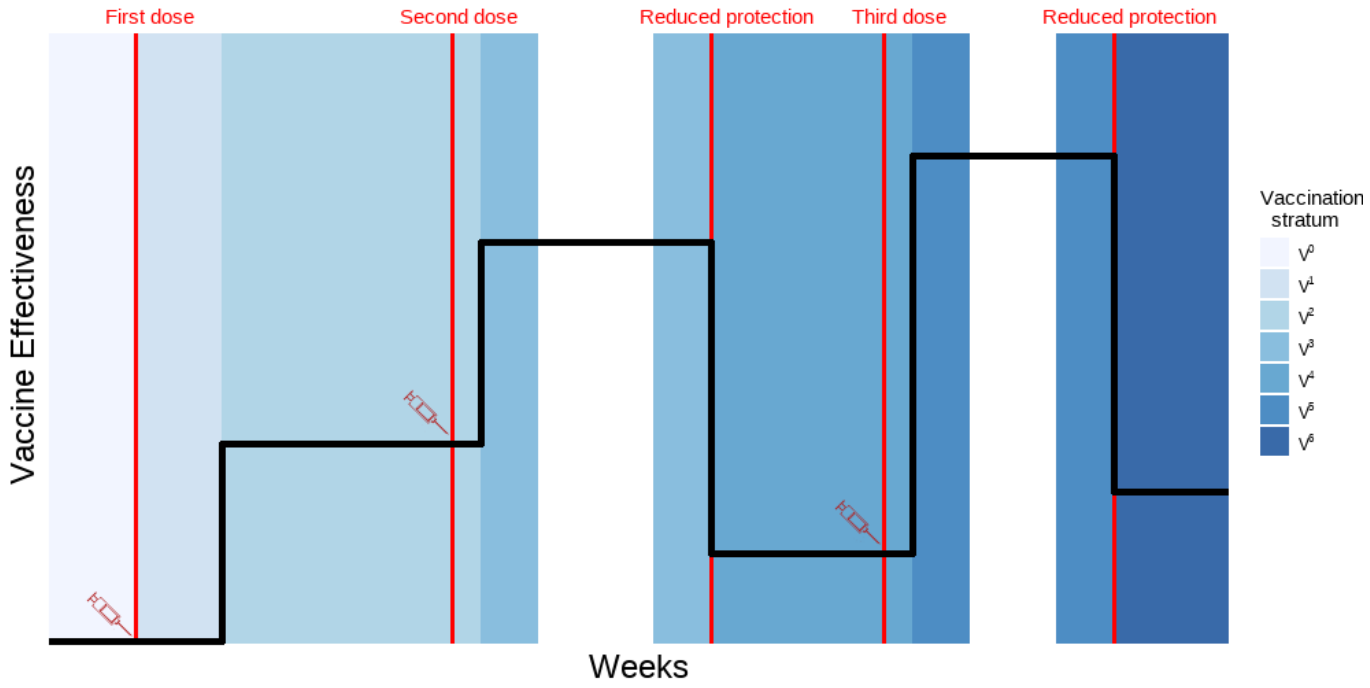


Figure A3: Vaccination strata duration and associated illustrative vaccine effectiveness. Red lines depict points at which a vaccine dose is administered. Assumed an average 24-weeks to waning. Y-axis illustrates changing vaccine effectiveness. Vaccination strata are defined in Table A2. Source: Perez-Guzman et al. 2023 [1].

As described in [32, 89], in addition to the exponentially-distributed delay of mean 21 days between the first dose and onset of VE, the model assumes a fixed 7-day delay between second and "booster" doses and onset of their VE (Figure A3). These assumptions were based on studies of immunogenicity after SARS-CoV-2 vaccination [100, 101].

Table A2: Vaccination strata considered for individuals, corresponding schedule, and vaccine effectiveness at each stage. Source: Perez-Guzman et al. 2023 [1].

Vaccination stratum	Number of doses	Vaccine effectiveness	Description	References
V_0	0	None	Non-vaccinated individuals.	Appendix A.2
V_1	1	None	Individuals who have had their first-dose but are in a delay period of average 3 weeks (exponential distribution) until effectiveness kicks in to transition out of V_1.	Appendix A.2
V_2	1	First dose effectiveness	Individuals with first-dose VE 3 weeks from date of first vaccination.	[100, 101]
V_3	2	Full second dose effectiveness	Individuals are fully protected 1 week from date of second vaccination.	[100, 101]
V_4	2	Reduced second dose effectiveness	Individuals with reduced second dose vaccine protection; transition from V_3 is randomly drawn from an exponential distribution with mean waning time of 24 weeks.	Appendix A.2
V_5	3	Booster dose effectiveness	Individuals are fully protected 1 week from date of third ("booster") vaccination.	[1]
V_6	3	Reduced "booster" dose effectiveness	Individuals with reduced "booster" dose vaccine protection; transition from V_5 is randomly drawn from an exponential distribution with mean waning time of 24 weeks.	Appendix A.2

Lastly, the model further assumes that individuals in V_3 (full second dose effectiveness) can

receive their booster dose and move to V_5 before their effectiveness has waned, thereby skipping V_4. The pool of individuals eligible for a booster dose is therefore those in V_3 and V_4, with no prioritisation of one over the other. All other strata movements are assumed to happen in a step-wise sequential order.

Vaccine effectiveness

The model assumed no differences in vaccine effectiveness by age, sex, or underlying health conditions [103,270]. Instead, it used values for vaccine effectiveness (VE) derived from vaccine efficacy measured in clinical trials and vaccine effectiveness studies (Table A3), mainly from UK based studies with a dosing schedule with an 12 week gap between doses. The efficacy was assumed to be the same against the Wildtype and Alpha variants, whereas variant-specific values were used for Delta and Omicron BA.1. Vaccine protection against symptomatic disease was assumed to be the same as that against infection [271].

Given the model explicitly simulations hospitalisations as well as cases that require hospitalisation remaining in the community Table A3 makes reference to vaccine effectiveness against *severe disease* and not *hospitalisation*. Vaccine effectiveness against severe disease, conditional on symptoms, acted on transition to both this compartment of individuals and those admitted to hospital.

Lastly, the model did not consider individual vaccines separately, but computed an age-dependent weighted mean of each vaccine's effectiveness. Weights for a given age group were the proportion of each vaccine type administered to that age group from vaccination linelist data in England as of the 24th of February 2022. Although, as mentioned above, the model assumed vaccine effectiveness did not vary by age, the weighted VE did vary across age groups given the proportion of each vaccine (PF, AZ or Mod) administered to each age group (from data)(Figure A4) and that VE varied between vaccines (Table A3).

Table A3: Vaccine effectiveness parameters for AstraZeneca (AZ), Pfizer (PF), and Moderna (Mod) by vaccine dose. "Infection" refers to vaccine effectiveness protecting an individual from being infected with SARS-CoV-2, whilst "transmission" refers to the vaccine effectiveness at preventing onward transmission by an infected individual. See Table A8 for VE assumptions for booster and booster waned sensitivity analysis for the Omicron variant. Source: Perez-Guzman et al. 2023 [1].

		Alpha		Delta		Omicron		Informed by
End point	Dose	AZ	PF/Mod	AZ	PF/Mod	AZ	PF/Mod	
Death	1	88%	89%	87%	89%	53%	53%	[102, 272]
	2 (Full protection)	99%	99%	99%	99%	97%	97%	[102, 272]
	2 (Waned protection)	83%	90%	82%	90%	56%	56%	[273–275]
	3 (Full protection)	99%	99%	99%	99%	96%	96%	Assume same as vs. severe disease.
	3 (Waned protection)					62%	62%	Assume same as vs. severe disease.
Severe disease	1	81%	89%	81%	89%	53%	53%	[276–278]
	2 (Full protection)	99%	99%	99%	99%	97%	97%	[279, 280]
	2 (Waned protection)	77%	90%	77%	90%	56%	56%	[273–275]
	3 (Full protection)	99%	99%	99%	99%	96%	96%	[281]
	3 (Waned protection)					62%	62%	[281]
Mild disease or infection	1	64%	79%	51%	51%	23%	37%	[103, 277, 279, 280, 282, 283]
	2 (Full protection)	92%	99%	87%	95%	41%	60%	[103, 280, 282–285]
	2 (Waned protection)	29%	77%	19%	49%	0%	0%	[273–275]
	3 (Full protection)	92%	92%	92%	92%	72%	74%	[286]
	3 (Waned protection)					0%	0%	[286]
Transmission	1	45%	45%	33%	33%	20%	20%	[102]
	2 (Full protection)	45%	45%	40%	40%	40%	40%	[102]
	2 (Waned protection)	40%	40%	19%	19%	0%	0%	[273–275]
	3 (Full protection)	40%	40%	40%	40%	40%	40%	[281]
	3 (Waned protection)					0%	0%	[281]

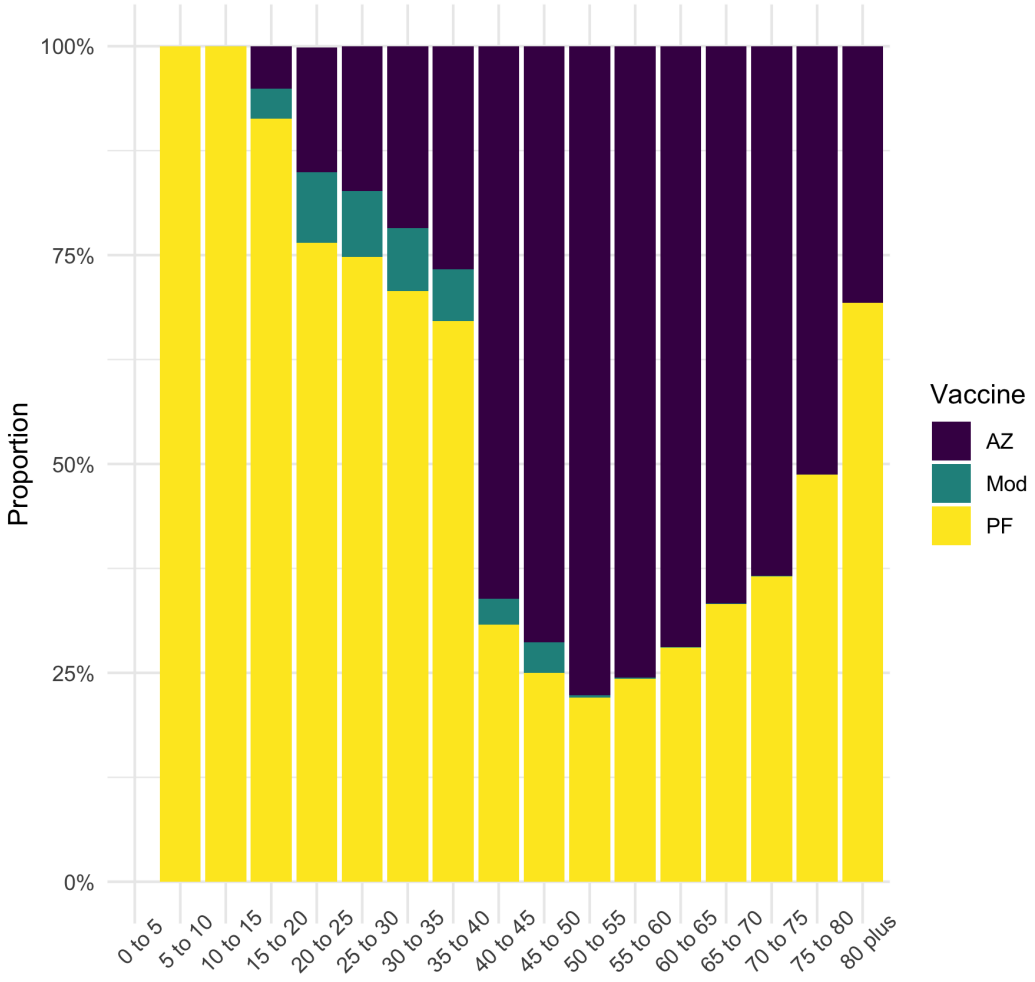


Figure A4: Proportion of each vaccine type: (Oxford-AstraZeneca (AZ), Pfizer-BioNTech (PF), Moderna (Mod)) dispensed to each five-year age band as of 24th February 2022. Data taken from UK Health Security Agency Immunisations database for vaccine delivery and ONS population estimates for each age group. Source: Perez-Guzman et al. 2023 [1].

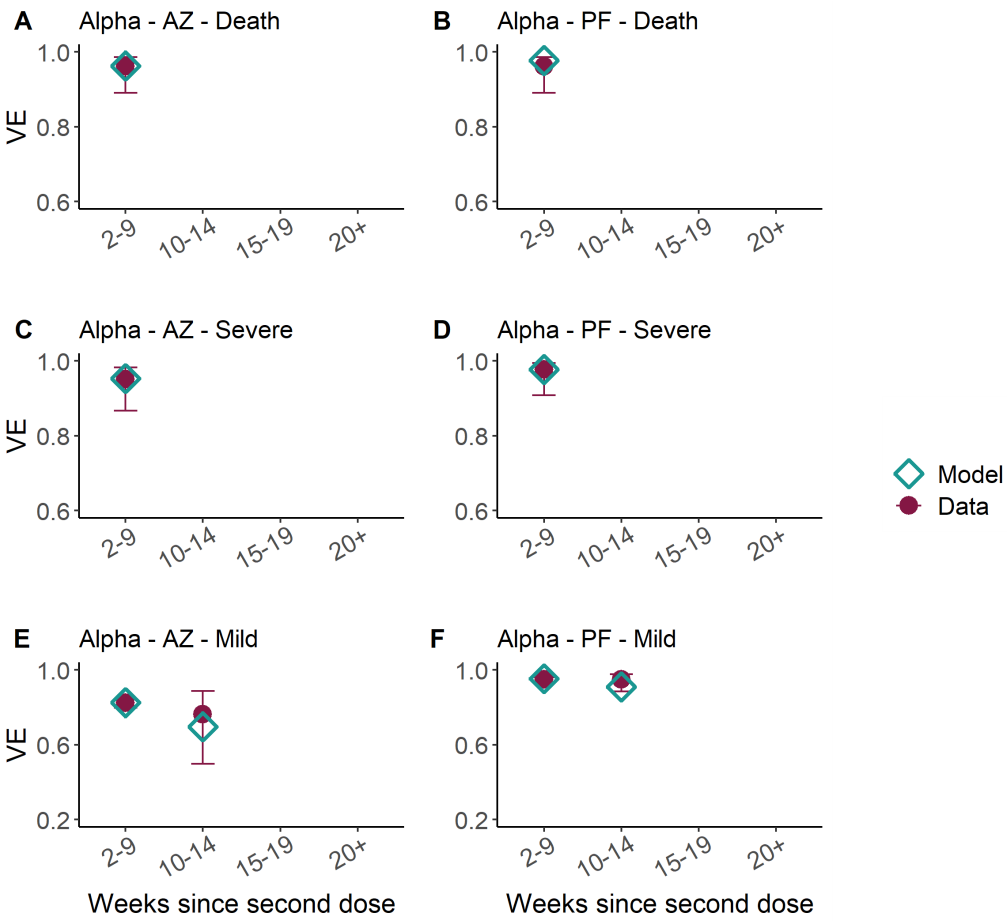


Figure A5: Vaccine effectiveness in weeks since second dose of AstraZeneca (AZ, left column) and Pfizer (PF, right column) vaccines against Alpha for death (top), severe disease, (middle) and mild disease/infection (bottom). Model assumed the same protection against infection and mild disease. Turquoise diamonds show model parameters, and the purple points estimates from data. Model assumed that the Moderna vaccine has the same VE as PF. Source: Perez-Guzman et al. 2023 [1].

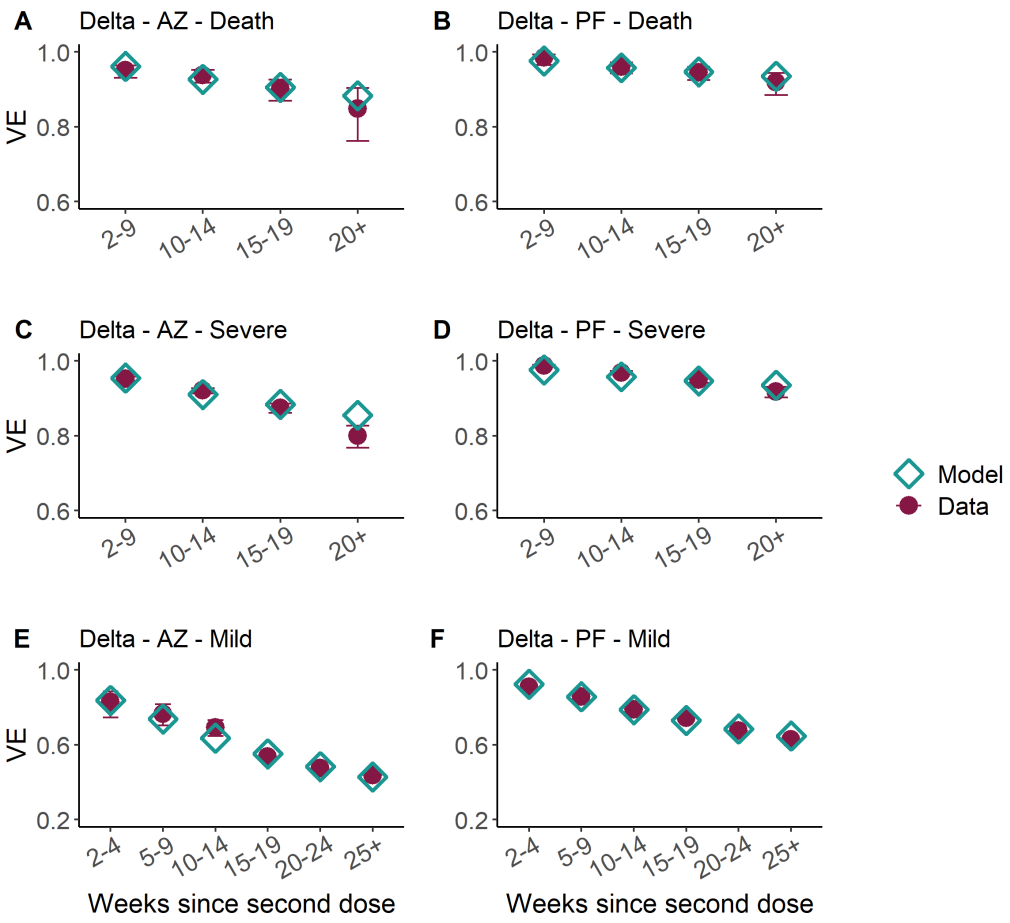


Figure A6: Vaccine effectiveness in weeks since second dose of AstraZeneca (AZ, left column) and Pfizer (PF, right column) vaccines against Delta for death (top), severe disease, (middle) and mild disease/infection (bottom). Model assumed the same protection against infection and mild disease. Turquoise diamonds show model parameters, and the purple points estimates from data. Model assumed that the Moderna vaccine has the same VE as PF. Source: Perez-Guzman et al. 2023 [1].

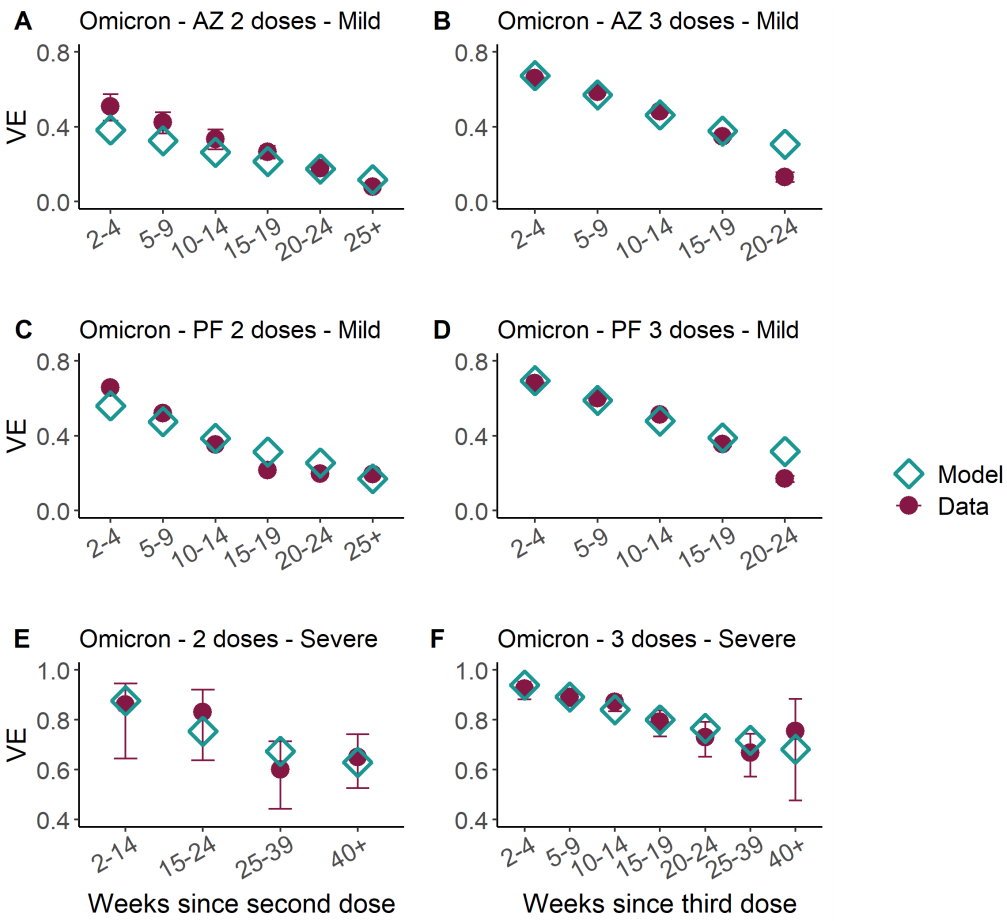


Figure A7: Vaccine effectiveness in weeks since second (left column) and third (booster, right column) dose of AstraZeneca (AZ, top row) and Pfizer (PF, middle row) vaccines against Omicron for mild disease (top and middle rows) and severe disease (bottom row). Model assumed the same protection against infection and mild disease. Turquoise diamonds show model parameters, and the purple points estimates from data. Model assumed that the Moderna vaccine has the same VE as PF. Source: Perez-Guzman et al. 2023 [1].

Waning of vaccine-induced immunity

The model assumed that, upon receiving a second or booster dose, individuals firstly progress to the V_3 or V_5 vaccination strata, respectively, granting full vaccine protection. The duration spent in these strata was stochastic, with an individual's rate of progression drawn from an exponential distribution with a mean duration of 24 weeks. It should be noted this was a simplifying assumption, as real-life studies showed that VE waning occurs gradually and continuously over time [273, 275].

The VE protection values from second and booster doses (V_3 and V_5 compartments), as well as reduced protection VE values (V_4 and V_6 compartments) were fitted (weighted least squares, using 1/width of the associated 95% CIs reported with the data as the weights) to data by vaccine type and outcome (Figures A5 to A7) [275, 286]. Due to the limited data available for waning effectiveness against the Alpha variant, a simplifying assumptions was made to fix the log-odds difference between full second dose VE and reduced second dose VE to be the same for each variant. That is, assuming the same proportion of protection drop-off from their initial full second dose VE. Mathematically, if $V(x)$ represents the VE of vaccine/health outcome x, then the log odds was defined as $L(x) = \log\left(\frac{V(x)}{1-V(x)}\right)$, and therefore for each health outcome

$$\begin{aligned} &L(\text{Reduced second dose VE vs. Delta}) - L(\text{Full second dose VE vs. Delta}) = \\ &L(\text{Reduced second dose VE vs. Alpha}) - L(\text{Full second dose VE vs.Alpha}). \end{aligned} \tag{189}$$

For first dose VE estimates, the model assumed the 16+ age group Delta VE values as presented in [275]. It should be noted that, due to prioritised groups being vaccinated during the Alpha wave, the reported VEs are unlikely to be generalisable to the general population. Thus, it was assumed that

$$\begin{aligned} &L(\text{First dose VE vs. Alpha}) - L(\text{Full second dose VE vs. Alpha}) = \\ &L(\text{First dose VE vs. Delta}) - L(\text{Full second dose VE vs.Delta}). \end{aligned} \tag{190}$$

Lastly, there was no VE data available for certain outcome-variant-dose combinations. Namely, booster dose VE vs. the Alpha variant, and booster dose VE against death for the Delta and Omicron variants. In these cases, values were fixed following the logical requirements that VE against more severe outcomes must be greater than that against less severe outcomes, and assuming that VE vs. the Alpha variant must be greater than the associated protection against the Delta variant.

Conditional dependencies of vaccine-immunity

As presented in Section 2.2.4, the model's severity transition pathways are a cascade of conditional probabilities (e.g. hospitalisation conditional on symptomatic disease), yet the VE values in Table A3 are unconditional on previous transitions to other severe outcomes. Thus, the latter were converted to conditional VE values for the model, as detailed in Table A4.

Table A4: Conditional vaccine effectiveness values modelled. Source: Perez-Guzman et al. 2023 [1].

VE vs.	Symbol / Calculation
Infection	e_{inf}
Severe disease	e_{SD}
Death	e_{death}
Severe disease given infection	$e_{SD\|inf} = \frac{e_{SD} - e_{inf}}{(1 - e_{inf})}$
Death given severe disease	$e_{death\|SD} = \frac{e_{death} - e_{SD}}{(1 - e_{inf})(1 - e_{SD\|inf})}$

Vaccine roll-out

Vaccination roll-out in the model followed the ordered priority set in the UK by the Joint Committee on Vaccination and Immunisation, whereby care home residents and care home workers were first, followed by other adult groups in decreasing age and clinical vulnerability [287, 288]. The model assumed vaccine doses were delivered in England as reported in age-stratified data [161].

A.3 Sensitivity analyses

In previous publications from the ICL-RTM team [32, 89], the effect of uncertainty around VE, and the time to waning of natural and vaccine-induced immunity for the Alpha and Delta variants had been explored, so these are not repeated herein. Here, to explore uncertainties around fixed model parameters during the period fitting to data from the Omicron, I performed the following sensitivity analyses:

1. *Cross-immunity against infection.*
2. *Cross-immunity against hospitalisation.*
3. *Cross-immunity against death.*
4. *Efficacy of boosters against severe disease and death.*
5. *Mean serial interval duration.*

Table A5 summarises cross-immunity sensitivity analyses explored, and table A8 the booster VE against Omicron. For vaccine efficacy, the central model parameters were as defined in Appendix A.2.

Table A5: Summary of sensitivity analyses explored. Source: Perez-Guzman et al. 2023 [1].

	Analysis	*Description*	*Source*
1	*Cross-immunity against infection*	Assumed a 25% protection against infection if infected with Omicron, given natural immunity (recovery from infection with Delta or other historic variant). Higher and a lower value for this assumption explored, at 35% and 20%, respectively.	Assumed
2	*Cross-immunity against hospitalisation*	Assumed a 55% protection against hospitalisation if infected with Omicron, given natural immunity (recovery from infection with Delta or other historic variant). Higher and a lower value for this assumption explored, at 63% and 48%, respectively.	[174]
3	*Cross-immunity against death*	Assumed an 18% protection against death if infected with Omicron, given natural immunity (recovery from infection with Delta or other historic variant). Higher and a lower value for this assumption explored, at 57% and 6%, respectively.	[174]
4	*Vaccine efficacy*	For VE *vs* the Alpha and Delta variants, assumed +/-10% for all VE parameters . Given greater uncertainty around VE of booster doses *vs* the Omicron variant (V_3 to V_6 strata), I explored the upper and lower bound from available literature. See Table A6, Table A7 and Table A8 for specific parameters.	Assumption and [281]
5	*Alternative μ_D*	Explored two alternative parameterisations of the μ_D piecewise linear function (see section 2.2.4), varying change points around the point of emergence of the Alpha and Delta variants.	Assumed
6	*Serial interval duration*	Central parameters assumed a decreasing SI for each variant in succession, compared to the Wildtype variant (Section 2.2.3). I re-fitted the model assuming a fixed SI of 5.2 days (Wildtype variant-like) and of 3.9 days (Omicron variant-like) for all variants.	Assumed

Vaccine efficacy sensitivity analysis

Table A6, table A7 and table A8 summarise VE values used for each scenario. Please see comparison of inferred intrinsic transmissibility and basic severity for each scenario, compared to central parameters, in Figure A8.

Table A6: Sensitivity analysis parameters for vaccine efficacy *vs* the Alpha variant. "Infection" refers to vaccine effectiveness protecting an individual from being infected with SARS-CoV-2, whilst "transmission" refers to the vaccine effectiveness at preventing onward transmission by an infected individual. Source: Perez-Guzman et al. 2023 [1].

		Central		Lower VE		Higher VE		Informed by
End point	Dose	AZ	PF/Mod	AZ	PF/Mod	AZ	PF/Mod	
Death	1	88%	89%	79%	80%	96%	98%	Assumed -10% lower, +10% higher VE.
	2 (Full protection)	99%	99%	89%	89%	99%	99%	
	2 (Waned protection)	83%	90%	75%	81%	91%	99%	
Severe disease	1	81%	89%	73%	80%	89%	98%	Assumed -10% lower, +10% higher VE.
	2 (Full protection)	98%	99%	89%	89%	99%	99%	
	2 (Waned protection)	77%	90%	70%	81%	85%	99%	
Mild disease or infection	1	64%	79%	58%	71%	70%	86%	Assumed -10% lower, +10% higher VE.
	2 (Full protection)	92%	99%	83%	89%	99%	99%	
	2 (Waned protection)	29%	77%	26%	69%	32%	85%	
Transmission	1	45%	45%	40%	40%	50%	50%	Assumed -10% lower, +10% higher VE.
	2 (Full protection)	45%	45%	40%	40%	50%	50%	
	2 (Waned protection)	40%	40%	36%	36%	44%	44%	

Table A7: Sensitivity analysis parameters for vaccine efficacy *vs* the Delta variant. "Infection" refers to vaccine effectiveness protecting an individual from being infected with SARS-CoV-2, whilst "transmission" refers to the vaccine effectiveness at preventing onward transmission by an infected individual. Source: Perez-Guzman et al. 2023 [1].

		Central		Lower VE		Higher VE		Informed by
End point	Dose	AZ	PF/Mod	AZ	PF/Mod	AZ	PF/Mod	
Death	1	87%	89%	78%	80%	96%	98%	Assumed -10% lower, +10% higher VE.
	2 (Full protection)	98%	99%	89%	89%	99%	99%	
	2 (Waned protection)	82%	90%	74%	81%	90%	99%	
	3 (Full protection)	99%	99%	89%	89%	99%	99%	
	3 (Waned protection)	89%	89%	81%	81%	99%	99%	
Severe disease	1	81%	89%	73%	80%	89%	98%	Assumed -10% lower, +10% higher VE.
	2 (Full protection)	98%	99%	89%	89%	99%	99%	
	2 (Waned protection)	77%	90%	70%	81%	85%	99%	
	3 (Full protection)	99%	99%	89%	89%	99%	99%	
	3 (Waned protection)	89%	89%	81%	81%	99%	99%	
Mild disease or infection	1	51%	51%	46%	46%	56%	56%	Assumed -10% lower, +10% higher VE.
	2 (Full protection)	87%	95%	79%	86%	96%	99%	
	2 (Waned protection)	19%	49%	17%	44%	21%	54%	
	3 (Full protection)	92%	92%	83%	83%	99%	99%	
	3 (Waned protection)	36%	36%	39%	39%	54%	54%	
Transmission	1	33%	33%	30%	30%	36%	36%	Assumed -10% lower, +10% higher VE.
	2 (Full protection)	40%	40%	36%	36%	44%	44%	
	2 (Waned protection)	19%	19%	17%	17%	21%	21%	
	3 (Full protection)	40%	40%	36%	36%	44%	44%	
	3 (Waned protection)	19%	19%	17%	17%	21%	21%	

Table A8: Sensitivity analysis parameters for vaccine effectiveness against the Omicron variant for AstraZeneca (AZ), Pfizer (PF), and Moderna (Mod) by vaccine dose. "Infection" refers to vaccine effectiveness protecting an individual from being infected with SARS-CoV-2, whilst "transmission" refers to the vaccine effectiveness at preventing onward transmission by an infected individual. Source: Perez-Guzman et al. 2023 [1].

		Central		Lower VE		Higher VE		Informed by
End point	Dose	AZ	PF/Mod	AZ	PF/Mod	AZ	PF/Mod	
Death	2 (Full protection)	97%	97%	97%	97%	97%	97%	Low and high values from Gov [281]
	2 (Waned protection)	56%	56%	56%	56%	56%	56%	
	3 (Full protection)	96%	96%	70%	70%	99%	99%	
	3 (Waned protection)	62%	62%	50%	50%	80%	80%	
Severe disease	2 (Full protection)	97%	97%	97%	97%	97%	97%	Low and high values from Gov [281]
	2 (Waned protection)	56%	56%	56%	56%	56%	56%	
	3 (Full protection)	96%	96%	70%	70%	99%	99%	
	3 (Waned protection)	62%	62%	40%	40%	75%	75%	
Mild disease or infection	2 (Full protection)	41%	60%	41%	60%	41%	60%	Low and high values from Gov [281]
	2 (Waned protection)	0%	0%	0%	0%	0%	0%	
	3 (Full protection)	72%	74%	50%	55%	75%	75%	
	3 (Waned protection)	0%	0%	0%	0%	10%	10%	
Transmission	2 (Full protection)	40%	40%	40%	40%	40%	40%	Assumed lower than for mild disease
	2 (Waned protection)	0%	0%	0%	0%	0%	0%	
	3 (Full protection)	40%	40%	30%	30%	50%	50%	
	3 (Waned protection)	0%	0%	0%	0%	0%	0%	

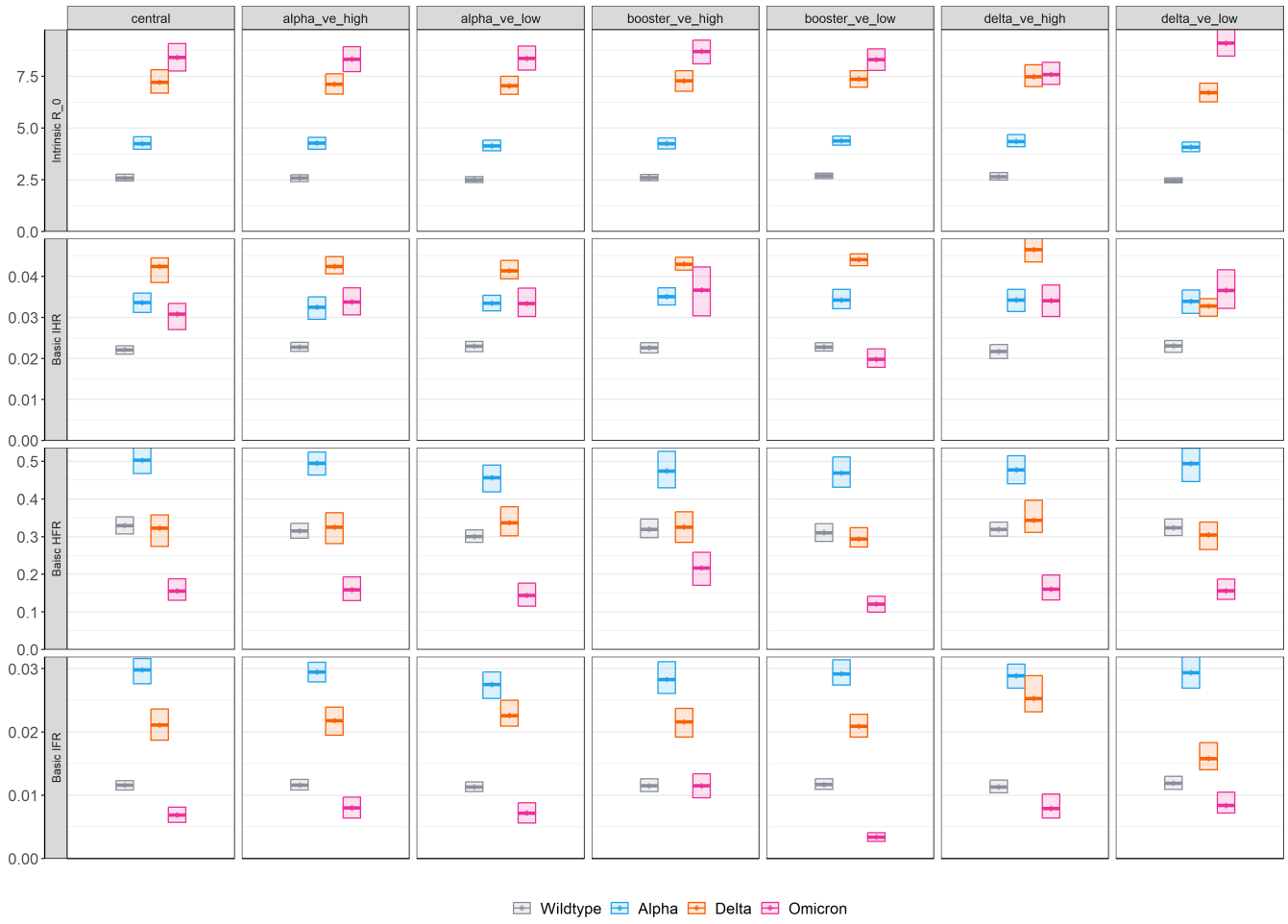


Figure A8: Sensitivity analysis of model inferred intrinsic R_0 and basic IHR, HFR and IFR of the variants. From left to right: Central, high and low vaccine efficacy (1st and 2nd dose, table A6) against the Alpha variant, high and low vaccine efficacy (1st, 2nd and booster, table 3.3) against the Delta variant, and high and low vaccine efficacy (2nd and booster dose only table A8) against the Omicron variant. Box plots show mean model-inferred properties and 95% CrI for the Wildtype (grey), Alpha (blue), Delta (orange) and Omicron (pink) variants. Source: Perez-Guzman et al. 2023 [1].

Cross-immunity sensitivity analysis

I varied central parameters of cross-immunity given prior infection with a historic variant *vs* an emerging variant (e.g. prior infection with Alpha or Delta *vs* Omicron variant). Each of cross-immunity against infection, hospitalisation or death was varied in a one-way analysis, simultaneously for all variants. For specific values and literature sources of parameters (where available) see table A5.

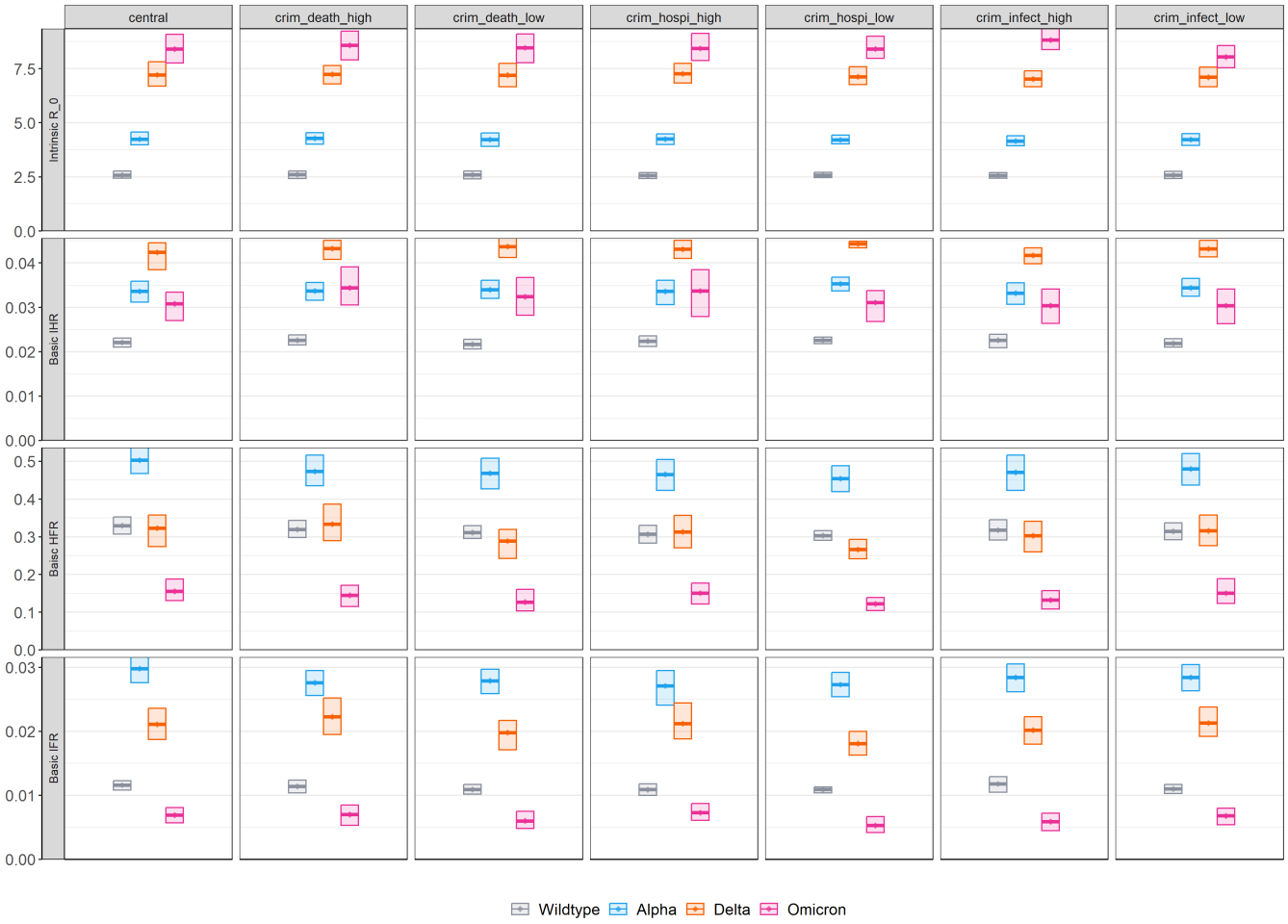


Figure A9: Sensitivity analysis for scenario description) of model inferred intrinsic R_0 and basic IHR, HFR and IFR of the variants. From left to right: high and low cross-immunity against death, high and low cross-immunity against hospitalisation, and high and low cross-immunity against infection (see values in table A5). Box plots show mean model-inferred properties and 95% CrI for the Wildtype (grey), Alpha (blue), Delta (orange) and Omicron (pink) variants. Source: Perez-Guzman et al. 2023 [1].

Other sensitivity analysis

The model allowed fitting time-varying changes in severity, independent of the inferred basic properties of the variants. I explored alternative parameterisations of the piece-wise linear function for changing $\mu_D(t)$ (see details in section 2.2.4. The specific aim of these scenarios was to infer whether the choice of change points influenced the inferred relative severity between the Alpha and Delta variants. Additionally, I varied the central serial interval parameters, by either fixing it at its highest (i.e. all variants with same value as Wildtype) and lowest (i.e. all variants with same value as Omicron BA.1) values.

Table A9: Sensitivity analysis of fitted change points for the time-varying $\mu_D(t)$. Source: Perez-Guzman et al. 2023 [1].

Scenario	Dates	Rationale
Central	01-04-2020 01-07-2020 to 15-09-2020 15-10-2020 to 01-12-2020 04-02-2021 01-04-2021 to 04-11-2021 31-12-2021	See table 2.4
$\mu_D(t)winter$	01-04-2020 01-07-2020 to 15-09-2020 01-12-2020 04-02-2021 01-04-2021 to 04-11-2021 31-12-2021	Explore a linear change in $\mu_D(t)$ from 09-2020 to 12-2020 to infer potential changes in severity given mounting healthcare demands around the date of Alpha emergence.
$\mu_D(t)summer$	01-04-2020 01-07-2020 to 15-09-2020 15-10-2020 to 01-12-2020 04-02-2021 01-04-2021 to 01-06-2021 04-11-2021 31-12-2021	Explore flat period in $\mu_D(t)$ between 04-2021 and 06-2021 during the period of Delta emergence, then linear change from 06-2021 to 11-2021 as Delta was dominant during a period of potential evolving hospital admission and triaging thresholds.

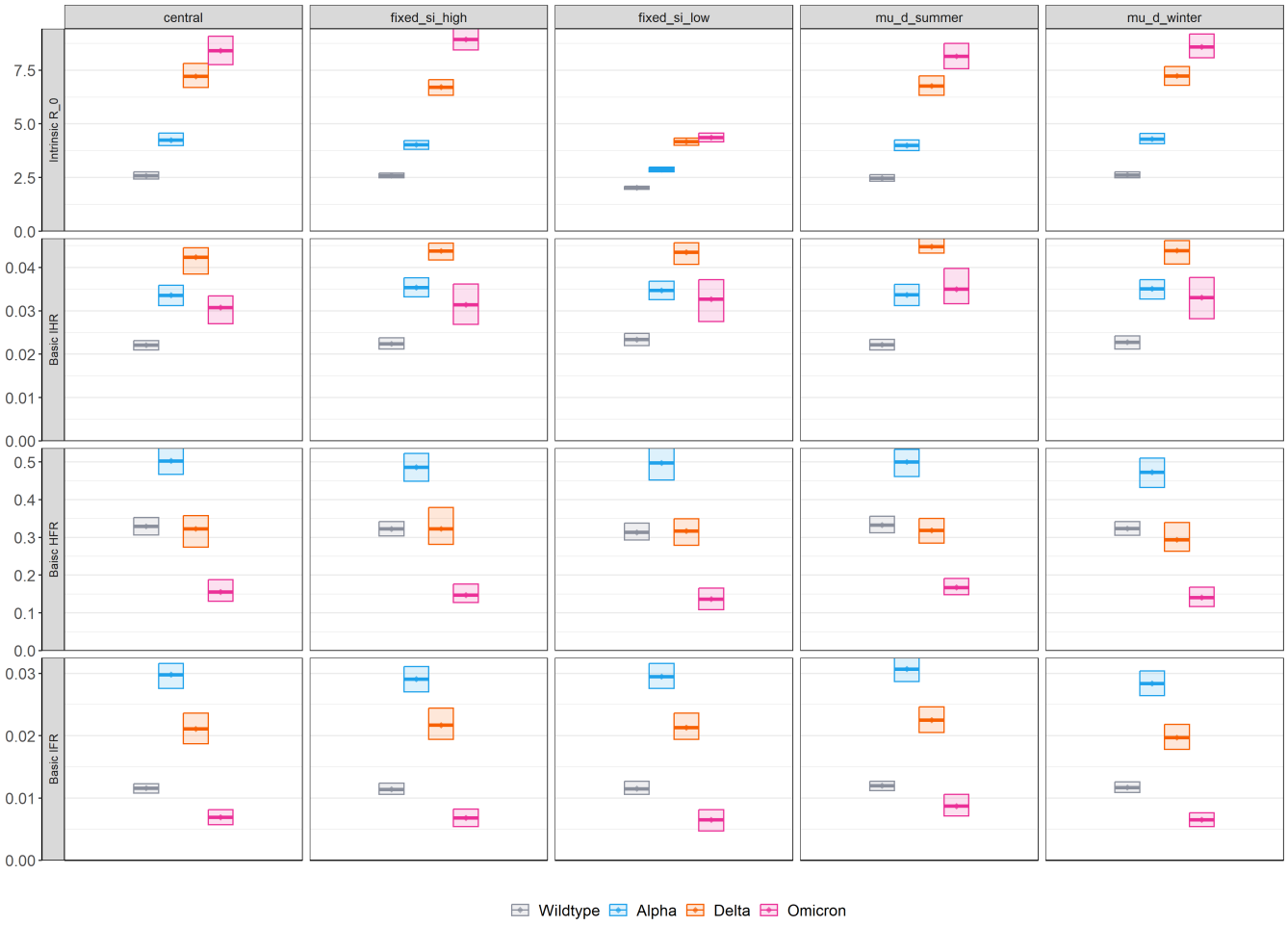


Figure A10: Sensitivity analysis of model inferred intrinsic R_0 and basic IHR, HFR and IFR of the variants. From left to right: Central, high and low fixed serial interval duration (see values in table A5), and alternative changepoints for $\mu_D(t)$ in the winter of 2020/21 and in the summer of 2021 (see values and rationale in table A9). Box plots show mean model-inferred properties and 95% CrI for the Wildtype (grey), Alpha (blue), Delta (orange) and Omicron (pink) variants. Source: Perez-Guzman et al. 2023 [1].

A.4 Fitted data trajectories

Plots in this section present the model-fitted trajectories to surveillance data streams in England during the study period. Where age-specific data fits are shown (Figures A20 to A56), the respective model comparison to aggregated data in Figure A11 are for illustrative purposes.

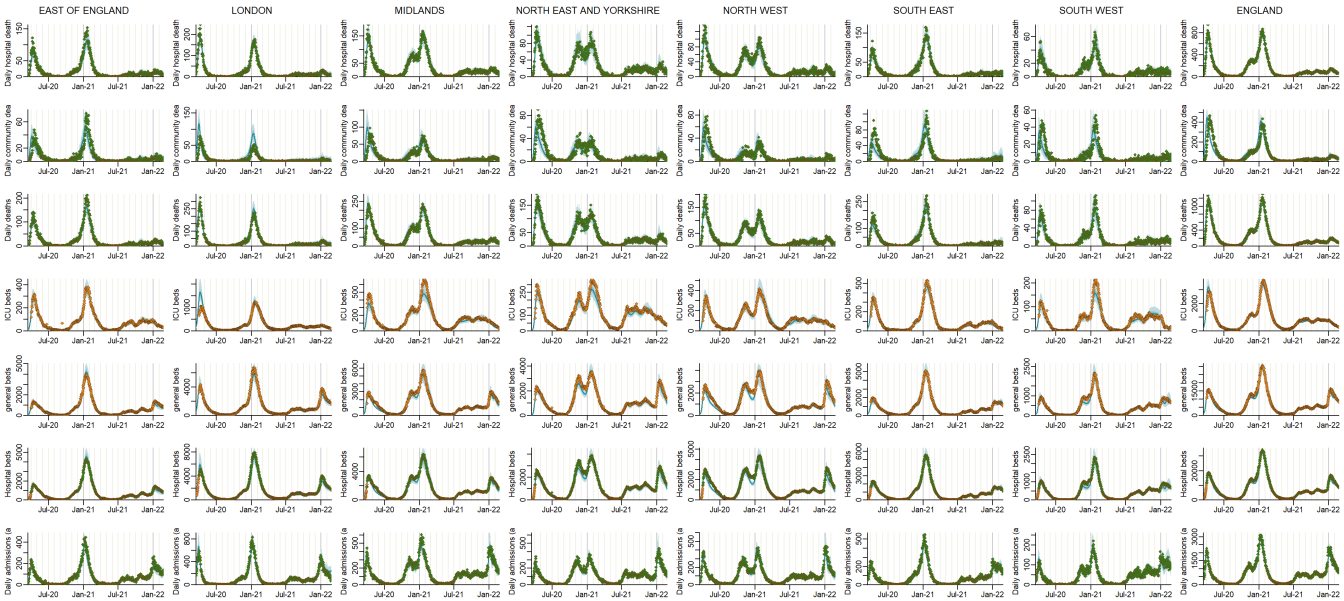


Figure A11: Fitted model trajectories to hospital and deaths data streams by England NHS regions. Source: Perez-Guzman et al. 2023 [1].

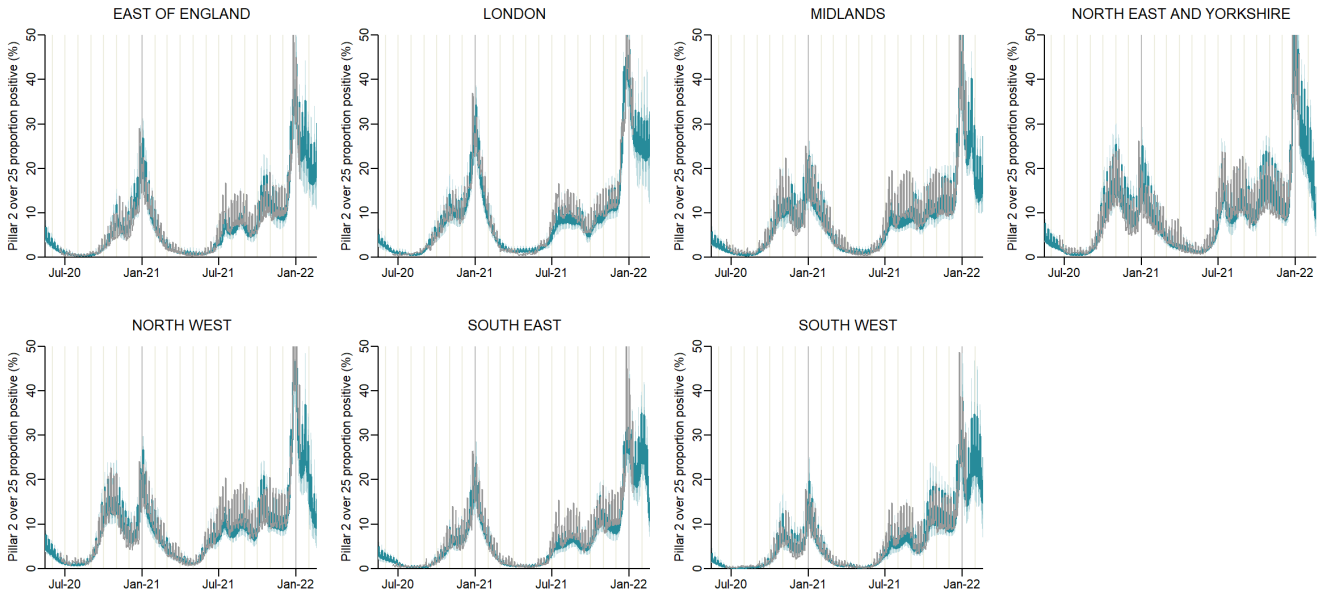


Figure A12: Fitted model trajectories to Pillar 2 (symptomatic case, PCR only) data by England NHS regions. Source: Perez-Guzman et al. 2023 [1].

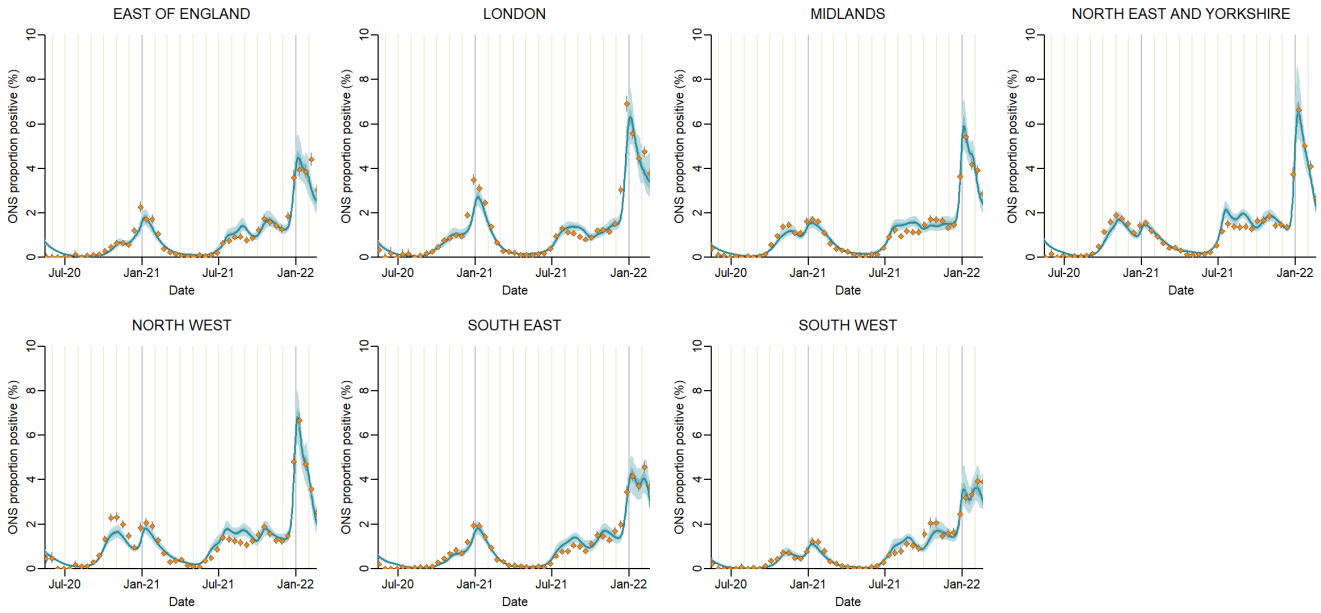


Figure A13: Fitted model trajectories to ONS infection prevalence data by England NHS regions. Source: Perez-Guzman et al. 2023 [1].

