Estimating the impact of interventions on viral epidemics using epidemiological and evolutionary models

Lily Geidelberg

Department of Infectious Disease Epidemiology
Imperial College London

A thesis submitted to Imperial College London for the degree of Doctor of Philosophy

January 2022

Declaration

I declare that the work contained in this thesis is my own. All secondary sources have been appropriately cited. The work was completed under the guidance and help from my supervisors Dr Erik Volz, Professor Marie-Claude Boily and Dr Kate Mitchell.

Chapters 2 and 3 are published together in a first author publication. Chapter 5 is published as a first author publication. Chapter 4 is published as half of a comparative analysis with a colleague; I am second author.

Elements of Chapter 6 are published as follows. Case study D features a phylodynamic analysis of England, split by lineage. This work was carried out by me within the Imperial College COVID-19 response team, and as such my analysis depended on work carried out by others. Specifically, Dr Erik Volz provided the sub-samples and maximum-likelihood phylogenetic trees. The tree dating and phylodynamic analyses were carried out by me. My analysis was featured in a publication in which several analyses agreed on the increased transmissibility of B.1.1.7; I am listed as a prominent author. Case study A features as a use-case for the *mlesky* method, and is submitted as a manuscript; I am second author.

All published and submitted manuscripts have had contributions from co-authors, most notably from my supervisors. All of the text, methods, figures and analyses are my own. All published work is under the Creative Commons Attribution (CC BY) licence.

Lily Geidelberg
January 2022

The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution-Non Commercial 4.0 International Licence (CC BY-NC). 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 and do not use it, or any derivative works, for a commercial purpose. 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

This thesis would not have been possible without the help and guidance from my supervisors, Professor Marie-Claude Boily, Dr Kate Mitchell and Dr Erik Volz. Thank you for teaching me so much, and for having so much patience throughout the years!

Thank you also to Professor Peter Vickerman and Professor Simon Gregson for acting as assessors and helping me re-shape the thesis following a rather unexpected twist in early 2020.

I have further benefited from academic help from many other members of the department, who have indeed also become friends. The little bits and pieces go a long way: Mathieu, Joe, Romain, Manon, Juan, Kris, Rob, Olivia, David, Igor, Fabricia, John, Rich, Joce, Alastair.

Thank you also for academic advice and friendship from outside the department: Katia, Fiona, Michel, Jasdeep, Charlotte, Tom.

The PhD process is of course far more than the research. Thank you therefore to all other staff who have supported me: Kate, Jo, Maria, Francesca, Renata, Ruth, Nousheen, Paul, Chris, Wes, Soraya.

I have made so many friends throughout my 6+ years in DIDE who have made the day-to-day so much fun. There are too many of you to name. I hope to see you again in the Monkey Puzzle.

To those special people who elected to endure countless hours of my self-indulgent comedy during lunch. I haven't forgotten you yet: Julia, Joel, Finlay, OJ, Izzy, Amy, Will, James.

To the other side of the continent: те сакам срце.

And finally to the most important of all, who has had to suffer me not only at work, but also at home through three lockdowns (and counting). It truly would not have been possible without you: Nora.

Abstract

Mathematical modelling is useful for understanding infectious disease epidemics, and for estimating the impact of interventions to control them. However, the appropriateness of different modelling approaches depends on the local context and availability of data. In this thesis I use two approaches, epidemiological and evolutionary modelling, to estimate the impact of interventions for controlling HIV and COVID-19.

After developing a dynamic transmission model fitted to local demographic, HIV prevalence and antiretroviral therapy (ART) coverage data, I estimated the impact of a 2-year real-world demonstration study of pre-exposure prophylaxis (PrEP) and ART treatment-as-prevention (TasP) among female sex workers (FSW) in Cotonou, Benin. I further estimated the impact of long-term scale-up scenarios of PrEP and TasP among FSW in Cotonou.

Next, based on data collected during the 2020 COVID-19 epidemic in Benin, I simulated scenarios of COVID-related disruptions to HIV services and sexual behaviour, using the same dynamic transmission model. I estimated the impact of these disruptions on HIV epidemiology in Cotonou.

Many non-pharmaceutical interventions (NPIs), including travel restrictions and lockdowns, were implemented globally to control COVID-19. Using model-based phylodynamic methods fitted to twenty SARS-CoV-2 genomes, I estimated the impact of NPIs in Weifang, China, one of the first COVID-19 epidemics outside of Wuhan.

Finally, I performed several non-parametric phylodynamic analyses on the COVID-19 epidemic in England, fitted to thousands of SARS-CoV-2 genomes. I evaluated the association of stringency of NPIs and mobility with viral transmission at different periods between spring 2020 and winter 2020-2021. I further investigated the relative growth rates between different viral lineages, including the B.1.1.7 ('alpha') variant.

The results from this thesis are useful principally to policymakers for designing effective local control measures for HIV and COVID-19. While modelling can be used to estimate the impact of interventions, model choice must be tailored to the local epidemiological context.

Contents

List of Figures **19**

List of Tables **35**

Nomenclature **37**

Overview **39**

1 Background **45**

1.1 Modelling of infectious disease dynamics 45

1.1.1 Uses of modelling in viral epidemic response 45

1.1.2 Different modelling frameworks 47

1.1.2.1 Dynamic transmission modelling 47

1.1.2.2 Phylodynamic inference 51

1.2 HIV 56

1.2.1 Natural history 56

1.2.2 Epidemiology 57

1.2.2.1 Global epidemiology 57

1.2.2.2 Key populations 58

1.2.3 HIV prevention interventions 59

1.2.3.1 Treatment as prevention (TasP) 60

1.2.3.2 Pre-exposure prophylaxis (PrEP) 61

1.3 SARS-CoV-2 63
1.3.1 Natural history 63
1.3.2 Epidemiology 64
1.3.2.1 Evolution 66
1.3.3 Interventions 67
1.4 Summary 68

2 Data summary and transmission model development for HIV in Cotonou 69
2.1 Overview 69
2.2 Introduction 70
2.2.1 The city of Cotonou and the Republic of Benin 70
2.2.2 HIV epidemiology in Cotonou 70
2.2.3 The response to HIV in Cotonou 72
2.2.4 The PrEP/TasP study 73
2.2.5 Previous studies modelling the HIV epidemic in Cotonou 75
2.2.6 Aims and justification 76
2.3 Data review 78
2.3.1 Main data sources 78
2.3.2 Summary of available data 79
2.3.2.1 Demographic parameters 79
2.3.2.2 Sexual behaviour parameters 81
2.3.2.3 Biological parameters 82
2.3.2.4 Intervention parameters 83
2.3.2.5 PrEP parameters 84
2.3.2.6 TasP parameters 85
2.3.2.7 Model fitting data 86
2.3.2.8 Cross-validation data 88
2.4 Model development 90

2.4.1 Model description 90
2.4.2 Equations 95
2.4.2.1 Ordinary differential equations (ODEs) 95
2.4.2.2 Entering the population 102
2.4.2.3 Turnover 103
2.4.2.4 Testing rates 109
2.4.2.5 ART initiation rates 110
2.4.2.6 Force of infection 110
2.4.2.7 Balancing of partnerships 111
2.4.2.8 Probability of sexual contact 113
2.4.2.9 HIV transmission probability per sex act 113
2.4.2.10 Relative risk of transmission 114
2.4.2.11 Disease progression and AIDS-related mortality 115
2.5 Model fitting 116
2.5.1 Pre-study model fitting (1986-2015) 117
2.5.2 PrEP/TasP study model fitting (2015-2017) 118
2.5.3 Scenarios 118
2.6 Discussion 118

3 Impact of pharmaceutical interventions for HIV in Cotonou 119
3.1 Overview 119
3.2 Methods 120
3.2.1 Scenarios 120
3.2.1.1 PrEP/TasP study scenario 120
3.2.1.2 Long-term 20-year intervention scenarios (2015-2035) . . 120
3.2.1.3 Counterfactual scenarios 121
3.2.2 Estimating impact 123
3.3 Results 123

3.3.1 Epidemic trajectory in Cotonou . . . 123
3.3.1.1 Model fitting up to 2017 . . . 123
3.3.1.2 Model projections from 2017 onwards . . . 126
3.3.2 Impact analysis . . . 127
3.3.2.1 Impact of the PrEP/TasP study . . . 127
3.3.2.2 Impact of long-term PrEP and TasP . . . 130
3.3.2.3 Sensitivity analyses . . . 135
3.4 Discussion . . . 141
3.4.1 What does the epidemic trajectory suggest? . . . 141
3.4.2 What was the impact of the PrEP/TasP study? . . . 142
3.4.3 What would be the impact of long term PrEP and TasP for FSW in Cotonou? . . . 143
3.4.4 Comparing impact with other studies . . . 144
3.4.5 Strengths and limitations . . . 145
3.4.6 Conclusions . . . 149

4 Estimating the impact of COVID-19 on HIV epidemiology in Cotonou 151
4.1 Overview . . . 151
4.2 Introduction . . . 152
4.2.1 The COVID-19 pandemic . . . 152
4.2.2 COVID-19 in Benin . . . 154
4.2.3 Possible impacts of COVID-19 on HIV epidemiology in Cotonou . 156
4.2.4 Previous studies estimating the impact of COVID-19 on HIV epidemiology . . . 157
4.2.5 Motivation and aims . . . 158
4.3 Methods . . . 159
4.3.1 Scenarios . . . 161
4.3.1.1 Plausible scenarios . . . 163

4.3.1.2 Exploratory scenarios . . . 164
4.3.1.3 Modelling the disruptions . . . 164
4.3.1.4 Estimating impact . . . 169
4.4 Results . . . 170
4.4.1 Plausible scenarios . . . 170
4.4.2 Exploratory scenarios . . . 171
4.4.2.1 Single scenarios . . . 171
4.4.2.2 Combination scenarios . . . 176
4.4.2.3 Group-specific scenarios . . . 179
4.5 Discussion . . . 182
4.5.1 COVID-related disruptions to HIV interventions . . . 182
4.5.1.1 Plausible scenarios . . . 182
4.5.1.2 More recent data . . . 183
4.5.2 The disproportionate impact on key populations . . . 184
4.5.3 Core group effect . . . 185
4.5.4 Other patterns in the impacts of disruptions . . . 186
4.5.5 Comparison to previous studies . . . 187
4.5.6 Strengths and limitations . . . 188
4.5.7 Implications . . . 188

5 Estimating the impact of non-pharmaceutical interventions for COVID-19 in Weifang 191

5.1 Introduction . . . 192
5.1.1 The COVID-19 epidemic in Weifang, China . . . 192
5.1.2 Motivation and aims . . . 193
5.2 Methods . . . 194
5.2.1 Epidemiological investigation, sampling, and genetic sequencing . . . 194
5.2.2 Mathematical model . . . 195

5.2.3 Model parameterisation . . . 198
5.2.4 Phylogenetic analysis . . . 198
5.3 Results . . . 200
5.3.1 Main analysis . . . 200
5.3.1.1 Model fitting . . . 200
5.3.1.2 Epidemic trajectory . . . 202
5.3.1.3 Phylogenetic analysis . . . 204
5.3.2 Sensitivity analyses . . . 207
5.3.2.1 Initial number of susceptibles . . . 207
5.3.2.2 Offspring distribution heterogeneity . . . 210
5.4 Discussion . . . 212
5.4.1 What was the true number of infections in Weifang? . . . 212
5.4.2 What was the impact of NPIs in Weifang? . . . 212
5.4.3 Limitations and strengths . . . 213
5.4.4 Conclusions . . . 215
5.5 Data availability . . . 215

6 Estimating the impact of non-pharmaceutical interventions for COVID-19 in England 217

6.1 Overview . . . 217
6.2 Introduction . . . 218
6.2.1 The COVID-19 epidemic in England . . . 218
6.2.2 SARS-CoV-2 sequence data . . . 220
6.2.2.1 The variant of concern B.1.1.7 . . . 221
6.2.3 Estimating the impact of lockdown . . . 223
6.2.3.1 The challenges . . . 223
6.2.3.2 Previous studies on the impact of lockdown . . . 224
6.2.3.3 Previous studies on the impact of the tier system . . . 226

6.2.4 Non-pharmaceutical interventions as a continuous variable 226
6.2.5 Aims and justification 228
6.3 Materials and Methods 230
6.3.1 Summarising NPI covariate data 230
6.3.2 S-gene target failure data 236
6.3.3 Sub-sampling the sequence alignment 237
6.3.4 Producing dated phylogenetic trees 240
6.3.5 Phylodynamic modelling 243
6.3.5.1 Implementation 243
6.3.5.2 Testing association between covariates and viral transmission 244
6.4 Results 247
6.4.1 Case study A: England during the early months of 2020 247
6.4.2 Case study B: Leicester during the spring and summer of 2020 . . . 253
6.4.3 Case study C: England during the winter of 2020-2021 260
6.4.4 Case study D: Variant of concern B.1.1.7 263
6.5 Discussion 269
6.5.1 The reliability of phylodynamic modelling for SARS-CoV-2 269
6.5.1.1 Robustness and corroboration from testing 269
6.5.1.2 Added evidence from analyses by lineage 270
6.5.1.3 Phylodynamic reconstruction is robust to using global ML tree 271
6.5.2 Impact of NPIs 271
6.5.2.1 How do we best evaluate association between NPI and viral transmission? 272
6.5.2.2 Did stricter restrictions control the virus better? 273
6.5.2.3 Did movement correlate with transmission? 274
6.5.2.4 Was lockdown 3 more effective than lockdown 2? 276
6.5.3 Limitations 276

6.5.4 Conclusions 278

7 Discussion **279**

7.1 Overview 279

7.2 Summary of this work 280

7.3 Implications for controlling HIV epidemics 283

7.3.1 Pharmaceutical interventions among FSW 283

7.3.2 Continued focus on key populations 285

7.3.3 Pharmaceutical vs. non-pharmaceutical interventions 286

7.3.4 Future work 287

7.4 Implications for controlling COVID-19 epidemics 289

7.4.1 Impact of non-pharmaceutical interventions 289

7.4.2 Real-time surveillance and variants of concern 291

7.4.2.1 Recommendations for sequencing efforts 292

7.4.3 Future work 293

7.5 Broader themes 295

7.5.1 Taking advantage of the available data 295

7.5.2 Overfitting and bias 297

7.5.3 Discerning complex trajectories 298

7.5.4 Considering natural history 299

7.5.4.1 Speed of outbreak 299

7.5.4.2 Virus evolution 300

7.5.5 Estimating the impact of interventions depends on context 301

7.5.5.1 When the intervention is easy to describe 301

7.5.5.2 When the intervention is hard to describe 302

7.6 Final conclusions 304

Bibliography **307**

Appendix A Supplementary Figures 353

List of Figures

2.1 Recruitment, movement and sexual mixing between groups, represented by green, black and blue lines respectively. Groups within the dashed rectangles form the sexual transmission network of Grand Cotonou. E_i represents the numbers of HIV-uninfected individuals in group i recruited to the model; P_1 represents the number of HIV-infected professional FSW recruited. Commercial partnerships are only formed between professional and part-time FSW and clients. NYSA = not yet sexually active. FSW = female sex worker. . . 91

2.2 The model flow diagram, representing progression through state of care (r) and stage of infection (s). PrEP initiation, HIV testing, and ART initiation are represented by blue, green and purple arrows respectively. Red arrows represent mortality and ageing; double red arrow indicates increased mortality due to HIV infection. Orange arrows represent HIV infection. All risk groups are represented by these compartments, but some compartments will always be empty for some groups (i.e. "on PrEP" for groups who are not professional FSW). 94

3.1 Model fitting to key targets up to 2017: The results show model predictions from the 111 posterior parameter sets, in grey, green and orange solid lines (median) and shaded regions (95% uncertainty intervals (UI) across all fits) representing pre-study, counterfactual (no PrEP or TasP) and 2-year PrEP/TasP study scenarios respectively, compared to available fitting HIV prevalence and ART and PrEP coverage data shown in thick bars before 2015 (grey) and during the 2-year PrEP/TasP intervention study between 2015-2017 (orange). Results are shown for HIV prevalence (%) among A) all women, B) all men, C) professional FSWs, D) clients of FSWs; ART coverage (%) among E) HIV-infected pFSW and F) all HIV-infected people; G) number of pFSWs on PrEP, and H) on ART. Additional cross-validation data not used at the fitting stage are also presented in thin bars. 125

3.2 Impact of the 2-year PrEP/TasP study for PrEP and TasP combined (scenario A1), PrEP alone (scenario A2) or TasP alone (scenario A3) alone, among professional FSWs ("pFSW"), clients and the whole population ("All") in terms of percentage of HIV infections prevented over (A) 2 years and (B) 20 years, and C) life-years gained (LYG) over 20 years, compared to counterfactual scenario (no PrEP; ART coverage 49% and 52% in 2017 and 2035 respectively). Boxplots represent the median impact (central horizontal line), 25th to 75th percentiles (box) and 95% UI (whiskers) across 111 parameter sets. Labels in panels A and B show median percentage (and number in brackets) of infections prevented; panel C shows median LYG. PrEP and ART coverage represent median posterior values across all model fits for each scenario as shown in Table 3.
† Median coverage of HIV-uninfected and HIV-infected for PrEP and ART respectively at end of the 2-year study.
‡ Median coverage in 2035, 20 years after start of PrEP/TasP study. 128

3.3 Impact of long-term 20-year PrEP (A, C) or TasP (B, D) intervention scenarios separately in terms of percentage of infections prevented (A, B) and life-years gained (C, D) among pFSW and the whole population ("All"), compared to the counterfactual scenario (no PrEP; 52% HIV-infected pFSW ART coverage in 2035). Boxplots represent the median impact (central horizontal line), 25th to 75th percentiles (box) and 95% UI (whiskers) across 111 parameter sets. The PrEP and ART coverages represent median posterior values in 2035 (among HIV-uninfected and HIV-infected pFSWs respectively) across all model fits for each scenario as shown in Table 3. Adherence to PrEP assumed either as observed in the 2-year study (green), or perfect (orange). Labels in panels A and B show median percentage (and number in brackets) of infections prevented; panels C and D show median life-years gained. 131

3.4 Impact of combined 20-year PrEP interventions and TasP scale-up (88% coverage HIV-infected pFSW in 2035), in terms of percentage of infections prevented (A) and life-years gained (B) among pFSW and the whole population ("All"), compared to the counterfactual scenario (no PrEP; 52% HIV-infected pFSW ART coverage in 2035). Boxplots represent the median impact (central horizontal line), 25th to 75th percentiles (box) and 95% UI (whiskers) across 111 parameter sets. PrEP coverages represent median posterior values in 2035 (among HIV-uninfected pFSW) across all model fits for each scenario as shown in Table 3; adherence assumed as observed in 2-year study. Labels in panel A show median percentage (and number in brackets) of infections prevented; panel B shows median life-years gained. . 134

3.5 Impact of long-term 20-year PrEP intervention scenarios in terms of percentage of infections prevented (A) and life-years gained (B) among pFSW and the whole population ("All"), compared to the counterfactual scenario (no PrEP). Boxplots represent the median impact (central horizontal line), 25th to 75th percentiles (box) and 95% UI (whiskers) across 111 parameter sets. The PrEP coverages represent median posterior values in 2035 (among HIV-uninfected pFSW) across all model fits for each scenario as shown in Table 3, as well as a scenario of 69% PrEP coverage. Adherence to PrEP assumed either as observed in the 2-year study (green), or perfect (orange). Labels in panels A shows median percentage (and number in brackets) of infections prevented; panel B shows median life-years gained. 136

3.6 Impact of long-term 20-year PrEP intervention scenarios in terms of percentage of infections prevented (A) and life-years gained (B) among pFSW and the whole population ("All"), compared to the counterfactual scenario (no PrEP). Boxplots represent the median impact (central horizontal line), 25th to 75th percentiles (box) and 95% UI (whiskers) across 111 parameter sets. The PrEP coverages represent median posterior values in 2035 (among HIV-uninfected pFSW) across all model fits for each scenario as shown in Table 3. The per sex-act reduction in transmission (efficacy) of PrEP in the intermediate adherence category (0.3 – 40 ng/ml) is assumed to be 30 – 49% (green), 0% (yellow) or 92% (purple). Labels in panels A shows median percentage (and number in brackets) of infections prevented; panel B shows median life-years gained. 138

3.7 Impact of long-term 20-year PrEP intervention scenarios in terms of percentage of infections prevented (A) and life-years gained (B) among pFSW and the whole population ("All"), compared to the counterfactual scenario (no PrEP). Boxplots represent the median impact (central horizontal line), 25th to 75th percentiles (box) and 95% UI (whiskers) across 111 parameter sets. The PrEP coverages represent median posterior values in 2035 (among HIV-uninfected pFSW) across all model fits for each scenario as shown in Table 3. The per sex-act reduction in transmission (efficacy) of PrEP in the intermediate adherence category (0.3 – 40 ng/ml) is assumed to be 30 – 49% (green), 0% (yellow) or 92% (purple). Labels in panels A shows median percentage (and number in brackets) of infections prevented; panel B shows median life-years gained. 140

4.1 Relative change in cumulative number of incident HIV infections and HIV-related deaths over 1 year among professional FSW (green), clients (purple) and overall (orange) under scenarios assuming 6-month single disruptions to all groups, compared to base-case (no COVID-19 nor COVID-19 response) in Cotonou. Squares and lines represent median and 95% percentiles of estimates across 111 model fits. 172

4.2 Relative change in cumulative number of incident HIV infections over 1 year and 5 years among professional FSW, clients and overall under single scenarios assuming 3 (light green), 6, 12 and 18 (dark green) month disruptions, compared to base-case in Cotonou. Squares and lines represent median and 95% percentiles of estimates across 111 model fits. 174

4.3 Relative change in cumulative number of HIV-related deaths over 1 year and 5 years among professional FSW, clients and overall under single scenarios assuming 3 (light green), 6, 12 and 18 (dark green) month disruptions, compared to base-case scenario (no COVID-19 nor COVID-19 response) in Cotonou. Squares and lines represent median and 95% percentiles of estimates across 111 model fits. 175

4.4 Relative change in cumulative number of new HIV infections and HIV-related deaths over 1 year among professional FSW, clients and overall under combination scenarios assuming 6 month disruptions, compared to base-case in Cotonou. For each panel, every new line indicates an additional disruption which occurs on top of the previous one. The level of reduction represented in each disruption is 10% (light blue), 25% or 50% (dark purple). Squares and lines represent median and 95% percentiles of estimates across 111 model fits. 178

4.5 Relative change in cumulative number of new HIV infections over 1 year among professional FSW (green), clients (purple) and overall (orange) under scenarios with group-specific 6-month disruptions, compared to base-case in Cotonou. Squares and lines represent median and 95% percentiles of estimates across 111 model fits. 180

4.6 Relative change in cumulative number of new HIV-related deaths over 1 year among professional FSW (green), clients (purple) and overall (orange) under scenarios with group-specific 6-month disruptions, compared to base-case in Cotonou. Squares and lines represent median and 95% percentiles of estimates across 111 model fits.. 181

5.1 Epidemiological model and sample times. A) A diagram representing the structure of the epidemiological SEIR model which was fitted in tandem with the time scaled phylogeny. Infected and infectious individuals may occupy a low (I) or a high (J) transmission rate state to account for high dispersion of the reproduction number. B) Sampling density of sequences from inside (yellow) and outside (grey) of Weifang respectively through time. 199

5.2 Epidemiological trajectory of the Weifang SARS-CoV-2 epidemic in 2020 when fitting the SEIR model to genetic data (blue) and sampling only from prior (grey). Solid lines and shaded area reflect posterior median and 95% HPD. The vertical green solid line represents the date of the final NPIs in Weifang. The vertical dashed line represents the date of the last sequence sampled in Weifang. A) Cumulative estimated infections through time compared to cumulative cases (yellow points) reported by Weifang CDC. B) Daily estimated infections through time compared to daily reported cases (yellow points). C) Effective reproduction number through time R_t. The horizontal dotted line indicates $R_t = 1$. 203

5.3 Estimated reporting frequency. The yellow points and grey bars reflect the mean and 95% HPD cumulative estimated proportion of cases that were identified through time respectively. Results shown separately for analyses conducted with and without the sequence data. 204

5.4 A maximum likelihood tree estimated using IQTree using the same data as used for the Bayesian analysis. Red and grey dots represent samples inside and outside Weifang respectively. 205

5.5 A time-scaled phylogeny (maximum clade credibility tree) co-estimated with epidemiological parameters of the SARS-CoV-2 epidemic in 2020. Red and grey tips correspond to samples from inside and outside Weifang, China respectively. The credible interval of time to most recent common ancestor (TMRCA) is shown as a blue bar for all nodes with more than 50% posterior probability support. 206

5.6 A root to tip regression (red and black points indicate sample and internal nodes respectively) showing approximately linear increase in diversity with time of sampling. The red text indicates the mean substitution rate. 207

5.7 Assuming a mean initial susceptible prior of S = 9,086,241, cumulative estimated infections and effective reproduction number through time are shown when fitting the SEIR model to genetic data (blue) and sampling only from prior (grey). Solid lines and shaded area reflect posterior median and 95% HPD. The vertical dashed line represents the date of the last sequence sampled in Weifang; the horizontal dashed line represents $R(t) = 1$. Cumulative cases (yellow points) reported by Weifang CDC. 209

5.8 Assuming $\tau = 13$ (corresponding to dispersion $k = 0.2$), cumulative and daily estimated infections and effective reproduction number through time are shown when fitting the SEIR model to genetic data (blue) and sampling only from prior (grey). Solid lines and shaded area reflect posterior median and 95% HPD. The vertical dashed line represents the date of the last sequence sampled in Weifang; the horizontal dashed line represents $R(t) = 1$. Cumulative cases (yellow points) reported by Weifang CDC. 211

6.1 National time series data of movement trends for different location types during spring 2020, published by Google. Thin black line represents raw data. Thick red and blue lines represent local regression smoothing with parameters of 20% and 4% respectively. The dashed grey lines indicate date of national lockdown. 231

6.2 Time series composite data of UK government interventions in the early months of 2020, estimated by Oxford COVID-19 Government Response Tracker (OxCGRT). Thin black lines represent raw data. Thick red and blue lines represent local regression smoothing with parameters of 20% and 4% respectively. The dashed grey lines indicate date of national lockdown. . . . 232

6.3 Time series data of movement trends for different location types in Leicestershire during spring and summer 2020, published by Google. Thin black line represents raw data. Thick red and blue lines represent local regression smoothing with parameters of 20% and 4% respectively. The dashed grey lines indicate dates of national and local lockdowns.
* Mobility data for Leicestershire as of 18th March 2021 had too much missingness; national data was instead used here 234

6.4 Time series composite data of national UK government interventions in the spring and summer months of 2020, estimated by Oxford COVID-19 Government Response Tracker (OxCGRT). Thin black lines represent raw data. Thick red and blue lines represent local regression smoothing with parameters of 20% and 4% respectively. The dashed grey lines indicate dates of national and local lockdowns. 234

6.5 Time series data of movement trends for different location types in England during the winter period 2020-2021, published by Google. Thin black line represents raw data. Thick red and blue lines represent local regression smoothing with parameters of 20% and 4% respectively. The dashed grey lines indicate dates of national lockdowns. 235

6.6 Time series composite data of UK government interventions in the winter of 2020-2021, estimated by Oxford COVID-19 Government Response Tracker (OxCGRT). Thin black lines represent raw data. Thick red and blue lines represent local regression smoothing with parameters of 20% and 4% respectively. The dashed grey lines indicate start dates of lockdowns. 236

6.7 The epidemiological trajectory of SARS-CoV-2 in England during spring 2020. Thick solid lines and shaded areas represent median and 95% quantiles respectively, estimated by *mlesky* (green) and *skygrowth* (purple). Red shaded area represents period of first national lockdown in England. Black dotted line represents daily confirmed cases (smoothed and rescaled). Panel A: effective population size ($N_e(t)$) through time. Panel B: reproduction number ($R(t)$) through time. Panel C: frequencies of tips in each sample week in the phylogenies used. 248

6.8 Scatterplot of the daily incremental changes in covariates against incremental changes in median $R(t)$ estimated by *mlesky* for England. Panels A-F display Google mobility data; panels G&H display OxCGRT index data. Grey line shows least squares linear regression; coefficients are displayed top left of each panel. No correlation was found to be significant to 5%. 249

6.9 Distribution of beta coefficients when testing for univariate association between Google mobility and changes in effective population size estimated in *mlesky*. Boxes represent median and interquartile range; whiskers denote 95% quantiles. 250

6.10 Distribution of beta coefficients when testing for univariate association between daily changes in OxCGRT indices and daily changes in effective population size (i.e. growth rate) in England estimated in *mlesky*. Boxes represent median and interquartile range; whiskers denote 95% quantiles. . 251

6.11 Distribution of beta coefficients when testing for univariate association between daily changes in OxCGRT health containment index and daily changes in effective population size (i.e. growth rate) in England estimated in *mlesky*. Boxes represent median and interquartile range; whiskers denote 95% quantiles. A positive delay of 10 represents testing an association between the covariate at time t and R at time $t+10$ days. 252

6.12 The epidemiological trajectory of SARS-CoV-2 in Leicester during the spring and summer months of 2020, using 70% of available sequences. Thick solid lines and shaded areas represent median and 95% quantiles respectively, estimated by *skygrowth* (purple) and *mlesky* (green). The black dotted lines show confirmed cases in Leicester according to PHE (smoothed and rescaled). Red shaded areas represent the first national lockdown in England and local lockdown in Leicester. Panel A: effective population size ($N_e(t)$) through time. Panel B: reproduction number ($R(t)$) through time. Panel C: frequencies of tips in each sample week in the phylogenies used. . 254

6.13 The epidemiological trajectory of SARS-CoV-2 in Leicester during the spring and summer months of 2020. Thick solid lines and thin dashed lines represent median and 95% quantiles respectively, estimated by *skygrowth*. Colour represents the fraction of samples analysed, ranging from 5% at the lowest (purple) to 70% at the highest (green). Legend displays sampling fraction and number of tips per alignment and tree. Red shaded areas represent the first national lockdown in England and local lockdown in Leicester. Panel A: effective population size ($N_e(t)$) through time. Panel B: reproduction number ($R(t)$) through time. Panel C: frequencies of tips in each sample week in the phylogenies used. 255

6.14 The epidemiological trajectory of SARS-CoV-2 in Leicester during the spring and summer months of 2020, cropped to the beginning of July, and only for sampling fractions of 30% and above. Thick solid lines and thin dashed lines represent median and 95% quantiles respectively, estimated by *mlesky*. Colour represents the fraction of samples analysed, ranging from 30% at the lowest (purple) to 70% at the highest (green). Legend displays sampling fraction and number of tips per alignment and tree. Red shaded areas represent the first national lockdown in England and local lockdown in Leicester. Panel A: effective population size ($N_e(t)$) through time. Panel B: reproduction number ($R(t)$) through time. Panel C: frequencies of tips in each sample week in the phylogenies used. 257

6.15 Scatterplot of the daily incremental changes in covariates in Leicester against incremental changes in median $R(t)$ estimated by *mlesky* at 70% sampling. Panels A-F display Google mobility data; panels G&H display OxCGRT index data. Grey line shows least squares linear regression; coefficients are displayed top left of each panel. No correlation was found to be significant to 5%. 258

6.16 Distribution of beta coefficients when testing for univariate association between Google mobility and growth rate in Leicester (70% sampling), estimated with *mlesky*. Boxes represent median and interquartile range; whiskers denote 95% quantiles. 259

6.17 Distribution of beta coefficients when testing for univariate association between daily changes in OxCGRT indices and daily changes in effective population size (i.e. growth rate) in Leicester (70% sampling), estimated with *mlesky*. Boxes represent median and interquartile range; whiskers denote 95% quantiles. 260

6.18 The epidemiological trajectory of SARS-CoV-2 in England during the 2020-2021 winter period. Thick solid lines and shaded areas represent median and 95% quantiles respectively, estimated by *mlesky* (green) and *skygrowth* (purple). The black dashed line represents daily confirmed cases from PHE (smoothed and rescaled). Red shaded areas represent periods of lockdowns and tier systems; a darker box reflects more strict restrictions. Panel A: effective population size ($N_e(t)$) through time. Panel B: reproduction number ($R(t)$) through time. Panel C: frequencies of tips in each sample week in the phylogenies used. 261

6.19 Scatterplot of the daily incremental changes in covariates against incremental changes in median $R(t)$ estimated by *mlesky* for England during the winter period. Panels A-F display Google mobility data; panels G&H display OxCGRT index data. The lines show least squares linear regression; coefficients are displayed top left of each panel. The red line shows a significant correlation for residential mobility. 262

6.20 Distribution of beta coefficients when testing for multivariate association between presence of lockdowns 2 and 3 against second order changes in effective population size estimated in *mlesky*. Boxes represent median and interquartile range; whiskers denote 95% quantiles. 263

6.21 The epidemiological trajectories of 2 lineages (B.1.1.7 and B.1.177) of SARS-CoV-2, as well as a control group including samples from any lineage outside of B.1.1.7, in England during the 2020-2021 winter period. Thick solid lines and shaded areas represent median and 95% quantiles respectively, estimated by *mlesky*. B.1.1.7, B.1.177 and control group are shown in green, orange and purple respectively. Red shaded areas represent periods of lockdowns and tier systems; a darker box reflects more strict restrictions. Panel A: effective population size ($N_e(t)$) through time. Panel B: reproduction number ($R(t)$) through time. Panel C: frequencies of tips in each sample week in the phylogenies used. 264

6.22 Effective population size of B.1.1.7 estimated by *mlesky* compared to TPP-corrected estimates of SGTF by epiweek. 265

6.23 Distribution of beta coefficients when testing for univariate association between changes in mobility estimated by Google and reproduction number estimated in *mlesky*. Boxes represent median and interquartile range; whiskers denote 95% quantiles. Lineages B.1.1.7, B.1.177 and control are shown in green, orange and purple boxes respectively. 266

6.24 Distribution of beta coefficients when testing for multivariate association between presence of lockdowns 2 and 3 against second order changes in effective population size estimated in *mlesky*. Boxes represent median and interquartile range; whiskers denote 95% quantiles. Lineages B.1.1.7, B.1.177 and control are shown in green, orange and purple boxes respectively. . . . 267

6.25 Distribution of beta coefficients when testing for univariate association between daily changes in OxCGRT indices and daily changes in effective population size (i.e. growth rate) estimated in *mlesky*. Boxes represent median and interquartile range; whiskers denote 95% quantiles. Lineages B.1.1.7, B.1.177 and control are shown in green, orange and purple boxes respectively. 268

6.26 Distribution of beta coefficients when testing for univariate association between daily changes in OxCGRT health containment index and daily changes in effective population size (i.e. growth rate) in England estimated in *mlesky*. Boxes represent median and interquartile range; whiskers denote 95% quantiles. A positive time delay of 10 days represents testing an association between the covariate at time t and R at time $t + 10$ days. Lineages B.1.1.7 and control are shown in green and purple boxes respectively. 269

A.1 Posterior model predictions for total population size of Grand Cotonou from the 111 posterior parameter sets, showing median (solid line) and 95%UI across all fits (shaded regions) of counterfactual (no PrEP or TasP) scenario, compared to available fitting historical demographic data shown in thick bars (black) and estimated future predictions assuming constant population growth rate (red). 354

A.2 Posterior model predictions for the percent of each group in their respective gender (e.g. the percent of women who are professional FSW), from the 111 posterior parameter sets, showing median (solid line) and 95%UI across all fits (shaded regions) of counterfactual (no PrEP or TasP) scenario, compared to available fitting historical demographic data shown in red lines. 355

A.3 Posterior model predictions of number of professional FSW from the 111 posterior parameter sets, showing median (solid line) and 95% UI across all fits (shaded regions) of counterfactual (no PrEP or TasP) scenario, compared to available fitting historical demographic data shown with a red bar. . . . 355

A.4 Posterior model predictions of the ratio of women to men on ART from the 111 posterior parameter sets, showing median (solid line) and 95% UI across all fits (shaded regions) of counterfactual (no PrEP or TasP) scenario, compared to available fitting intervention data shown with a red bar. 356

A.5 Posterior model predictions from the 111 posterior parameter sets, showing median (solid line) and 95% UI across all fits (shaded regions) of counterfactual (no PrEP or TasP) scenario, of annual deaths due to HIV infection compared to available cross-validation UNAIDS Reference Group Spectrum model estimates (orange bars) in Cotonou. 356

A.6 Posterior model predictions from the 111 posterior parameter sets, showing median (solid line) and 95% UI across all fits (shaded regions) of counterfactual (no PrEP or TasP) scenario, of proportion of HIV-positive people on ART and virally suppressed, compared to available cross-validation UNAIDS Reference Group Spectrum model estimates (orange bars) in Cotonou. . . . 357

A.7 Posterior percent of commercial and non-commercial partnerships (shown in green and orange respectively) between professional FSW and clients protected by condoms from the 111 posterior parameter sets, showing median (solid line) and 95% UI across all fits (shaded regions) of counterfactual (no PrEP or TasP) scenario, compared to available historical behavioural data in bars. 357

A.8 Posterior percent of commercial and non-commercial partnerships (shown in green and orange respectively) protected by condoms between part-time FSW and clients from the 111 posterior parameter sets, showing median (solid line) and 95% UI across all fits (shaded regions) of counterfactual (no PrEP or TasP) scenario, compared to available historical behavioural data in bars. 358

A.9 Posterior percent of non-commercial partnerships between low-risk women and their partners (clients and low-risk men) protected by condoms from the 111 posterior parameter sets, showing median (solid line) and 95% UI across all fits (shaded region) of counterfactual (no PrEP or TasP) scenario, compared to available historical behavioural data in orange bars. 358

A.10 Model predictions from the 111 posterior parameter sets, in solid lines (median) and shaded regions (95%UI across all fits). Error bars represent cross-sectional HIV prevalence data up to 2017 by risk group. Projections are shown until 2035 in the absence of the PrEP/TasP intervention (counterfactual scenario). The “women” and “men” panels include all women and men in Cotonou, including key populations. 359

A.11 Daily incremental changes in the transformed time series data from Google mobility (blue lines) and OxCGRT indices (red lines). The re-scaled daily incremental changes in median $R(t)$ estimated with *mlesky* is superimposed on each panel (grey lines). 360

A.12 Scatterplot of the daily incremental changes in transformed covariates against incremental changes in median $R(t)$ shifted 10 days earlier estimated by *mlesky*. Panels A-F display Google mobility data; panels G&H display OxCGRT index data. Grey line shows least squares linear regression; coefficients are displayed top left of each panel. No correlation was found to be significant to 5% significance. 361

A.13 Scatterplot of the daily incremental changes in transformed covariates against incremental changes in median $R(t)$ shifted 10 days later estimated by *mlesky*. Panels A-F display Google mobility data; panels G&H display OxCGRT index data. Grey line shows least squares linear regression; coefficients are displayed top left of each panel. No correlation was found to be significant to 5% significance. 362

A.14 The epidemiological trajectory of SARS-CoV-2 in England during spring 2020. Thick solid lines and shaded areas represent median and 95% quantiles respectively, estimated by *mlesky* with (purple) and without (green) covariates. Red shaded area represents period of first national lockdown in England. Black dotted line represents daily confirmed cases (smoothed and rescaled). Panel A: effective population size ($N_e(t)$) through time. Panel B: reproduction number ($R(t)$) through time. Panel C: frequencies of tips in each sample week in the phylogenies used. 363

A.15 The epidemiological trajectory of SARS-CoV-2 in Leicester during the spring and summer months of 2020. Thick solid lines and thin dashed lines represent median and 95% quantiles respectively, estimated by *mlesky*. Colour represents the fraction of samples analysed, ranging from 5% at the lowest (purple) to 70% at the highest (green). Legend displays sampling fraction and number of tips per alignment and tree. Red shaded areas represent the first national lockdown in England and local lockdown in Leicester. Panel A: effective population size ($N_e(t)$) through time. Panel B: reproduction number ($R(t)$) through time. Panel C: frequencies of tips in each sample week in the phylogenies used. 364

A.16 The epidemiological trajectory of SARS-CoV-2 in Leicester during until July 2020. Thick solid lines and thin dashed lines represent median and 95% quantiles respectively, estimated by *mlesky*. Colour represents the fraction of samples analysed, ranging from 5% at the lowest (purple) to 70% at the highest (green). Legend displays sampling fraction and number of tips per alignment and tree. Red shaded areas represent the first national lockdown in England and local lockdown in Leicester. Panel A: effective population size ($N_e(t)$) through time. Panel B: reproduction number ($R(t)$) through time. Panel C: frequencies of tips in each sample week in the phylogenies used. . 365

A.17 The epidemiological trajectory of B.1.1.7 (red) and control (blue) lineages in England during the 2020-2021 winter period. Thick solid lines and shaded areas represent median and 95% quantiles respectively, estimated by *mlesky*. Effective population size, reproduction number and the ratio of reproduction numbers (B.1.1.7/control) are shown through time. Panel A: using the global ML tree Panel B: reconstructing an ML tree from the sequences 366

A.18 Comparison of effective population size of B.1.177 lineage against TPP-corrected SGTF case numbers by epiweek. 367

A.19 Comparison of effective population size of control lineages against TPP-corrected SGTF case numbers by epiweek. 368

A.20 Distribution of beta coefficients when testing for univariate association between changes in mobility estimated by Google (smoothed with loess parameter 20%) and reproduction number estimated in *mlesky*. Boxes represent median and interquartile range; whiskers denote 95% quartiles. Lineages B.1.1.7, B.1.177 and control are shown in orange, white and grey boxes respectively. 369

A.21 Distribution of beta coefficients when testing for univariate association between daily changes in OxCGRT (smoothed with loess parameter 4%) health containment index and daily changes in effective population size (i.e. growth rate) in England estimated in *mlesky*. Boxes represent median and interquartile range; whiskers denote 95% quartiles. A positive time delay of 10 days represents testing an association between the covariate at time t and R at time $t+10$ days. Lineages B.1.1.7 and control are shown in green and purple boxes respectively. 369

List of Tables

2.1 Demographic parameters for Cotonou 80

2.2 Sexual behaviour parameters for Cotonou 81

2.3 Biological parameters for Cotonou 82

2.4 Intervention parameters for Cotonou 83

2.5 PrEP parameters for Cotonou 85

2.6 TasP parameters for Cotonou 85

2.7 Outcomes and target range or estimate used for the A) pre-PrEP/TasP study model fitting (before 2015) and B) the PrEP/TasP study scenario model fitting (2015-2017). 87

2.8 Cross-validation data for Cotonou 89

2.9 The relative risk parameter employed according to the care state and stage of infection of the transmitter. 115

3.1 Short and long-term scenarios investigated. A) the 2-year PrEP/TasP study and (B) various extended and scaled up 20-year PrEP and TasP interventions. 122

4.1 The posterior distributions (median and 95% UI (2.5th and 97.5th percentiles) for relevant parameters in January 2020, following model calibration in Chapter 2. 160

4.2 Disruption scenarios simulated in Cotonou from January 2020 162

5.1 Summary of prior distributions for primary epidemiological and evolutionary parameter. 198

5.2 Summary of primary parameters for the main analysis, including mean estimated posterior and effective sample size due to autocorrelation. 201

5.3 Summary of primary parameters for the main analysis but sampling from the prior, including mean estimated posterior and effective sample size due to autocorrelation. 201

6.1 Characteristics of the time-scale phylogenies used in each case study 242

Nomenclature

Acronyms / Abbreviations

ACE2 Angiotensin converting enzyme 2

AIC Akaike information criterion

AIDS Acquired immune deficiency syndrome

ART Antiretroviral therapy

CD4 Cluster of differentiation 4

CFR Case fatality ratio

COG-UK COVID-19 Genomics UK Consortium

COVID-19 Coronavirus disease 2019

Ct Cycle threshold

DHS Demographic and Health Surveys

DNA Deoxyribonucleic acid

FSW Female sex worker

GC Grand Cotonou

GISAID Global initiative on sharing avian influenza data

HIV Human immunodeficiency virus

HPTN HIV Prevention Trials Network

IBBS Integrated bio-behavioural surveys

ICU Intensive care unit

IFR Infection fatality ratio

INSAE Institut national de la statistique et de l'analyse économique

ITU Intensive therapy unit

LTLA Lower Tier Local Authority

MEMS Micro-electromechanical systems

ML Maximum likelihood

MSM Men who have sex with men

NHS National Health Service

NPI Non-pharmaceutical interventions

NYSA Not yet sexually active

ODE Ordinary differential equation

PAF Population attributable fraction

PCR Polymerase Chain Reaction

pFSW Professional female sex worker

PHE Public Health England

PLHIV People living with HIV

PMTCT Prevention of mother-to-child transmission

PrEP Pre-exposure prophylaxis

RNA Ribonucleic acid

RR Relative risk

RT-qPCR Quantitative reverse transcription polymerase chain reaction

SARS Severe acute respiratory syndrome

SARS-CoV-2 Severe acute respiratory syndrome coronavirus 2

SGTF S-gene target failure

SIDA Syndrome d'immunodéficience acquise (AIDS)

SIR Susceptible-Infected-Recovered

STD Sexually transmitted disease

STI Sexually transmitted infection

STP Sustainability and transformation partnerships

TasP Treatment-as-prevention

UNAIDS The Joint United Nations Programme on HIV/AIDS

UTLA Upper Tier Local Authority

VMMC Voluntary medical male circumcision

VOC Variant of concern

VOI Variant of interest

VUI Variant under investigation

WHO World Health Organisation

Overview

Motivation

Since its origins over one hundred years ago [1], mathematical modelling of infectious disease epidemics has become increasingly important [2]. There are now large bodies of academic literature dedicated to modelling all major global infectious pathogens, from understanding basic epidemiology to estimating the impact of control measures [3]. For HIV, malaria and tuberculosis, among the main causes of global deaths by infectious disease [4], mathematical modelling forms an instrumental contribution to almost all policy decisions regarding their control [5].

In the era of COVID-19, mathematical modelling of infectious disease epidemics has reached unprecedented influence and attention [6]. There has been a vast expansion in the abundance, availability and diversity of data describing the COVID-19 pandemic, collected and published by governments and public health bodies across the world [7]. As a result SARS-CoV-2, the viral agent causing COVID-19 [8], quickly became an intensively studied virus [9], and the academic literature has since produced numerous publications aiming to understand the epidemic in various contexts [10].

Further, the responses of many governments to the COVID-19 pandemic has relied heavily on modelling studies, including real-time analysis of the impact of lockdowns, school closures and various other restrictions [11]. Although modelling has previously been used in outbreak response, notably for pandemic influenza, Ebola and Zika [12, 13], its scope and influence for COVID-19 policy, particularly for Western governments, is significantly higher [11]. This marks an advancement in evidence-led policy for infectious disease control, but also raises many questions regarding the utility of modelling in different situations [14–16].

In this thesis, I explore different mathematical models for estimating the impact of interventions for viral epidemics. To do this, I present two distinct approaches. The first is to use a transmission dynamic model to understand the HIV epidemic in Cotonou, Benin, and to

estimate the impacts of pharmaceutical interventions and COVID-related disruptions on HIV epidemiology. The second is to use phylodynamic models to understand the COVID-19 epidemics of England and China, and to estimate the impact of non-pharmaceutical interventions on COVID-19 epidemiology.

The main contributions of this thesis are specific to each local disease context: HIV control in Benin, and COVID-19 control in China and England. Secondarily, given the current unprecedented importance of mathematical modelling, I also explore the broader implications of my research for the modelling of infectious diseases, in light of the diversity in disease dynamics, data, and techniques employed.

Aims

The primary aim of this thesis is to use two approaches, epidemiological and evolutionary modelling, to understand the spread of disease and assess the impact of interventions for two different pathogens. A secondary aim of this thesis is to explore common themes emerging from these diverse analyses, the strengths and limitations of each, and to understand considerations for model choice in any given epidemiological context.

In the first half of the thesis, I explore the epidemiological impact of interventions for controlling HIV in Cotonou, Benin, including antiretroviral treatment-as-prevention (TasP) and pre-exposure prophylaxis (PrEP) among female sex workers (FSW). I use an epidemiological model to do this, specifically a dynamic transmission model, fitted to decades of local demographic and epidemiological data.

The specific objectives in these chapters are to answer the following research questions:

1. *Is the city of Cotonou, Benin on track to meet the 2030 UNAIDS goal of eliminating HIV/AIDS as a global health threat?*

2. *What was the impact of a two-year intervention of PrEP and TasP among FSW in Cotonou?*

3. *What would be the impact of scaling up PrEP and TasP among FSW over twenty years in Cotonou?*

4. *What was the impact of disruptions to HIV prevention services and sexual behaviour, caused by the COVID-19 epidemic in Cotonou?*

In the second half of the thesis, I investigate the epidemiological impact of non-pharmaceutical interventions (NPIs), such as lockdowns and travel restrictions, imposed to control the

COVID-19 epidemics in China and England. In these analyses, I employ evolutionary models: both parametric and non-parametric phylodynamic methods. These models are fitted to whole genome sequences of SARS-CoV-2, the virus responsible for COVID-19.

The specific objectives in these chapters are to answer the following research questions:

5. *What was the true number of COVID-19 infections in Weifang, China, during January and February 2020?*

6. *What was the impact of NPIs on COVID-19 transmission in Weifang?*

7. *What was the impact of NPIs on COVID-19 transmission in England?*

8. *Was the B.1.1.7 variant spreading faster than previous variants in England?*

Each modelling approach utilised in this thesis represents a different way to understand the spread of disease, and to estimate the impact of interventions. I discuss the appropriateness, robustness and utility of each modelling approach given the available data, the epidemiological context, and the specific research question.

Summary of chapters

In this section, I provide a brief description of the objectives and contents of each chapter in the thesis.

Chapter 1 introduces background information on the historical usage of modelling to understand infectious disease dynamics. In particular, there is a detailed discussion of the dynamic transmission and phylodynamic modelling methods. Further, there is an introduction to the natural history, epidemiology and interventions for HIV and SARS-CoV-2.

Chapter 2 builds the foundation for the transmission model analysis for HIV in Cotonou, Benin. It begins by providing a background to the local HIV epidemic. Next, an analysis and summary of the available data is described. Finally, a mathematical description of the dynamic transmission model development is provided, as well as the procedure for fitting the model to available data. The dynamic transmission model described at the end of this chapter is used for the analyses in Chapter 3 and Chapter 4.

Chapter 3 evaluates the impact of pharmaceutical interventions among FSW for HIV in Cotonou. First, the trajectory of the HIV epidemic of Cotonou is estimated by fitting the model to available data. Next, the impact of a real-world two-year PrEP/TasP demonstration study that was conducted among FSW is estimated, as well as long-term PrEP and TasP

scale-up scenarios with varying adherence and coverage. The fitted model resulting from this chapter is extended for analysis in Chapter 4.

Chapter 4 explores the impact of COVID-19 on the HIV epidemic of Cotonou. First there is a description of the COVID-19 epidemic in Benin, and the NPIs implemented by the government. Next, based on available data, disruptions to already-present HIV interventions are simulated, and the consequent impacts on HIV transmission are estimated.

Chapter 5 evaluates the impact of NPIs in response to the COVID-19 epidemic in Weifang, China. It begins by introducing the information known about the local COVID-19 epidemic, and describes the NPIs implemented. It then estimates the epidemic trajectory using a parametric phylodynamic model fitted to available genomic data, and explores the effects of physical distancing and lockdowns near the beginning of the outbreak.

Chapter 6 evaluates the impact of NPIs in response to the COVID-19 epidemic in England. There is a detailed discussion of the COVID-19 epidemic in England, including the emergence of viral genetic variants. The chapter is then divided into four case studies: A) England in spring 2020, B) Leicester in summer 2020, C) England in winter 2020-2021, D) England in winter 2020-2021 by genetic variant. For each analysis, the epidemic trajectory and the impacts of NPIs are estimated using non-parametric phylodynamic models fitted to available genomic data.

Chapter 7 discusses the implications of the thesis for responding to the HIV epidemic in Cotonou, and the COVID-19 epidemics in Weifang and England. Further, this chapter brings together themes common to all of the analyses. There is a discussion of the validity and applicability of the modelling frameworks used in this thesis, and a comparison of their strengths and limitations. Finally, broader implications of the research are explored.

Chapter 1

Background

1.1 Modelling of infectious disease dynamics

1.1.1 Uses of modelling in viral epidemic response

Mathematical analysis of epidemics dates back to the middle of the 18th century, when Daniel Bernoulli estimated the effectiveness of immunising uninfected patients against smallpox [17]. Later, a discrete time model was implemented by WH Hamer in the early 20th century to investigate cycles of measles outbreaks [1]. The field of research grew throughout this period, and by the 1990s, there was a steady number of academic papers featuring models of infectious disease epidemics, averaging around 100 articles per year [2]. From the turn of the millennium the number of papers increased dramatically, particularly in top journals such as Nature and Science [2].

Mathematical modelling is now used to understand and predict epidemics of most major pathogens around the world. For slower moving or endemic diseases, such as HIV or malaria, policymakers benefit from modelling to decide upon the most impactful or cost-effective interventions [18–21]. This is especially crucial in places such as Sub-Saharan Africa, where resources are very limited [22]. One example of the influence of mathematical models for designing policy is an article published in The Lancet by Granich and colleagues in 2009 [23]. Using both a stochastic and deterministic model of HIV transmission, they found that a massively scaled-up test-and-treat strategy could cause a marked decline in HIV incidence in South Africa, moving high-burden settings into an elimination phase, if high coverage could be achieved. As a result, UNAIDS made this an official policy for their 90-90-90 agenda, to help bring the AIDS epidemic to an end [24, 25].

For fast-paced epidemics, mathematical modelling has also been impactful. Before the arrival of the 2009 H1N1 influenza pandemic, mathematical models had already been developed to assess its potential impact and to aid preparedness in the UK [26]. The UK government's response, which included school and port closures, was informed by advice from mathematical modellers [27]. On an international level, these same models, which included seasonality, viral antigenic changes, severity of infection, and impact and cost-effectiveness of pharmaceutical interventions, provided evidence to the World Health Organisation (WHO) for operational response and policy [15].

In 2014, near the beginning of the West African Ebola outbreak, early modelling studies from the United States Center for Disease Control and Prevention (CDC) predicted an unmitigated scenario in which over one million cases of Ebola were estimated [28]. As a result of these analyses, the political will to respond quickly was generated [13]. Other model analyses provided estimates of potential future incidence, bed requirements, risk factors of increased transmission and case fatality ratio [29], all of which were invaluable to the WHO for budgeting and strategic planning [30]. Further, ring vaccination trials were designed with the help of mathematical models to evaluate efficacy during periods of reduced incidence [31, 32]. Several years later, during the Ebola outbreak in DRC in 2018-2020, these previously developed methods aided in designing impactful vaccination strategies to contain the virus [12].

The COVID-19 pandemic has seen a significant increase in the influence of mathematical modelling, domestically and globally [16]. From January 2020, simple epidemiological models were used to estimate the total number of infected cases of COVID-19 (then termed '2019-nCoV') in Wuhan, the mean time to detection of a case, and the incubation period [33]. Later, as reports from China indicated a very serious epidemic was emerging, mathematical and statistical modelling were used to estimate transmissibility and severity of COVID-19 [34, 35], during a time when lockdowns were imposed in Hubei and other provinces. Phylogenetic and phylodynamic analyses were used to further estimate disease burden, viral growth rates, heterogeneity in secondary infections, and evolutionary clock rate [36].

In the UK context, mathematical modelling of COVID-19 epidemics have been very influential for response policy. For example, model analyses were assessed by the Scientific Advisory Group for Emergencies (SAGE), which recommended restrictive measures for viral control to the British government [37].

On 16th March 2020, Imperial College published a modelling study estimating the impact of non-pharmaceutical interventions (NPIs) on COVID-19 mortality and healthcare demand [38]. With an individual-based simulation model, they estimated that a completely unmitigated

COVID-19 epidemic in the UK would cause over 500,000 deaths, due to the number of severe cases overwhelming healthcare capacity at the time. They argued that social distancing, case isolation and quarantines could control the epidemic and avoid this scenario. On the same day as the publication, SAGE strongly recommended a swift introduction of restrictive social distancing measures to the UK government [37]. One week later, on 23rd March 2020, the Prime Minister implemented the first national lockdown in the UK, following the evidence from these modelling studies [39].

In parallel with SAGE, other groups such as the Joint Committee on Vaccination and Immunisation (JCVI) and New and Emerging Respiratory Virus Threats Advisory Group (NERVTAG) provide counsel to the British authorities regarding COVID-19 policy [40, 41], with the benefit of model analyses. Notably, NERVTAG advised further restrictions in December 2020 following evidence from phylogenetic analysis that the novel variant of concern, later named 'B.1.1.7' or the 'alpha variant', was more transmissible than previously circulating lineages [42]. As of 11th August 2021, the UK government continues to receive evidence from mathematical models to navigate the best course of action regarding the COVID-19 pandemic.

1.1.2 Different modelling frameworks

There are myriad approaches to modelling infectious disease epidemics. In this section I describe two families of models: dynamic transmission modelling and phylodynamic modelling. I provide the theoretical basis, strengths and limitations of each. Given that a model analysis is limited by the quality of the data, I also discuss the ideal data requirements for reliable epidemic representation.

1.1.2.1 Dynamic transmission modelling

As a result of being spread from person to person, a communicable disease has unique epidemiological properties. Among these is the probability of a susceptible individual becoming infected, which depends on the current number (or density) of infectious individuals [43]. The number of infectious individuals varies through time, thus the rate of infection, also termed the 'force of infection', follows a non-linear pattern [2].

The most studied framework to capture this epidemiological non-linearity is the 'SIR' model, a form of compartmental model [44]. In this model susceptible (S), infected (I) and recovered (R) individuals are treated as separate compartments, usually represented by ordinary dif-

ferential equations [44]. From this framework, useful parameters for understanding disease spread can be calculated. For example, the basic reproduction number (R_0), defined as the average number of secondary infections an infectious individual will cause in a fully susceptible population, is an effective measure of the potential of a pathogen to spread [45]. Given that the rate of infection changes with time in the SIR model, it can be classed as a 'dynamic transmission' model.

The basic SIR framework has been adapted and extended to better represent the natural history of pathogens. In this section, I focus on models of HIV epidemics, but similar principles apply to other pathogens. For example, compartments relating to the acute, chronic and AIDS stages of HIV infection are commonly modelled [46]; the chronic stage can be further divided into stages of progressively smaller CD4 T-cell counts, representing the pathogenic influence of HIV on the immune system [47]. The advent of antiretroviral therapy (ART), with its curtailing of disease progression, mortality and transmission, has led to further elaborations of the transmission dynamic model to include compartments of ART and viral suppression [18, 48].

Most major infectious diseases have been modelled with the dynamic transmission modelling framework, due to its flexibility to adapt to specific demography, natural history and interventions. Different age groupings [49], sexes [48], risk groups [50] and race [51] can all be incorporated in the model, represented by separate compartments. For HIV epidemics, where sexual transmission is the main driver, these delineations are often necessary to accurately represent patterns of sexual activity [52]. Within-host dynamics can also be modelled [53, 54], which can be nested within a between-host model if desired [55]. The deterministic model can be translated into a stochastic process [2], allowing for potential epidemic extinction [56]. With the increased complexity of models in recent years, there is a tendency to forego analytical solutions for numerical ones [57].

Another key advantage of the dynamic transmission modelling framework is its power to capture chains of disease transmission. As mentioned previously, the force of infection applied to susceptibles depends on the density of infected individuals (i.e. disease prevalence). Therefore, any infection incurred in the present could cause further infections as time progresses. When modelling the impact of HIV interventions for example, secondary infections averted in the long term can be added to primary infections averted in the short term, providing a fuller estimate of impact. Conversely, a comparative modelling analysis showed that a static model of HIV transmission, which calculates the distribution of incident HIV infections by subgroup over one year based on the sexual behaviour and HIV prevalence in their partners [58], will estimate a much lower impact for HIV prevention interventions

[59]. By including secondary infections, the dynamic transmission model is a powerful and flexible tool for measuring a fuller estimate of the impact of disease control interventions.

Data requirements for dynamic transmission modelling

A dynamic transmission model with more compartments will require more data. If, for instance, there are twenty compartments reflecting individuals with HIV in different phases of infection or at different stages in the HIV care cascade, then there are twenty unique parameterisations that must be defined *a priori*. Another elaboration, as discussed earlier, is the inclusion of separate compartments that divide demographics of individuals, such as age, sex, risk group, ethnicity, etc. These compartments may differ in terms of birth and mortality rates. In HIV models, they may interact differently with other groups in terms of sexual partnership. All of these different parameterisations require data.

Further, there are transitions between compartments in the model at given rates, for example moving from HIV undiagnosed, to diagnosed with HIV, to diagnosed with HIV and on ART [2, 50, 18]. The more compartments there are, the more transitions there are, and each rate will require data to for parameterisation. Further, in many models, individuals can transition from one compartment into several others, meaning that the number of transitions increases geometrically with the number of compartments. For the example of HIV compartments at different stages of infection or in different states of engagement in the HIV care cascade, data that describe transition rates may be clinical studies on disease progression (or rate of depletion of CD4 T-cells) [60], or average waiting times until receiving an HIV test or initiating ART [61].

In the simple SIR model, transmission can be modelled as a simple constant per capita rate (β). However, in many cases (such as in HIV modelling) this rate is made more complex to capture temporal variance in sexual activity (e.g. number of partnerships and condom use), prevalence of co-morbidities (e.g. HSV-2), circumcision, pharmaceutical interventions (e.g. PrEP and ART), among many others [62–64]. These elaborations require data to parameterise: not only trends through time, by group or by age, but also their estimated impact on per sex act probability of transmission [65].

A final yet indispensable data requirement is that for model fitting, also known as calibration. Model fitting is an optimisation process to find sets of parameters that minimise the sum of errors between a model output and empirical observations [66]. Put another way, model fitting maximises the extent to which a model agrees with data [67]. The posterior parameter sets (or distributions) obtained as a result of model fitting can themselves be useful (e.g. to

estimate basic reproduction number [68]), but a fitted model can also be used to project an epidemic into the future, or to simulate various intervention scenarios [38].

In the case of dynamic transmission models, they are usually fitted to epidemiological data, such as disease prevalence [48], case counts [69], or disease-related deaths [70], but can also be fitted to demographic, behavioural and intervention data [18, 50]. Various approaches have been developed to achieve model fits. One approach is using Latin Hypercube Sampling (LHS) in combination with trajectory matching [50]. This involves sampling all free parameters from pre-defined uniform distributions, creating thousands or millions of independent parameter sets. Next, these parameter sets are used as inputs, and those which produce model outcomes that meet pre-defined criteria are kept as model fits. Parameter sets producing outcomes that do not meet the criteria are discarded.

Another approach is using Bayesian Markov chain Monte Carlo (MCMC) methods [71]. Briefly, the posterior parameter space is explored by attempting iterative rounds of newly proposed parameter values, and evaluating a likelihood function. Each parameter value is proposed based on a Markov Chain, i.e. that it depends on its previously proposed value. A typical proposition might be sampled from a normal distribution centred on the previous value. Parameters may be sampled one at a time, or several at a time.

A version of the Bayesian approach can also be implemented, called incremental mixture importance sampling (IMIS) [72]. After sampling from the prior, importance weights are calculated from an initial likelihood. A multivariate normal distribution is centred on the point with the highest weight at each iteration of the sampler. The sampling distribution thereby becomes a combination of the weighted likelihood function and the prior, allowing for more sampling of underrepresented areas of parameter space [73].

In summary, the dynamic transmission model is an effective way to represent an epidemic and to synthesise the information from many data sources. It is important to note that although I have discussed examples of using data to parameterise and fit models of HIV transmission, the same principle applies to any transmissible pathogen.

In other chapters in this thesis, there is further discussion of the historical use of dynamic transmission modelling to estimate the impact of interventions. Chapter 2 features a detailed discussion of previous modelling studies to estimate the impact of HIV prevention interventions in Cotonou, Benin. Chapter 4 presents previous modelling of the impact of COVID-related disruptions on HIV epidemiology in Sub-Saharan Africa. Chapter 6 provides a summary of the previous modelling studies estimating the impact of NPIs on COVID-19 epidemiology.

Compartmental models vs. individual-based models

Another family of epidemiological models is individual-based (or agent-based) models [2]. These methods explicitly model a number of individuals as separate entities, whose traits and behaviours are drawn in a probabilistic fashion from distributions. These individuals can interact with each other within the model, for example disease transmission from an infectious individual to a susceptible one [74].

At small population sizes, stochastic effects can have large influence on outcomes, such as the survival or extinction of an epidemic [75]. Implementing an individual-based model can therefore be advantageous to capture this uncertainty. However, this comes at a computational cost, particularly when there are large numbers of parameters [76]. In cases where population sizes are large, and where there is a complex model with many parameters, the population-level compartmental model may be a preferable choice in situations of limited computing power [77].

1.1.2.2 Phylodynamic inference

Reconstruction of evolutionary history has been carried out by scientists for around 150 years. In the mid 19th century, the German zoologist Ernst Haeckel coined the term *phylogeny* [78], a description of the historical relationship and shared ancestry of living creatures. The field of phylogenetics has since been developed to systematise this process, using various algorithms and data sets to reconstruct the common ancestors of animals, plants, bacteria and viruses [79].

Although this has traditionally relied on morphological characteristics, the advent of genetic sequencing has enabled the production of more detailed and reliable phylogenies [80]. The accompanying increase in high powered computing has led to more sophisticated methods, such as maximum likelihood optimisation, which can exploit the evolutionary information encoded in larger numbers of base pairs within larger numbers of genetic sequences [81].

At the same time, research into population genetics has revealed the effects that population dynamics can have on genes and genomes [82]. Genetic drift, natural selection, population growth and shrinkage all leave signature imprints on phylogenies, and various models and frameworks have been designed to disentangle and describe these relationships [83].

One such model is the coalescent, mostly attributed to Kingman in the 1980s [84], which sought to understand the historical evolutionary relationships of sampled individuals (or

alleles) as a stochastic random process going backwards in time. Under this framework, alleles coalesce together to form hypothetical ancestral copies in the past. Repeating this process eventually leads to the most recent common ancestor of the whole sample. In the classic coalescent model, the time to coalescence - i.e. how long it takes for two randomly sampled individuals to "merge" into their common ancestor - is geometrically distributed [85]. Time to coalescence can be approximated by an exponential distribution at large population sizes [85]. Some common assumptions of the coalescent model include no natural selection, genetic recombination, population structure or changing population size; various extensions of the model have been explored which relax some or all of these assumptions [86]. One important feature of this framework is the establishment of the relationship between total population size and time to coalescence of a random sample of individuals.

With this relationship in mind, it is possible to infer population size from the time to coalescence of sampled individuals. If these individuals are pathogens sampled from infected patients during an epidemic, a viral phylogeny can be constructed and the coalescent framework can be co-opted to infer viral population size [87]. The exploitation of genetic sequence data for the inference of epidemiological parameters and processes is known as phylodynamic inference [88]. Some common assumptions of viral phylodynamic inference include random sampling, a bottleneck at transmission, negligible diversity within-host compared to between-hosts, and no sampling of ancestors [87–90].

Non-parameteric phylodynamic inference

The simplest coalescent-based phylodynamic models assume a constant or exponentially growing viral population, and the parameters of these models can be inferred from the phylogeny. However, the phylodynamic framework has been expanded to be able to estimate a more detailed description of complex real life epidemiological trajectories. One of the first of these was the *skyline* model which allowed population size to be estimated at the time points of every coalescent event [91]. Given that *skyline* assumed no prior population dynamic function, it is classified as a 'non-parametric' model. The signature pattern of the effective population size (N_e) trajectory resembled a city skyline, hence its namesake, but many deemed this too noisy and overfitted [92]. The model has been modified in order to compensate for such overfitting, with one example being the *skyride* model [93], which assumes the existence of autocorrelation between effective population sizes, and uses a Gaussian Markov random field (GMRF) to smooth the trajectory.

There has also been interest in exploring the association of epidemiological parameters with relevant non-genetic factors ('covariates'), which may include disease incidence or prevalence in the case of HIV [94], or geographic structure and ecology for influenza [95]. The *skygrid* model was developed in order to estimate N_e from multiple loci simultaneously [96], and the model was later extended to incorporate covariate data, by affecting the smoothing with a GMRF [94]. This model divides the epidemic into grids, creating a piece-wise population dynamic function. These grid points are flexible, and can be chosen to correspond to observations in the covariate data. The inclusion of covariates into phylodynamic analyses can shed light on external factors associated with viral spread [96].

Inspired by the division of the epidemic into flexibly defined grid points, and the inclusion of covariates, Volz and Didelot developed the *skygrowth* model [97]. The authors observed that effective population size, whose increments are modelled as a stochastic process (such as Brownian Motion) with stationary increments, is strongly determined by the choice of demographic prior. Because the Brownian Motion prior will penalise changes in effective population size, it will bias it towards being constant, even if the true underlying population is exponentially growing [98]. This is especially true when genetic data are relatively uninformative, such as during the period near the most recent sample. The key difference of *skygrowth* is the assumption that the stochastic process is defined not on effective population size, but rather on its growth rate. As a result, as genealogical data becomes uninformative in the recent past, the growth rate estimated in the previous period is likely to remain unchanged, and therefore the underlying exponential dynamics are continued. In analysing a rabies virus genetic data set, Volz and Didelot manage to avoid the constant population size in the recent past estimated by previous models, instead estimating a decline, a result much more consistent with independent surveillance data from the Centers for Disease Control [97].

Although non-parameteric models such as those discussed offer significantly more flexibility than parameteric alternatives where the demographic function is pre-specified, nevertheless the demographic prior has a very strong influence on estimated epidemic trajectories. For example, both *skygrid* and *skygrowth* define different stochastic processes *a priori*, and as a result predict different estimates of effective population size through time. The smoothing parameter (τ), which strongly affects the degree of penalisation of changing values in the GMRF stochastic process (whether on population size or growth rate), is of particular importance. A prior distribution as uninformative as possible is normally assumed for τ, however results can be sensitive to this arbitrary choice [93], highlighting the importance of this parameter.

As a result, Didelot and Volz designed the *mlesky* model, a frequentist approach to phylodynamic modelling [99]. They implemented a Markov Chain where deviation from the previous element is penalised according to a GMRF, but there is flexibility as to whether this applies to effective population size, growth rate, or rate of change of growth rate. Importantly, for the smoothing process, the smoothing parameter τ is selected using a cross-validation method, rather than specified *a priori*. As in previous models, covariates can be incorporated into the fitting framework. Further, the number of grid points is chosen using an AIC model comparison criterion. Reproduction number through time ($R(t)$) can be estimated from the growth rate and serial interval.

The authors argued that the *mlesky* approach is advantageous when it is difficult to establish non-arbitrary priors for smoothing and for grid resolution. They showed success when using simulated data, as well as real data from HIV and cholera epidemics. The *mlesky* model, like *skygrowth*, is able to handle fitting to phylogenetic trees with thousands of tips within a reasonable time frame. Both models have therefore been implemented for the modelling of the SARS-CoV-2 epidemic in New York and England respectively [100, 101].

Model-based phylodynamic inference

An alternative methodological approach to inferring population dynamics from a tree is possible, where the demographic function is specified *a priori*. This group of methods is called parametric or 'model-based' phylodynamics. The simplest case would be where a fixed value of N_e is the only parameter estimated from the tree. An alternative model is that of exponential growth, where there are two parameters: initial N_e and growth rate. The topology of the phylogenetic tree may indicate the most appropriate demographic model to represent the epidemic [88].

However, as discussed earlier, infectious disease dynamics are usually complex and non-linear. The 'SIR' model assumes that the rate of infection depends on the number of infectious individuals; this rate is analogous to the birth rate in population genetics models. For such non-linear models, however, understanding the relationship between N_e and disease prevalence is not trivial [102]. Volz et al. proposed a coalescent rate ($\lambda_n(t)$) assuming the SIR demographic model [90], shown below:

$$\lambda_n(t) = \binom{n}{2} \frac{2\beta S(t)}{I(t)} \tag{1.1}$$

Where n is the number of extant lineages, β is the transmission rate, and $S(t)$ and $I(t)$ are the numbers of susceptible and infected individuals through time.

This result is similar to the analytical result from the Kingman coalescent assuming a standard Wright-Fisher model with discrete, nonoverlapping generations [103]:

$$\lambda_n(t) = \binom{n}{2} \frac{1}{N(t)} \tag{1.2}$$

However, the key difference is the number of susceptibles in the numerator in Equation 1.1. Estimates of N_e using the Kingman coalescent are only proportional if S is invariable; in such cases $\binom{n}{2} 2\beta S(t)$ collapses into a constant. During the early phase of exponential growth, where approximately 100% of the population is susceptible and decreases very slowly, this is a reasonable assumption. However, this assumption becomes untenable as the epidemic progresses, or if the pathogen becomes endemic.

Building on the coalescent rate which takes into account varying birth rates (Equation 1.1), Volz proposed a model which can reflect different demographic structures [104], wherein coalescence can take place within and between states. The rate of coalescence can therefore vary by state and time. The likelihood of a tree given a complex population model is given, allowing the incorporation of prior information about epidemiological parameters such as population size or birth rates. Importantly, the 'true' number of infections can be estimated with relatively little bias.

This flexible framework was implemented in BEAST2 under the name *PhyDyn*, where the parameters for the tree and epidemiological model can be co-estimated [105]. The population structure is specified *a priori* by the user. It can therefore be used to represent different mixing patterns of infected hosts, such as risk groups for HIV transmission (e.g. men who have sex with men vs. heterosexual men and women) [106]. Alternatively, it can reflect natural history of infection such as for Ebola virus (with an 'exposed' but not infectious compartment) [105]. Finally it has been implemented for SARS-CoV-2, where individuals can move from the exposed compartment to two infectious compartments, accounting for high dispersion in secondary transmissions [107, 108].

Data requirements for phylodynamic inference

Just as with dynamic transmission models, phylodynamic models are also fitted to data. One difference, however, is that phylodynamic models are fitted to genomic data, usually summarised in one or several phylogenetic trees.

The reliability of the phylogenetic tree, and therefore of epidemic representation after fitting, depends on the quality of the alignment of genomic samples. Therefore, the greater the number of sequences, the more bootstrapping (or other cross-validation) can be performed on the tree, and the greater the confidence in the topology and branch lengths of the tree [109]. Coalescent models assume that the sample was randomly selected from the wider population, but there are many factors that could bias sampling. For example, viral genomes extracted from hospital patients could bias towards certain demographics or virulent lineages [110]. Another departure from randomness is a sample generated from contact tracing [111], as more closely related viruses may be disproportionately likely to be sampled, and the genetic diversity of the wider population might be underestimated [112]. The ideal genomic sample for phylodynamic modelling will therefore be as random as possible from the community. Further, the genomic sample must be dated in order estimate the molecular clock [113], and to infer dates of ancestral nodes, from which the coalescent model can estimate epidemiological parameters.

Certain natural history data may be necessary to parameterise model-based phylodynamic models [105]; using a mechanistic model will require the same parameters as discussed in the dynamic transmission model section. Non-parametric models may also require parameterisation of the serial interval in order to estimate reproduction number [114]. Finally, if exploring the association of covariates with viral growth rates, as is possible with the *skygrid*, *skygrowth* and *mlesky* models, data on covariates are required [94, 97, 99].

1.2 HIV

1.2.1 Natural history

The Human Immunodeficiency Virus (HIV) was first isolated in the Pasteur Insitute in Paris in 1983 [115], and was given its name in 1986 by the International Committee on Taxonomy of Viruses [116]. Phylogenetic analysis has confirmed that HIV-1, globally the most common subtype, is a retrovirus related to the Simian Immunodeficiency Virus prevalent among chimpanzees in Western and Central Africa [117]. A spillover event from other primates to humans is therefore thought to be the origin of HIV [118].

The viral genome is coded within two single strands of RNA, held within a conical capsid [119]. Each strand has a length of just under 10,000 base pairs, organised into 9 genes coding for 15 proteins [120]. Among the genes are *gag*, *pol* and *env*, responsible for the inner shell, reproduction and infection of the virus respectively [121].

Viral infection is achieved primarily through sexual intercourse, but also occurs via childbirth, breastfeeding, and injecting drug use [122]. A virion, an entire virus particle with genome and protein shell, will attach to microphages and CD4+ T-cells, involved with the body's immune system [123]. The RNA will then enter the cell, and with the help of certain enzymes such as reverse transcriptase and integrase, will be spliced into the host's chromosomes [77]. Viral proteins will then be translated through the host ribosomes, which will lead to the assembly and release of more viral particles to infect other cells. A cell infected with HIV will die; the depletion of CD4+ T-cells is what leaves HIV-infected patients highly vulnerable to opportunistic infections that would otherwise be easily managed [124].

The process of viral integration into the host chromosomes is highly error-prone, with frequent recombination between the two strands of RNA [125]. It is believed that the fast within-host evolutionary rate stems from this, causing the rapid development of drug resistance to antiretroviral monotherapy [126]. Modern antiretroviral therapy is therefore a combination of more than one drug, usually targetting different parts of the viral life cycle [127].

1.2.2 Epidemiology

1.2.2.1 Global epidemiology

HIV continues to be a global public health threat [128]. According to UNAIDS, 37.7 million people were living with HIV in 2020 [129]. UNAIDS also estimated that 1.5 million people are newly infected each year, with 870,000 (58%) of these in sub-Saharan Africa [129]. Globally, HIV infections have decreased by 31% since 2010 [130]. A total of 680,000 people died in 2020 from AIDS-related illnesses, denoting a decline of 47% compared to 2010 [130]. In 2016, UNAIDS teamed up with partners to end AIDS as a public health threat, which they defined as a reduction of 90% in new HIV infections and HIV-related deaths by 2030, compared to 2010 [131].

Globally, an estimated 27.5 million people are on antiretroviral treatment, representing 73% of those living with HIV [130]. ART coverage in Sub-Saharan Africa is estimated to be close to the global average, at around 76% of people living with HIV [130] in 2020.

1.2.2.2 Key populations

According to UNAIDS, key populations are defined as subgroups who are key to the epidemic's dynamics or to its response [132]. According to WHO, they are defined as groups who are at increased risk of HIV, often with intersecting legal and social issues [133]. This increased risk is attributed to specific high-risk behaviours and unmet needs in HIV prevention and treatment ('prevention gaps') [134]. The main five key populations as of the 2016 WHO update include 1) men who have sex with men, 2) people who inject drugs, 3) people in prisons, 4) sex workers and 5) transgender people [133].

In most countries in Central Asia, Europe and North America, over 90% of new HIV infections in 2014 were among people in key populations [135]. In the regions of Asia and the Pacific, Latin America and the Caribbean, they accounted for two thirds of newly acquired HIV infections [135]. In Sub-Saharan Africa where HIV can also be prevalent in the wider population, key populations nevertheless account for over 20% of new HIV infections [135]. It is argued that key populations are integral to the establishment of an epidemic, and continue to contribute to its persistence [136, 137].