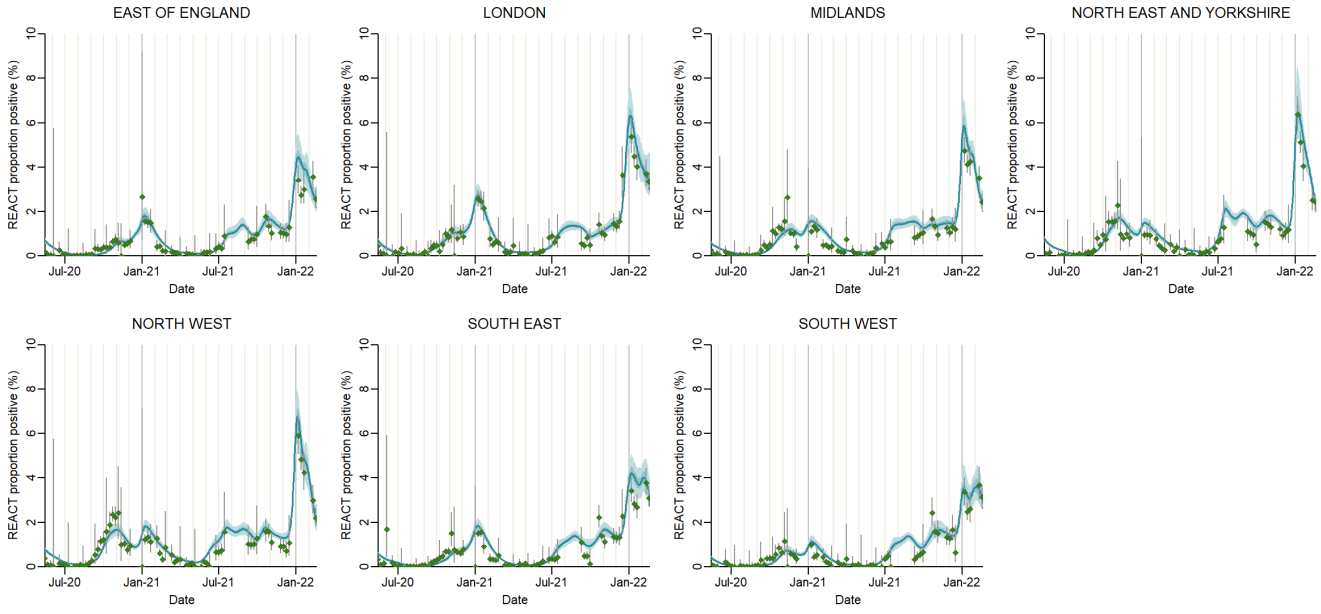


Figure A14: Fitted model trajectories to REACT-1 infection prevalence data by England NHS regions. Source: Perez-Guzman et al. 2023 [1].

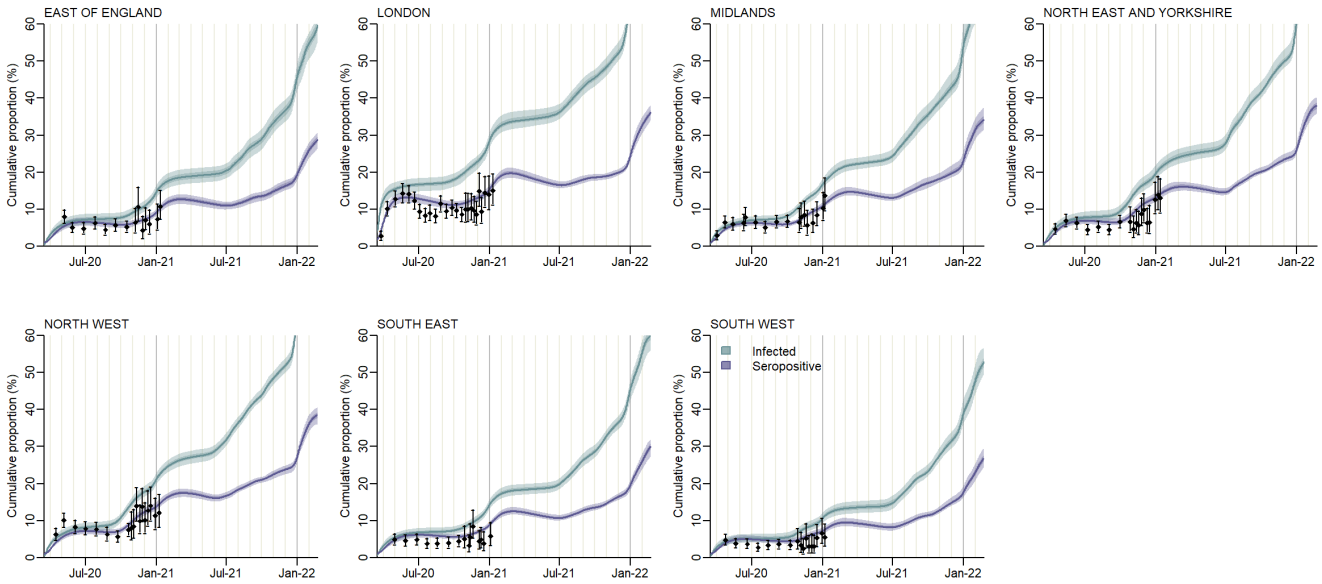


Figure A15: Fitted model trajectories to Euroimmun serology prevalence data by England NHS regions. Source: Perez-Guzman et al. 2023 [1].

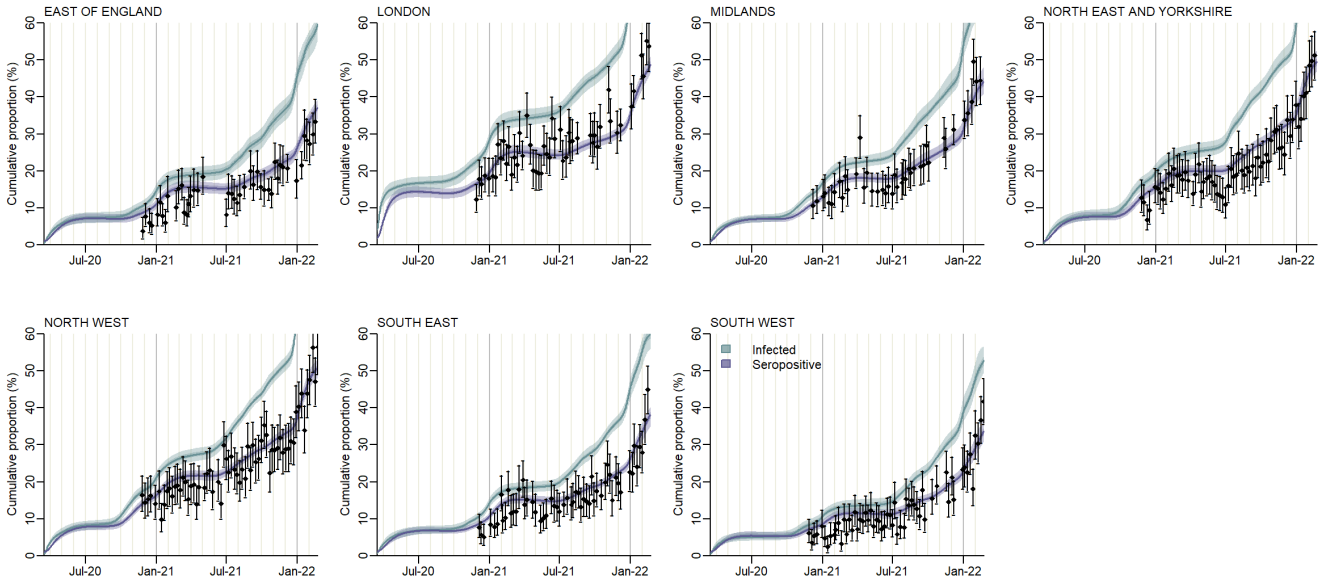


Figure A16: Fitted model trajectories to Roche serology data by England NHS regions. Source: Perez-Guzman et al. 2023 [1].

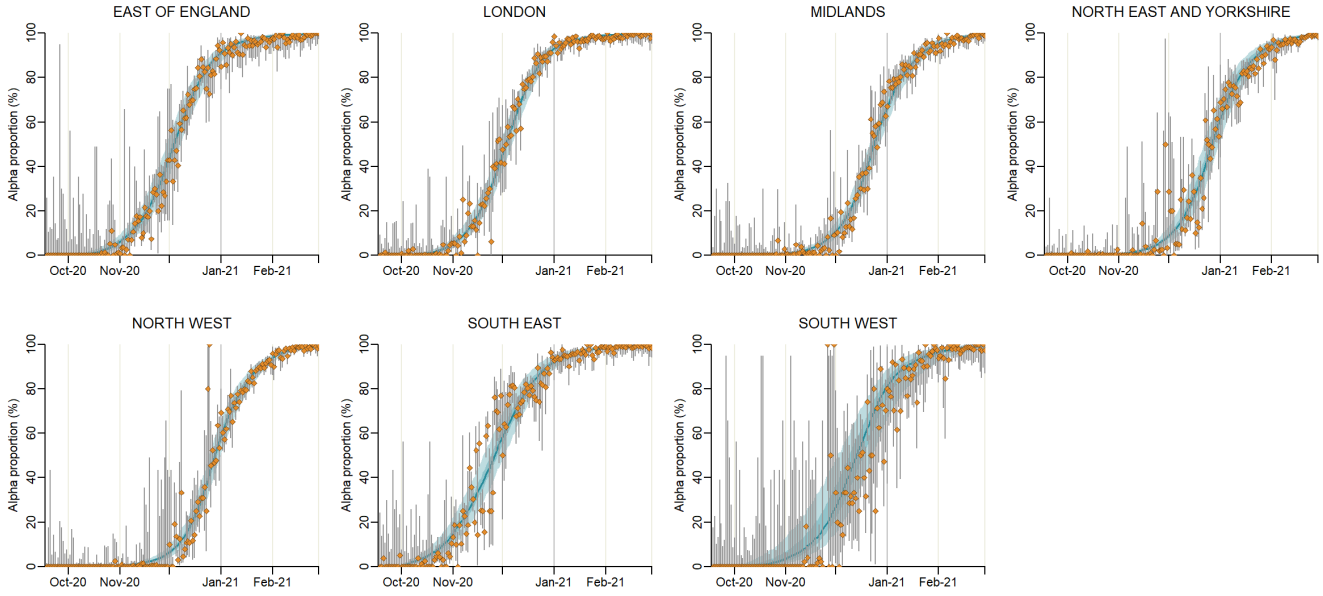


Figure A17: Fitted model trajectories to Alpha/Wildtype frequency from VAM data by England NHS regions. Source: Perez-Guzman et al. 2023 [1].

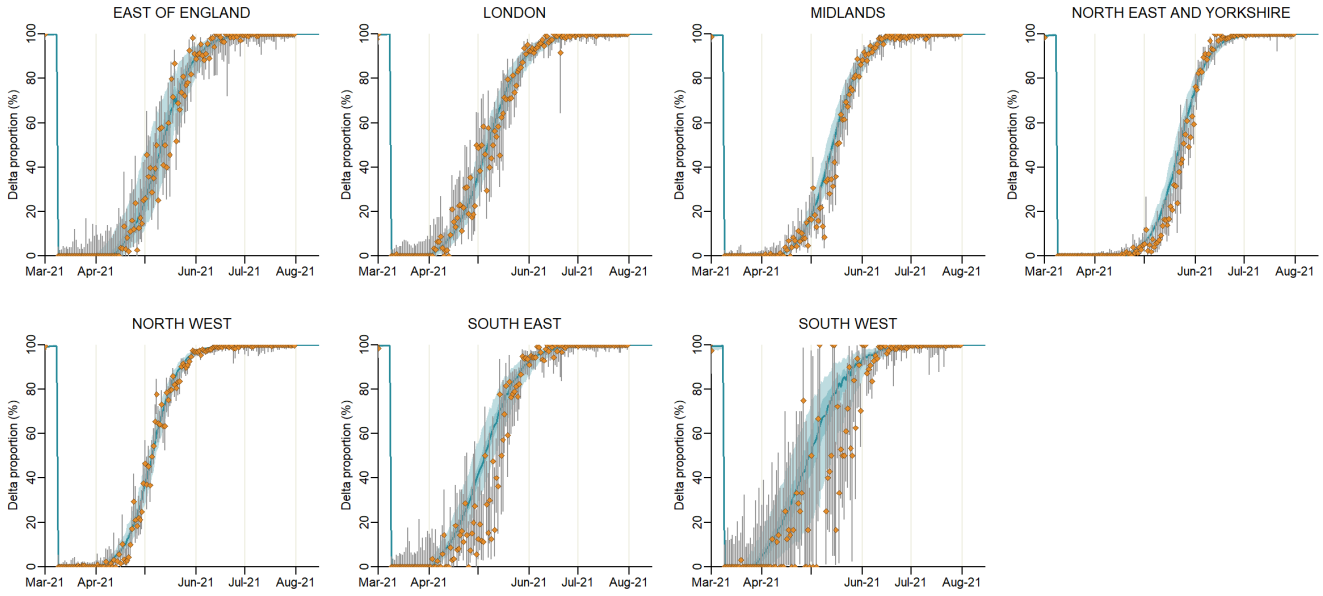


Figure A18: Fitted model trajectories to Delta/Alpha frequency from VAM data by England NHS regions. Source: Perez-Guzman et al. 2023 [1].

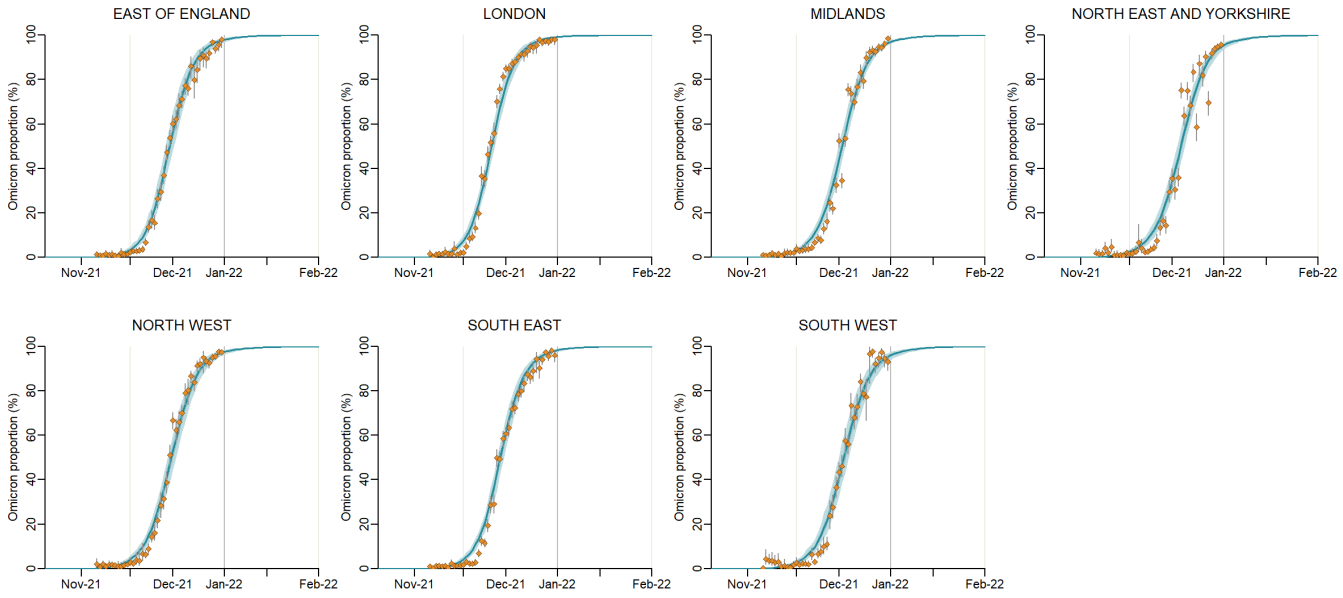


Figure A19: Fitted model trajectories to Omicron/Delta frequency from VAM data by England NHS regions. Source: Perez-Guzman et al. 2023 [1].

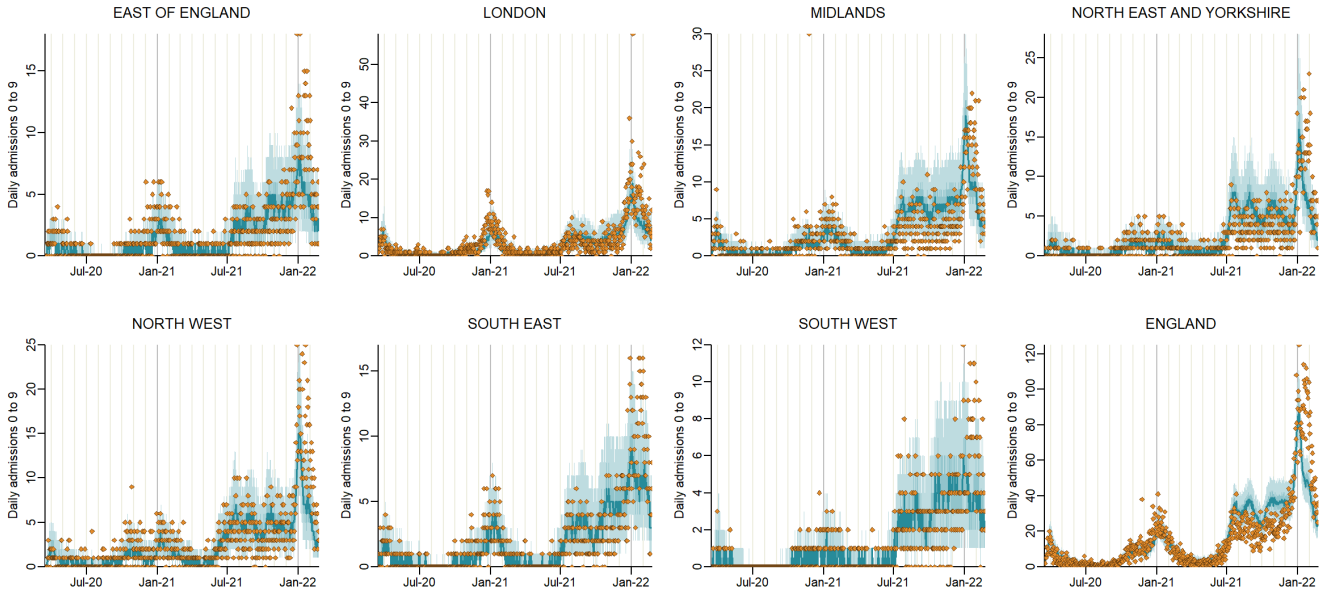


Figure A20: Fitted model trajectories to hospital admissions of 0 to 9 year-olds data by England NHS regions. Source: Perez-Guzman et al. 2023 [1].

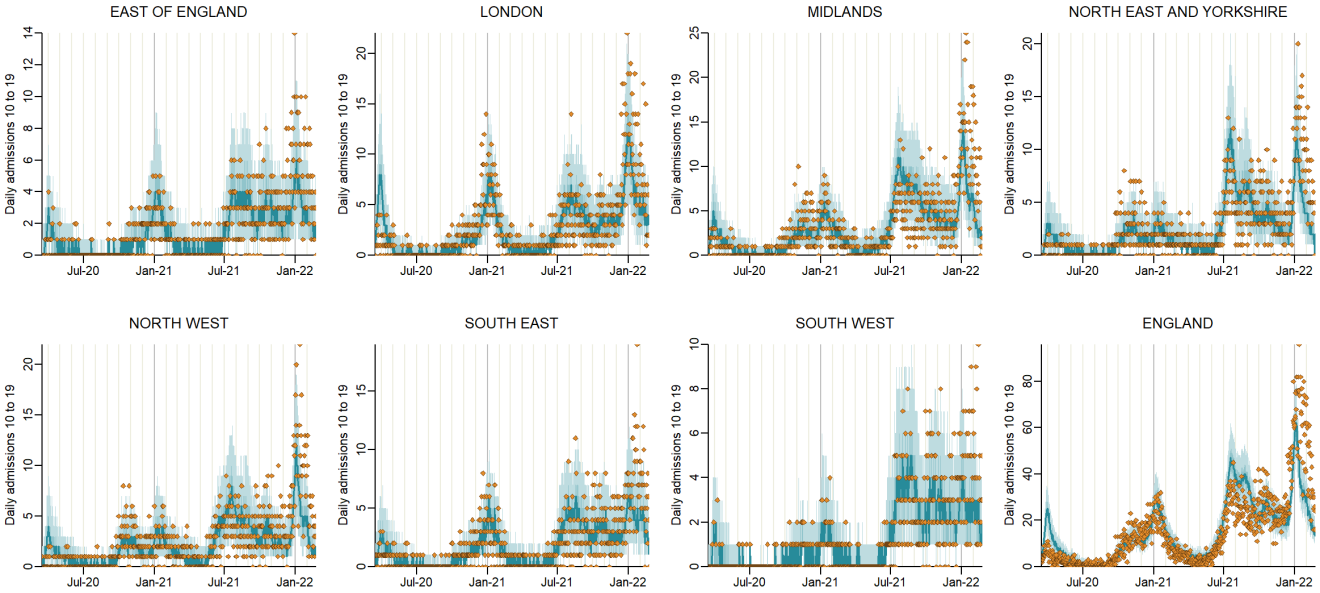


Figure A21: Fitted model trajectories to hospital admissions of 10 to 19 year-olds data by England NHS regions. Source: Perez-Guzman et al. 2023 [1].

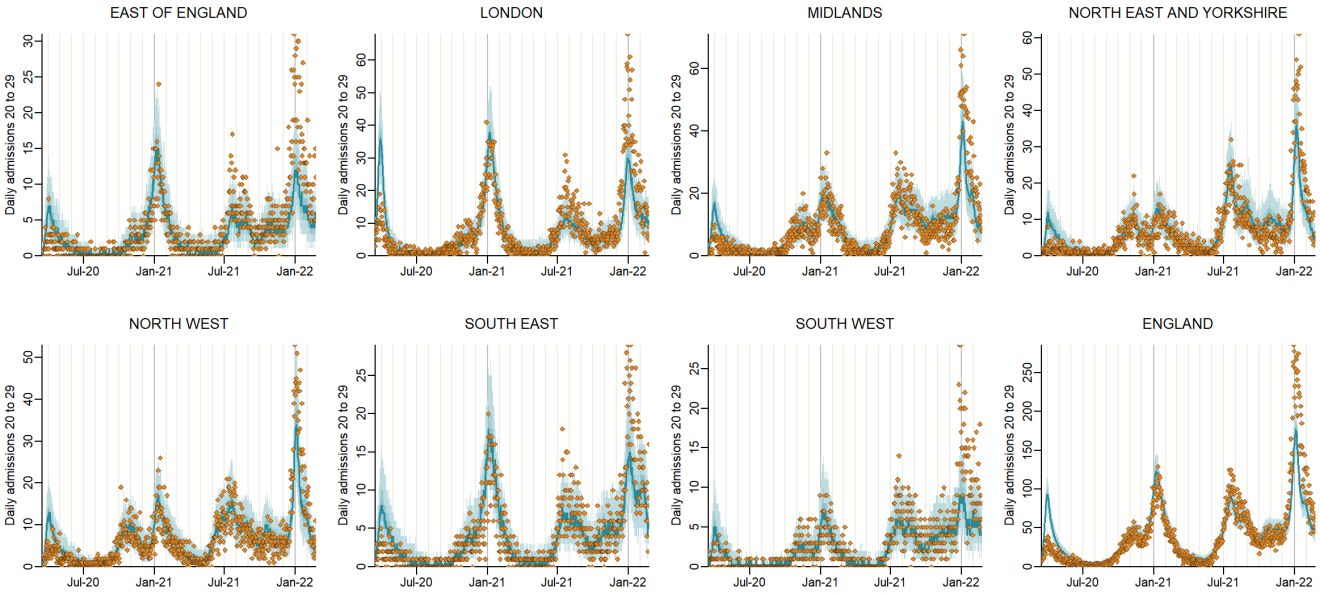


Figure A22: Fitted model trajectories to hospital admissions of 20 to 29 year-olds data by England NHS regions. Source: Perez-Guzman et al. 2023 [1].

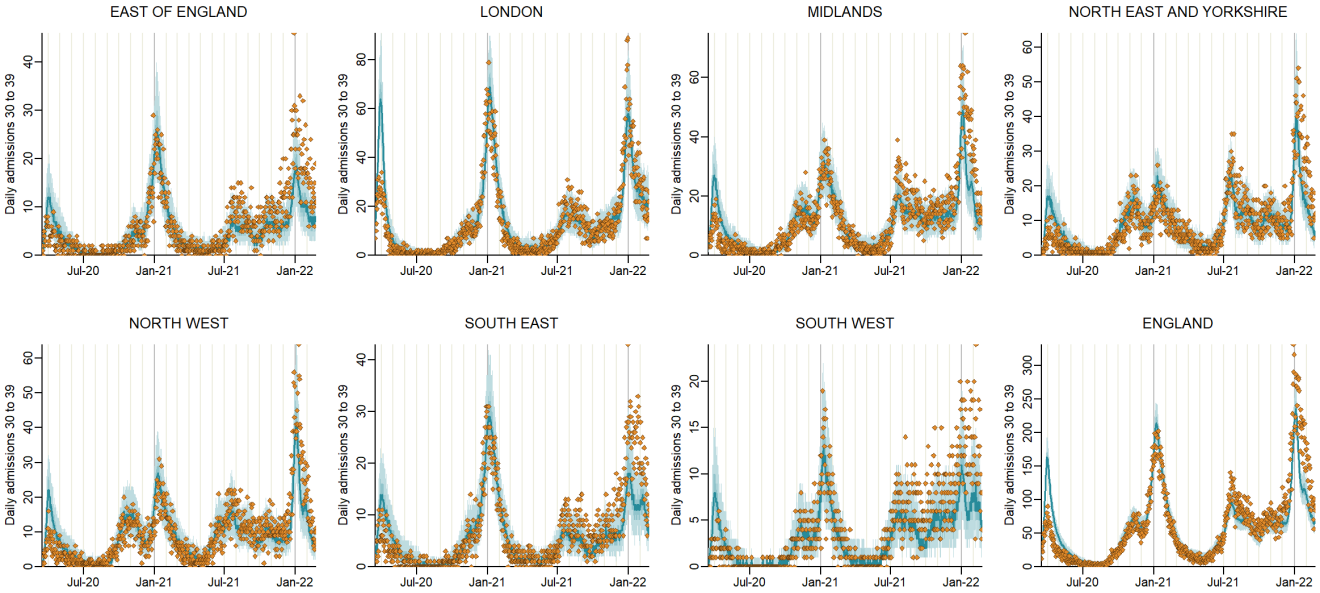


Figure A23: Fitted model trajectories to hospital admissions of 30 to 39 year-olds data by England NHS regions. Source: Perez-Guzman et al. 2023 [1].

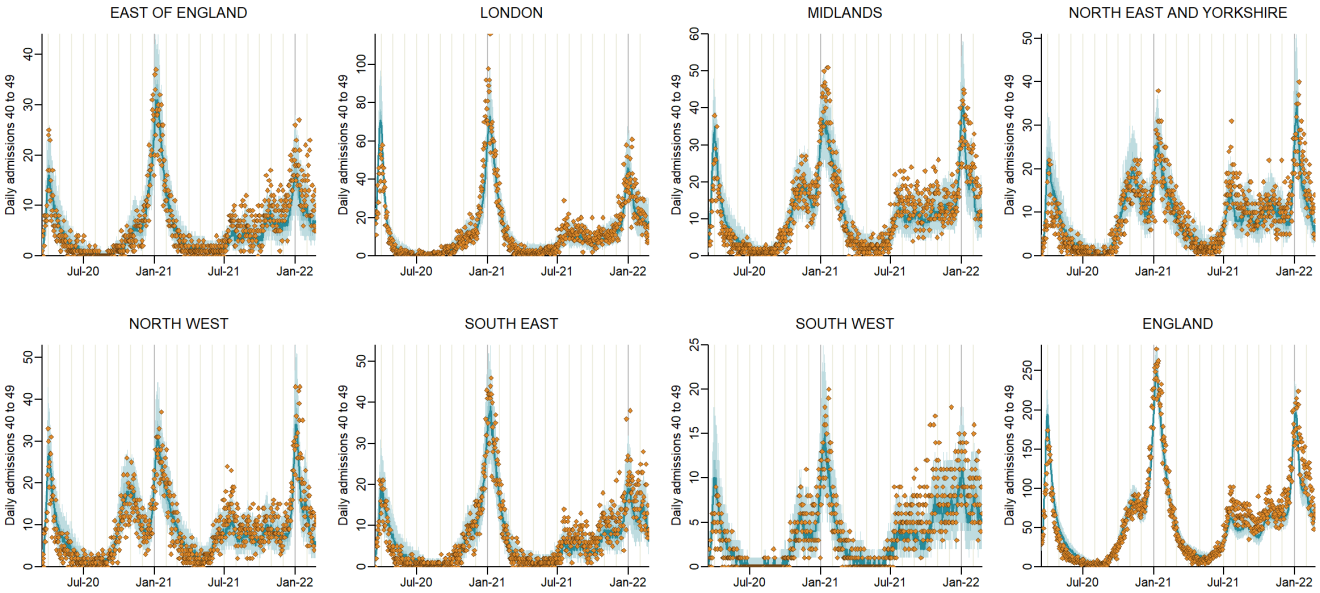


Figure A24: Fitted model trajectories to hospital admissions of 40 to 49 year-olds data by England NHS regions. Source: Perez-Guzman et al. 2023 [1].

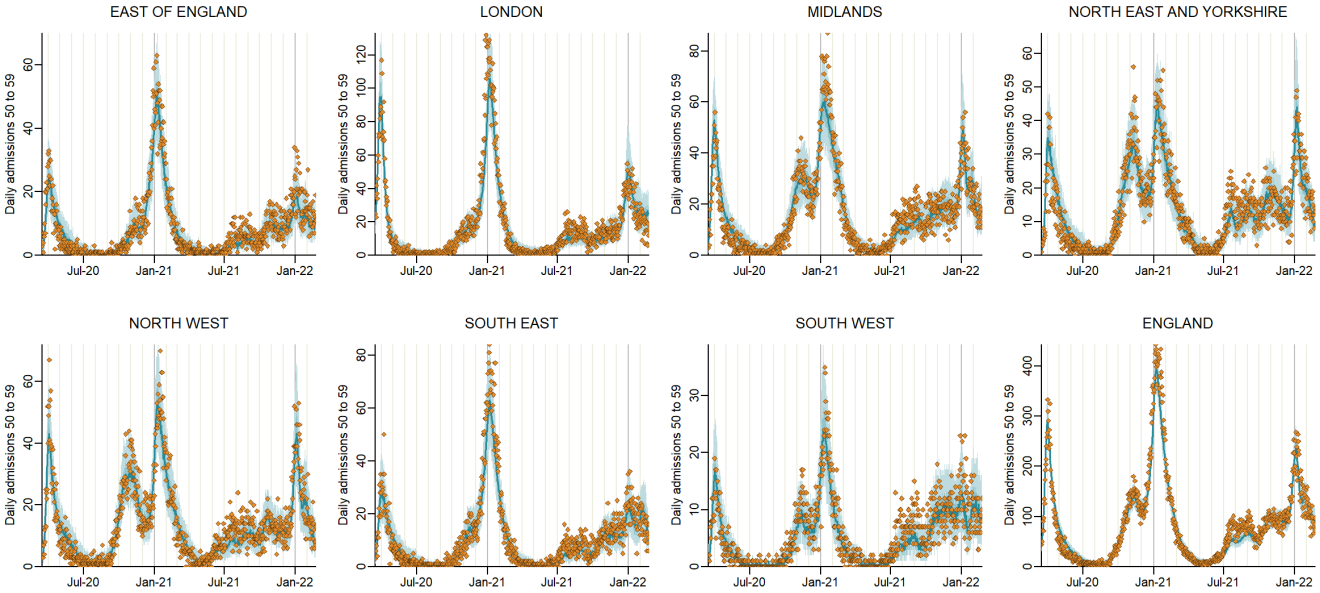


Figure A25: Fitted model trajectories to hospital admissions of 50 to 59 year-olds data by England NHS regions. Source: Perez-Guzman et al. 2023 [1].

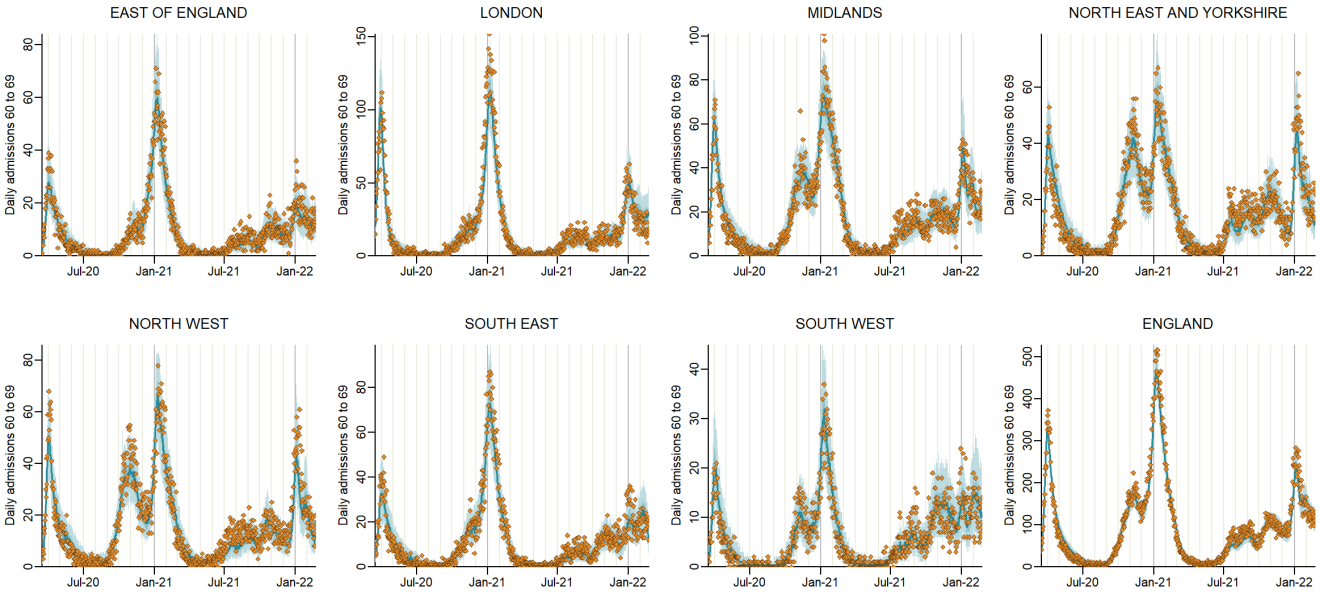


Figure A26: Fitted model trajectories to hospital admissions of 60 to 69 year-olds data by England NHS regions. Source: Perez-Guzman et al. 2023 [1].

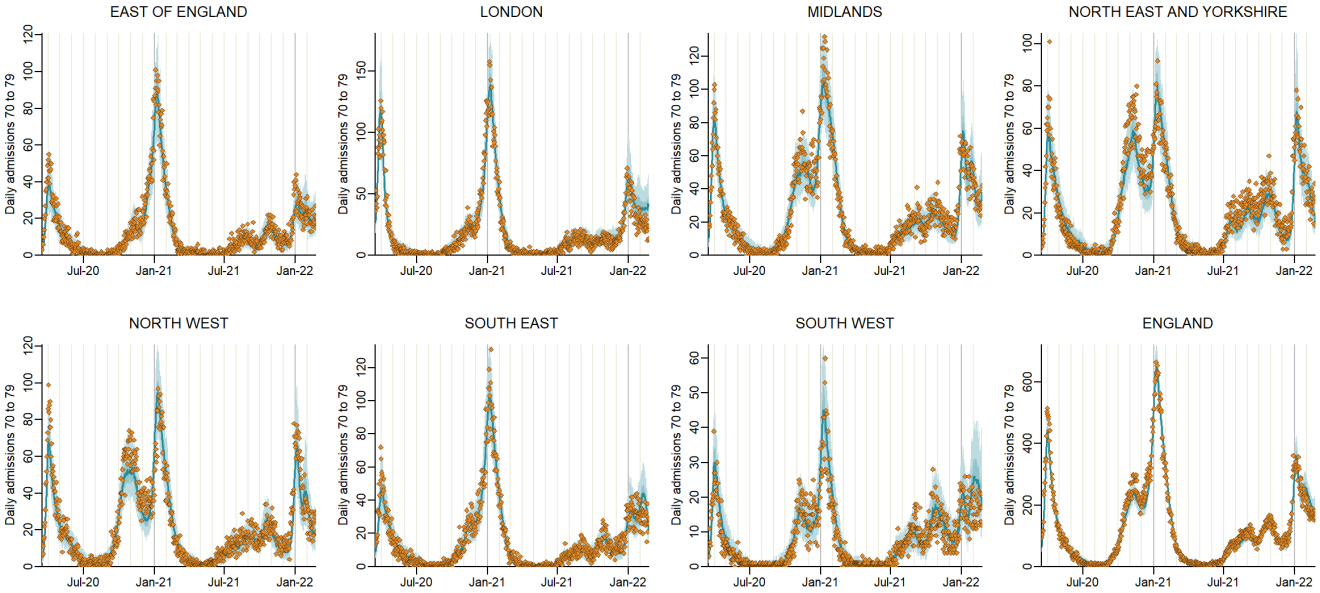


Figure A27: Fitted model trajectories to hospital admissions of 70 to 79 year-olds data by England NHS regions. Source: Perez-Guzman et al. 2023 [1].

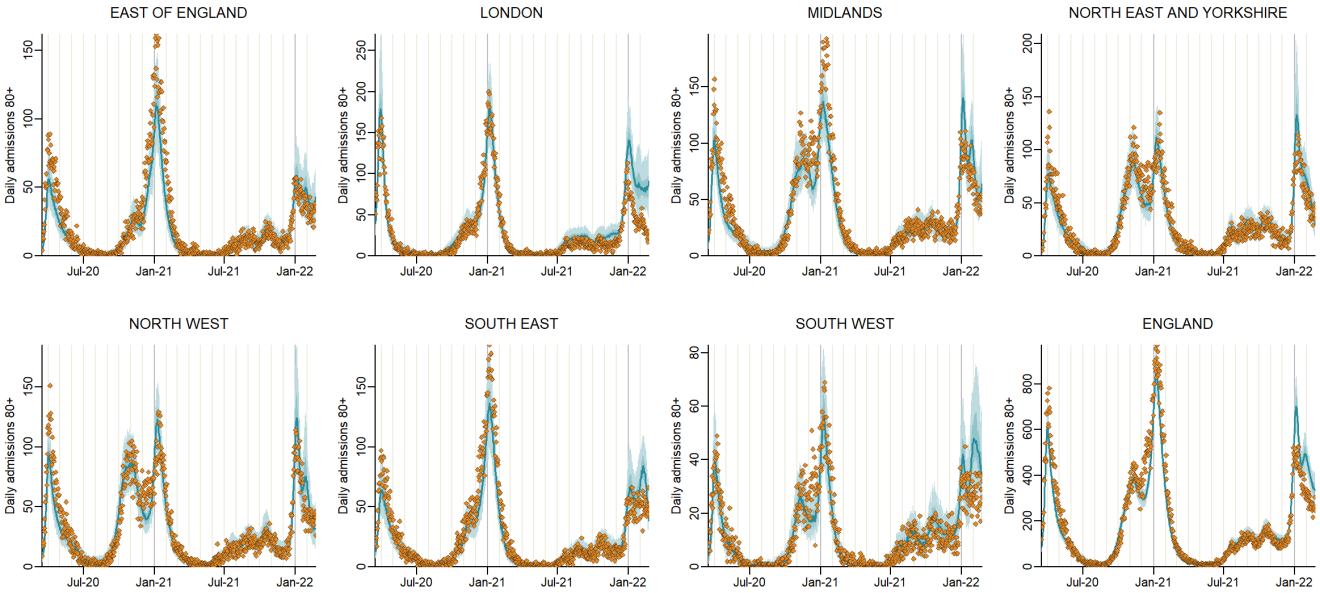


Figure A28: Fitted model trajectories to hospital admissions of 80+ year-olds data by England NHS regions. Source: Perez-Guzman et al. 2023 [1].

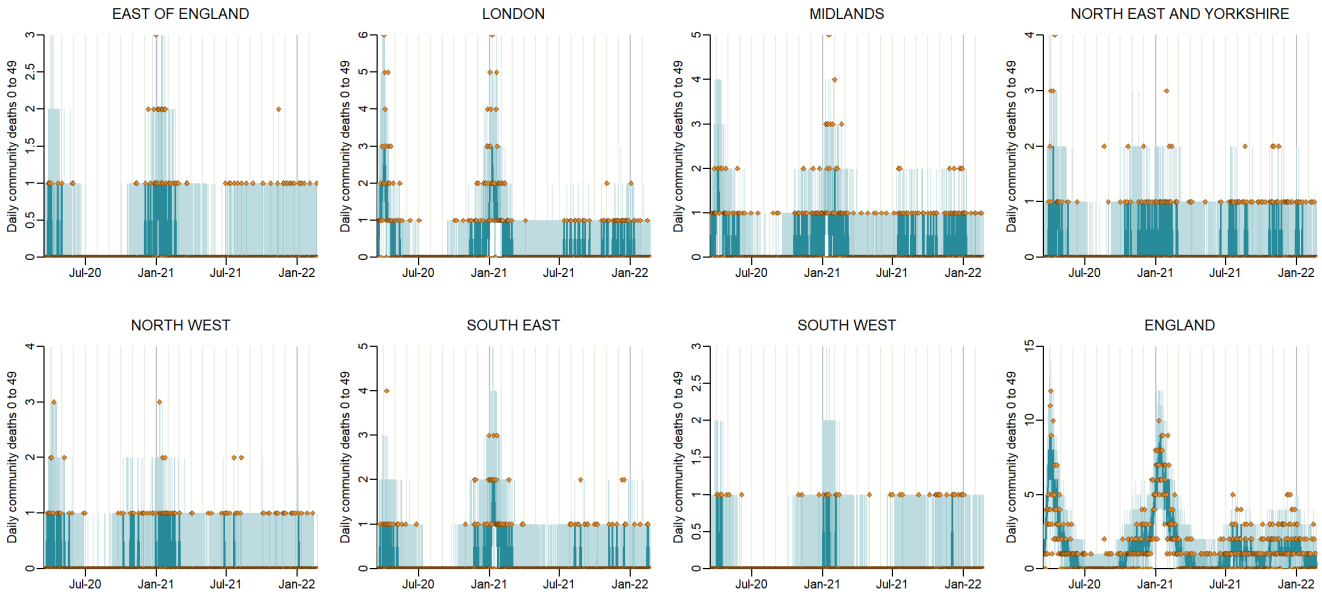


Figure A29: Fitted model trajectories to community deaths of 0 to 49 year-olds data by England NHS regions. Source: Perez-Guzman et al. 2023 [1].

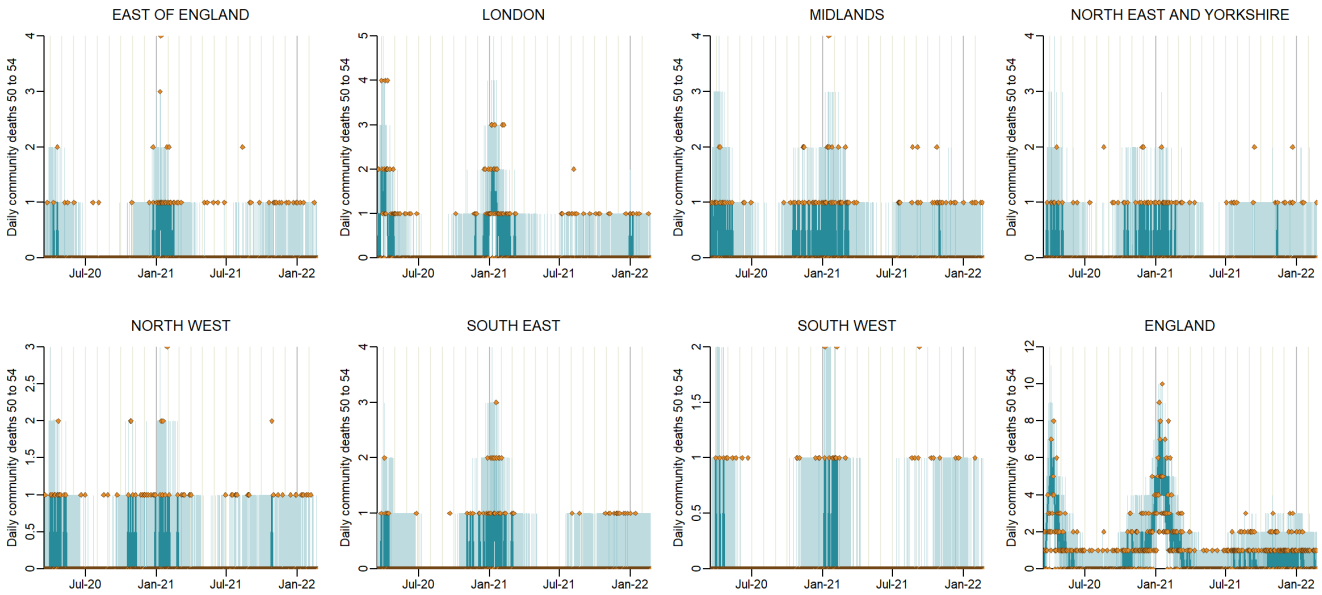


Figure A30: Fitted model trajectories to community deaths of 50 to 54 year-olds data by England NHS regions. Source: Perez-Guzman et al. 2023 [1].

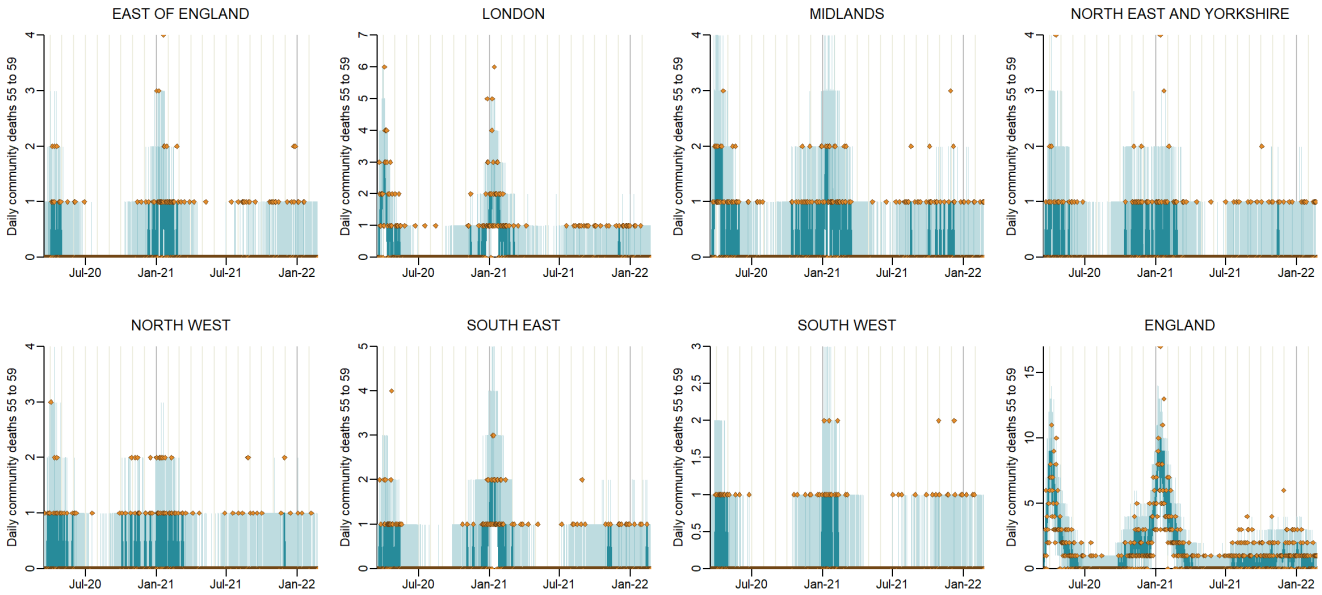


Figure A31: Fitted model trajectories to community deaths of 55 to 59 year-olds data by England NHS regions. Source: Perez-Guzman et al. 2023 [1].

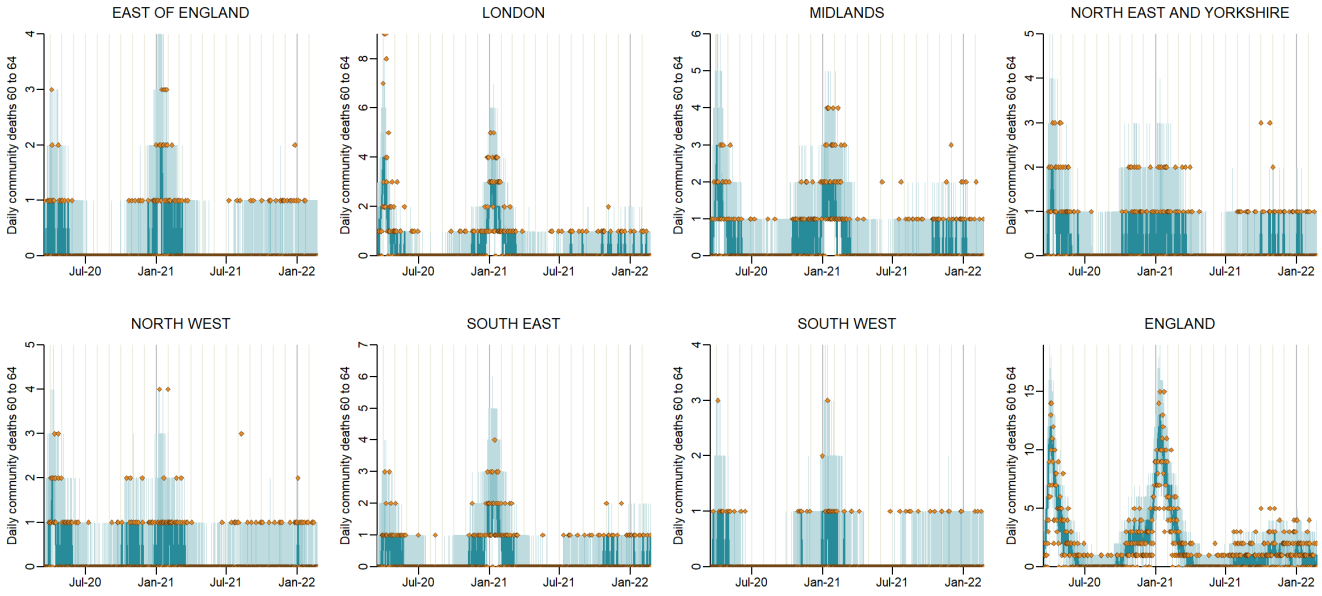


Figure A32: Fitted model trajectories to community deaths of 60 to 64 year-olds data by England NHS regions. Source: Perez-Guzman et al. 2023 [1].

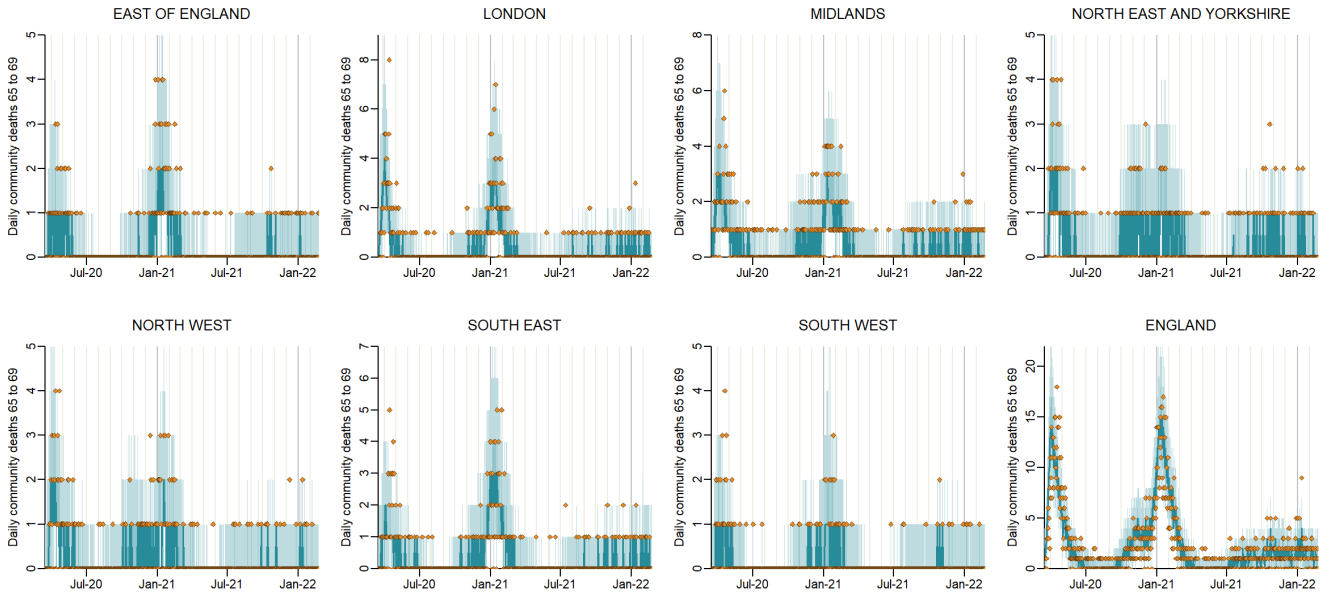


Figure A33: Fitted model trajectories to community deaths of 65 to 69 year-olds data by England NHS regions. Source: Perez-Guzman et al. 2023 [1].