Female sex workers

Among the most at risk of HIV infection are female sex workers (FSW), defined as women who exchange sex for money or other goods [132]. They vary in terms of visibility, the organisation of their workplace (street-based, brothel-based), and the proportion of their income coming from sex work, all depending on local social, historical, economic and legal factors [138, 139]. These, along with violence, stigma and other structural factors contribute to varying levels of risk of HIV infection [140]. The proportion of individuals engaged in sex work is estimated to be up to 4.3% in Sub-Saharan Africa [138], where the bulk of HIV infections are found [135]. HIV prevalence is estimated to be 30.7% among female sex workers from a pooled estimate of several countries with generalised epidemics in 2007 - 2011 [141], mostly in countries in Sub-Sarahan Africa, but also in Asian and Carribean countries. The WHO estimates in 2016 a global HIV prevalence of 12% among female sex workers [133].

Contribution of FSW to HIV epidemics

One common method for identifying epidemiological associations between exposures and outcomes is the population attributable fraction (PAF), defined as the fraction of cases in a population that is attributable to a specific exposure, or the fraction of cases that would not have occurred had this exposure not existed [142]. The PAF of a key population may be calculated as the fraction of cases that would have been avoided were there no transmission from this group [143]. Commonly, dynamic transmission modelling is employed to estimate this [59].

As discussed earlier, dynamic transmission modelling takes into account the chain of transmission. It has been shown that even in epidemics where high HIV prevalence is also found in the wider population, epidemics can be driven by high risk behaviours in key populations [144]. To calculate the number of infections attributable to FSW, dynamic transmission models are capable of accounting for infections in their clients, as well as onward transmissions to the partners of clients and into the wider population [50].

There have been several estimates of the PAF of FSW for HIV. In South India, from the beginning of the epidemic to 2008, over 90% of incident HIV infections in men and 10-20% in women were attributable to commercial sex either directly or indirectly through bridging infections to the wider population from clients [145]. In Benin, it is estimated that commercial sex has contributed to 93% (90% credible interval (CrI): 84–98%) of all cumulative infections in heterosexual populations from 1985 – 1993, and 52% (90% CrI: 33–71%) from 1993 – 2008, including direct and secondary transmissions [50]. In Côte d'Ivoire, FSWs contributed 95% (95% CrI: 91–97%) of all new infections from 1976 – 1985, and 19% (95% CrI: 11–27%) of infections from 2005 – 2015 [48].

Due to their disproportionate burden and contribution to HIV infection, HIV prevention interventions have often been focused on FSW to maximise effectiveness in the long term [146].

1.2.3 HIV prevention interventions

In this section I introduce recent forms of bio-medical HIV prevention interventions. Moreover, I discuss available information regarding their feasibility among FSW in local settings, including adherence and retention.

1.2.3.1 Treatment as prevention (TasP)

Treatment as prevention (TasP) is the early provision of ART to HIV-infected patients with the intent to lower viral load, and therefore onward transmission, as soon as possible [147]. The WHO and UNAIDS have both included TasP as a recommended HIV prevention tool to policymakers, following a successful trial (HPTN-052) which demonstrated a 96% reduction in onward HIV transmission from those receiving early ART [147]. However, large real-world community-based randomised trials (CBRT) of TasP in Sub-Saharan Africa have returned mixed results [148]. The 'TasP study' (ANRS 12249), was a cluster-randomised trial aiming to estimate the impact of TasP in KwaZulu-Natal, South Africa, where HIV prevalence was around 19% among individuals between 15-49 years of age in the low-risk population [149]. In this rural setting, they found that fewer than 40% of HIV-infected individuals were able to access an HIV clinic within 3 months of referral. There was no statistically significant difference in HIV incidence between the group assigned TasP and the one assigned to the national guidelines (eligibility at CD4 <350 cells/mm^3) [150].

Another CBRT in South Africa and Zambia, the PopART study (HPTN 071), had three arms [151]. The first (A) offered a full package, including ART initiation at any CD4 T-cell count, HIV testing and a variety of healthcare access services. The second (B) provided the same package except for immediate eligibility. Eligibility for ART was at national guidelines, which was initially at CD4 <350 cells/mm^3, changing half way through the study to universal ART, which rendered arm A and B identical. The control arm (C) provided the standard of care according to the Ministries of Health of the two countries. The final results were counter-intuitive: no difference in HIV incidence was observed between arms A and C. However, arm B had a significantly lower incidence (around 30%) than both of the others, including arm A which had earlier ART initiation than arm B [152]. The authors offer various reasons for this surprising result, including unrecognised demographic differences at baseline, migration, and random chance [152].

Two more notable CBRT also provided mixed results. The 'SEARCH' trial, carried out in rural Uganda and Kenya, showed no effect of TasP on HIV incidence [153]. The 'Ya Tsie' trial in rural and peri-urban communities in Botswana found TasP reduced HIV incidence by 31%, but this result was not statistically significant at a threshold of 5% [154]. The success of real-life delivery of TasP is therefore uncertain, particularly in rural African communities with poor access to HIV testing and linkage to care [148].

1.2.3.2 Pre-exposure prophylaxis (PrEP)

Pre-exposure prophylaxis (PrEP) is the practice of administering antiretroviral drugs to HIV-uninfected individuals, lowering the risk of HIV infection at times of HIV exposure [155]. The WHO have recommended that PrEP be considered for anyone with substantial risk of HIV infection [156], particularly those who struggle to maintain consistent condom use. The different forms in which PrEP can be delivered are detailed in the following sections.

Oral PrEP

Oral PrEP involves the provision of tablets containing active antiretroviral drugs, commonly tenofovir in combination with emtricitabine [157]. Clinical trials of oral PrEP have shown a large variation in individual level effectiveness between sub-populations. Among men who have sex with men (MSM), the results have been encouraging, showing significant reduction in HIV incidence due to PrEP [158, 155, 159], with trial effectiveness nearing 90%. Among women however, the results are more mixed, with two trials discontinued for lack of effectiveness [160, 161]. This was attributed to poor adherence, sparking a stronger focus on monitoring adherence in subsequent trials. The Partners PrEP trial for discordant couples showed that PrEP can achieve high levels of efficacy with good adherence in both women and men [162], suggesting the success of PrEP hinges on behaviour, not physiology. However, the uptake of the protective molecules seems more efficient in rectal than vaginal tissues, meaning that good protection may require better adherence for women than for men [163].

Several demonstration projects testing the real-life effectiveness of PrEP for FSW in Sub-Saharan Africa have been completed. In Kenya, the feasibility of daily oral PrEP was investigated for FSW, MSM and young women [164]. They found that retention was poor for all groups, including for FSW, whose retention after one, three and six months was 40%, 26% and 14% respectively. The most commonly cited reasons for drop out were side effects, self-perception of risk, a virally suppressed sexual partner, stigma and myths about PrEP.

In South Africa, the TAPS study was carried out, evaluating the feasibility of PrEP and TasP among FSW [165]. They also found poor retention, with only 22% of enrolled FSW presenting to clinics after one year. Although reasons for discontinuation were not recorded, the authors speculate that it could be due to the establishment of a national PrEP programme for FSW during the time of their study. Self-reported PrEP adherence was between 70-

85%, but the authors believe this could be subject to a bias; no objective PrEP adherence measurements were made.

In another setting with high HIV prevalence, similarly poor retention among FSW was observed. The PrEP demonstration study in fishing villages in Uganda, where HIV prevalence was around 40% [166], found that FSW were 42% more likely to drop out from the study than women in HIV-discordant couples. For all women in this rural setting, median retention on PrEP was only 45 days [167], but adherence was not measured.

Conversely, not all settings have poor PrEP retention among FSW. In Senegal, a demonstration study found that 73% of FSW recruited were retained after 12 months [168]. The authors suggest that high retention overall was due to the use of Ministry of Health clinics. This good retention was complemented by no observed HIV infections among all participants, even though 4 infections were expected given the baseline incidence rate of 14 cases per 1000 person-years. High mobility was cited as the main reason for those that did drop out. The authors measured PrEP adherence using electronic recordings of each time the pill container was opened ("MEMS cap"), and using blood plasma concentration samples. With the electronic caps, daily adherence started off at 80%, but soon dropped to 50% then rose to 65% where it stabilised. For blood plasma concentrations, the active drug tenofovir was detected in 52% of samples (>0.3 ng/ml); 32% of samples had a concentration consistent with daily adherence (>35.5 ng/ml).

Finally, a demonstration study of daily oral PrEP was carried out among FSW in Cotonou, Benin [169]. Retention after two years was 43%, and tenofovir was detected in 53% of visits. More details about this demonstration study are available in Chapter 2, and the study impact is explored in Chapter 3.

Long-acting PrEP

More recently, injectable options for PrEP are moving towards regulatory approval [170]. Long-acting injectable cabotegravir (CAB-LA) is an injection of antiretroviral therapy into the gluteal muscles every eight weeks [171]. There have been clinical trials for CAB-LA, among MSM and transgender women (HPTN-083) and among heterosexual women (HPTN-084). The HPTN-084 trial was carried out in seven Sub-Saharan African countries, and showed CAB-LA to be 89% (95% CI 68-96%) more effective than daily oral PrEP, as the challenge of daily adherence to pills was alleviated [172].

As a result of these successful trials, CAB-LA was awarded Breakthrough Therapy designation by the United States Food and Drug Administration (FDA) [171]. Following full regulatory approval, it is hoped that long-acting PrEP may help alleviate some of the issues surrounding adherence, particularly for high-risk women [173]. Real-world implementation studies of CAB-LA are yet to be conducted [174].

1.3 SARS-CoV-2

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is a single stranded RNA virus that causes the coronavirus disease 2019 (COVID-19) [8]. It emerged in late 2019 in Wuhan, Hubei Province, China, and has since spread throughout the world [175]. Initially called 'nCoV-2019', phylogenetic analysis found that it is related to the virus causing the first SARS epidemic in 2002-2004 [176], and is thought to have emerged through a zoonotic spillover event from bats, due to its genetic similarities to bat viruses found near Wuhan [177].

1.3.1 Natural history

SARS-CoV-2 has four structural proteins. Spike (S), envelope (E) and membrane (M) form the viral envelope, while nucleocapsid (N) binds to and encapsulates the viral RNA genome [178]. The genome contains around 30,000 base pairs [179].

The spike protein has strong affinity to the angiotensin converting enzyme 2 (ACE2), an enzyme attached to cell membranes found throughout the body [180]. Once a virion has attached to ACE2, the spike protein is broken open and the virion can enter the cell protected by an endosome [181]. The endosome will break, and the virion will release its genome into the cell where it will replicate [182], producing many more virions for further infection of other cells [183].

When the virus has spread to enough cells, the disease COVID-19 can occur. The main symptoms are a dry continuous cough, fever, headache, fatigue, breathlessness and an inability to taste or smell [184]. Symptom onset is estimated to be between one and fourteen days after exposure [185], but is typically around 5-6 days [186, 187]. The severity of symptoms is strong associated with age [188]; older patients are significantly more likely to develop serious or fatal illness [189]. A certain proportion of individuals do not notice any

symptoms, although it is believed asymptomatic cases can still transmit the virus to others [190].

The infection fatality ratio (IFR), i.e. the proportion of all infected individuals that die from COVID-19, is estimated at 0.004% for individuals aged 0-34 years old, but rises to 28% for individuals over 85 years [191]. Other risk factors for mortality include male sex [192], ethnic minority status [193], and comorbidities such as diabetes, hypertension and asthma [194]. Where medical infrastructure is poor or overwhelmed with cases, mortality from COVID-19 may also increase [195]. For all of these demographic and economic reasons, there are different observed IFRs for COVID-19 across the world [175].

Case fatality ratio (CFR) is the proportion of cases that result in death; this number tends to be higher than IFR as the denominator will exclude asymptotic and unreported cases. Global CFR is estimated to be 2.10% as of 17th August 2021 [196].

As of 17th August 2021, there is no effective cure for COVID-19, and treatment for severe cases usually involves management of symptoms until the virus is cleared by the body [197], although neutralising monoclonal antibodies and other antivirals have been found to improve clinical outcomes [198]. For those that have trouble breathing, oxygen support and prone positioning is recommended, and ventilation and dexamethasone for patients with very low oxygen levels [199]. As of 17th August 2021, several effective vaccines are available which greatly reduce morbidity and mortality from COVID-19, although efficacy against infection and transmission seems to be more modest [200–202].

It is believed that COVID-19 is spread through droplets and aerosols, released by infected individuals while they exhale from breathing, talking, singing, coughing and sneezing [203]. New infections occur mainly when infected particles are inhaled, and less commonly when they come into contact with the eyes, nose or mouth [204]. It is important to note that individuals infected with SARS-CoV-2 are able to transmit to others prior to symptom onset, which has implications for effective epidemic control [205].

1.3.2 Epidemiology

Since its emergence in Wuhan in December 2019, SARS-CoV-2 quickly spread throughout Hubei Province [206]. In January 2020, cases appeared throughout China following movement for the Chinese New Year [207]. China implemented restrictions in late January and early February across the country to contain the virus, most notably the complete quarantine of Wuhan on 23rd January 2020 [208].

By the end of January 2020, cases had also been identified outside of China, including in Japan, South Korea, United States, Singapore, Thailand, France and Italy [209]. Italy overtook China with the highest number of COVID-related deaths in March 2020, as it swept through Lombardy in the north of the country. The WHO declared COVID-19 a pandemic on 11th March 2020 [210].

As of 17th August 2021, there have been over 200 million confirmed cases of COVID-19 globally, and over 4 million COVID-related deaths; over 187 million confirmed cases have led to a recovery. The COVID-19 pandemic has spread to every country in the world [175].

Comparing the epidemiological situation between countries is problematic for many reasons. Firstly, the definition of a case has varied by country and across time [211]. Secondly, the availability of tests for SARS-CoV-2 is not homogeneous, with developed nations being more capable of implementing mass testing [212]. Thirdly, demographic differences between countries, particularly age distributions, affects the proportion of symptomatic cases (therefore reporting frequency), severity of illness (therefore hospitalisations), and mortality [213]. Finally, there are striking differences in the types and levels of restrictions implemented across the world, from very little intervention to very severe restrictions [214].

Nevertheless, as of 17th August 2021, the United States is the country with the highest number of confirmed cases and COVID-related deaths at 37 million and 600,000 respectively. The highest number of COVID-related deaths per capita is in Peru, at over 600 deaths per 100,000. Peru also has the highest case fatality ratio at 9.3% [196].

Many European countries have also suffered substantial mortality due to COVID-19. As of 17th August 2021, many countries in Eastern Europe have high deaths per capita, notably Hungary, Bosnia and Herzegovina, Czechia, North Macedonia and Bulgaria [196]. In Western Europe the situation is less severe, although high COVID-related mortality is observed particularly in Belgium and Italy. More detail on the COVID-19 epidemic in England is available in Chapter 6.

Comparatively few cases and deaths have been confirmed in Sub-Saharan Africa, particularly in the first few months, attributed to underreporting [215, 216]. It is likely that the age distribution in this continent, with a much younger population compared to developed nations, contributes to lower overall mortality [217]. More detail on the COVID-19 epidemic in Benin is available in Chapter 4.

Since the implementation of COVID-19 vaccines at the end of 2020, vaccine coverage has increased globally, but generally at a higher pace in more developed countries. As of 17th August 2021, nearly 2.5 billion people (31.7%) have had at least one dose of a COVID-19

vaccine [175]. Some countries, including Malta, United Arab Emirates and Iceland, have achieved over 80% coverage [175].

1.3.2.1 Evolution

Genomic data for SARS-CoV-2 has been collected since December 2019, when the virus was first identified [218]. Since then, the virus has been sequenced across the world, mostly deposited on the GISAID database [179]. SARS-CoV-2 quickly became the most sequenced virus in history [219], as the sequencing efforts accelerated in the UK, US and elsewhere. As of 17th August 2021, over 2.8 million individual isolates have been uploaded to the GISAID database [179].

The vast amount of sequence data has enabled close monitoring of viral evolution. Early phylogenetic studies estimated a substitution rate of $1.19-1.31\times10^{-3}$ and $1.5-3.3\times10^{-3}$ per site per year respectively [220, 221], although later studies estimate an approximately two-fold decrease in substitution rate over time to 8.99×10^{-4} [222]; similar substitution rates are observed in other related coronaviruses [223]. There is evidence of a number of recombination events in the evolutionary history of the SARS-CoV-2 genome [224].

Since spreading from China, the evolutionary and epidemiological patterns of SARS-CoV-2 are observable on the same timescale [225]. There have been multiple proposals for naming and classifying SARS-CoV-2 lineages. One is the 'Pango' nomenclature system which uses a phylogenetic framework [226], but whose labelling depends on number of samples, geographical location and genetic difference [227]. Crucially, the names of lineages are updated according to the dynamics of the pandemic; the entire evolutionary history of the lineage is therefore not captured in the label, which would get far too complicated over a short time frame.

This dynamic phylogenetic framework can identify lineages which are emerging and growing. A lineage which is growing especially fast, coupled with any evidence of increased virulence, immune escape or resistant to interventions, is designated a variant of concern by Public Health England (PHE) [228]. This nomenclature is useful to policymakers in England and elsewhere, whose recommendations for epidemic control are influenced by ongoing viral evolution.

One prominent variant of concern was B.1.1.7 (later also termed the 'alpha variant'), discovered towards the end of 2020 [101]. This lineage was estimated to spread faster than previously circulating lineages. A detailed background on this lineage is available in Section 6.2.2.1, and an analysis exploring its increased transmissibility is described in Section 6.4.4.

Monitoring evolution of SARS-CoV-2 is therefore of great importance for understanding epidemiological patterns through time.

1.3.3 Interventions

In this section I introduce interventions implemented by various governments to control the spread of COVID-19. In Chapter 6, I discuss previous mathematical modelling studies exploring the impact of non-pharmaceutical interventions (NPIs).

There have been numerous NPIs implemented since the beginning of the outbreak to control its spread. Beginning in China, a cordon sanitaire was placed around Wuhan on 23rd January 2020, followed by travel restrictions to and from other cities in the region and elsewhere [229]. They further implemented diagnosis, case isolation and contact tracing, particularly of travellers from Wuhan [230]. Finally, there was a suite of restrictions aiming to reduce social contact, including school closures, stay at home orders, cancellation of public events and gatherings, closure of workplaces and other non-essential buildings [231]. In Chapter 5 I introduce the NPIs implemented in Weifang, a city in Shandong Province, China.

Similar measures were implemented in Europe, with most countries ordering their public to stay at home during March 2020 [232]. In some countries there was a greater emphasis on contact tracing, particularly in Germany [233], while countries in Southern Europe relied on a stricter lockdown with curfews. In France, Spain and Italy, the public were required to justify the necessity before leaving the house during lockdown, enforced by the police [234–236]. In Sweden, although gatherings over 500 people were banned, the stay at home orders were not mandatory, but advisory [237]. In Chapter 6 I introduce the NPIs implemented in England, both locally and nationally.

In Sub-Saharan Africa, there was even more stark diversity in the COVID-19 response. Although many countries implemented lockdowns early, sometimes before the appearance of any confirmed cases [214], extensive and thorough restrictions were impossible to maintain in most of them. Mostly, this is because a large proportion of the workforce is in informal work (e.g. open air traditional markets) [238]. Further, the poor economic situation in most African countries prevents large amounts of borrowing, spending, and economic relief, and accepting the opportunity costs of closing businesses [239]. In Chapter 4 I discuss in detail the non-pharmaceutical interventions implemented in the country of Benin in West Africa.

1.4 Summary

In this chapter, I have introduced the background material required to understand the context of the research in this thesis, including the modelling of infectious diseases, and the pathogens HIV and SARS-CoV-2. However, the analyses in this thesis are diverse in terms of metholodological approach and disease context. There are therefore further detailed introductions at the beginning of each new chapter, to give further necessary background for each analysis.

Chapter 2

Data summary and transmission model development for HIV in Cotonou

2.1 Overview

In the first main section of this chapter, I introduce the local context of Cotonou: the demographic and epidemiological patterns specific to HIV in the city, and its history of HIV interventions beginning from the early 1990s. I further explore previous research studies which have modelled the HIV epidemic in Cotonou.

Next, I begin the process of answering my research questions by reviewing and summarising available data in Cotonou, as the analytical and predictive power of any mathematical model depends on the data used to parameterise and fit it.

In the next main section of this chapter, I describe the mathematical details of the model structure.

Finally, I describe the model fitting process, involving a sample from a Latin Hypercube and retaining parameter sets which meet the fitting criteria. The fitted model described at the end of this chapter is used for the analyses in Chapter 3 and Chapter 4.

The content in this chapter is published with the following citation, mostly in the supplementary materials.

Geidelberg, L., Mitchell, K. M., Alary, M., Mboup, A., Béhanzin, L., Guédou, F., ... & Boily, M. C. (2021). Mathematical Model Impact Analysis of a Real-Life Pre-Exposure

Prophylaxis and Treatment-As-Prevention Study Among Female Sex Workers in Cotonou, Benin. *Journal of Acquired Immune Deficiency Syndromes*, 86(2), e28.

2.2 Introduction

2.2.1 The city of Cotonou and the Republic of Benin

The Republic of Benin (hereafter referred to as 'Benin') is a French speaking country in West Africa [240]. The southern tip of Benin forms its coastline, a narrow stretch of land called the Bight of Benin, home to the majority of its population of around 12 million (2018) [241]. The rest of the country stretches inland, bordering Togo to the west, Burkina Faso to the north-west, Niger to the north-east and Nigeria to the east [240]. The main ethnic groups of Benin are Fon (38%), Adja & Mina (15%) and Yoruba (12%), and the main religions are Roman Catholicism (52%), Islam (25%) and traditional local faiths (18%) [242–245].

Although Porto-Novo is the capital city of Benin, the nation's seat of government is located in its largest city and economic hub, Cotonou [246]. The population of Cotonou has increased dramatically in the past decades not only from its growing population, but also because of its expanding urban sprawl engulfing previously distinct towns [247]. The city of Cotonou had a population of 320,000 according to the census in 1979, rising to 679,000 in 2013, reflecting a growth rate of around 5% per year [242–245]. However, the recent inclusion of surrounding areas such as Abomey-Calavi and Seme-Kpodji into 'Greater Cotonou' resulted in a total population estimate of over 1.5 million in 2013 [245].

The Autonomous Port of Cotonou generates over 60% of the GDP of Benin, and is responsible for 90% of its international trade [248]. Because of the various trade links to neighbouring countries, Cotonou has had a significant amount of migration in the past several decades, mostly from other African nations, but also from Lebanon [249]. Despite the abundance of trade and promising growth, Benin remains one of the world's poorest countries by most economic indicators [250].

2.2.2 HIV epidemiology in Cotonou

The first confirmed cases of HIV-1 in Benin were detected in blood samples from Cotonou between 1986-1987 [251]. An analysis of 3478 blood donors found seropositivity in 4 individuals (0.12%) [252], although a separate survey in the same year found an HIV

prevalence of 3% (7/215) among female sex workers (FSW) [251], indicating that the beginning of the epidemic might have occurred earlier. The most recent estimate of HIV prevalence across Benin was among pregnant women in 2015, which showed that Cotonou had a greater HIV prevalence than any other region in Benin [253].

FSW have a particularly high HIV prevalence in Cotonou; in 1993, it was estimated at 53% (95% confidence interval: 48-58) [254]. In contrast, the national AIDS programme found in the same year that among pregnant women, HIV prevalence was estimated at only 0.5% [255]. In 2017, HIV prevalence had declined to 10% (7-13) among FSW in Cotonou [256].

FSW in Cotonou can be classified into two categories. The first is 'professional FSW' (pFSW), which describes women for whom sex work is their main source of income [256, 257]. These women tend to work in brothels or on the street, exchanging sex with clients for money, gifts and favours [256]. The second type, 'part-time FSW' or 'bargirls', is a broad category of women for whom sex work does not provide the majority of their income [258]. These women tend to work in bars, clubs or restaurants where they earn a wage, but also where they may find clients [259]. Professional FSW have significantly higher numbers of commercial sexual partners, and consequently have a much higher HIV prevalence than part-time FSW [256, 259].

As a group, professional FSW in Cotonou represent a diversity of nationalities. In 1993, 97% of these women were foreign-born, mostly from Ghana, Nigeria and Togo [255]. Although more recent surveys revealed a greater proportion of FSW from Benin (e.g. 55% in 2015) [256, 259], it is possible that Beninese women in earlier years did not reveal their status as FSW due to intense stigma [254]. Many foreign-born FSW are drawn to Cotonou, as the port provides ample commercial activity and therefore many travelling male clients [256, 259]. Further, the more severe legal prohibitions on prostitution in neighbouring countries may have incentivised these women to migrate instead to Benin [260], where prostitution is more generally accepted, albeit still illegal [261]. Many women return to their home countries after a few years of sex work in Cotonou, creating a high turnover of sex workers in the city [256, 259, 255].

Both types of FSW are considered to be 'key populations', i.e. groups whose risk of HIV acquisition and mortality are elevated compared to other groups [133, 132]. Other key populations (since 2017) in Cotonou include i) the male sexual partners of FSW: both their clients and their long-term boyfriends or husbands, ii) men who have sex with men, iii) young people (particularly women), iv) people who inject drugs, v) prisoners, vi) truck drivers and vii) tuberculosis patients [259].

The epidemiological pattern of disproportionate HIV prevalence among certain groups has been commonly termed "concentrated", as opposed to a "generalised epidemic" which would be characterised by high HIV prevalence among low-risk groups (see Chapter 1) [137]. Cotonou has been described as having a concentrated epidemic [137, 50], due to the disproportionate burden and spread among FSW. As a result, there has been much effort over several decades to respond to the HIV epidemic among this group in Cotonou [254, 262, 50], as well as among other key populations [263].

2.2.3 The response to HIV in Cotonou

Ever since the early 1990s, there has been substantial activity responding to the crisis in HIV and STI burden among FSW in Cotonou, in cooperation with the Canadian government and other funders [254, 262]. All HIV prevention activities among FSW from 1993-2008 in Cotonou are collectively known as the "SIDA 1/2/3" project [50]. The response package in 1993 included visits by community leaders and healthcare professionals, both to the homes and workplaces of FSW, for the prevention of HIV and STIs [254]. Further, a dedicated clinic was established to care for FSW, originally offering screening and treatment of STIs, but in subsequent years including treatment for HIV, as antiretroviral therapy (ART) became widely available [169]. Other STIs were also major problems in the early 1990s, such as *Neisseria gonorrhoeae* (gonorrhoea) at 43% prevalence [254]. The imperative to treat these other diseases stemmed not only from their own morbidity, but also from the increased risk of HIV seroconversion with STI co-infection [264].

Further activities included the promotion of condom use, particularly for FSW during commercial sex. In 1993, the FSW surveyed reported that 62% of sex acts with clients in the past week were protected with a condom which, after years of awareness campaigning, rose to 81% in 1998 [254]. By 2015, condom use with clients was estimated between 97-99% [259, 265], although remained low with their long-term non-commercial partners (23-44%). HIV monitoring and prevention were increased among the male sexual partners of FSW in 2000, after it was observed in 1998 that HIV prevalence among clients and boyfriends were 8% and 16% respectively [266]. Male circumcision was not required as an HIV prevention intervention, as over 99% of men in Cotonou were already circumcised for cultural reasons [267]. HIV prevalence and sexual behaviour have been monitored every few years since 1993 among professional FSW, bargirls and their clients, using integrated bio-behavioural surveys (IBBS) [256, 259, 254, 262].

Since 2002, when ART was first introduced in Cotonou [268], a high rate of viral suppression has been one of the key targets for the HIV response [253]. As elsewhere, eligibility for ART has evolved over time. Originally, only those with less than 200 CD4 T-cells per mm^3 of blood were eligible [269]. From 2012, this was expanded to anyone below 350 cells per mm^3 [270], and from 2015 anyone below 500 cells per mm^3 [271]. From 2016 onwards [253, 169], anyone who is HIV-infected became eligible for ART; this policy became common throughout the world [272], since it was discovered that viral suppression was beneficial not only for mortality and morbidity for the HIV-infected patient, but was also itself a form of HIV prevention for their sexual partners [273].

Treating HIV-infected patients as soon as possible, with the intention of preventing onward transmission, will be henceforth called in this thesis "treatment-as-prevention", or "TasP" [147]. PrEP is the practise of administering antiretroviral medications to at-risk HIV-uninfected individuals, such that their risk of HIV acquisition is lessened with the presence of these drugs [155]. Detailed introductions to TasP and PrEP are found in Chapter 1.

2.2.4 The PrEP/TasP study

In late 2014, a two-year demonstration project (henceforth called the "PrEP/TasP study") was conducted to assess the real-life feasibility and usefulness of adding TasP and daily oral PrEP to the existing clinical, behavioural and structural combination prevention package already offered to FSW in Cotonou [169]. The study focused on pFSW from a designated catchment area in Greater Cotonou, representing approximately 28% and 51% of all HIV-uninfected and HIV-infected pFSW respectively in Cotonou and its surroundings [169, 259, 257]. Prior to the study, there were extensive tailored community outreach activities, preparedness and adherence counselling by researchers and community workers who were familiar with this population of women in the past decades of HIV response. In the study, uptake, adherence and retention to PrEP and TasP and self-reported sexual behaviour were measured over two years (2015-2017) at quarterly visits in face-to-face interviews. Further to this, pill counts and biological markers were used to measure adherence every 6 months. For PrEP, the concentration of the active antiviral molecule, tenofovir, was measured in the blood plasma. A low or undetectable concentration (<0.3 ng/ml) suggests poor or no adherence, while a high concentration (over 40 ng/ml) was considered good or daily adherence. For TasP, CD4 T-cell counts and viral load measurements were taken as an indicator of successful adherence. A viral load of <1000 copies/ml and <40 copies/ml represented states of viral suppression and undetectable viral load, respectively.

In the years prior to the study, there was already significant community engagement with FSW in Cotonou to enable access to the sexual health clinic. As such, the population of FSW was well known to the local researchers. They designed a catchment area within Cotonou in which an estimated 600 FSW were located, expecting to reach around 500 of them. They recruited FSW within this catchment area until the desired numbers on the PrEP (approximately 250) and TasP arms (approximately 100) were achieved [169].

Uptake for PrEP and TasP were both encouraging. Of the 290 FSW eligible for PrEP (i.e. those who were HIV-uninfected and not pregnant, breastfeeding or with active Hepatitis B), 256 (88%) were recruited. For TasP, of the 110 who were eligible (i.e. those who were HIV-infected but without HIV-1/2 dual infection, and not on ART), 105 (95.5%) were recruited [169].

Although self-reported adherence to PrEP was initially promising, as 78% of FSW reported daily adherence during the 14 days before the first visit, it declined to 57% by the end of the study. Furthermore, drug concentration measurements suggested that adherence was poorer than what was self-reported by the participants. Over the course of the study, tenofovir was detectable in blood plasma in only 53% of 569 visits by PrEP participants. Indeed, there were two seroconversions during the course of the study among PrEP participants, and both events were associated with poor adherence [274]. This translated to an incidence rate of 0.8/100 person-years in the PrEP arm [275]. For TasP, adherence was better than PrEP: at the final visit 87% of participants had measurable viral suppression [169].

Retention was also a challenge for the study. By the end of the study, only 47% and 59% of PrEP and TasP participants made the final visit. There were many reasons for dropping out of the study, with the most common by far being movement: many of these FSW returned to their town, province or country of origin (50% and 79% of PrEP and TasP dropouts). Other reasons for leaving the study included no longer being interested, becoming pregnant and therefore ineligible (for PrEP), and leaving sex work [274].

In 2015, prior to the PrEP/TasP study, only 24% of FSW living with HIV in Cotonou were on ART. At the end of the study in 2017, ART coverage rose to 83% of FSW [276]. More generally in Benin, only 66% of women and 45% of men living with HIV were estimated to be on ART in 2017 [277], far below UNAIDS 2020 90-90-90 targets (90% of people living with HIV (PLHIV) diagnosed, 90% of diagnosed on ART, and 90% on ART virally suppressed) [278]. The imperfect ART coverage among all populations may prevent Benin achieving the UNAIDS goal of eliminating HIV/AIDS as a global health threat by 2030 [131]; as such a continued focus and improvements in HIV prevention strategies remain important.

2.2.5 Previous studies modelling the HIV epidemic in Cotonou

There have been several studies modelling the epidemiological dynamics and impacts of interventions in Cotonou. In 2002, using a mathematical model, Boily et al. estimated the effectiveness of HIV interventions among key populations (termed 'core groups'), depending on the phase of the HIV epidemic [137]. They modelled an open population of high and low risk sexually active heterosexual adults; high risk females and males were interpreted as FSW and their clients respectively. They integrated sexual mixing patterns which featured a parameter controlling the extent of assortativity between high and low risk groups. They modelled both HIV and STD transmission, with the assumption that STD infection can increase risk of HIV acquisition [264]. The model simulations reflected data from various HIV and STD prevalence estimates among all groups, including from studies on sexually active men and women, antenatal clinics, and the SIDA1/2 project funded by the Canadian International Development Agency.

They estimated that the SIDA1/2 project, by screening 50% of FSW for gonorrhoea every year and increasing commercial condom use from 50% to 60%, prevented 21% and 33% of HIV infections among all women and men respectively during the first year of intervention, rising to 48% and 51% if continued over 20 years. They found that this project prevented 45% and 71% of infections among FSW alone over 1 and 20 years respectively. They concluded that focusing HIV prevention interventions on key populations in Cotonou, owing to their disproportionate role in HIV transmission, could have large impacts in the whole population.

Lowndes et al. further explored the extent to which FSW and their clients were driving the HIV epidemic of Cotonou, by examining the epidemiological and behavioural data at the time [255]. They estimated the number of non-FSW female partners of clients exposed to HIV or STI infection through unprotected sex with these men. They estimated that eight non-FSW women per 100 clients were exposed to HIV through unprotected sex, and 65 were exposed to STIs. They concluded that most HIV infections that are found in the general population are not generated from within these groups, but are rather transmitted via 'bridging groups' such as clients of FSW, and that HIV interventions focusing on core groups are therefore disproportionately effective. Core groups can be interpreted as equivalent to 'key populations' in today's literature.

In 2006, Vickerman et al. published another mathematical modelling study of HIV/STI transmission in Cotonou [279]. Their dynamic model was fitted to available prevalence data of HIV, gonorrhoea, chlamydia and genital ulcer disease among FSW and the general population. They estimated that SIDA1/2 prevented 18% of HIV infections among FSW in

Cotonou over 4 years (1995–8) inclusive, and 22% in the overall population, with around three quarters of the impact due to increased condom use, and one quarter due to STI treatments.

They further argued that the 'syndromic approach' to treating STIs in Sub-Saharan Africa, whereby eligibility for antibiotics depends on the appearance of symptoms, has poor sensitivity and specificity, causing many false negative women to remain untreated. Moreover, the overuse of antibiotics in false positive cases may lead to resistance. They aimed to estimate the impact of an alternative strategy, called 'point of care' (POC) tests, which screen for infection prior to treatment. Being more sensitive, these tests could successfully diagnose more STI infections and thereby reduce risk of HIV acquisition. They estimated that if SIDA2 had used POC tests with a 70% sensitivity, 37% more STI infections would have been treated, and 42% more HIV infections would have been averted.

One final modelling studied conducted in 2014 by Williams et al. aimed to explore the potential contribution of interventions to the observed decrease in HIV prevalence among most groups in Cotonou from 1993-2008 [50]. For instance, the reduction in HIV prevalence in FSW (from 53% in 1993 to 30% in 2008) was followed by a significant decline in prevalence in men aged 15-49 from 3.4% in 1998, to 2.0% in 2008 [280]. A similar decline had however not been observed in women of the same age, whose prevalence did not significantly change (3.5% to 4.0% in 1998 and 2008 respectively) [280]. Williams and colleagues parameterised and calibrated a compartmental transmission model with newer behavioural and disease prevalence data (HIV and gonorrhoea), as well as ART uptake data that had become available at the time. They estimated that the SIDA1/2/3 interventions prevented 62% of new infections in FSW between 1993 and 2008, and would continue to reduce FSW prevalence to 11% in 2040, assuming condom use continues at the same levels. They found that commercial sex may have contributed to 93% of all cumulative heterosexual HIV infections between the start of the epidemic until 1993, and 52% between 1993 and 2008, adding further justification to the focus of HIV interventions on key populations.

2.2.6 Aims and justification

The first four objectives in this thesis are as follows:

1. Is the city of Cotonou, Benin on track to meet the 2030 UNAIDS goal of eliminating HIV/AIDS as a global health threat?

2. What was the impact of a two-year intervention of PrEP and TasP among FSW in Cotonou?

3. What would be the impact of scaling up PrEP and TasP among FSW over twenty years in Cotonou?

4. What was the impact of disruptions to HIV prevention services and sexual behaviour, caused by the COVID-19 epidemic in Cotonou?

The UNAIDS goal of eliminating HIV/AIDS as a global health threat is measured as a 90% reduction in both new HIV infections and HIV-related deaths by 2030 compared to 2010 [131]. Impact of interventions and disruptions can also be measured with these outcomes, or with others such as life-years gained [281]. The modelling approach that I use to answer these four objectives must therefore be able to calculate such outcomes.

Given that the HIV epidemic of Cotonou is concentrated among key populations [255], who likely form 'bridges' of infections into the general population [137], it is important that the method used to measure these outcomes incorporates not only direct infections, but also secondary ones [59]. As such, a dynamic transmission model is a very useful and robust method to represent these chains of transmission. Furthermore, there is a wealth of data in Cotonou spanning several decades; some of these data are already presented in the modelling studies described in this chapter, but there have since been further published HIV prevalence estimates, as well as data on HIV testing, ART coverage and dropout. A large selection of data provides an excellent basis for a dynamic transmission model, as the epidemic can be closely recapitulated with careful model fitting.

The recent PrEP/TasP study, conducted among professional female sex workers in Cotonou, Benin, is also an interesting and useful source of data [169]. In Chapter 1, I discussed the mixed results for real-world PrEP and TasP effectiveness among women, particularly high-risk women such as FSW [160–162, 148, 151, 153, 154]. A model impact analysis based on the rich dataset provided by the PrEP/TasP study would provide a reliable and timely evaluation of recently available pharmaceutical interventions. Indeed, at the time of the analysis, it was the first published impact study of PrEP among FSW in Sub-Saharan Africa based on real-life uptake, adherence and retention data. The real-life setting of the study, combined with uniquely detailed biological markers for adherence provide excellent data to accurately reflect the context and impact of the intervention.

To answer these four objectives, however, the rich data set of Cotonou must be reviewed and summarised in a useful format. This represents the aim of the next section of this chapter. The following section describes the process of designing the transmission model with two principles in mind: to maximise the utility of the data, and to best answer the research questions. The final section in this chapter describes the process of fitting the model to the summarised data in Cotonou. The results of the model fitting, epidemic trajectory

and PrEP/TasP impact analysis are presented in Chapter 3. The results of the impact of COVID-related disruptions are presented in Chapter 4.

For this model impact analysis, I received approval from the Imperial College Research Ethics Committee (ICREC reference: 18IC4641) to carry out my secondary data analysis on data from the PrEP/TasP demonstration study (2014-2017). This model impact analysis involved no additional data collection.

2.3 Data review

2.3.1 Main data sources

I divided the data into eight categories. The first four are for model parameterisation: demographic, sexual behaviour, biological and HIV intervention data. The fifth and sixth categories represent data relevant to the PrEP and TasP arms of the PrEP/TasP study respectively, which are also used for model parameterisation. For all parameters I present a brief description of the calculation of the fixed estimate or range (see Section 'Model fitting'), based on the original data sources.