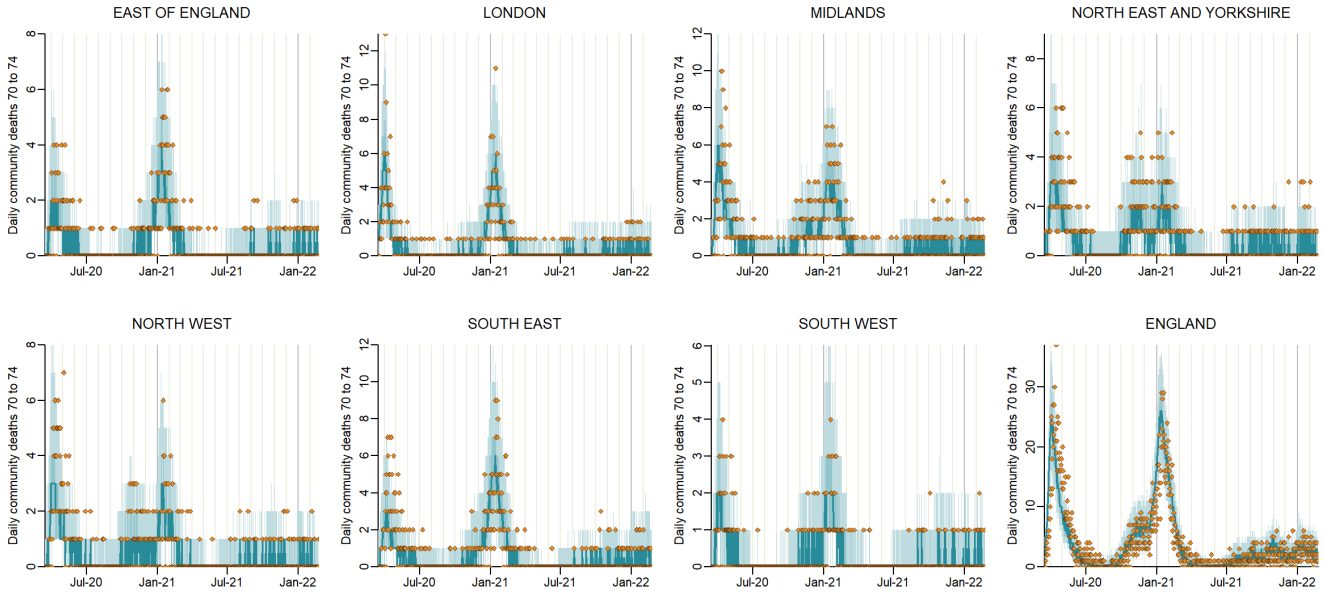


Figure A34: Fitted model trajectories to community deaths of 70 to 74 year-olds data by England NHS regions. Source: Perez-Guzman et al. 2023 [1].

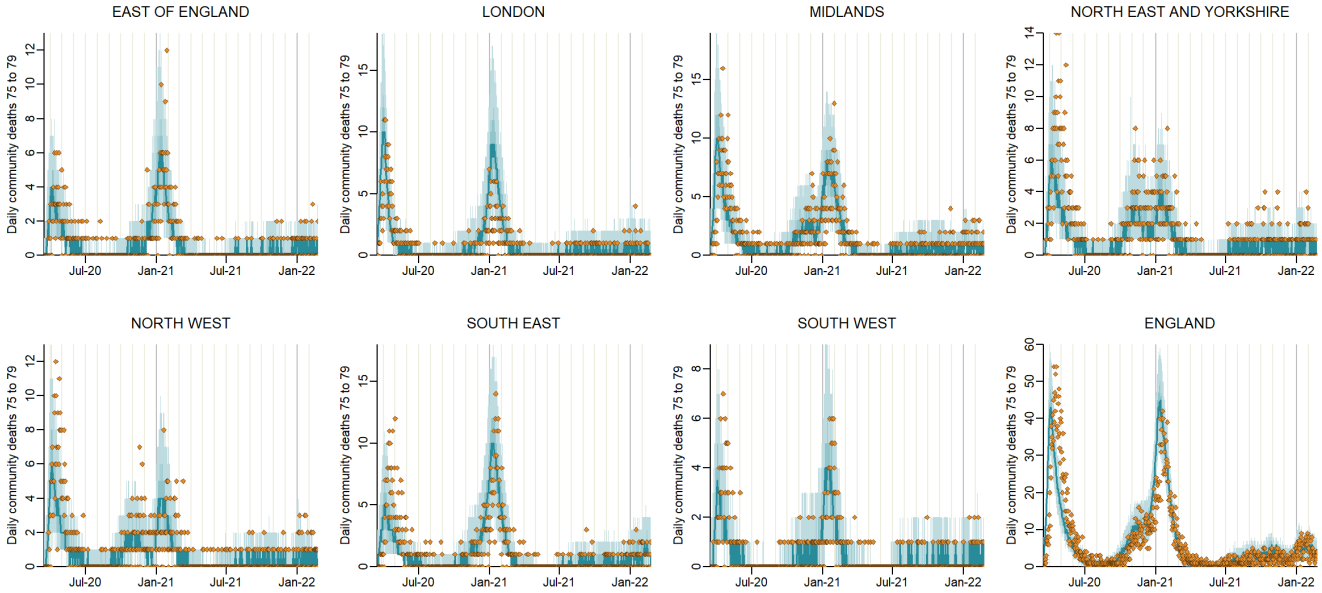


Figure A35: Fitted model trajectories to community deaths of 75 to 79 year-olds data by England NHS regions. Source: Perez-Guzman et al. 2023 [1].

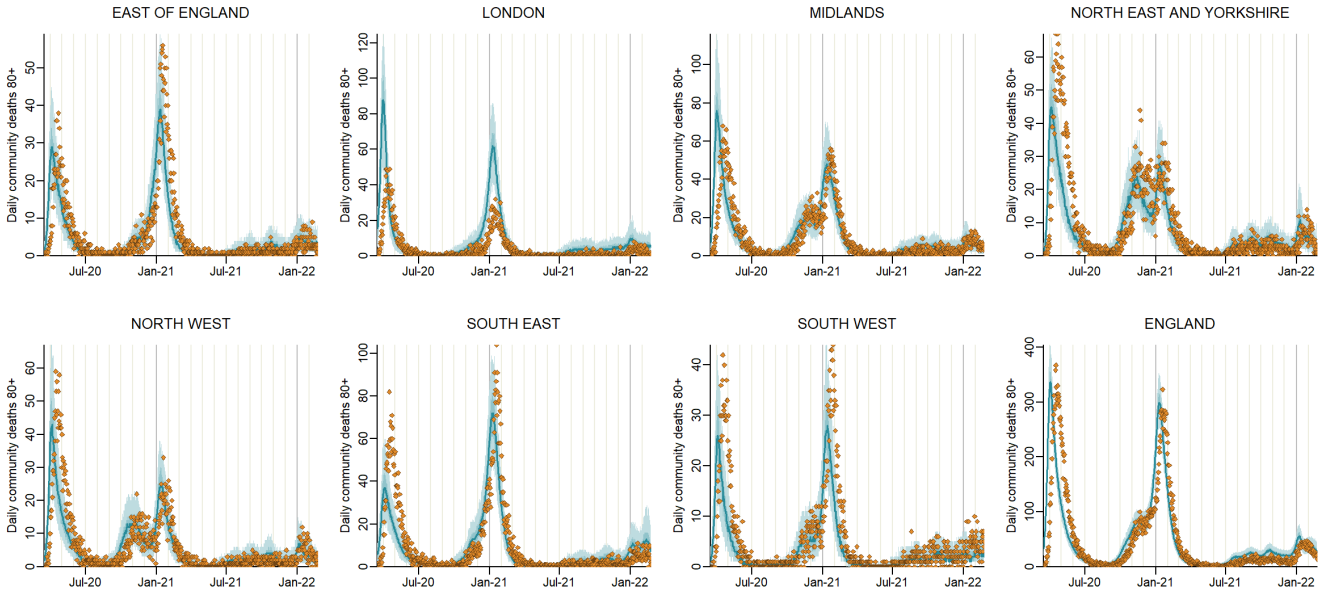


Figure A36: Fitted model trajectories to community deaths of 80 year-olds data by England NHS regions. Source: Perez-Guzman et al. 2023 [1].

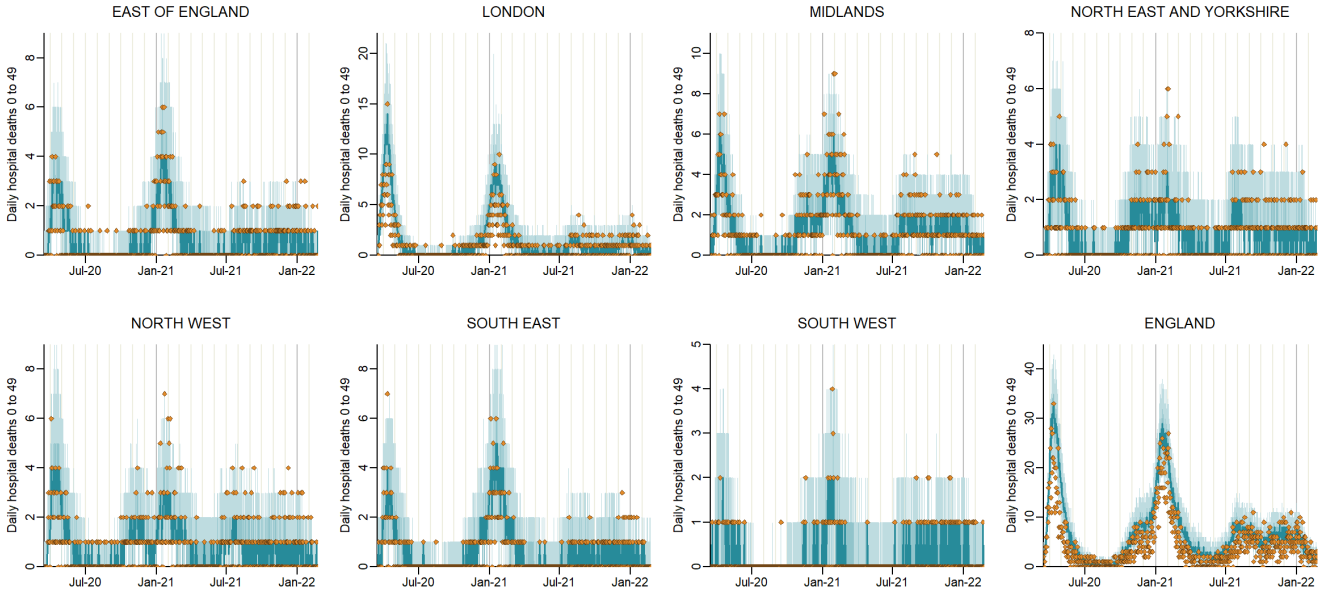


Figure A37: Fitted model trajectories to hospital deaths of 0 to 49 year-olds data by England NHS regions. Source: Perez-Guzman et al. 2023 [1].

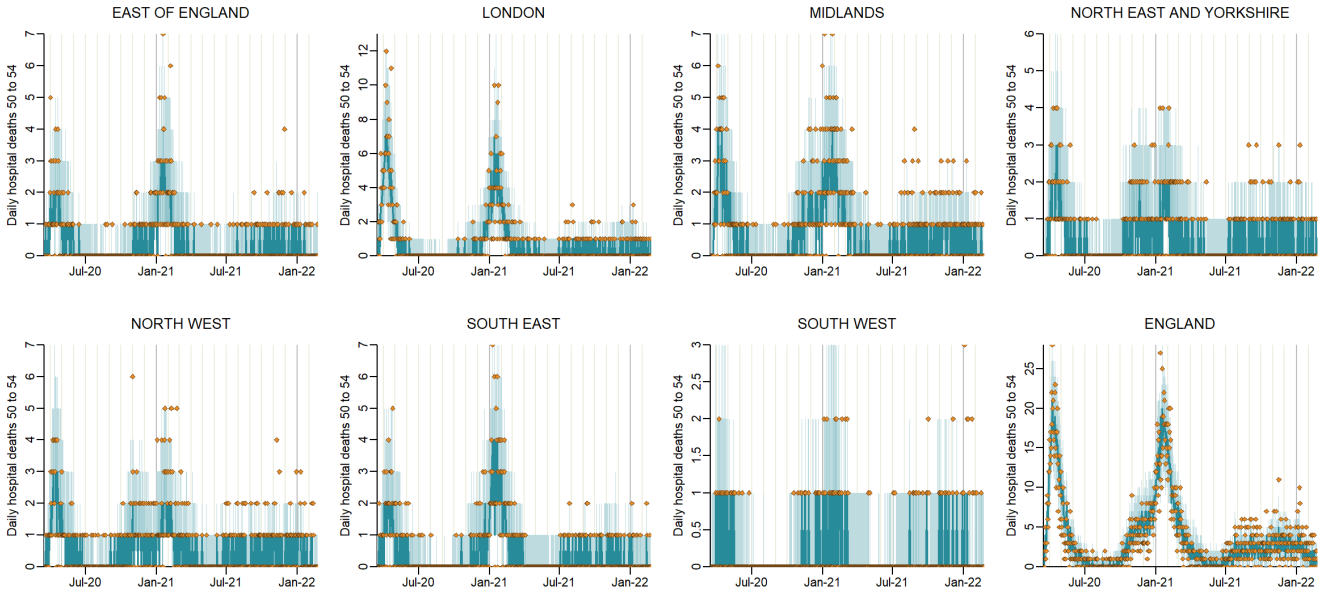


Figure A38: Fitted model trajectories to hospital deaths of 50 to 54 year-olds data by England NHS regions. Source: Perez-Guzman et al. 2023 [1].

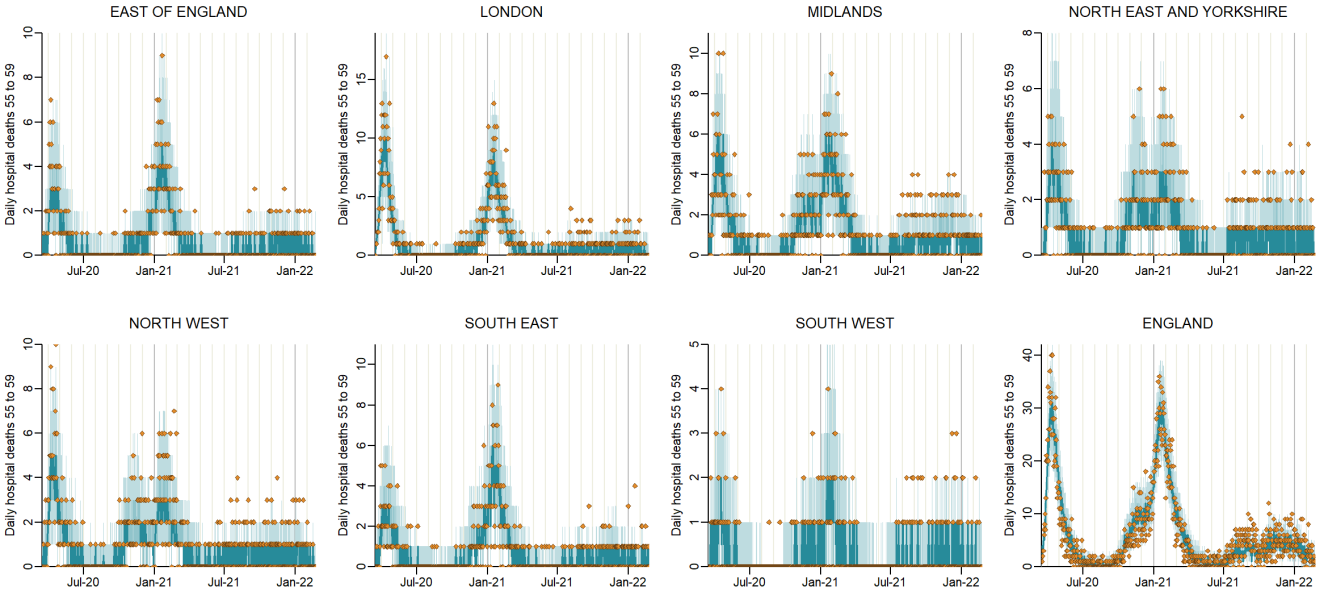


Figure A39: Fitted model trajectories to hospital deaths of 55 to 59 year-olds data by England NHS regions. Source: Perez-Guzman et al. 2023 [1].

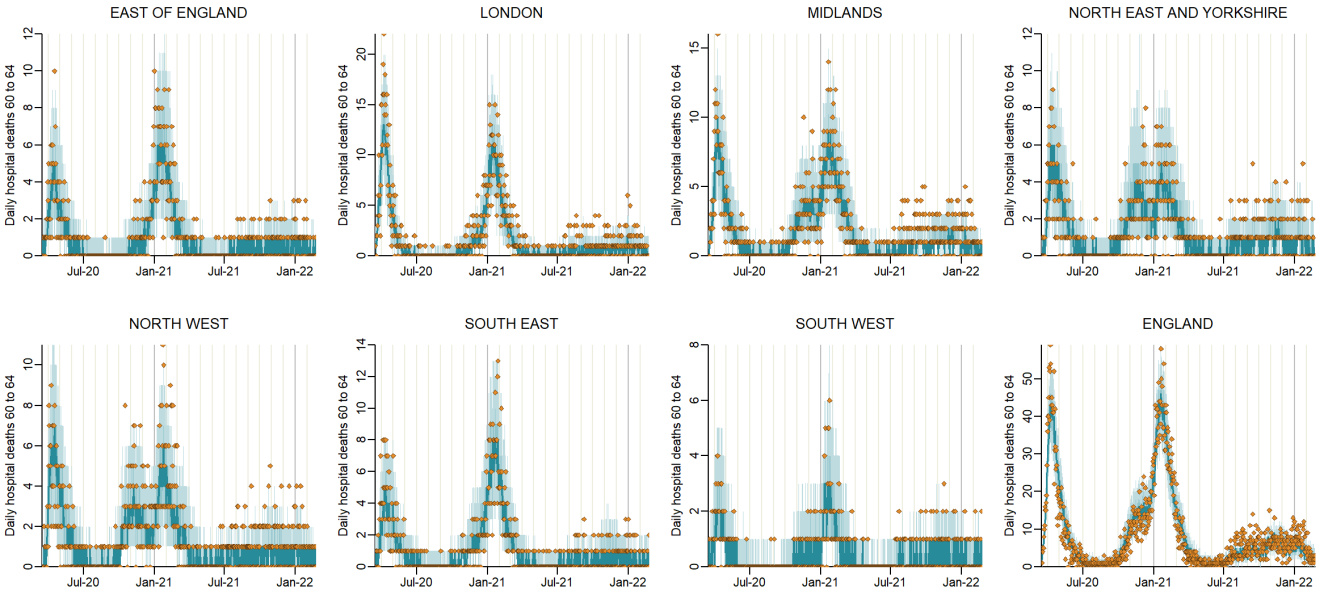


Figure A40: Fitted model trajectories to hospital deaths of 60 to 64 year-olds data by England NHS regions. Source: Perez-Guzman et al. 2023 [1].

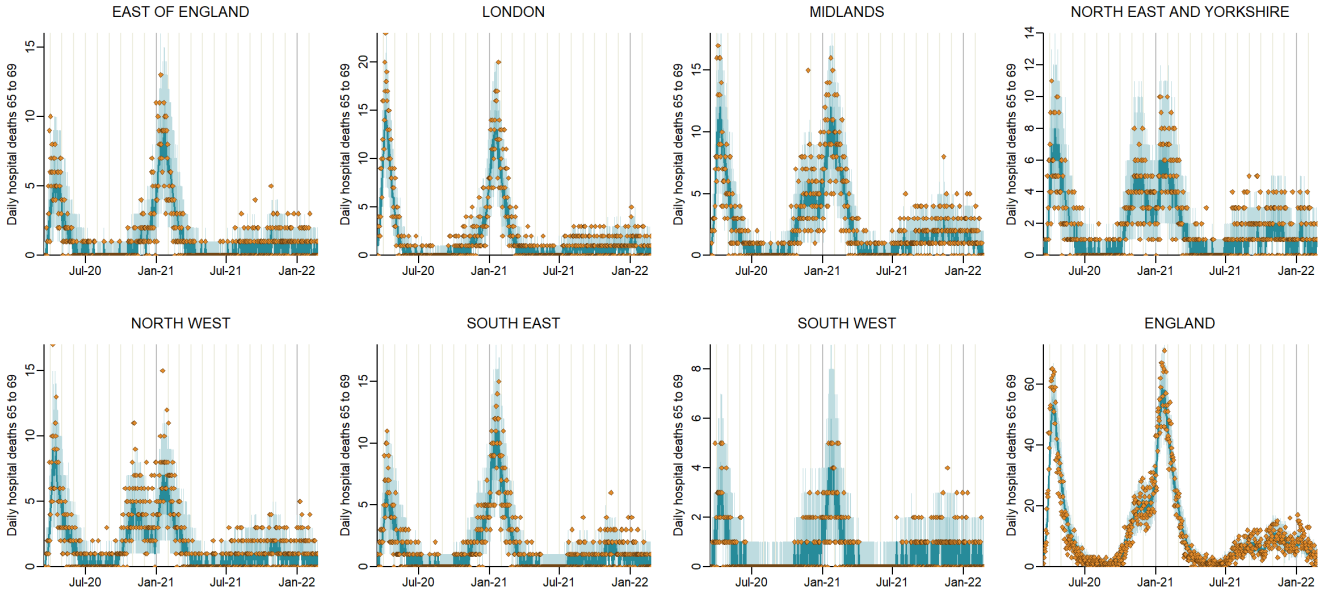


Figure A41: Fitted model trajectories to hospital deaths of 65 to 69 year-olds data by England NHS regions. Source: Perez-Guzman et al. 2023 [1].

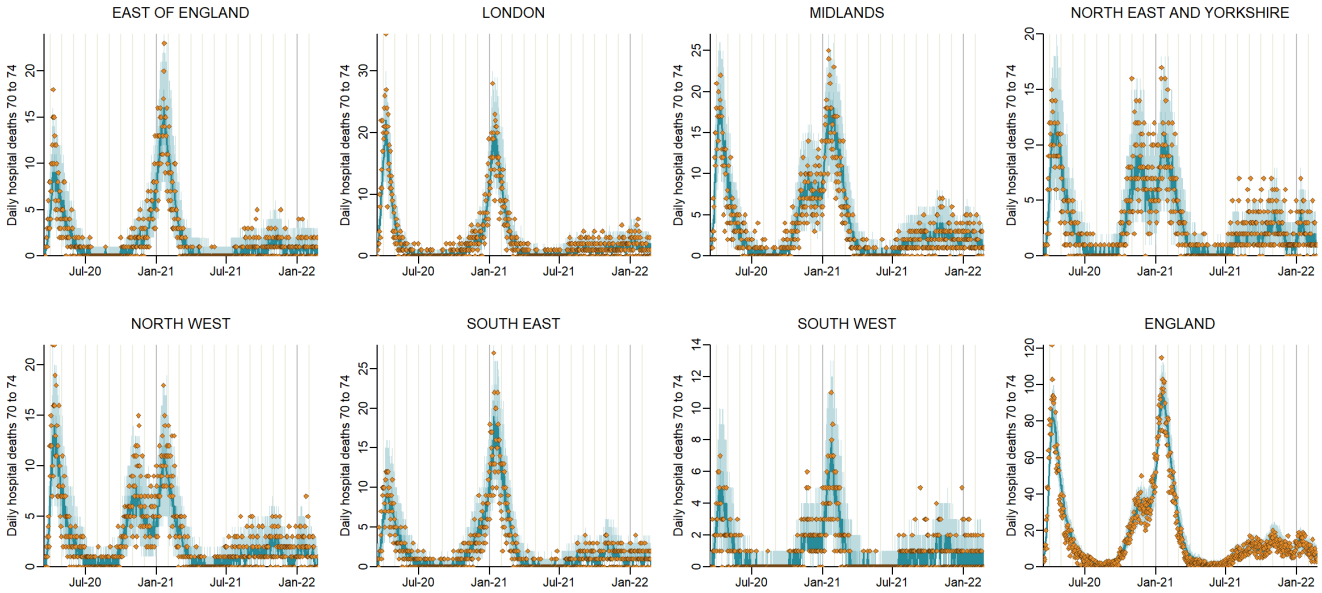


Figure A42: Fitted model trajectories to hospital deaths of 70 to 74 year-olds data by England NHS regions. Source: Perez-Guzman et al. 2023 [1].

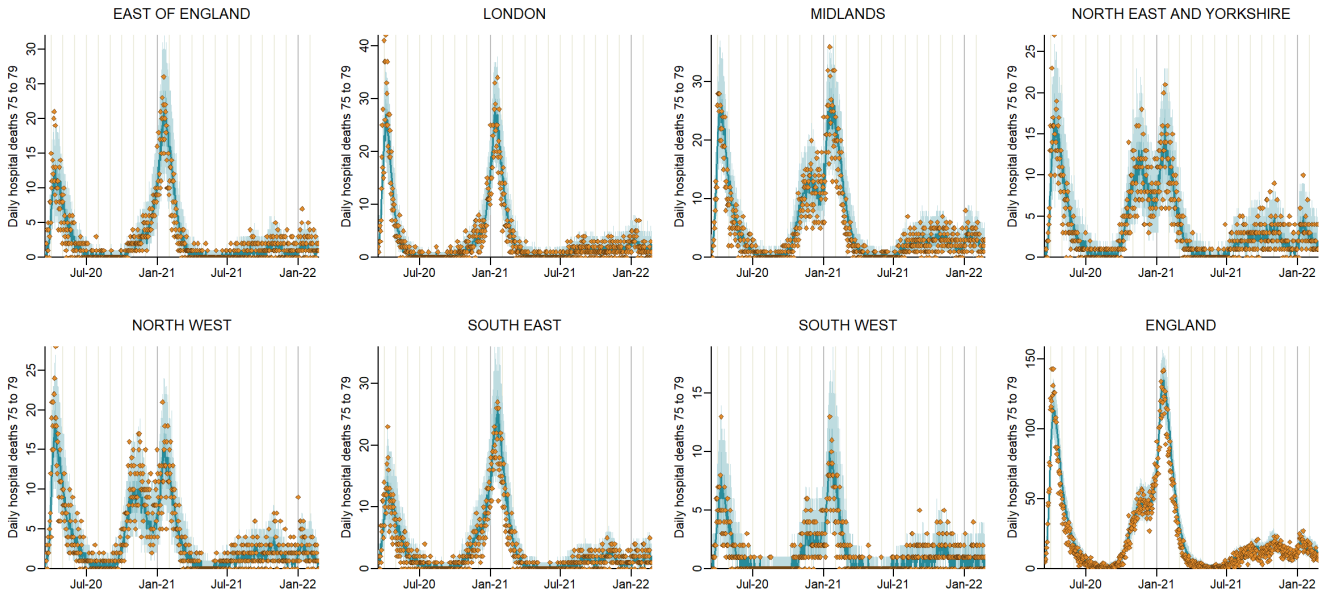


Figure A43: Fitted model trajectories to hospital deaths of 75 to 79 year-olds data by England NHS regions. Source: Perez-Guzman et al. 2023 [1].

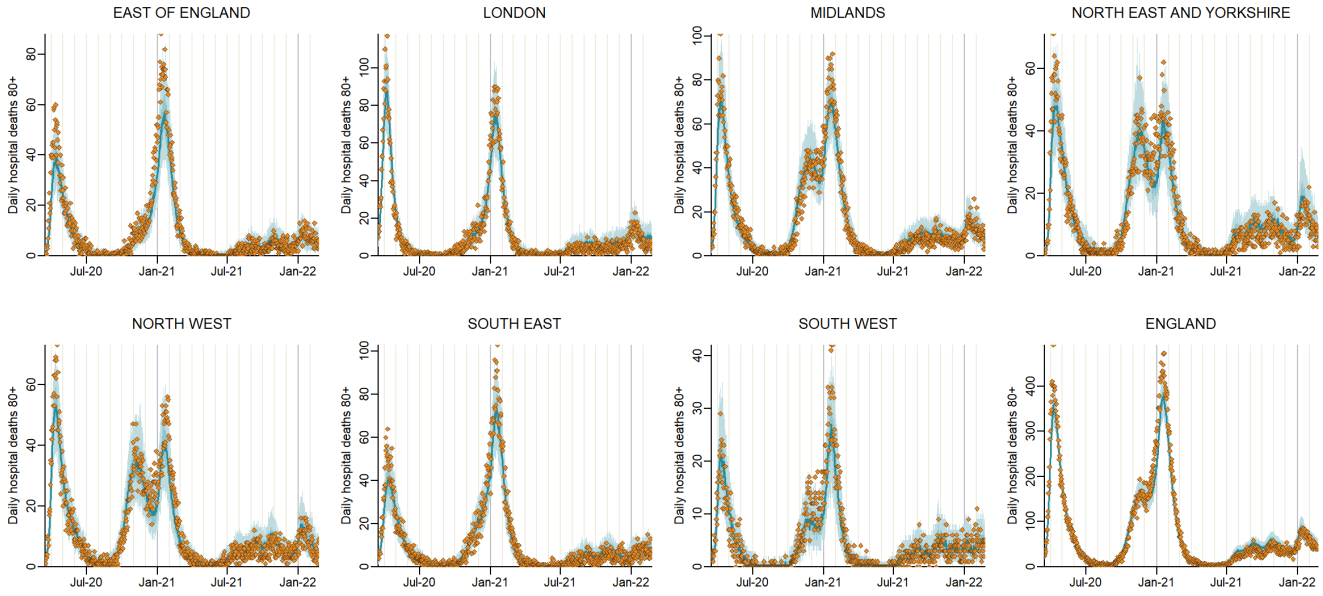


Figure A44: Fitted model trajectories to hospital deaths of 80 year-olds data by England NHS regions. Source: Perez-Guzman et al. 2023 [1].

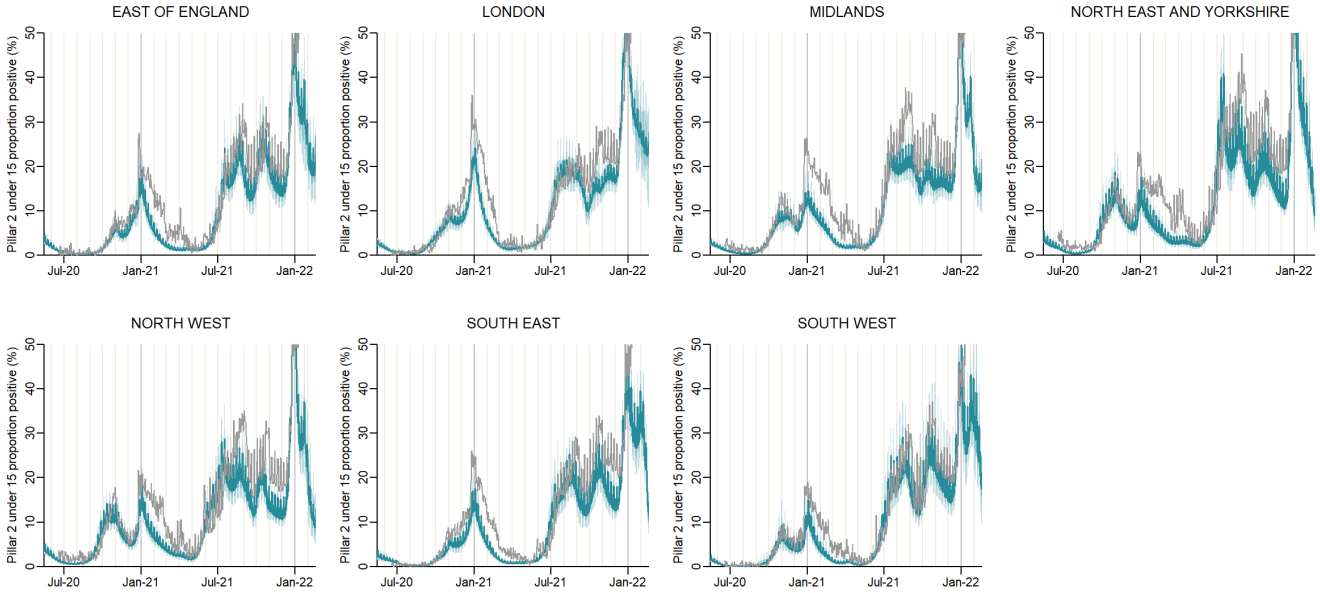


Figure A45: Fitted model trajectories to Pillar 2 cases of 0 to 14 year-olds data by England NHS regions. Source: Perez-Guzman et al. 2023 [1].

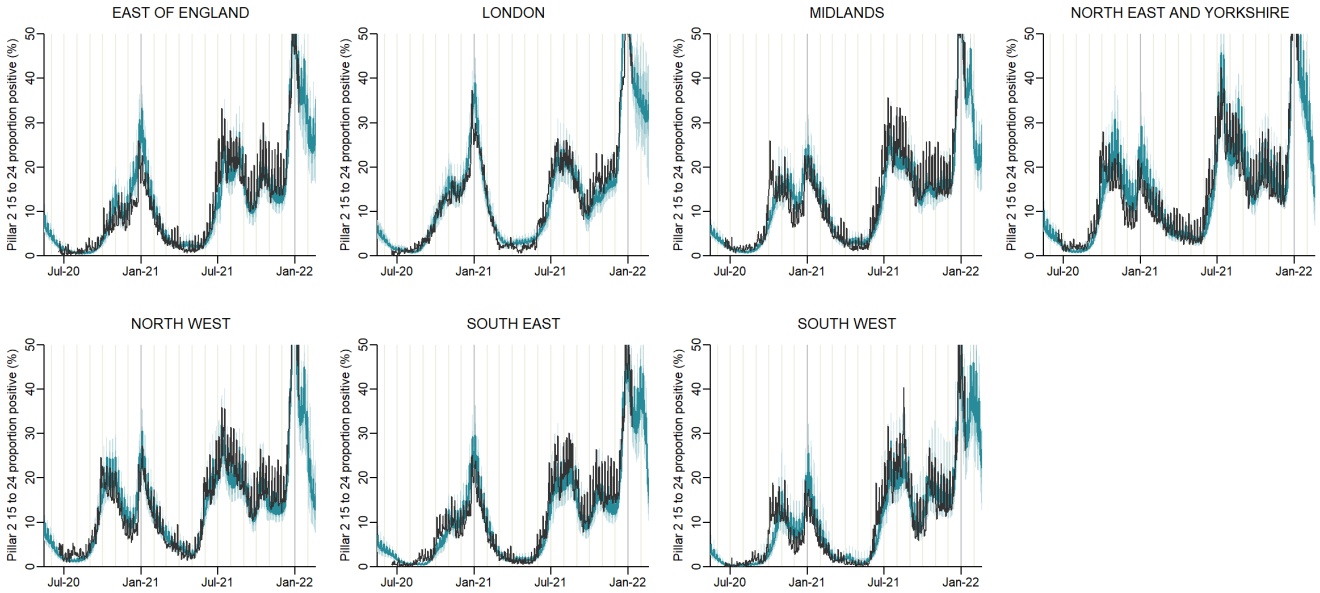


Figure A46: Fitted model trajectories to Pillar 2 cases of 15 to 24 year-olds data by England NHS regions. Source: Perez-Guzman et al. 2023 [1].

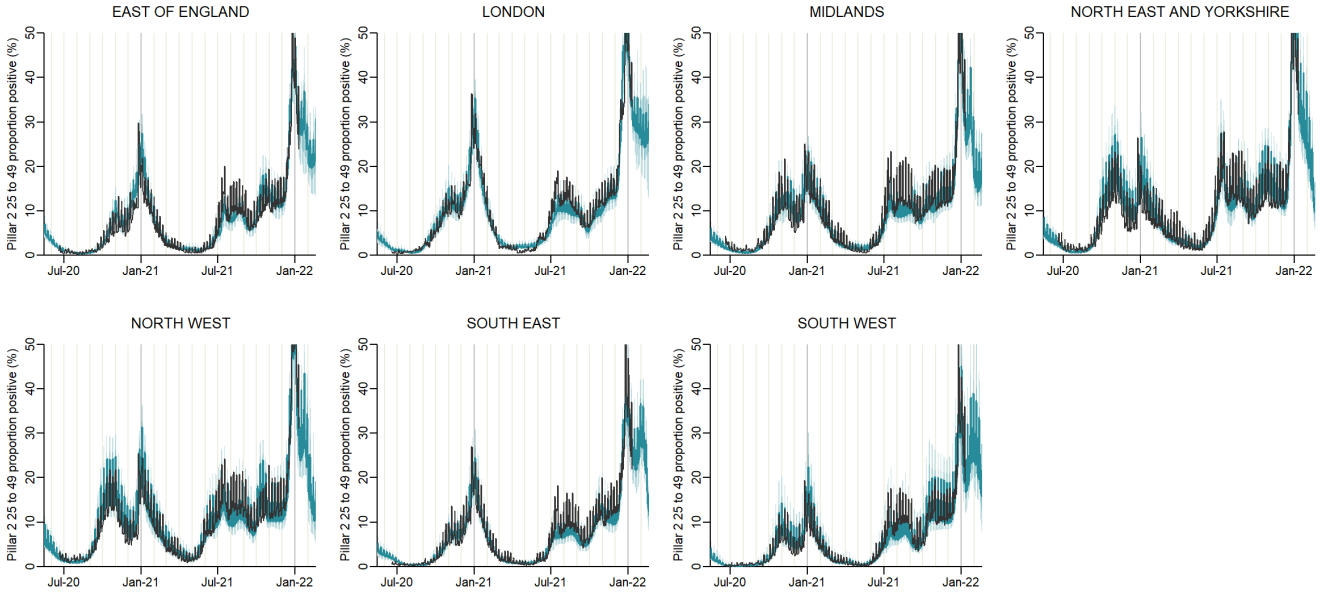


Figure A47: Fitted model trajectories to Pillar 2 cases of 25 to 49 year-olds data by England NHS regions. Source: Perez-Guzman et al. 2023 [1].

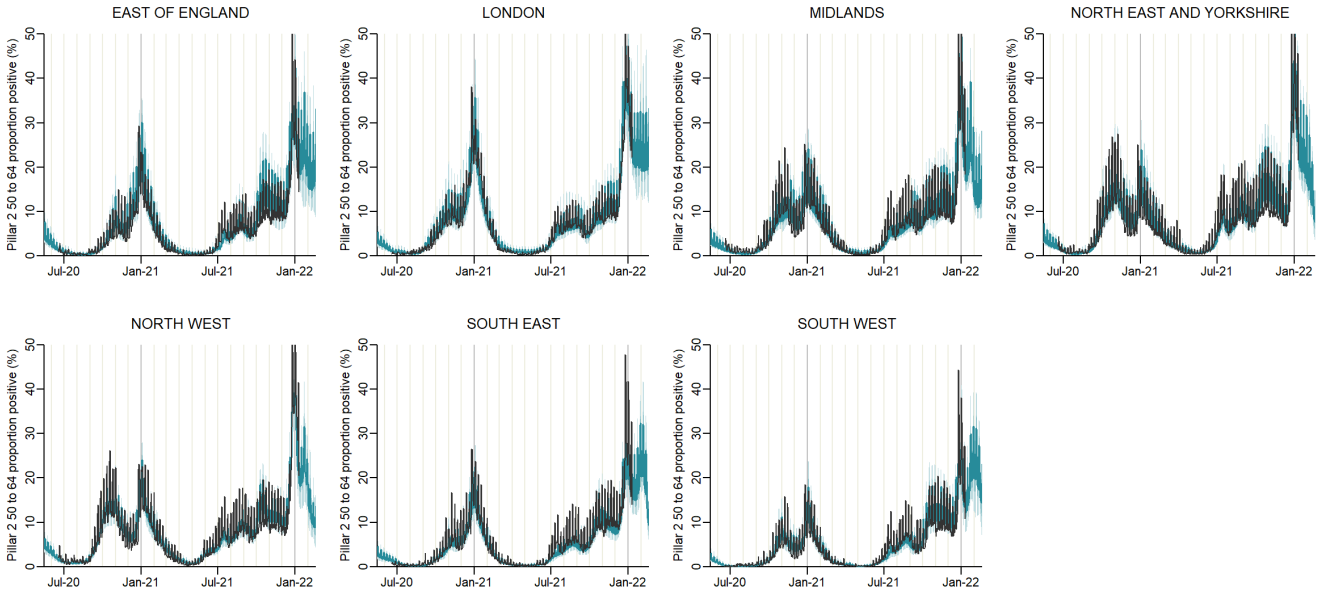


Figure A48: Fitted model trajectories to Pillar 2 cases of 50 to 64 year-olds data by England NHS regions. Source: Perez-Guzman et al. 2023 [1].

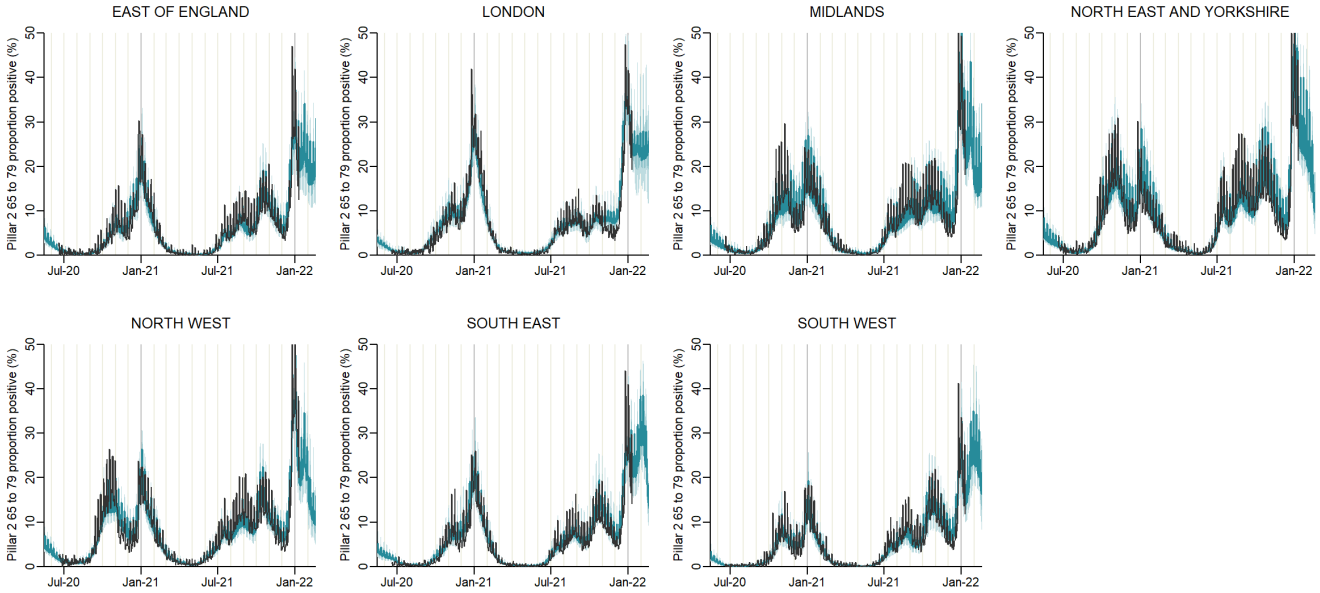


Figure A49: Fitted model trajectories to Pillar 2 cases of 65 to 79 year-olds data by England NHS regions. Source: Perez-Guzman et al. 2023 [1].

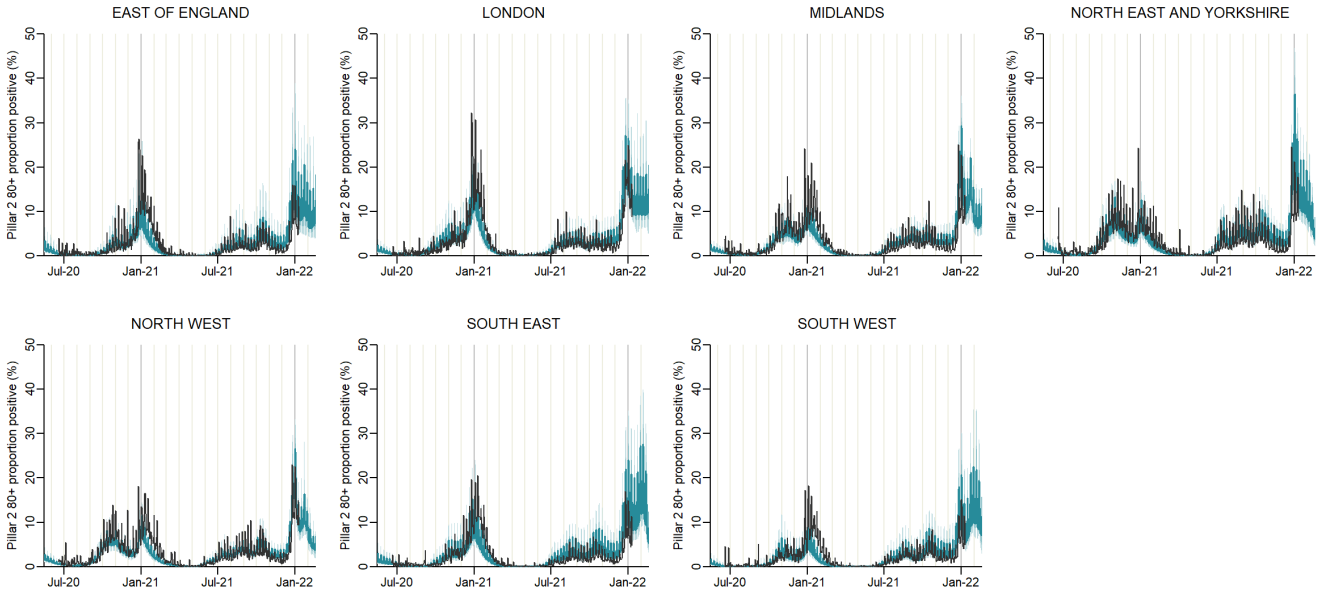


Figure A50: Fitted model trajectories to Pillar 2 cases of 80+ year-olds data by England NHS regions. Source: Perez-Guzman et al. 2023 [1].

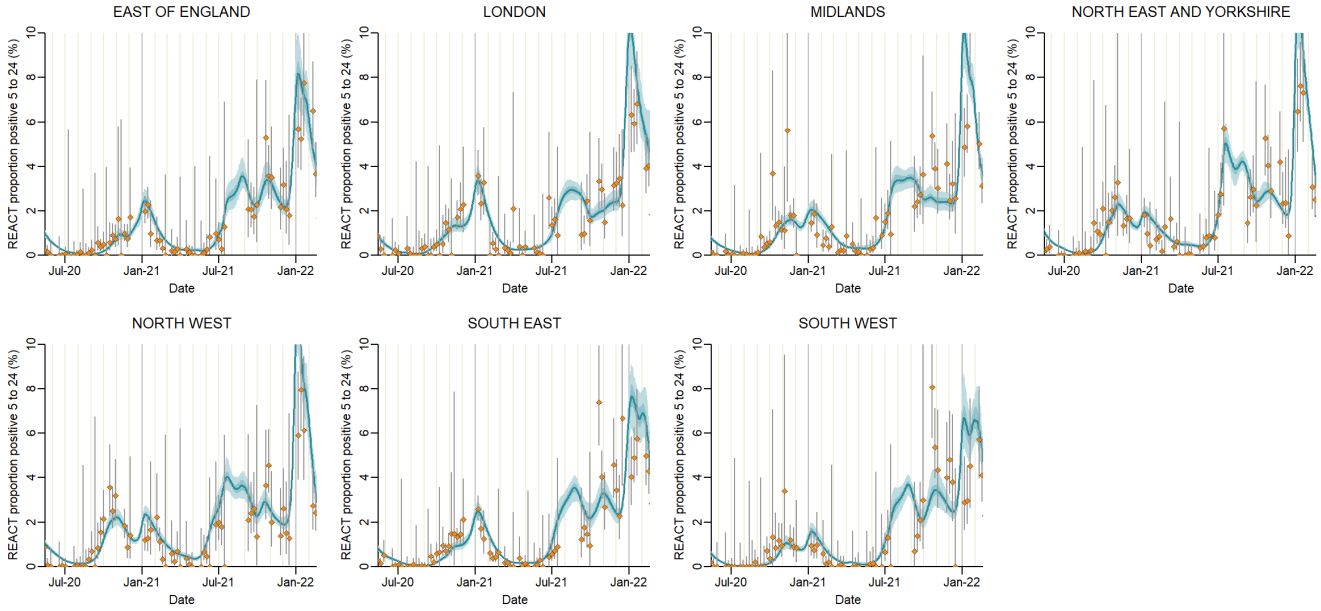


Figure A51: Fitted model trajectories to REACT-1 infection prevalence of 5 to 24 year-olds data by England NHS regions. Source: Perez-Guzman et al. 2023 [1].

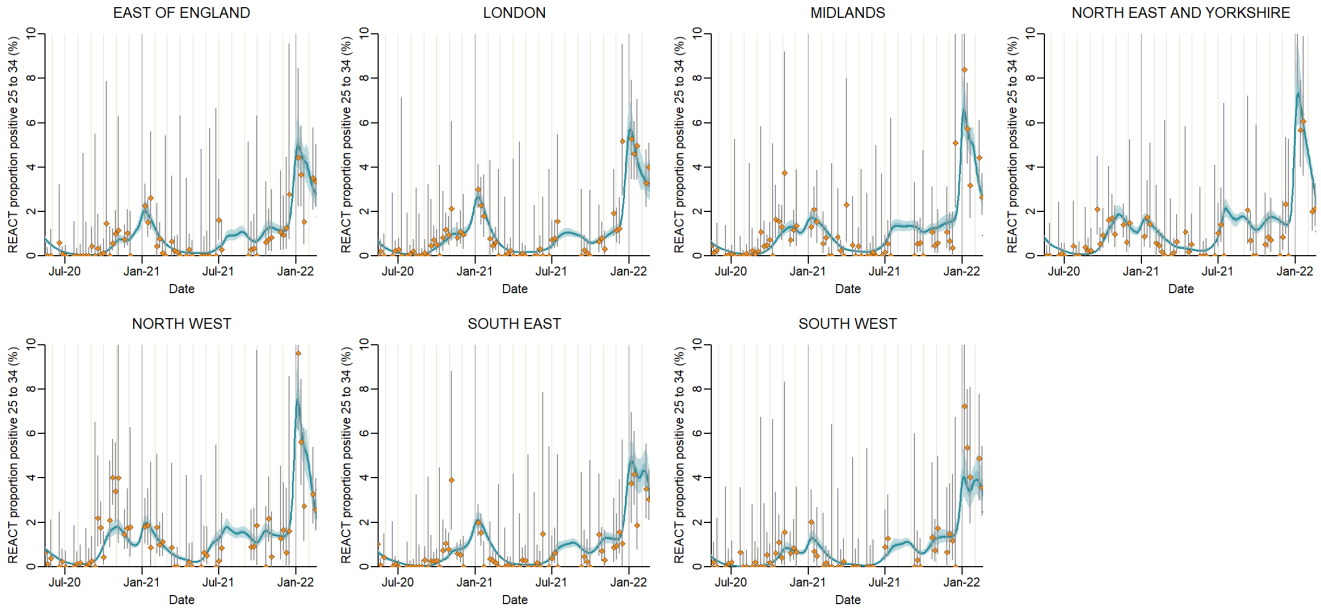


Figure A52: Fitted model trajectories to REACT-1 infection prevalence of 25 to 34 year-olds data by England NHS regions. Source: Perez-Guzman et al. 2023 [1].

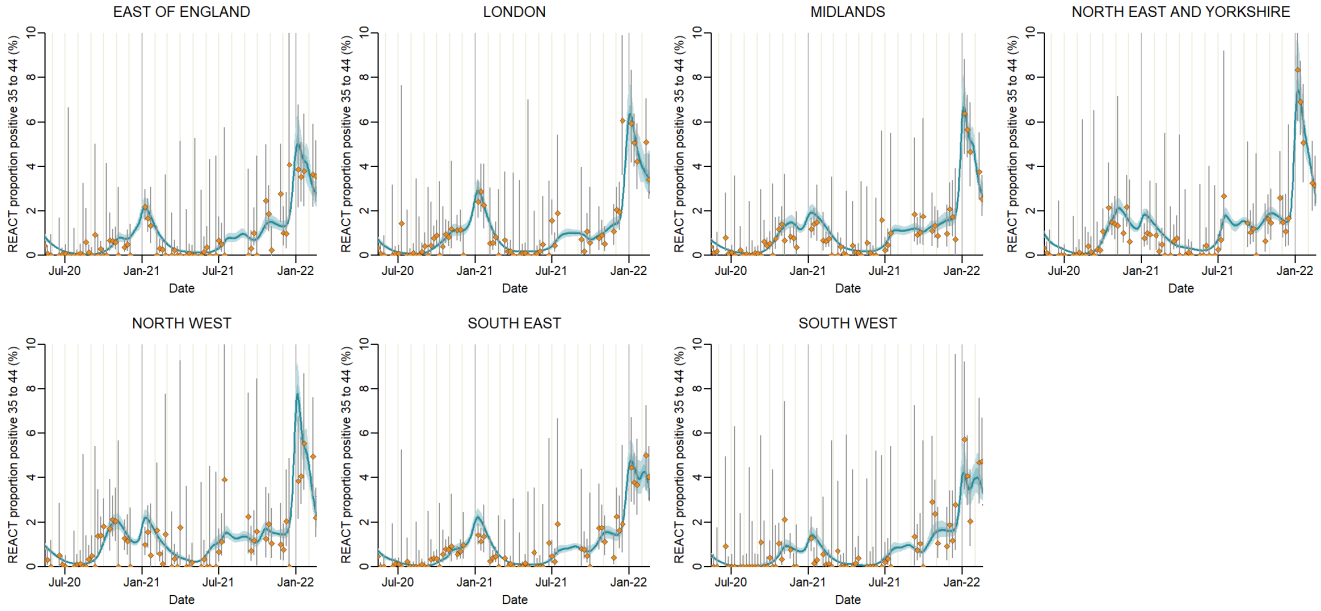


Figure A53: Fitted model trajectories to REACT-1 infection prevalence of 35 to 44 year-olds data by England NHS regions. Source: Perez-Guzman et al. 2023 [1].