The model should reflect key epidemiological outcomes to accurately represent the HIV epidemic of Cotonou. As such, the seventh and eighths categories are data those used for model fitting. The seventh category summarised demographic, epidemiological and intervention data which serve as targets for the model fitting. Finally, data used for cross-validation are presented in the eighth section.

There were many data sources which informed parameters and model fitting targets. In this section I briefly describe some of the key sources.

Demographic parameters and model fitting targets were informed heavily by four censuses carried out in Benin, in 1979, 1992, 2002 and 2013 [242–245]. Other sources included the UN World Population Prospect [241], and studies estimating the size of pFSW (2012) and client populations (1998, 2008) [255, 254, 257, 138].

Many parameters were informed by several rounds of FSW IBBS (1993, 1995, 1998, 2002, 2005, 2008, 2012, 2015) [254, 262, 258, 259]. These IBBS were conducted using a cluster-based sampling method, based on mapping studies of FSW in Cotonou [262]. Sex work sites were identified in Cotonou, and all FSW above the age of 15 were asked to participate in the surveys. For surveys between 1993-2002 inclusive, the number of sites ranged between 47-97, and the number of FSW recruited ranged between 350-593 [262]. For surveys

conducted between 2005-2015 inclusive, over 100 sites were sampled and over 1000 FSW were recruited [262, 258, 259]. The same sex work sites identified for FSW surveys were randomly sampled and used to recruit clients in the same fashion [282], to produce client IBBS (1998-2015) [255, 265, 283].

These IBBS were used for data in many categories. For example, they were used for sexual behaviour parameters, such as number of annual partners or condom use. Further, they were used for intervention data, such as HIV testing, ART uptake and coverage, adherence (viral suppression) and dropout. Finally, they were used to estimate HIV prevalence through time, used in model fitting and cross-validation.

For low-risk groups, there were two household-based surveys conducted in Cotonou in 1998 and 2008 [280], from which considerable parameterisation and model fitting data were extracted. Census areas were determined in Cotonou using probabilities proportional to size and number of households; in both surveys there were 38 census areas. The households chosen for each survey were randomly sampled from these census areas, producing 950 and 1,070 households in 1998 and 2008, respectively. All residents in these households between 15-49 years old were asked to participate in the studies [280]. These surveys provided data on sexual behaviour, HIV testing, ART coverage, and HIV prevalence.

Finally, the PrEP/TasP study itself was a major source of data for parameterisation and model fitting.

2.3.2 Summary of available data

2.3.2.1 Demographic parameters

Table 2.1 Demographic parameters for Cotonou

Name (units)	Risk group	Year	Symbol	Value/Range	Justification/Source
Year of seeding of epidemic	NA	NA	t_0	1986	The earliest detection of HIV-1 in Cotonou (and in Benin) was in 1986-1987. Bigot et al. 1987.
Total population aged 15-59 at seed	All	NA	N_0	286114	Calculated using the census population estimate of all adults 15-59 in Grand Cotonou at 1979, and the growth rate between 1979 and 1992. Benin censuses 1979 and 1992.
Population growth rate (per year)	All	1979-1992	ε	0.059	Calculated assuming exponential growth between census population estimates. Note that the period 1993-2002 has a prior range that is sampled in the Latin Hypercube; this was done to achieve more uncertainty in projected population size. Benin censuses 1979, 1992, 2002, 2013.
	All	1993-2002		0.048 - 0.058	
	All	2002-2013		0.027	
	All	2014-		0.027	
Proportion of population that is female at seeding	NA	1986-2035	$prop_F$	0.512 - 0.520	Little variation in this estimate for each census. Range is the lowest (1992) to highest (2013) estimate. Benin censuses 1979, 1992, 2002, 2013.
Proportion of women that are professional FSW at seeding	Pro FSW	1986-2035	$prop_{PFSW}$	0.0024 - 0.00715	The lower bound is estimated from the number of FSW estimated in 2012 (Mapping study, Equite en Santé); the upper bound is estimated from Vandepitte et al. 2006.
Ratio of women that are part-time FSW to Pro FSW at seeding	Part-time FSW	1986-2035	$ratio_{LFSW}$	1 - 2	Estimated from mapping study of FSW in Cotonou (Equité en Santé).
Proportion of women that are former FSW at seeding	Former FSW in Cotonou	1986-2035	$prop_{FFSW}$	Same as proportion that are professional FSW	The number of former FSW in 1986 is unknown so I assumed it is the same number of active professional FSW.
Proportion of men that are clients at seeding	Client	1986-2035	$prop_C$	0.066 - 0.30	Lower bound taken from Lowndes et al. 2002; Upper bound from West Africa HIV/AIDS Epidemiology and Response Synthesis report 2008
Fraction of the 15-59 population that is not yet sexually active at seeding	Low-risk women	1986-2035	$prop_{FV}$	0.079 - 0.20	Lower bound taken from household survey in Cotonou 1998; upper bound from DHS 2011
	Low-risk men	1986-2035	$prop_{MV}$	0.070 - 0.17	Lower bound taken from household survey in Cotonou 1998; upper bound from DHS 2011
Prevalence of HIV at seeding	FSW	NA	π_{FSW}	0.0132 - 0.0659	Earliest estimate of HIV prevalence among FSW in Cotonou. 7/215 FSW (3.26%) in survey tested as positive for HIV-1. Prior range is the 95% CI for an exact binomial test. Bigot et al. 1987.
	Client	NA	π_C	0.000313 - 0.00294	4/3478 (0.12%) blood donors found to be positive in Cotonou between 1986 and 1987. No estimates specific to men or women so assumed to be same. Prior range is the 95% CI for an exact binomial test. Boyle and Way 1990.
	Low-risk women	NA	π_F		
	Low-risk men	NA	π_M		
Rate of leaving by ageing (per year)	All	1986-2035	ν	0.022	Given population modelled is 15-59, an average of 45 years, this rate is the inverse 1/45 per year.
Background mortality rate (per year)	Women	1986-2035	μ	0.0187 - 0.0200	Life expectancy estimates for 1980 - 2015 separately for men and women aged 15. The inverse of these durations gives the mortality rate of those aged 15. The highest and the lowest estimates of the mortality rate give the upper and lower bounds separately for men and women. UN World Population Prospects.
	Men			0.0194 - 0.0220	
Rate of leaving of sex work (per year)	Pro FSW & part-time FSW	1986-2035	$1 / dur_{FSW}$	0 - 0.55	FSW reported how long they have been doing sex work; this gives the lower bound for the duration. I assumed an upper bound of infinity. The rate of leaving is the inverse of these durations. Alary et al 2002. Behanzin et al 2013.
Proportion of FSW that are non-Beninese	Pro FSW	1986-2035	$frac_{FSWforeign}$	0.5 - 0.9	From the IBBS, this is the fraction of FSW that are foreign born. Alary et al 2002. Behanzin et al 2013.
HIV prevalence of incoming professional FSW	Pro FSW	1986 - 1993	$Prev_{1986}$	0 - 0.163	The HIV prevalence of low-risk women in neighbouring Ghana, Nigeria and Togo were weighted by the fractions of FSW in Benin of each nationality. Alary et al 2002. Behanzin et al 2013.UNAIDS Reference Group.
		1993 - 2015	$Prev_{1993}$	Linearly interpolated	
		2015 -	$Prev_{2015}$	0 - 0.046	
Rate of leaving of client category (per year)	Client	1986-2035	$1 / dur_{client}$	0 - 0.295	Estimated from the average number of years since first buying sex. The upper bound is the shortest average duration of being a client. The lower bound assumes clients never stop purchasing sex once they start. Lowndes et al 2002.
Rate of entering the sexual population (per year)	Not yet sexually active female	1986-2035	χ_F	0.2 - 0.5	Derived from the 25 and 75% quantiles of the age of sexual debut of those who were not sexually active at 15, giving the rate of entry to the sexually active population. Behanzin et al. 2013.
	Not yet sexually active male	1986-2035	χ_M	0.2 - 0.5	
Proportion of people that are sexually active at 15	Women	1986-2035	$frac_{WSA}$	0.12 – 0.17	Household based surveys in 1998 and 2008
	Men	1986-2035	$frac_{MSA}$	0.18 – 0.35	Household based surveys in 1998 and 2008

2.3.2.2 Sexual behaviour parameters

Table 2.2 Sexual behaviour parameters for Cotonou

Name	Group	Year	Symbol	Value	Justification/Source
Commercial partner change rate (per year)	Professional FSW	1993	c^{C}_{PFSW}	192 - 1277	Value estimated in 1993 and 2005. Between 1986-1993 has the value assumed for 1993. Between 1993 and 2005 it is linearly interpolated. From 2005 onwards it is kept constant. Alary et al 2002. Behanzin et al 2013. ESDG-2012. ESDG-2015.
		1993 - 2005		Linearly interpolated	
		2005 -		81 - 562	
	Part-time FSW	1986-2035	c^{C}_{LFSW}	26 - 78	Personal communication from Michel Alary, MD.
	Client	1986 - 1998	c^{C}_{C}	8.4 - 32	Value estimated in 1998 and 2002. Between 1986-1998 has the value assumed for 1998. Between 1998 and 2002 it is linearly interpolated. After 2002 onwards it is kept constant. Lowndes et al 2000. ESDG-2015.
		1998 - 2002		Linearly interpolated	
		2002 -		11.1 - 19.8	
Non-commercial partner change rate (per year)	Professional FSW	1986-2035	c^{N}_{PFSW}	0.31 – 0.86	Combination of all casual and long term partnerships estimated in IBBS for professional FSW (1993-2015) which are sexual but not contractual. Alary et al 2002. Behanzin et al 2013. ESDG-2012. ESDG-2015.
	Part-time FSW	1986-2035	c^{N}_{LFSW}	0.41 – 1.04	Estimated from 'bargirls' IBBS 2012 and 2015. ESDG-2012. ESDG-2015.
	Client	1986-2035	c^{N}_{C}	1.6 – 3.3	Estimated from clients IBBS 1998-2015. Lowndes et al 2000. ESDG-2012. ESDG-2015.
	Low-risk women	1986 - 1998	c^{N}_{F}	0.93 - 0.99	Household based surveys 1998 and 2008. Behanzin et al 2013.
		1998 - 2008		Linearly interpolated	
		2008 -		0.77 - 0.82	
	Low-risk men	1986 - 1998	c^{N}_{M}	1.25 - 1.43	Household based surveys 1998 and 2008. Behanzin et al 2013.
		1998 - 2008		Linearly interpolated	
		2008 -		0.73 - 0.84	
Commercial number of sex acts per partnership (per year)	Professional FSW	1986-2035	n^{C}_{PFSW}	1 - 3	Estimates from FSW IBBS 1993-2015. Alary et al 2002. Behanzin et al 2013. ESDG-2012. ESDG-2015.
	Part-time FSW	1986-2035	n^{C}_{LFSW}	1	Assumption based on advice from experts familiar with the population in Cotonou, Michel Alary MD.
Non-commercial number of sex acts per partnership (per year)	Professional and part-time FSW	1986 - 2002	n^{N}_{PFSW}	13.0 - 20.0	Value estimated in 2002 and 2015. Between 1986-2002 has the value assumed for 2002. Between 2002 and 2015 it is linearly interpolated. After 2015 onwards it is kept constant. Alary et al 2002. Behanzin et al 2013. ESDG-2012. ESDG-2015.
		2002 - 2015		Linearly interpolated	
		2015 -		38.2 - 60.0	
	Clients	NA	n^{N}_{C}	Matching the partner in their respective partnerships	The number of sex acts per partnerships has to match the values estimated among professional and part-time FSW
	Low-risk women - Low-risk men	1998	n^{N}_{F}	35 - 44	Value estimated in 1998 and 2011. Between 1986-1998 has the value assumed for 1998. Between 1998 and 2011it is linearly interpolated. After 2011 onwards it is kept constant. Census 2013. Behanzin et al 2013.
		1998 - 2011		Linearly interpolated	
		2011		29 - 38	
Proportion of commercial sex acts protected by condom	Professional FSW & clients	1986	$fc^{C}_{PFSWclient}$	0 - 0.18	There is a general upward trend through time. The proportion is estimated at various time points and linearly interpolated between them. After 2008 the value is kept constant. Alary et al 2002. Behanzin et al 2013. ESDG-2012. ESDG-2015. Lowndes et al 2000.
		1986 - 1993		Linearly interpolated	
		1993		0.18 - 0.33	
		1993 - 1998		Linearly interpolated	
		1998		0.4 - 0.73	
		1998 - 2002		Linearly interpolated	
		2002		0.61 - 0.99	
		2002 - 2008		Linearly interpolated	
		2008		0.86 - 0.99	
	Part-time FSW & clients	1986	$fc^{C}_{LFSWclient}$	0	Estimated in 2015. Assumed zero in 1986. Linearly interpolated in between. After 2015 the value is kept constant Alary et al 2002. Behanzin et al 2013. ESDG-2012. ESDG-2015. Lowndes et al 2000.
		1986 - 2015		Linearly interpolated	
		2015		0.25 – 0.52	
Proportion of non-commercial sex acts protected by condom	Professional FSW & clients	1986	fc^{N}_{FSW}	0	Estimated in 2002. Assumed zero in 1986. Linearly interpolated in between. After 2002 the value is kept constant. Alary et al 2002. Behanzin et al 2013. ESDG-2012. ESDG-2015. Lowndes et al 2000.
		1986 - 2002		Linearly interpolated	
		2002 -		0.19 - 0.62	
	Part-time FSW & clients	1986	$fc^{C}_{LFSWclient}$	0	Estimated in 2015. Assumed zero in 1986. Linearly interpolated in between. After 2015 the value is kept constant. Alary et al 2002. Behanzin et al 2013. ESDG-2012. ESDG-2015. Lowndes et al 2000.
		1986 - 2015		Linearly interpolated	
		2015		0.138 – 0.383	
	Clients	1986-2035	fc^{N}_{C}	Same as FSW or GPF in respective partnerships	Condom use must be equal on both sides of the partnership.
	Low-risk women	1986 - 1998	fc^{N}_{F}	0.033 - 0.05	Estimated in 1998 and 2011. Linearly interpolated in between. After 2011 the value is kept constant. Census 2013. Behanzin et al 2013.
		1998 - 2011		Linearly interpolated	
		2011 -		0.16 - 0.26	
	Low-risk men	Same as GPF	fc^{N}_{M}	Same as GPF	

2.3.2.3 Biological parameters

Table 2.3 Biological parameters for Cotonou

Name	Group	Year	Symbol	Value	Justification/Source
HIV baseline transmission probability male to female per sex act	All	1986-2035	β	0.0006 - 0.00109	From the 95% confidence interval. Boily et al 2009.
RR HIV transmission risk female to male	All	1986-2035	$RR^{\beta FM}$	0.53 - 2	From the 95% confidential interval. Boily et al 2009.
RR transmission probability with HSV-2 infection for HIV acquisition in commercial sex	Professional FSW	1986-2035	RRj^{HSV2}	0.9 – 2.3	Pooled estimates for studies reporting prevalent HSV-2 exposure. Looker et al 2017.
RR transmission probability with HSV-2 infection for HIV acquisition in commercial sex	Women who are not professional FSW	1986-2035	RRj^{HSV2}	1.8 – 3.4	
RR transmission probability with HSV-2 infection for HIV acquisition in commercial sex	Clients	1986-2035	RRj^{HSV2}	1.5 – 2.2	
RR transmission probability with HSV-2 infection for HIV acquisition in commercial sex	Men who are not clients	1986-2035	RRj^{HSV2}	2.2 – 4.3	
Prevalence of HSV-2	Pro FSW & part-time FSW	1986-2035	πj^{HSV2}	0.87 - 0.94	n = 275, backcalculated for confidence intervals. Lowndes et al 2002.
Prevalence of HSV-2	Clients	1986-2035	πj^{HSV2}	0.18 - 0.28	Table 3 +- 5%. Weiss et al 2001.
Prevalence of HSV-2	Low-risk women	1986-2035	πj^{HSV2}	0.27 - 0.32	n = 1015 backcalculated for confidence intervals. Lowndes et al 2002.
Prevalence of HSV-2	Low-risk men	1986-2035	πj^{HSV2}	0.098 - 0.14	n = 928, backcalculated for confidence intervals. Lowndes et al 2002.
RR transmission circumcision from female-to-male transmission	All	1986-2035	$RR^{\beta circum}$	0.34 - 0.72	From the 95% confidence interval. Mills et al 2008. Siegfried et al 2009.
RR of acquiring HIV when partner is in primary phase	All	1986-2035	$RR^{primary}$	4.5 - 18.8	From the 95% confidence interval. Boily et al 2009.
RR of acquiring HIV when partner is in AIDS (CD4 <200) phase	All	1986-2035	RR^{AIDS}	4.5 - 11.9	From the 95% confidence interval. Boily et al 2009.
Efficacy of ART if virally suppressed regarding reduction in onward transmission	All	1986-2035	eff^{ART}	0.96 - 0.99	Converted efficacy to RR using RR = 1-efficacy. Cohen et al 2016.
Fraction of other groups virally suppressed	Low-risk	1986-2035	$V2_{-9}$	0.42 - 0.85	42% is from serology report by department, Benin 2015; 85% is from ART audit report 2017.
Fraction of professional FSW virally suppressed	Professional FSW	1986 - 2015	V1		
Condom efficacy	All	1986-2035	ec	0.58-0.95	From the 95% confidence interval. Weller et al 2002.
Duration between seroconversion and death (years)	All	1986-2035	$dur_{SCdeath}$	8.7 - 12.3	Glynn et al 2007. Lavreys et al 2006. May et al 2009. Morgan et al 2002. Pantazis et al 2012. Peters et al 2007. Todd et al 2007. Van der Paal et al 2007.
Duration of primary phase (years)	All	1986-2035	$1/\gamma_{01}$	0.25 - 0.42	Lavreys et al 2006. Boily et al 2009.
Duration of CD4 200-349 (years)	All	1986-2035	$1/\gamma_{04}$	2.3 - 4.4	May et al 2009. Pantazis et al 2012. Bakari et al 2008.
Duration in CD4 <200 (years)	All	1986-2035	$1/\alpha_{05}$	0.58 - 3.17	Based on several estimates. May et al 2009. Morgan et al 2002.
Duration of CD4 >500 (years)	All	1986-2035	$1/\gamma_{02}$	Derived from other sampled parameters	Derived: ($dur_{SCdeath}$ - γ_{01} - γ_{04} - α_{05}) gives the time after primary phase until CD4 200-349. Dividing this time by 2 gives the duration of both CD4 >500 and CD4 350-500, assuming they have the same duration.
Duration of CD4 350-500 (years)	All	1986-2035	$1/\gamma_{03}$		
Death rate CD4 350-500 (per year)	All	1986-2035	α_{03}	0.01 - 0.05	Mortality rate from source minus average background mortality rate (0.02 years^-1). UN World Population Prospects. Badri et al 2006.
Death rate CD4 200-349 (per year)	All	1986-2035	α_{04}	0.03 - 0.10	
Reduction in progression for individuals on ART	All	1986-2035	$prog_{ART}$	4.82 - 10.2	The same factor is applied to both the slowing of mortality and slowing of progression. Keiser et al 2004.
Reduction in mortality for individuals on ART	All	1986-2035	$mort_{ART}$		

2.3.2.4 Intervention parameters

Table 2.4 Intervention parameters for Cotonou

Name	Group	Year	Symbol	Value	Justification/Source
Probability of having been tested in last year	Professional FSW	1986 - 2005	τ	0	No widespread testing until 2005 according to local experts. Batona et al 2015. ESDG-2015.
		2005 - 2013		Linearly interpolated	
		2013		0.65	
		2013 - 2015		Linearly interpolated	
		2015 -		0.68	
	Low-risk women	1986 - 2001		0	No widespread testing until 2002 according to local experts. UNAIDS report 2016. House hold surveys 1998, 2008.
		2001 - 2006		Linearly interpolated	
		2006		0.14	
		2006 - 2008		Linearly interpolated	
		2008		0.21	
		2008 - 2011		Linearly interpolated	
		2011 -		0.33	
	Low-risk men	1986 - 2001		0	No widespread testing until 2001 according to local experts. UNAIDS report 2016. House hold surveys 1998, 2008.
		2001 - 2006		Linearly interpolated	
		2006		0.098	
		2006 - 2008		Linearly interpolated	
		2008		0.1	
		2008 - 2011		Linearly interpolated	
		2011 -		0.058	
Relative increase in testing at CD4 stage <200	All	1986-2035	$RR_{test200}$	1 - 5.4	Calculated from proportion who commenced ART that had CD4 count <200 cells/mm^3. Ekouevi et al 2010.
ART initiation rate (per year)	Professional FSW	1986-2035	ρ	0.25 - 6	From Ministry of Health "File Active" database 2017.
	Rest of population	1986-2035		6 - 12	
ART dropout rate (per year)	Professional FSW	1986-2035	ϕ_{PFSW}	0.023 - 0.11	Based on TasP study retention as well as 11 clinical centres; no differentiation between males and females Ekouevi et al 2010. UNAIDS report 2016.
	Client	1986-2035	ϕ_C	0.023 - 0.11	
	Low-risk women	1986-2035	ϕ_F		
	Low-risk men	1986-2035	ϕ_M		
ART re-initiation rate (per year)	Professional FSW	1986-2035	ι_{FSW}	0.25 - 1.5	From Ministry of Health "File Active" database 2017.
	Low-risk	1986-2035	ι_{rest}	0.25 - 1.5	
PrEP/TasP study testing rate (per year)	Professional FSW	2015 - 2017	T_{PrEP}	0.5 - 2	Assumption.

2.3.2.5 PrEP parameters

The Partners PrEP efficacy trial (Donnell et al. 2014 [284]) was conducted among heterosexual serodiscordant African couples, evaluating the efficacy of tenofovir disoproxil fumarate (TDF) and TDF with emtricitabine (FTC/TDF), the same prophylactic component in the PrEP/TasP demonstration study.

They observed that most study participants displayed blood plasma tenofovir concentrations which were either below the assay limit of quantification (< 0.3 ng/ml), or consistent with daily adherence (>40 ng/ml). Few participants had a concentration between 0.3 – 40 ng/ml, from as little as 8% to at most 14%, depending on the visit. The study authors calculated PrEP efficacy at different thresholds of adherence: >0.3 ng/ml, >10 ng/ml and >40 ng/ml. Because so few participants had an intermediate adherence, those that had detectable tenofovir were almost always also daily adherers and as a result the estimated efficacy of PrEP for each threshold were very similar.

I assumed PrEP efficacy for daily adherers in this analysis was 90 – 95% (modelled as per sex act reduction in HIV transmission), consistent with the adjusted risk reduction for women on FTC/TDF at tenofovir concentration >40 ng/ml estimated by Donnell et al. I assumed that undetectable tenofovir conferred 0% PrEP efficacy. Based on expert opinion, I assumed that intermediate adherence (0.3 – 40 ng/ml) was associated with 30 – 49% efficacy. This assumption was explored in a sensitivity analysis (Chapter 3).

The proportions of visits in the perfect, intermediate and non-adherent categories are shown in Table 2.5. Also shown are the assumed efficacies of PrEP. Given these, I calculated the average effectiveness of PrEP as 41%, by computing the sum of product (average adherence × efficacy) across all adherence levels.

Table 2.5 PrEP parameters for Cotonou

Name	Group	Year	Symbol	Value	Justification/Source
Testing rate when on PrEP	Professional FSW	2015 - 2017	τ_i^{11} & τ_i^{00}	4	FSW are tested 4x a year on PrEP. Giguère et al 2019. Mboup et al 2018.
PrEP offering rate (per year)	Professional FSW	2015 - 2017	ξ_i	0.1 - 0.2	298 FSW offered PrEP from 902 susceptibles over 2 years in the PrEP/TasP study. Mboup et al 2018.
PrEP acceptance rate (per year)	Professional FSW	2015 - 2017	σ_i	86%	Of 298 FSW eligible for PrEP, 256 were recruited. Mboup et al 2018.
Fraction in perfect PrEP adherence category	Professional FSW	2015 - 2017	ζ_a	36.9%	Proportion of visits tenofovir concentration > 40ng/ml. Mboup et al 2018.
Fraction in intermediate PrEP adherence category	Professional FSW	2015 - 2017	ζ_b	16.5%	Proportion of visits tenofovir concentration 0.3-40 ng/ml Mboup et al 2018.
Fraction in PrEP non-adherence category	Professional FSW	2015 - 2017	ζ_c	46.6%	Proportion of visits tenofovir concentration < 0.3 ng/ml. Mboup et al 2018.
PrEP drop out rate (per year)	Professional FSW	2015 - 2017	κ	0.45 - 0.64	Number of drop outs divided by PrEP person-years. Mboup et al 2018.
PrEP efficacy	Daily adherence	2015 - 2017	eP^{1a}	0.9 - 0.95	Based on Donnell et al. efficacies at different tenofovir blood plasma concentrations, and expert opinion.
	Intermediate adherence	2015 - 2017	eP^{1b}	0.3 - 0.49	
	Non-adherence	2015 - 2017	eP^{1c}	0	Undetectable PrEP assumed zero efficacy. Marrazzo et al 2015. Van Damme et al 2012.

2.3.2.6 TasP parameters

Table 2.6 TasP parameters for Cotonou

Name	Group	Year	Symbol	Value	Justification/Source
TasP initiation rate (per year)	Professional FSW	1986 - 2015	$init_{TasP}$	0	No TasP intervention
		2015 - 2017		0.5 - 5	Initiations over person-years of HIV+ professional FSW in the PrEP/TasP study. Mboup et al 2018.
Fraction of professional FSW virally suppressed	Professional FSW	2015 -	VS_{pFSW}	0.75 - 0.87	Observed in study. Mboup et al 2018. From Ministry of Health audit report 2017.

2.3.2.7 Model fitting data

Table 2.7 presents the data to which the transmission model is fitted. The model was fitted in two stages: A) pre-study model fitting and B) PrEP/TasP study model fitting. Further details are available in the 'Model Fitting' section.

Table 2.7 Outcomes and target range or estimate used for the A) pre-PrEP/TasP study model fitting (before 2015) and B) the PrEP/TasP study scenario model fitting (2015-2017).

Fitting outcomes	Year (start)	Target range/estimate	Justification/Source
A) Pre-study fitting			
Stage 1 (1986-2015): Demographic			
Total population size of sexually active adult population of Grand Cotonou (N)	1992	343,705 – 465,013	The population size of Cotonou aged 15-59 was estimated by censuses in 1992, 2002 and 2013. Assuming a constant growth rate equal to the one between 2002-2013, I projected estimates of population sizes in 2020 and 2030.
	2002	579,325 – 783,793	
	2013	776,076 – 1,049,985	
	2020	959,418 – 1,298,036	
	2030	1,210,305 – 1,637,471	
Percentage of women who are professional FSW (%)	All	0.19 – 0.72	The lower bound is estimated from the number of FSW in Cotonou estimated in 2012 (Equite en Santé); the upper bound is estimated from Vandepitte et al. 2006, assuming that half of these women are pFSW.
Number of professional FSW (N)	2012	889 – 1,391	From the mapping study (Equite en Santé)
Percentage of women who are active FSW (professional + part-time) (%)	All	0.48 – 1.4	Lower bound from the mapping study (Equite en Santé); upper bound is estimated from Vandepitte.
Percentage of men who are clients (%)	All	7.4 – 30	Lower bound taken from Lowndes et al. 2002; Upper bound from West Africa HIV/AIDS Epidemiology and Response Synthesis report 2008.
Percentage of women who are not yet sexually active (%)	All	7.9 – 20	Lower bound taken from household survey in Cotonou 1998; upper bound from DHS 2011.
Percentage of men who are not yet sexually active (%)	All	7 – 17	Lower bound taken from household survey in Cotonou 1998; upper bound from DHS 2011.
Stage 2 (1986-2015): Epidemiological			
HIV prevalence professional FSW (%)	1993	48 – 58	Several IBBS among professional FSW. Alary et al 2002. Behanzin et al 2013. ESDG-2012. ESDG-2015.
	2002	32 – 46	
	2008	25 – 34	
	2015	14 – 22†	
	2017	7 – 13 †	
HIV prevalence clients (%)	2002	6.8 – 12.0	IBBS among clients. Lowndes et al 2000. ESDG-2015.
HIV prevalence all women (%)	2011	1.3 – 3.5	DHS survey 2011.
HIV prevalence all men (%)	2011	0.75 – 2.9	DHS survey 2011.
HIV incidence rate professional FSW (% infected per person-year)	2015	0 – 3	Estimated by Diabaté et al. 2018.
Stage 2 (1986-2015): ART coverage			
Professional HIV infected FSW on ART (N)	2015	42 – 56	"File Active": anonymised database of current female sex workers on ART in Benin. Ministry of Health.
Men and women on ART combined (N)	2017	8524 – 18273	UNAIDS Country progress report 2019.
ART coverage of men and women combined (%)	2011	33 – 52	UNAIDS Country progress report 2016.
ART coverage of men and women combined (%)	2017	60 – 91	UNAIDS Country progress report 2019.
Ratio of women to men on ART	2017	1.5 – 3.0	Ministry of Health audit report.
B) PrEP/TasP study scenario (2015-2017)			
Person-years on PrEP (years)	2015-2017	250	Data from the PrEP/TasP study. Mboup et al 2018.
PrEP initiations (N)	2015	256	
Number on PrEP at end of study (N)	2017	121	
TasP initiations (N)	2015	105	
All professional FSW on ART at end of first year (N)	2016	122	
All professional FSW on ART at end of study (N)	2017	137	

† Parameter sets which produced model outcomes meeting at least one of the empirical estimates of HIV prevalence among professional FSW in 2015 and 2017 were accepted.

2.3.2.8 Cross-validation data

Table 2.8 presents the cross-validation data. Although the model was not calibrated to these data, they served as independent verification of how closely my simulations reproduced the context in Cotonou.

Table 2.8 Cross-validation data for Cotonou

Outcome	Year	Value	Source
HIV prevalence professional FSW (%)	1995	43.0 – 54.4	From FSW IBBS. Alary et al 2002. Behanzin et al 2013. ESDG-2012. ESDG-2015.
	1998	36.6 – 44.7	
	2005	30.4 – 39.4	
	2012	23.0 – 32.2	
HIV prevalence clients (%)	1998	5.9 – 11.6	From client IBBS. Lowndes et al 2000. ESDG-2015.
	2002	6.8 – 11.6	
	2008	3.5 – 9	
	2012	1.3 – 5.2	
	2015	0.6 – 3.5	
HIV prevalence all women (%)	1998	2.4 – 4.8	From household surveys 1998, 2008.
	2008	3.0 – 5.3	
HIV prevalence all men (%)	1998	2.3 – 4.7	
	2008	1.2 – 3.0	
Deaths due to AIDS (all groups)	1990	100 – 300	From AIDSinfo, UNAIDS.
	1991	100 – 300	
	1992	100 – 300	
	1993	100 – 300	
	1994	100 – 300	
	1995	100 – 400	
	1996	100 – 400	
	1997	100 – 500	
	1998	200 – 600	
	1999	200 – 900	
	2000	200 – 1200	
	2001	500 – 1500	
	2002	500 – 1500	
	2003	500 – 1500	
	2004	500 – 1500	
	2005	500 – 2000	
	2006	500 – 2000	
	2007	500 – 2000	
	2008	500 – 2000	
	2009	200 – 900	
	2010	200 – 900	
	2011	200 – 800	
	2012	200 – 1300	
	2013	200 – 1300	
	2014	200 – 1300	
	2015	200 – 1300	
	2016	100 – 700	
Proportion of HIV+ on ART and virally suppressed (all groups combined)	2015	14 – 30	From AIDSinfo, UNAIDS.
	2016	15 – 32	
	2017	28 – 60	

2.4 Model development

2.4.1 Model description

I developed a deterministic model of sexual HIV transmission and HIV interventions (PrEP, ART, condom use) in the heterosexual adult population of Grand Cotonou (abbreviated GC; comprising Cotonou, Abomey-Calavi and Seme-Kpodji). The model represents an open and growing population aged 15-59 years old, stratified in 9 risk groups (subscript i): 1) professional FSW, 2) part-time FSW, 3) low-risk women, 4) former FSW in Cotonou, 5) clients, 6) low-risk men, 7) not-yet-sexually-active (NYSA) women, 8) NYSA men, 9) former FSW outside of Cotonou.

New susceptible individuals can join the population at time-varying rates that reflect changes in the total population size of Cotonou. They are distributed between six risk groups (NYSA women & men, low-risk women & men and professional & part-time FSW) according to relative population sizes. E_i represents the number of new susceptible recruits joining the population in each group i. NYSA women and men enter the sexually active but susceptible low-risk population at per capita rates u_7 and u_8 determined by the average age at sexual debut, respectively. Professional and part-time FSW are recruited directly from the low-risk women population of GC at per capita rates $v_{1,3}$ and $v_{2,3}$, respectively, or from outside of GC (from both inside and outside Benin) at rates E_1 and E_2 respectively, reflecting HIV prevalence of neighbouring countries at time t. Professional and part-time FSW cease sex work and join the former FSW categories at rates u_1 and u_2, reflecting duration in sex work, respectively; the factor $p_{foreign}$ determines the fraction that leaves GC upon ceasing sex work. Low-risk population males become clients at a per capita rate u_6 and return at per capita rate u_5, reflecting duration of paying for sex. Individuals leave the modelled population due to aging, HIV-related and unrelated mortality. Movement between risk groups applies equally to all care states and stages of infection. The movement between risk groups are summarised in Figure 2.1.

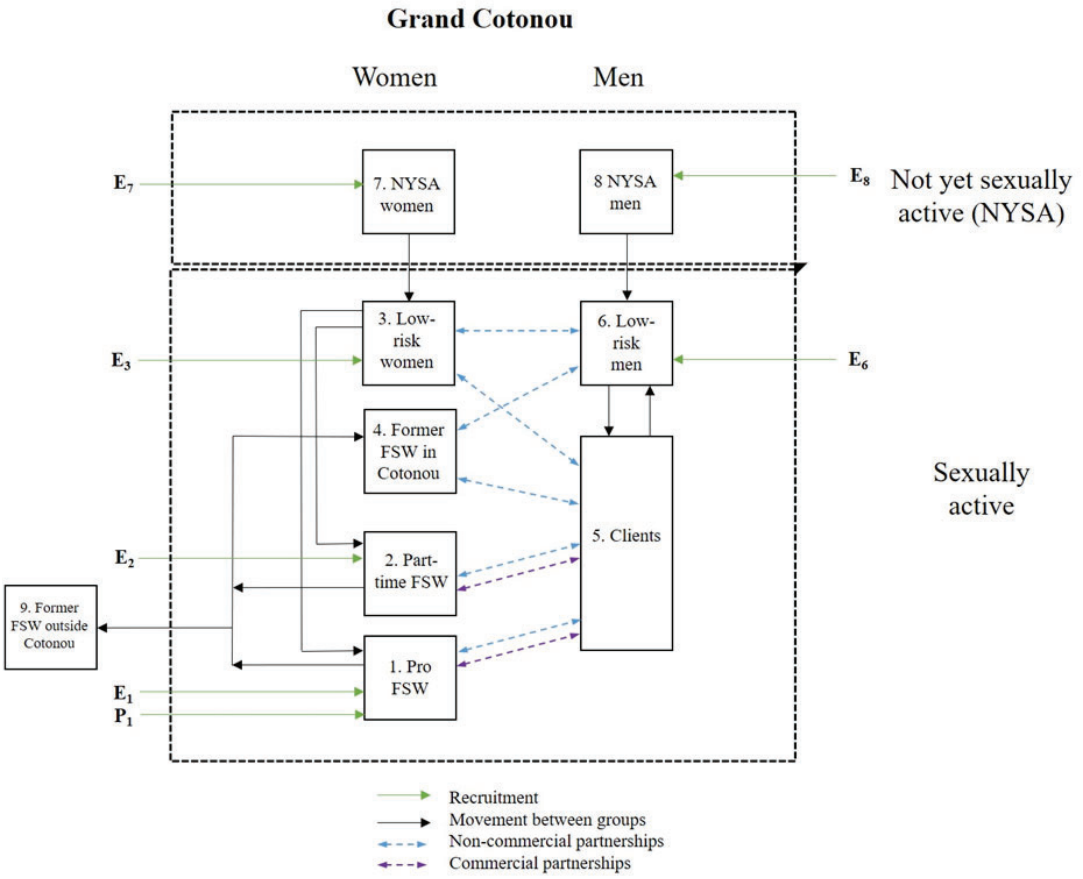


Figure 2.1 Recruitment, movement and sexual mixing between groups, represented by green, black and blue lines respectively. Groups within the dashed rectangles form the sexual transmission network of Grand Cotonou. E_i represents the numbers of HIV-uninfected individuals in group i recruited to the model; P_1 represents the number of HIV-infected professional FSW recruited. Commercial partnerships are only formed between professional and part-time FSW and clients. NYSA = not yet sexually active. FSW = female sex worker.

Apart from a fraction of professional FSW, all individuals enter the population in the susceptible off PrEP category (S_i^0). Susceptible professional FSW can initiate PrEP at per capita rate ζ_i, depending on background testing rate $\tau_i(t)$ of group i at time t, testing rate due to the PrEP/TasP study $\eta_i(t)$, PrEP offering rate $\xi_i(t)$ and PrEP acceptance rate σ_i, and the fraction fP_i^a, fP_i^b, fP_i^c, that are initially perfectly, intermediately and non-adherent (though receiving pills), moving into states S_i^{1a}, S_i^{1b} and S_i^{1c} respectively. The data indicates decreasing PrEP adherence over time; thus PrEP users can move into lower adherence categories (at rates ψ^a and ψ^b moving from perfect to intermediate and from intermediate to non adherence respectively) or cease to use PrEP and move to the dropout PrEP category, at per capita rates

κ^a, κ^b, κ^c. Professional FSW can retire from sex work and move to the former FSW groups inside or outside GC; in this case they do not move to their associated PrEP category, but rather into their susceptible off PrEP category (S_i^0).

Sexually active susceptible individuals are assumed to become infected at a per capita force of infection λ_{ij}^{rs} that depends on their PrEP status, number and type of sexual partners, HIV prevalence among partners, sexual mixing patterns between risk groups, the fraction of sex acts effectively protected by condoms, the levels of PrEP use and efficacy (related to adherence status), the partner's infectiousness (i.e., varying by disease stage, ART treatment). The protective effect of male circumcision on HIV acquisition risk was captured in the per-act HIV transmission probability since more than 99% of men in GC are circumcised [255]. Infected individuals are represented in the model as $I_i^{r,s}$, for risk group i, care state r and stage of infection s.

Following HIV infection, the model represents the course of disease progression stratified by CD4 cell count levels and different levels of engagement in the care and treatment cascade. Untreated undiagnosed HIV-infected individuals progress through a short highly infectious primary infection stage followed by 4 longer disease stages defined based on CD4 cell counts, at rates $\gamma^{r,s}$ that reflect care status (r) and stage of infection (s). Individuals in infected stages CD4 350-500, CD4 200-349 and CD4 <200 (s = 3, 4, 5 respectively) die at AIDS mortality rate α^{rs}.

When infected, individuals on PrEP enter a distinct primary phase of infection to reflect the fact that they are tested at an increased rate. These individuals leave this compartment at testing rate $\tau_i^{1,1}(t)$, dropout of PrEP at rate κ^1, or progress to CD4 >500 at rate $\gamma_i^{1,1}$.