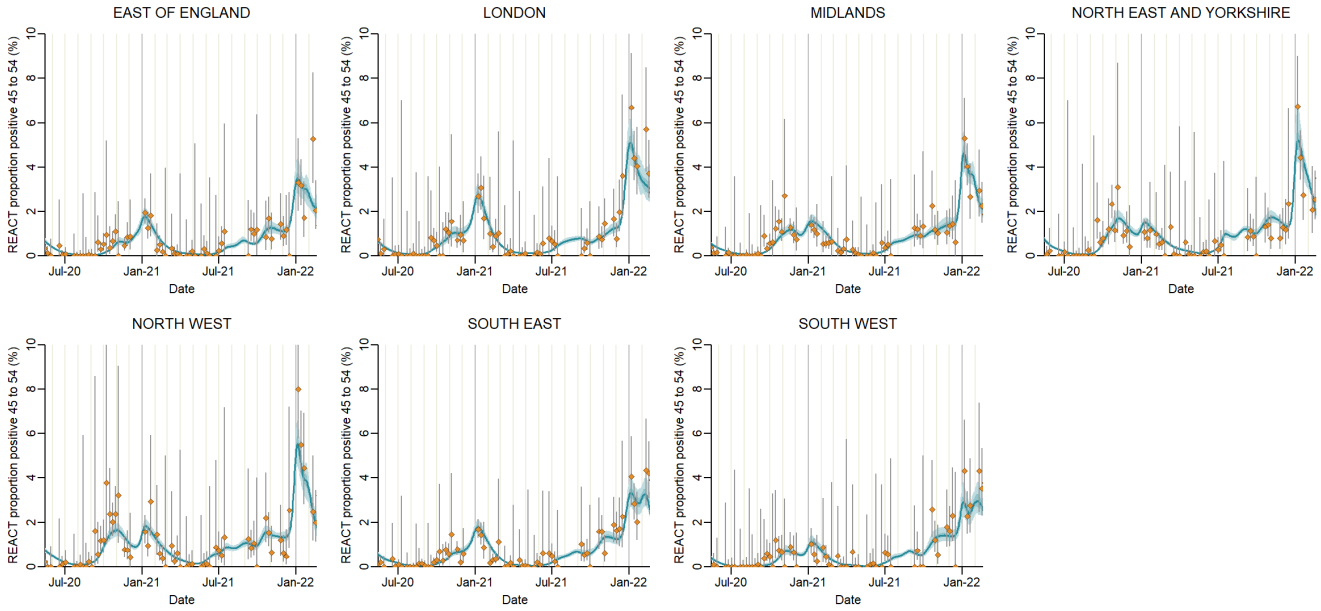


Figure A54: Fitted model trajectories to REACT-1 infection prevalence of 45 to 54 year-olds data by England NHS regions. Source: Perez-Guzman et al. 2023 [1].

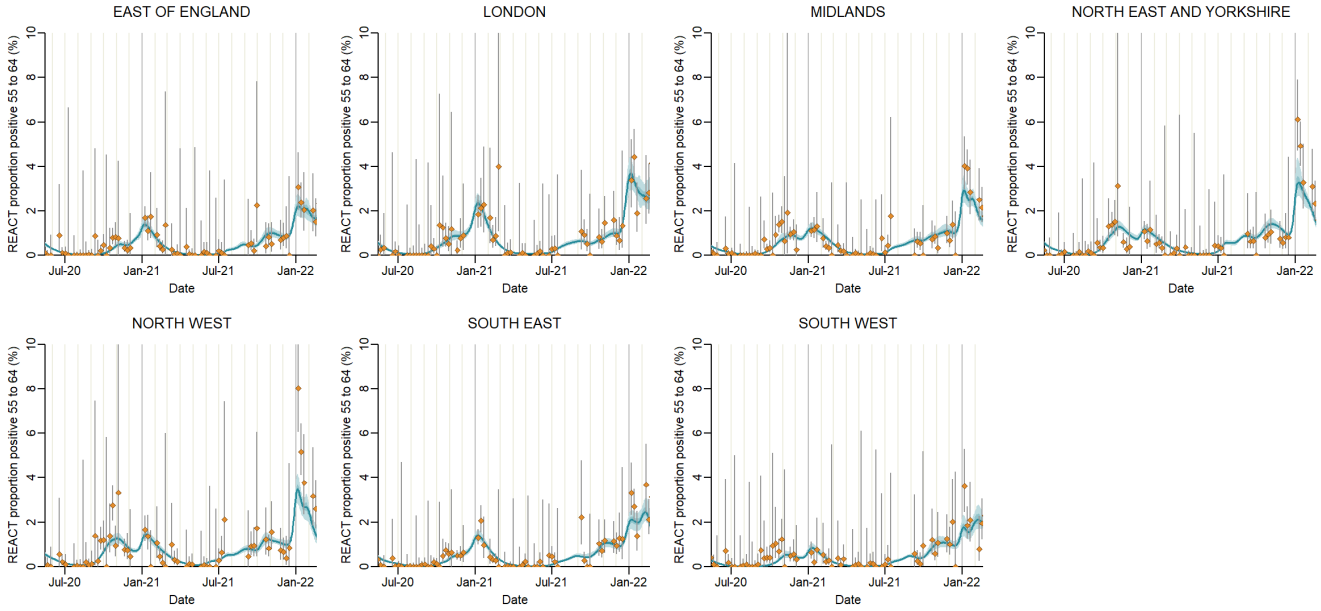


Figure A55: Fitted model trajectories to REACT-1 infection prevalence of 55 to 64 year-olds data by England NHS regions. Source: Perez-Guzman et al. 2023 [1].

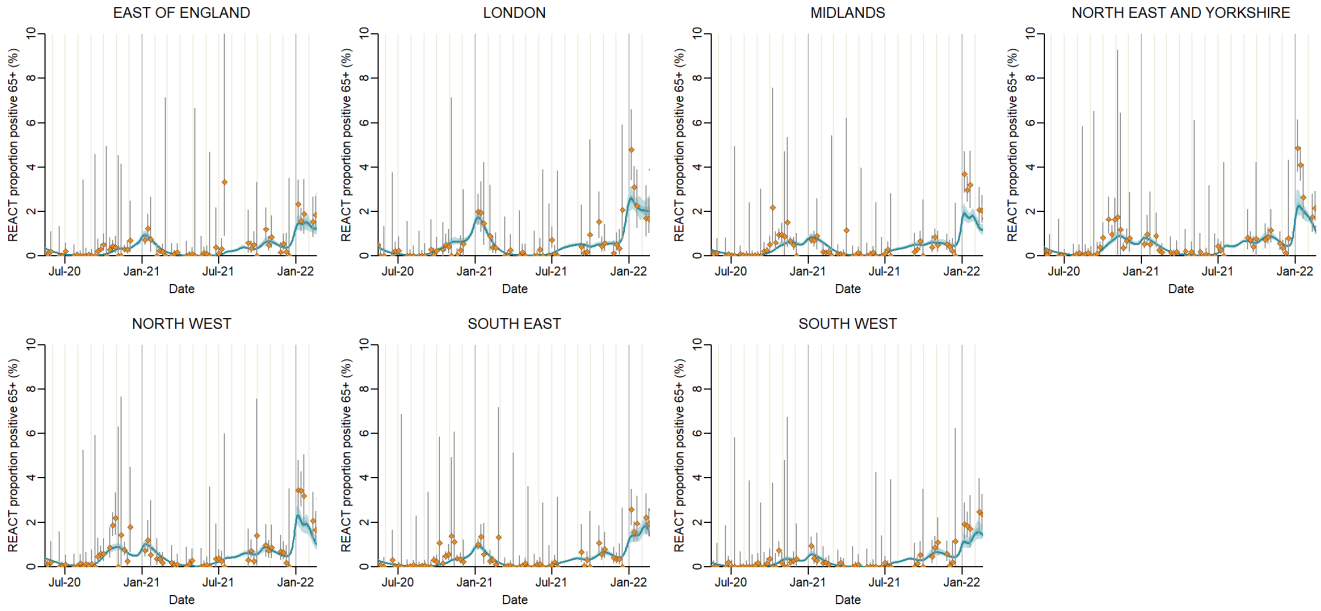


Figure A56: Fitted model trajectories to REACT-1 infection prevalence of 65+ year-olds data by England NHS regions. Source: Perez-Guzman et al. 2023 [1].

A.5 Fitted parameters

This sections presents plots of the posterior distribution of fitted parameters, pre-tuned MCMC traceplots and final fitted pMCMC traceplots by NHS England region.

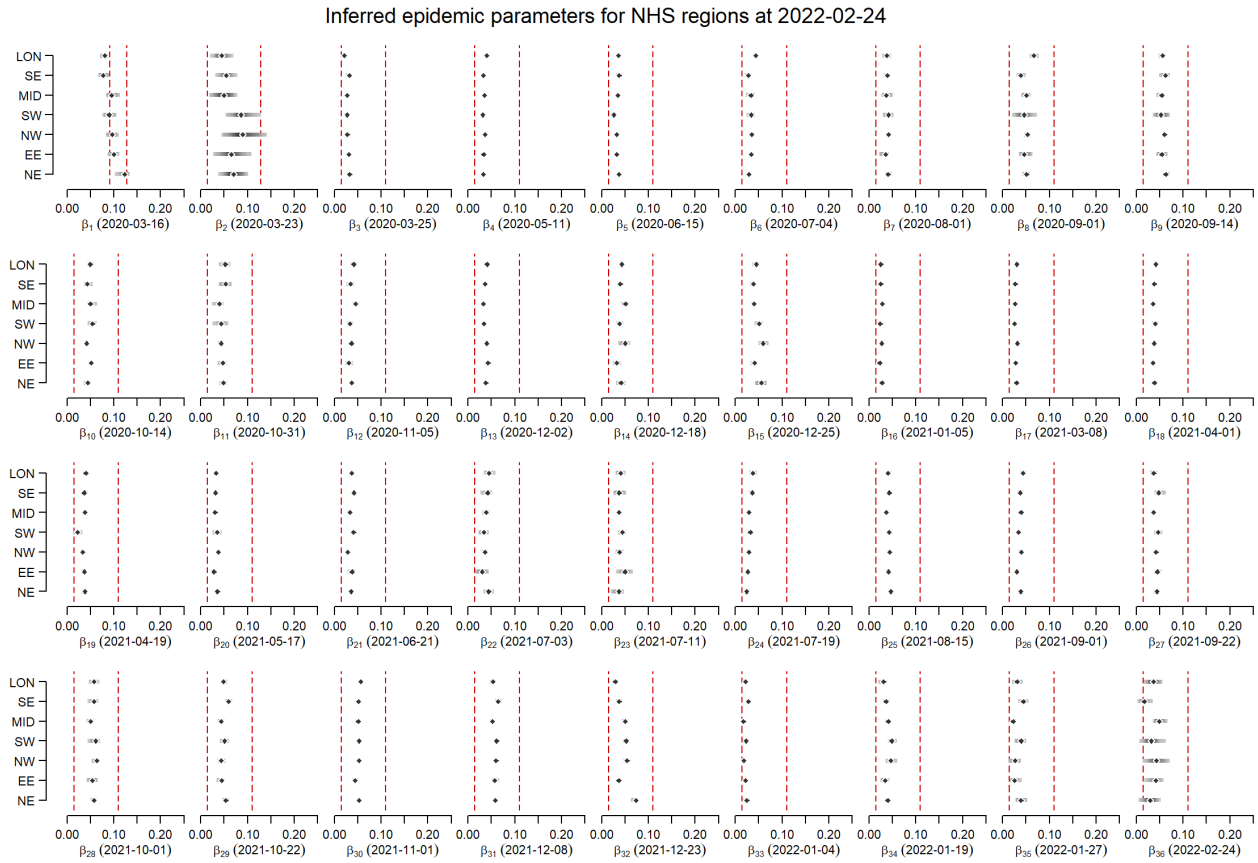


Figure A57: Posterior distributions of $\beta(t)$ values for all change-points (piece-wise linear form) by England NHS regions. Red lines show prior distributions; where no prior lines are shown, a uniform (non-informative) prior was assumed. Source: Perez-Guzman et al. 2023 [1].

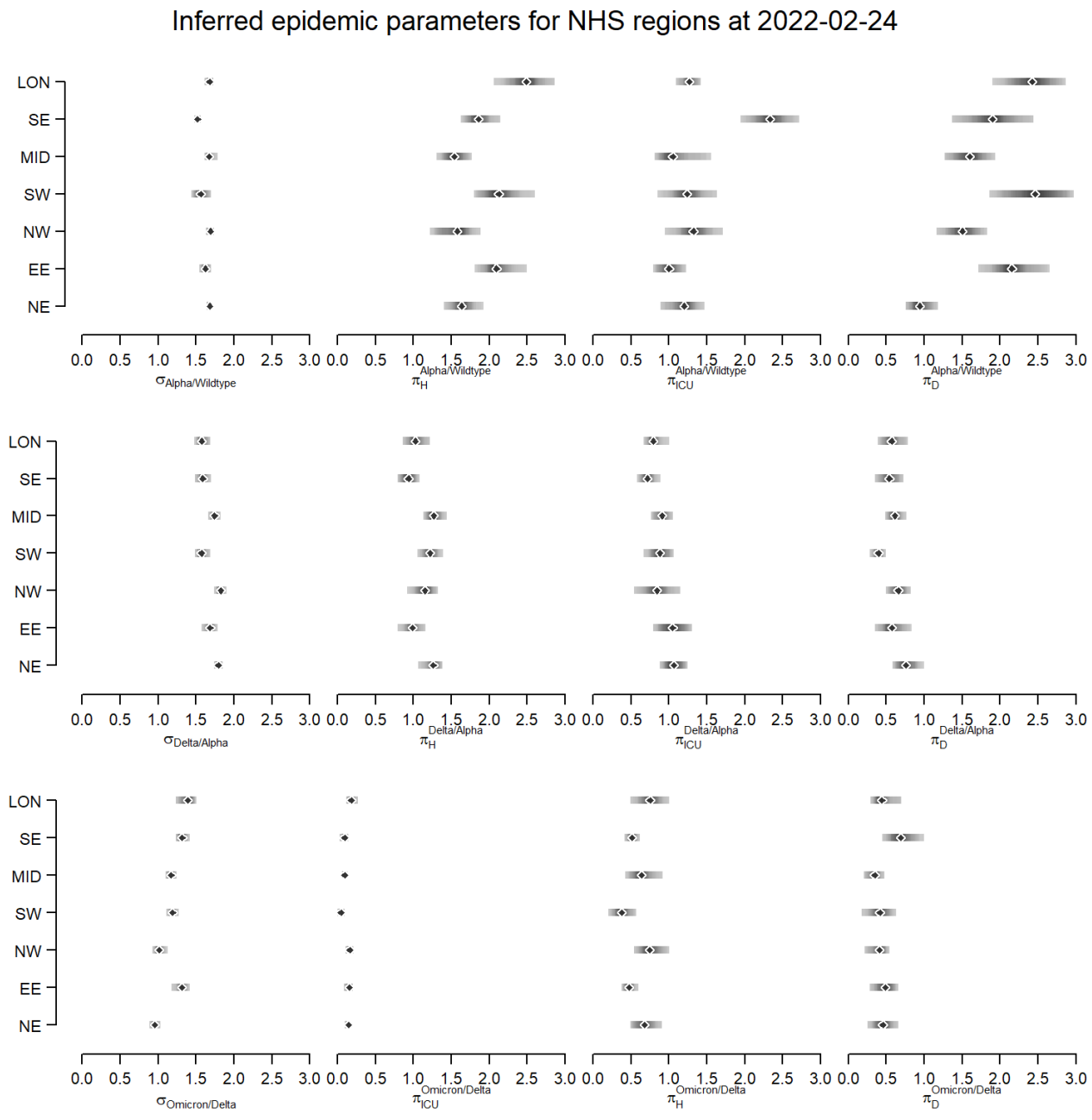


Figure A58: Posterior distributions of variant-specific transmission and severity parameters. A uniform (non-informative) prior was assumed for all parameters. Source: Perez-Guzman et al. 2023 [1].

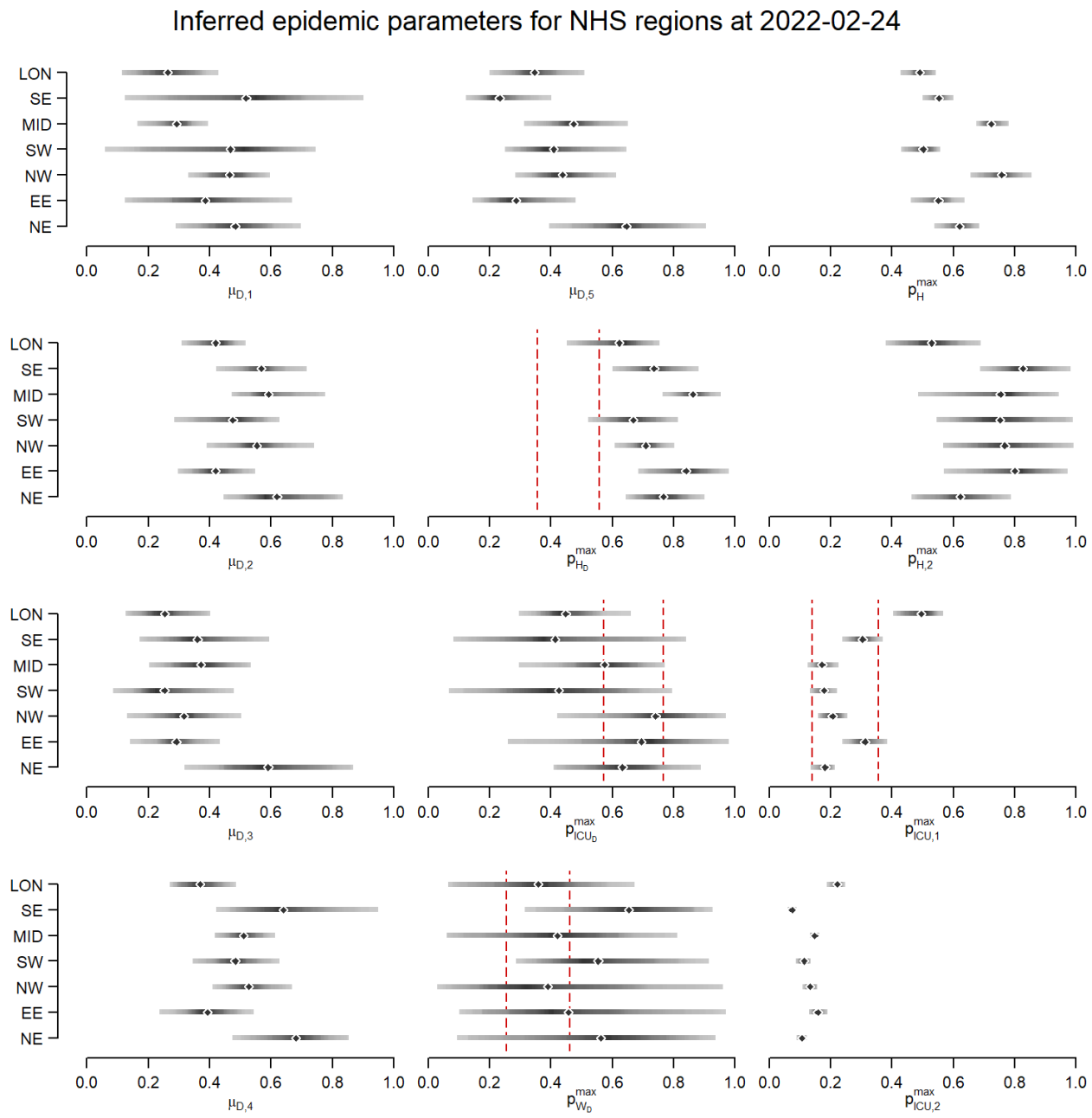


Figure A59: Posterior distributions of time-varying severity multiplier parameters, with piece-wise linear form. Red lines show prior distributions; where no prior lines are shown, a uniform (non-informative) prior was assumed. Source: Perez-Guzman et al. 2023 [1].

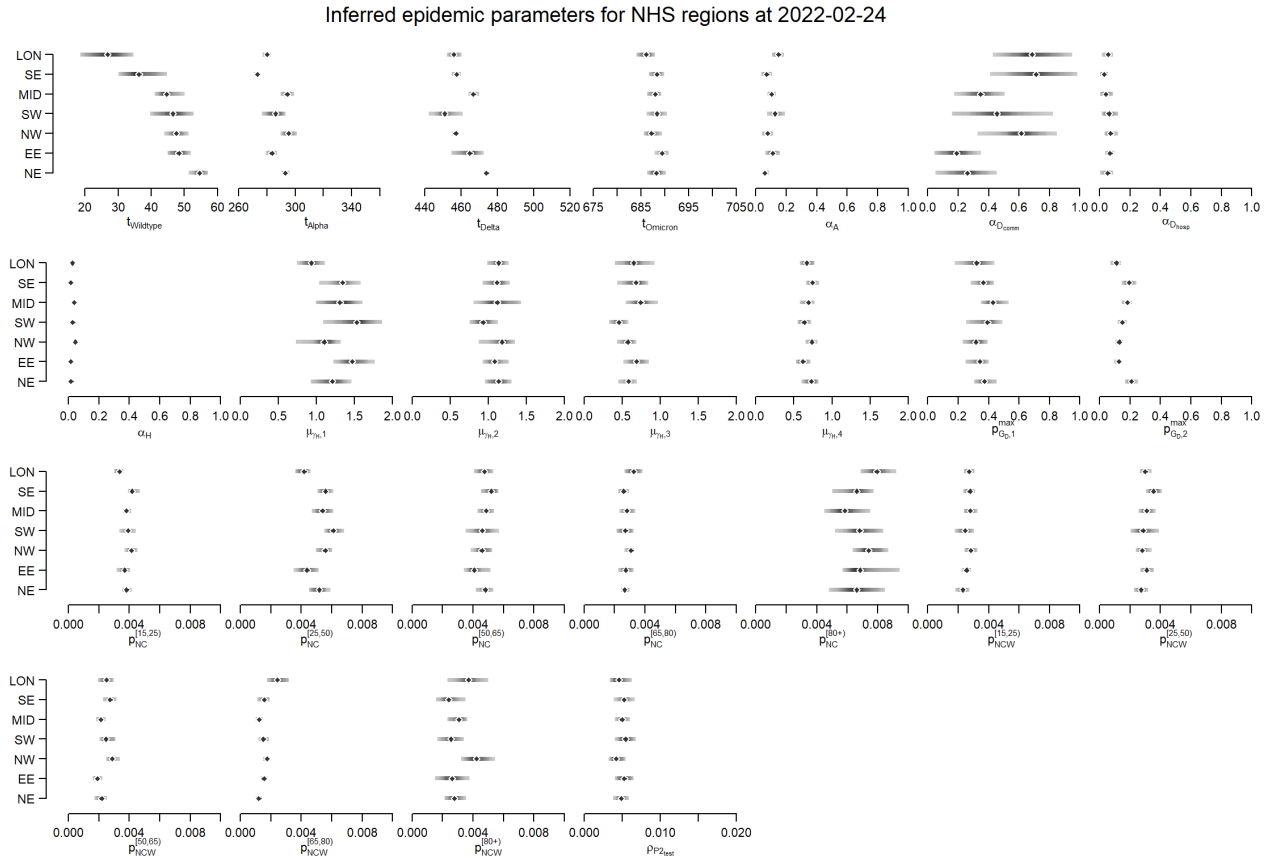


Figure A60: Posterior distributions of all other fitted parameters. Red lines show prior distributions; where no prior lines are shown, a uniform (non-informative) prior was assumed. Source: Perez-Guzman et al. 2023 [1].



Figure A61: MCMC traceplots for model fit to East of England. Source: Perez-Guzman et al. 2023 [1].

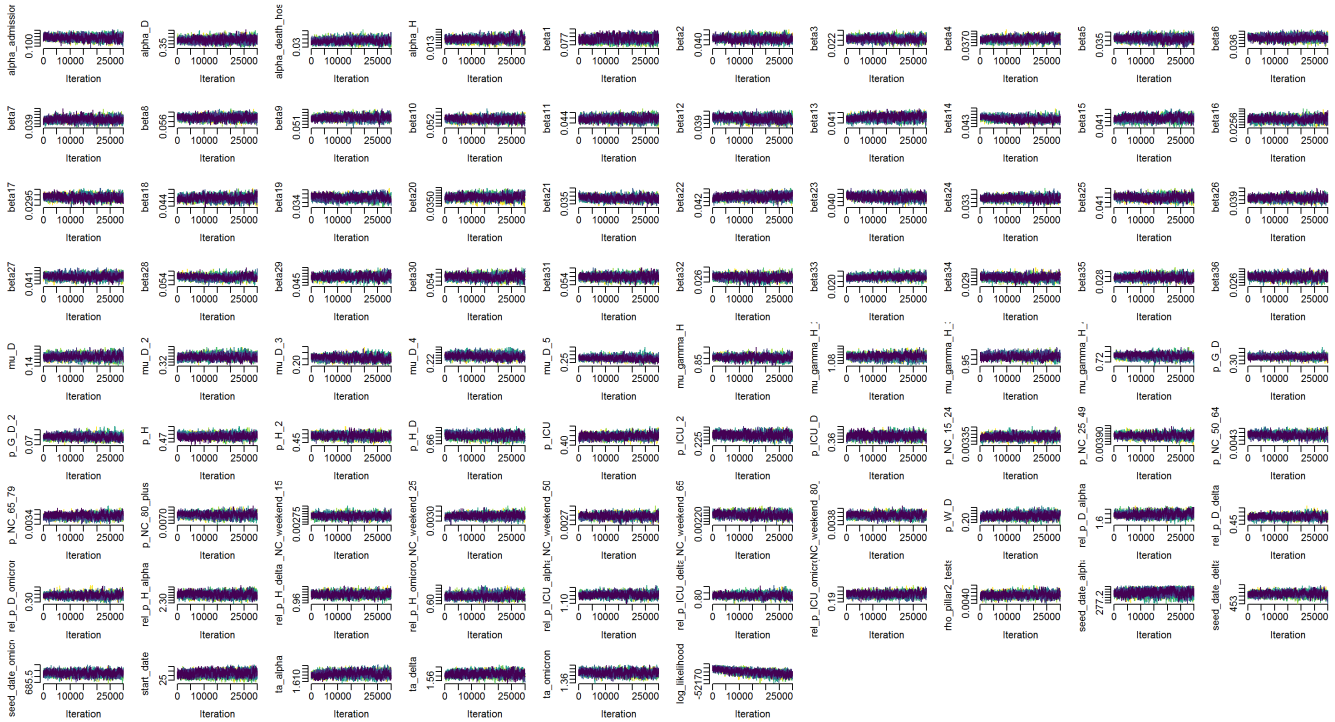


Figure A62: MCMC traceplots for model fit to London. Source: Perez-Guzman et al. 2023 [1].

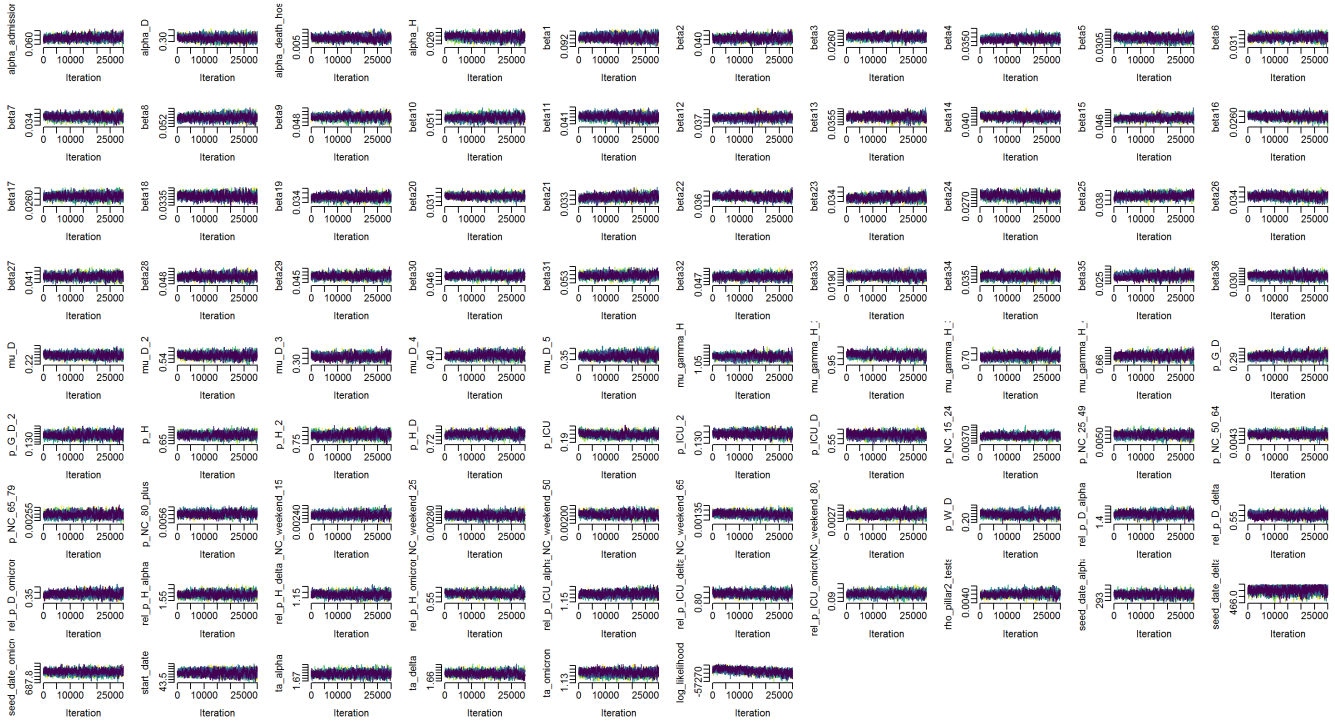


Figure A63: MCMC traceplots for model fit to the Midlands. Source: Perez-Guzman et al. 2023 [1].

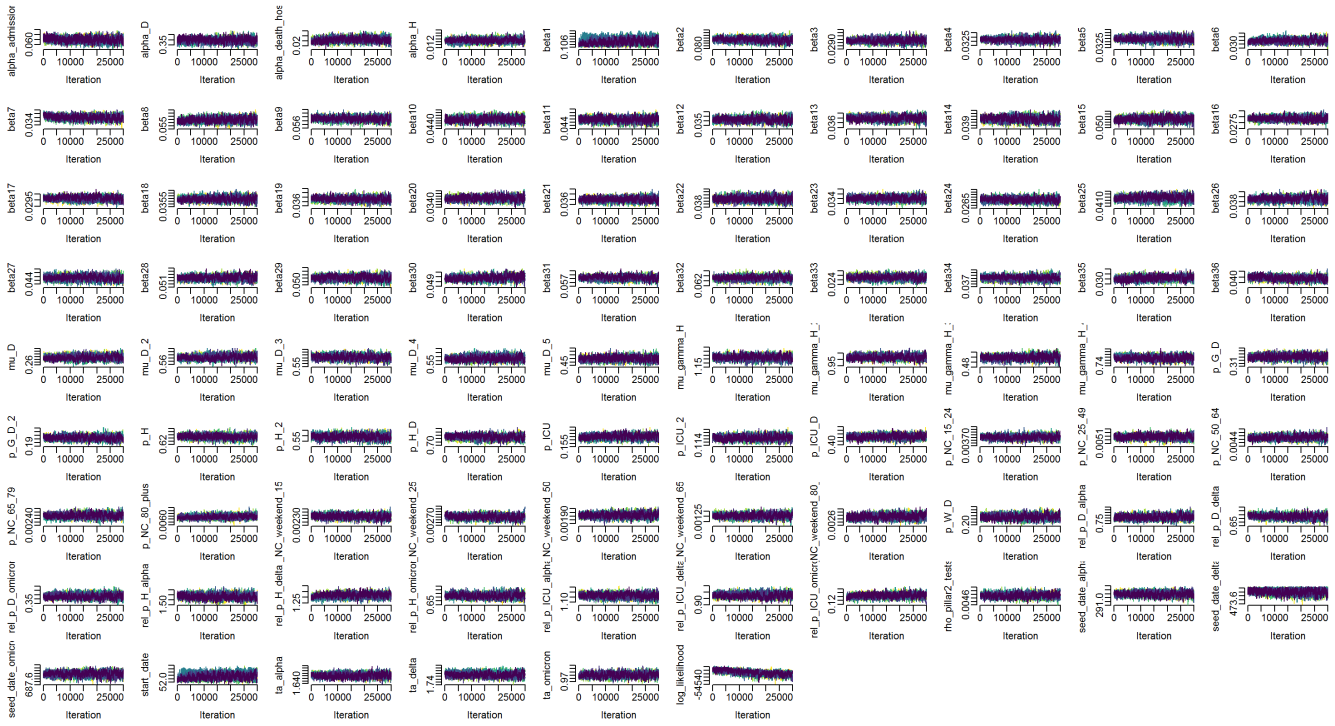


Figure A64: MCMC traceplots for model fit to North East and Yorkshire. Source: Perez-Guzman et al. 2023 [1].

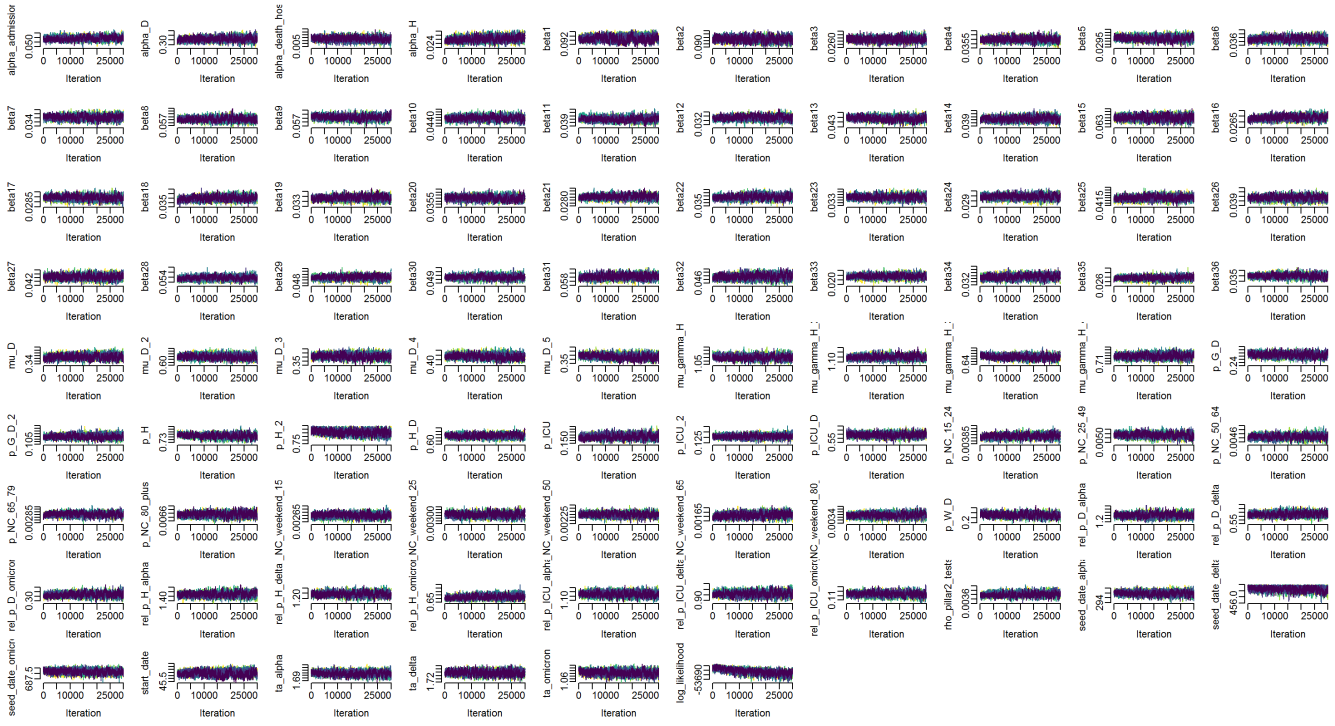


Figure A65: MCMC traceplots for model fit to North West. Source: Perez-Guzman et al. 2023 [1].

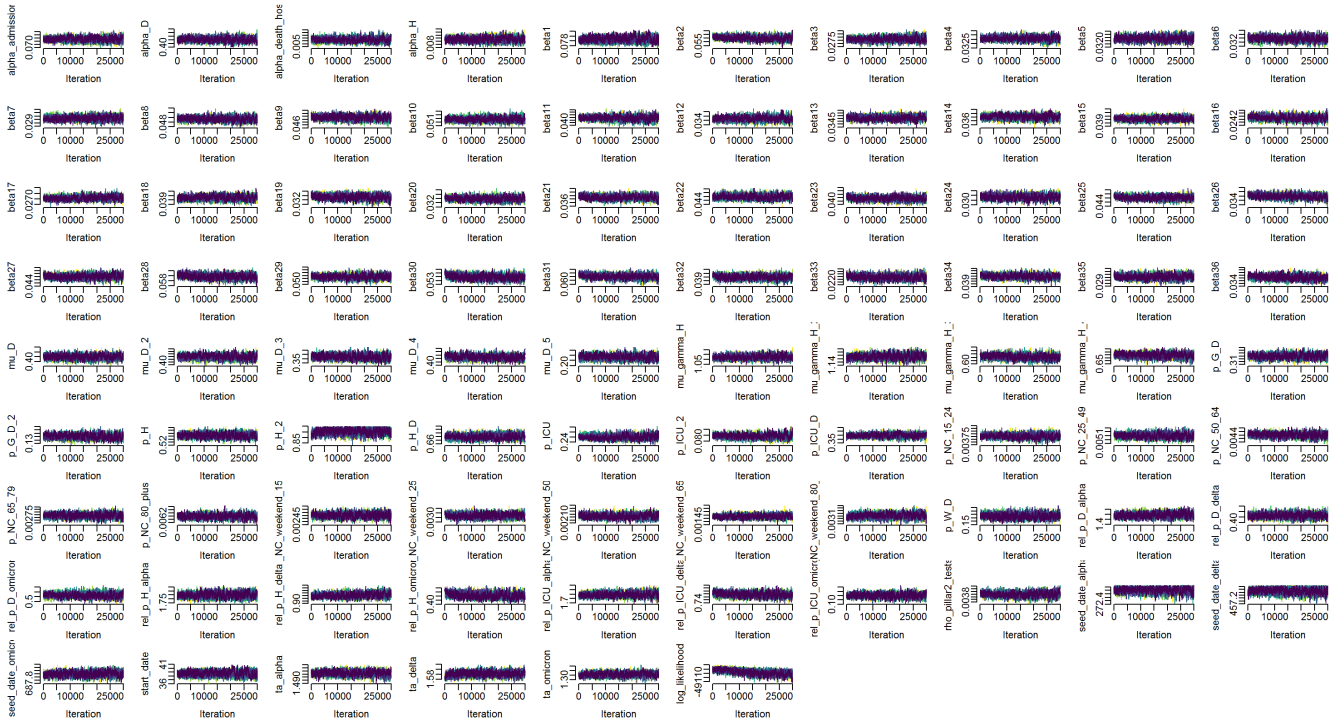


Figure A66: MCMC traceplots for model fit to South East. Source: Perez-Guzman et al. 2023 [1].

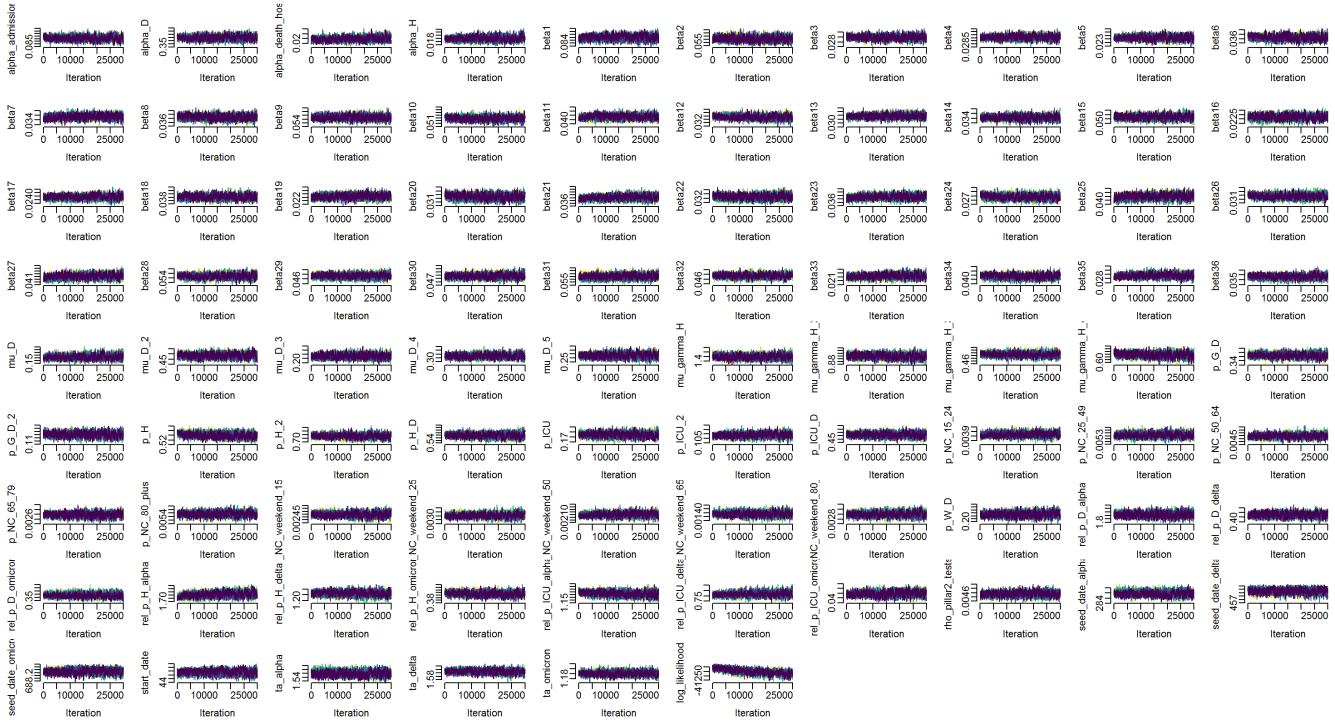


Figure A67: MCMC traceplots for model fit to South West. Source: Perez-Guzman et al. 2023 [1].

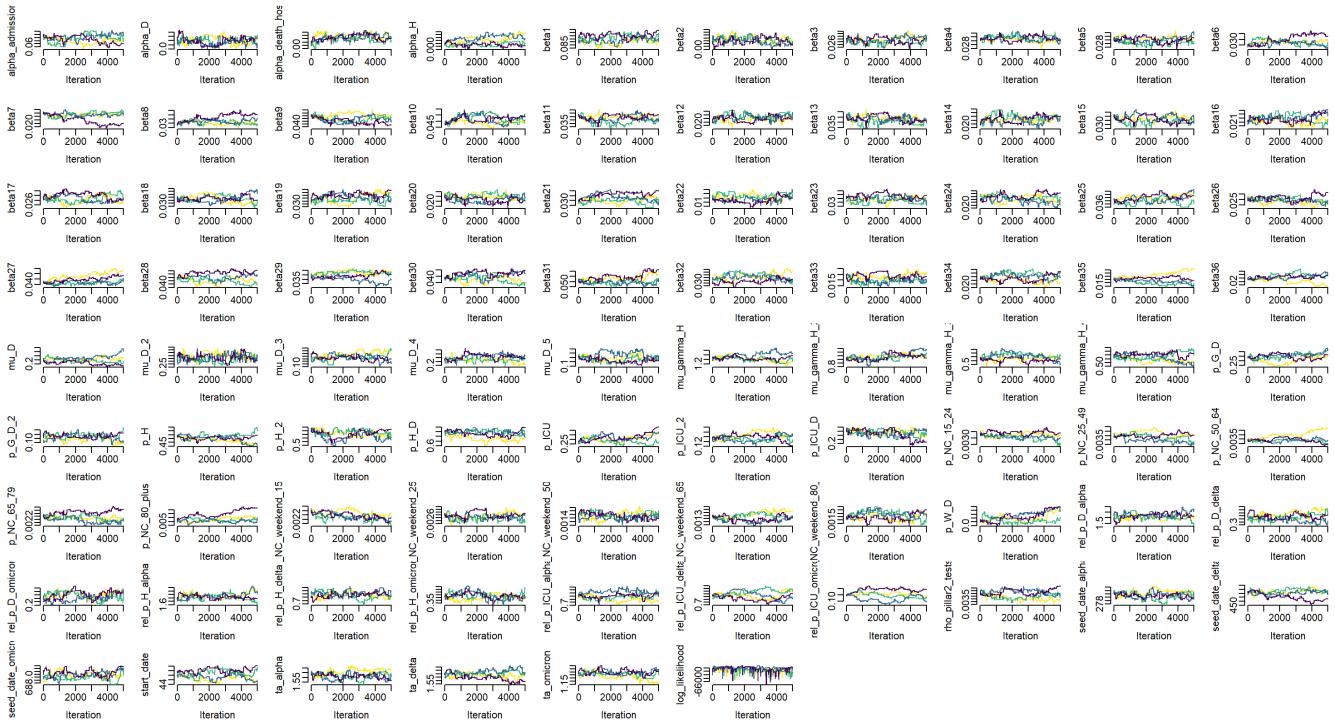


Figure A68: pMCMC traceplots for model fit to East of England. Source: Perez-Guzman et al. 2023 [1].



Figure A69: pMCMC traceplots for model fit to London. Source: Perez-Guzman et al. 2023 [1].

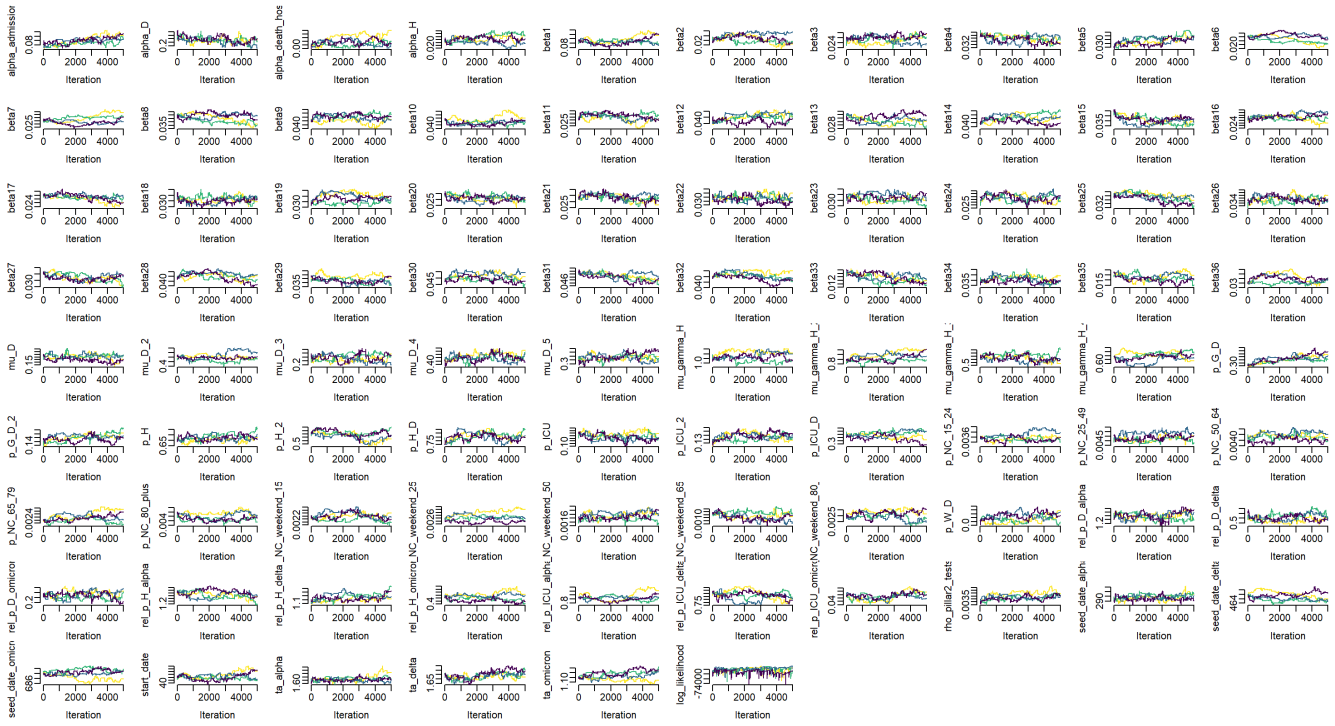


Figure A70: pMCMC traceplots for model fit to the Midlands. Source: Perez-Guzman et al. 2023 [1].



Figure A71: pMCMC traceplots for model fit to North East and Yorkshire. Source: Perez-Guzman et al. 2023 [1].



Figure A72: pMCMC traceplots for model fit to North West. Source: Perez-Guzman et al. 2023 [1].

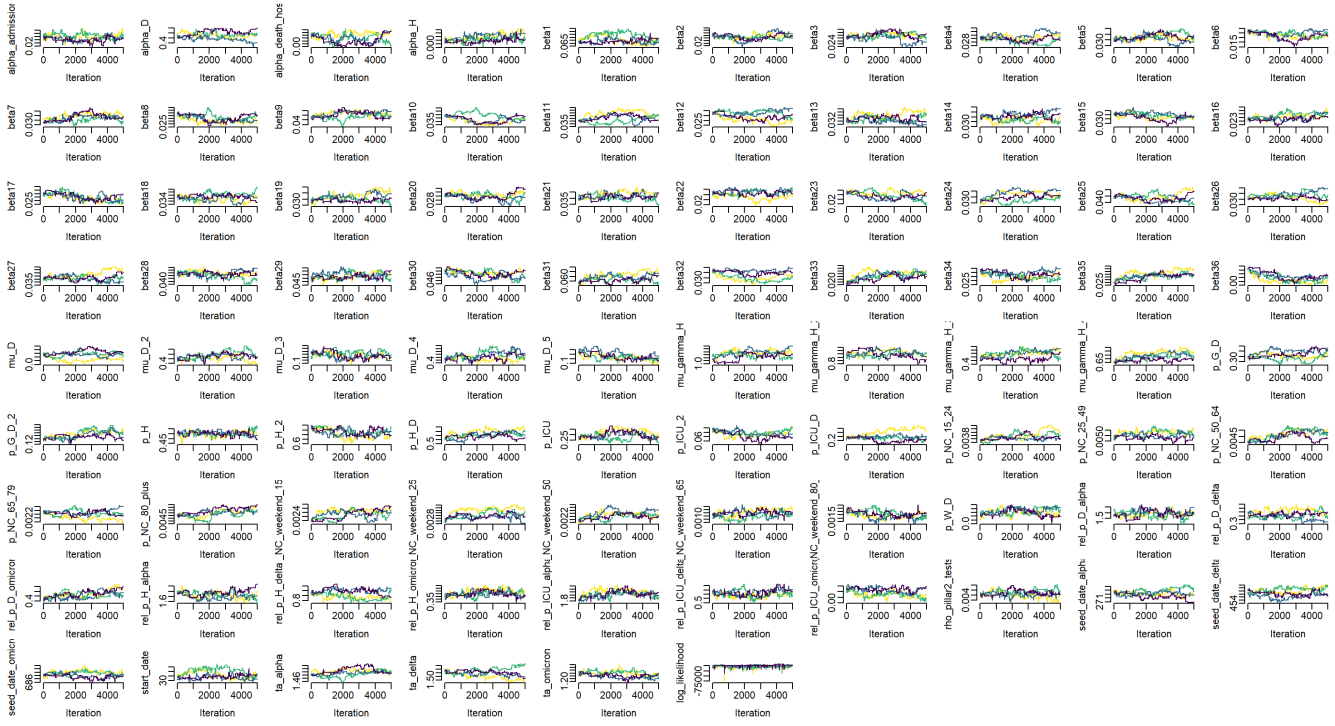


Figure A73: pMCMC traceplots for model fit to South East. Source: Perez-Guzman et al. 2023 [1].

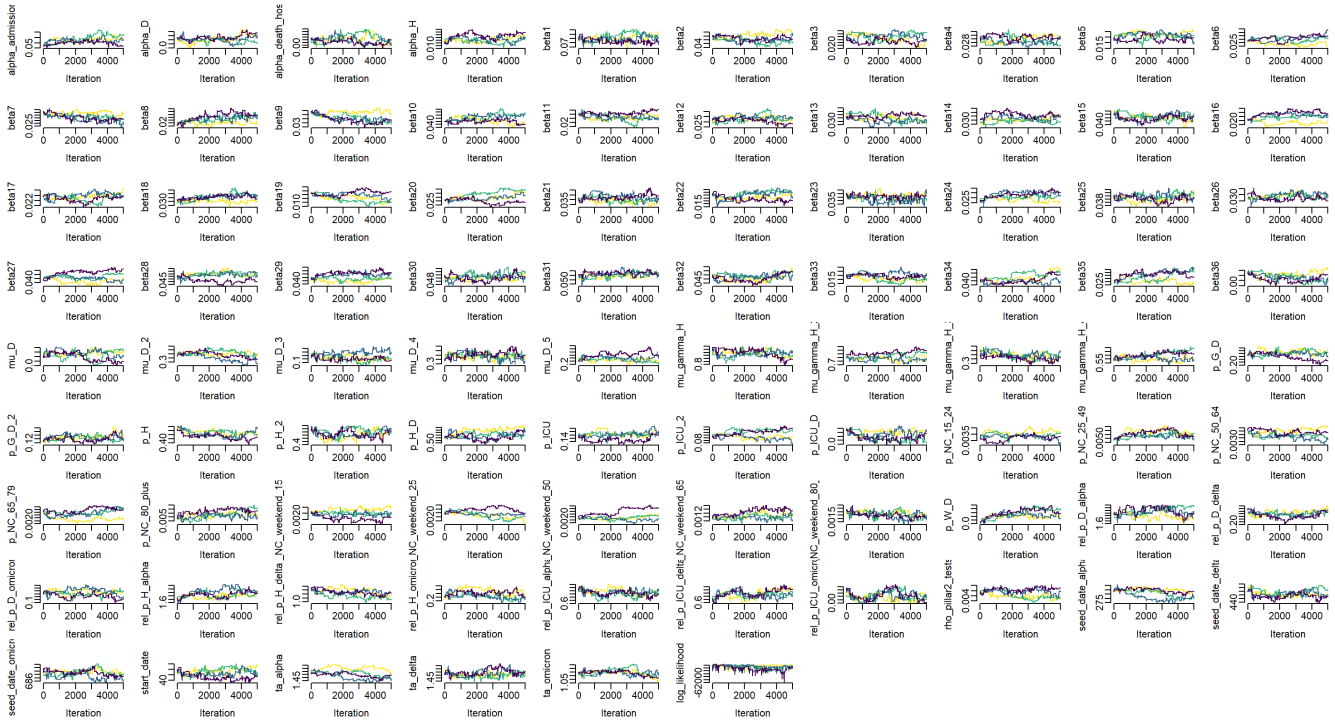


Figure A74: pMCMC traceplots for model fit to South West. Source: Perez-Guzman et al. 2023 [1].

A.6 Inferred trajectories

This section presents model-inferred (given model fit to data and pMCMC-inferred fitted parameters) epidemiological trajectories.



Figure A75: Inferred trajectories of $\beta(t)$ by England NHS regions. Source: Perez-Guzman et al. 2023 [1].