All undiagnosed individuals in these 5 stages are tested for HIV at a per capita rate τ_i^{rs}, depending on group i, care status r and stage of infection s, moving to the diagnosed off ART category, who experience the same disease progression as undiagnosed individuals. When modelling the intervention, individuals experience an increased testing rate of $\eta_i(t)$.

Diagnosed individuals initiate treatment at a per capita rate ρ_i^s that depends on group i, calendar year (due to changes in eligibility criteria) and stage of infection s. When modelling the intervention, individuals experience an increased ART initiation rate of $\rho'(t)$. Compared to those not on ART, treated individuals have slower rates of disease progression and reduced HIV-related mortality, represented by factors $prog_{ART}$ and $mort_{ART}$ respectively. Finally, some individuals may experience a therapeutic failure or discontinue ART, at rate φ^s, in which case the disease follows its natural progression unless these individuals re-initiate treatment at rate ι_i^s.

Individuals leave the population at per capita rates ν, μ^{rs} and α^{rs}, which represent rate of aging, background mortality and AIDS-related mortality, respectively. The full system of ODEs is presented in this chapter in Section 2.4.2. The model structure is presented in Figure 2.2.

HIV was seeded at the initial prevalence π_i, and the system of ODEs was solved numerically from 1986 – 2035 using the lsoda algorithm (Petzhold and Hindmarsh) with a variable time step. The model was coded in R v3.6.2, using "odin" (github.com/mrc-ide/odin), a wrapper around the deSolve package for solving ordinary differential equations (https://cran.r-project.org/web/packages/deSolve/index.html).

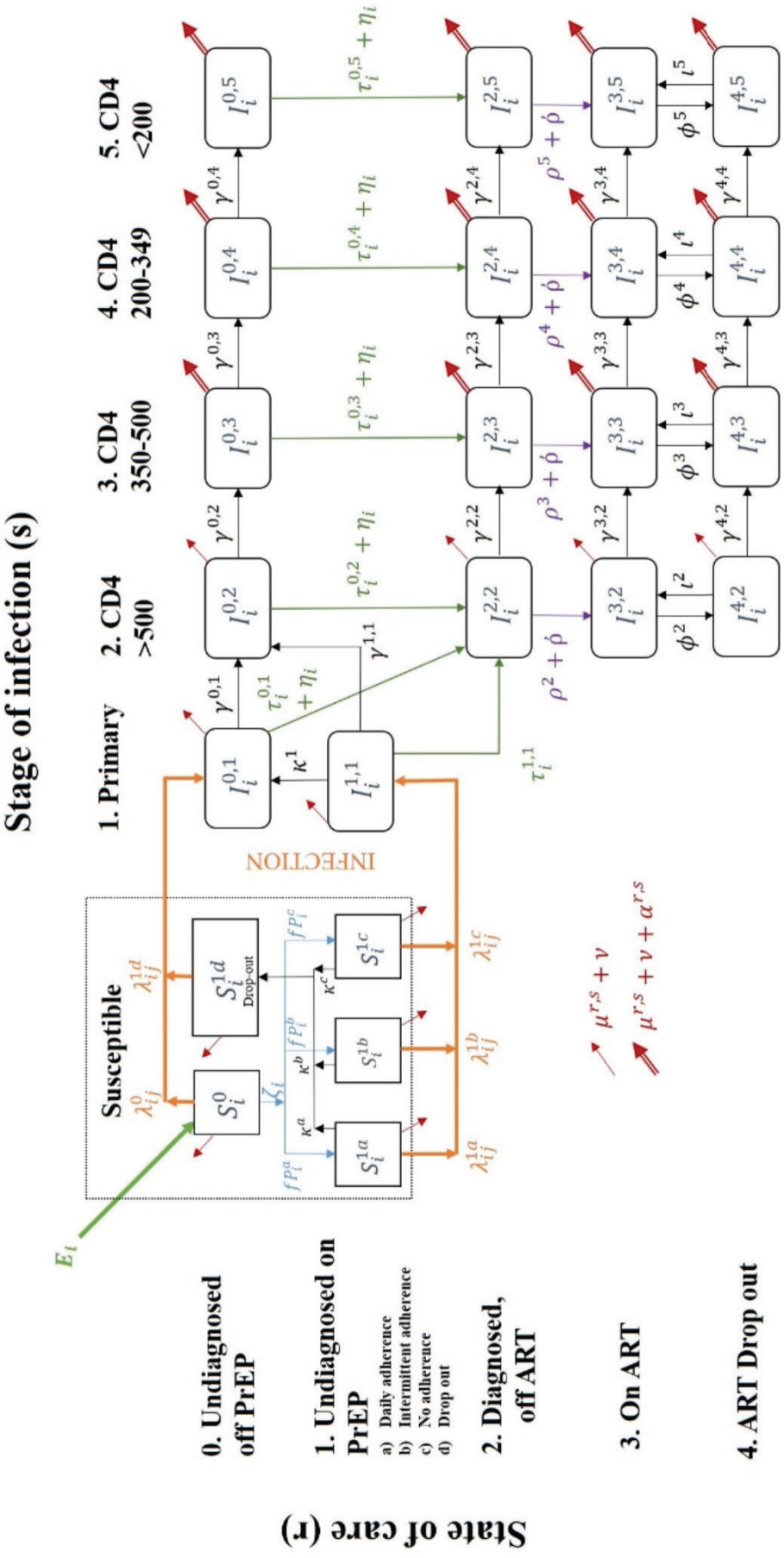


Figure 2.2 The model flow diagram, representing progression through state of care (r) and stage of infection (s). PrEP initiation, HIV testing, and ART initiation are represented by blue, green and purple arrows respectively. Red arrows represent mortality and ageing; double red arrow indicates increased mortality due to HIV infection. Orange arrows represent HIV infection. All risk groups are represented by these compartments, but some compartments will always be empty for some groups (i.e. "on PrEP" for groups who are not professional FSW).

2.4.2 Equations

2.4.2.1 Ordinary differential equations (ODEs)

Figure 2.2 shows the care state r and stage of infection s categories for each group i. S_i^r and $I_i^{r,s}$ represent susceptible and infected individuals respectively.

Testing/ART care-state (r):

0. Off PrEP, undiagnosed
1. On PrEP, undiagnosed
   (a) Perfect adherence (daily dosing)
   (b) Intermediate adherence
   (c) Non-adherence (undetectable)
   (d) PrEP Dropout
2. Diagnosed, off ART
3. Diagnosed, on ART
4. ART dropout

Stage of infection (s):

1. Primary phase
2. CD4 >500 cells/μl
3. CD4 350-500 cells/μl
4. CD4 200-349 cells/μl
5. CD4 <200 cells/μl

$$N(t) = \sum_{i=1}^{i=8} \sum_{r=0}^{r=4} \left(S_i^r(t) + \sum_{s=1}^{s=5} I_i^{rs}(t) \right) \tag{2.1}$$

$N(t)$ is the total population of GC at time t; and $S_i^r(t)$ and $I_i^{rs}(t)$ are the susceptible and infected populations respectively, of group i, care state r and disease state s. Note that $N(t)$ does not include former FSW outside of GC ($i = 9$).

The following equations describe the recruitment of new individuals $(E(t))$ into the population. Due to the population growth of GC, new people are added to the population every timestep dependent on the growth rate $\varepsilon(t)$.

$$E_i(t) = (\varepsilon(t) + \mu_i + \nu)\ N(t)\ \omega_i, \qquad i \neq 1 \tag{2.2}$$

Where $E(t)$ is the number of new susceptible people entering the sexually active population into group i; $\varepsilon(t)$ is the total population growth rate; μ_i is the natural mortality rate; ν is the rate of exit due to ageing; $N(t)$ is the total population size at time t as defined in Equation 2.2; ω_i is the proportion of new people distributed into group i.

For the numbers of professional FSW to be maintained, they have an extra term to account for those who retire from sex work and leave Grand Cotonou. Additionally, professional FSW enter the model at a prevalence of the neighbouring countries to Benin (Nigeria, Togo and Ghana) at time t:

$$F_i(t) = (\varepsilon(t) + \mu_i + \nu)\ N(t)\ \omega_i + \frac{1}{dur_{PFSW}} * N_{PFSW}(t) * frac_{FSW\,foreign}, \qquad i = 1 \tag{2.3}$$

$$E_i(t) = F_i(t) * \left(1 - prev_{FSW_{foreign}}(t)\right), \qquad i = 1 \tag{2.4}$$

$$P_i(t) = F_i(t) * prev_{FSW_{foreign}}(t), \qquad i = 1 \tag{2.5}$$

$$F_i(t) = 0, \qquad i \neq 1 \tag{2.6}$$

$$P_i(t) = 0, \qquad i \neq 1 \tag{2.7}$$

Where $F_i(t)$ is the total number of new individuals in group i entering the model at time t, $E_i(t)$ is the number of susceptible individuals in group i entering the model at time t, $P_i(t)$ is the number of infected individuals in group i entering the model at time t, dur_{PFSW} is the duration of sex work for professional FSW, $N_{PFSW}(t)$ is the number of professional FSW at time t, $frac_{FSW\,foreign}$ is the fraction of FSW that are foreign-born and $prev_{FSW_{foreign}}(t)$ is the prevalence of FSW in neighbouring countries.

Susceptible:

Susceptible off PrEP:

$$\frac{dS_i^0(t)}{dt} = E_i(t) - S_i^0(t)\left(\sum_j \left(\lambda_{ij}^0(t)\right) + \mu_i + \zeta_i + \nu\right) + T_i^0 \tag{2.8}$$

Where $E(t)$ is the number of new susceptible people entering the sexually active population into group i as described in Equation 2.4; μ_i is the background mortality rate of group i; ζ_i is the PrEP initiation rate; $\lambda_{ij}^0(t)$ is the force of infection from group j to group i; ν is the rate of exit due to ageing; T_i^{rs} is the sum of the movement of people to and from group i.

The PrEP initiation rate ζ_i is composed of background testing rate $\tau_i(t)$ of group i at time t, testing rate due to PrEP/TasP intervention $\eta_i(t)$, PrEP offering rate $\xi_i(t)$ and PrEP acceptance rate σ_i.

$$\zeta_i = (\tau_i(t) + \eta_i(t)) * \xi_i(t) * \sigma_i \tag{2.9}$$

Susceptible on PrEP (perfect adherence):

$$\frac{dS_i^{1a}(t)}{dt} = \zeta_i(t)\ fP_i^a\ S_i^0(t) - S_i^{1a}(t)\left(\sum_j \left(\lambda_{ij}^{1a}(t)\right) + \mu_i + \ \kappa_i^a + \nu\right) + T_i^{1a} \tag{2.10}$$

Where fP_i^a is the fraction of individuals in the perfect PrEP adherence category and κ_i^r is rate of PrEP dropout.

Susceptible on PrEP (intermediate adherence):

$$\frac{dS_i^{1b}(t)}{dt} = \ \zeta_i(t)\ fP_i^b\ S_i^0(t) - S_i^{1b}(t)\ \left(\sum_j \left(\lambda_{ij}^{1b}(t)\right) + \mu_i + \kappa_i^b + \nu\right) + T_i^{1b,0} \tag{2.11}$$

Where fP_i^b is the fraction of individuals in the intermediate PrEP adherence category.

Susceptible on PrEP (non-adherence):

$$\frac{dS_i^{1c}(t)}{dt} = \ \zeta_i(t)\ fP_i^c\ S_i^0(t) - S_i^{1c}(t)\left(\sum_j \left(\lambda_{ij}^{1c}(t)\right) + \mu_i + \ \kappa_i^c + \nu\right) + T_i^{1c,0} \tag{2.12}$$

Where fP_i^c is the fraction of individuals in the non-adherence to PrEP category.

Susceptible PrEP drop-out:

$$\frac{dS_i^{1d}(t)}{dt} = \kappa_i^a(t)\ S_i^{1a}(t) + \kappa_i^b(t)\ S_i^{1b}(t) + \kappa_i^c(t)\ S_i^{1c}(t) - S_i^{1d}(t)\left(\sum_j \left(\lambda_{ij}^{1d}(t)\right) + \mu_i + \nu\right) + T_i^{1d,0} \tag{2.13}$$

Note that $\lambda_{ij}^{0}(t)$ and $\lambda_{ij}^{1d}(t)$ are equivalent, as those who have dropped out of PrEP do not have any PrEP protection.

Primary infection:

As detailed in Equation 2.4, the number of new HIV-infected professional FSW entering the model depends on changing HIV prevalence in neighbouring countries over time. There are no HIV-infected individuals entering the model for any other group. The HIV-infected professional FSW entering the model are divided into the 5 stages of infection in the undiagnosed care state, proportional to the relative duration of each disease state (d^s).

Primary infection off PrEP:

$$\begin{aligned}\frac{dI_i^{0,1}(t)}{dt} = d^1\ P_i(t) + S_i^0(t)\sum_j\left(\lambda_{ij}^0(t)\right) + S_i^{1d}(t)\sum_j\left(\lambda_{ij}^{1d}(t)\right) - \\ I_i^{0,1}(t)\ \left(\gamma^{0,1} + \tau_i^{0,1}(t) + \eta_i(t) + \ \mu_i + \nu\right) + I^{1,1}(t)\ \kappa^1 + T_i^{0,1}(t)\end{aligned} \tag{2.14}$$

Where γ^{rs} is disease progression from care state r and disease stage s; $\tau_i^{rs}(t)$ is testing rate; κ^1 is the PrEP dropout rate for infected individuals.

Primary infection on PrEP:

$$\begin{aligned}\frac{dI_i^{1,1}(t)}{dt} = S_i^{1a}(t)\sum_j\left(\lambda_{ij}^{1a}(t)\right) + S_i^{1b}(t)\sum_j\left(\lambda_{ij}^{1b}(t)\right) + S_i^{1c}(t)\sum_j\left(\lambda_{ij}^{1c}(t)\right) - \\ I_i^{1,1}(t)\ \left(\gamma^{1,1} + \kappa^1 + \tau_i^{1,1}(t) + \ \mu_i + \nu\right) + T_i^{1,1}(t)\end{aligned} \tag{2.15}$$

Undiagnosed off PrEP:

CD4 > 500 off PrEP:

$$\frac{dI_i^{0,2}(t)}{dt} = d^2\, P_i(t) + \gamma^{0,1}\, I_i^{0,1}(t) + \gamma^{1,1}\, I_i^{1,1}(t) - I_i^{0,2}(t)\left(\gamma^{0,2} + \tau_i^{0,2}(t) + \eta_i(t) + \mu_i + \nu\right) + T_i^{0,2}(t) \tag{2.16}$$

CD4 350 - 500 off PrEP:

$$\frac{dI_i^{0,3}(t)}{dt} = d^3\, P_i(t) + \gamma^{0,2}\, I_i^{0,2}(t) - I_i^{0,3}(t)\left(\gamma^{0,3} + \tau_i^{0,3}(t) + \eta_i(t) + \alpha_i^{0,3} + \mu_i + \nu\right) + T_i^{0,3}(t) \tag{2.17}$$

CD4 200 - 349 off PrEP:

$$\frac{dI_i^{0,4}(t)}{dt} = d^4\, P_i(t) + \gamma^{0,3}\, I_i^{0,3}(t) - I_i^{0,4}(t)\left(\gamma^{0,4} + \tau_i^{0,4}(t) + \eta_i(t) + \alpha_i^{0,4} + \mu_i + \nu\right) + T_i^{0,4}(t) \tag{2.18}$$

CD4 < 200 off PrEP:

$$\frac{dI_i^{0,5}(t)}{dt} = d^5\, P_i(t) + \gamma^{0,4}\, I_i^{0,4}(t) - I_i^{0,5}(t)\left(\tau_i^{0,5}(t) + \eta_i(t) + \alpha_i^{0,5} + \mu_i + \nu\right) + T_i^{0,5}(t) \tag{2.19}$$

Diagnosed infection, not on ART:

CD4 > 500 diagnosed, off ART:

$$\begin{aligned}\frac{dI_i^{2,2}(t)}{dt} = \left(\tau_i^{0,1}(t) + \eta_i(t)\right)\, I^{0,1}(t) + \tau_i^{1,1}(t) I^{1,1}(t) + \left(\tau_i^{0,2}(t) + \eta_i(t)\right)\, I^{0,2}(t) - \\ I_i^{2,2}(t)(\gamma^{2,2} + \rho_i^2(t) + \rho_i'(t) + \mu_i + \nu) + T_i^{2,2}(t)\end{aligned} \tag{2.20}$$

Where $\rho_i^s(t)$ is rate of ART uptake in group i, disease state s.

CD4 350 - 500 diagnosed, off ART:

$$\frac{dI_i^{2,3}(t)}{dt} = \gamma^{2,2} I_i^{2,2}(t) + \left(\tau_i^{0,3}(t) + \eta_i(t)\right) I^{0,3}(t) - I_i^{2,3}(t)(\gamma^{2,3} + \rho_i^3(t) + \rho_i'(t) + \alpha_i^{0,3} + \mu_i + \nu) + T_i^{2,3}(t) \tag{2.21}$$

CD4 200 - 349 diagnosed, off ART:

$$\frac{dI_i^{2,4}(t)}{dt} = \gamma^{2,3} I_i^{2,3}(t) + \left(\tau_i^{0,4}(t) + \eta_i(t)\right) I^{0,4}(t) - I_i^{2,4}(t)(\gamma^{2,4} + \rho_i^4(t) + \rho_i'(t) + \alpha_i^{0,4} + \mu_i + \nu) + T_i^{2,4}(t) \tag{2.22}$$

CD4 < 200 diagnosed, off ART:

$$\frac{dI_i^{2,5}(t)}{dt} = \gamma^{2,4} I_i^{2,4}(t) + \left(\tau_i^{0,5}(t) + \eta_i(t)\right) I^{0,5}(t) - I_i^{2,5}(t)(\rho_i^5(t) + \rho_i'(t) + \alpha_i^{0,5} + \mu_i + \nu) + T_i^{2,5}(t) \tag{2.23}$$

Diagnosed infection, on ART:

CD4 > 500 diagnosed, on ART:

$$\frac{dI_i^{3,2}(t)}{dt} = \rho_i^2(t) + \rho_i'(t) I_i^{2,2} t + \iota_i^2 I_i^{4,2}(t) - I_i^{3,2}(t)(\gamma^{3,2} + \phi_i^2(t) + \mu_i + \nu) + T_i^{3,2}(t) \tag{2.24}$$

Where $\iota_i^s(t)$ is rate of ART re-initialisation in group i, disease state s; φ_i^s is rate of ART dropout.

CD4 350 - 500 diagnosed, on ART:

$$\frac{dI_i^{3,3}(t)}{dt} = \gamma^{3,2} I_i^{3,2}(t) + \rho_i^3(t) + \rho_i'(t) I_i^{2,3} t + \iota_i^3 I_i^{4,3}(t) - I_i^{3,3}(t)(\gamma^{3,3} + \phi_i^3(t) + \alpha_i^{3,3} + \mu_i + \nu) + T_i^{3,3}(t) \tag{2.25}$$

CD4 200 - 349 diagnosed, on ART:

$$\frac{dI_i^{3,4}(t)}{dt} = \gamma^{3,3} I_i^{3,3}(t) + \rho_i^4(t) + \rho_i'(t) I_i^{2,4}(t) + \iota_i^4 I_i^{4,4}(t) - I_i^{3,4}(t)(\gamma^{3,4} + \phi_i^4(t) + \alpha_i^{3,4} + \mu_i + \nu) + T_i^{3,4}(t) \tag{2.26}$$

CD4 < 200 diagnosed, on ART:

$$\frac{dI_i^{3,5}(t)}{dt} = \gamma^{3,4} I_i^{3,4}(t) + \rho_i^5(t) + \rho_i'(t) I_i^{2,5}(t) + \iota_i^5 I_i^{4,5}(t) - I_i^{3,5}(t)(\phi_i^5(t) + \alpha_i^{3,5} + \mu_i + \nu) + T_i^{3,5}(t) \tag{2.27}$$

Diagnosed infection, ART dropout / therapeutic failure:

CD4 > 500 ART drop out / therapeutic failure:

$$\frac{dI_i^{4,2}(t)}{dt} = \varphi_i^2(t)\ I_i^{3,2}(t) - I_i^{4,2}(t)\left(\gamma^{4,2} + \iota_i^2(t) + \mu_i + \nu\right) + T_i^{4,2}(t) \tag{2.28}$$

CD4 350 - 500 ART drop out / therapeutic failure:

$$\frac{dI_i^{4,3}(t)}{dt} = \gamma^{4,2}\ I_i^{4,2}(t) + \varphi_i^3(t) I_i^{3,3}(t) - I_i^{4,3}(t)\left(\gamma^{4,3} + \iota_i^3(t) + \alpha_i^{0,3} + \ \mu_i + \nu\right) + T_i^{4,3}(t) \tag{2.29}$$

CD4 200 - 349 ART drop out / therapeutic failure:

$$\frac{dI_i^{4,4}(t)}{dt} = \gamma^{4,3}\ I_i^{4,3}(t) + \varphi_i^4(t) I_i^{3,4}(t) - I_i^{4,4}(t)\left(\gamma^{4,4} + \iota_i^4(t) + \alpha_i^{0,4} + \ \mu_i + \nu\right) + T_i^{4,4}(t) \tag{2.30}$$

CD4 < 200 ART drop out / therapeutic failure:

$$\frac{dI_i^{4,5}(t)}{dt} = \gamma^{4,4}\ I_i^{4,4}(t) + \varphi_i^5(t) I_i^{3,5}(t) - I_i^{4,5}(t)\left(\iota_i^5(t) + \alpha_i^{0,5} + \ \mu_i + \nu\right) + T_i^{4,5}(t) \tag{2.31}$$

2.4.2.2 Entering the population

New people entering the population are distributed into the 9 groups at fractions determined by the parameter ω_i.

$$\omega_{PFSW} = frac_F * frac_{FPFSW} * frac_{FSWforeign} \tag{2.32}$$

$$\omega_{LFSW} = frac_F * frac_{FPFSW} * ratio_{LFSW} * frac_{FSWforeign} \tag{2.33}$$

$$\omega_{GPF} = frac_F - \omega_{PFSW} - \omega_{LFSW} - \omega_{VF} \tag{2.34}$$

$$\omega_{GPM} = (1 - frac_F) * frac_{MSA} \tag{2.35}$$

$$\omega_{VF} = frac_F * (1 - frac_{FSA}) \tag{2.36}$$

$$\omega_{VM} = (1 - frac_F) * (1 - frac_{MSA}) \tag{2.37}$$

$$\omega_C = \omega_{FFSWinCot} = \omega_{FFSWoutCot} = 0 \tag{2.38}$$

$$\sum_{i=1}^{i=9} \omega_i = 1 \tag{2.39}$$

Where ω_i is the fraction of new individuals entering group i; $PFSW$, $LFSW$, GPF, GPM, VF, VM, C, $FFSWinCot$ and $FFSWoutCot$ represent professional female sex workers, part-time female sex workers, low risk women and men, not yet sexually active women and men, clients, former female sex workers in GC and former female sex workers outside of GC respectively; $frac_F$ is the fraction of individuals that are female; $frac_{FPFSW}$ is the fraction of women that are professional sex workers; $ratio_{LFSW}$ is the ratio of number of part-time sex workers to professional sex workers; $frac_{FSWforeign}$ is the fraction of sex workers that are non-Beninese; $frac_{MSA}$ is the fraction of men that are sexually active; $frac_{FSA}$ is the fraction of women that are sexually active.

2.4.2.3 Turnover

Turnover between compartments was a critical component of the model for spreading infections from high risk to low risk groups, and was also important for maintaining the correct demographic patterns in the population. Turnover is represented by two separate parameters, moving out (u_i) and moving in (v_{ij}). The former is the rate of leaving a group i, thus need only apply to the group in question. The latter is the rate of entering a group i from group j which captures the correct numbers in care-state r and stage of disease s.

Turnover applies equally to both susceptible and infected states (whereby $X = S$ and $X = I$ respectively), and applies to all care states and stages of infection. Thus, the movement into states S_i^0 and $I_i^{4,5}$ will arrive from states S_j^0 and $I_j^{4,5}$ respectively. As such, when considering the movement of people from group j to group i, it will be with the HIV prevalence of group j. Note however that professional FSW on PrEP do not move to their associated PrEP category, but rather into the susceptible off PrEP category (S_i^0) of the destination group.

Movements in and out are combined into the turnover $T_i^{rs}(t)$ as follows:

$$T_i^{r,s}(t) = -X_i^{r,s}(t) * u_i + \sum_{j=1}^{j=9} v_{ij} * X_j^{r,s}(t) \tag{2.40}$$

Where $T_i^{r,s}(t)$ is the turnover of group i, care-state r and disease state s; $X_i^{r,s}(t)$ is the current number of individuals in the state, u_i is the rate of leaving of group i; v_{ij} is the rate of entering in group i from group j.

Groups (i):

1. Professional FSW
2. Part-time FSW
3. Low-risk women
4. Former FSW in GC
5. Clients
6. Low-risk men
7. Not yet sexually active women
8. Not yet sexually active men

9. Former FSW outside model

Not yet sexually active men and women

Not yet sexually active (NYSA) men and women enter the sexually active population. NYSA women $(i = 7)$ move into the low-risk women category $(i = 3)$; NYSA men $(i = 8)$ move into the low-risk men category $(i = 6)$.

Rate of leaving NYSA woman group (u_7):

$$u_7 = \frac{1}{\chi_7 - 15} \tag{2.41}$$

Rate of leaving NYSA men group (u_8):

$$u_8 = \frac{1}{\chi_8 - 15} \tag{2.42}$$

Where u_7 and u_8 are the rate leaving of NYSA women and men groups respectively; χ_i is the average age of sexual debut of group i. Fifteen is the youngest age in the model, thus $\chi_8 - 15$ represents the length of time spent NYSA.

This results in the movement of NYSA women and men to be expressed as follows:

$$T_7^{r,s}(t) = -X_7^{r,s}(t) * u_7 \tag{2.43}$$

$$T_8^{r,s}(t) = -X_8^{r,s}(t) * u_8 \tag{2.44}$$

Where $T_7^{r,s}(t)$ and $T_7^{r,s}(t)$ are the movements of NYSA women and men respectively of care state r and stage of infection s; $X_i^{r,s}$ is either a susceptible $(X = S)$ or infected $(X = I)$ individual in group i, care state r and stage of infection s at time t.

Professional and part-time female sex workers

Professional and part-time FSW are recruited from either outside of the sexual network of GC, or recruited from the low-risk population. Both FSW groups retire to former FSW groups (both inside and outside GC).

Rate of leaving sex work is equal to the inverse of the duration of sex work ($\frac{1}{dur_{PFSW}}$ and $\frac{1}{dur_{LFSW}}$ for professional and part-time FSW respectively).

Rate of leaving professional sex work (u_1):

$$u_1 = \frac{1}{dur_{PFSW}} \tag{2.45}$$

Rate of leaving part-time sex work (u_2):

$$u_2 = \frac{1}{dur_{LFSW}} \tag{2.46}$$

Where dur_{PFSW} and dur_{LFSW} is the duration of professional and part-time sex work respectively.

The movement rates of FSW from the low-risk women group are calculated as follows:

Rate of entering professional sex work from low-risk women:

$$v_{1,3} = \left(\frac{1}{dur_{PFSW}} + \mu_F + \nu\right) \frac{N_1(0)}{N_3(0)} \left(1 - frac_{FSW\,foreign}\right) \tag{2.47}$$

Rate of entering part-time sex work from low-risk women:

$$v_{2,3} = \left(\frac{1}{dur_{LFSW}} + \mu_F + \nu\right) \frac{N_2(0)}{N_3(0)} \left(1 - frac_{FSW\,foreign}\right) \tag{2.48}$$

Where μ_F is female baseline mortality; ν is the rate of ageing; $N_i(0)$ is the initial population size of group i; $frac_{FSW\,foreign}$ is the fraction of sex workers that are non-Beninese. If $frac_{FSW\,foreign} = 0$, all FSW are recruited from the low-risk women group of GC; if $frac_{FSW\,foreign} = 1$, all FSW are recruited from outside GC within the rate of entry E_1 and E_2 for professional and part-time sex workers respectively. The ratio of professional FSW and part-time FSW to low-risk women ($\frac{N_1(0)}{N_3(0)}$ and $\frac{N_2(0)}{N_3(0)}$ respectively) is required to ensure that the approximately the same numbers of FSW that retire from sex work are recruited from the low-risk women group. μ_F and ν are also required to recruit FSW at the rate at which they age and die.

Total movement for professional and part-time FSW ($T_1^{r,s}(t)$ and $T_2^{r,s}(t)$ respectively), including leaving and entering, are thus expressed as follows:

$$T_1^{r,s}(t) = -X_1^{r,s}(t) * u_1 + v_{1,3} * X_3^{r,s}(t) \tag{2.49}$$

$$T_2^{r,s}(t) = -X_2^{r,s}(t) * u_2 + v_{2,3} * X_3^{r,s}(t) \tag{2.50}$$

Where u_1 and u_2 are rates of leaving professional and part-time sex work respectively, which are multiplied by the number of individuals in each care state r and stage of infection s at time t ($X_1^{r,s}(t)$ and $X_2^{r,s}(t)$ respectively); $v_{1,3}$ and $v_{2,3}$ are the rates of entering professional and part-time sex work respectively, which are multiplied by the number of individuals in the low-risk women ($X_3^{r,s}(t)$) of care state r and stage of infection s.

Low-risk women

The rate of leaving of low-risk women group (u_3) is the combination of the rates $v_{1,3}$ and $v_{2,3}$:

$$\begin{aligned} u_3 = &\left(\frac{1}{dur_{PFSW}} + \mu_F + \nu\right) \frac{N_1(0)}{N_3(0)} \left(1 - frac_{FSW\,foreign}\right) + \\ &\left(\frac{1}{dur_{LFSW}} + \mu_F + \nu\right) \frac{N_2(0)}{N_3(0)} \left(1 - frac_{FSW\,foreign}\right) \end{aligned} \tag{2.51}$$

The rate of entering of the low-risk women group is equal to the rate of leaving of the NYSA women group as shown in the following equation:

$$v_{3,7} = \frac{1}{\chi_7 - 15} \tag{2.52}$$

Total movement for low-risk women ($T_3^{r,s}(t)$), including leaving and entering, is expressed as follows:

$$T_3^{r,s}(t) = -X_3^{r,s}(t) * u_3 + v_{3,7} * X_7^{r,s}(t) \tag{2.53}$$

Where $X_3^{r,s}(t)$ is the number of low-risk women individuals in care state r and stage of infection s at time t; u_3 is the leaving rate of low-risk women; $v_{3,7}$ is the entering rate of low-risk women from the NYSA women group; $X_7^{r,s}(t)$ is the number of NYSA women in the same care state and stage of infection.

Former FSW inside and outside Grand Cotonou

Professional and part-time FSW retire to become former FSW, either inside or outside of GC. The parameter $frac_{FSW\,foreign}$ represents the fraction of FSW that were foreign-born, which determines how many FSW leave GC upon retirement.

Rate of entering former FSW in GC from professional FSW category:

$$v_{4,1} = \frac{1}{dur_{PFSW}} \left(1 - frac_{FSW\,foreign}\right) \tag{2.54}$$

Rate of entering former FSW in GC from part-time FSW category:

$$v_{4,2} = \frac{1}{dur_{LFSW}} \left(1 - frac_{FSW\,foreign}\right) \tag{2.55}$$

Rate of entering former FSW outside GC from professional FSW category:

$$v_{9,1} = \frac{1}{dur_{PFSW}} \, frac_{FSW\,foreign} \tag{2.56}$$

Rate of entering former FSW outside GC from part-time FSW category:

$$v_{9,2} = \frac{1}{dur_{LFSW}} \, frac_{FSW\,foreign} \tag{2.57}$$

Total movement for former FSW inside ($T_4^{r,s}(t)$) and outside GC ($T_9^{r,s}(t)$) are calculated as follows:

$$T_4^{r,s}(t) = v_{4,1} * X_1^{r,s}(t) + v_{4,2} * X_2^{r,s}(t) \tag{2.58}$$

$$T_9^{r,s}(t) = v_{9,1} * X_1^{r,s}(t) + v_{9,2} * X_2^{r,s}(t) \tag{2.59}$$

Where $X_1^{r,s}(t)$ and $X_2^{r,s}(t)$ represent the number of professional and part-time FSW respectively, in care state r and stage of infection s at time t.

Clients

Clients are recruited only from the low-risk men population, at a rate inversely proportional to the duration of being a client ($\frac{1}{dur_{client}}$).

Rate of leaving client group (u_5):

$$u_5 = \frac{1}{dur_{client}} \tag{2.60}$$

To keep the proportion of men that are clients approximately the same over time, the rate of entering the client group from the low-risk men group ($v_{5,6}$) scales the rate of leaving of the client group by the fraction of men that are clients at seeding ($\frac{N_5(0)}{N_6(0)}$). The background male mortality and ageing rates (μ_M and ν respectively) are included to replace the clients that are leaving due to death or ageing. The rate of entering the client category from the low-risk men ($v_{5,6}$) is calculated as follows:

$$v_{5,6} = \left(\frac{1}{dur_{client}} + \mu_M + \nu\right)\frac{N_5(0)}{N_6(0)} \tag{2.61}$$

Total movement for clients is thus calculated as follows:

$$T_5^{r,s}(t) = -X_5^{r,s}(t) * u_5 + v_{5,6} * X_6^{r,s}(t) \tag{2.62}$$

Where $X_6^{r,s}(t)$ is the number of low-risk men in care state r and stage of infection s at time t.

Low-risk men

Low-risk men leave at rate u_6, which is calculated as follows:

$$u_6 = \left(\frac{1}{dur_{client}} + \mu_M + \nu\right)\frac{N_5(0)}{N_6(0)} \tag{2.63}$$

Rate of entering low-risk men from the client category ($v_{6,5}$) is calculated as follows:

$$v_{6,5} = \frac{1}{dur_{client}} \tag{2.64}$$

Rate of entering low-risk men from the NYSA men category ($v_{6,8}$) is calculated as follows:

$$v_{6,8} = \frac{1}{\chi_8 - 15} \tag{2.65}$$

Where χ_8 is the average age of sexual debut of NYSA men.

Total movement for the low-risk men category ($T_6^{r,s}(t)$) combines the turnover with clients with the movement from NYSA men, and is calculated as follows:

$$T_6^{r,s}(t) = -X_6^{r,s}(t) * u_6 + v_{6,5} * X_5^{r,s}(t) + v_{6,8} * X_8^{r,s}(t) \tag{2.66}$$

Where $X_5^{r,s}(t)$ and $X_8^{r,s}(t)$ are the numbers of clients and NYSA men respectively, in care state r and stage of infection s.

2.4.2.4 Testing rates

Household, IBBS and DHS surveys provided estimates on the probability of being tested in the last year for FSW and men and women in GC. After interpolating over time yearly estimates of testing coverage (see parameter table), every time step I estimated the rate of undergoing testing with the following equation:

$$\tau_i^{r,s} = -\ln(1 - ptest_i) \tag{2.67}$$

Where $\tau_i^{r,s}$ is the rate of testing of group i; $ptest_i$ is the probability of being tested last year of group i.

There are two cases when this is different:

1. Individuals in $I_i^{0,5}$ have an increased testing rate ($\tau_i^{0,5}$) as calculated below

$$\tau_i^{0,5} = -\ln(1 - ptest_i) * RR_{testCD4200} \tag{2.68}$$

Where $RR_{testCD4200}$ is the increase in testing due to a CD4 count <200 cells/mm^3, which is associated with symptoms of AIDS.

2. For individuals on PrEP but infected in primary phase ($I^{1,1}$), testing rate $\tau_i^{1,1}$ is fixed to 4 ($year^{-1}$), as those under PrEP surveillance are tested every 3 months.

$$\tau_i^{1,1} = 4 \tag{2.69}$$

I assumed no difference in test sensitivity by PrEP status or stage of infection.

Furthermore, increased testing in the TasP/PrEP intervention was reflected in increased testing when simulating the intervention scenario. This additional testing rate, represented by $\eta_i(t)$, was only applied during the 2 years of the study, or for longer when simulating an intervention scale up.

2.4.2.5 ART initiation rates

ART initiation rate is represented by ρ_i^s, varying by group i and stage of infection s. Eligibility of ART has historically depended on CD4 count according to national and WHO recommendations; the rates of ART initiation in the model reflected this. Prior to 2012, only those with a CD4 count below 350 cells/mm^3 were eligible for ART ($s = 4,5$); in 2015 this was extended to those below 500 ($s = 3,4,5$), and in 2016 all stages of infection were eligible.

Increased ART initiation due to the TasP/PrEP intervention was also represented in the model, with the parameter $\rho_i'(t)$, which applies only to professional FSW.

The assumed proportion of all pFSW on ART that are virally suppressed was a weighted average of observed viral suppression among pFSW on TasP (87%), and the viral suppression among pFSW outside of TasP (75%).

2.4.2.6 Force of infection

The force of infection ($\lambda_{ij}^r(t)$) is defined as the per capita rate at which susceptible individuals of risk group i and care state r are infected by individuals of risk group j. It is derived from the total risk of infection from commercial (C) and non-commercial partnerships (N), all care states r and stages of infection s, including sex acts with and without protection from condoms and PrEP.

$$\lambda_{ij}^r(t) =$$
$$p_{ij}^C c_i^C \sum_{r=0}^{r=4}\sum_{s=1}^{s=5}\left(\frac{I_j^{r,s}(t)}{\sum_{r=0}^{r=4}\sum_{s=1}^{s=5}\left(I_j^{r,s}(t)+S_j^{r,s}(t)\right)}\left(1-\left(1-(1-e_P^r)\,\beta_{ij}\,R^{r,s}\right)^{n_{ij}^C(t)\left(1-f_{ij}^C(t)\right)}\times\left(1-\left(1-ec_{ij}\right)\,(1-e_P^r)\,\beta_{ij}\,R_j^{r,s}\right)^{n_{ij}^C(t)\,f_{ij}^C(t)}\right)\right)$$
$$+$$
$$p_{ij}^N c_i^N \sum_{r=0}^{r=4}\sum_{s=1}^{s=5}\left(\frac{I_j^{r,s}(t)}{\sum_{r=0}^{r=4}\sum_{s=1}^{s=5}\left(I_j^{r,s}(t)+S_j^{r,s}(t)\right)}\left(1-\left(1-(1-e_P^r)\,\beta_{ij}\,R^{r,s}\right)^{n_{ij}^N(t)\left(1-f_{ij}^N(t)\right)}\times\left(1-\left(1-ec_{ij}\right)\,(1-e_P^r)\,\beta_{ij}\,R_j^{r,s}\right)^{n_{ij}^N(t)\,f_{ij}^N(t)}\right)\right) \tag{2.70}$$

Where p_{ij}^C and p_{ij}^N are the probability of sexual contact between groups i and j for commercial and non-commercial partnerships respectively; c_i^C and c_i^N are the rates of new partner acquisition for commercial and non-commercial partnerships respectively; β_{ij}^C and β_{ij}^N are the commercial and non-commercial probability of HIV transmission per sex act from j to i respectively; $R_j^{r,s}$ is a matrix containing the scaling factors for risk of transmission based on care state r and stage of infection s of transmitter j; $n_{ij}^C(t)$ and $n_{ij}^N(t)$ are the commercial and non-commercial number of sex acts per partnership, $f_{ij}^C(t)$ and $f_{ij}^N(t)$ are the proportion of commercial and non-commercial sex acts protected by a condom, ec_{ij} is the condom efficacy, and e_P^r is PrEP efficacy of care state r (perfect, intermediate or non-adherence to PrEP). For individuals who are not using PrEP, or who are not adherent, efficacy of PrEP is assumed to be 0.

The total force of infection experienced by individuals of group i, and care state r $(\lambda tot_i^r(t))$ is the sum of the forces of infection contributed by each type of partnership they have.

$$\lambda tot_i^r(t) = \sum_j \lambda_{ij}^r(t) \tag{2.71}$$

2.4.2.7 Balancing of partnerships

Commercial partnerships

The total number of commercial partnerships as declared by women ($i = 1,2$) must match the total number of commercial partnerships as declared by men ($i = 5$), such that the following equation is satisfied:

$$c_1^C N_1 + c_2^C N_2 = c_5^C N_5 \tag{2.72}$$

Where c_i^C is the commercial partner change rate of group i; N_i is the total number of individuals in group i. I used two methods (one of which is randomly chosen in each simulation) in order to ensure that this equation is fulfilled:

1. Either I changed the partner change rate of the professional FSW every timestep in order to balance

$$c_1^{C\prime} = \frac{c_5^C N_5 - c_2^C N_2}{N_1} \tag{2.73}$$

Where $c_1^{C\prime}$ is the balanced partner change rate of professional FSW, which will be applied in the force of infection.

2. Or I changed the partner change rate of the clients every timestep in order to balance

$$c_5^{C\prime} = \frac{c_1^C N_1 + c_2^C N_2}{N_5} \tag{2.74}$$

Where $c_5^{C\prime}$ is the balanced partner change rate of clients, which will be applied in the force of infection.

Non-commercial partnerships

The number of non-commercial partnerships as declared by women ($i = 1,2,3,4$) must match the number of partnerships as declared by men ($i = 5,6$), such that the following equation is satisfied:

$$\sum_{i=1}^{i=4} c_i^N N_i = \sum_{i=5}^{i=6} c_i^N N_i c_1^N N_1 + c_2^N N_2 + c_3^N N_3 + c_4^N N_4 = c_5^N N_5 + c_6^N N_6 \tag{2.75}$$

I changed the partner change rate of low-risk women ($i = 3$) and former FSW ($i = 4$) in GC every timestep to balance partnerships:

$$c_3^{N\prime} = c_4^{N\prime} = \frac{c_5^N N_5 + c_6^N N_6 - c_1^N N_1 - c_2^N N_2}{N_3 + N_4} \tag{2.76}$$

Low-risk women and former FSW in GC share the same non-commercial partner change rate ($c_3^{N\prime}$ and $c_4^{N\prime}$ respectively).

Sex acts per partnership

This parameter is equal between each pairing, i.e. professional FSW have the same number of sex acts per commercial partnerships with clients, as clients have with professional FSW. The value of this parameter is drawn from a range whose uncertainty encompasses estimates from professional FSW and client data.

2.4.2.8 Probability of sexual contact

The contact matrices M_{ij}^C and M_{ij}^N determine whether commercial and non-commercial partnerships respectively are formed between groups i and j.

The values in the matrices take value either 1 or 0, representing sexual contact or not, respectively. For commercial partnerships (M_{ij}^C), clients ($i = 5$) have sexual contact with professional FSW ($i = 1$) and part-time FSW ($i = 2$). The corresponding elements in the matrix have value 1; all the rest have value 0.

For non-commercial partnerships (M_{ij}^N), clients have sexual contact with professional FSW, part-time FSW, low-risk women, and former FSW in Cotonou. Low-risk men have sexual contact with low-risk women and former FSW in Cotonou. The corresponding elements in the matrix have value 1; the rest have 0.

The probabilities of commercial and non-commercial sexual contact of an individual from group i with an individual from group j (p_{ij}^C and p_{ij}^C respectively) are determined by the proportion of partnerships declared by group j out of the total number of partnerships, calculated as follows:

$$p_{ij}^C(t) = \frac{c_j^C N_j}{\sum_{j=1}^{9} c_j^C N_j M_{ij}^C}, \qquad M_{ij}^C = 1 \tag{2.77}$$

$$p_{ij}^C(t) = 0, \qquad M_{ij}^C = 0 \tag{2.78}$$

$$p_{ij}^N(t) = \frac{c_j^N N_j}{\sum_{j=1}^{9} c_j^N N_j M_{ij}^N}, \;\; M_{ij}^N = 1 \tag{2.79}$$

$$p_{ij}^N(t) = 0, \;\; M_{ij}^N = 0 \tag{2.80}$$

Where c_j^C and c_j^N represent the commercial and non-commercial partner change rates of group j respectively, after balancing.

2.4.2.9 HIV transmission probability per sex act

From Boily et al. 2009, I obtained ranges for the baseline probability of HIV transmission from men to women (β) [65]. I estimated the per sex act probability of transmission from

men to women (and vice versa) which depended on: relative risk of HIV transmission probability (RR) with concurrent HSV-2 infection in the acquirer, female-to-male transmission, circumcision, and HSV-2 prevalence in the acquirer. I assumed all males were circumcised [255].

HIV transmission probability per sex act from men to women (β_{MF}):

$$\beta_{MF} = \beta * (1 + (RR_j^{HSV2} - 1) * \pi_j^{HSV2}) \tag{2.81}$$

Where RR_j^{HSV2} is the RR of transmission probability if acquirer group j is HSV-2 infected; π_j^{HSV2} is the prevalence of HSV-2 in group j.

There are other STIs which affect HIV transmission, such as gonorrhoea, chlamydia, and various other genital infections. However, prevalence of HSV-2 is at much higher prevalence than other STIs in Cotonou. HSV-2 was present in 87-94% of FSW in Cotonou in 1998 [255], while syphilis, gonorrhoea and chlamydia were at 1.5%, 20.5% and 5.1% respectively in 1998 [254]. In order to control model complexity, I aimed to reduce the number of parameters wherever possible, and so did not explicitly model the other STIs.

HIV transmission probability per sex act from women to men (β_{FM}):

$$\beta_{FM} = \beta * \ RR^{\beta\ FM} * \ RR^{\beta\ circum} * \left(1 + (RR_j^{HSV2} - 1) * \pi_j^{HSV2}\right) \tag{2.82}$$

Where $RR^{\beta\ FM}$ is the RR of female-to-male transmission; $RR^{\beta\ circum}$ is RR of male circumcision.

2.4.2.10 Relative risk of transmission

Table 2.9 shows how the relative risk of transmission ($R_j^{r,s}$) per sex act varies with care state (r) and the stage of infection (s) of the transmitter (j).

Table 2.9 The relative risk parameter employed according to the care state and stage of infection of the transmitter.

Relative risk of transmission ($R_j^{r,s}$)		Stage of infection (s) of the transmitter (j)				
		1. Primary	2. CD4 >500	3. CD4 350 - 500	4. CD4 200 - 349	5. CD4 <200
Care state (r) of the transmitter (j)	0. Off PrEP	$RR^{primary}$	1	1	1	RR^{AIDS}
	1. On PrEP	$RR^{primary}$	-	-	-	-
	2. Diagnosed off ART	-	1	1	1	RR^{AIDS}
	3. On ART	-	RR_j^{ART}	RR_j^{ART}	RR_j^{ART}	RR_j^{ART}
	4. ART Drop out	-	1	1	1	RR^{AIDS}

Where $RR^{primary}$ is the relative risk of infection per sex act if the transmitter is in the primary stage of infection; RR^{AIDS} is the relative risk of infection per sex act if the transmitter is in AIDS stage; RR^{ART} is the relative risk of infection if the transmitter is on ART. RR_j^{ART} is calculated as follows:

$$\mathrm{RR}_j^{ART} = 1 - VS_j * \ eff^{ART} \tag{2.83}$$

Where VS_j is the proportion of individuals that are virally suppressed in group j, and eff^{ART} is the efficacy of ART when virally suppressed.

2.4.2.11 Disease progression and AIDS-related mortality

Progression through the stages of infection occurs at rate $\gamma^{r,s}$ for each care state r and stage of infection s.

The duration of the entire infection ($dur_{SCdeath}$, seroconversion to death) was estimated, as well as the duration of primary phase ($\frac{1}{\gamma^{01}}$), infection stage CD4 200-349 ($\frac{1}{\gamma^{04}}$) and CD4 <200 ($\frac{1}{\alpha^{05}}$). The durations of stages CD4 >500 ($\frac{1}{\gamma^{02}}$) and CD4 350-500 ($\frac{1}{\gamma^{03}}$) are calculated as follows:

$$\frac{1}{\gamma^{02}} = \frac{1}{\gamma^{03}} = dur_{SCdeath} - \frac{1}{\gamma^{01}} - \frac{1}{\gamma^{04}} - \frac{1}{\alpha^{05}} \tag{2.84}$$

Progression happens at the same rate when undiagnosed off PrEP, undiagnosed on PrEP, diagnosed off PrEP or in ART dropout ($r = 0, 1, 2, 4$); progression is slowed by a factor of $prog_{ART}$ for those on ART ($r = 3$) such that:

$$\gamma^{32} = \frac{\gamma^{02}}{prog_{ART}} * VS_j + \gamma^{02}\left(1 - VS_j\right) \tag{2.85}$$

$$\gamma^{33} = \frac{\gamma^{03}}{prog_{ART}} * VS_j + \gamma^{03}\left(1 - VS_j\right) \tag{2.86}$$

$$\gamma^{34} = \frac{\gamma^{04}}{prog_{ART}} * VS_j + \gamma^{04}\left(1 - VS_j\right) \tag{2.87}$$

Where VS_j is the proportion of individuals that are virally suppressed in group j.

If an individual on PrEP is infected ($I_i^{1,1}$), they progress at the same rate as if not on PrEP:

$$\gamma^{1,1} = \gamma^{0,1} \tag{2.88}$$

There is no AIDS-related mortality for those who are susceptible, primary phase of infection, or CD4 >500 ($s = 0, 1, 2$). It occurs for CD4 350-500, CD4 200-249, CD4 <200 ($s = 3, 4, 5$). AIDS related mortality for those on ART ($r = 3$) is slowed by a factor of $mort_{ART}$, such that:

$$\alpha^{33} = \frac{\alpha^{03}}{mort_{ART}} * VS_j + \alpha^{03}\left(1 - VS_j\right) \tag{2.89}$$

$$\alpha^{34} = \frac{\alpha^{04}}{mort_{ART}} * VS_j + \alpha^{04}\left(1 - VS_j\right) \tag{2.90}$$

$$\alpha^{35} = \frac{\alpha^{05}}{mort_{ART}} * VS_j + \alpha^{05}\left(1 - VS_j\right) \tag{2.91}$$

2.5 Model fitting

The model was calibrated within a Bayesian framework. Based on the summarised data described earlier in the chapter, I defined pre-specified plausible prior ranges for 87 demographic, behavioural, biological and intervention model parameters (Tables 2.1-2.4), and 31 different demographic, epidemiological and intervention fitting targets before and after the PrEP/TasP study (Table 2.7).

Some parameters were sampled in a Latin Hypercube from a prior uniform distribution of the range given in the table, others were fixed and not sampled. For parameters that were sampled at multiple timepoints (e.g. commercial partner change range), I assumed that the values varied linearly between them. Parameter estimates sourced from the published literature assumed the 95% CI as the prior range. If the 95% CI was not used for a certain parameter, its range is described in the parameter tables.

The model was fitted in two stages. The first was to all data up until the beginning of the study (2015); this is called "pre-study model fitting", and the data is presented in Table 2.7A. Note that this stage was itself completed in two steps (see: pre-study model fitting section below). The second stage, PrEP/TasP study model fitting, reflected observed data from the 2-year study between 2015-2017, as shown in Table 2.7B.

2.5.1 Pre-study model fitting (1986-2015)

The pre-study model fitting was itself completed in two steps using Latin Hypercube Sampling (LHS) [285]. In step 1, I sampled my pre-specified demographic priors to form 7,000 parameter sets, and retained those which produced model estimates meeting the 11 pre-specified demographic model fitting targets (442 parameter sets retained) (Table 2.7A). In step 2 (epidemiological), the posterior ranges of these 442 parameter sets were sampled 79 million times, retaining those which produced model outcomes meeting all 25 fitting targets before 2015 (111 parameter sets retained). On important precision is that all trajectories had to fall strictly within the bounds provided in Table 2.7, except for HIV prevalence among professional FSW in 2015 and 2017. As it was impossible to obtain a trajectory which met those two data points, I relaxed the conditions such that trajectories were permitted if they fell within the bounds of at least one of them.

Following this procedure, the modelled epidemic trajectories therefore reflected the empirical data in Cotonou. Some fitting targets (e.g. men and women on ART combined) are after 2015, but these were insignificantly impacted by the PrEP/TasP study, hence I called this stage "pre-study" for simplicity. Additional available data not used for model fitting were used for cross-validation (Table 2.8).

In terms of HIV prevalence, there were multiple estimates from survey data at different time points for each risk group. It was not possible to include all points as strict fitting data without discarding all parameter sets. I therefore chose *a priori* which data points would be used for fitting or for cross-validation. In collaboration with experts in on the ground in Cotonou who have studied the setting for decades, the data points reflecting surveys in which

they had most confidence were chosen as fitting points. For example, due to the difficulties of finding a fully representative sample of part-time FSW, their HIV prevalence estimates were only used as cross-validation.

2.5.2 PrEP/TasP study model fitting (2015-2017)

I next fitted the model to the 2-year PrEP/TasP targets, reflecting the observed study data (adherence, retention, coverage). First, I defined prior ranges for six PrEP and TasP parameters (Table 2.5 and Table 2.6). For each of the 111 parameter sets retained after the pre-study fitting, I sampled the ranges of the six PrEP/TasP parameters 100 times separately and retained the set that produced model results best matching (using a least squares method) the study targets (Table 2.7B).

2.5.3 Scenarios

After fitting to the 2-year PrEP/TasP study, the model was now fully calibrated to the HIV epidemic in Cotonou, from 1986-2017. Using this fitted model, I could simulate various scenarios. Chapter 3 describes scenarios to estimate the impact of PrEP and TasP among pFSW in Cotonou. Chapter 4 describes scenarios to estimate the impact of COVID-19 on HIV epidemiology.

2.6 Discussion

In this chapter, I have introduced the HIV epidemic of Cotonou, and previous modelling studies describing it. I summarised data used to parameterise and fit the HIV dynamic transmission model in Cotonou, which is also described in this chapter. In Chapter 3, I show that the model captures the HIV epidemic in Cotonou, by comparing epidemic trajectories to model fitting targets. I also estimate the impact of the PrEP/TasP study, as well as scale-up scenarios of PrEP and TasP. In Chapter 4, I build upon the fitted model described in this chapter, to generate scenarios of COVID-related disruptions to HIV prevention services and sexual behaviour. I use these scenarios to estimate the impact of COVID-19 on HIV epidemiology.

Chapter 3

Impact of pharmaceutical interventions for HIV in Cotonou

3.1 Overview

In Chapter 2, I introduced the HIV epidemiological context of Cotonou and the various prevention and treatment interventions that have been implemented since the early 1990s. I presented and summarised the sources of data for the parameterisation and fitting of a transmission model to the Cotonou HIV epidemic. Next, I mathematically described the transmission model. Finally, I detailed the model fitting process.

In this chapter, I describe the scenarios relevant to the PrEP/TasP impact analysis. I present the results of the model fitting by plotting the simulated epidemic trajectories against demographic, epidemiological and intervention fitting data and cross-validation data. Next, I present the results of the impact analysis of the PrEP/TasP study, as well as scaled-up PrEP and TasP interventions. Finally, I discuss the implications of these results firstly in the wider context of HIV interventions in Cotonou, and secondly with respect to the appropriateness of the transmission model framework, to respond to the aims of this thesis.

The content in this chapter is published in the same article as the content in Chapter 2.

Geidelberg, L., Mitchell, K. M., Alary, M., Mboup, A., Béhanzin, L., Guédou, F., ... & Boily, M. C. (2021). Mathematical Model Impact Analysis of a Real-Life Pre-exposure Prophylaxis and Treatment-As-Prevention Study Among Female Sex Workers in Cotonou, Benin. *Journal of Acquired Immune Deficiency Syndromes*, 86(2), e28.

3.2 Methods

3.2.1 Scenarios

3.2.1.1 PrEP/TasP study scenario

After completing the two stages of model fitting described in Chapter 2, I used the 111 parameter sets obtained and produced the PrEP/TasP study scenario, reflecting what happened during the two years of study, as well as after the study end in 2017 [169]. This specific scenario assumed that HIV testing, PrEP and ART eligibility returned to the pre-study levels after 2 years (at study end), i.e. that PrEP was unavailable, and ART eligibility for pFSW returned to CD4<500 cells/μl. I assumed that pFSW recruited to TasP remained on ART post-study. I simulated the HIV epidemic from 1986-2035 using these assumptions (PrEP/TasP study scenario A1, Table 3.1). The PrEP and TasP arms of the study were also simulated alone (scenarios A2 and A3 respectively) to estimate the relative impact of each. These scenarios (along with all other scenarios) are summarised in Table 3.1.

3.2.1.2 Long-term 20-year intervention scenarios (2015-2035)

I also simulated long-term PrEP and/or TasP "extension" and "scale-up" scenarios, which offered interventions to all eligible pFSW for 20 years (2015-2035) instead of 2 years. Details are available in Table 3.1. The "extension" scenarios maintained over twenty years the same HIV testing, PrEP/ART initiation and dropout rates as the PrEP/TasP study scenario. PrEP and TasP extensions are labelled scenarios B1 and B4 respectively. "Scale-up" scenarios assumed a PrEP or TasP initiation rate twice as high as those observed in the study. PrEP scale-up also allowed PrEP re-initiations, while TasP scale-up also assumed a four-fold decrease in ART dropout rate. PrEP and TasP scale-ups are labelled scenarios B2 and B5 respectively. These scenarios are within limits of what was deemed realistic by local experts (Michel Alary, MD). I assumed no reduction in condom use for those on PrEP, as observed [255].

I further simulated combination scenarios of PrEP and TasP. Scenario B6 is a PrEP extension combined with TasP scale-up. Scenario B7 is a PrEP scale-up combined with a TasP scale-up.

Finally, I also simulated an "optimistic" PrEP scale-up scenario, which assumed a PrEP initiation rate five-fold higher than that observed in the study (scenario B3). This scenario is less realistic according to local experts (Michel Alary, MD), as such it is not considered a

part of the main analysis. However, the impact of this scenario is explored in a sensitivity analysis in Section 3.3.2.3.

In all long-term intervention scenarios, pFSW HIV testing equalled the PrEP/TasP study rate. When modelling PrEP alone, I assumed that ART eligibility remained at CD4<500 cells/μl, and that PrEP daily adherence patterns were either (i) equal to the 2-year study or (ii) "perfect" adherence. For long-term TasP intervention scenarios, all HIV-infected pFSW were eligible for ART for 20 years.

3.2.1.3 Counterfactual scenarios

The main counterfactual scenario (scenario C0) simulated the HIV epidemic over 1986-2035 without PrEP or TasP among pFSW. I assumed that all parameters (including HIV testing, ART initiation and dropout rates) remain constant from 2015 onwards, PrEP never became available, and ART eligibility remained at CD4<500 cells/μl (reflecting guidelines in 2015) for pFSW until 2035. Despite all HIV-infected key populations later becoming eligible for ART [169], CD4<500 was the "standard of care" at the time of the study, and was therefore used as the counterfactual to estimate the impact of TasP (the combined effect of increased ART eligibility, initiation and HIV testing), to answer my primary modelling objective. I ensured the number of pFSW on ART predicted in the model counterfactual in 2017 was consistent with the observed number within the Beninese health system. Most of the main results in Chapter 3 assume this as the counterfactual scenario.

Finally, I explored an alternative counterfactual scenario (scenario C1), which assumed that all HIV-infected FSW were eligible for ART from 2017 onwards, even if the study had never happened. I repeated all analyses using this counterfactual scenario. This facilitated understanding the impact of increased ART initiation and HIV testing resulting from TasP, but not the effects of changing eligibility. A fuller explanation of this scenario, and the impact it has on the results, are presented as a sensitivity analysis in Section 3.3.2.3.

Table 3.1 Short and long-term scenarios investigated. A) the 2-year PrEP/TasP study and (B) various extended and scaled up 20-year PrEP and TasP interventions.

	Scenario	Increased HIV testing among FSW†	PrEP among FSW		ART among FSW		Median coverage across 111 fits (%)		
			Initiation rate‡	Adherence	Eligibility criteria (from 2015)	TasP initiation & dropout rates‡		2017	2035
C0	Main counterfactual (no PrEP or TasP)	-	-	-	CD4 < 500	-	PrEP	0%	0%
							ART	49%	52%
C1	Alternative counterfactual	-	-	-	Any CD4	-	PrEP	0%	0%
							ART	64%	67%
PrEP/TasP study scenarios (2015-2017)§									
A1	Observed PrEP/TasP study	As observed	As observed	As observed (Table 1E)	Any CD4 (until 2017)¶	As observed	PrEP	9%	0%
							ART	83%	53%
A2	PrEP study arm alone	As observed	As observed	As observed	CD4 < 500	-	PrEP	9%	0%
							ART	As counterfactual	
A3	TasP study arm alone	As observed	-	-	Any CD4 (until 2017)¶	As observed	PrEP	As counterfactual	
							ART	83%	53%
Long-term 20-year interventions scenarios (2015-2035)									
Long-term PrEP alone									
B1	PrEP extension††	As observed	As observed	i) As observed ii) Perfect	CD4 < 500	-	PrEP	24%‡	20%
							ART	As counterfactual	
B2	PrEP scale-up††	As observed	2x higher than observed; allow re-initiations	i) As observed ii) Perfect	CD4 < 500	-	PrEP	37%	47%
							ART	As counterfactual	
B3	Optimistic PrEP scale-up††	As observed	5x higher than observed; allow re-initiations	i) As observed ii) Perfect	CD4 < 500	-	PrEP	53%	69%
							ART	As counterfactual	
Long-term TasP alone									
B4	TasP extension	As observed	-	-	Any CD4	As observed	PrEP	As counterfactual	
							ART	83%	81%
B5	TasP scale-up	As observed	-	-	Any CD4	Initiation 2x higher & dropout 4x lower than observed	PrEP	As counterfactual	
							ART	89%	88%
Long-term PrEP and TasP combined									
B6	PrEP extension†† + TasP scale-up	As observed	As observed	i) As observed ii) Perfect	Any CD4	Initiation 2x higher & dropout 4x lower than observed	PrEP	As B1	
							ART	As B4	
B7	PrEP scale-up†† + TasP scale-up	As observed	2x higher than observed; allow re-initiations	i) As observed ii) Perfect	Any CD4	Initiation 2x higher & dropout 4x lower than observed	PrEP	As B2	
							ART	As B5	

† Increased testing among pFSW applied over the duration of the intervention (PrEP/TasP study scenarios: 2015-2017; long-term scenarios: 2015-2035)

‡ PrEP and TasP initiation rates among HIV-uninfected and HIV-infected pFSW (PrEP/TasP study scenarios as observed: 2015-2016; long-term scenarios: 2015-2035)

§ PrEP and TasP were available between 2015–2017, and impact measured both over 2 years (2015-2017) and over 20 years (2015-2035)

¶ Eligibility returns to CD4<500 in 2017

†† All 20-year scenarios involving PrEP are simulated twice: either assuming adherence as observed or assuming all FSWs adhere perfectly

‡‡ Note long-term PrEP initiation rate assumed constant from 2015 onwards (c.f. in the study where initiation only lasted one year: 2015-2016

3.2.2 Estimating impact

For each model fit and scenario, I estimated the annual number of incident HIV infections and life-years lived (number of individuals at each mid-year) for each risk group.

Infections averted (IA_i^W) were calculated as the difference in cumulative infections between intervention w (CI_i^W) and the main counterfactual scenario (CI_i^C) (Table 3.1).

$$IA_i^W = CI_i^W - CI_i^C \tag{3.1}$$

Life years gained were calculated in a similar fashion as above.

The fraction of infections averted (fIA_w) was calculated as the proportion of cumulative infections in the counterfactual that were averted by intervention w:

$$fIA_i^W = \frac{IA_i^W}{CI_i^C} \tag{3.2}$$

I calculated HIV infections prevented and life-years gained among professional FSW only ($i = 1$), clients ($i = 5$) and in the total population of Grand Cotonou (e.g. $\sum_{i=1}^{i=8} IA_i^W$), measured over 2 year (2015-2017) and 20 year (2015-2035) time horizons; the latter allowed observation of long term impact such as accumulated life-years and prevention of secondary HIV transmissions. I reported the median impact and the 2.5th and 97.5th percentile (95% uncertainty interval (UI)) from all model fits.

3.3 Results

3.3.1 Epidemic trajectory in Cotonou

3.3.1.1 Model fitting up to 2017

After model fitting, simulations based on the 111 posterior parameter sets agreed well with available data from Cotonou. Figure 3.1 shows how the fitted model compares to fitting and cross-validation data, for a series of key targets, including HIV prevalence among pFSW, clients, all men and all women. Model results suggest that the HIV epidemic in Cotonou peaked in the mid-1990s for key populations: at 57.4% (95% uncertainty interval (UI): 52.1-63.2%) in 1996 for pFSW and at 9.4% (6.7-15.2%) in 1997 for clients. The peaks in

HIV prevalence occurred later for other groups, and at lower values: 3.0% (2.2-3.9%) in 2003 among all women and 2.8% (1.9-3.9%) in 2000 among all men. By the end of the PrEP/TasP study (2017), model posterior HIV prevalence among pFSW was estimated at 11% (10-14), with prevalence decreasing in all groups in the long term.

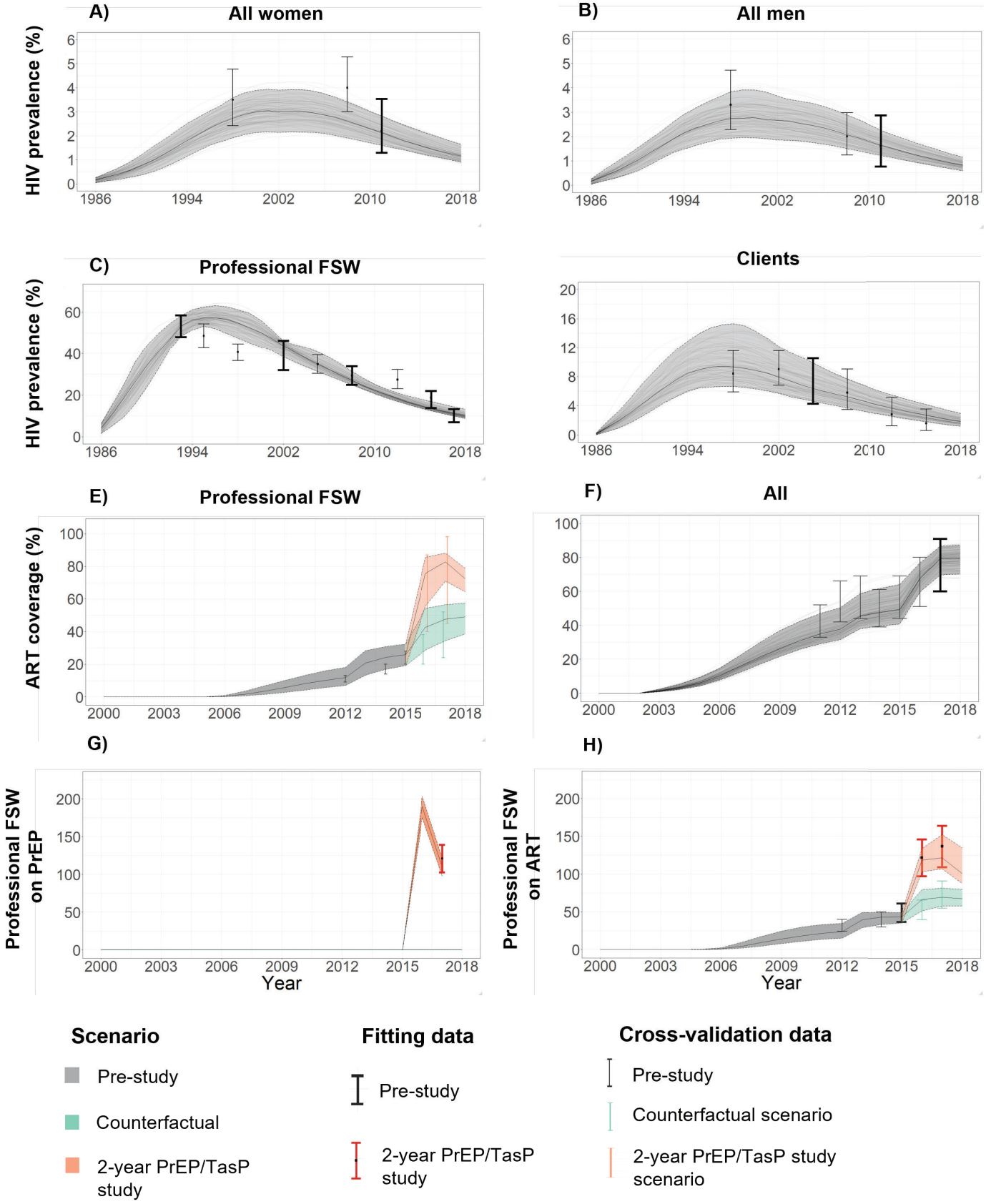


Figure 3.1 Model fitting to key targets up to 2017: The results show model predictions from the 111 posterior parameter sets, in grey, green and orange solid lines (median) and shaded regions (95% uncertainty intervals (UI) across all fits) representing pre-study, counterfactual (no PrEP or TasP) and 2-year PrEP/TasP study scenarios respectively, compared to available fitting HIV prevalence and ART and PrEP coverage data shown in thick bars before 2015 (grey) and during the 2-year PrEP/TasP intervention study between 2015-2017 (orange). Results are shown for HIV prevalence (%) among A) all women, B) all men, C) professional FSWs, D) clients of FSWs; ART coverage (%) among E) HIV-infected pFSW and F) all HIV-infected people; G) number of pFSWs on PrEP, and H) on ART. Additional cross-validation data not used at the fitting stage are also presented in thin bars.

Figure 3.1 also demonstrates that the model closely matched patterns in ART coverage among pFSW and overall, which show a broad increase over time. ART coverage is shown as an absolute number (Panel H) and percentage (Panel E) among pFSW, and the difference in coverage resulting from the PrEP/TasP study (orange) compared to the counterfactual (no PrEP or TasP, green) is clearly observable. Similarly, the difference in PrEP coverage between the two scenarios is shown in Panel G; no PrEP was available in the counterfactual scenario (C0). From 2015-2017, the model replicated closely the observed numbers of pFSW on PrEP and ART (including TasP) in the 2-year study scenario, and the observed number on ART but not on TasP in the main counterfactual scenario (Figure 3.1H). At study end (2017), median posterior ART coverage was 83% of HIV-infected pFSW, of which 34% were those recruited in the TasP arm of the study. Median posterior PrEP coverage was 9% among HIV-uninfected pFSW.

Other fitting targets were also met. Appendix Figure A.1 shows that the expanding population size of Cotonou was well reflected in the model simulations, both those observed in the census and those predicted assuming constant growth rates (Chapter 2, Table 2.7). Appendix Figure A.2 shows the percent of each group in their respective gender (e.g. the percent of women who are pFSW) through time, which agreed with the demographic fitting data. Appendix Figure A.3 shows the modelled number of pFSW in 2012 compared well to data from the Équité en Santé mapping study in Cotonou. Finally, Appendix Figure A.4 shows that the observed ratio of women to men on ART was well reflected in the model simulations after fitting.

Cross validation to data from UNAIDS Reference Group Spectrum model estimates are shown in Appendix Figures A.5 and A.6 for HIV-related deaths and proportion of HIV-infected individuals on ART and virally suppressed respectively. The model simulations for the counterfactual scenario (i.e. no PrEP or TasP, scenario C0) are shown. There is strong agreement between my model simulations and the independent cross-validation data.

Condom use data from IBBS among several risk groups were used in model parameterisation. Appendix Figures A.7, A.8 and A.9 show posterior parameter estimates for condom use by professional FSW, part-time FSW and low-risk women respectively. The parameterisation through time agrees well with the general trend of increasing condom use.

3.3.1.2 Model projections from 2017 onwards

Appendix Figure A.10 shows long-term model projections of the epidemic trajectory in terms of HIV prevalence through time, from 1986 until 2035, assuming the main counterfactual

scenario (no PrEP or TasP). Even if the PrEP/TasP study had not happened, and if all parameters including condom use remained constant from 2015 onwards, my fitted model projections estimate that HIV prevalence would continue to decline in all risk groups.

By 2020, the fitted model predicts that pFSW would have a median HIV prevalence of 8.2% (95%UI: 6.9-11.1%), falling to 3.7% (2.2-5.6%) by 2030. HIV prevalence in clients is an estimated 1.5% (0.9-2.4%) and 0.5% (0.3-0.9%) in 2020 and 2030 respectively, representing a dramatic decline and approaching that of the lower-risk population. Indeed, among all women and men, there is an estimated HIV prevalence of 0.9% (0.7-1.4%) and 0.7% (0.5-0.9%) in 2020 respectively, and 0.4% (0.2-0.6%) and 0.2% (0.2-0.4%) in 2030.

In terms of HIV incidence, the fitted model also estimated a sharp decline through time among all groups. In 2010, there were an estimated 539 (401-760) infections in Cotonou overall, which declined to 97 (57-166) in 2030. This reflects a decrease in HIV incidence of 82%. There were an estimated 642 (441-923) HIV-related deaths overall in 2010, and 160 (118-224) in 2030, representing a decrease of 75%. In the absence of any PrEP or TasP intervention, Cotonou is predicted to not meet the UNAIDS 2030 targets.

3.3.2 Impact analysis

In this section I present the estimated impact of the PrEP/TasP study, and various long-term PrEP and TasP scenarios, among pFSW, clients and overall, either over 2 years or 20 years. All scenarios presented are described in Table 3.1. Note that to determine impact, infections and life-years gained in each intervention scenario are compared to the main counterfactual scenario (scenario C0), unless stated otherwise.

3.3.2.1 Impact of the PrEP/TasP study

In the 2-year PrEP/TasP study scenario, there were an estimated 539 (95%UI: 401-760) HIV infections in 2010, and 93 (56-163) in 2030. This represents an 83% decline in new HIV infections. There were an estimated 642 (441-923) deaths in 2010, and 157 (115-218) in 2030. This represents a 75% decline in new HIV-related deaths.

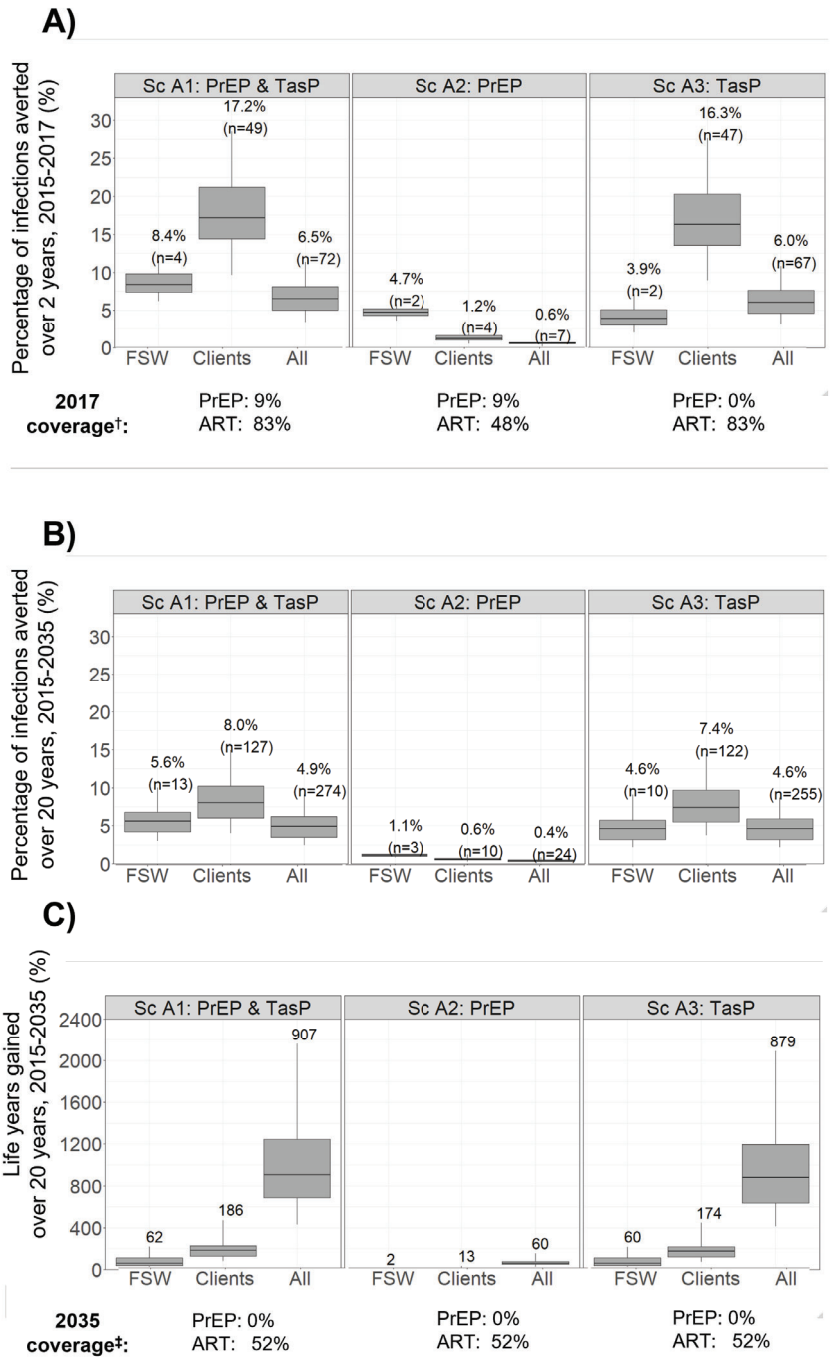


Figure 3.2 Impact of the 2-year PrEP/TasP study for PrEP and TasP combined (scenario A1), PrEP alone (scenario A2) or TasP alone (scenario A3) alone, among professional FSWs ("pFSW"), clients and the whole population ("All") in terms of percentage of HIV infections prevented over (A) 2 years and (B) 20 years, and C) life-years gained (LYG) over 20 years, compared to counterfactual scenario (no PrEP; ART coverage 49% and 52% in 2017 and 2035 respectively). Boxplots represent the median impact (central horizontal line), 25th to 75th percentiles (box) and 95% UI (whiskers) across 111 parameter sets. Labels in panels A and B show median percentage (and number in brackets) of infections prevented; panel C shows median LYG. PrEP and ART coverage represent median posterior values across all model fits for each scenario as shown in Table 3.
† Median coverage of HIV-uninfected and HIV-infected for PrEP and ART respectively at end of the 2-year study.
‡ Median coverage in 2035, 20 years after start of PrEP/TasP study.

Infections averted

Figure 3.2 shows the estimated impact of the 2-year PrEP/TasP study, among pFSW, clients and overall, in terms of infection averted (over 2 and 20 years) and life-years gained (over 20 years). In the two years between 2015-2017, I estimated that the PrEP/TasP study itself (scenario A1) prevented a median of 8% (6-12%), 17% (10-29%), and 7% (3-11%) of new HIV infections among all pFSW, clients, and in the whole population respectively. In absolute terms, this represents 4 (2-5), 49 (27-91) and 72 (41-119) infections prevented (Figure 3.2A).