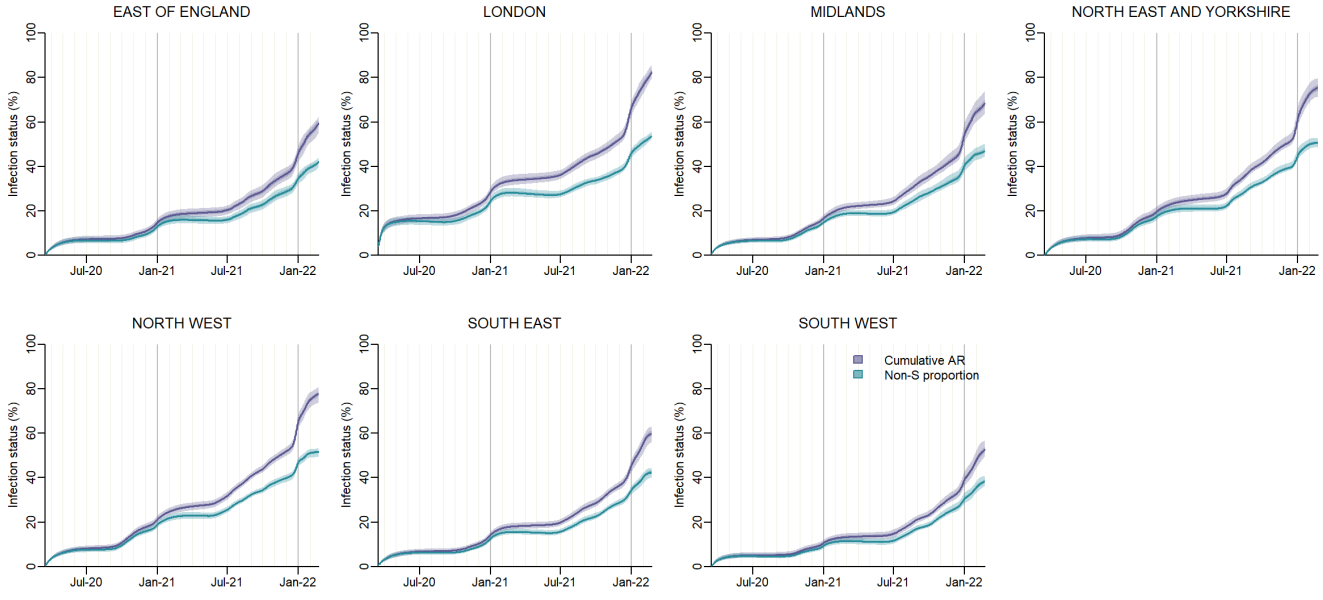


Figure A76: Inferred trajectories of the cumulative infection attack rate and proportion of the population not in susceptible classes by England NHS regions. Source: Perez-Guzman et al. 2023 [1].

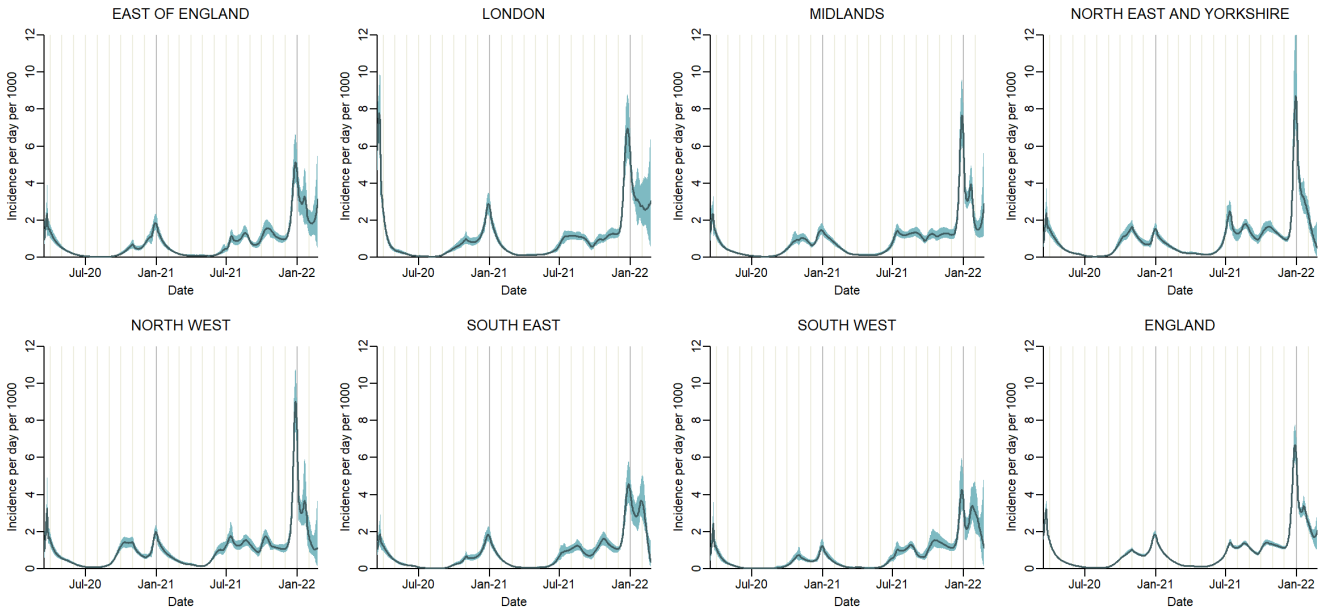


Figure A77: Inferred trajectories of the daily infection incidence rate per-1,000 population by England NHS regions. Source: Perez-Guzman et al. 2023 [1].

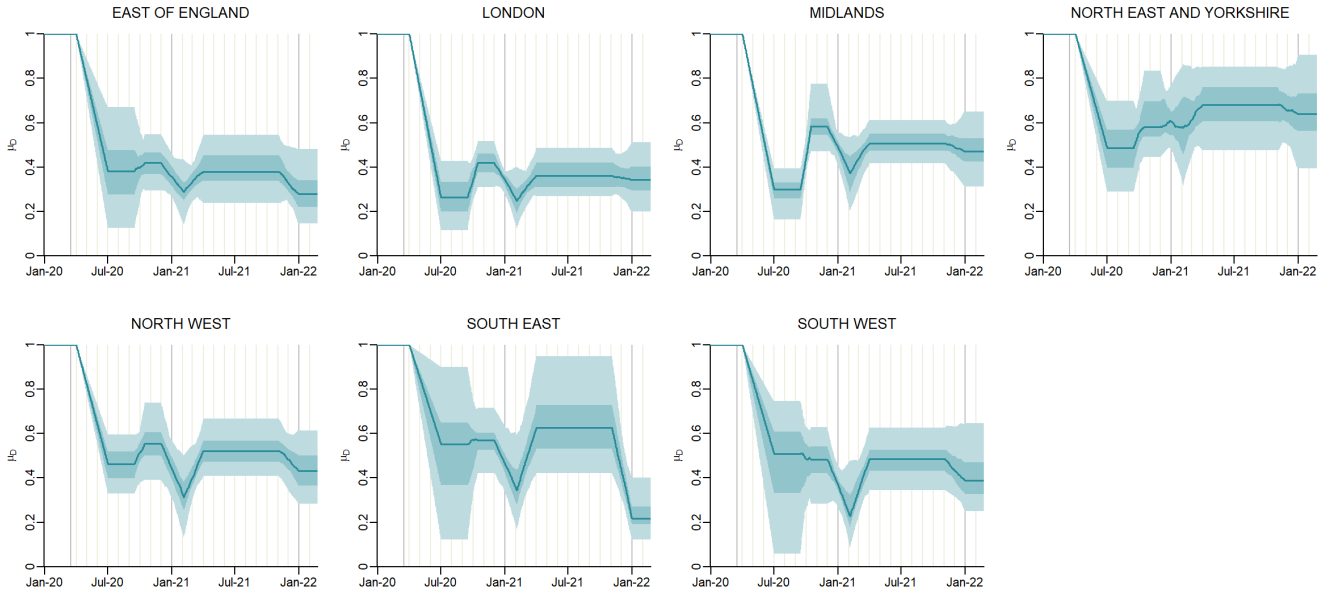


Figure A78: Inferred trajectories of the time-varying probability of death μ_D severity multiplier by England NHS regions. Source: Perez-Guzman et al. 2023 [1].

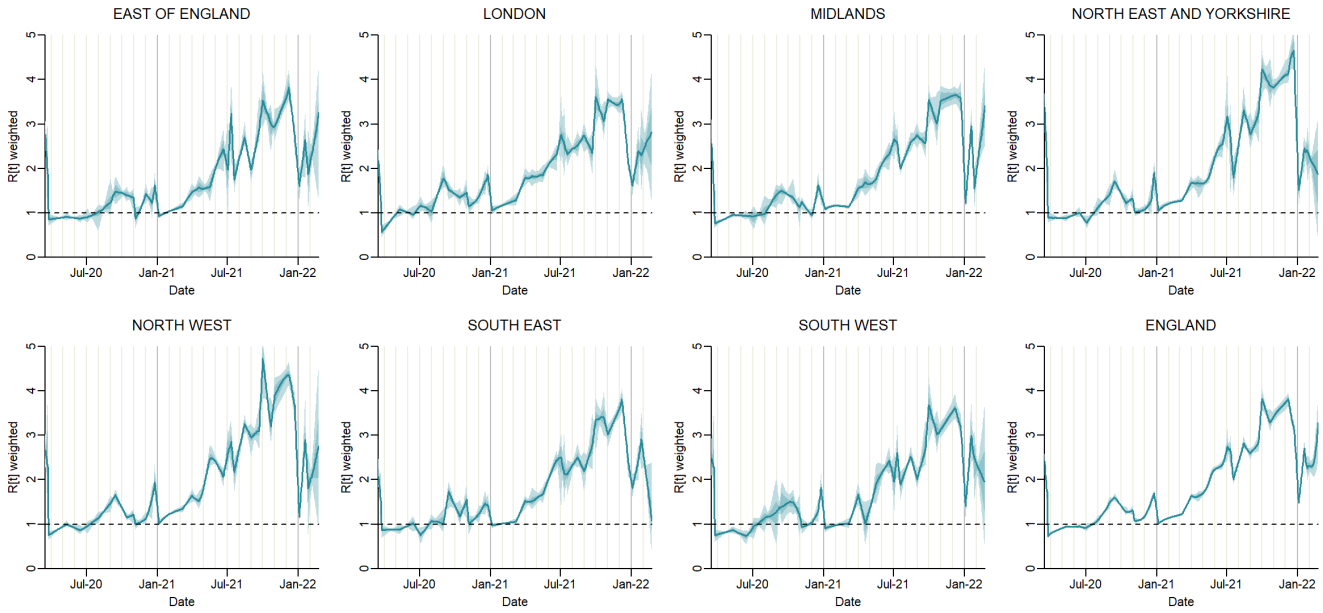


Figure A79: Inferred trajectories of the time-varying reproduction number in the absence of immunity R_t by England NHS regions. Source: Perez-Guzman et al. 2023 [1].

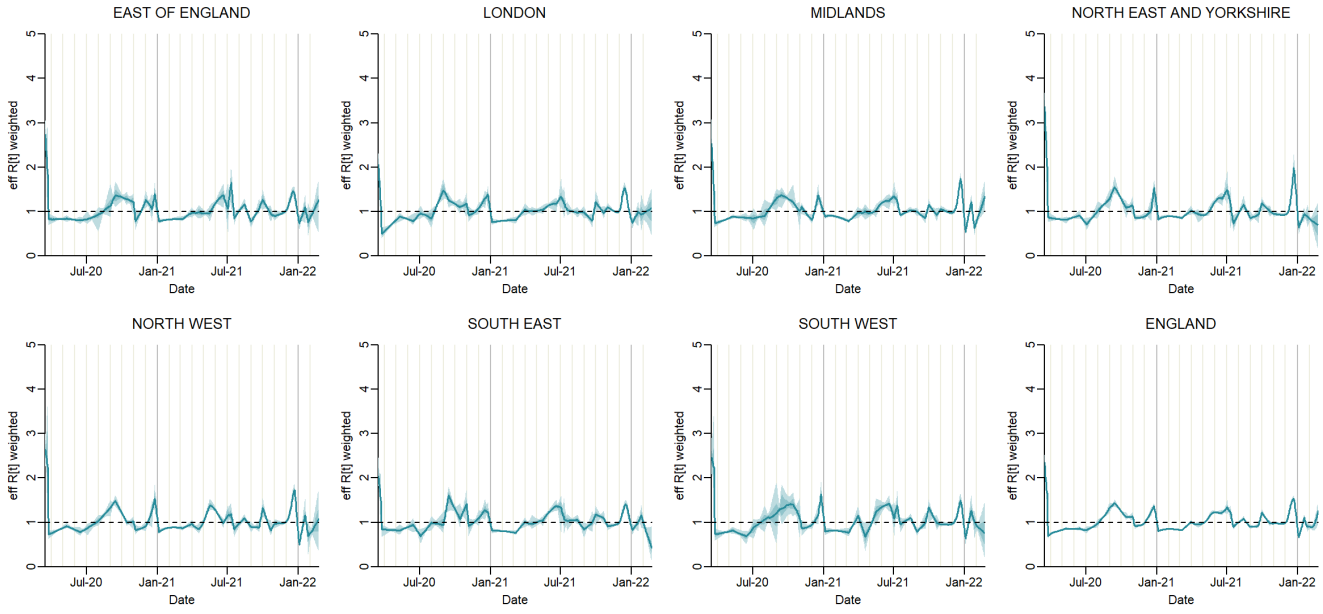


Figure A80: Inferred trajectories of the time-varying reproduction number accounting for immunity R_t^{eff} by England NHS regions. Source: Perez-Guzman et al. 2023 [1].

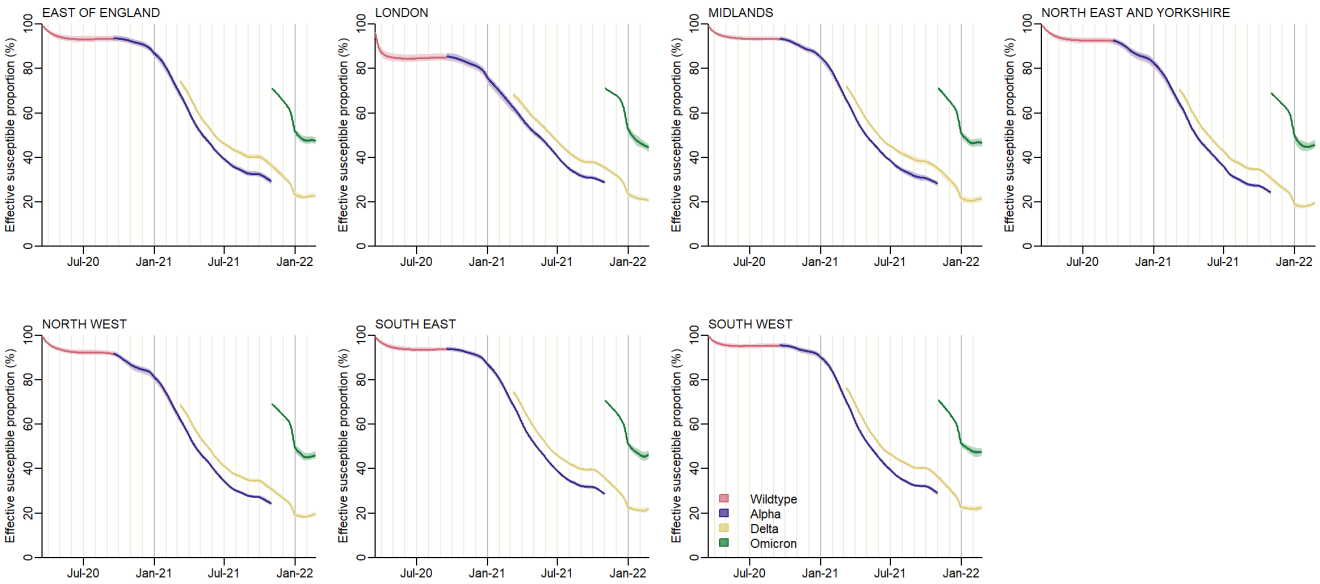


Figure A81: Inferred trajectories of the proportion of the population with susceptibility to infection by variant and England NHS regions. Source: Perez-Guzman et al. 2023 [1].

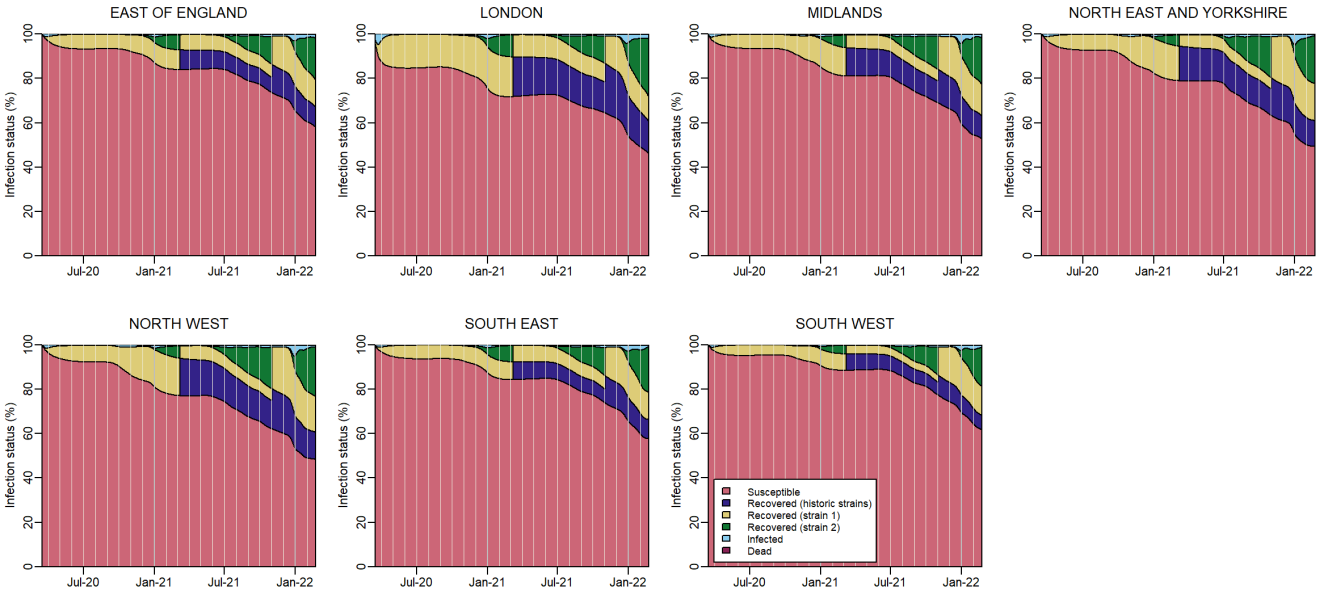


Figure A82: Inferred trajectories of the cumulative proportion of the population by infection status and England NHS regions. Source: Perez-Guzman et al. 2023 [1].

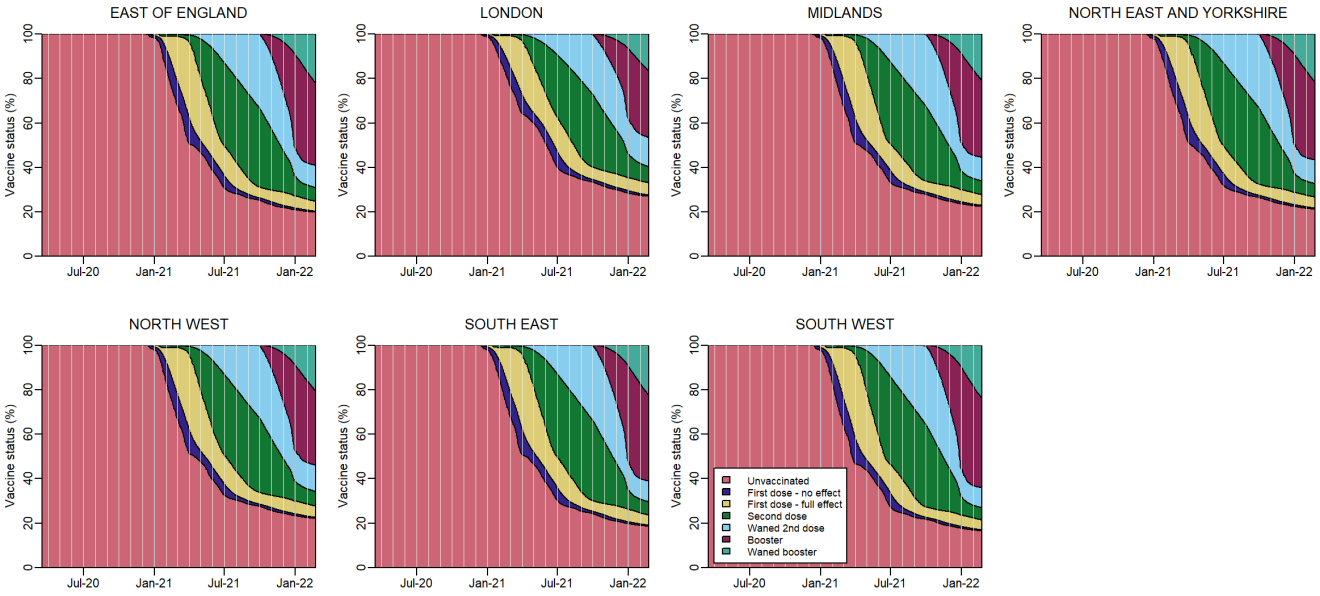


Figure A83: Inferred trajectories of the cumulative proportion of the population by vaccination status and England NHS regions. Source: Perez-Guzman et al. 2023 [1].

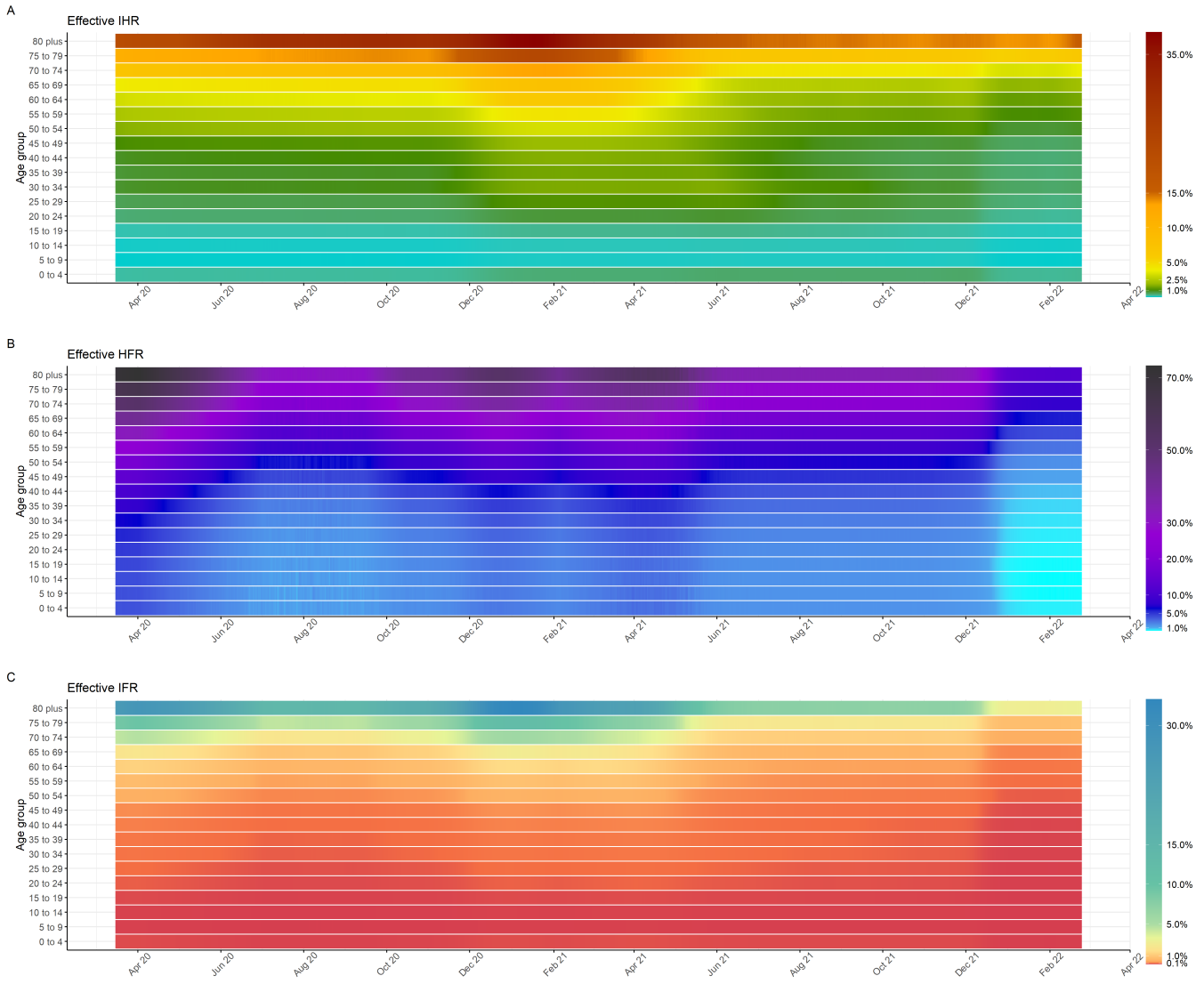


Figure A84: Inferred trajectories of the effective severity by age and England NHS regions. Note bespoke scales of heatmap colours, which were chosen to emphasise differences across age groups. Source: Perez-Guzman et al. 2023 [1].

A.7 Comparison to non-fitted trajectories

This section presents a comparison of key non-fitted trajectories, as inferred by my model, to available data. I performed these comparisons as means of model validation, and also to unpick key areas of model uncertainty.



Figure A85: Comparison of model inferred (not fitted) trajectories to data on hospital admissions by vaccination status and England NHS regions. Source: Perez-Guzman et al. 2023 [1].

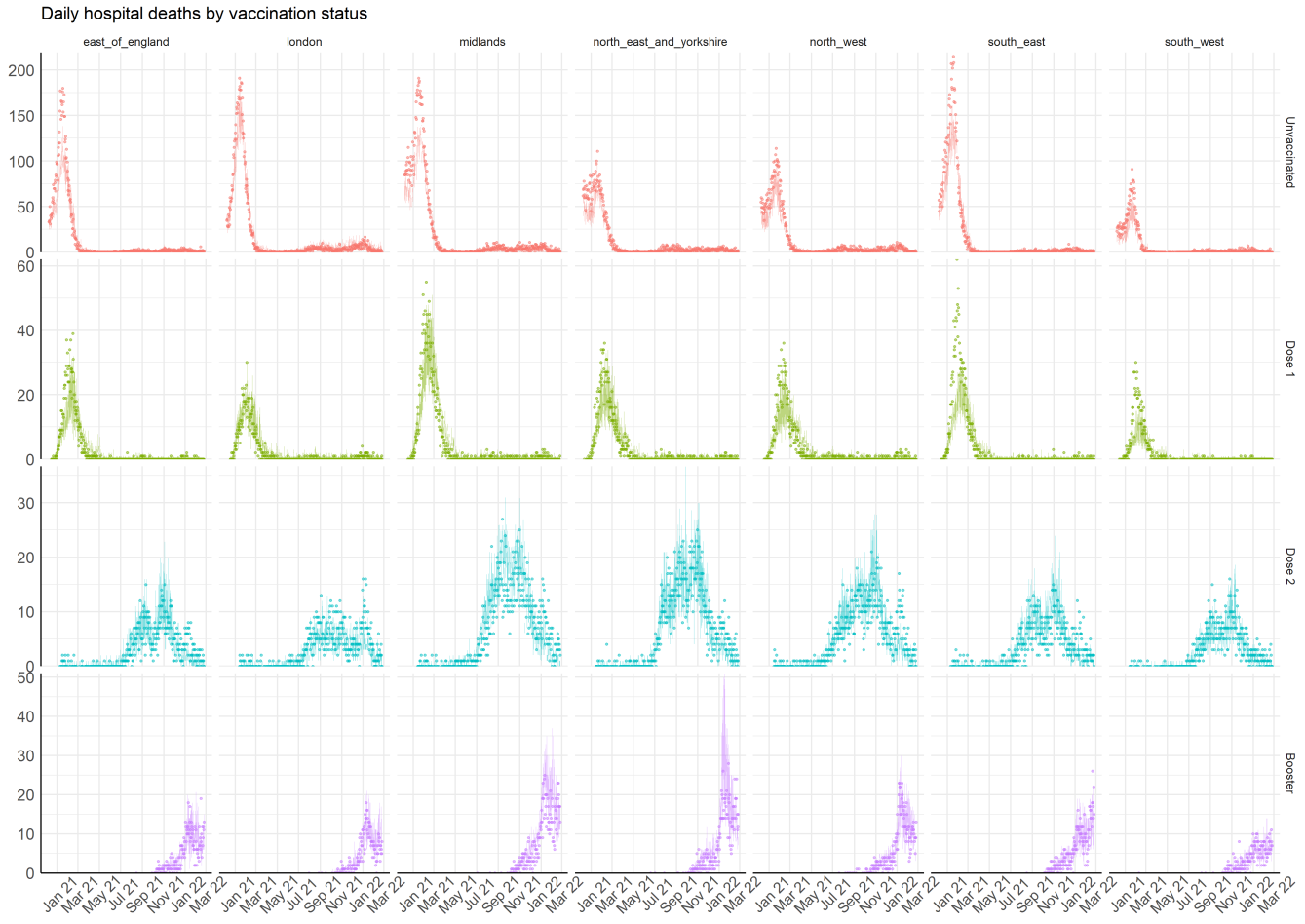


Figure A86: Comparison of model inferred (not fitted) trajectories to data on hospital deaths by vaccination status and England NHS regions. Source: Perez-Guzman et al. 2023 [1].

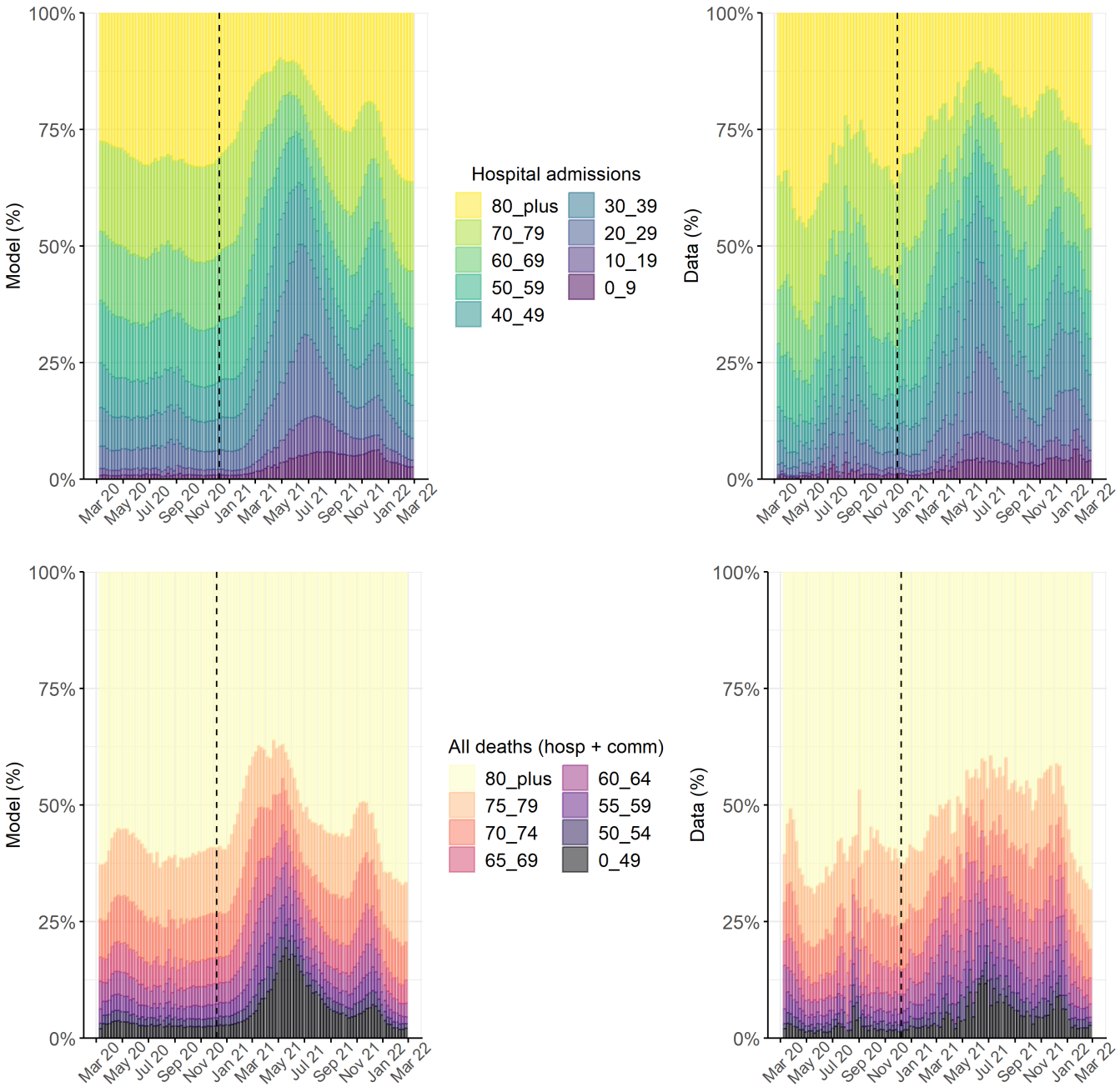


Figure A87: Comparison of model inferred (not fitted) trajectories of the age distribution of hospital admissions and deaths at the national level. Source: Perez-Guzman et al. 2023 [1].

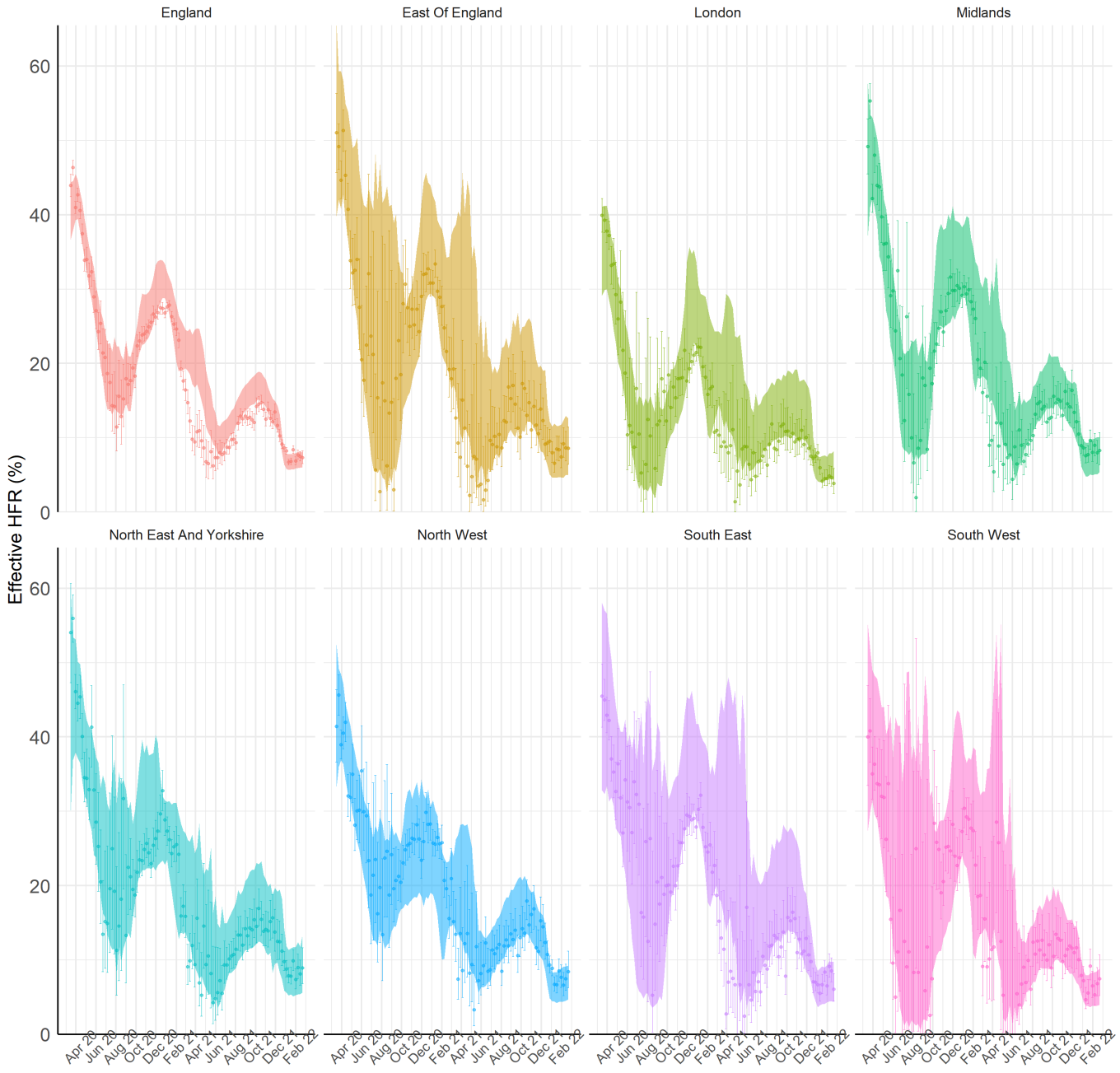


Figure A88: Comparison of model inferred (not fitted) trajectories of the hospital fatality ratio by NHS England region. Source: Perez-Guzman et al. 2023 [1].

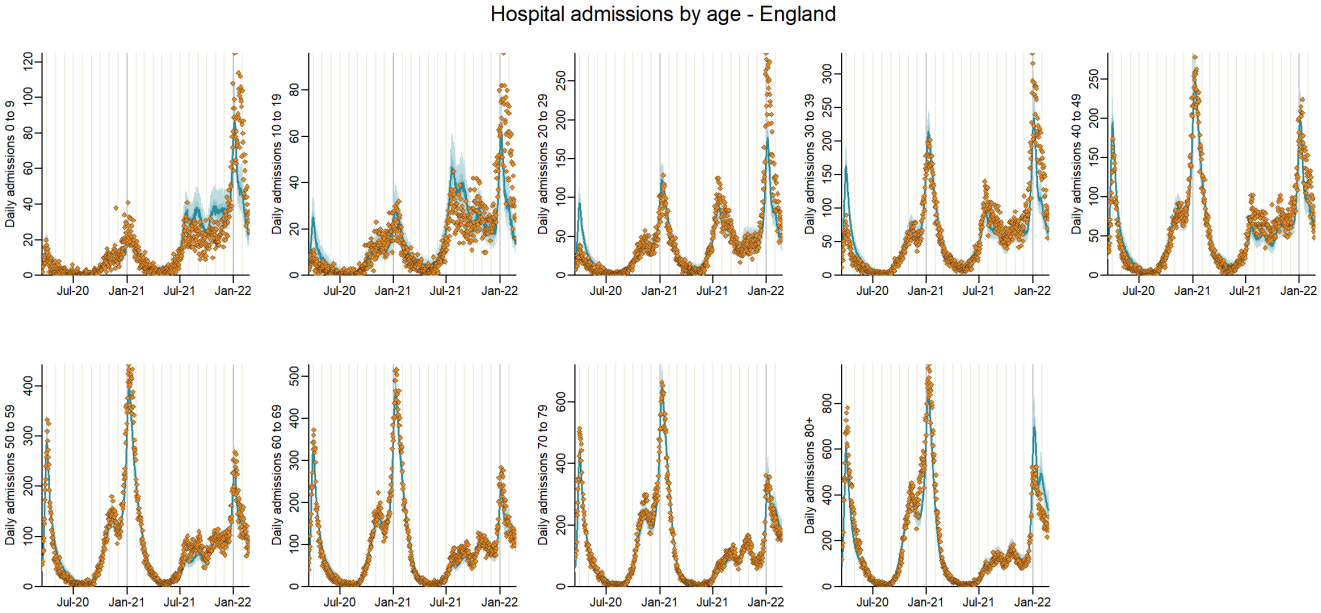


Figure A89: Comparison of model inferred (not fitted) trajectories of hospital admissions by age aggregated at the national level. Source: Perez-Guzman et al. 2023 [1].

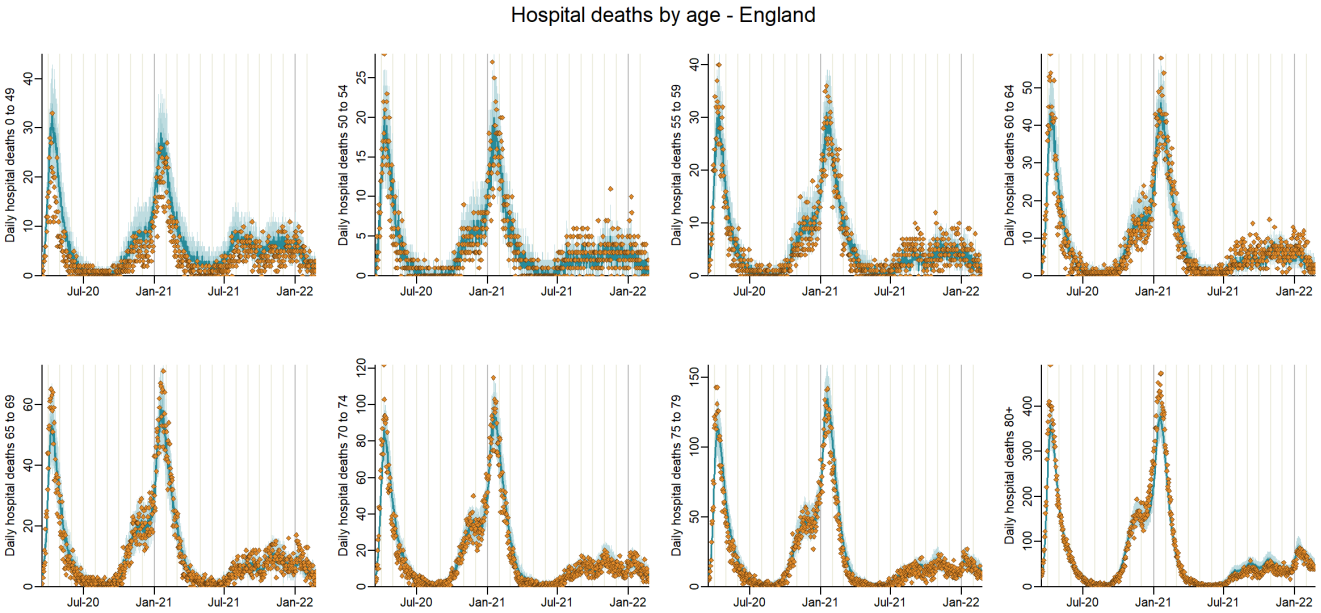


Figure A90: Comparison of model inferred (not fitted) trajectories of hospital deaths by age aggregated at the national level. Source: Perez-Guzman et al. 2023 [1].

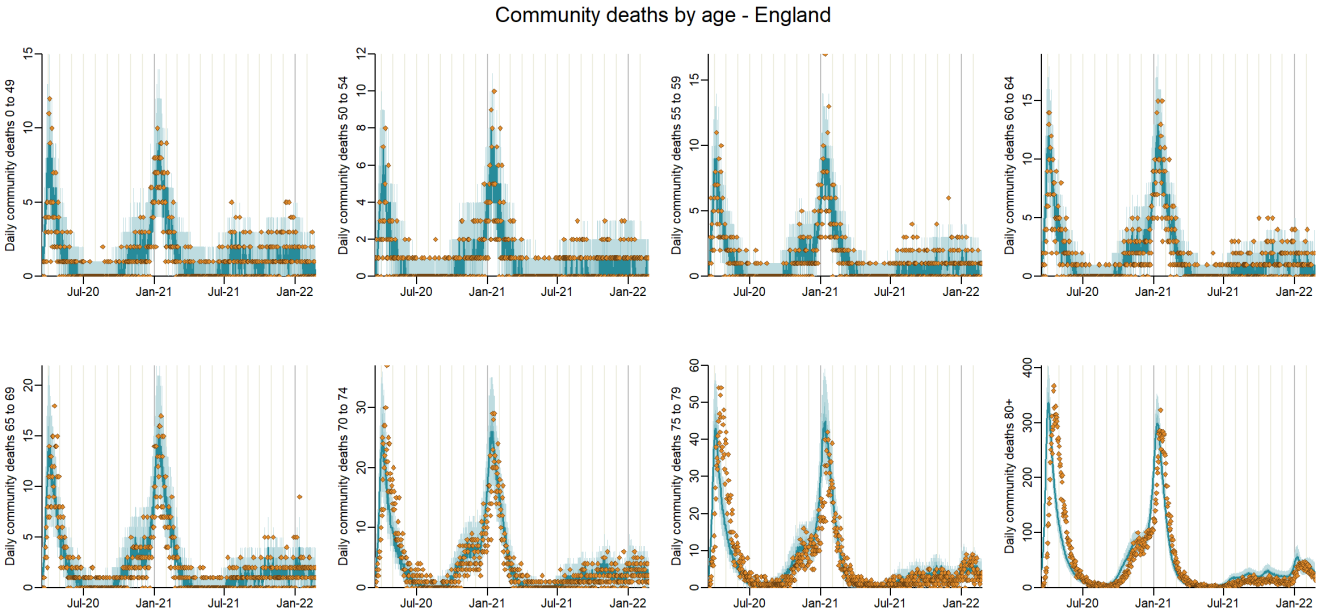


Figure A91: Comparison of model inferred (not fitted) trajectories of community deaths by age aggregated at the national level. Source: Perez-Guzman et al. 2023 [1].

B Kabwe model technical appendix

B.1 Model demography and infections seed

I initialised the model on the 1st of January 2020, assuming the entire population of Kabwe was susceptible to SARS-CoV-2. I used total population estimates for Kabwe district from the 2022 national population census [39]. However, given no information on the age distribution in Kabwe was available from this source, I assumed this was the same as that at the national level (Table B1), which I derived from the United Nations' World Population Prospects [289].

I then seeded the epidemic, assuming a single day of introduction of 20 infected cases per million population $\delta^{i,j=0}(t)$, by fitting a seeding date parameter t_{seed} with a non-informative (flat) prior and a range between February 13 and March 5, 2020. That is, assuming the local epidemic in Kabwe started within three weeks prior to the date of first detection of SARS-CoV-2 in the Southern Africa region [290].

Table B1: Assumed population distribution in Kabwe at model initialisation. Total population as per latest national population census [39]. Population estimates as of 2022 adjusted downwards for a 3.3% growth rate per year (as reported in same source) to obtain population in 2020. Age distribution assumed to be the same as national level for Zambia [289]. Source: Perez-Guzman et al. 2024 [65].

Age group	n
0-4	44,842
5-9	41,351
10-14	36,961
15-19	32,288
20-24	26,587
25-29	21,815
30-34	18,163
35-39	15,093
40-44	12,300
45-49	9,217
50-54	6,599
55-59	4,987
60-64	3,615
65-69	2,531
70-74	1,716
75+	1,718
Total	279,783

B.2 Model parameters

Fitted parameter priors

I used informative prior distributions for $\beta(t)$, t_{seed} and $p_{G_D}^{max}$, and uninformative (flat) distributions for the rest of parameters (Table B2).

Table B2: Inferred model parameter notations and prior distributions

Parameter	Description	Prior distribution	95% probability interval	Rationale
t_{seed}	Start date of local pandemic (dd/mm/2020)	$U[13/02, 05/03]$	(13/02 - 05/03)	Wide range of dates in early 2020 before our first data point
$\beta(t)$	Transmission rate (pp) at $t = dd/mm/yy$			
β_1	01/03/20	$\Gamma(74, 0.00128)$	(0.09,0.0128)	Range corresponding to a basic reproduction numbers between 2.0 and 3.1 consistent with [121]
β_2	15/03/20	$\Gamma(74, 0.00128)$	(0.09,0.0128)	Maintaining R_0 range
β_3	29/03/20	$\Gamma(74, 0.00128)$	(0.09,0.0128)	As above
β_4	12/04/20	$\Gamma(74, 0.00128)$	(0.09,0.0128)	As above
β_5	26/04/20	$\Gamma(74, 0.00128)$	(0.09,0.0128)	As above
β_6	10/05/20	$\Gamma(74, 0.00128)$	(0.09,0.0128)	As above
β_7	24/05/20	$\Gamma(74, 0.00128)$	(0.09,0.0128)	As above
β_8	07/06/20	$\Gamma(74, 0.00128)$	(0.09,0.0128)	As above
β_9	21/06/20	$\Gamma(74, 0.00128)$	(0.09,0.0128)	As above
β_{10}	05/07/20	$\Gamma(74, 0.00128)$	(0.09,0.0128)	As above
β_{11}	19/07/20	$\Gamma(24, 0.00269)$	(0.129,0.00269)	Broader range to allow R_t between 1.1 and 3.1
β_{12}	02/08/20	$\Gamma(24, 0.00269)$	(0.129,0.00269)	As above
β_{13}	16/08/20	$\Gamma(24, 0.00269)$	(0.129,0.00269)	As above
β_{14}	30/08/20	$\Gamma(24, 0.00269)$	(0.129,0.00269)	As above
β_{15}	13/09/20	$\Gamma(24, 0.00269)$	(0.129,0.00269)	As above
β_{16}	27/09/20	$\Gamma(24, 0.00269)$	(0.129,0.00269)	As above
β_{17}	11/10/20	$\Gamma(24, 0.00269)$	(0.129,0.00269)	As above
β_{18}	25/10/20	$\Gamma(24, 0.00269)$	(0.129,0.00269)	As above

Continued on next page

Table B2 – continued from previous page

	Description	Prior distribution	95% probability interval	Rationale
β_{19}	08/11/20	$\Gamma(24, 0.00269)$	(0.129,0.00269)	As above
β_{20}	22/11/20	$\Gamma(24, 0.00269)$	(0.129,0.00269)	As above
β_{21}	06/12/20	$\Gamma(24, 0.00269)$	(0.129,0.00269)	As above
β_{22}	20/12/20	$\Gamma(24, 0.00269)$	(0.129,0.00269)	As above
β_{23}	03/01/21	$\Gamma(24, 0.00269)$	(0.129,0.00269)	As above
β_{24}	17/01/21	$\Gamma(24, 0.00269)$	(0.129,0.00269)	As above
β_{25}	31/01/21	$\Gamma(24, 0.00269)$	(0.129,0.00269)	As above
β_{26}	14/02/21	$\Gamma(24, 0.00269)$	(0.129,0.00269)	As above
β_{27}	28/02/21	$\Gamma(24, 0.00269)$	(0.129,0.00269)	As above
β_{28}	14/03/21	$\Gamma(24, 0.00269)$	(0.129,0.00269)	As above
β_{29}	28/03/21	$\Gamma(24, 0.00269)$	(0.129,0.00269)	As above
β_{30}	11/04/21	$\Gamma(24, 0.00269)$	(0.129,0.00269)	As above
β_{31}	25/04/21	$\Gamma(24, 0.00269)$	(0.129,0.00269)	As above
β_{32}	09/05/21	$\Gamma(24, 0.00269)$	(0.129,0.00269)	As above
β_{33}	23/05/21	$\Gamma(24, 0.00269)$	(0.129,0.00269)	As above
β_{34}	06/06/21	$\Gamma(24, 0.00269)$	(0.129,0.00269)	As above
β_{35}	20/06/21	$\Gamma(24, 0.00269)$	(0.129,0.00269)	As above
β_{36}	04/07/21	$\Gamma(24, 0.00269)$	(0.129,0.00269)	As above
β_{37}	18/07/21	$\Gamma(24, 0.00269)$	(0.129,0.00269)	As above
β_{38}	01/08/21	$\Gamma(24, 0.00269)$	(0.129,0.00269)	As above
β_{39}	15/08/21	$\Gamma(24, 0.00269)$	(0.129,0.00269)	As above
β_{40}	29/08/21	$\Gamma(24, 0.00269)$	(0.129,0.00269)	As above
β_{41}	12/09/21	$\Gamma(24, 0.00269)$	(0.129,0.00269)	As above
β_{42}	26/09/21	$\Gamma(24, 0.00269)$	(0.129,0.00269)	As above

Continued on next page

Table B2 – continued from previous page

	Description	Prior distribution	95% probability interval	Rationale
$p_{G_D}^{max}$	Probability of death given disease for the group with the largest baseline probability (see Section 4.2.4)	$\beta(8.3, 82.4)$	(0.041,0.092)	Range to incorporate plausible values from literature, consistent with [119]
$\mu_{D,1}, \mu_{D,2}$	Probability of death given disease multipliers due to changes in pandemic severity at different timepoints (see Section 4.2.4)	$U(0, 3)$	(0.075,2.925)	Uninformative
α_D	Overdispersion for all-cause excess deaths data	$U(0, 1)$	(0.025,0.975)	Uninformative

B.3 Model equations

As described in [65] I employed a stochastic, space-state model. To clearly illustrate its compartmental transition dynamics, however, I first present a deterministic (ordinary differential) equivalent set of equations (191)-(204), followed by its stochastic implementation in (208)-(235). Please refer to Chapter 4 for the model diagram and definitions of its compartments and parameters.

Deterministic equations

I use $\sum_i$ to refer to the sum across age groups (i.e. $i \in \{[0,5), \ldots, [70,75), [75+)\}$), $\sum_j$ refers to the sum across all vaccination strata ($j \in \{0, 1, ..., 4\}$), and $k \in \{1, 2\}$ to Erlang-distributed durations of stay in certain compartments (e.g. $E^{i,j}$) as the sum of independent exponential distributions. [265] For a detailed explanation of vaccine class transition parameters $\zeta^{i,j}$ see sections 5.2.2 and 5.2.3.

$$\frac{dS^{i,j}(t)}{dt} = \zeta^{i,j-1}(t)S^{i,j-1}(t) - (\lambda^{i,j}(t) + \zeta^{i,j}(t))S^{i,j}(t) - \delta^{i,j=0}(t) + \gamma_R R^{i,j}(t) \tag{191}$$

$$\begin{aligned}\frac{dE^{i,j,k=1}(t)}{dt} &= \zeta^{i,j-1}(t)E^{i,j-1,k=1}(t) + \lambda^{i,j}(t)S^{i,j}(t) + \delta^{i,j=0}(t) \\ &\quad + (1-\eta(t))\lambda^{i,j}(t)R^{i,j}(t) - (\gamma_E + \zeta^{i,j}(t))E^{i,j,k=1}(t)\end{aligned} \tag{192}$$

$$\frac{dE^{i,j,k=2}(t)}{dt} = \zeta^{i,j-1}(t)E^{i,j-1,k=2}(t) + \gamma_E E^{i,j,k=1}(t) - (\gamma_E + \zeta^{i,j}(t))E^{i,j,k=2}(t) \tag{193}$$

$$\frac{dI_A^{i,j}(t)}{dt} = \zeta^{i,j-1}(t)I_A^{i,j-1}(t) + (1-p_C^i)\gamma_E E^{i,j,k=2}(t) - (\gamma_A + \zeta^{i,j}(t))I_A^{i,j} \tag{194}$$

$$\frac{dI_P^{i,j}(t)}{dt} = \zeta^{i,j-1}(t)I_P^{i,j-1}(t) + p_C^i\gamma_E E^{i,j,k=2}(t) - (\gamma_P + \zeta^{i,j}(t))I_P^{i,j}(t) \tag{195}$$

$$\frac{dI_{C_1}^{i,j}(t)}{dt} = \gamma_P I_P^{i,j}(t) - \gamma_{C_1} I_{C_1}^{i,j}(t) \tag{196}$$

$$\frac{dI_{C_2}^{i,j}(t)}{dt} = \gamma_{C_1} I_{C_1}^{i,j}(t) - \gamma_{C_2} I_{C_2}^{i,j}(t) \tag{197}$$

$$\frac{dG_D^{i,j,k=1}(t)}{dt} = \gamma_{C_2} p_{G_D}^{i,j}(t) I_{C_2}^{i,j} - \gamma_{G_D} G_D^{i,j,k=1}(t) \tag{198}$$

$$\frac{dG_D^{i,j,k=2}(t)}{dt} = \gamma_{G_D} G_D^{i,j,k=1}(t) - \gamma_{G_D} G_D^{i,j,k=2}(t) \tag{199}$$

$$\frac{dD^{i,j}(t)}{dt} = \gamma_{G_D} G_D^{i,j,k=2}(t) \tag{200}$$

$$\begin{aligned}\frac{dR^{i,j}(t)}{dt} &= \zeta^{i,j-1}(t)R^{i,j-1}(t) + \gamma_{C_2}(1 - p_{G_D}^{i,j}(t))I_{C_2}^{i,j}(t) \\ &\quad - (\gamma_R + \zeta^{i,j}(t) + (1-\eta(t))\lambda^{i,j}(t))R^{i,j}(t)\end{aligned} \tag{201}$$

$$\frac{dT_{sero_{pre}}^i(t)}{dt} = \sum_j \gamma_E E^{i,j,k=2}(t) - \gamma_{sero_{pre}} T_{sero_{pre}}^i(t) \tag{202}$$

$$\frac{dT_{sero_{pos}}^i(t)}{dt} = p_{sero_{pos}}\gamma_{sero_{pre}} T_{sero_{pre}}^i(t) - \gamma_{sero_{pos}} T_{sero_{pos}}^i(t) \tag{203}$$

$$\frac{dT^i_{sero_{neg}}(t)}{dt} = \left(1 - p_{sero_{pos}}\right) \gamma_{sero_{pre}} T^i_{sero_{pre}}(t) + \gamma_{sero_{pos}} T^i_{sero_{pos}}(t) \tag{204}$$

Stochastic equations

I used the tau-leap method [266] to create a stochastic, time-discretised version of the afore-described deterministic equations. [65] I took four update steps per day ($dt = 0.25$ days), at each of which the process described below was followed to update model compartments.