However, when considering the impacts of PrEP and TasP individually, there are key differences. Over 2 years, the percentage of infections averted among pFSW was similar for the PrEP and TasP arms at 5% (4-6%) and 4% (2-7%) respectively, but among clients it was significantly different at 1% (1-2%) and 16% (9-28%), and overall it was 1% (0-1%) and 6% (3-11%). It is important to note that the median posterior coverage of PrEP and TasP at study end (2017) were 9% and 34% respectively among pFSW; ART coverage including those not on TasP was 83%.

Over 20 years, the percentages of HIV infections prevented by the PrEP/TasP study were lower, as intervention efforts ceased after study completion: 6% (3-10%), 8% (4-15%) and 5% (2-9%) among pFSW, clients and overall respectively (Figure 3.2B). However, due to the cumulative effects of prevention of secondary transmissions over a longer time period, the impact on absolute infections is higher: 13 (6-26), 127 (60-262) and 274 (142-608) among pFSW, clients and overall respectively. When comparing the two pharmaceutical interventions over twenty years, the TasP arm prevented 4x, 13x and 11x more infections than the PrEP arm among pFSW, clients and overall respectively.

Life-years gained

In terms of life-years gained, the impact of the PrEP/TasP study over two years was 2 (1-6) among pFSW, 0 (0-0) among clients and 2 (1-6) overall (not shown); the very small impact is because the 2-year time horizon is not long enough for the mortality effects to accumulate. Over 20 years however, the 2-year study gained 62 (25-219), 186 (80-473) and 907 (430-2157) life-years among pFSW, clients, and overall between 2015-2035, respectively (Figure 3.2C). The relative impact of TasP over PrEP is very clear: TasP gained 32x, 13x and 15x more life-years, among pFSW, clients and overall respectively over twenty years compared to PrEP.

Although TasP broadly had greater impact overall than PrEP, the picture is heterogeneous when considering different key populations. Whether measured over either 2 or 20 years (Figure 3.2A, B), the PrEP study arm prevented a greater percentage of infections among

pFSW than clients or overall, while the TasP arm prevented more infections among clients than pFSW.

3.3.2.2 Impact of long-term PrEP and TasP

In this sub-section I explore the impact of 20 year PrEP and TasP interventions, as described in Table 3.1. To summarise, a PrEP or TasP "extension" is a scenario where I used the same HIV testing, PrEP/ART initiation and dropout rates as were observed in the study, but instead of stopping after 2 years, I extended them for 20 years. A "scale-up" scenario is one where the initiation rate is increased (and dropout rate is decreased for TasP) to achieve higher coverages than those estimated in the extension scenarios.

I present the estimated impact of both PrEP and TasP extension and scale-up interventions alone, while also exploring PrEP adherence. Next, I estimate the impact of combining a PrEP extension or scale-up with a TasP scale-up. After this, I consider more optimistic PrEP scenarios with increased coverage, as well as a sensitivity analysis on the efficacy of PrEP. Finally, I perform a sensitivity analysis on the impact of TasP by exploring an alternative counterfactual scenario (scenario C1).

Impact of 20-year long-term PrEP or TasP interventions alone evaluated over 20 years

Figure 3.3 compares the long-term impact over 20 years of extending and scaling-up PrEP (Figure 3.3A, C), and TasP (Figure 3.3B, D) interventions alone.

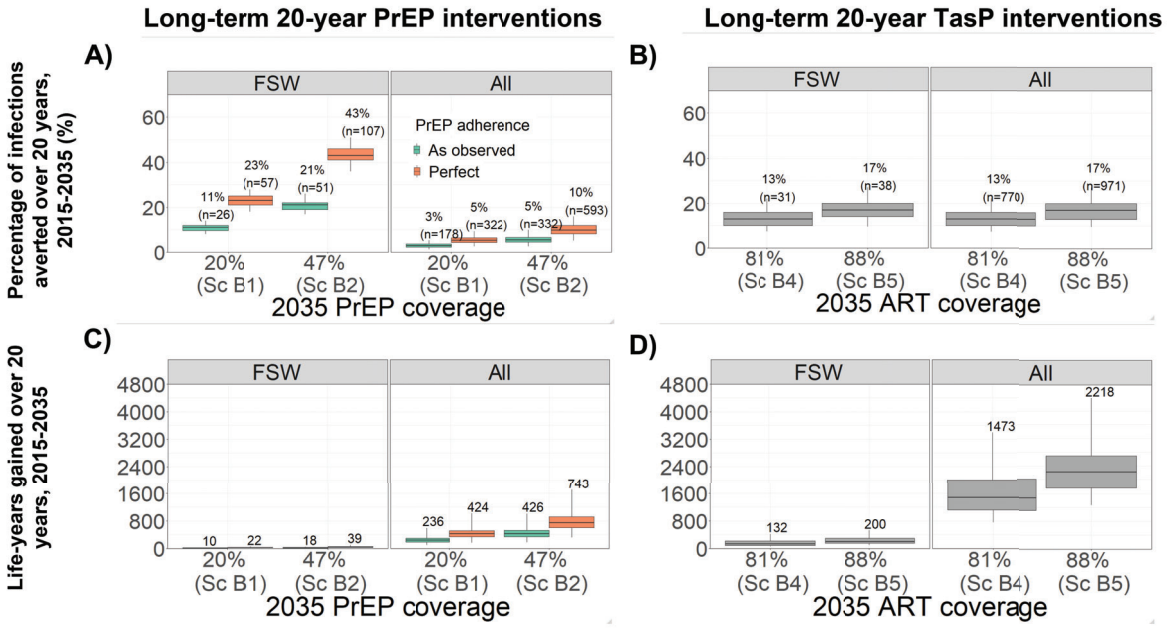


Figure 3.3 Impact of long-term 20-year PrEP (A, C) or TasP (B, D) intervention scenarios separately in terms of percentage of infections prevented (A, B) and life-years gained (C, D) among pFSW and the whole population ("All"), compared to the counterfactual scenario (no PrEP; 52% HIV-infected pFSW ART coverage in 2035). Boxplots represent the median impact (central horizontal line), 25th to 75th percentiles (box) and 95% UI (whiskers) across 111 parameter sets. The PrEP and ART coverages represent median posterior values in 2035 (among HIV-uninfected and HIV-infected pFSWs respectively) across all model fits for each scenario as shown in Table 3. Adherence to PrEP assumed either as observed in the 2-year study (green), or perfect (orange). Labels in panels A and B show median percentage (and number in brackets) of infections prevented; panels C and D show median life-years gained.

Infections averted

Assuming the observed study PrEP dropout rate (including migration), extending PrEP for 20 years achieved a modest 20% coverage in 2035 among HIV-uninfected pFSW (scenario B1), preventing 11% (8-14) and 3% (1-5) of new infections over 20 years among pFSW and in the whole population, respectively, if adherence remains as observed in the study (green bars in Figure 3.3A). Compared to PrEP extension (scenario B1), the percentage of infections prevented over 20 years among pFSW and overall was doubled by either increasing 2035 PrEP coverage to 47% (scenario B2), or assuming perfect PrEP adherence (orange bars in Figure 3.3A), and increased by 3-4 fold to 43% (36-51) and 10% (5-16) in pFSW and overall, by achieving both.

In contrast, extending TasP (scenario B4) achieved a 2035 ART coverage of 81% (compared to 52% in the counterfactual), and prevented 13% (7-22) infections among pFSW and overall (Figure 3.3B). Scaling-up TasP to 88% coverage (scenario B5) prevented 17% (10-27) infections among pFSW and overall. The TasP extension and scale-up scenarios respectively have a 3- & 4-fold higher impact than the 20-year impact of the 2-year TasP study arm (scenario A3), among pFSW and overall.

Similar to the results for the 2-year study, TasP broadly had a larger impact than PrEP overall. Although PrEP and TasP extensions each prevented similar infections among pFSW (11-13%), the population-level impact was 4x higher for TasP (13% vs. 3%). PrEP always had greater impact among pFSW than overall, and TasP always had greater impact overall than among pFSW. Among pFSW, the impact in terms of infections of PrEP was greater than that of TasP if coverage was increased to 47% or if PrEP adherence was assumed to be perfect. However, for the overall impact, even increasing both coverage and adherence to these levels did not make PrEP more impactful than TasP.

Life-years gained

With PrEP extension, 10 (5-24) and 236 (93-596) life-years were gained among pFSW and overall (Figure 3.3C), assuming adherence was as observed in the study. Similar to infections averted, if either PrEP coverage was increased to 47% or if PrEP adherence was assumed to be perfect, then life-years gained were doubled, and increased 3-4 fold if both were achieved.

At 81% ART coverage (TasP extension, scenario B4), 132 (58-419), and 1473 (758-3391) life-years were gained in pFSW and overall respectively (Figure 3.3D), and 200 (95-523) and 2218 (1274-4407) at 88% coverage. Compared to PrEP scale-up scenario (scenario B2: 47% coverage; perfect adherence), the TasP scale-up scenario led to 5x and 3x more life-years gained among pFSW and overall respectively.

Impact of 20-year combined PrEP and TasP scale-ups evaluated over 20 years

Infections averted

A TasP scale-up alone (achieving 88% ART coverage of HIV-infected pFSW in 2035, scenario B5, Figure 3.4A) is estimated to avert 17% (10-27%) of infections among pFSW and 17% (10-27%) overall. Combining this TasP scale-up with a PrEP extension (20% coverage of HIV-uninfected pFSW in 2035) would be substantially impactful among pFSW, with infections averted rising from 17% to an estimated 24% (18-34%) (Figure 3.4A); but not overall, with infections averted rising from 17% to 18% (11-30%) of infections. This

pattern is reproduced with even higher PrEP coverage (47% in 2035), rising to 32% (25-41%) among pFSW and overall to 20% (12-31). Compared to TasP scale-up alone, PrEP and TasP scale-up combined prevented approximately twice the infections among pFSW, but only 18% higher overall.

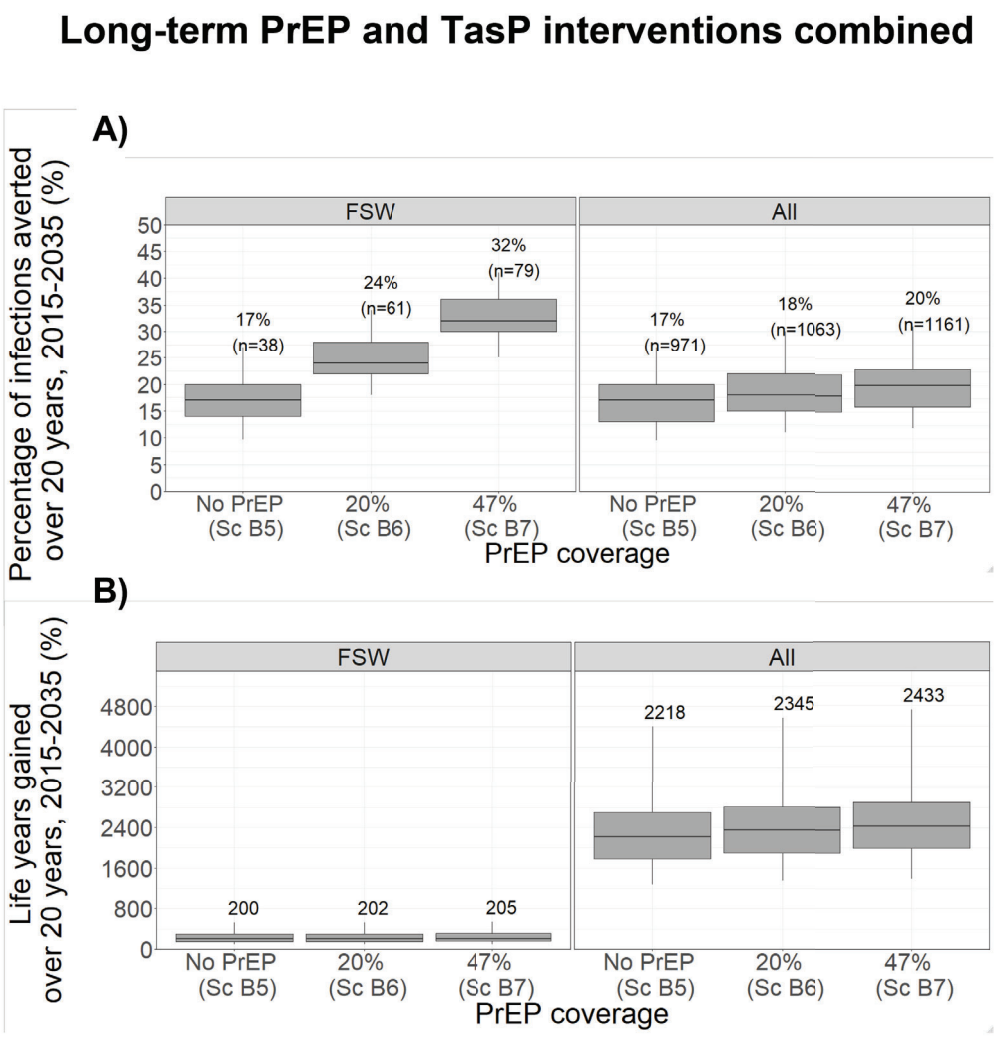


Figure 3.4 Impact of combined 20-year PrEP interventions and TasP scale-up (88% coverage HIV-infected pFSW in 2035), in terms of percentage of infections prevented (A) and life-years gained (B) among pFSW and the whole population ("All"), compared to the counterfactual scenario (no PrEP; 52% HIV-infected pFSW ART coverage in 2035). Boxplots represent the median impact (central horizontal line), 25th to 75th percentiles (box) and 95% UI (whiskers) across 111 parameter sets. PrEP coverages represent median posterior values in 2035 (among HIV-uninfected pFSW) across all model fits for each scenario as shown in Table 3; adherence assumed as observed in 2-year study. Labels in panel A show median percentage (and number in brackets) of infections prevented; panel B shows median life-years gained.

Life-years gained

In terms of life-years gained, the impact of adding PrEP incrementally onto TasP was even more slight. Compared to TasP scale-up alone, PrEP and TasP scale-up combined (scenario B7) only increased life-years gained by 2.5% among pFSW and 10% overall (Figure 3.4B).

3.3.2.3 Sensitivity analyses

Exploring a further increase in PrEP coverage

In this chapter I have estimated the impact of scaling up PrEP to higher coverage (47% of HIV-uninfected pFSW in 2035), which was achieved by doubling the PrEP initiation rate observed in the study. This was deemed a realistic scenario by experts who are familiar with the FSW population of Cotonou (Michel Alary, MD). Despite the fast observed study initiation rate, where 256 HIV-uninfected FSW were recruited in a year among around 900 eligible women (28%), achieving high coverage was difficult due to the high dropout rate (median 54% per year). Doubling the initiation rate therefore only achieved 47% coverage in 2035.

However, as an exploratory scenario, I investigated the potential impact of PrEP if high coverage were achieved (scenario B3). By multiplying five-fold the PrEP initiation rate from the study, and enabling re-initiation of those who dropped out, 69% PrEP coverage could be achieved among HIV-uninfected pFSW. I simulated this scenario with i) PrEP adherence as observed, and ii) perfect PrEP adherence.

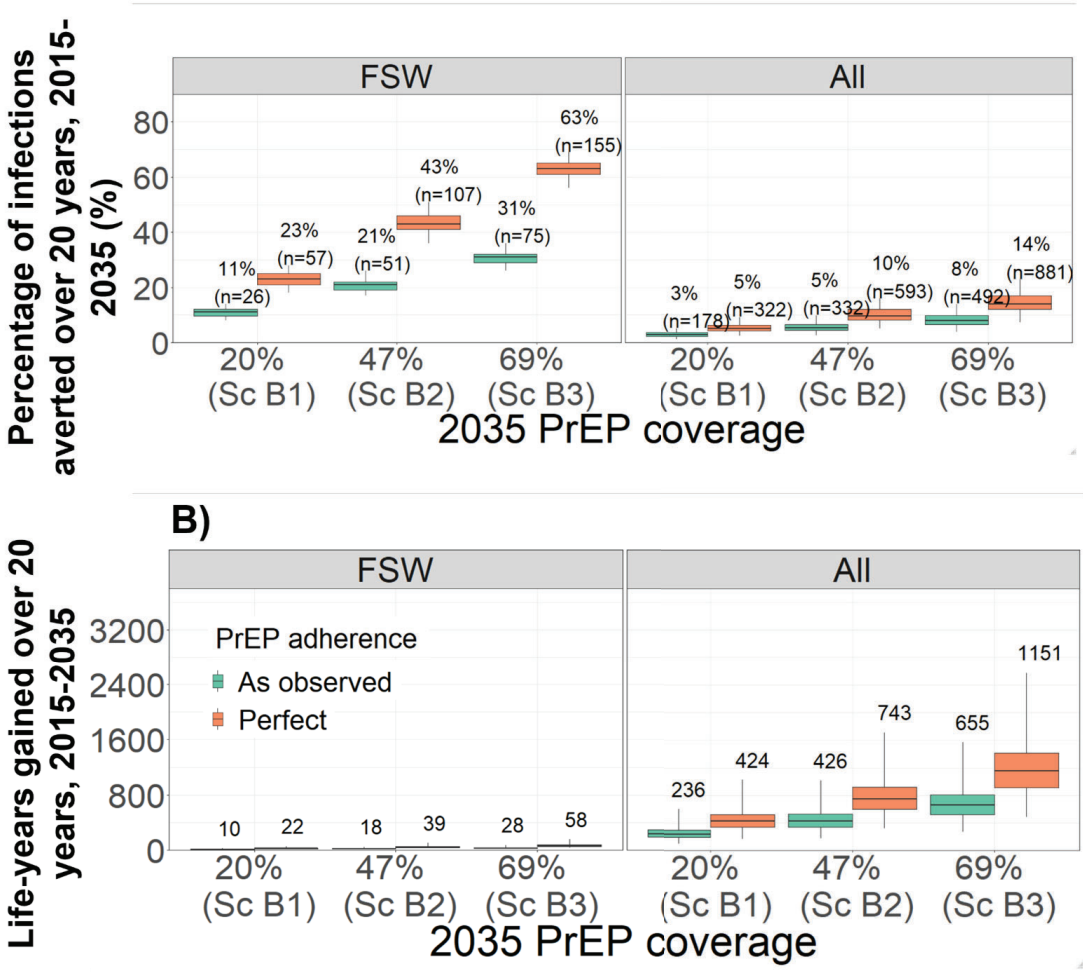


Figure 3.5 Impact of long-term 20-year PrEP intervention scenarios in terms of percentage of infections prevented (A) and life-years gained (B) among pFSW and the whole population ("All"), compared to the counterfactual scenario (no PrEP). Boxplots represent the median impact (central horizontal line), 25th to 75th percentiles (box) and 95% UI (whiskers) across 111 parameter sets. The PrEP coverages represent median posterior values in 2035 (among HIV-uninfected pFSW) across all model fits for each scenario as shown in Table 3, as well as a scenario of 69% PrEP coverage. Adherence to PrEP assumed either as observed in the 2-year study (green), or perfect (orange). Labels in panels A shows median percentage (and number in brackets) of infections prevented; panel B shows median life-years gained.

Figure 3.5 shows the impact of this PrEP coverage compared to the more realistic main scenarios. With adherence as observed, raising PrEP coverage to 69% prevented 31% (26-36%) of infections among pFSW and 8% (4-14) overall, rising to 63% and 14% respectively if adherence were perfect. The increase from 47% to 69% represents a ∼50% increase in coverage, which leads to an improvement of ∼50% on the impact, suggesting a broadly

linear relationship between coverage and impact. However, when considering the overall population impact, even with perfect adherence the number of infections averted is only marginally bigger than when simulating a TasP extension (14% vs 13% infections averted overall), and is less impactful than a TasP scale-up (17%).

With perfect adherence, 69% PrEP coverage prevented 58 (32-153) and 1151 (483-2576) life-years among pFSW and overall, each of which are lower than that achieved by a TasP extension (132 and 1473 respectively). The most optimistic PrEP scenarios, even with a hugely intensified effort for recruitment and perfect adherence, would still have smaller impact than a more realistic TasP scenario.

Exploring different assumptions about PrEP efficacy

As described in Chapter 2, PrEP efficacy was assumed to be dependent on adherence. In summary, I split the on PrEP population into three adherence categories: daily, intermediate and non-adherence, representative of different tenofovir blood plasma concentrations. These three categories were assumed to have efficacies of 90-95%, 30-49% and 0% respectively, derived from the Partners PrEP study (Donnell et al. 2014 [284]). In this sub-section I describe a sensitivity analysis on the efficacy of PrEP for the intermediate category.

In the PrEP/TasP study, intermediate PrEP adherence (by tenofovir concentrations between 0.3 – 40 ng/ml) was recorded in 16.5% of all visits.

Figure 3.6 shows the results of the PrEP extension (Sc B1) and PrEP scale-up (Sc B2) scenarios, exploring different assumptions for PrEP efficacy for the intermediate adherence category:

- 30 – 49% (assumption in main analysis)
- 0% (pessimistic)
- 92% (optimistic)

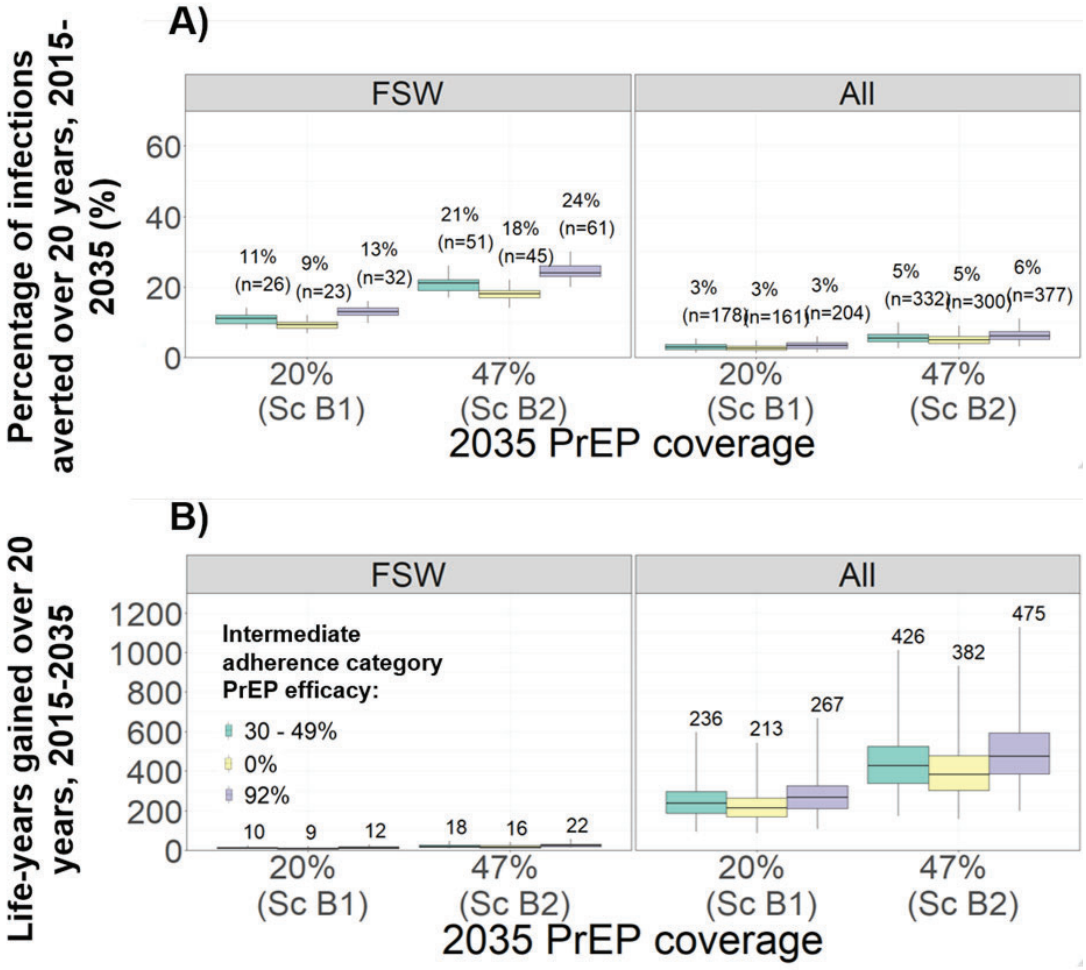


Figure 3.6 Impact of long-term 20-year PrEP intervention scenarios in terms of percentage of infections prevented (A) and life-years gained (B) among pFSW and the whole population ("All"), compared to the counterfactual scenario (no PrEP). Boxplots represent the median impact (central horizontal line), 25th to 75th percentiles (box) and 95% UI (whiskers) across 111 parameter sets. The PrEP coverages represent median posterior values in 2035 (among HIV-uninfected pFSW) across all model fits for each scenario as shown in Table 3. The per sex-act reduction in transmission (efficacy) of PrEP in the intermediate adherence category (0.3 – 40 ng/ml) is assumed to be 30 – 49% (green), 0% (yellow) or 92% (purple). Labels in panels A shows median percentage (and number in brackets) of infections prevented; panel B shows median life-years gained.

In the PrEP extension scenario (scenario B1) with 20% PrEP coverage in 2035, the main analysis assumption prevented 11% infections among pFSW and 3% overall (Figure 3.6A). The pessimistic scenario (0% efficacy in intermediate adherence category) had similar impacts, 9% and 3% among pFSW and overall respectively, while the optimistic scenario had impacts of 13% and 3%. In a similar fashion, little difference was observed between these

three assumptions for the PrEP scale-up scenario (scenario B2). In terms of life-years gained (Figure 3.6B), likewise, there was little variation resulting from the different assumptions on PrEP efficacy for this adherence category.

Even though I explored the effect of varying efficacy of intermediate PrEP adherence from 0 to 92%, I found that the impact of PrEP interventions did not substantially vary. This is because intermediate PrEP adherence was observed in only 16.5% of visits, with the vast majority of blood plasma tenofovir concentration samples being consistent either with daily adherence or non-adherence (undetectable).

Exploring ART eligibility

The main counterfactual scenario (scenario C0) assumed that from 2015 onwards, only HIV-infected pFSW who had a count CD4 T-cell count < 500 cells/mm^3 blood were eligible for ART. This was the standard of care before the start of the PrEP/TasP study in 2015 [271]. The TasP arm of the PrEP/TasP study involved three components: increased HIV testing, increased ART initiation and increased eligibility for ART of all pFSW HIV-infected regardless of CD4 count. The main objective of my analysis (indeed also an objective of the PrEP/TasP study) was to evaluate the full impact of the PrEP/TasP study compared to the standard of care at the time. To assess the impact of all three components of TasP, the counterfactual scenario had to maintain the 2015 eligibility criteria for ART (CD4 < 500 cells/mm^3) from 2015 onwards.

However, following the PrEP/TasP study's completion in 2017, the eligibility criteria changed such that all HIV-infected individuals were eligible for ART. Had I assumed this change in the counterfactual scenario, then the full impact of TasP, including increased ART eligibility, would not have been estimated. In this section I describe a second counterfactual scenario (Table 3.1, scenario C1), whereby ART eligibility is expanded to all HIV-infected pFSW in 2015. I compared incident infections and life-years between this new counterfactual scenario with the long-term 20-years TasP scenarios (TasP extension and scale-up scenarios B4 and B5 respectively), to evaluate the added impact of increased HIV testing and initiation, but without the difference in ART eligibility.

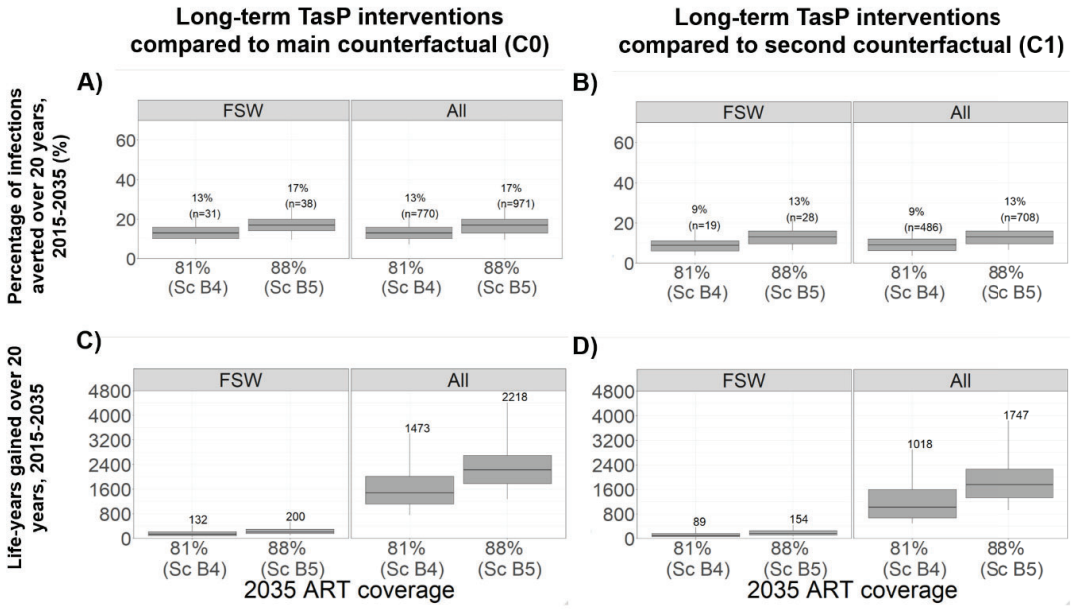


Figure 3.7 Impact of long-term 20-year PrEP intervention scenarios in terms of percentage of infections prevented (A) and life-years gained (B) among pFSW and the whole population ("All"), compared to the counterfactual scenario (no PrEP). Boxplots represent the median impact (central horizontal line), 25th to 75th percentiles (box) and 95% UI (whiskers) across 111 parameter sets. The PrEP coverages represent median posterior values in 2035 (among HIV-uninfected pFSW) across all model fits for each scenario as shown in Table 3. The per sex-act reduction in transmission (efficacy) of PrEP in the intermediate adherence category (0.3 – 40 ng/ml) is assumed to be 30 – 49% (green), 0% (yellow) or 92% (purple). Labels in panels A shows median percentage (and number in brackets) of infections prevented; panel B shows median life-years gained.

I estimated that expanding ART eligibility to any HIV-infected pFSW in 2015 (second counterfactual scenario C1) would lead to a 2035 ART coverage of 67% of HIV-infected pFSW in this new counterfactual scenario, compared to 52% in the main original counterfactual scenario. The impacts of TasP extension and scale-up (to median 2035 ART coverage of 81% and 88% respectively), relative to both counterfactual scenarios, are shown in Figure 3.7. When compared to the alternative counterfactual scenario (C1), TasP extension prevented 9% (4-18%) and 9% (4-17%) infections in FSW and overall respectively, and a TasP scale-up prevented 13% (6-22%) and 13% (7-22%).

The impact of TasP when assuming the alternative counterfactual is around 30% smaller than when assuming the main counterfactual. This can be interpreted as the relative contribution (in HIV infections prevented) of increased ART eligibility to all HIV-infected FSW in the TasP intervention, and 70% as attributable to increased HIV testing and ART initiation. In terms of life-years gained, the relative contribution of increased ART eligibility is 20-25%.

3.4 Discussion

Pharmaceutical interventions are among the many tools in the HIV prevention toolkit [286]. The motivation to treat HIV-infected individuals is driven not only by marked improvements to their mortality and morbidity from viral suppression, but also from the effectiveness of test-and-treat for preventing onward transmission [272]. As PrEP is by design a medication for HIV-uninfected individuals, there is no therapeutic benefit to the patient. Instead, PrEP is a purely preventative measure which, in its various modalities, has been seen as crucial for achieving the UNAIDS prevention target of a 90% reduction in HIV infections by 2030 [287]. Indeed, PrEP has been granted regulatory approval in dozens of countries worldwide [288]. However, clinical trials have shown mixed results for the effectiveness of PrEP among heterosexual women [160, 161], and the real-life evidence of uptake, adherence and continuation in Sub-Saharan Africa was not available at the time of the PrEP/TasP study towards the end of 2014 [289, 290].

To understand the feasibility of implementing PrEP and TasP among professional FSW in Cotonou, the PrEP/TasP study was carried out from 2015-2017 [169]. In this chapter, I performed an impact analysis of the two-year PrEP/TasP study, estimating the numbers of incident HIV infections averted and life-years gained, among pFSW, clients and overall. Further, I estimated the impact of 20-year interventions of PrEP and TasP, either alone or together, with varying levels of adherence and coverage.

3.4.1 What does the epidemic trajectory suggest?

By fitting a transmission model to multiple sets of cross-sectional data, I obtained 111 simulations of HIV prevalence through time. The median and 95% UI of these fits are summarised for groups separately, from 1986-2035, in Appendix Figure A.10. The fitting process provided model simulations which reflected closely the observed data. Note that this assumes no PrEP/TasP interventions, i.e. it is the main counterfactual scenario.

The epidemic trajectories show that the interventions present in Cotonou before the PrEP/TasP study were already highly effective in decreasing HIV burden among all groups. I estimated that the HIV epidemic in Cotonou rose dramatically between the mid-1980s and the mid-1990s. Since 1992, various HIV prevention packages have been implemented, in collaboration between the Canadian government and other funders [254]. Following these interventions, which initially focused on STI and HIV prevention among professional FSW, HIV prevalence slowed down within this group, peaked in 1996, then started to decline. Soon after, in 1997,

HIV prevalence peaked in clients. From the mid-1990s onwards, HIV has continued to decline with every new survey, and is estimated to further decline until 2035 according to my model projections. The combination of condom use, HIV testing and ART initiation has led to considerable progress in responding to the HIV epidemic in Cotonou.

The peaks in HIV for all men and women occurred around the turn of the century, a few years after those among pFSW and clients. In other words, both the rise and decline in HIV prevalence in core groups was followed by the rise and decline among lower-risk groups. This is consistent with the argument that infections in Cotonou among lower-risk groups are mostly seeded from more risky groups involved in commercial sex. A reduction in infections from key populations therefore has secondary benefits for the rest of the population, which supports the impetus to prioritise HIV prevention and treatment among these groups.

Projecting the fitted model to 2030, I estimated that HIV prevalence will continue to decline in the long term. The group with the highest HIV estimated prevalence in 2030 is professional FSW, at 3.7% (2.2-5.6%), compared to all women at 0.4% (0.2-0.6%) and all men at 0.2% (0.2-0.4%).

Within the Sustainable Development Goals set out by the United Nations, UNAIDS and partners aimed to ending AIDS as a public health threat, which was defined as a reduction of 90% in both new HIV infections and HIV-related deaths by 2030, compared to 2010 [131]. My fitted model estimated 539 (401-760) HIV infections in 2010, which declined to 97 (57-166) in 2030, representing an 82% decrease. Assuming no further intervention efforts from 2015 onwards (i.e. no PrEP/TasP study), my model analysis suggests that Cotonou would fall slightly short of the UNAIDS goals for 2030.

3.4.2 What was the impact of the PrEP/TasP study?

I estimated that the PrEP/TasP demonstration study (2015-2017) in Cotonou averted 8% and 6% of HIV infections among pFSW over 2 and 20 years, respectively, and 7% and 5% in the whole population. In the 2-year PrEP/TasP study scenario, I estimated that there would be 93 (56-163) HIV infections in 2030. Compared to 539 (401-760) HIV infections in 2010, this represents an 83% decline. The 2-year study therefore is estimated to bring Cotonou closer to the UNAIDS goals of 90% reduction in HIV incidence for 2030, but would not completely achieve it.

Over 20 years, the study led to an estimated 62 and 907 life-years gained among pFSW and overall, respectively. Unsurprisingly, the PrEP arm had its greatest preventative impact among pFSW, as the direct protective benefits are incurred by the group taking the medication.

In contrast, the TasP arm had its greatest impact among clients due to viral suppression among pFSW directly reducing HIV transmission to clients; the TasP arm prevented an estimated 16% and 7% of infections among clients over 2 and 20 years, respectively.

Overall, there was greater impact (in terms of infections prevented and life years gained) from the TasP arm compared to PrEP for several reasons. First, I assumed that PrEP was unavailable post-study since it was not recommended by national guidelines, while I assumed those on TasP continued treatment (as observed) [169]. Second, TasP had a higher coverage (34% of HIV-infected) than PrEP (9% of HIV-uninfected) after 2 years among pFSW in Cotonou, due to higher uptake and retention; FSW on PrEP had high dropout due to migration and leaving sex work. Third, there was better adherence and thus higher preventative efficacy for TasP than PrEP (averaging 81% pFSW on ART virally suppressed vs. 37% of visits with perfect PrEP adherence) [169, 275, 291]. Finally, TasP reduces mortality [292], thus gaining life-years among HIV-infected pFSW.

3.4.3 What would be the impact of long term PrEP and TasP for FSW in Cotonou?

Similar to the 2-year study results, I predicted greater impact for the extended and scaled-up TasP interventions than PrEP. Compared to a 20-year TasP extension, PrEP extension prevented similar percentages of infections among pFSW (11-13%), but 4-times fewer infections overall; even PrEP scale-up to 47% HIV-uninfected pFSW with perfect adherence had less impact overall. Long-term TasP always gained more life-years than PrEP. Increasing PrEP coverage beyond 50% would be unfeasible according to country-level experts, mainly due to healthcare and funding constraints, and high observed turnover of pFSW in Cotonou [255, 259, 256]. With TasP scale-up to 88% HIV-infected pFSW in 2035, my analysis suggests little additional impact would be gained by simultaneously increasing PrEP coverage to 47%.

In a sensitivity analysis I explored a second counterfactual, whereby ART is available for all HIV-infected pFSW from 2015 (as in long-term TasP scenarios). I found that removing the added effect of increased eligibility reduces the impact of TasP on infections prevented by around 30%, however TasP extension still prevents more infections overall than PrEP, unless 47% coverage and perfect adherence can be achieved. All long-term TasP scenarios still gain more life-years among all groups than any long-term PrEP scenario. This sensitivity analysis suggests that the increased eligibility for ART as a result of test-and-treat policies make a

substantial impact. Further, it suggests that even if ART is universally available, increasing ART coverage is still more impactful than PrEP.

My results suggest that efforts to improve PrEP adherence may be as beneficial as increasing coverage. For example, achieving perfect adherence with a PrEP coverage of 20% had similar impact to increasing PrEP coverage to 47% with study level adherence. The observed PrEP adherence in the demonstration study was low despite the important components of outreach and adherence counselling. However, long-acting PrEP (injectable cabotegravir "CAB LA") was recently shown to have high efficacy among women [172], men who have sex with men and transgender women [293], reducing the barriers to adherence with daily pills [172]. Assuming that women would attend visits for regular injections of CAB-LA, then my perfect PrEP adherence scenario could be interpreted as the impact of long-acting PrEP.

While there is good efficacy of daily oral PrEP under good adherence, the degree of efficacy is less well understood for women who have intermediate adherence [284]. However, given the observed biological concentrations of tenofovir in women in the study, it seems that this group is relatively small: only 16.5% of visits showed a concentration between 0.3-40 ng/ml. My sensitivity analysis showed that the overall impact of PrEP was robust to changes to the assumption of intermediate PrEP efficacy.

Finally, this model analysis shows that the impact of PrEP is hampered by low coverage. This was observed in the study because of poor retention and high mobility, as 53% of PrEP study participants dropped out and of these 58% left either Cotonou or sex work [169]. This is a reality of this key population, who are in unstable and clandestine employment, and whose families may still live in their home nation. Scaled up PrEP interventions among this or similar groups would be more effective when identifying factors associated with better adherence and retention, or when aiding these women in achieving more stable lives.