For clarity in describing transitions involving multiple onward compartments, I write

$$\left(d^{i,j}_{E,E},\ d^{i,j,k=1}_{E,v}\right) \sim \text{Multinom}\left(E^{i,j,k=1}(t),\ q^{i,j}_{E,E},\ q^{i,j,k=1}_{E,v}\right) \tag{205}$$

instead of

$$\left(d^{i,j}_{E,E},\ d^{i,j,k=1}_{E,v},\ d^{i,j}_{nomove}\right) \sim \text{Multinom}\left(E^{i,j,k=1}(t),\ q^{i,j}_{E,E},\ q^{i,j,k=1}_{E,v},\ 1 - \sum_{x \in E,v} q^{i,j}_{E,x}\right) \tag{206}$$

where $d^{i,j,k}_{nomove}$ is a dummy variable counting the number of individuals remaining in compartment $E^{i,j,k=1}$. I also omit the time dependency i.e. use $d^{i,j}_{E,I_A}$ or $q^{i,j}_{E,I_A}$ instead of $d^{i,j}_{E,I_A}(t)$ or $q^{i,j}_{E,I_A}(t)$.

Of note, in the following equations I use $\mathbb{1}_A$ as an indicator function, such that

$$\mathbb{1}_A(j) \coloneqq \begin{cases} 1 & \text{if } j \in A, \\ 0 & \text{if } j \notin A. \end{cases} \tag{207}$$

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

$$q^{i,j}_{S,E} = 1 - e^{-\lambda^{i,j}(t)dt} \tag{208}$$

$$q^{i,j}_{S,v} = 1 - e^{-\zeta^{i,j}(t)dt} \tag{209}$$

$$d^{i,j}_{S,E} \sim \text{Binom}\left(S^{i,j}(t),\ q^{i,j}_{S,E}\right) \tag{210}$$

$$d^{i,j}_{seed} \sim \min\left(\text{Poisson}\left(\delta^{i,j}(t)dt\right),\ S^{i,j}(t) - d^{i,j}_{S,E}\right) \tag{211}$$

$$d^{i,j}_{S,v} \sim \text{Binom}\left(S^{i,j}(t) - d^{i,j}_{seed} - d^{i,j}_{S,E},\ q^{i,j}_{S,v}\right) \tag{212}$$

$$\left(q^{i,j}_{E,E},\ q^{i,j,k=1}_{E,v}\right) = \left(1 - e^{-\gamma_E dt},\ e^{-\gamma_E dt}\left(1 - e^{-\zeta^{i,j}(t)dt}\right)\right) \tag{213}$$

$$\left(d^{i,j}_{E,E},\ d^{i,j,k=1}_{E,v}\right) \sim \text{Multinom}\left(E^{i,j,k=1}(t),\ q^{i,j}_{E,E},\ q^{i,j,k=1}_{E,v}\right) \tag{214}$$

$$q^{i,j}_{E,I_A} = \left(1 - p^i_C\right)\left(1 - e^{-\gamma_E dt}\right) \tag{215}$$

$$q^{i,j}_{E,I_P} = p^i_C\left(1 - e^{-\gamma_E dt}\right) \tag{216}$$

$$\left(d^{i,j}_{E,I_A},\ d^{i,j}_{E,I_P},\ d^{i,j,k=2}_{E,v}\right) \sim \text{Multinom}\left(E^{i,j,k=2}(t),\ q^{i,j}_{E,I_A},\ q^{i,j}_{E,I_P}\right) \tag{217}$$

$$\left(q^{i,j}_{I_A,R},\ q^{i,j}_{I_A,v}\right) = \left(1-e^{-\gamma_A dt},\ e^{-\gamma_A dt}\left(1-e^{-\zeta^{i,j}(t)dt}\right)\right) \tag{218}$$

$$\left(d^{i,j}_{I_A,R},\ d^{i,j}_{I_A,v}\right) \sim \text{Multinom}\left(I^{i,j}_A(t),\ q^{i,j}_{I_A,R},\ q^{i,j}_{I_A,v}\right) \tag{219}$$

$$\left(q^{i,j}_{I_P,I_{C_1}},\ q^{i,j}_{I_P,v}\right) = \left(1-e^{-\gamma_P dt},\ e^{-\gamma_P dt}\left(1-e^{-\zeta^{i,j}(t)dt}\right)\right) \tag{220}$$

$$\left(d^{i,j}_{I_P,I_{C_1}},\ d^{i,j}_{I_P,v}\right) \sim \text{Multinom}\left(I^{i,j}_P(t),\ q^{i,j}_{I_P,I_{C_1}},\ q^{i,j}_{I_P,v}\right) \tag{221}$$

$$d^{i,j}_{I_{C_1},I_{C_2}} \sim \text{Binom}\left(I^{i,j}_{C_1}(t),\ 1-e^{-\gamma_{C_1}dt}\right) \tag{222}$$

$$q^{i,j}_{I_{C_2},G_D} = p^{i,j}_{G_D}(t)\left(1-e^{-\gamma_{C_2}dt}\right) \tag{223}$$

$$q^{i,j}_{I_{C_2},R} = \left(1-p^{i,j}_{G_D}(t)\right)\left(1-e^{-\gamma_{C_2}dt}\right) \tag{224}$$

$$\left(d^{i,j}_{I_{C_2},G_D},\ d^{i,j}_{I_{C_2},R}\right) \sim \text{Multinom}\left(I^{i,j}_{C_2}(t),\ q^{i,j}_{I_{C_2},G_D},\ q^{i,j}_{I_{C_2},R}\right) \tag{225}$$

$$d^{i,j}_{G_D,G_D} \sim \text{Binom}\left(G^{i,j,k=1}_D(t),\ 1-e^{-\gamma_{G_D}dt}\right) \tag{226}$$

$$d^{i,j}_{G_D,D} \sim \text{Binom}\left(G^{i,j,k=2}_D(t),\ 1-e^{-\gamma_{G_D}dt}\right) \tag{227}$$

$$d^{i,j}_{R,S} \sim \text{Binom}\left(R^{i,j}(t),\ 1-e^{-\gamma_R dt}\right) \tag{228}$$

$$d^{i,j}_{R,E} \sim \text{Binom}\left(R^{i,j}(t)-d^{i,j}_{R,S},\ 1-e^{-(1-\eta(t))\lambda_{i,j}(t)dt}\right) \tag{229}$$

$$q^{i,j}_{R,v} = 1-e^{-\zeta^{i,j}(t)dt} \tag{230}$$

$$d^{i,j}_{R,v} \sim \text{Binom}\left(R^{i,j}(t)-d^{i,j}_{R,S}-d^{i,j}_{R,E},\ q^{i,j}_{R,v}\right) \tag{231}$$

$$q^{i}_{T_{sero_{pre}},T_{sero_{pos}}} = p_{sero_{pos}}\left(1-e^{-\gamma_{seropre}dt}\right) \tag{232}$$

$$q^{i}_{T_{sero_{pos}},T_{sero_{neg}}} = \left(1-p_{sero_{pos}}\right)\left(1-e^{-\gamma_{seropre}dt}\right) \tag{233}$$

$$\left(d^{i}_{T_{sero_{pre}},T_{sero_{pos}}},\ d^{i}_{T_{sero_{pre}},T_{sero_{neg}}}\right) \sim \text{Multinom}\left(T^{i}_{sero_{pre}}(t),\ q^{i}_{T_{sero_{pre}},T_{sero_{pos}}},\ q^{i}_{T_{sero_{pre}},T_{sero_{neg}}}\right) \tag{234}$$

$$d^{i}_{T_{sero_{pos}},T_{sero_{neg}}} \sim \text{Binom}\left(T^{i}_{sero_{pos}}(t), 1-e^{-\gamma_{seropos}dt}\right) \tag{235}$$

Model compartments are then updated as follows:

$$S^{i,j}(t+dt) \coloneqq S^{i,j}(t) + d^{i,j}_{R,S}(t) + d^{i,j-1}_{S,v} - d^{i,j}_{S,E} - d^{i,j}_{seed} - d^{i,j}_{S,v} \tag{236}$$

$$E^{i,j,k=1}(t+dt) \coloneqq E^{i,j,k=1}(t) + d^{i,j}_{S,E} + d^{i,j}_{seed} + d^{i,j}_{R,E} + d^{i,j-1,k=1}_{E,v} - d^{i,j}_{E,E} - d^{i,j,k=1}_{E,v} \tag{237}$$

$$E^{i,j,k=2}(t+dt) \coloneqq E^{i,j,k=2}(t) + d^{i,j}_{E,E} + d^{i,j-1,k=2}_{E,v} - d^{i,j}_{E,I_A} - d^{i,j}_{E,I_P} - d^{i,j,k=2}_{E,v} \tag{238}$$

$$I^{i,j}_A(t+dt) \coloneqq I^{i,j}_A(t) + d^{i,j}_{E,I_A} + d^{i,j-1}_{I_A,v} - d^{i,j}_{I_A,R} - d^{i,j}_{I_A,v} \tag{239}$$

$$I^{i,j}_P(t+dt) \coloneqq I^{i,j}_P(t) + d^{i,j}_{E,I_P} + d^{i,j-1}_{I_P,v} - d^{i,j}_{I_P,I_{C_1}} - d^{i,j}_{I_P,v} \tag{240}$$

$$I^{i,j}_{C_1}(t+dt) \coloneqq I^{i,j}_{C_1}(t) + d^{i,j}_{I_P,I_{C_1}} - d^{i,j}_{I_{C_1},I_{C_2}} \tag{241}$$

$$I^{i,j}_{C_2}(t+dt) \coloneqq I^{i,j}_{C_2}(t) + d^{i,j}_{I_{C_1},I_{C_2}} - d^{i,j}_{I_{C_2},G_D} \tag{242}$$

$$G_D^{i,j,k=1}(t+dt) \coloneqq G_D^{i,j,k=1}(t) + d^{i,j}_{I_{C_2},G_D} - d^{i,j}_{G_D,G_D} \tag{243}$$

$$G_D^{i,j,k=2}(t+dt) \coloneqq G_D^{i,j,k=2}(t) + d^{i,j}_{G_D,G_D} - d^{i,j}_{G_D,D} \tag{244}$$

$$D^{i,j}(t+dt) \coloneqq D^{i,j}(t) + d^{i,j}_{G_D,D} \tag{245}$$

$$R^{i,j}(t+dt) \coloneqq R^{i,j}(t) + d^{i,j}_{I_A,R} + d^{i,j}_{I_{C_2},R} + d^{i,j-1}_{R,v} - d^{i,j}_{R,S} - d^{i,j}_{R,E} - d^{i,j}_{R,v} \tag{246}$$

$$T^i_{sero_{pre}}(t+dt) \coloneqq T^i_{sero_{pre}}(t) + \sum_j \left(d^{i,j}_{E,I_P} + d^{i,j}_{E,I_A}\right) - d^i_{T_{sero_{pre}},T_{sero_{pos}}} - d^i_{T_{sero_{pre}},T_{sero_{neg}}} \tag{247}$$

$$T^i_{sero_{pos}}(t+dt) \coloneqq T^i_{sero_{pos}}(t) + d^i_{T_{sero_{pre}},T_{sero_{pos}}} - d^i_{T_{sero_{pos}},T_{sero_{neg}}} \tag{248}$$

$$T^i_{sero_{neg}}(t+dt) \coloneqq T^i_{sero_{neg}}(t) + d^i_{T_{sero_{pos}},T_{sero_{neg}}} + d^i_{T_{sero_{pre}},T_{sero_{neg}}} \tag{249}$$

B.4 Additional results: chapter 4

Figures B1 and B2 show the posterior distributions of fitted parameters. Red lines shown the prior distribution; where no prior lines are shown, a uniform (non-informative) prior was assumed. Figures B3 and B4 show the MCMC and pMCMC traceplots, respectively, of fitted parameters.

Figures B5 to B8 present the sensitivity analysis I conducted of impact of key model parameters on the model's performance to recover the available data streams and infer the effective transmissibility and severity of the local pandemic.

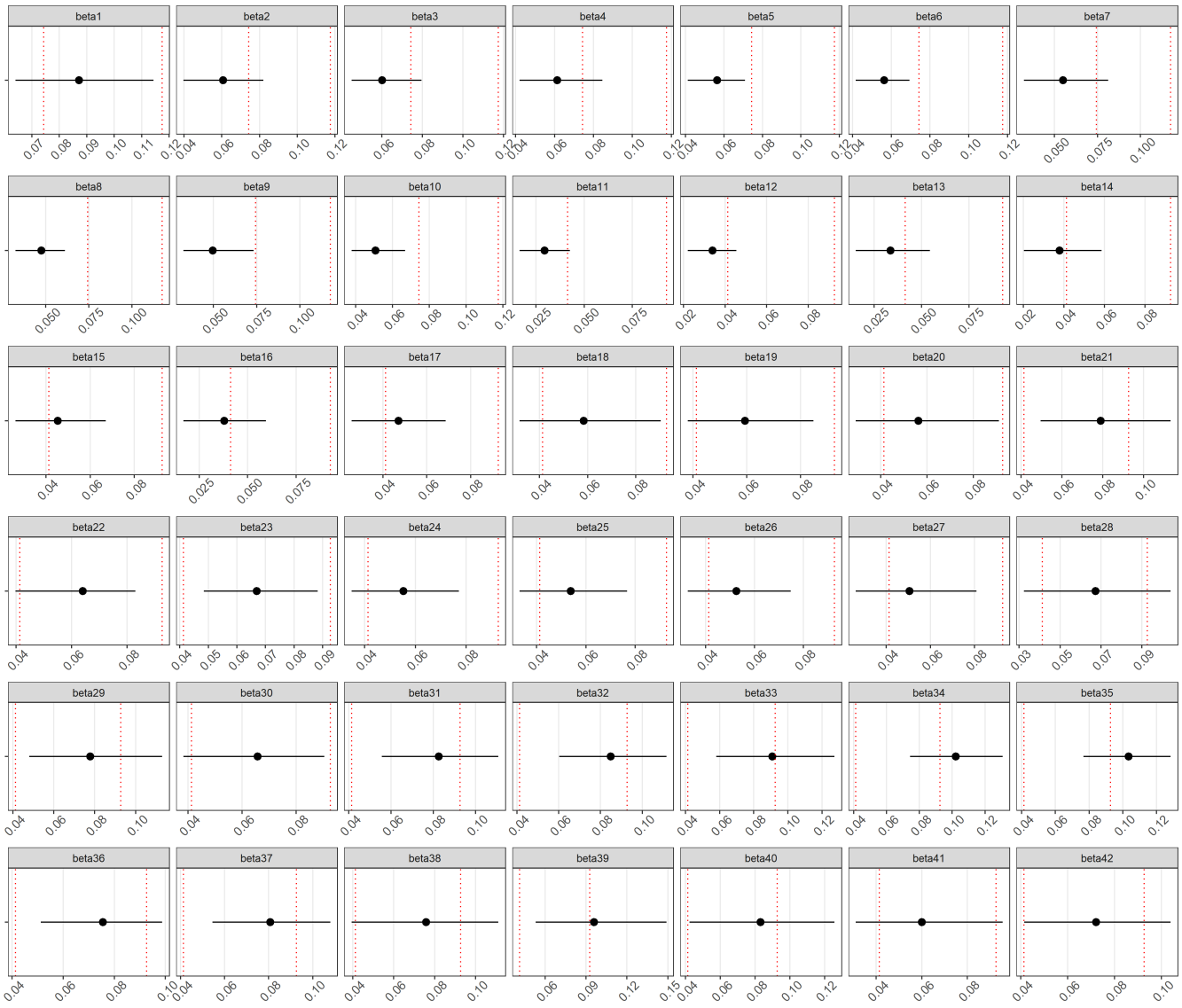


Figure B1: Posterior distributions of fitted values of $\beta(t)$ (piece-wise linear form).

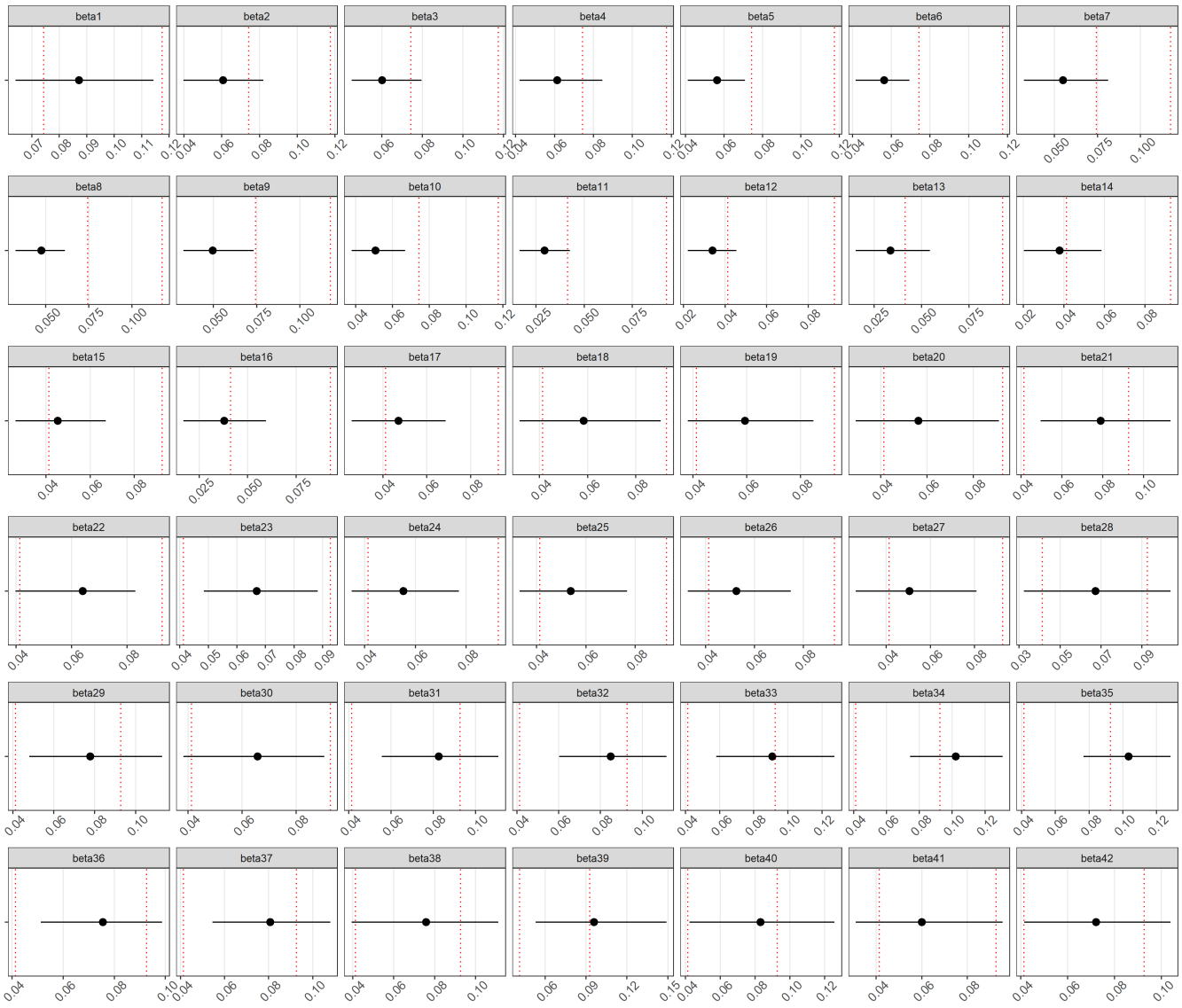


Figure B2: Posterior distributions of all other fitted parameters.

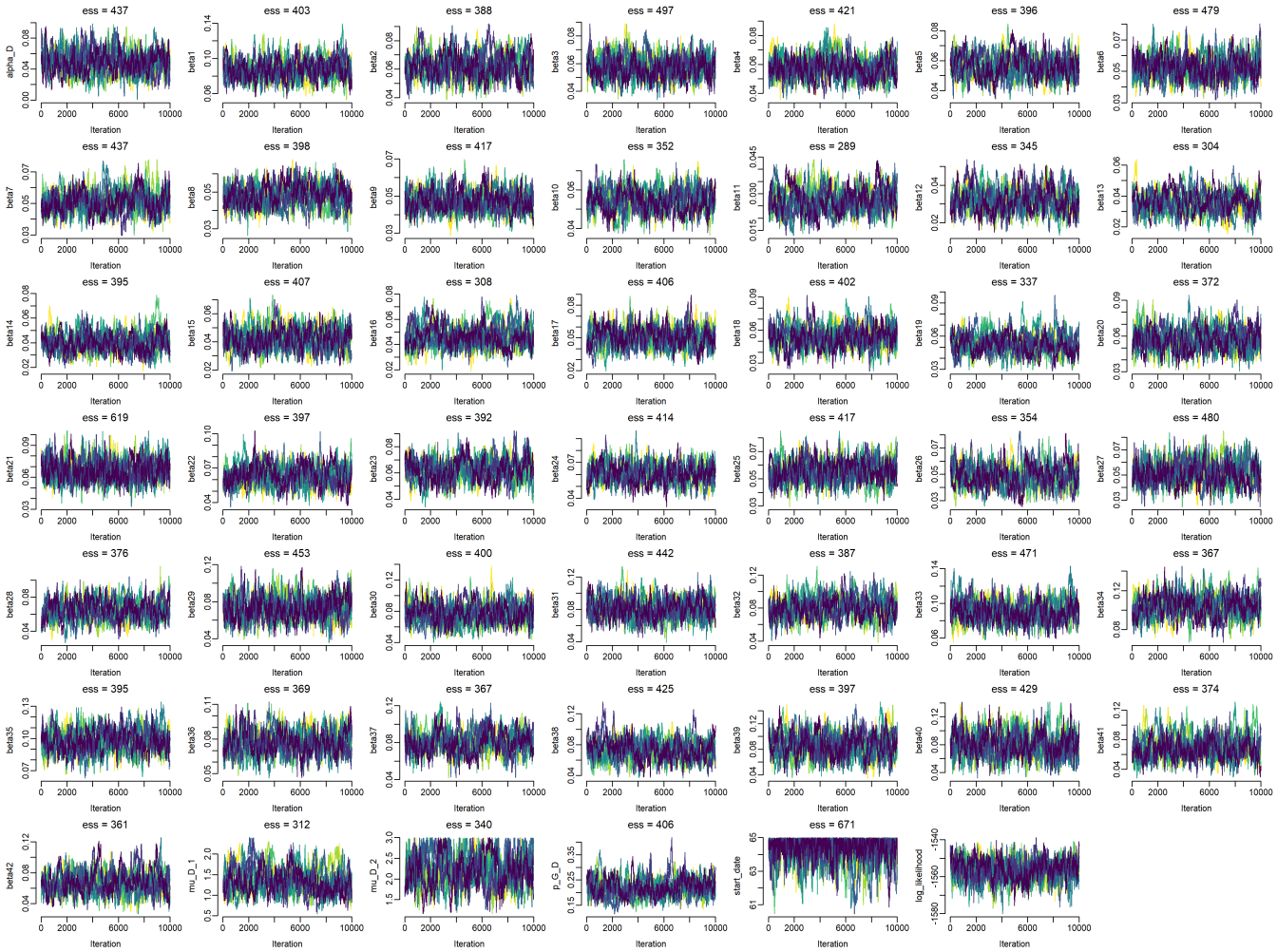


Figure B3: MCMC traceplots of all fitted parameters.

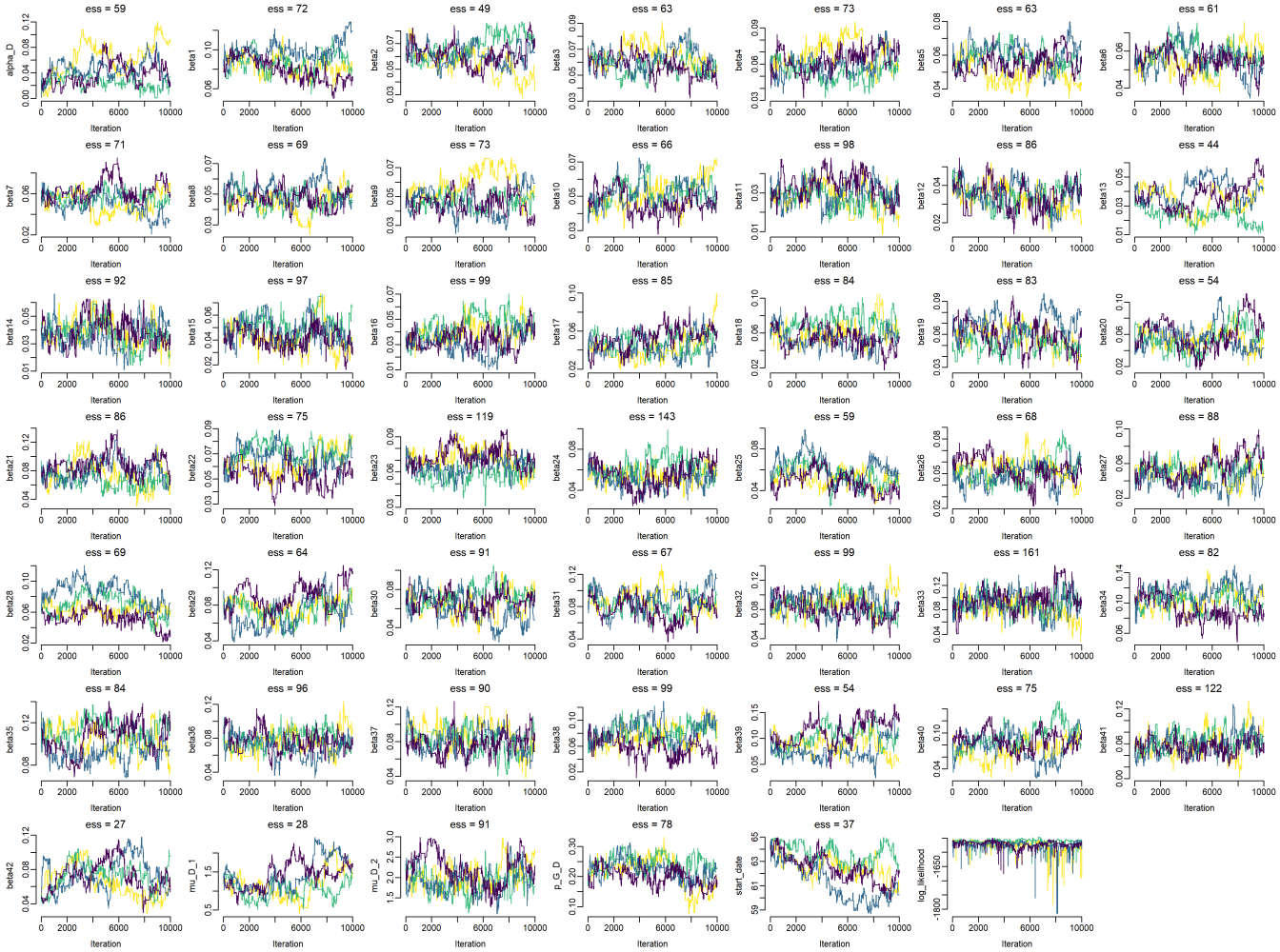


Figure B4: pMCMC traceplots of all fitted parameters.

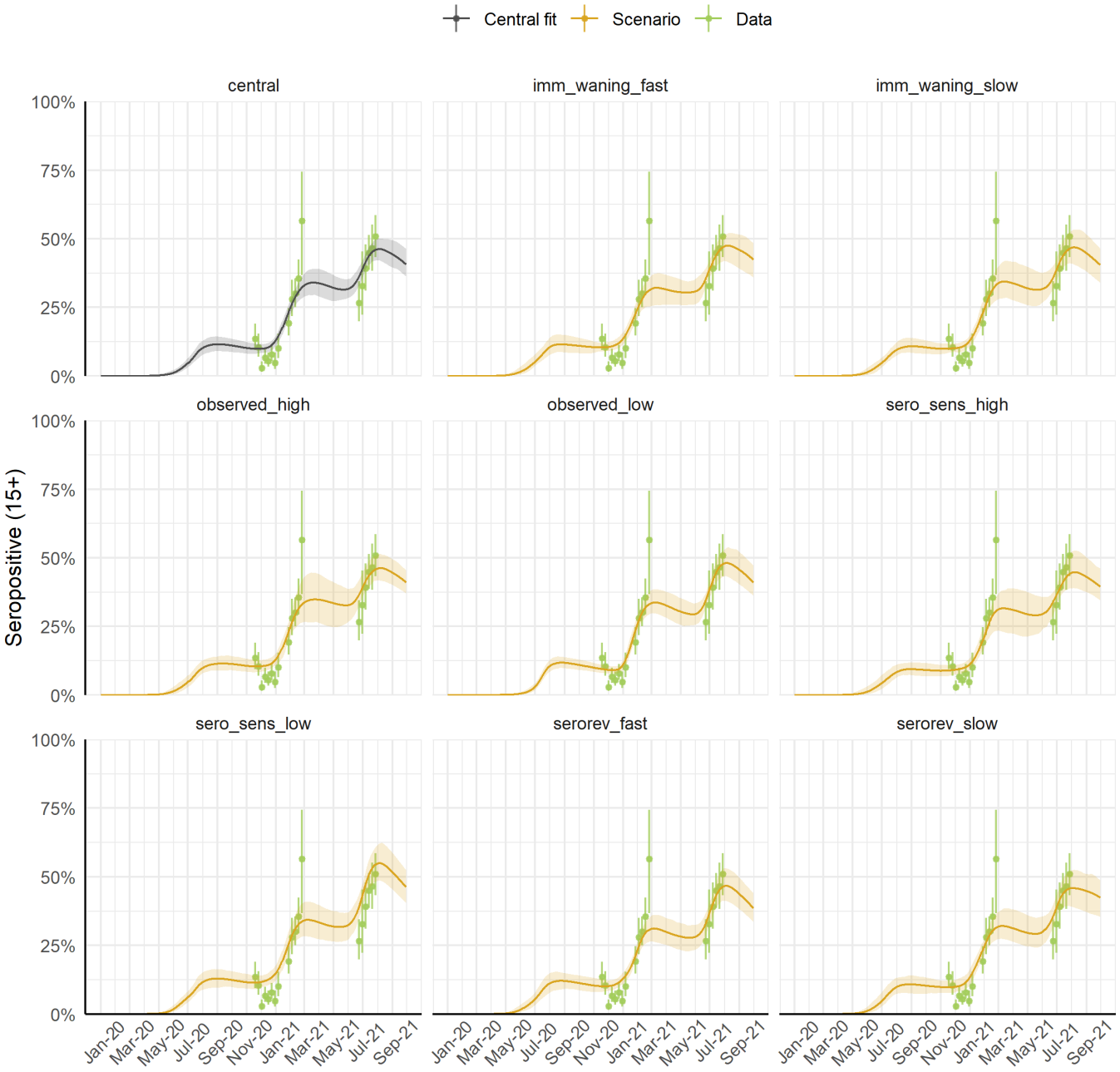


Figure B5: Sensitivity analysis of the fit to seroprevalence data with variations in model parameters of the rate of waning infection-induced immunity, assumed underreporting ('observed' low and high) of pre-pandemic all-cause mortality data for calculating baseline mortality in the model, sensitivity of the serologic test and rate of sero-reversion (negativisation of serologic positive individuals).

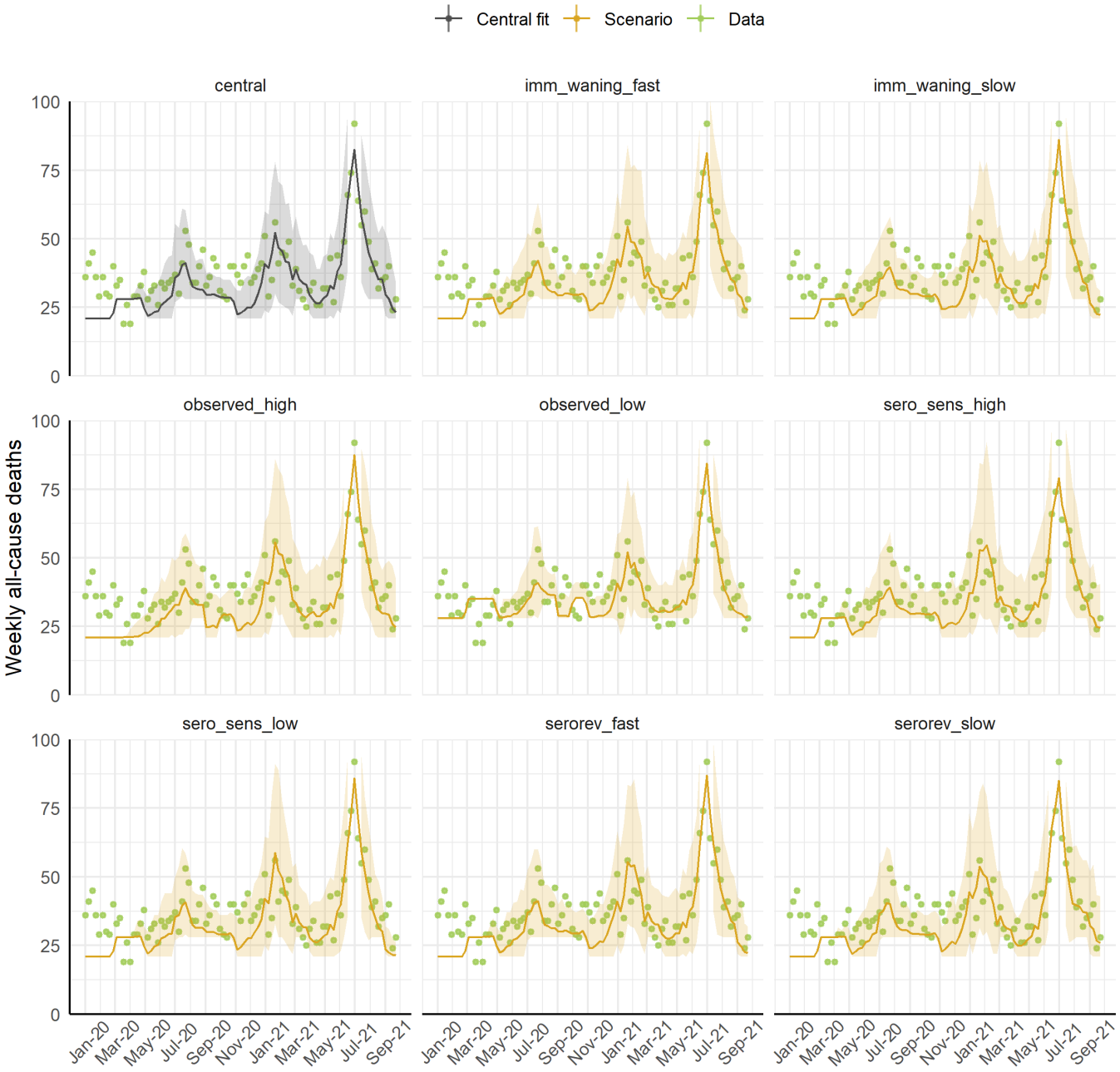


Figure B6: Sensitivity analysis of the fit to all-cause excess mortality data with variations in model parameters of the rate of waning infection-induced immunity, assumed underreporting ('observed' low and high) of pre-pandemic all-cause mortality data for calculating baseline mortality in the model, sensitivity of the serologic test and rate of sero-reversion (negativisation of serologic positive individuals).



Figure B7: Impact of sensitivity analysis on model-inferred trajectory of the reproduction number accounting for immunity R_t^{eff} (eff_Rt_general, red) and void of it R_t (Rt_general, blue), given variations in model parameters of the rate of waning infection-induced immunity, assumed underreporting ('observed' low and high) of pre-pandemic all-cause mortality data for calculating baseline mortality in the model, sensitivity of the serologic test and rate of sero-reversion (negativisation of serologic positive individuals).

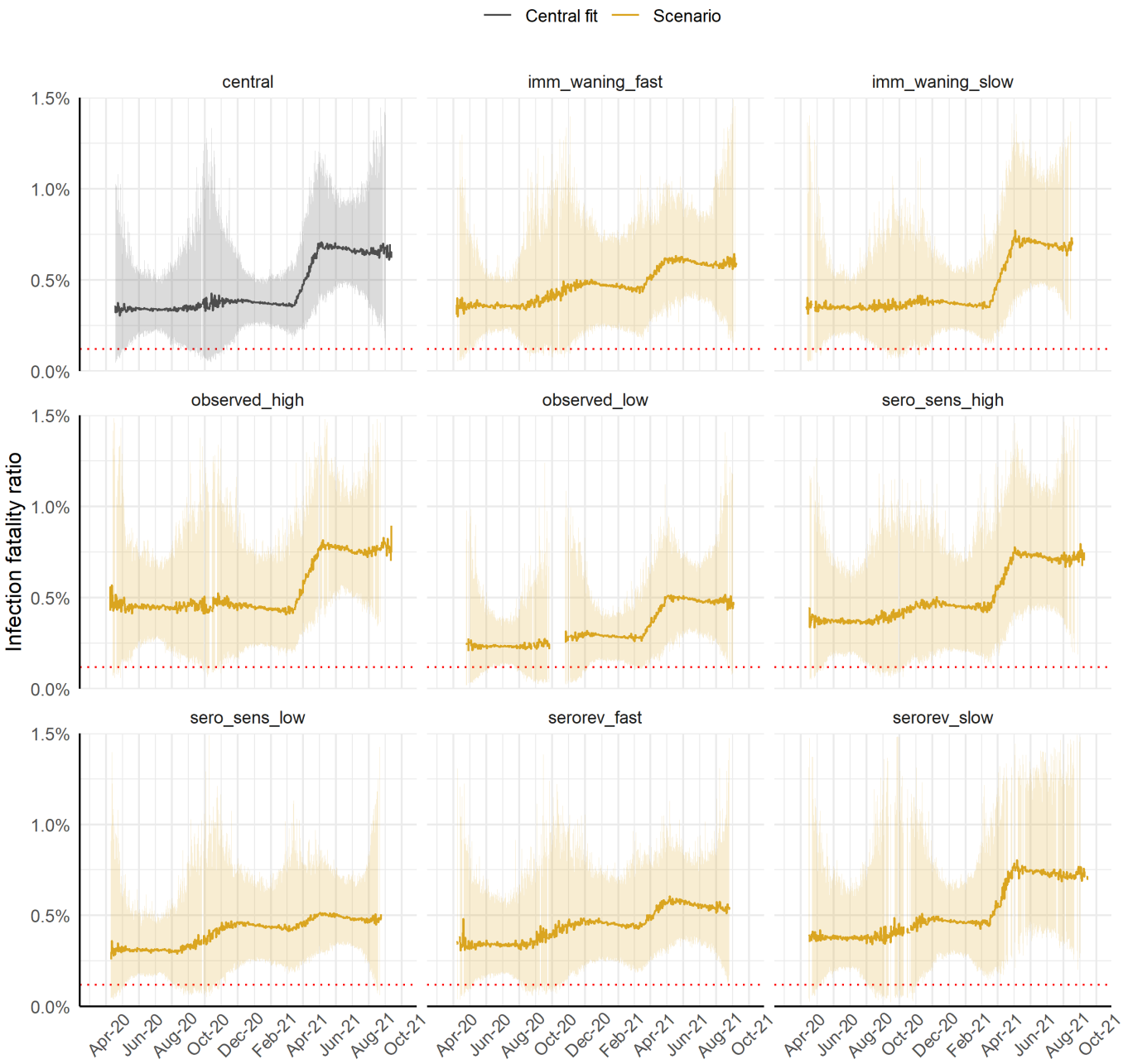


Figure B8: Impact of sensitivity analysis on model-inferred trajectory of the infection fatality ratio over time IFR_t, given variations in model parameters of the rate of waning infection-induced immunity, assumed underreporting ('observed' low and high) of pre-pandemic all-cause mortality data for calculating baseline mortality in the model, sensitivity of the serologic test and rate of sero-reversion (negativisation of serologic positive individuals).

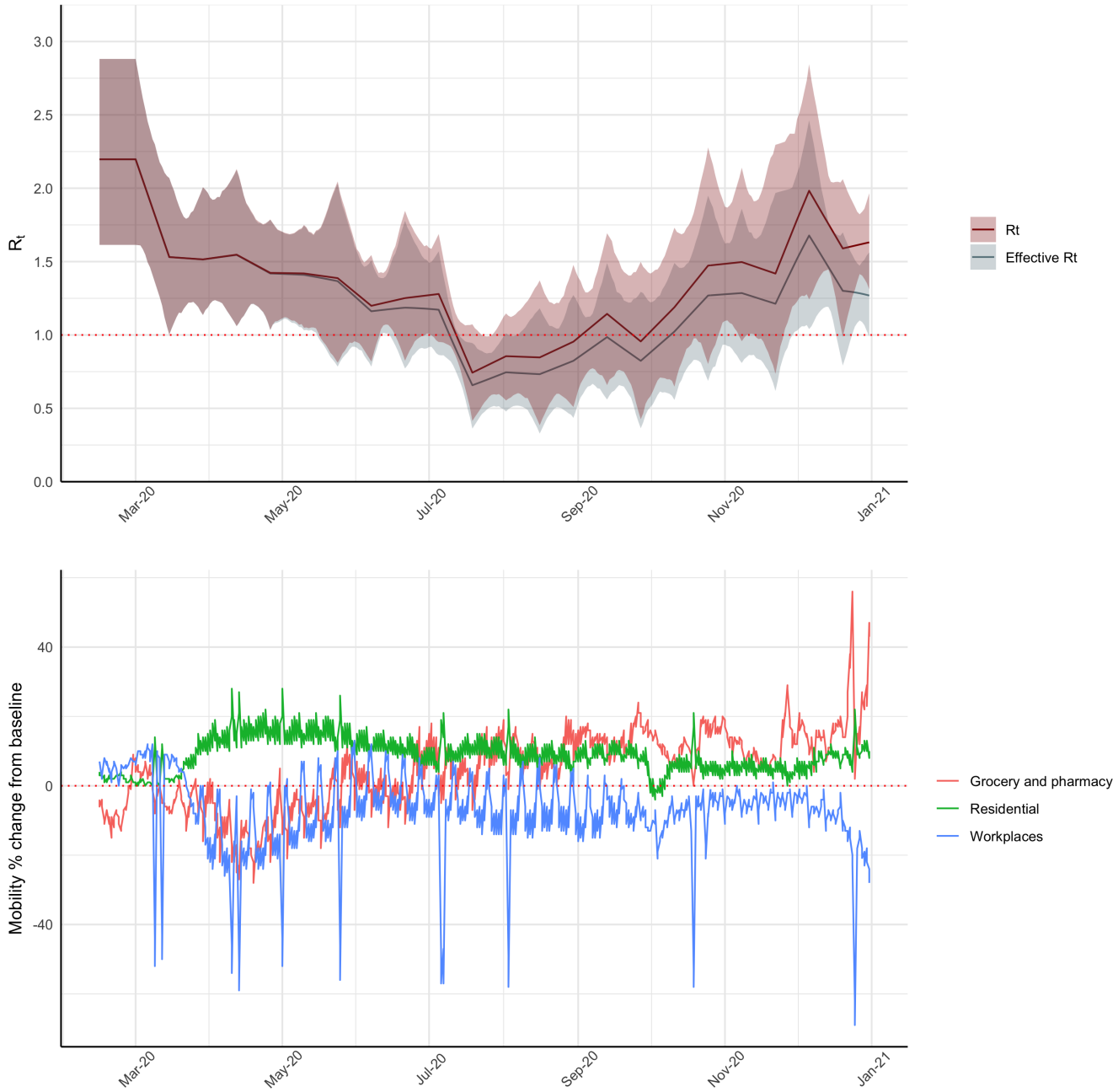


Figure B9: Comparison of model-inferred R_t and R_t^{eff} (without and with accounting for immunity against infection, respectively) in Kabwe, Zambia during 2020 and Google mobility data [193] for Zambia. It should be noted the latter is representative of mobility within Lusaka, the capital of Zambia, rather than Kabwe district. Thus, this comparison is merely intended to be illustrative of the extend to which mobility in the population may or not have resulted in changes in virus transmission, particularly during the first wave, as discussed in Section 4.4.

B.5 Additional results: chapter 5

Figures B16 to B17 show the SARS-CoV-2 pandemic trajectories inferred across simulated counterfactual scenarios, as described in Chapter 5. The model showed that counterfactual scenarios of an earlier roll-out of vaccination with a higher vaccination capacity would have attained more favourable population-level profile of immunity against infection in Kabwe (Figure B19), thus, leading to more infections averted (Figure B21). As expected, on the other hand, the healthcare improvement scenarios had no effect in driving changes in population-level immunity (Figure B20).

Figures B22 to B25 present the sensitivity analysis I conducted of varying the uptake of vaccines on the inferred counterfactual outcomes and population level immunity profile.

Figure B26 shows the posterior distribution (200 stochastic simulations per scenario) of the incremental cost-effectiveness ratio and net monetary benefit across counterfactual scenarios of vaccination and healthcare improvement between March 2020 and September 2021 in Kabwe, Zambia, as described in Chapter 5, section 5.2.4. In panels C and D, dashed line shows the lowest plausible opportunity cost-based willingness to pay threshold, at $57.69 or 6% of the GDPpc of Zambia. The highest threshold value considered, at $1,146.64 2020-USD or 120% the GDPpc of Zambia, is not shown, as all simulated interventions were below it (Table B4).

Figures B27 and B28 show heatmaps of the disability-adjusted life years (DALYs) lost accrued by age group across vaccination and healthcare improvement scenarios compared to factual scenario, as described in Chapter 5. The 40-44 year-old group accrued the most DALYs as a result of the age-specific severity dynamics of SARS-CoV-2 and the age distribution of the population of Kabwe.

Tornado plots B29 to B32 show the results of the one-way sensitivity analysis, varying the costs of vaccination, inpatient or outpatient care across counterfactual vaccination and healthcare improvement scenarios, as described in Chapter 5.

Lastly, tables B3 to B6 present the median and 95%CrI across 200 stochastic trajectories simulated for each the 30 counterfactual scenarios of vaccination roll-out and healthcare improvement, as described in Chapter 5.

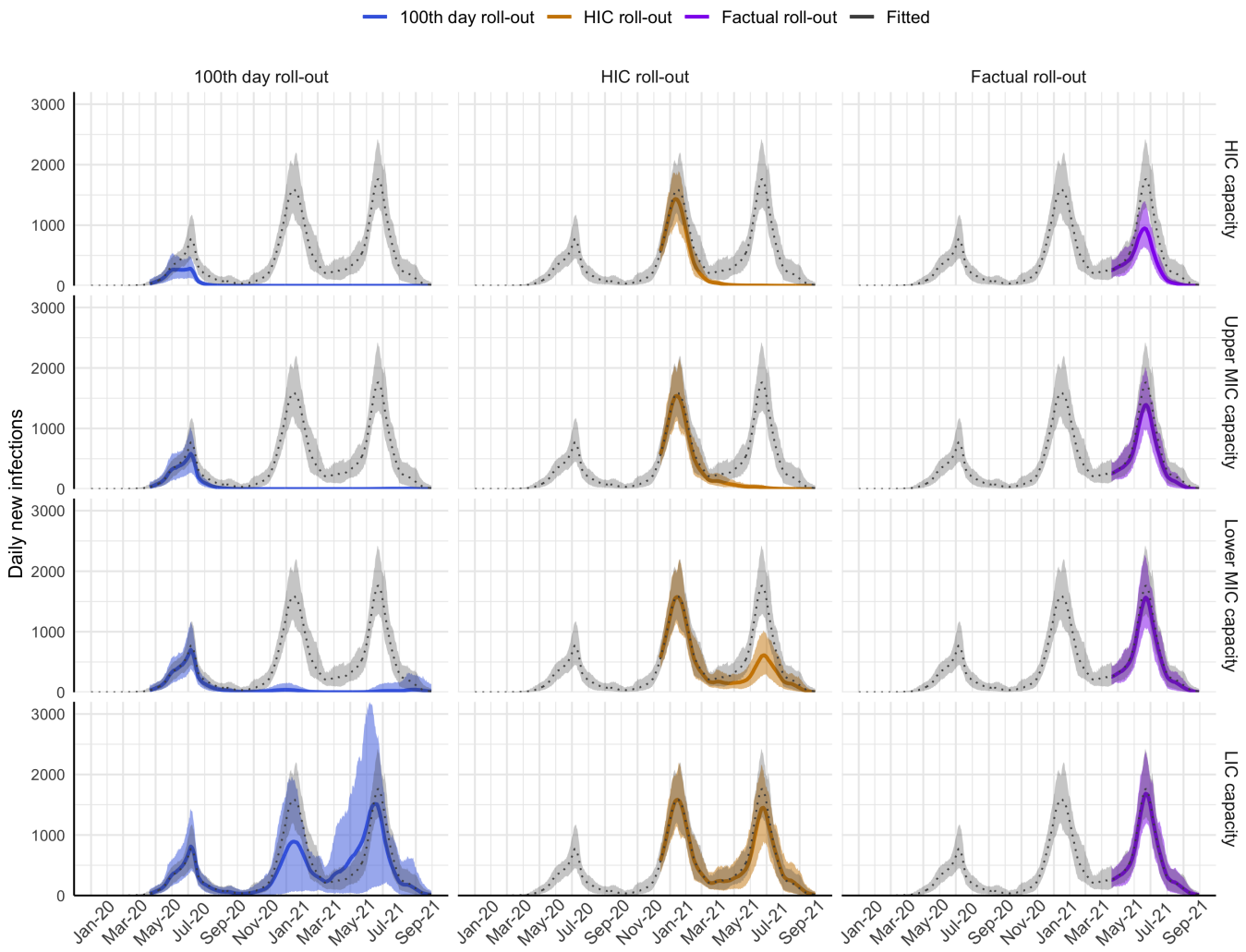


Figure B10: Daily infections in counterfactual vaccination scenarios.

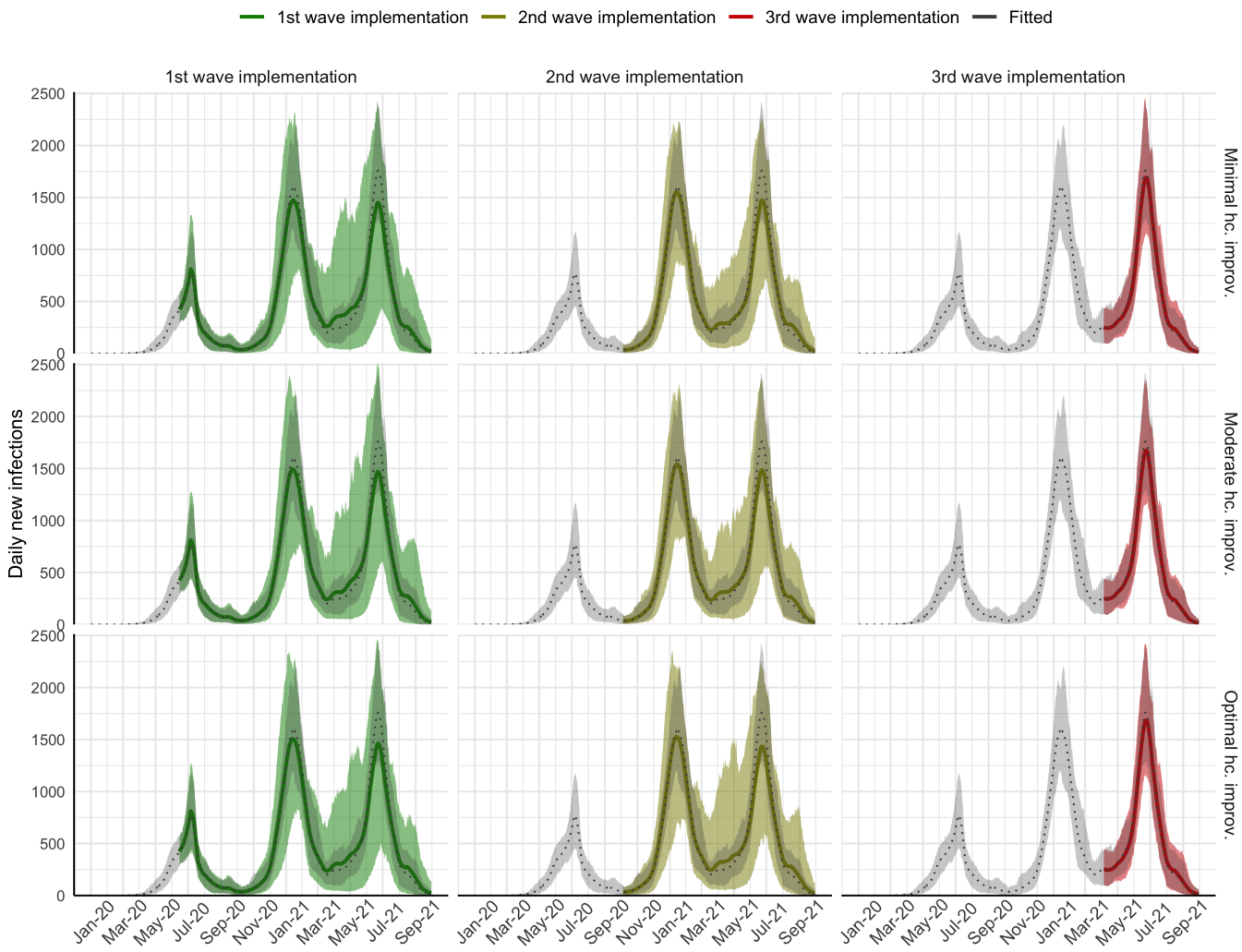


Figure B11: Daily infections in counterfactual healthcare improvement scenarios.

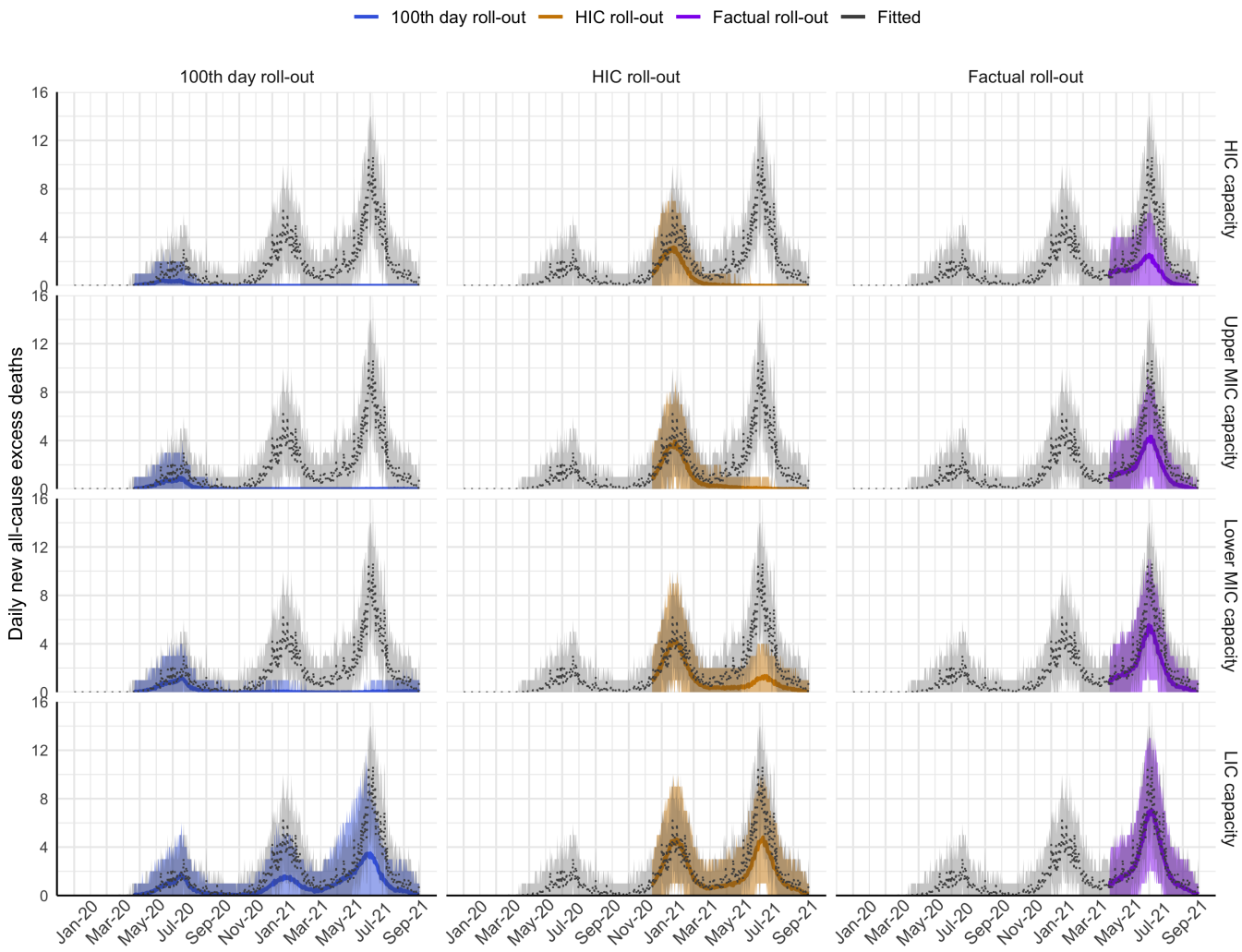


Figure B12: Daily all-cause deaths in counterfactual vaccination scenarios.

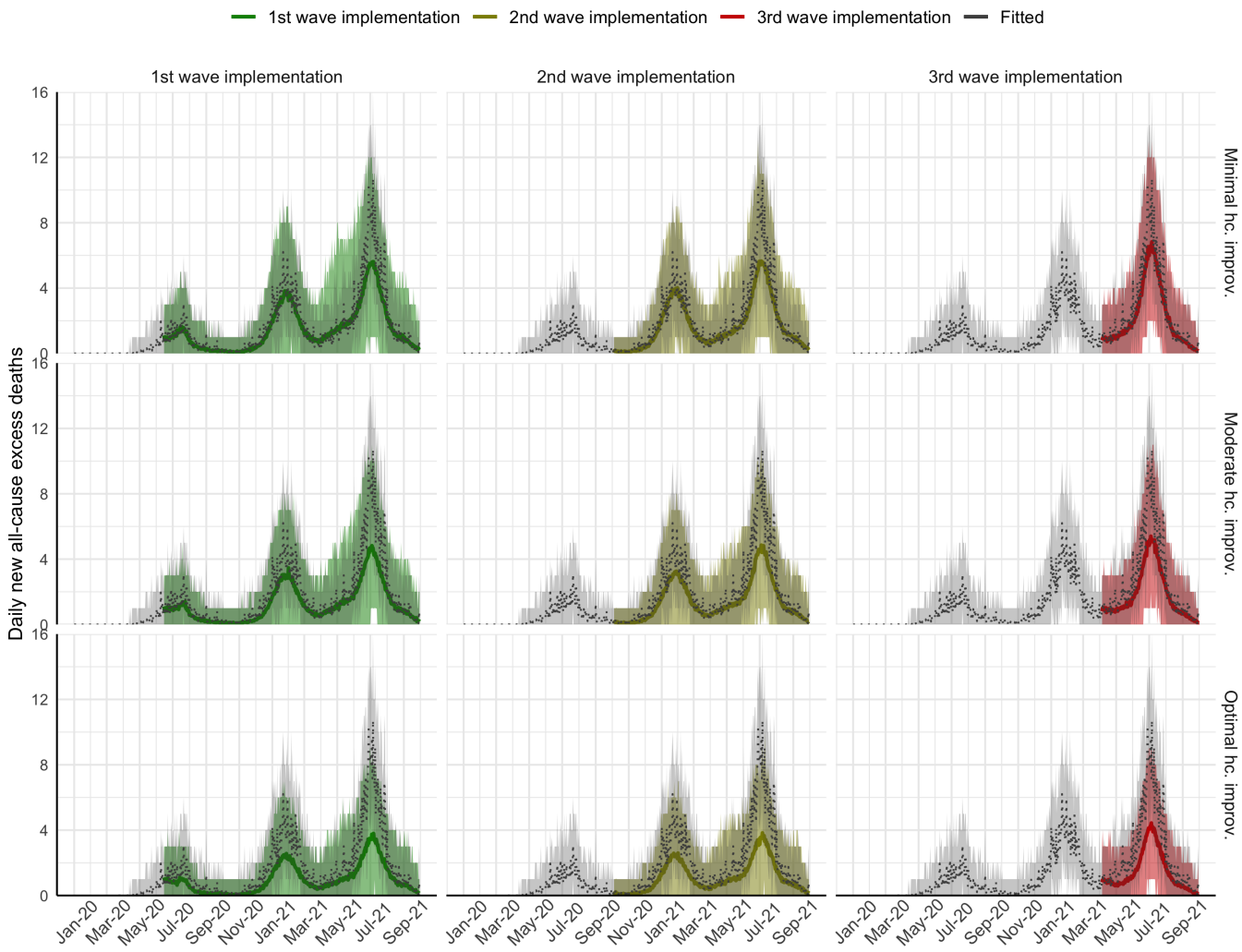


Figure B13: Daily all-cause deaths in counterfactual healthcare improvement scenarios.



Figure B14: Inferred trajectories of the effective reproduction number in the absence of and accounting for immunity, R_t and R_t^{eff}, respectively, in counterfactual vaccination scenarios.



Figure B15: Inferred trajectories of the effective reproduction number in the absence of and accounting for immunity, R_t and R_t^{eff}, respectively, in counterfactual healthcare improvement scenarios.

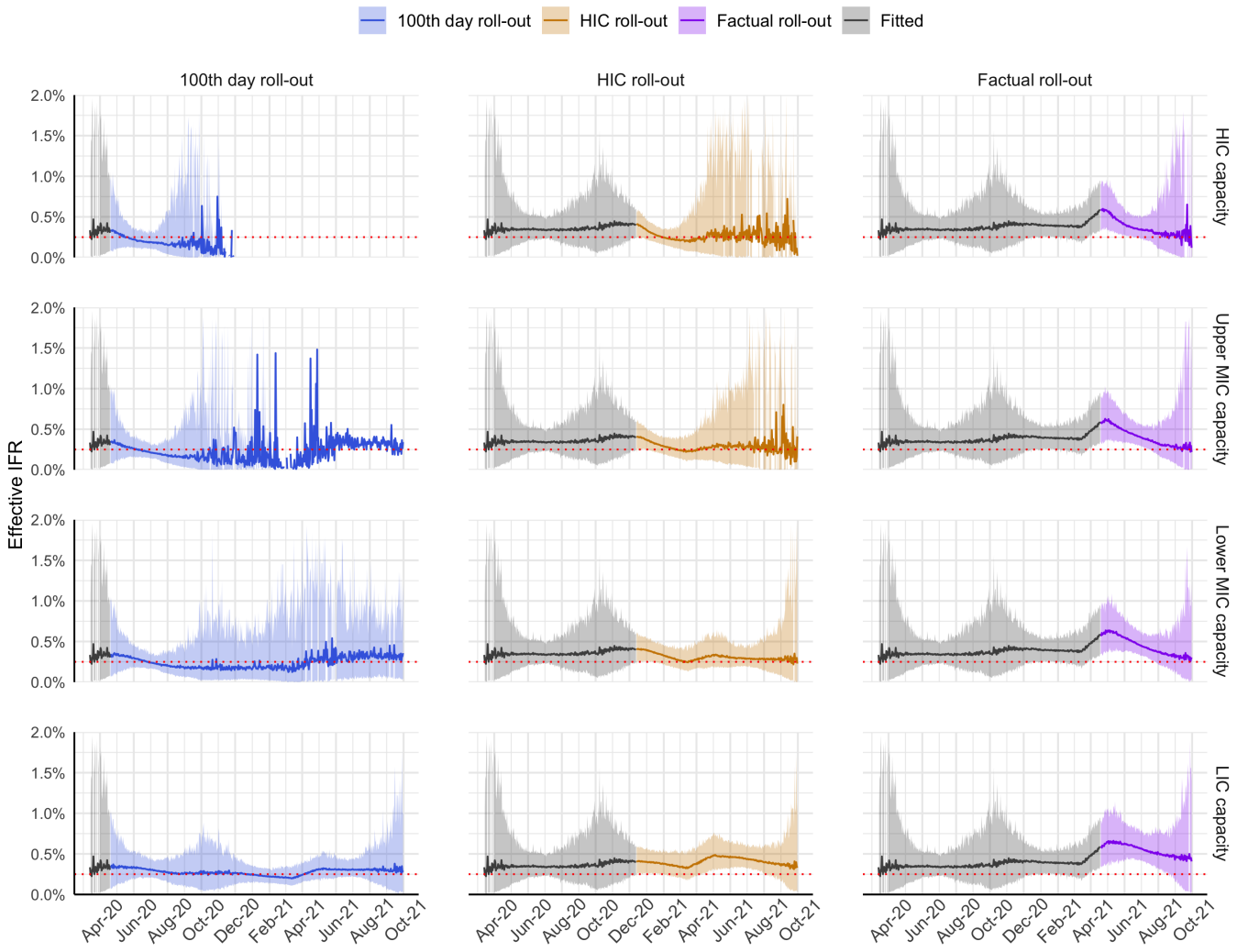


Figure B16: Inferred trajectories of the effective infection fatality ratio, IFR_t, in counterfactual vaccination scenarios.

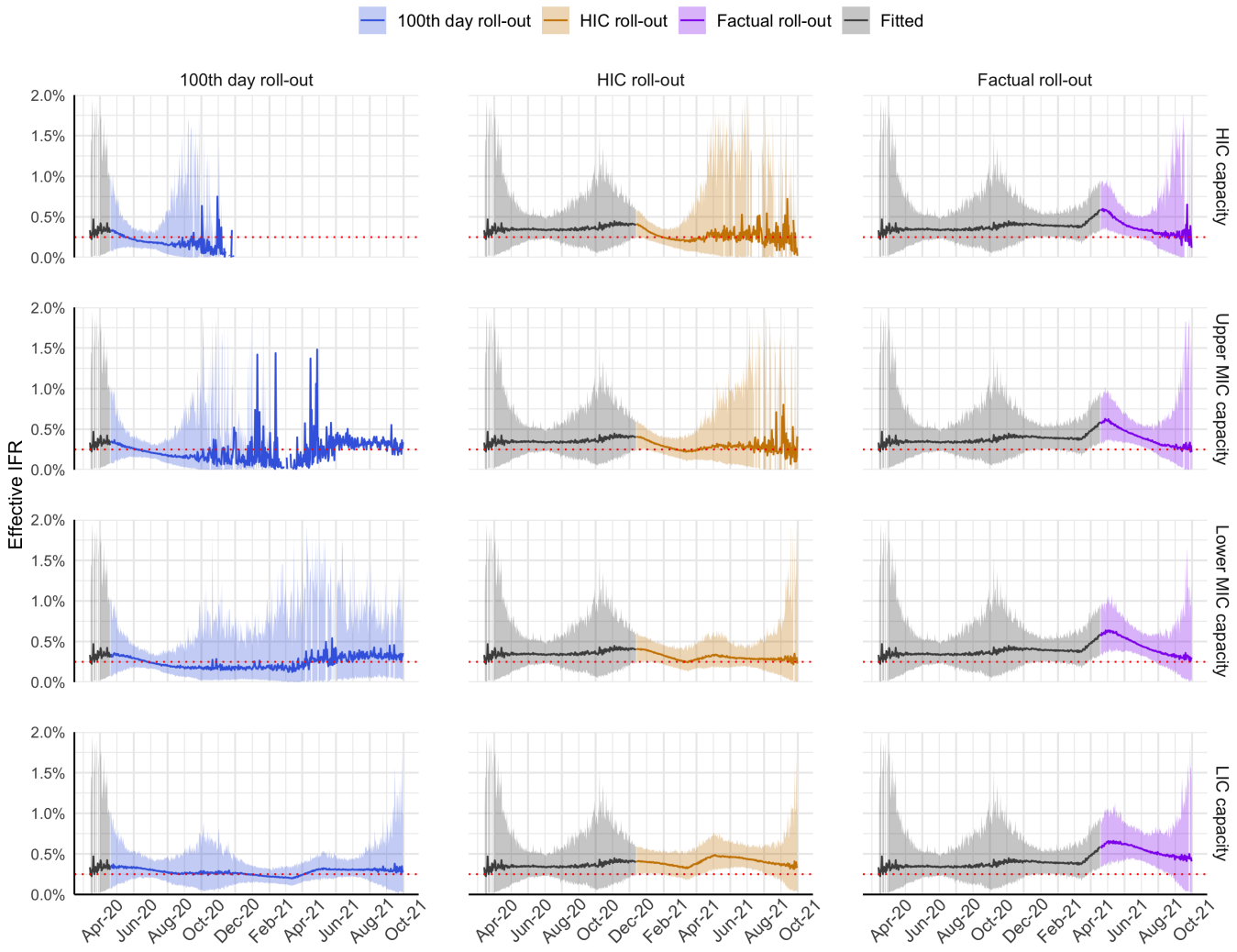


Figure B17: Inferred trajectories of the effective infection fatality ratio, IFR_t, in counterfactual healthcare improvement scenarios.

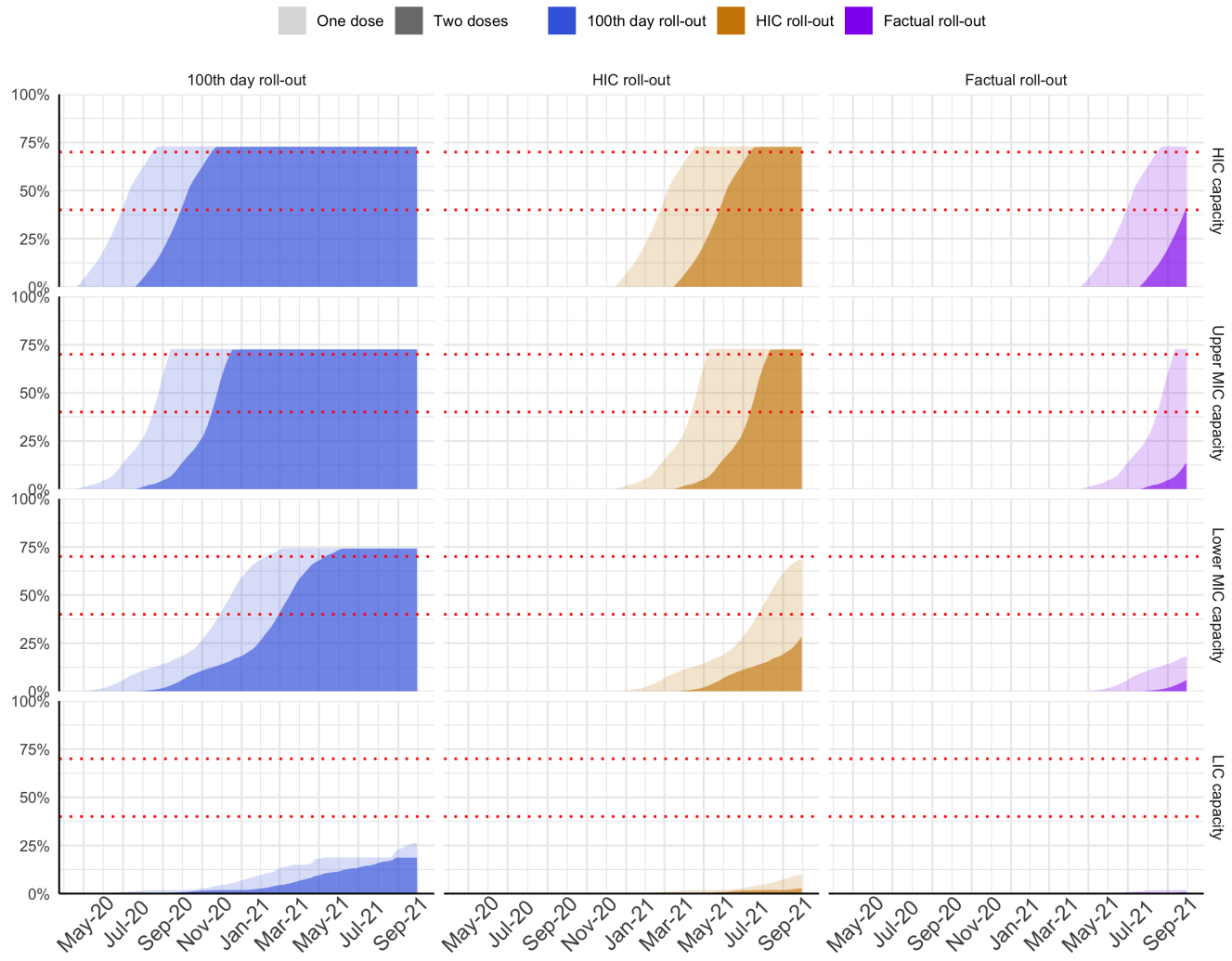


Figure B18: Proportion of the population vaccinated with one and two doses in Kabwe across counterfactual vaccination scenarios.

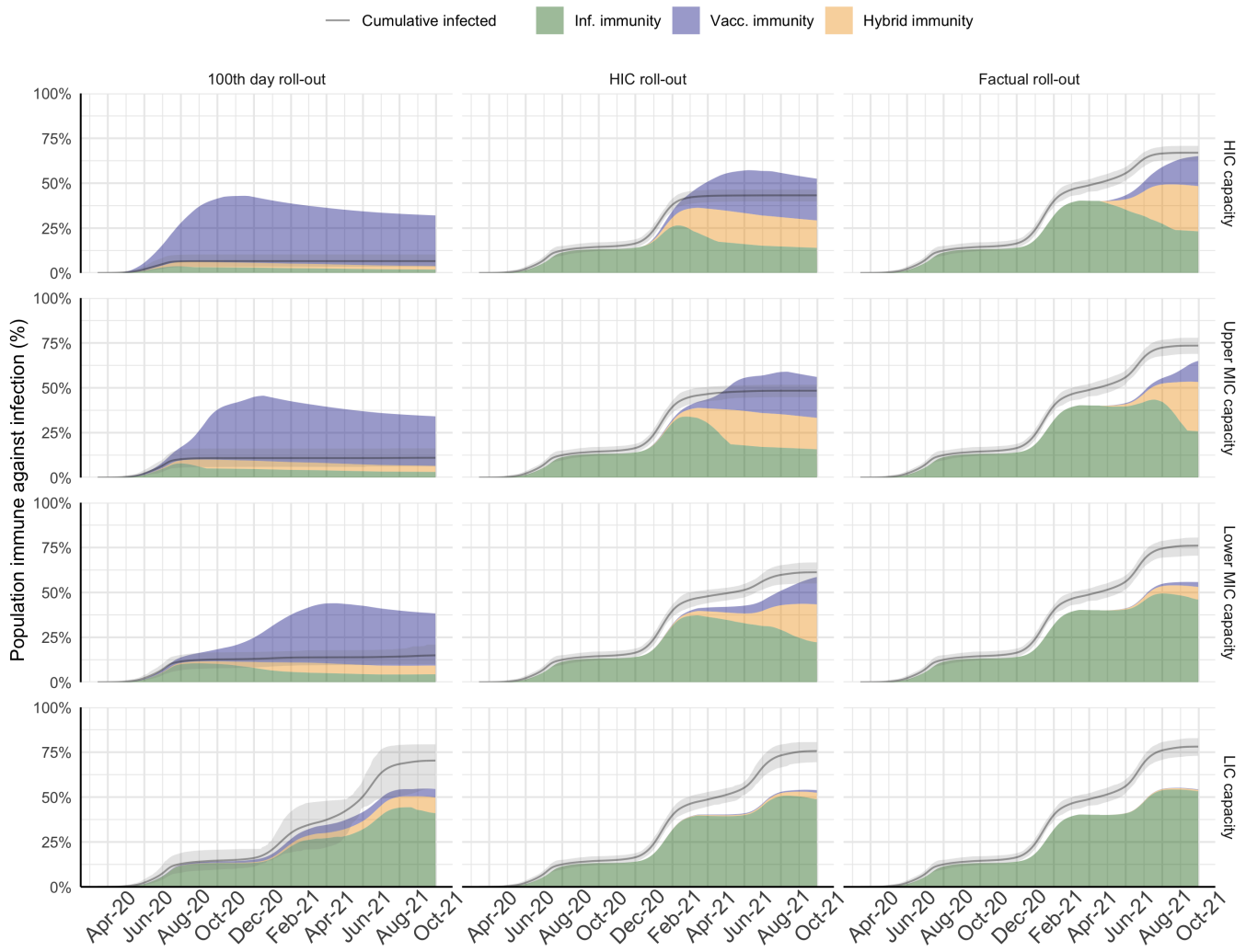


Figure B19: Population-level profile of immunity against infection in the population of Kabwe with counterfactual vaccination scenarios.

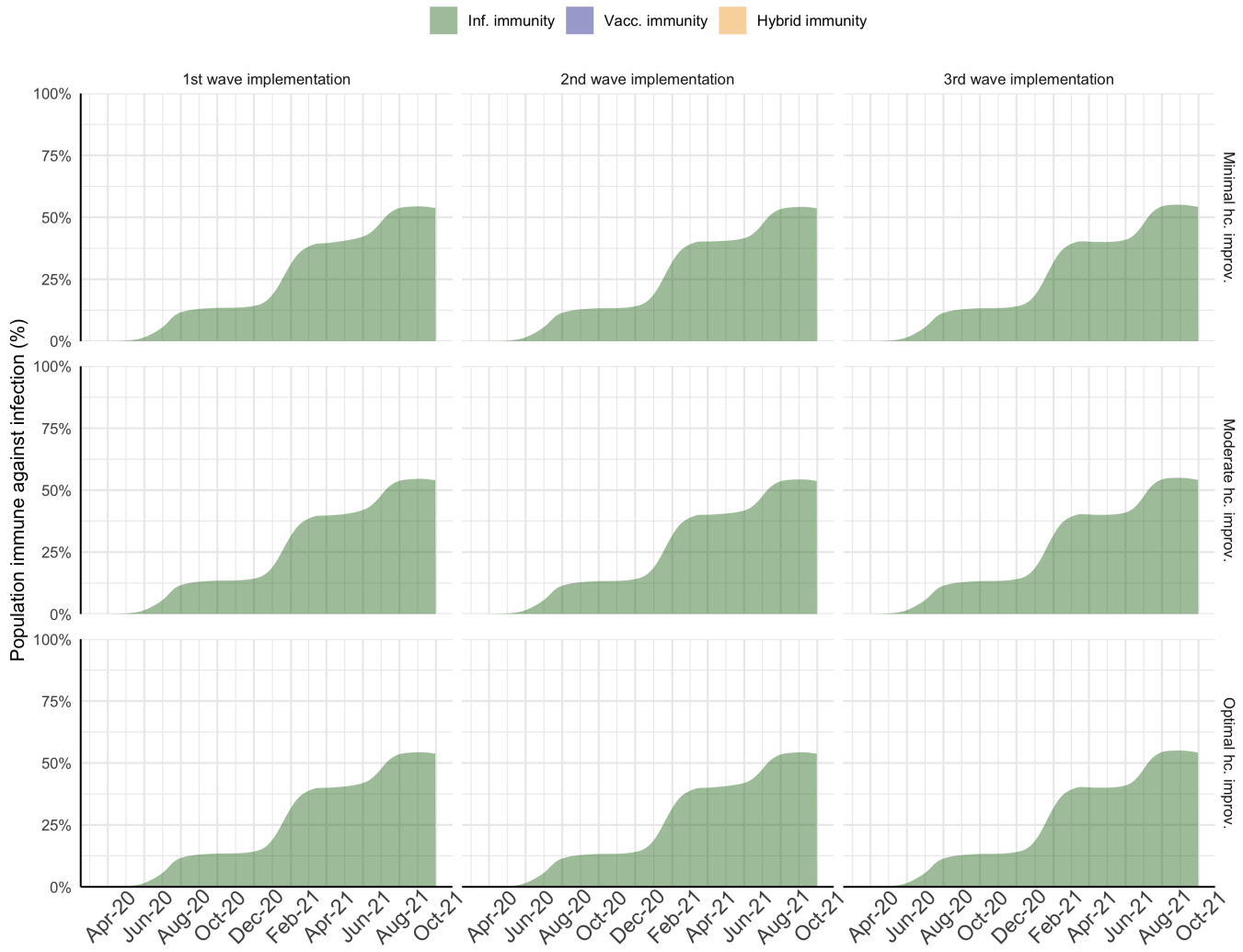


Figure B20: Population-level profile of immunity against infection in the population of Kabwe with counterfactual healthcare improvement scenarios.

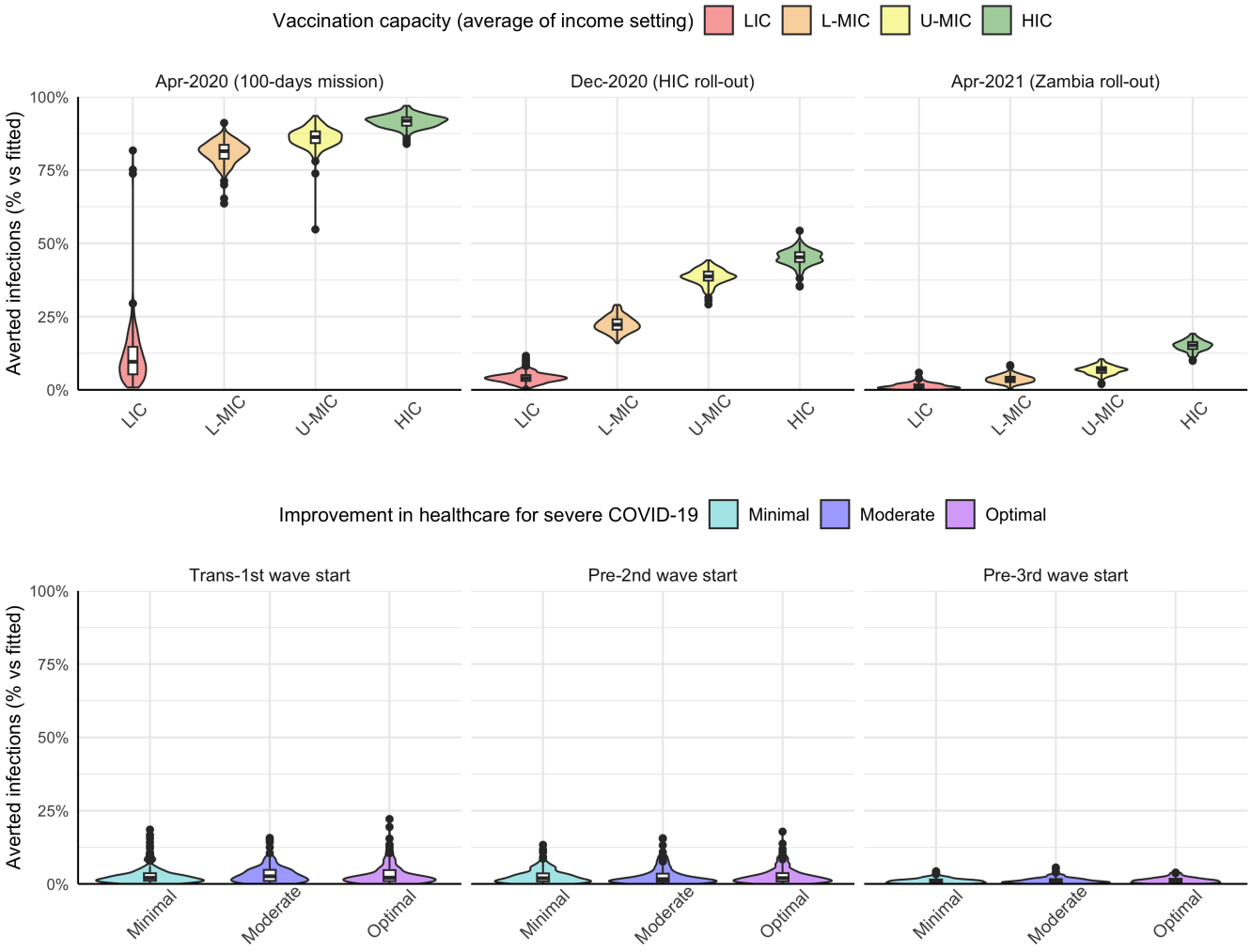


Figure B21: Accrued infections averted in the population of Kabwe in counterfactual scenarios.

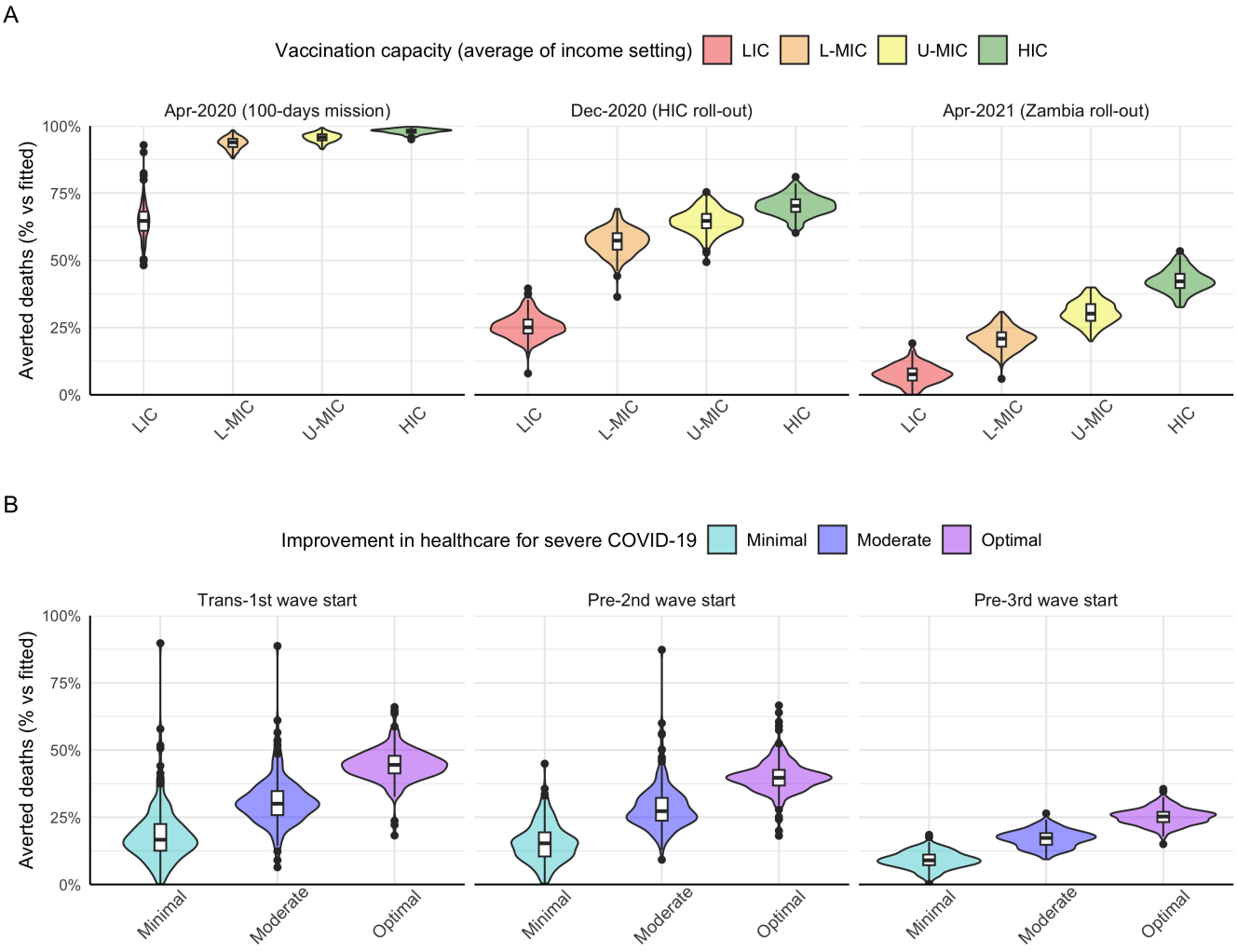


Figure B22: Accrued deaths averted in the population of Kabwe in counterfactual scenarios, assuming a higher uptake of vaccines, at 95%, 90% and 80% amongst 50+, 40-49-, and 12–39-year-olds, respectively. Note that in healthcare improvement scenarios plots reflect solely the effect of stochastic variation, as no vaccination is assumed in them.

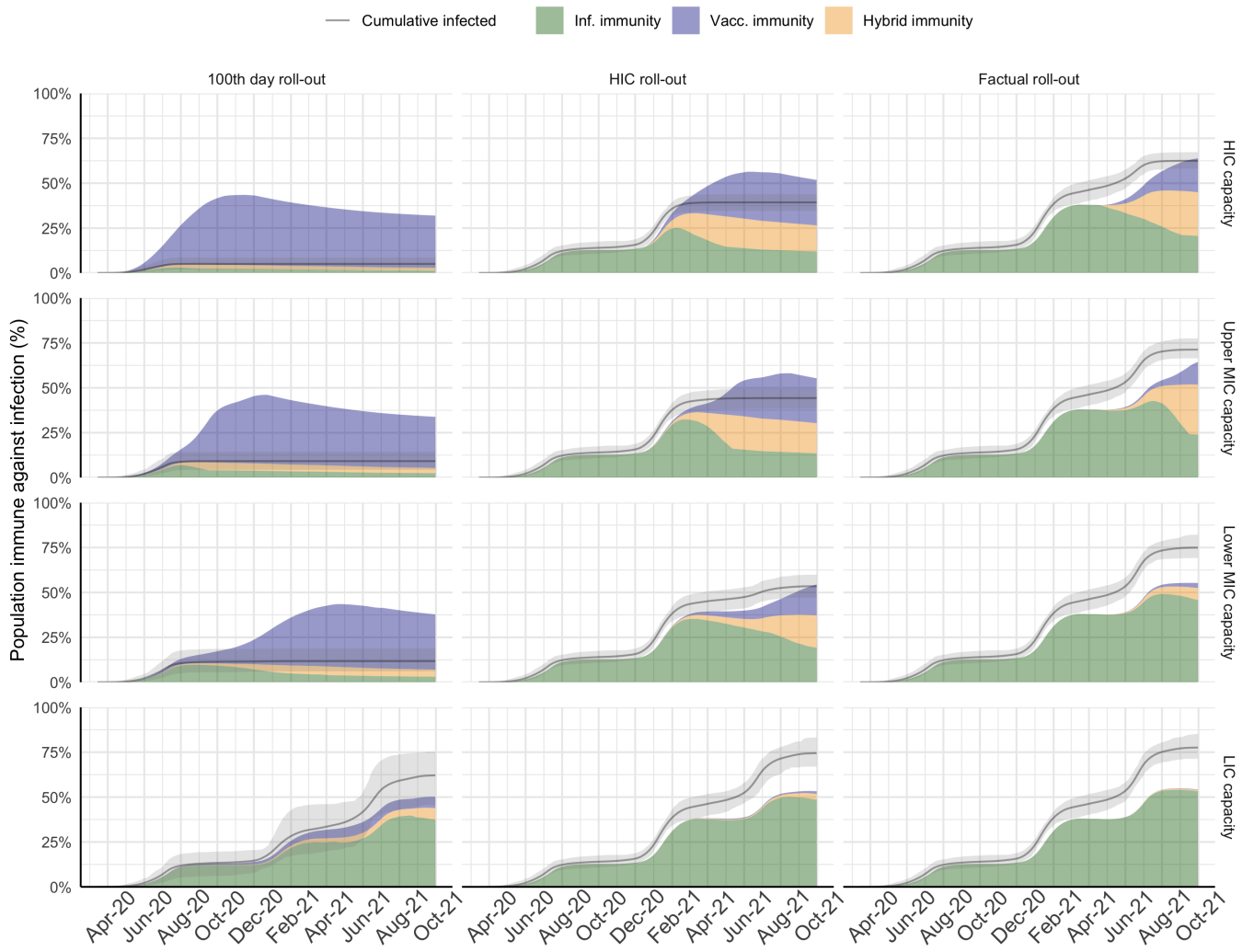


Figure B23: Accrued profile of immunity in the population of Kabwe in vaccination counterfactual scenarios, assuming a higher uptake of vaccines, at 95%, 90% and 80% amongst 50+, 40-49-, and 12–39-year-olds, respectively.

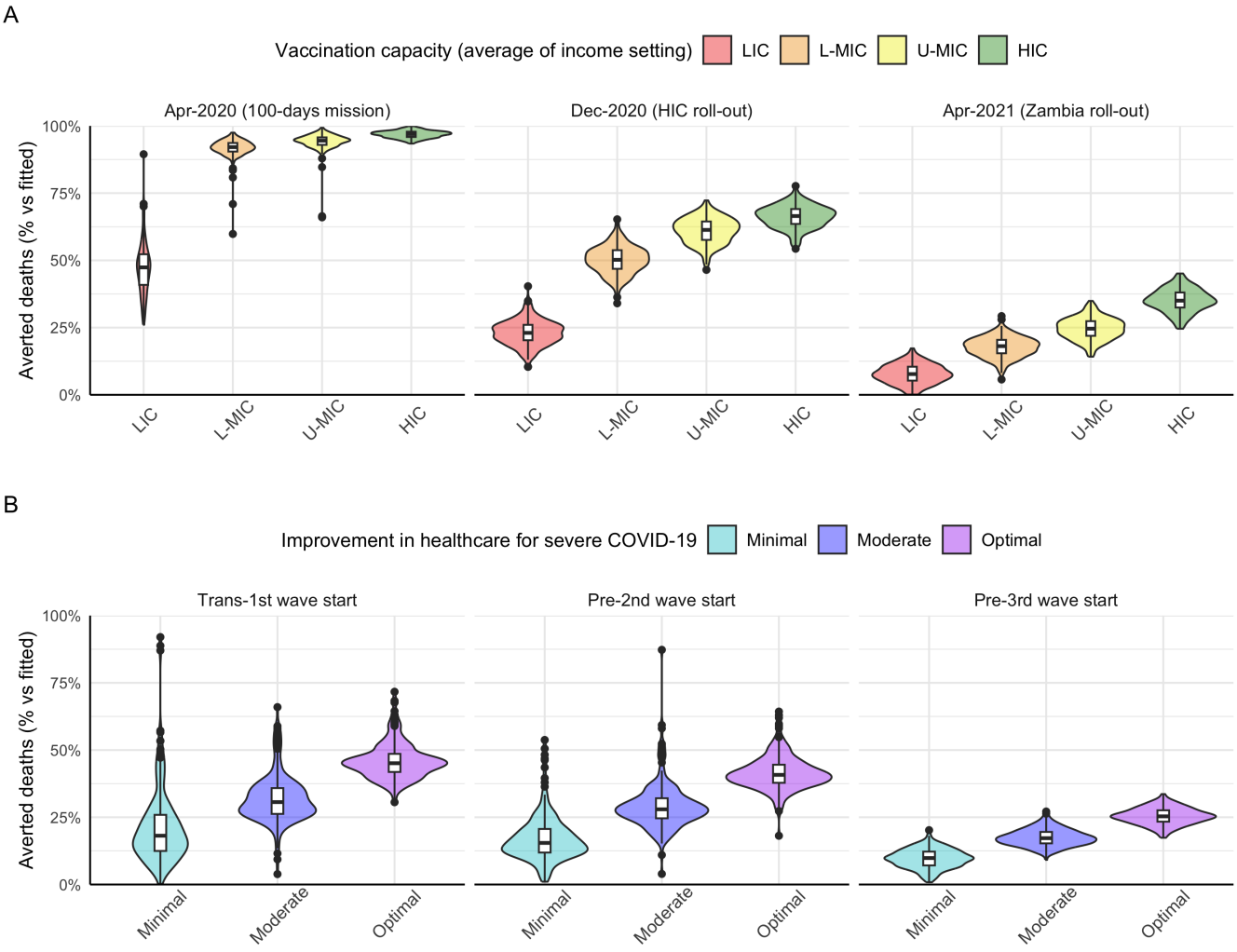


Figure B24: Accrued deaths averted in the population of Kabwe in counterfactual scenarios, assuming a lower uptake of vaccines, at 50% across all age groups. Note that in healthcare improvement scenarios plots reflect solely the effect of stochastic variation, as no vaccination is assumed in them.

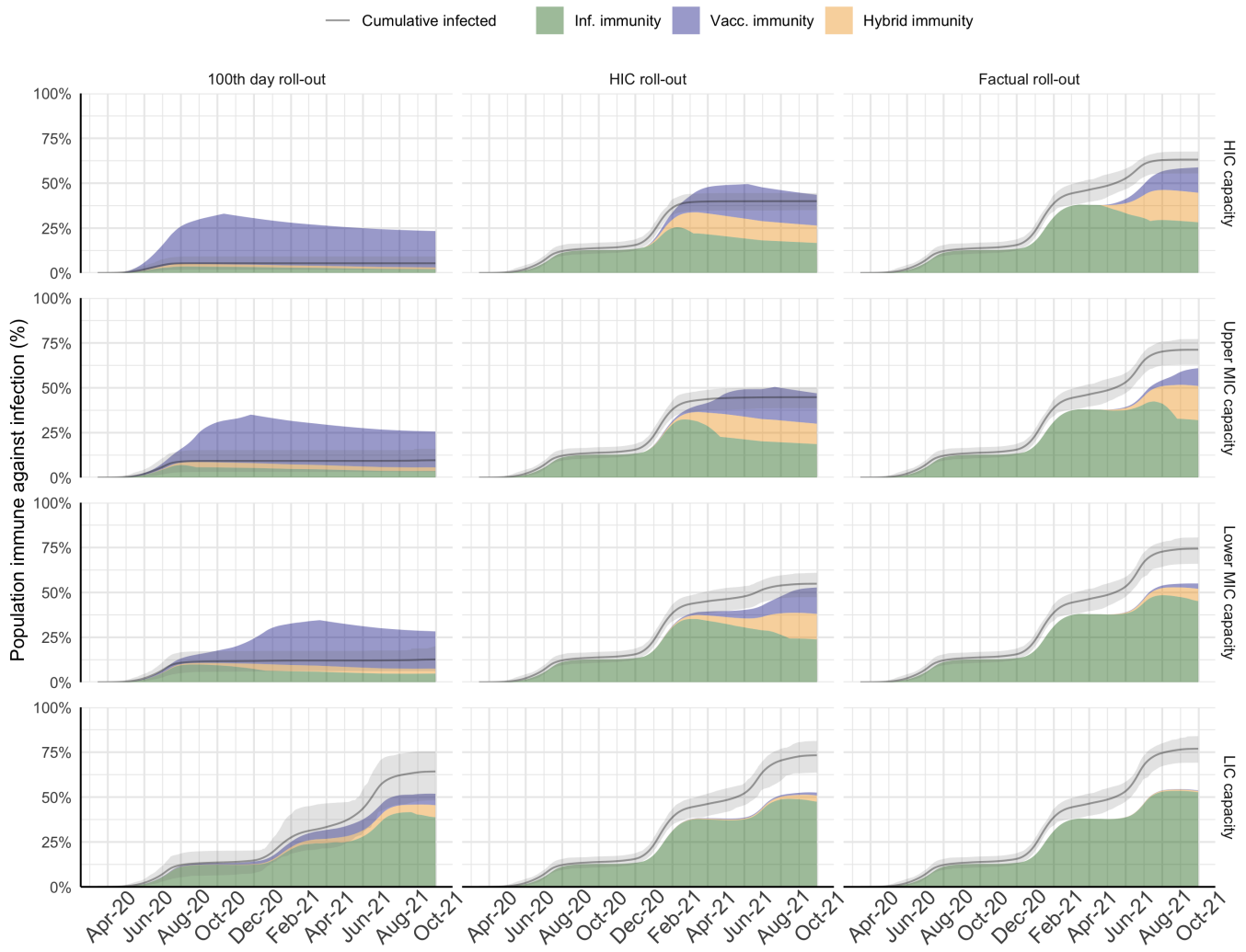


Figure B25: Accrued profile of immunity in the population of Kabwe in vaccination counterfactual scenarios, assuming a lower uptake of vaccines, at 50% across all age groups.

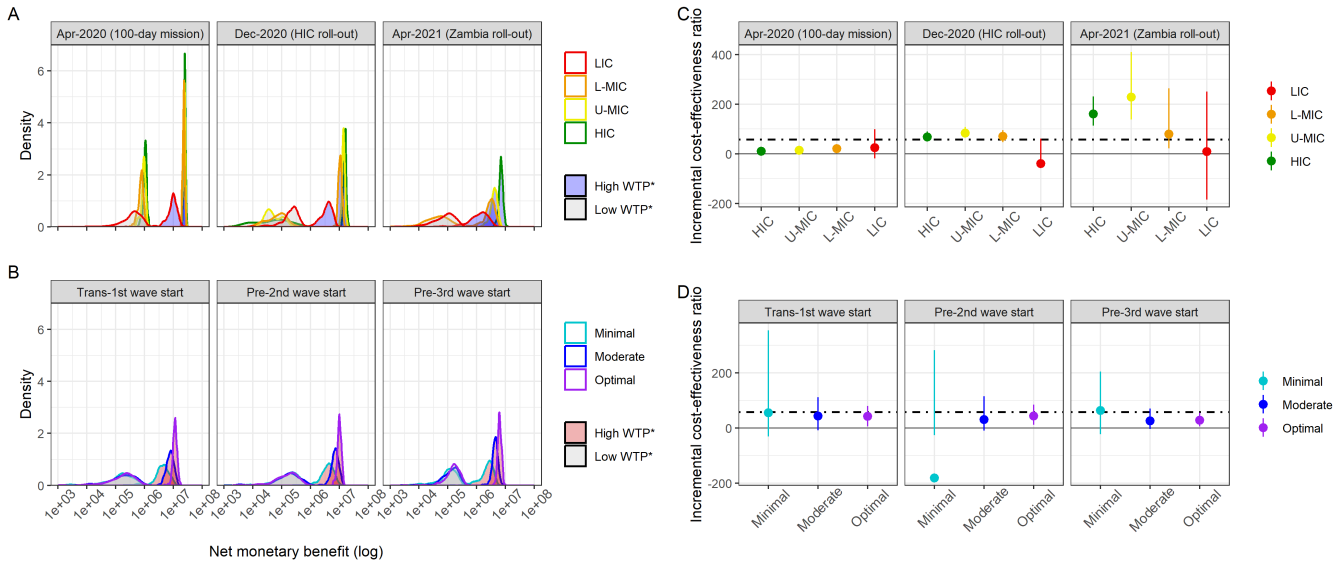


Figure B26: Distribution plots of the incremental cost-effectiveness ratio and net monetary benefit across counterfactual scenarios.

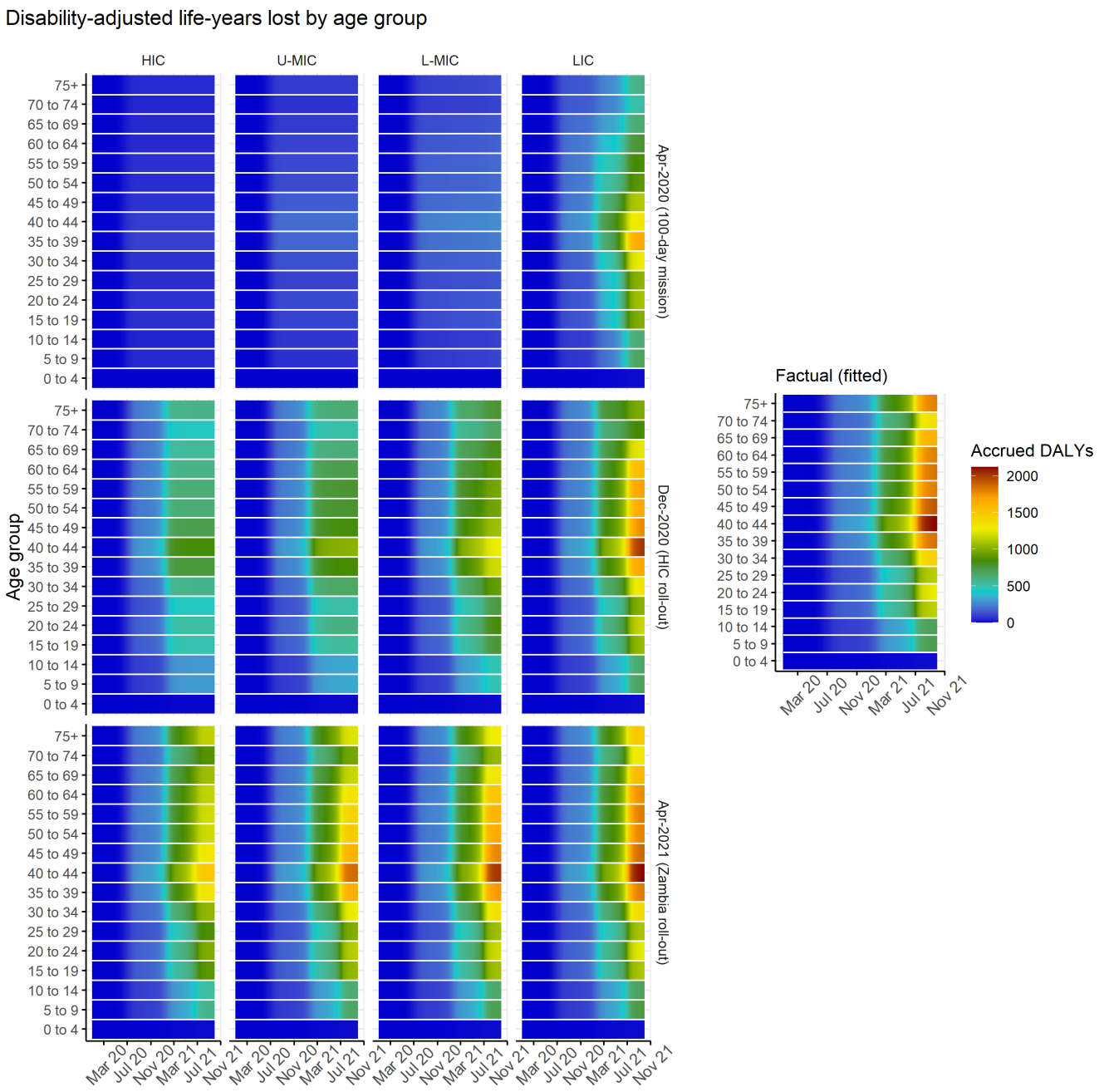


Figure B27: Disability-adjusted life years accrued by age group in factual vs counterfactual vaccination scenarios.

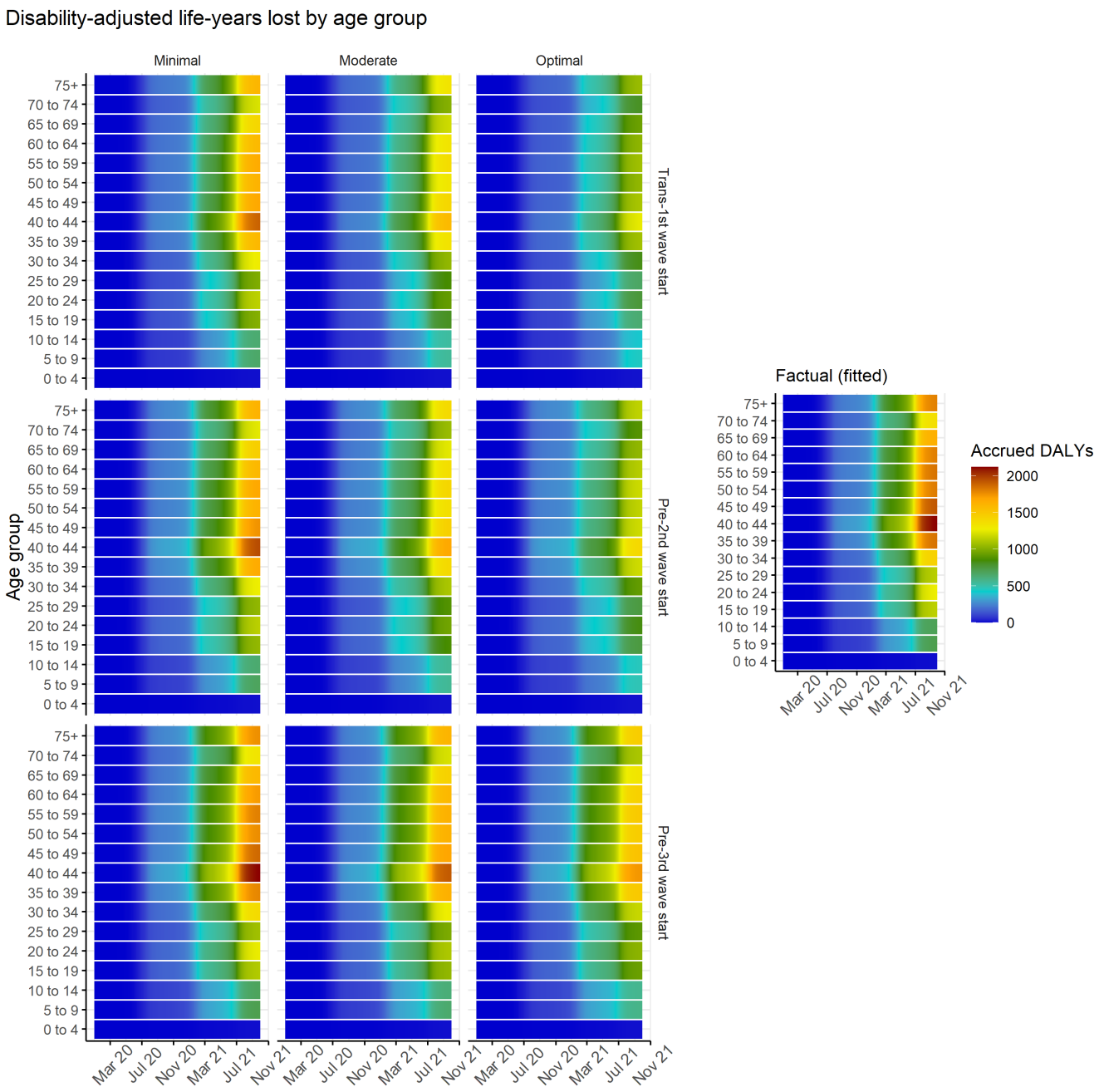


Figure B28: Disability-adjusted life years accrued by age group in factual vs counterfactual healthcare improvement scenarios.

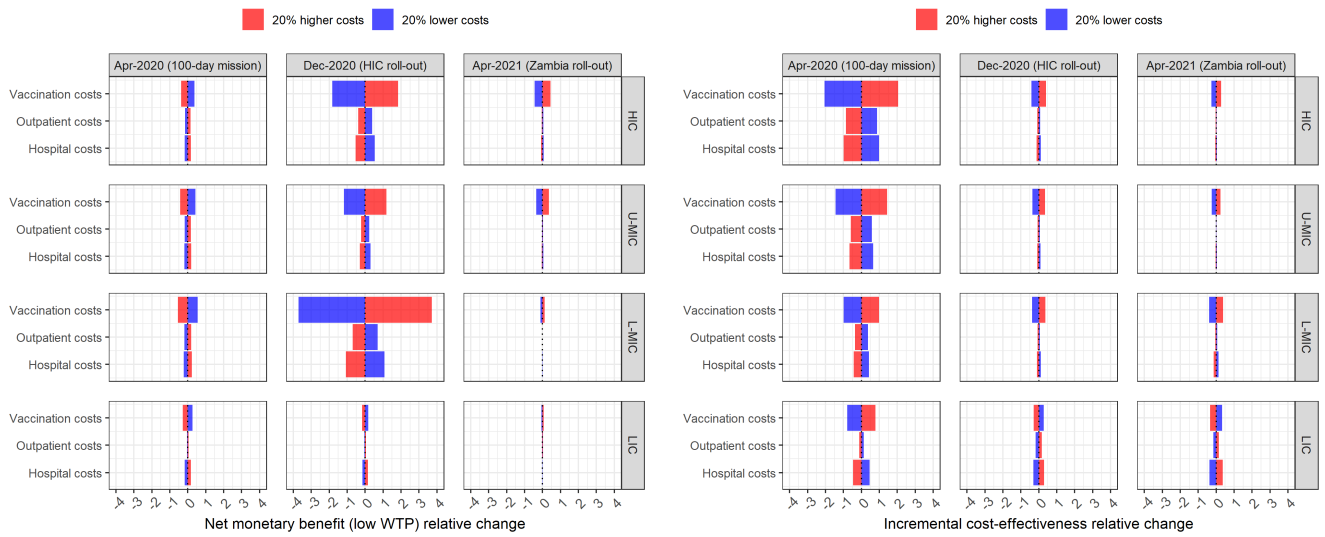


Figure B29: Relative change (% with respect to central parameters) in the net monetary benefit and incremental cost-effectiveness ratio of counterfactual vaccination interventions in one-way sensitivity analysis varying vaccination, outpatient and hospital costs, assuming the lowest plausible ocWTP threshold.

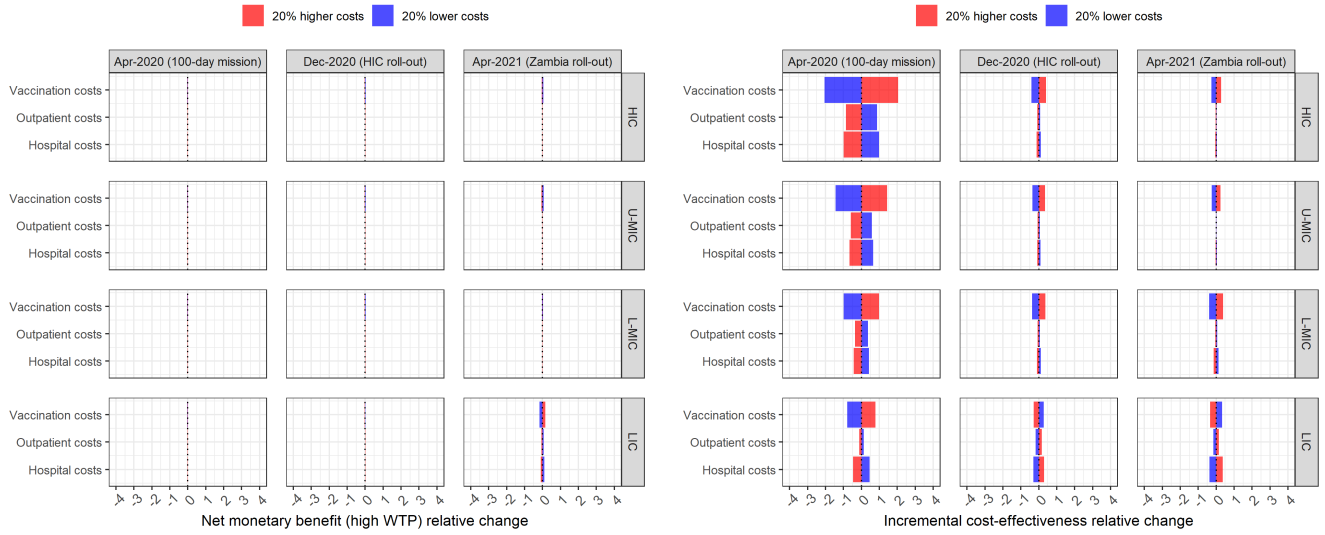


Figure B30: Relative change (% with respect to central parameters) in the net monetary benefit and incremental cost-effectiveness ratio of counterfactual vaccination interventions in one-way sensitivity analysis varying vaccination, outpatient and hospital costs, assuming the highest plausible ocWTP threshold.

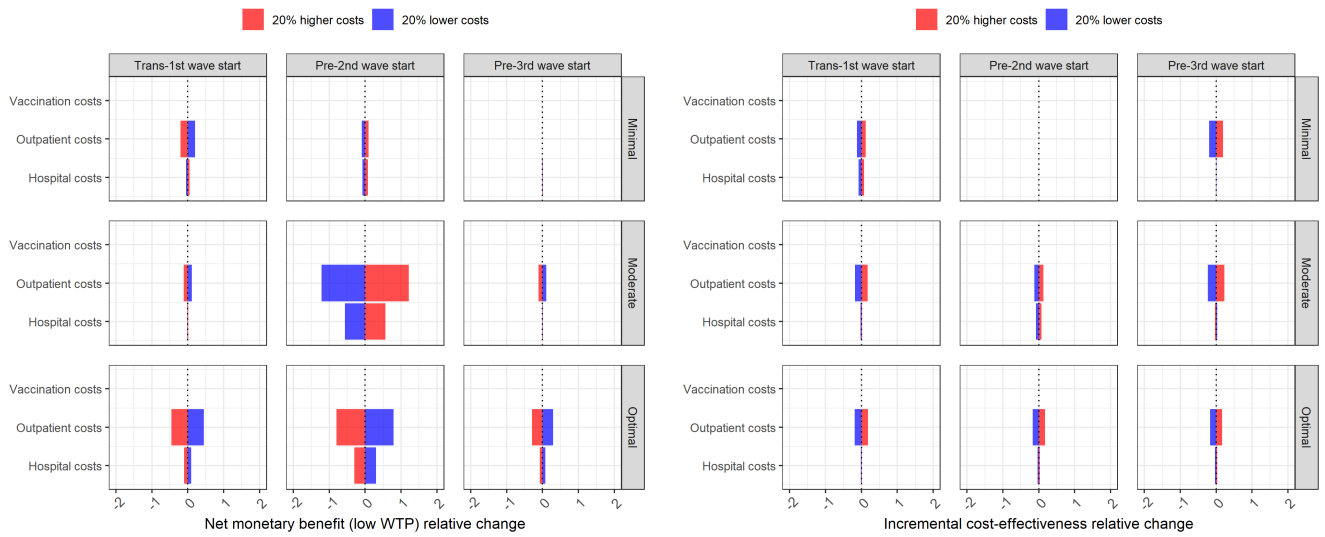


Figure B31: Relative change (% with respect to central parameters) in the net monetary benefit and incremental cost-effectiveness ratio of counterfactual healthcare improvement interventions in one-way sensitivity analysis varying vaccination, outpatient and hospital costs, assuming the lowest plausible ocWTP threshold.

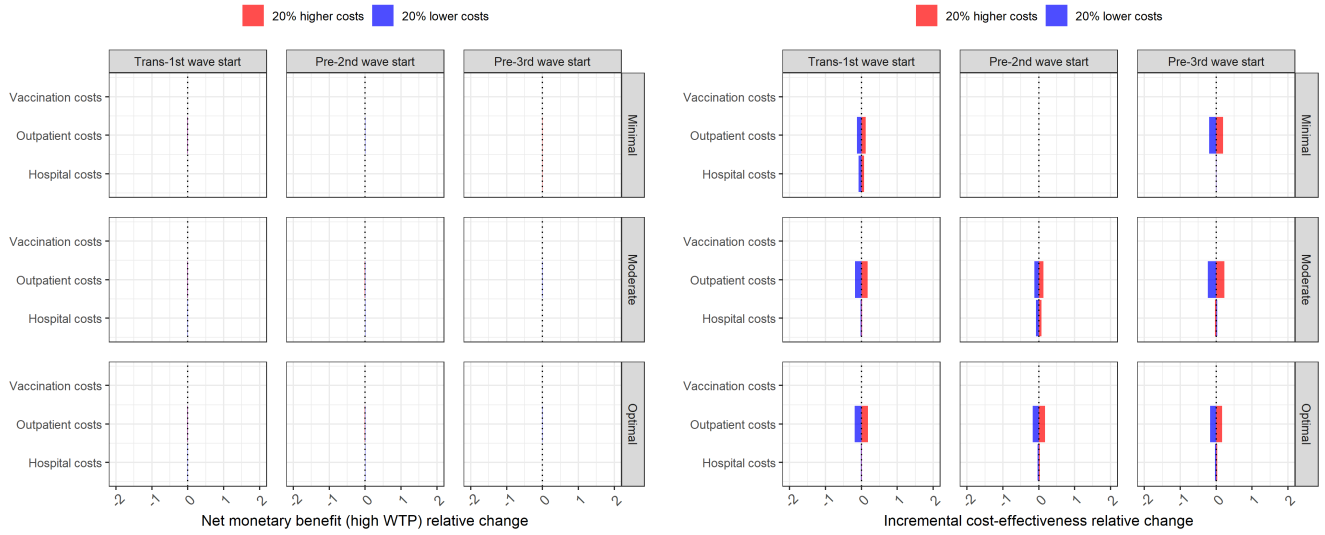


Figure B32: Relative change (% with respect to central parameters) in the net monetary benefit and incremental cost-effectiveness ratio of counterfactual healthcare improvement interventions in one-way sensitivity analysis varying vaccination, outpatient and hospital costs, assuming the highest plausible ocWTP threshold.

Table B3: DALYs and DALYs averted across counterfactual scenarios of vaccination roll-out and improved healthcare in Kabwe, Zambia between March 2020 and September 2021, as described in Chapter 5. Values shown are median and 95% credible interval across 200 stochastic trajectories simulated per scenario.

Start	Rate	Accrued DALYs	DALYS averted
Apr-2020 (100-day mission)	HIC	833 (300 to 1,642)	21,729 (18,725 to 24,623)
Apr-2020 (100-day mission)	U-MIC	1,529 (626 to 2,686)	21,019 (17,753 to 23,699)
Apr-2020 (100-day mission)	L-MIC	2,351 (1,228 to 4,122)	20,221 (17,429 to 23,066)
Apr-2020 (100-day mission)	LIC	13,914 (8,143 to 19,356)	8,556 (3,230 to 14,380)
Dec-2020 (HIC roll-out)	HIC	8,173 (6,597 to 10,100)	14,256 (12,064 to 17,058)
Dec-2020 (HIC roll-out)	U-MIC	9,556 (7,807 to 11,784)	12,851 (10,300 to 15,926)
Dec-2020 (HIC roll-out)	L-MIC	12,892 (10,306 to 15,052)	9,750 (7,662 to 12,384)
Dec-2020 (HIC roll-out)	LIC	19,290 (15,330 to 22,802)	3,391 (1,006 to 6,046)
Apr-2021 (Zambia roll-out)	HIC	15,490 (13,219 to 18,109)	6,981 (5,113 to 8,947)
Apr-2021 (Zambia roll-out)	U-MIC	18,325 (15,049 to 21,059)	4,308 (2,341 to 6,127)
Apr-2021 (Zambia roll-out)	L-MIC	19,788 (16,832 to 23,277)	2,733 (893 to 4,598)
Apr-2021 (Zambia roll-out)	LIC	21,661 (18,506 to 24,988)	898 (-778 to 2,839)
Trans-1st wave start	Minimal	18,904 (12,429 to 22,516)	3,771 (490 to 8,387)
Trans-1st wave start	Moderate	15,521 (10,769 to 19,072)	7,046 (4,194 to 11,068)
Trans-1st wave start	Optimal	12,380 (8,410 to 14,719)	10,275 (7,547 to 13,767)
Pre-2nd wave start	Minimal	19,013 (14,539 to 22,940)	3,530 (73 to 6,853)
Pre-2nd wave start	Moderate	16,072 (12,020 to 19,693)	6,405 (3,465 to 9,921)
Pre-2nd wave start	Optimal	13,584 (10,492 to 16,468)	8,961 (6,179 to 11,379)
Pre-3rd wave start	Minimal	20,412 (17,084 to 23,520)	2,193 (412 to 3,833)
Pre-3rd wave start	Moderate	18,675 (15,841 to 21,011)	3,899 (2,429 to 5,791)
Pre-3rd wave start	Optimal	16,996 (14,591 to 19,715)	5,492 (4,288 to 7,144)

Table B4: Incremental cost-effectiveness ratios (cost in 2020-USD per DALY averted) of counterfactual scenarios of vaccination roll-out and improved healthcare in Kabwe, Zambia between March 2020 and September 2021, as described in Chapter 5. Values shown are median and 95% credible interval across 200 stochastic trajectories simulated per scenario.

Start	Rate	ICER
Apr-2020 (100-day mission)	HIC	10 (4 to 21)
Apr-2020 (100-day mission)	U-MIC	14 (6 to 28)
Apr-2020 (100-day mission)	L-MIC	20 (12 to 33)
Apr-2020 (100-day mission)	LIC	13 (-18 to 99)
Dec-2020 (HIC roll-out)	HIC	66 (52 to 89)
Dec-2020 (HIC roll-out)	U-MIC	82 (64 to 108)
Dec-2020 (HIC roll-out)	L-MIC	70 (47 to 96)
Dec-2020 (HIC roll-out)	LIC	-10 (-36 to 61)
Apr-2021 (Zambia roll-out)	HIC	159 (111 to 231)
Apr-2021 (Zambia roll-out)	U-MIC	211 (138 to 409)
Apr-2021 (Zambia roll-out)	L-MIC	61 (21 to 263)
Apr-2021 (Zambia roll-out)	LIC	-16 (-190 to 250)
Trans-1st wave start	Minimal	28 (-32 to 353)
Trans-1st wave start	Moderate	41 (-8 to 111)
Trans-1st wave start	Optimal	42 (5 to 80)
Pre-2nd wave start	Minimal	25 (-28 to 282)
Pre-2nd wave start	Moderate	36 (-11 to 115)
Pre-2nd wave start	Optimal	43 (12 to 84)
Pre-3rd wave start	Minimal	17 (-23 to 205)
Pre-3rd wave start	Moderate	24 (-4 to 70)
Pre-3rd wave start	Optimal	29 (6 to 55)

Table B5: Incremental costs factual scenario and total costs (in 2020-USD) accrued in counterfactual scenarios of vaccination roll-out and improved healthcare in Kabwe, Zambia between March 2020 and September 2021, as described in Chapter 5. Values shown are median and 95% credible interval across 200 stochastic trajectories simulated per scenario.

Start	Rate	Incremental cost	Total cost
Apr-2020 (100-day mission)	HIC	210,549 (83,193 to 404,993)	2,001,460 (1,943,759 to 2,082,642)
Apr-2020 (100-day mission)	U-MIC	286,515 (132,591 to 496,870)	2,079,473 (1,974,625 to 2,177,710)
Apr-2020 (100-day mission)	L-MIC	411,744 (269,356 to 589,109)	2,201,691 (2,078,719 to 2,319,373)
Apr-2020 (100-day mission)	LIC	109,020 (-245,795 to 390,301)	1,893,699 (1,552,420 to 2,166,626)
Dec-2020 (HIC roll-out)	HIC	953,879 (829,889 to 1,099,680)	2,736,051 (2,666,020 to 2,847,531)
Dec-2020 (HIC roll-out)	U-MIC	1,070,739 (938,855 to 1,198,582)	2,845,872 (2,766,038 to 2,959,849)
Dec-2020 (HIC roll-out)	L-MIC	674,277 (562,660 to 786,344)	2,460,232 (2,338,832 to 2,576,281)
Dec-2020 (HIC roll-out)	LIC	-29,855 (-154,815 to 90,056)	1,760,940 (1,557,972 to 1,896,673)
Apr-2021 (Zambia roll-out)	HIC	1,106,338 (996,102 to 1,197,143)	2,893,689 (2,762,347 to 2,980,940)
Apr-2021 (Zambia roll-out)	U-MIC	906,436 (807,878 to 984,014)	2,690,617 (2,543,182 to 2,809,321)
Apr-2021 (Zambia roll-out)	L-MIC	171,900 (91,842 to 262,833)	1,964,971 (1,783,870 to 2,065,093)
Apr-2021 (Zambia roll-out)	LIC	-12,784 (-101,324 to 59,647)	1,770,022 (1,577,219 to 1,908,270)
Trans-1st wave start	Minimal	126,664 (-167,503 to 325,338)	1,916,107 (1,486,183 to 2,109,367)
Trans-1st wave start	Moderate	295,640 (-90,753 to 480,146)	2,078,622 (1,618,943 to 2,280,108)
Trans-1st wave start	Optimal	438,480 (59,593 to 640,832)	2,231,877 (1,766,977 to 2,454,977)
Pre-2nd wave start	Minimal	103,783 (-151,200 to 271,292)	1,891,068 (1,578,654 to 2,069,985)
Pre-2nd wave start	Moderate	236,564 (-82,456 to 423,141)	2,033,932 (1,658,268 to 2,211,283)
Pre-2nd wave start	Optimal	386,477 (120,058 to 583,106)	2,183,358 (1,853,620 to 2,392,674)
Pre-3rd wave start	Minimal	38,742 (-54,912 to 114,783)	1,827,172 (1,595,985 to 1,952,022)
Pre-3rd wave start	Moderate	90,614 (-19,051 to 189,114)	1,881,486 (1,653,110 to 2,035,895)
Pre-3rd wave start	Optimal	159,275 (37,034 to 267,776)	1,944,537 (1,691,316 to 2,118,420)

Table B6: Net monetary benefit (in 2020-USD) of counterfactual scenarios of vaccination roll-out and improved healthcare in Kabwe, Zambia between March 2020 and September 2021, as described in Chapter 5. Values shown are median and 95% credible interval across 200 stochastic trajectories simulated per scenario.

Start	Rate	max_nmb	min_nmb
Apr-2020 (100-day mission)	HIC	24,755,471 (21,107,854 to 28,159,994)	1,046,825 (700,270 to 1,345,663)
Apr-2020 (100-day mission)	U-MIC	23,833,539 (19,866,331 to 26,908,694)	935,542 (554,688 to 1,174,599)
Apr-2020 (100-day mission)	L-MIC	22,786,132 (19,514,165 to 26,160,492)	755,480 (438,051 to 1,045,927)
Apr-2020 (100-day mission)	LIC	9,736,557 (3,419,384 to 16,733,488)	381,703 (-177,935 to 1,083,028)
Dec-2020 (HIC roll-out)	HIC	15,457,488 (12,848,923 to 18,603,267)	-123,686 (-386,270 to 88,859)
Dec-2020 (HIC roll-out)	U-MIC	13,702,934 (10,777,907 to 17,218,594)	-320,911 (-553,997 to -94,830)
Dec-2020 (HIC roll-out)	L-MIC	10,501,223 (8,034,799 to 13,539,124)	-121,116 (-306,479 to 129,062)
Dec-2020 (HIC roll-out)	LIC	3,900,088 (1,019,237 to 6,966,092)	219,456 (-11,400 to 471,746)
Apr-2021 (Zambia roll-out)	HIC	6,883,461 (4,694,751 to 9,318,862)	-705,489 (-879,215 to -490,881)
Apr-2021 (Zambia roll-out)	U-MIC	4,027,523 (1,726,138 to 6,156,612)	-656,543 (-823,742 to -485,631)
Apr-2021 (Zambia roll-out)	L-MIC	2,947,406 (770,385 to 5,150,923)	-9,461 (-185,495 to 161,582)
Apr-2021 (Zambia roll-out)	LIC	1,061,523 (-954,398 to 3,359,479)	67,472 (-98,421 to 250,304)
Trans-1st wave start	Minimal	4,215,617 (310,397 to 9,808,182)	107,753 (-272,734 to 632,248)
Trans-1st wave start	Moderate	7,757,961 (4,346,967 to 12,619,420)	113,722 (-224,553 to 719,753)
Trans-1st wave start	Optimal	11,344,466 (8,046,270 to 15,410,559)	157,455 (-171,710 to 683,024)
Pre-2nd wave start	Minimal	4,014,238 (-213,043 to 7,928,514)	104,056 (-223,899 to 515,672)
Pre-2nd wave start	Moderate	7,114,805 (3,555,591 to 11,363,955)	140,139 (-217,143 to 632,611)
Pre-2nd wave start	Optimal	9,884,661 (6,675,817 to 12,820,779)	134,075 (-180,652 to 526,035)
Pre-3rd wave start	Minimal	2,489,011 (389,923 to 4,411,301)	92,542 (-60,308 to 265,397)
Pre-3rd wave start	Moderate	4,388,245 (2,606,854 to 6,570,738)	132,641 (-30,052 to 324,329)
Pre-3rd wave start	Optimal	6,150,499 (4,755,828 to 8,089,343)	158,728 (13,699 to 346,769)

B.6 Additional results: chapter 6

Figures B33 and B34 present the distribution of estimated DALYs averted and total costs of all 180 simulated intervention across 250 stochastic simulations of each. Figure B35 is a faceted version of Figure 6.2 in Chapter 6, to clearly demonstrate which scenarios were either not, weakly or strongly dominated.

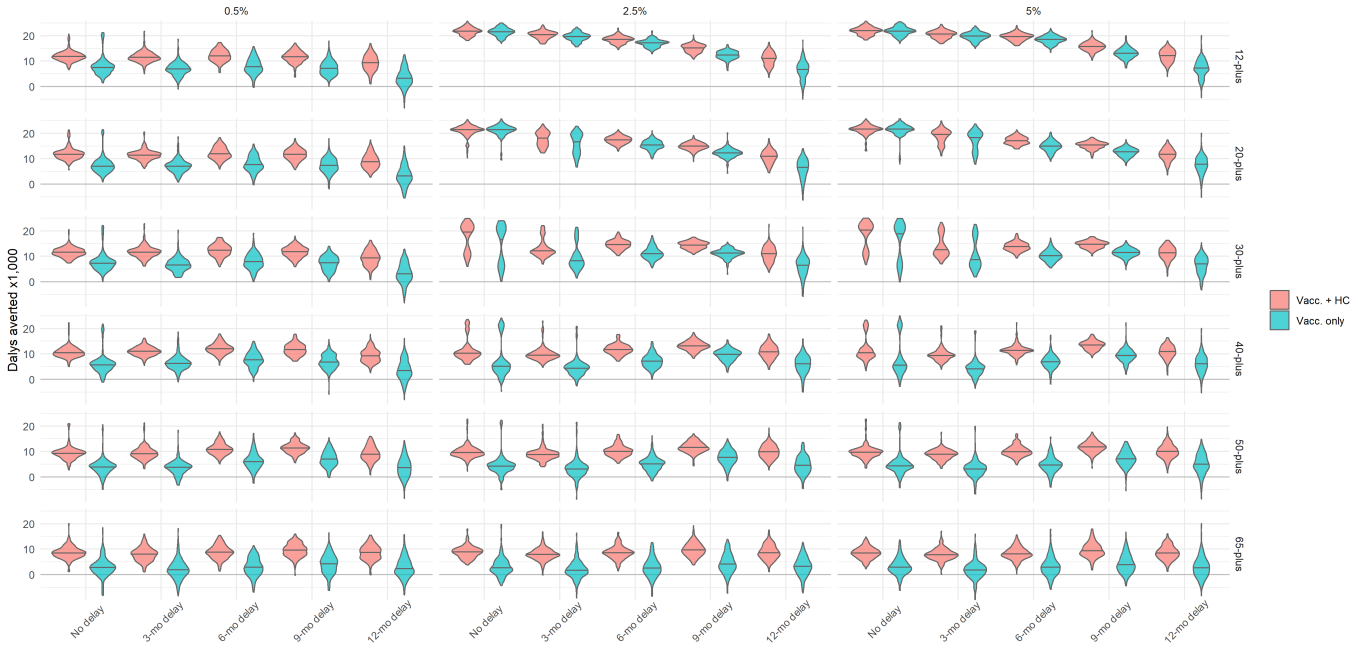


Figure B33: DALYs averted across scenarios in extended dominance analysis with the implementation of COVID-19 vaccination only vs. vaccination and healthcare improvement in May 2020 in Kabwe, Zambia as per CEPI's 100-day mission (*No delay)*, or at quarterly delays up to a year thereafter.

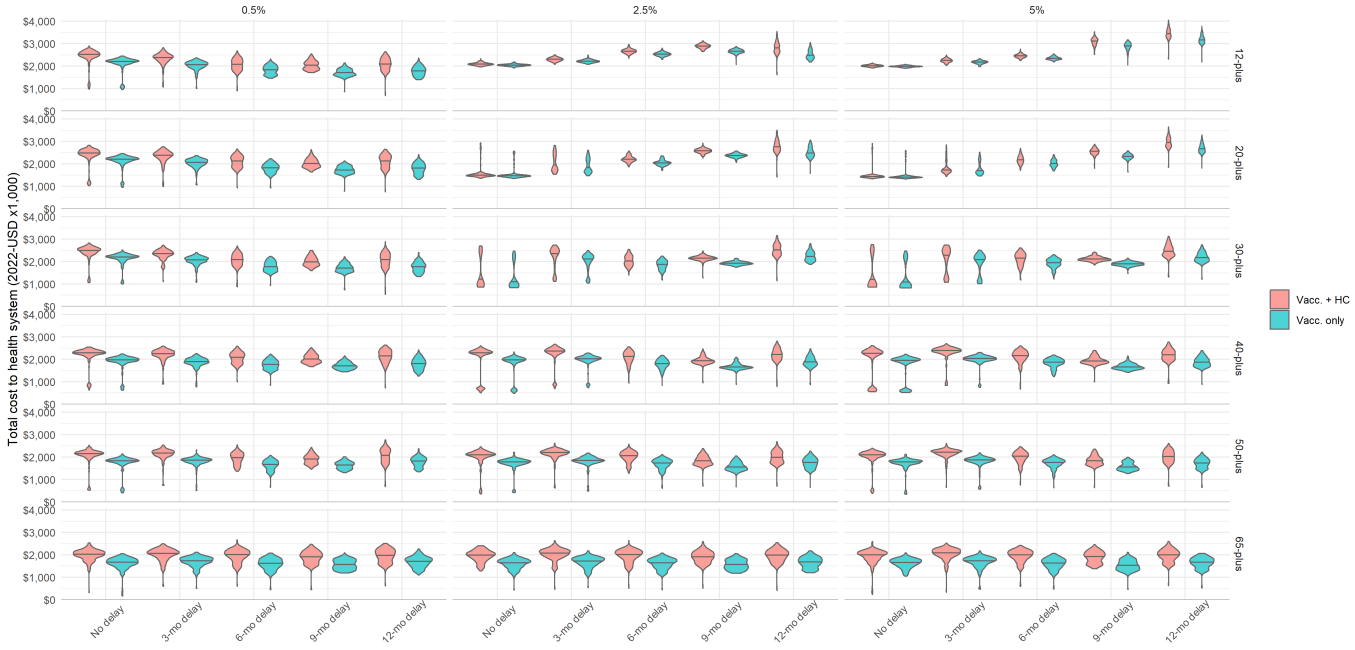


Figure B34: Total costs across scenarios in extended dominance analysis with the implementation of COVID-19 vaccination only vs. vaccination and healthcare improvement in May 2020 in Kabwe, Zambia as per CEPI's 100-day mission (*No delay)*, or at quarterly delays up to a year thereafter.

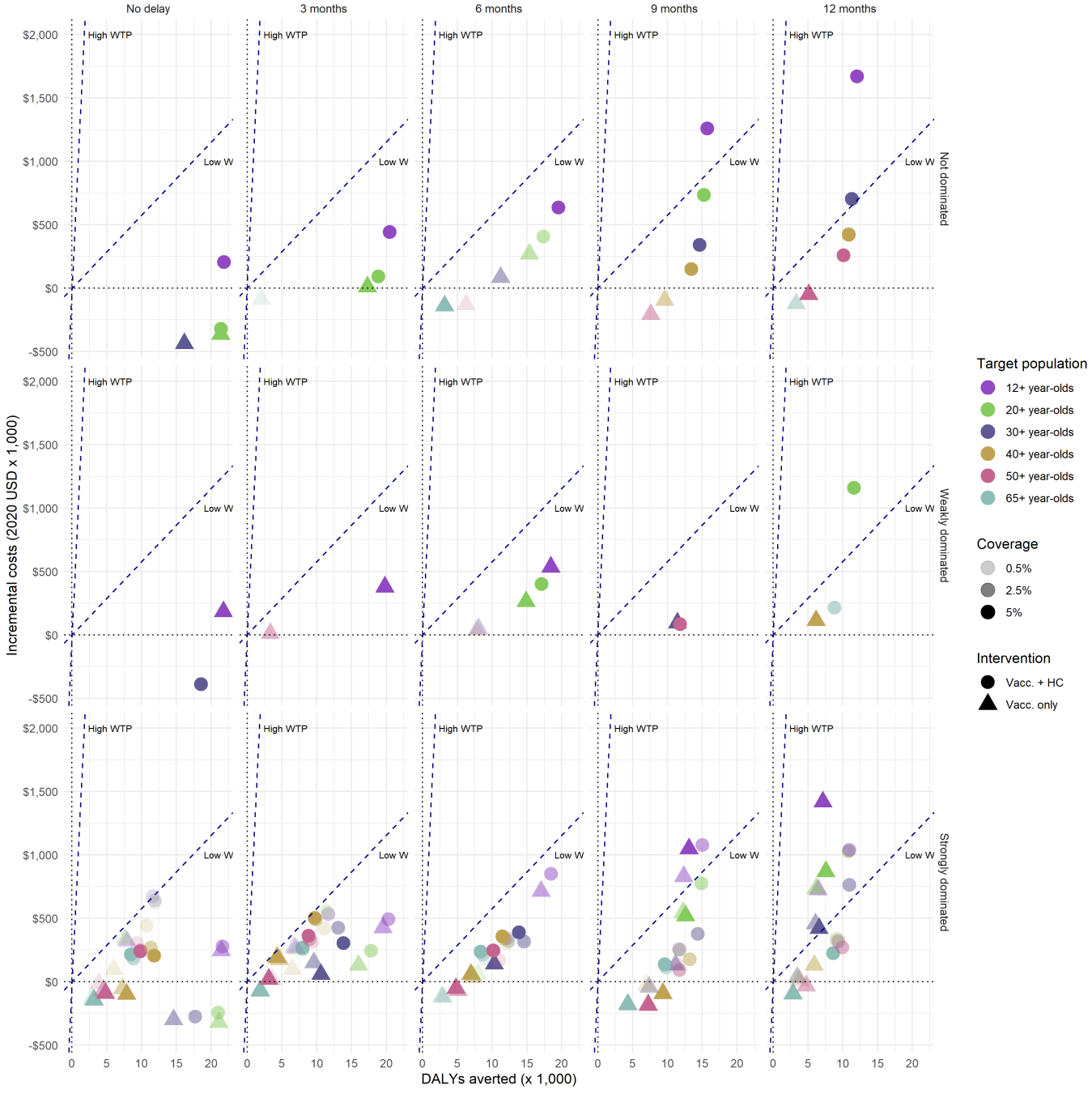


Figure B35: Cost-effectiveness plane of scenarios in extended dominance analysis of the implementation of COVID-19 vaccination only vs. vaccination and healthcare improvement in May 2020 in Kabwe, Zambia, as per CEPI's 100-day mission (*No delay*) or at quarterly delays up to a year thereafter. Diagonal dotted line show the low and high opportunity cost-based willingness to pay threshold, at $57.69 and $1,146.64 2020-USD per DALY-averted, respectively, or between 6% and 120% the GDPpc of Zambia per DALY averted, respectively, which were calculated as per [199, 200].

Table B7: League table of not dominated and weakly dominated interventions, ranged by total cost in extended dominance analysis and by delay window. *outcomes compared to factual (model-fitted) outcomes. **Outcomes compared to next more costly, not dominated intervention.

Delay	Age eligible	Intervention	Coverage	Status	Total cost	DALYs averted (fact.*)	ICER (fact.)*	ICER ext. dom.**
No delay	30+ year-olds	Vacc. only	5%	Not dominated	1341457	16167.275	-24.957757	NA
No delay	20+ year-olds	Vacc. only	5%	Not dominated	1426333	21382.356	-15.625785	16.275062
No delay	20+ year-olds	Vacc. + HC	5%	Not dominated	1461970	21441.124	-14.025813	606.405971
No delay	12+ year-olds	Vacc. + HC	5%	Not dominated	1990672	21880.655	9.575639	1202.876950
No delay	30+ year-olds	Vacc. + HC	5%	Weakly dominated	1386482	18552.289	-7.699999	NA
No delay	12+ year-olds	Vacc. only	5%	Weakly dominated	1971948	21760.730	8.662419	NA
3 months	65+ year-olds	Vacc. only	0.5%	Not dominated	1701209	2044.850	-1.163174	NA
3 months	20+ year-olds	Vacc. only	5%	Not dominated	1800621	17254.758	5.569700	6.536055
3 months	20+ year-olds	Vacc. + HC	5%	Not dominated	1876115	18844.401	8.040681	47.490833
3 months	12+ year-olds	Vacc. + HC	5%	Not dominated	2227203	20466.998	21.919076	216.374027
3 months	50+ year-olds	Vacc. only	2.5%	Weakly dominated	1800149	3294.488	103.576060	NA
3 months	12+ year-olds	Vacc. only	5%	Weakly dominated	2166732	19764.188	19.550645	NA
6 months	65+ year-olds	Vacc. only	5%	Not dominated	1649641	3170.066	-131.664876	NA
6 months	50+ year-olds	Vacc. only	0.5%	Not dominated	1659720	6258.269	21.211822	3.263702
6 months	30+ year-olds	Vacc. only	2.5%	Not dominated	1863258	11235.156	10.582246	40.896582
6 months	20+ year-olds	Vacc. only	2.5%	Not dominated	2059599	15360.204	18.427229	47.597282
6 months	20+ year-olds	Vacc. + HC	2.5%	Not dominated	2190092	17393.022	23.864620	64.193345
6 months	12+ year-olds	Vacc. + HC	5%	Not dominated	2421567	19519.714	32.955911	108.842795
6 months	40+ year-olds	Vacc. only	0.5%	Weakly dominated	1824592	7717.147	14.837673	NA
6 months	30+ year-olds	Vacc. only	0.5%	Weakly dominated	1825354	7983.386	-61.423464	NA
6 months	12+ year-olds	Vacc. only	0.5%	Weakly dominated	1836732	8070.147	19.413078	NA
6 months	20+ year-olds	Vacc. only	5%	Weakly dominated	2053752	14875.307	19.224855	NA
6 months	20+ year-olds	Vacc. + HC	5%	Weakly dominated	2185479	17125.719	24.223245	NA

Table B7: League table of not dominated and weakly dominated interventions, ranged by total cost in extended dominance analysis and by delay window. *outcomes compared to factual (model-fitted) outcomes. **Outcomes compared to next more costly, not dominated intervention.

Delay	Age eligible	Intervention	Coverage	Status	Total cost	DALYs averted (fact.*)	ICER (fact.)*	ICER ext. dom.**
6 months	12+ year-olds	Vacc. only	5%	Weakly dominated	2324223	18420.760	29.547676	NA
9 months	50+ year-olds	Vacc. only	2.5%	Not dominated	1580824	7602.709	-27.098220	NA
9 months	40+ year-olds	Vacc. only	2.5%	Not dominated	1695331	9591.887	-7.112707	57.565321
9 months	40+ year-olds	Vacc. + HC	5%	Not dominated	1932764	13421.459	12.469024	61.999794
9 months	30+ year-olds	Vacc. + HC	5%	Not dominated	2115925	14631.067	23.799382	151.421368
9 months	20+ year-olds	Vacc. + HC	5%	Not dominated	2519026	15257.869	49.024294	643.107731
9 months	12+ year-olds	Vacc. + HC	5%	Not dominated	3044239	15727.421	81.690784	1118.541175
9 months	30+ year-olds	Vacc. only	5%	Weakly dominated	1872610	11437.205	9.345335	NA
9 months	50+ year-olds	Vacc. + HC	5%	Weakly dominated	1873403	11800.981	10.567996	NA
12 months	65+ year-olds	Vacc. only	2.5%	Not dominated	1665378	3300.846	-46.734538	NA
12 months	50+ year-olds	Vacc. only	5%	Not dominated	1737816	5106.138	-29.691441	40.125201
12 months	50+ year-olds	Vacc. + HC	5%	Not dominated	2046675	10101.496	36.424201	61.829312
12 months	40+ year-olds	Vacc. + HC	5%	Not dominated	2202724	10847.017	49.412616	209.315199
12 months	40+ year-olds	Vacc. + HC	2.5%	Not dominated	2210280	10864.469	47.762477	432.958287
12 months	30+ year-olds	Vacc. + HC	5%	Not dominated	2478127	11326.069	72.695847	580.257949
12 months	12+ year-olds	Vacc. + HC	5%	Not dominated	3454357	12068.829	148.671195	1314.326027
12 months	40+ year-olds	Vacc. only	5%	Weakly dominated	1903608	6160.282	47.164957	NA
12 months	65+ year-olds	Vacc. + HC	0.5%	Weakly dominated	1996037	8704.030	-6.940699	NA
12 months	65+ year-olds	Vacc. + HC	2.5%	Weakly dominated	1999401	8826.028	40.409246	NA
12 months	20+ year-olds	Vacc. + HC	5%	Weakly dominated	2945401	11608.313	110.511942	NA

Table B8: Median (95%CrI) net-monetary benefit of the implementation COVID-19 vaccination only vs. vaccination and healthcare improvement interventions in Kabwe, Zambia, assuming a low opportunity cost-based willingness to pay threshold, at $57.69 2022-USD or 6% of the per-capita gross domestic product of Zambia per DALY averted.

Age eligible	Intervention	Vacc. rate	No delay	3-mo delay	6-mo delay	9-mo delay	12-mo delay
12-plus	Vacc. + HC	0.5%	6 (-388 - 1459)	88 (-347 - 955)	313 (-258 - 933)	415 (-145 - 907)	208 (-540 - 922)
12-plus	Vacc. + HC	2.5%	984 (631 - 1240)	701 (285 - 935)	228 (-107 - 505)	-190 (-543 - 49)	-367 (-1063 - 106)
12-plus	Vacc. + HC	5%	1074 (743 - 1362)	746 (375 - 1081)	518 (109 - 758)	-366 (-799 - 91)	-904 (-1591 - -510)
12-plus	Vacc. only	0.5%	50 (-343 - 1930)	131 (-359 - 700)	368 (-265 - 1094)	433 (-8 - 1023)	135 (-510 - 902)
12-plus	Vacc. only	2.5%	1019 (665 - 1268)	714 (346 - 984)	281 (-18 - 545)	-97 (-463 - 119)	-295 (-1209 - 308)
12-plus	Vacc. only	5%	1091 (725 - 1355)	793 (390 - 1050)	546 (197 - 797)	-311 (-714 - 258)	-968 (-1705 - -485)
20-plus	Vacc. + HC	0.5%	-5 (-382 - 1759)	63 (-349 - 806)	309 (-243 - 1021)	435 (-178 - 911)	128 (-420 - 884)
20-plus	Vacc. + HC	2.5%	1549 (126 - 1830)	922 (-106 - 1509)	624 (233 - 927)	99 (-237 - 350)	-356 (-1109 - 159)
20-plus	Vacc. + HC	5%	1614 (91 - 1904)	1208 (-70 - 1585)	605 (112 - 1033)	172 (-285 - 410)	-440 (-1174 - 16)
20-plus	Vacc. only	0.5%	1 (-344 - 1868)	128 (-318 - 908)	349 (-283 - 1048)	498 (-128 - 966)	136 (-606 - 1015)
20-plus	Vacc. only	2.5%	1602 (208 - 1870)	1007 (-278 - 1476)	648 (156 - 936)	175 (-170 - 416)	-332 (-1306 - 229)
20-plus	Vacc. only	5%	1660 (47 - 1922)	1211 (-122 - 1619)	596 (93 - 1085)	221 (-110 - 478)	-356 (-1236 - 85)
30-plus	Vacc. + HC	0.5%	-47 (-395 - 470)	114 (-381 - 949)	395 (-277 - 1034)	452 (-160 - 907)	186 (-439 - 944)
30-plus	Vacc. + HC	2.5%	1897 (-321 - 2299)	89 (-320 - 1840)	514 (21 - 1095)	476 (88 - 739)	-118 (-819 - 458)
30-plus	Vacc. + HC	5%	2054 (-300 - 2411)	108 (-316 - 1976)	386 (-27 - 1050)	524 (129 - 771)	-17 (-781 - 454)
30-plus	Vacc. only	0.5%	15 (-391 - 1849)	76 (-345 - 740)	380 (-221 - 1091)	502 (-151 - 955)	142 (-624 - 957)
30-plus	Vacc. only	2.5%	1855 (-354 - 2266)	172 (-301 - 1832)	528 (83 - 1137)	539 (129 - 781)	-50 (-1005 - 578)
30-plus	Vacc. only	5%	2004 (-335 - 2331)	196 (-332 - 1962)	436 (22 - 1051)	585 (180 - 793)	49 (-855 - 548)
40-plus	Vacc. + HC	0.5%	127 (-175 - 809)	210 (-190 - 738)	361 (-157 - 1031)	446 (-100 - 982)	198 (-326 - 909)
40-plus	Vacc. + HC	2.5%	133 (-308 - 2561)	-2 (-325 - 1874)	286 (-132 - 1125)	605 (99 - 968)	226 (-476 - 798)
40-plus	Vacc. + HC	5%	147 (-353 - 2504)	-13 (-329 - 1686)	274 (-165 - 1090)	647 (69 - 988)	224 (-451 - 778)
40-plus	Vacc. only	0.5%	176 (-350 - 2201)	231 (-257 - 1197)	383 (-189 - 1065)	397 (-79 - 967)	138 (-546 - 991)
40-plus	Vacc. only	2.5%	159 (-391 - 2547)	48 (-369 - 1791)	361 (-134 - 1077)	704 (65 - 993)	250 (-634 - 850)

Table B8: Median (95%CrI) net-monetary benefit of the implementation COVID-19 vaccination only vs. vaccination and healthcare improvement interventions in Kabwe, Zambia, assuming a low opportunity cost-based willingness to pay threshold, at $57.69 2022-USD or 6% of the per-capita gross domestic product of Zambia per DALY averted.

Age eligible	Intervention	Vacc. rate	No delay	3-mo delay	6-mo delay	9-mo delay	12-mo delay
40-plus	Vacc. only	5%	197 (-384 - 2592)	17 (-464 - 1672)	316 (-241 - 1069)	652 (95 - 1071)	262 (-547 - 885)
50-plus	Vacc. + HC	0.5%	203 (-327 - 898)	178 (-271 - 807)	412 (-41 - 1151)	525 (15 - 1035)	153 (-457 - 927)
50-plus	Vacc. + HC	2.5%	272 (-106 - 834)	122 (-309 - 1174)	320 (-115 - 1092)	596 (58 - 1047)	283 (-312 - 929)
50-plus	Vacc. + HC	5%	265 (-149 - 2269)	154 (-309 - 615)	306 (-122 - 1142)	590 (6 - 1104)	289 (-352 - 1002)
50-plus	Vacc. only	0.5%	237 (-313 - 704)	209 (-362 - 778)	449 (-181 - 1302)	591 (-34 - 1152)	215 (-439 - 1088)
50-plus	Vacc. only	2.5%	285 (-245 - 2585)	145 (-427 - 1000)	363 (-168 - 1071)	681 (6 - 1162)	292 (-397 - 1092)
50-plus	Vacc. only	5%	299 (-247 - 2474)	143 (-488 - 830)	311 (-338 - 1117)	598 (14 - 1171)	344 (-479 - 1026)
65-plus	Vacc. + HC	0.5%	250 (-297 - 949)	192 (-398 - 997)	271 (-333 - 1094)	400 (-210 - 1058)	285 (-351 - 1055)
65-plus	Vacc. + HC	2.5%	314 (-160 - 1098)	152 (-361 - 906)	238 (-320 - 1140)	422 (-141 - 1157)	255 (-334 - 1041)
65-plus	Vacc. + HC	5%	264 (-214 - 895)	159 (-321 - 949)	202 (-316 - 970)	339 (-255 - 1199)	229 (-381 - 995)
65-plus	Vacc. only	0.5%	276 (-475 - 1078)	154 (-540 - 1164)	283 (-437 - 1006)	425 (-402 - 1127)	176 (-415 - 1056)
65-plus	Vacc. only	2.5%	251 (-269 - 1105)	146 (-405 - 1104)	220 (-362 - 1214)	378 (-260 - 1226)	313 (-501 - 1080)
65-plus	Vacc. only	5%	297 (-295 - 1237)	157 (-533 - 988)	282 (-332 - 1139)	361 (-218 - 1174)	216 (-518 - 1106)

Table B9: Median (95%CrI) net-monetary benefit of the implementation COVID-19 vaccination only vs. vaccination and healthcare improvement interventions in Kabwe, Zambia, assuming a high opportunity cost-based willingness to pay threshold, at $1,146.64 2020-USD or 120% of the per-capita gross domestic product of Zambia per DALY averted.

Age eligible	Intervention	Vacc. rate	No delay	3-mo delay	6-mo delay	9-mo delay	12-mo delay
12-plus	Vacc. + HC	0.5%	12799 (8348 - 20858)	12529 (8442 - 17681)	13438 (8769 - 19117)	13296 (7354 - 18123)	10435 (2478 - 17496)
12-plus	Vacc. + HC	2.5%	24704 (21170 - 28018)	23057 (18994 - 26344)	20403 (17081 - 23417)	16377 (12446 - 19214)	11950 (5573 - 17077)

Table B9: Median (95%CrI) net-monetary benefit of the implementation COVID-19 vaccination only vs. vaccination and healthcare improvement interventions in Kabwe, Zambia, assuming a high opportunity cost-based willingness to pay threshold, at $1,146.64 2020-USD or 120% of the per-capita gross domestic product of Zambia per DALY averted.

Age eligible	Intervention	Vacc. rate	No delay	3-mo delay	6-mo delay	9-mo delay	12-mo delay
12-plus	Vacc. + HC	5%	25019 (21281 - 28251)	23237 (19501 - 26534)	21958 (18050 - 25002)	16737 (12561 - 20757)	12376 (6616 - 16972)
12-plus	Vacc. only	0.5%	8157 (2994 - 23801)	7685 (2116 - 13436)	8724 (2381 - 16764)	8193 (3586 - 15197)	3515 (-3786 - 12503)
12-plus	Vacc. only	2.5%	24500 (20601 - 27724)	22057 (18242 - 25365)	19059 (15834 - 21848)	13492 (8803 - 16586)	7023 (-2443 - 14247)
12-plus	Vacc. only	5%	24835 (20948 - 28327)	22452 (18469 - 25700)	20686 (17164 - 23962)	13915 (9091 - 19015)	6788 (-899 - 12889)
20-plus	Vacc. + HC	0.5%	12774 (8734 - 22776)	12448 (7825 - 17660)	13343 (8568 - 19601)	13178 (7758 - 18224)	9738 (4472 - 16760)
20-plus	Vacc. + HC	2.5%	24900 (16381 - 28228)	20496 (13873 - 25811)	19601 (16179 - 22897)	16370 (13057 - 19076)	11595 (5318 - 17039)
20-plus	Vacc. + HC	5%	25199 (17277 - 28559)	22650 (14072 - 26765)	19324 (15405 - 22634)	17050 (12665 - 19620)	12413 (5966 - 17767)
20-plus	Vacc. only	0.5%	7614 (3125 - 23719)	7952 (2129 - 15194)	8586 (1858 - 16419)	8498 (2504 - 14414)	3489 (-4426 - 13581)
20-plus	Vacc. only	2.5%	25024 (13119 - 28207)	19391 (8380 - 24672)	17394 (12214 - 21272)	13482 (9891 - 17129)	6915 (-4036 - 13683)
20-plus	Vacc. only	5%	25246 (12219 - 28412)	21724 (9988 - 25537)	16808 (11922 - 21641)	14116 (9982 - 17482)	8035 (-778 - 13823)
30-plus	Vacc. + HC	0.5%	12488 (8396 - 16999)	12914 (7725 - 18956)	13992 (7861 - 19504)	13338 (7829 - 17577)	10362 (4046 - 17349)
30-plus	Vacc. + HC	2.5%	23918 (8202 - 28782)	13142 (8883 - 25069)	16526 (12108 - 20891)	16295 (12956 - 19144)	11945 (5311 - 17652)
30-plus	Vacc. + HC	5%	24719 (9379 - 29015)	13475 (8852 - 25821)	15459 (11645 - 20259)	16630 (13013 - 19696)	12548 (5882 - 17487)
30-plus	Vacc. only	0.5%	7940 (2937 - 23323)	7248 (2117 - 13438)	8856 (2358 - 16840)	8811 (2175 - 14207)	3546 (-5940 - 12287)
30-plus	Vacc. only	2.5%	22695 (3047 - 27602)	8874 (3417 - 23853)	12488 (7633 - 19089)	12816 (8495 - 16107)	7236 (-4176 - 14195)
30-plus	Vacc. only	5%	23842 (3067 - 28524)	9111 (3434 - 25507)	11464 (7268 - 17416)	13135 (8538 - 16927)	7802 (-2113 - 13756)
40-plus	Vacc. + HC	0.5%	11607 (7681 - 17130)	12190 (8488 - 17416)	13517 (8985 - 19141)	13132 (8088 - 19149)	10224 (5279 - 16600)
40-plus	Vacc. + HC	2.5%	11281 (6482 - 26965)	10313 (6528 - 22824)	13071 (8613 - 19597)	14890 (10531 - 19451)	11955 (6165 - 18136)
40-plus	Vacc. + HC	5%	11491 (5775 - 26842)	10243 (5960 - 21017)	12647 (8234 - 18585)	15410 (11228 - 19457)	12143 (5581 - 17736)
40-plus	Vacc. only	0.5%	6501 (-206 - 23561)	7059 (1739 - 15992)	8606 (2313 - 16070)	7777 (2479 - 14599)	3754 (-4948 - 13185)
40-plus	Vacc. only	2.5%	5941 (-836 - 26981)	5005 (-814 - 20635)	8127 (2912 - 15464)	11506 (5385 - 15627)	6848 (-2900 - 13744)
40-plus	Vacc. only	5%	6290 (-732 - 27375)	4634 (-1448 - 19091)	7855 (1149 - 14575)	10949 (4914 - 16041)	6874 (-1329 - 14428)
50-plus	Vacc. + HC	0.5%	10359 (4930 - 17010)	10156 (5710 - 15398)	12262 (7733 - 18413)	12964 (8007 - 17628)	9956 (4045 - 17073)

Table B9: Median (95%CrI) net-monetary benefit of the implementation COVID-19 vaccination only vs. vaccination and healthcare improvement interventions in Kabwe, Zambia, assuming a high opportunity cost-based willingness to pay threshold, at $1,146.64 2020-USD or 120% of the per-capita gross domestic product of Zambia per DALY averted.

Age eligible	Intervention	Vacc. rate	No delay	3-mo delay	6-mo delay	9-mo delay	12-mo delay
50-plus	Vacc. + HC	2.5%	10759 (6352 - 16262)	9699 (4973 - 17639)	11396 (6789 - 18040)	13098 (8884 - 17833)	11088 (5795 - 17353)
50-plus	Vacc. + HC	5%	10926 (5732 - 22937)	9915 (5260 - 13690)	11182 (6620 - 16920)	13399 (8570 - 18622)	11140 (4966 - 17895)
50-plus	Vacc. only	0.5%	4702 (-2193 - 10535)	4366 (-2417 - 10830)	7113 (61 - 16067)	8184 (1242 - 15015)	4245 (-3379 - 13571)
50-plus	Vacc. only	2.5%	4928 (-1756 - 24984)	3594 (-3768 - 12123)	6125 (-275 - 13054)	9019 (1613 - 15128)	5318 (-2005 - 14334)
50-plus	Vacc. only	5%	5029 (-1514 - 23538)	3512 (-4475 - 10302)	5552 (-2076 - 13310)	8422 (1841 - 15445)	5723 (-3780 - 13989)
65-plus	Vacc. + HC	0.5%	9607 (4485 - 15508)	9051 (2725 - 15962)	9817 (3170 - 16421)	10936 (4823 - 16478)	9962 (3656 - 16441)
65-plus	Vacc. + HC	2.5%	10022 (5166 - 16679)	8723 (3224 - 15045)	9571 (3482 - 16356)	11029 (5496 - 17565)	9484 (4103 - 16626)
65-plus	Vacc. + HC	5%	9453 (4117 - 15123)	8568 (3597 - 15352)	9075 (4030 - 15470)	10328 (4434 - 18189)	9490 (3670 - 15804)
65-plus	Vacc. only	0.5%	3166 (-6198 - 12097)	2175 (-5925 - 12395)	3440 (-5363 - 11210)	5066 (-4696 - 13051)	2492 (-4723 - 12780)
65-plus	Vacc. only	2.5%	2910 (-3190 - 12588)	1949 (-4767 - 12223)	2941 (-4286 - 12826)	4578 (-2439 - 14229)	3900 (-5318 - 12549)
65-plus	Vacc. only	5%	3422 (-3471 - 13154)	2160 (-6536 - 11192)	3357 (-3877 - 12637)	4407 (-1956 - 13869)	3196 (-4639 - 12899)