3.4.4 Comparing impact with other studies

At the time of analysis, this study was the first model impact analysis of PrEP and TasP among FSW reflecting observed real-life uptake, adherence and retention data [276]. Previous modelling studies on the impact of PrEP assumed parameters which were based on clinical trials, not on implementation studies in real-life contexts [281, 19, 294]. Indeed, the investigation of PrEP and TasP adherence with biological tests in this real-life implementation study provided excellent and reliable data on which to perform a model analysis [276].

Although this is the first impact analysis of PrEP among FSW reflecting real-life demonstration study data on uptake, adherence and retention, the results are comparable with

previous studies that explore similar parameter values, despite differences in epidemiological conditions across settings. Assuming 30% annual PrEP uptake and 40% effectiveness, Bekker et al estimated ∼8% infections prevented over 10 years among FSW in South Africa [294]. I estimated that with similar annual PrEP uptake (26%) and average effectiveness (41%, Section 2.3.2.5), my PrEP extension scenario (scenario B1) prevented 11% (8-14%) of infections in pFSW over 20 years. The slight differences are partially explained by the 10% condom use reduction assumed for South Africa, whereas no condom migration was observed in pFSW in Cotonou [291].

Low et al estimated that providing TasP to FSW in Burkina Faso achieving an ART coverage of 79% (compared to 37% in their base-case scenario) would prevent 15% of infections over 20 years overall [281]. This result is comparable to my TasP extension scenario (scenario B4) at 81% long-term coverage preventing 13% (7-22%) of infections overall, slightly lower because my counterfactual scenario has an estimated 52% long-term ART coverage. It is important to note that while these other studies explore the impact of PrEP and TasP under many coverage and adherence scenarios, my analysis estimates long-term coverage based on either observed, or plausible increases in, uptake, retention and adherence. Nevertheless, the comparable results between all studies underlines the robustness and generalisability of the conclusions, that if FSW have PrEP and TasP adherence and retention as observed in Cotonou, the population-level impact of TasP is higher.

3.4.5 Strengths and limitations

Data availability

There are several strengths to this modelling analysis. First, I had access to a large amount of detailed data (see Chapter 2) including censuses [242–245], size estimate studies [257, 255, 254, 138], and house-hold based surveys [280]. Importantly, I also had access to several rounds of historical cross-sectional data among several risk groups, including IBBS among pFSW from 1993-2017 [254, 262, 258, 259] and their clients from 1998-2017 [255, 265, 283]. These studies informed demographic and sexual behaviour parameters, and HIV prevalence estimates over time for all groups. The modelled trajectories reflected empirical data from Cotonou and therefore were an accurate representation of its demography and HIV epidemiology. Other data were used for cross-validation for further confidence. Second, the PrEP/TasP study data provided estimates for key PrEP and TasP parameters including uptake, retention and individual biological measures of adherence every 6 months: blood

plasma tenofovir concentration for PrEP, and viral load for TasP [169]. The data used in this model analysis included outreach, adherence counselling, migration, leaving sex work, and other legal and structural barriers to PrEP/TasP use among pFSW. As such, this analysis is significantly more realistic compared to previous studies which did not reflect such data [281, 19, 294].

Limitations to this analysis are mostly due to data uncertainties. Self-reported data used to inform the sexual behaviour parameters may have been subject to social desirability or recall biases, although these would be mitigated by assigning wide prior parameter ranges capturing uncertainty. I fixed adherence levels over 20 years in PrEP extension/scale-ups, which ignores the observed decrease in adherence over the 2-year study. However, this is justifiable since an extended PrEP intervention would allow more new initiations and dropouts, potentially maintaining a stable overall adherence pattern. There remains substantial uncertainty about the efficacy of PrEP for women especially at intermediate adherence levels, based on clinical trial data and low biological availability of tenofovir in vaginal tissue [163]. I performed a sensitivity analysis showing that this uncertainty did not influence impact substantially, as intermediate adherence represented only 17% of samples. Finally, although I estimated the impact of PrEP and TasP among professional FSW in Cotonou, I did not include a cost-effectiveness analysis. Such an analysis would be important to consider in a resource-limited setting such as Benin, and is included in possible further work (see Chapter 7).

A further source of bias may come from specific parameters which are difficult to estimate. One example of this is the average duration each pFSW spends in active sex work. In the questionnaires, the women are asked how long they have been paid for sex. The median length was assumed to be the lower bound for the duration. It is however impossible to receive a response from a pFSW who has already retired from sex work. I assumed therefore as an upper bound that they never retire from sex work (or that the rate of leaving was equal to zero). This could potentially lead to a bias of over-estimating the duration of sex work, as the uniform prior distribution extends to decades. However, the model analysis is likely robust to this, as the number of pFSW in the model is forced to carefully reflect survey estimates. In other words, if my posterior duration of sex work was overestimated, then the number of pFSW would be overestimated, and the model fitting process would discard that run.

Another parameter that is difficult to estimate includes the date of HIV introduction into Benin. This can have important implications on the epidemiological trajectory. Although I did not vary the date of seeding, I sampled HIV prevalence in 1986, separately for professional FSW and for other groups. The HIV prevalence among professional FSW at seeding of

my model in 1986 was sampled between 1.3% and 6.6% according to the available data (Table 2.1). A lower and higher prevalence at this time may implicitly suggest a later and earlier introduction respectively. I also varied the transmission rate and HIV prevalence of incoming migrants. Therefore, although I did not explicitly sample different dates of introduction, the uncertainty surrounding the early years of the epidemic in Cotonou was explored. Nevertheless, there are limitations regarding the uncertainty for the date of initial HIV introduction.

A final source of bias may come from the fitting and cross-validation data. As discussed in Chapter 2, the data in which local experts had most confidence was used as fitting data, while the rest was used as cross-validation. Finding a representative sample of part-time FSW to survey is difficult, as they work in bars, restaurants, and a number of different venues. Further, as sex work is not their main source of income by definition, they have numerous other occupations and represent a diverse group of women. As such, estimates of HIV prevalence were not used as strict fitting data. The posterior epidemic trajectories estimated HIV prevalence below the observed data (Figure A.10). However, this is not likely to be a serious limitation. Firstly, these data points are not very reliable due to the aforementioned difficulties in defining and sampling this group. Secondly, the important research questions focused on professional FSW, for whom accurate model representation is more important and was maintained.

The transmission modelling framework

The transmission model framework itself comes with many benefits. The numbers of individuals who are susceptible and infected are counted and calibrated to data such as HIV prevalence; as a result, the model can output predicted incident infections and life-years (or deaths) which, when analysing the impact of interventions, is a tangible and useful outcome for policymakers. Further, with the inclusion of secondary transmissions, the full impact of PrEP and TasP can be evaluated. This is particularly important when investigating interventions that are given to core groups, under the understanding that this may indirectly protect other populations. For answering my specific research questions, in terms useful to policymakers, the transmission model framework is well suited. Indeed, my model analysis suggests that the PrEP/TasP study will bring Cotonou slightly closer to meeting the 2030 UNAIDS goals for ending AIDS as a public health threat; explicitly modelling this scenario and the past historical epidemic was vital for achieving this result. It is worth noting that

other modelling frameworks, including some presented later in this thesis, cannot produce outcomes such as disease incidence or prevalence.

Chapter 2 outlined the rich set of data included in the analysis. Indeed, this was only possible because the transmission model framework was able to incorporate such data. For example, sexual partnerships were explicitly modelled, so numbers and types of partnerships according to various surveys could be directly used as parameters. The same is true for several other important parameters relating to sexual behaviour, such as probability of transmission per sex act, condom use, condom efficacy, HSV-2 infection, and many more. In summary, a transmission model can be carefully shaped to input several sources of data of many varieties, and in doing so, combines all of their information to create a simulated epidemic closely resembling reality.

Furthermore, the parameterisation of rates between compartments not only permits utilising data such as PrEP and ART initiation and dropout rates, but also provides a convenient framework with which to simulate scenarios. For example, I calculated the PrEP initiation rate from the observed number of individuals who began PrEP in the study, and the number of susceptible professional FSW who were eligible. Based on this rate, I generated the PrEP study scenario, as well as several long-term PrEP scenarios. I also generated a counterfactual scenario, where the PrEP initiation rate was set to zero. By comparing intervention scenarios to the counterfactual, and counting the number of incident HIV infections, the impact of PrEP could be evaluated, akin to measuring the difference in outcomes between two arms of a clinical trial. The comparison is made clear in Figure 3.1, where the orange and green shaded areas represent the study and counterfactual scenarios respectively. The transmission model framework simulates not only the scenario that happened in reality, but also a control scenario against which to compare it. It is again worth noting that other modelling analyses presented later in this thesis, such as those using phylodynamic inference, are not capable of estimating the impact of interventions by simulating two scenarios for comparison.

Indeed, evaluating the population-level impact of the PrEP/TasP study is made possible by this dynamic transmission model framework. For practical and ethical reasons, there were no control arms in the study, and so while incidence (and other outcomes) can be measured in those in the study, there was no other ideal group against which to compare them. It is possible to compare the women who initiated PrEP with those who refused it or who were ineligible, but this is likely to be incurring bias as the two groups of women were not randomly assigned.

Finally, as discussed in Chapter 1, individual-based models are able to capture stochastic events that are not possible with a deterministic transmission model such as the one I

used. However, the large population sizes of individuals in Cotonou, coupled with the very large number of parameters for various risk groups and through time, the computational requirement for an individual-based model would have been too large to be practically possible. Therefore, the deterministic compartmental model implemented is justified.

In Chapter 7, I compare the strengths of my approach in this chapter for estimating the impact of interventions, against others described in other chapters. I also discuss future work that could add to the results of this analysis.

3.4.6 Conclusions

My model analysis suggests that the 2-year PrEP/TasP study had a modest impact among pFSW and overall in Cotonou over 2 and 20 years, but will aid the city in getting slightly closer to the UNAIDS 2030 goals. The observed migration and poor adherence were key barriers to high PrEP coverage and effectiveness. As a result, the TasP arm was substantially more impactful than the PrEP arm in reducing HIV infections and mortality. Indeed, achieving high PrEP coverage over 20 years is only possible with greatly increased PrEP initiation, and it would still prevent fewer infections and gained fewer life-years overall than a TasP extension, even with perfect PrEP adherence. In resource-limited circumstances, this analysis suggests that scaling up TasP is a more effective population-level strategy in the HIV prevention toolkit, particularly among key populations, for both preventing direct and indirect transmissions and increasing survival. PrEP should be made available where possible within combination prevention strategies to improve individual choices, but strong efforts should be made to ensure good adherence and retention, particularly in highly vulnerable groups such as FSW. The transmission modelling approach employed was very effective at exploiting the rich data set available, carefully representing the real-life epidemic and simulating several scenarios for comparison.

Chapter 4

Estimating the impact of COVID-19 on HIV epidemiology in Cotonou

4.1 Overview

In Chapter 3, I presented a transmission dynamics model analysis of the HIV epidemic in Cotonou, Benin, in which I estimated the impact of the PrEP/TasP study, and long-term PrEP and TasP interventions. In this Chapter, I describe the outbreak of COVID-19 in Benin, three years after the end of the PrEP/TasP study. I discuss the potential disruptions that COVID-19 may have on already-present HIV interventions in Cotonou. Using the fitted model described in Chapter 2 and Chapter 3, I simulate various scenarios of COVID-related disruptions to HIV prevention and treatment services, and to sexual behaviour. I estimate the impact of these disruptions on HIV epidemiology. Finally, I discuss the implications of these disruptions for pharmaceutical and non-pharmaceutical interventions for HIV in Cotonou.

This analysis was originally performed in October 2020, and the results form a significant contribution to the following publication.

Silhol, R., **Geidelberg, L.**, Mitchell, K. M., Mishra, S., Dimitrov, D., Bowring, A., ... & Boily, M. C. (2021). Assessing the potential impact of disruptions due to COVID-19 on HIV among key and lower-risk populations in the largest cities of Cameroon and Benin. *Journal of Acquired Immune Deficiency Syndromes*, 87(3), 899.

4.2 Introduction

4.2.1 The COVID-19 pandemic

As of 17th August 2021, there have been over 200 million confirmed cases of COVID-19 globally, and over 4 million COVID-related deaths; over 187 million confirmed cases have led to a recovery [175]. The COVID-19 pandemic has spread to every country in the world [175]. Governments across the world have implemented various non-pharmaceutical response strategies to control the spread of COVID-19, including enforced social distancing, venue closures, border closures, local or national lockdowns, economic relief packages, among many others [214]. Both the SARS-CoV-2 pandemic, and the international response to it, have had profound impacts on global health, economies and societies [295].

On top of the direct threat of COVID-19 to public health, indirect impacts have also emerged through the disruption of public and private life [296–299]. For example, the governments of many countries ordered the evacuation of hospital beds and cancellation of elective appointments to increase capacity for acute COVID-19 patients [214]. Many healthcare professionals and scientists are therefore concerned about the long-term incidence of other illnesses, for example preventable cancers, for which routine screenings may be delayed or abandoned [300, 301].

Also of concern is HIV, a virus that infects CD4+ T cells [123], which play a large role in the acquired immune system [124]. Untreated HIV infection, characterised by suppressed immunity and vulnerability to opportunistic infections, has a $\sim$2x increased relative risk of COVID-19 mortality compared to HIV-uninfected individuals [302]. The impact of COVID-19 on excess HIV-related mortality is therefore of great concern [303].

Further to this direct impact on health, the disruptive consequences of the COVID-19 pandemic on healthcare systems are also problematic [304]. During the spring and summer of 2020, there were concerns about COVID-related disruptions to HIV prevention and treatment services [305], and the potentially significant indirect effect on HIV transmission and mortality [306]. Sub-Saharan African countries, generally suffering from greater HIV burden and poorer economic and healthcare resilience, were likely to be disproportionately affected [307, 308].

Indeed there were concerns regarding the entire HIV care continuum. Firstly, HIV testing was at risk of being disrupted: the Ministry of Health and Child Care in Zimbabwe estimated that, from April to June 2020, there was a 59% decline in patients who were tested for HIV and received their results [309]. Secondly, there was evidence of reduced attendance at

healthcare clinics: In Kisumu County, Kenya, where the general population HIV prevalence was 17.5% in 2018 [310], average attendance across 60 Ministry of Health clinics decreased by 51% from March to April 2020 [311]. Further, a qualitative study in Uganda revealed that engagement with healthcare facilities during April-May 2020 was hampered by fear of catching COVID-19, and by government restrictions (mostly closure of business and public transport) [312]. Thirdly, there were concerns about ART provision: a survey conducted by the World Health Organisation between April and June 2020 revealed that 36% of African countries were at risk of anti-retroviral disruptions [313], and indeed 7 countries in Sub-Saharan Africa reported stocks of less than 3 months of first-line drugs. Fourthly, ART adherence was also at risk of being reduced: over 10,000 patients living with HIV in Guateng, South Africa missed collection of their treatment medication during lockdown from March to May 2020 [314].

Various responses have been implemented in certain Sub-Saharan African countries to prevent disruptions to HIV services. In Kisumu County, Kenya, the Ministry of Health increased their stocks of ART by ordering extra shipments, and contacted patients to come to the clinic to receive an early refill of a 3-month ART supply [311]. Of the $\sim$28,000 patients who were eligible for this extra supply of drugs, 84% presented at the clinic to pick it up [311]. Further, evidence from South Africa suggests that the networks and health infrastructure rapidly developed for the COVID-19 response could prove useful for the HIV response, with improved medical triage, home delivery of medication and longer refills to aid in adherence and retention [315]. The current HIV healthcare infrastructure, largely comprising in-person clinic appointments, HIV testing, medication retrieval and check-ups, is however still not adapted to cope with the social and legal implications of the COVID-19 pandemic and response [308]. An increase in HIV infections is likely where there are interruptions in diagnosis, delays in treatment uptake, and decreases in viral suppression [306].

Conversely, the physical distancing observed in all countries, either organically developed from fear of infection, or mandated by government interventions, may lead to an interruption of risky sexual behaviours [316], which may in turn reduce incident HIV infections [317]. In Kenya for example, where there were venue closures, curfews and travel restrictions, female sex workers were forced to adapt to different working conditions [318]. Among the strategies were bringing clients to their home, going to the clients' homes, and working outside of 'formal' workday hours. However, the changes in work were accompanied with serious challenges: stigma, police harassment and fines [318]. The result was a reduction in income from reduced commercial sexual contact [318]. There are two potential consequences: it may indirectly decrease HIV incidence through fewer commercial sexual encounters, or it

may increase incidence by compensation with more risky sex [319], or with the association with police harassment and violence [320].

At the time of this analysis, reliable quantitative data describing the indirect impacts of COVID-19 on the HIV care cascade were limited, particularly in Sub-Saharan Africa [306]. In these countries the scarce economic resources, and the prioritisation of acute healthcare and of food and financial security, were a barrier for data collection during a rapidly evolving and chaotic period [321]. Further, patient behaviour and healthcare systems have changed over the course of the year, with a move away from face-to-face interactions. This, coupled with varying paper or electronic recording systems [322] and physically distant staff [313], has led to delays, and unreliable and inconsistent data collection [323]. All of these challenges have resulted in a scarcity of empirical evidence for the impact of COVID-19 on HIV epidemiology.

4.2.2 COVID-19 in Benin

During the year of 2020, the Republic of Benin had a less severe COVID-19 epidemic than other nearby countries. The first confirmed case of SARS-CoV-2 infection was reported on the 16th March 2020 [324], and as of 28th October 2020, the whole country had reported 2557 cumulative confirmed cases and 41 deaths [175]. However, the number of confirmed cases is likely to be an underestimate [325]; at the beginning of the epidemic there was only one laboratory in the country (the haemorrhagic fever lab in Cotonou) capable of carrying out tests for SARS-CoV-2, which ran at a maximum capacity of 300 tests a day [326]. This lack of testing could also limit the ability to attribute deaths to COVID-19, and as such, this figure is also likely to be an underestimate. Indeed in March 2020, healthcare capacity to cope with acute medical cases was found to be insufficient in Benin [326], as it was in other Sub-Saharan nations [327].

On 29th March 2020, the President, Patrice Talon, explained that harsh measures, including lockdowns, were economically unviable for the country [328], and with the majority of its workforce in informal labour, would be devastating to the livelihood of the population [329]. Instead, the Government responded by declaring a two-week cordon sanitaire around several of the urban agglomerations, including the cities of Cotonou and Porto-Novo, the largest city and capital city respectively, where non-essential travel into and out of these areas was restricted [328]. Further, groupings of larger than 10 people were banned, physical distancing of at least one metre was required, public transport was cancelled, taxis were limited to one passenger, and bars and other venues were closed. Although these restrictions were initially

only meant to last two weeks, they were extended [330]. Mandatory mask wearing was implemented on 27th April 2020 [331], and restrictions were only further relaxed in June 2020, with the reopening of bars, religious venues and public transport [332].

Following the lack of COVID-19 testing capacity early in the epidemic, efforts were made to increase it, such that by October 2020 there were 13 testing centres, with at least one located in each of Benin's twelve districts [326]. As of 23rd October 2020, around 8,000 tests were carried out each week, with a cumulative total of more than 222,000 [326]. The European Investment Bank provided 30 million Euros to supply health centres in Benin with medical equipment to care for severe acute cases of COVID-19 [333].

However, at the time of analysis there was great uncertainty regarding the extent of disruption to HIV prevention services and sexual activity in Benin. Although stocks of ART were maintained throughout 2020 (personal communication: Michel Alary, MD and Luc Béhanzin, PhD), data on initiations, adherence and retention following the pandemic were unavailable.

Although general population data was very limited, some data was available for FSW in Cotonou. A survey was conducted in October 2020 among 407 FSW, investigating the impact of COVID-19 on sex work and related activities [334]. This survey was embedded within a wider study investigating the feasibility and implementation of HIV self-testing [335]. FSW were recruited from 31 sex work sites known to local researchers to achieve a suitable number of individuals for the study (around 400). The FSW were interviewed face to face, and asked questions about their lives during the COVID-19 epidemic, with respect to sexual behaviour and HIV services [334].

Among these women, 47% had no other source of income outside of sex work. The definition of a professional FSW, as described in Chapter 2, is a woman whose main source of income is sex work. Although slightly less than half of the surveyed women declared to be exclusively sex workers, it is possible that the others still meet the definition of professional sex worker, despite having other sources of income.

The women surveyed described two main changes to the nature of sex work. Firstly, the vast majority (88% n=358) reported that clients were much harder to find since the outbreak of COVID-19. Secondly, a large minority (36% n=145) revealed a reduction in the price of each sex act, presumably a result of lessened client demand. A smaller number of FSW reported a deliberate reduction in the number of clients (8%) and a changing in sexual position for the sex act (7%). A small minority (7%) found that COVID-19 had no impact on their commercial sex work. Although some have speculated that FSW may switch to riskier (and more remunerative) sex as a result of fewer clients [319], only 0.5% of the FSW in Cotonou reported using fewer condoms or opting for anal sex. It is however important to note

that reporting on such activities may have a social desirability bias, as it was a face-to-face interview; such a bias has been previously observed in Cotonou [336].

With regards to HIV prevention and treatment services, COVID-19 proved to have an impact for FSW in Cotonou. Twenty-two percent of FSW attended fewer HIV/STI awareness sessions, and 21% saw a reduction in the availability of free condoms. Further, 12% and 9% of women reported a decrease in attendance at clinic visits and HIV testing appointments, with the most common reasons being a lack of funds and fear of contracting COVID-19. Among the 399 women who responded to the question, only 7% filled in a new prescription in the pharmacies in the last 6 months, suggesting a very significant or near total pause in new ART initiations. Among HIV-positive women already on ART however, only 4% declared a pause, although 15% did not respond to the question. The clinics reported no lack of ART following the first cases of COVID-19, due to a resupply in March 2020, although medications for treating other STIs were not fully stocked (personal communication, Michel Alary, MD).

4.2.3 Possible impacts of COVID-19 on HIV epidemiology in Cotonou

Although there was limited data demonstrating general disruptions to HIV prevention services in Cotonou, the unexpected arrival of the epidemic, as well as governmental restrictions on movement, are likely to have had effects all along the HIV care cascade, as has been observed in other Sub-Saharan countries [309, 311–314]. Indeed the FSW survey in Cotonou has indicated a decrease in condom distribution, clinic visits, HIV testing and ART initiation among these women [334]; if these disruptions reflect a broader pattern then an increase in HIV infections (and later deaths) can be anticipated.

On the other hand, the cordon sanitaire, and more importantly the closure of bars and venues, are likely to reduce commercial sexual activity in Cotonou, as these are locations in which professional and part-time FSW meet clients [258, 259, 256]. Commercial sex partnerships are disproportionately important for the HIV epidemic of Cotonou [50], as FSW and their clients represent crucial roles as a ‘core’ and ‘bridging’ groups that drive the epidemic [137]; an estimated 52% of all HIV infections between 1993-2008 were attributable to commercial sex [50]. Despite the successful condom promotion in Cotonou since 1993 (“SIDA 1/2/3” interventions), HIV prevalence remains significantly higher among FSW: HIV prevalence in Cotonou among professional and part-time FSW were 10% and 7% in 2017 respectively, compared to 1% in the general Beninese population in 2015 [256, 253, 277]. If commercial sex activity decreases following venue closures, this could reduce new incident HIV infections

among key populations, and therefore reduce overall incidence from secondary transmissions [276].

However, regarding commercial sex, other less favourable scenarios are possible. The recently reduced availability of paying sexual partners and encounters following restrictions may force FSW to engage in higher paid (but riskier) sex without a condom, which carries a much higher risk of HIV acquisition [65]. Further, clients who normally frequent part-time FSW in bars or clubs may instead bring their custom to professional FSW, who could thereby be exposed to new and risky partnerships.

4.2.4 Previous studies estimating the impact of COVID-19 on HIV epidemiology

Early modelling studies have investigated the potential impact of the pandemic on HIV in Sub-Saharan Africa. Hogan et al. modelled the effects of disruptions to heath services for HIV, tuberculosis and malaria in low- and middle-income countries (South Africa, Malawi, Mozambique, Zimbabwe, eSwatini, Lesotho, Uganda, Kenya, Botswana, Tanzania, Cameroon, Côte d'Ivoire, Nigeria and Sub-Saharan Africa as a whole) [337]. They simulated several scenarios of COVID-19 response, ranging from no action to total suppression for 1 year. For HIV-related activities, they assumed that COVID-19 suppression would involve reductions in new ART initiations, viral suppression, PrEP enrolment, voluntary medical male circumcision (VMMC), and sexual partnerships. In their worst-case scenario, they estimated that there would be 10% more HIV-related deaths over 5 years as a result of these disruptions.

Jewell et al. presented a study of five separate models of HIV epidemiology simulating a range of scenarios related to COVID-19 and the COVID-19 response [306]. These included interruptions to VMMC, condom availability, prevention of mother-to-child transmission (PMTCT) programmes, HIV testing, ART initiation, viral load testing and ART retention. They estimated that a 6-month interruption of ART supplies for half of people living with HIV would result in nearly 300,000 excess HIV-related deaths in Sub-Saharan Africa over the course of a year. An equal disruption to condom programmes would lead to an increase of around 20% of new incident HIV infections over one year. These impacts are estimated to be partially offset by disruptions to sexual activity resulting from physical distancing.

4.2.5 Motivation and aims

The impact of COVID-19 on HIV epidemiology is complicated, multi-faceted, and depends on the specifics of the local HIV epidemic in question. This chapter focuses on the impact of COVID-19 on the HIV epidemic of Cotonou, Benin. Previous modelling studies have already explored this in several countries [306, 337], but there were two key reasons why a model analysis was important in Cotonou.

Firstly, the epidemic trajectory of COVID-19, and the responses of national governments, have both varied dramatically between countries, even in the same geographical location [338–340]. For instance, the response of the Beninese government was very different to that of Nigeria, despite being a neighbouring country. Further, given that every country's HIV epidemic is unique [136], the same restrictions or disruptions in HIV care or sexual activity could affect places in very different ways. The epidemiological context of Cotonou, distinct not only from other African countries but also from the less densely populated regions of Benin [265, 258, 253], is therefore deserving of its own analysis.

Second, previous studies have reported potential impact of COVID-related disruptions only among the general populations of each country. Given our understanding of the crucial role of key populations in transmission [50], particularly in concentrated epidemics such as in Cotonou [255], and their unique vulnerability experiencing higher levels of stigma and violence [341], it is important to explore the impact of COVID-19 among these groups in particular. Cotonou is an excellent locality within which to do this: the HIV epidemic of Cotonou has been studied for over thirty years [251, 336], providing a wealth of detailed data concerning prevalence, sexual behaviour and interventions, among key populations and low-risk groups.

Therefore, the key research question I aim to answer is, as outlined in the aims of this thesis: what was the impact of disruptions to HIV prevention services and sexual behaviour, caused by the COVID-19 epidemic in Cotonou? I further aim to answer these secondary questions: are key populations disproportionately affected by these disruptions? What can these disruptions tell us about the effectiveness and utility of already-present HIV prevention and treatment interventions?

To answer these questions, I used the transmission model calibrated to the HIV epidemic of Cotonou described in Chapters 2 and 3. I simulated some plausible scenarios reflecting the most likely disruptions to occur in Cotonou based on available data and reports, as well as other exploratory disruption scenarios. I measured impact in terms of changes in

HIV infections and HIV-related deaths among professional FSW, their clients, and in the population overall.

This model analysis was completed in October 2020, as part of a model comparison with a model describing the HIV epidemic in Yaoundé, Cameroon. The other model analysis was performed by Dr Romain Silhol. The exploratory scenarios described in Section 4.3.1.2 were identical between the two analyses to facilitate comparison. This model comparison is published [317]. At the time of publication, it was the first model analysis explicitly investigating the impact of COVID-19 on HIV epidemiology among key populations.

4.3 Methods

Chapter 2 provided the description of the transmission model used for this analysis. I assumed there was no PrEP in Cotonou, as observed [275, 291]. There was otherwise no difference to the model structure. Details on how I modelled COVID-related disruptions are given in Section 4.3.1.3.

Table 4.1 summarises posterior distributions for relevant demographic, biological, epidemiological and intervention parameters, and outcomes from the model fitting in Chapter 2.

Table 4.1 The posterior distributions (median and 95% UI (2.5th and 97.5th percentiles) for relevant parameters in January 2020, following model calibration in Chapter 2.

Population structure	
% FSW among all females	0.30% (0.27-0.38%)
% clients among all males	21.3% (12.6-25.9%)
Sexual behaviours	
FSW:	
Number of commercial sex acts (per year)	648 (194-1546)
Number of non-commercial sex acts (per year)	27 (15-44)
Clients:	
Number of sex acts (per year)	121 (89-162)
Lower-risk populations:	
Number of sex acts (per year)	38 (29-47)
Fraction of sex acts which are with casual partners.	20%
HIV prevalence	
FSW	8.5% (7.1-11.2%)
Clients	1.5% (0.9-2.2%)
Lower-risk females	1.0% (0.8-1.5%)
Lower-risk males	0.6% (0.3-0.9%)
% condom use at last sex	
FSW during paid sex	91.4% (86.0-98.2%)
FSW during unpaid sex	39.9% (21.7-59.9%)
Lower-risk populations	20.7% (16.2–25.5%)
ART and viral suppression coverage in PLHIV	
RR for disease progression and AIDS-related mortality	5.1 (7.2-10.0)
FSW: % on ART	72.0% (63.1-80.6%)
% virally suppressed	57.1% (49.7-66.2%)
Clients: % on ART	58.8% (42.0-73.4%)
% virally suppressed	35.8% (20.2-55.8%)
Lower-risk females: % on ART	74.9% (64.7-83.2%)
% virally suppressed	46.8% (31.4-65.4%)
Lower-risk males: % on ART	68.3% (50.4-80.6%)
% virally suppressed	42.0% (23.3-63.0%)

4.3.1 Scenarios

I explored several disruption scenarios, informed by data, informal reports and expert opinion, to determine the potential impact of COVID-19.

A list of scenarios is presented in Table 4.2. The base-case scenario assumes no COVID-related disruptions, where parameters remain unchanged from 2020 onwards. The rest of the scenarios are divided into two types: i) plausible scenarios and ii) exploratory scenarios. The exploratory scenarios include 'Single disruptions', 'Combined disruptions' and 'Group-specific disruptions'; section 4.3.1.2 provides further details.

Each scenario reflects different disruptions to HIV testing, ART initiation, viral suppression, condom use, or sexual behaviour. The scenarios are explained in more detail in the following sections.

For each of the 111 fits described in Chapter 3, the epidemic was simulated from 1986-2025 assuming parameter sets from each scenario in Table 4.2. All disruptions were modelled to begin on 1st January 2020 for computational reasons, and lasted for either 3, 6, 12 or 18 months, unless stated otherwise.

Table 4.2 Disruption scenarios simulated in Cotonou from January 2020

Scenario number	Description
A0: Base-case scenario	No COVID and no COVID response*
Plausible scenarios	
A1: Data-driven†	- ART initiations paused among all groups
	- 10% reduction in viral suppression among all groups
	- 10% less condom use among all groups
	- 10% fewer casual partners among all groups
	- 10% fewer commercial partners among all groups
A2: Policy-driven†	100% reduction in client demand for commercial sex with part-time FSWs; demand not transferred.
Exploratory scenarios: single disruptions (all groups)	
B1	Reduction in HIV testing by 10%, 25%, or 50%
B2	All HIV testing paused
B3	All ART initiations paused
B4	Reduction in proportion of those on ART who are virally suppressed by 10%, 25%, or 50%
B5	Reduction in all condom use by 10%, 25%, or 50%
B6	Reduction in non-commercial condom use by 10%, 25%, or 50%
B7	Reduction in commercial condom use by 10%, 25%, or 50%
B8	Reduction in number of casual partners by 10%, 25%, or 50%
B9	Reduction in all non-commercial partnerships by 10%, 25%, or 50%
B10	Reduction in all commercial partnerships by 10%, 25%, or 50%
B11	Reduction in number of all partners by 10%, 25%, or 50%
Exploratory scenarios: combined disruptions (all groups)	
C1	- No new ART initiations
C2	- No new ART initiations
	- Reduction in proportion virally suppressed by 10%, 25%, or 50%
C3	- No new ART initiations
	- Reduction in proportion virally suppressed by 10%, 25%, or 50%
	- Reduction in condom use frequency by 10%, 25%, or 50%
C4	- No new ART initiations
	- Reduction in proportion virally suppressed by 10%, 25%, or 50%
	- Reduction in condom use frequency by 10%, 25%, or 50%
	- Reduction in casual partners by 10%, 25%, or 50%
C5	- No new ART initiations
	- Reduction in proportion virally suppressed by 10%, 25%, or 50%
	- Reduction in condom use frequency by 10%, 25%, or 50%
	- Reduction in casual partners by 10%, 25%, or 50%
	- Reduction in commercial partners by 10%, 25%, or 50%
Exploratory scenarios: group-specific disruptions	
D1	Among clients, no new ART initiations and viral suppression halved
D2	Among professional FSW, no new ART initiations and viral suppression halved
D3	Among lower-risk populations, no new ART initiations and viral suppression halved
D4	All condom use between pro FSW and clients halved
D5	All condom use of lower-risk populations halved
D6	Commercial sex in bars stops but client demand not transferred
D7	Commercial sex in bars stops and so clients go to professional FSW in brothels

∗ All parameters remain as they were in 2020

† All disruptions in the plausible scenarios last 3 months

4.3.1.1 Plausible scenarios

I developed two plausible scenarios to best understand what happened in Cotonou during the COVID-19 epidemic in 2020. The first 'data-driven' scenario (Scenario A1) assumes several disruptions based on behavioural and intervention data from a FSW-specific survey in Cotonou, and general population data from other nearby countries. The second 'policy-driven' scenario contains one specific disruption to sexual activity, which reflects the likely effects of the government policy to close bars in Benin.

1. Data-driven scenario

Although local experts report that stocks of ART were maintained throughout 2020 in Benin, there were no data on initiation, adherence and retention. The FSW survey in October 2020 found that only 7% of FSW filled in a new prescription of any kind [334]; I therefore assumed that ART initiations stopped temporarily. The same survey found that 21% of FSW received fewer free condoms due to the COVID-19 response, and 8% report deliberately reducing the number of clients. Based on this, I assume there was a 10% reduction in commercial partnerships, as well as a 10% reduction in condom use during commercial sex acts. For lack of data and for simplicity, I assumed also 10% reductions in non-commercial partnerships and condom use too. In Yaoundé, Cameroon, there was a 10% reduction in viral suppression in 2020 compared to 2019 [52]; I therefore assumed a 10% decrease in viral suppression. Given that restrictions that began in March 2020 ended in June of the same year, I assumed this scenario lasted for 3 months.

2. Policy-driven scenario

Among the COVID-19 response activities announced by the President Patrice Talon on 29th March 2020, the policy to shut bars, clubs and restaurants is notable for having a likely direct impact on commercial sex. As described in Chapter 2, a portion of commercial sex in Cotonou is performed in such venues, mostly by part-time FSW. These FSW engage in less sexual activity than professional FSW, with a lower frequency of condom use [258, 259]. Based on this government policy, this scenario explores the impact of a total cessation of commercial sex by part-time FSW. Given most (88%) of professional FSW experienced fewer clients during the COVID-19 period, I further assume that the clients who could not

find commercial sex partnerships in bars did not transfer their custom instead to professional FSW.

4.3.1.2 Exploratory scenarios

Given that data was limited for Cotonou during this period, I designed several exploratory scenarios which reflect potential disruptions to various HIV prevention services and sexual behaviours. These scenarios were split into three sets: single disruptions affecting all groups, combined disruptions affecting all groups, and group-specific disruptions (Table 4.2). All of these disruptions were implemented for 3, 6, 12 and 18 months, and the level of reduction was explored at 10%, 25% and 50%, unless stated otherwise. For the purposes of clarity I do not present the results of every single combination.

The first set, 'single scenarios' contained those which feature only one disruption applied to all risk groups in Cotonou. An example of this is scenario B1, where the frequency of HIV testing among all groups was reduced.

The second set, 'combined scenarios' were scenarios where two or more disruptions were applied simultaneously. These disruptions were all individually represented in the 'single scenario' set, and as before were applied to all risk groups. For example, scenario C1 involved a reduction in viral suppression, combined with a pause in all ART initiations.

The third set, 'group-specific scenarios', were those where only particular groups were affected. For example, scenario D1 explores the impact when commercial sex partnerships were paused for part-time FSW, but the client demand was instead transferred to professional FSW.

The next section describes how each disruption was modelled mathematically.

4.3.1.3 Modelling the disruptions

HIV testing

I described in Chapter 2 an analysis of historical intervention data in Cotonou, including integrated bio-behavioural surveys, household-based surveys, and demographic and health surveys. A common question in these surveys asked the participant if they have been tested for HIV in the last year.

The following is a summary of how testing was modelled in Chapter 3. I extracted the probability of annual testing for each risk group over time and transformed this probability into a rate using the following equation:

$$\tau_i^{r,s} = -ln(1 - ptest_i) \tag{4.1}$$

Where $\tau_i^{r,s}$ is the rate of testing of group i; $ptest_i$ is the probability of being tested last year of group i.

To model the impact of COVID-related disruptions to testing, I modified this formula to the following:

$$\tau_i^{r,s} = -ln(1 - ptest_i * (1 - D)) \tag{4.2}$$

Where D is the disruption factor due to COVID-19.

A reduction of D to the annual probability of being testing reduces the rate of testing. In this analysis, I explored scenarios where D takes values of 10%, 25% and 50%.

ART initiation

As described in Chapter 2, disease progression and transmission risk are both reduced when an HIV-positive individual is on ART. ART initiation rate is represented by ρ_i^s, varying by group i and stage of infection s. Eligibility of ART depends on the stage of infection and time, which is reflected in the model.

In this analysis, a pause in ART initiation is implemented by setting all groups to be ineligible for ART, thereby setting $\rho_i^s = 0$ for any i and any s, during the period of disruption.

Viral suppression

The effectiveness of ART depends strongly on viral suppression; the proportion of individuals virally suppressed is therefore a key parameter.

As described in Chapter 2, I assume that when a person living with HIV is on ART, the transmission probability per sex act in the force of infection is multiplied by RR_j^{ART}, which takes a value below 1. RR_j^{ART} is calculated as follows:

$$RR_j^{ART} = 1 - VS_j * eff^{ART} \tag{4.3}$$

Where VS_j is the proportion of individuals that are virally suppressed in group j, eff^{ART} is the efficacy of ART when virally suppressed. I adapted this formula to model disruptions to viral suppression:

$$RR_j^{ART} = 1 - VS_j * (1-D) * eff^{ART} \tag{4.4}$$

Where D is the factor of disruption resulting from COVID-19.

Likewise, progression through the stages of infection is slowed by a factor of $prog_{ART}$ for those on ART ($r = 3$) such that:

$$\gamma^{32} = \frac{\gamma^{02}}{prog_{ART}} * VS_j + \gamma^{02}\left(1 - VS_j\right) \tag{4.5}$$

$$\gamma^{33} = \frac{\gamma^{03}}{prog_{ART}} * VS_j + \gamma^{03}\left(1 - VS_j\right) \tag{4.6}$$

$$\gamma^{34} = \frac{\gamma^{04}}{prog_{ART}} * VS_j + \gamma^{04}\left(1 - VS_j\right) \tag{4.7}$$

Where γ^{32}, γ^{33} and γ^{34} are the rates of disease progression of individuals on ART in stages CD4 > 500, CD4 350 – 500 and CD4 200 – 349 cells/mm^3 respectively, γ^{02} , γ^{03} and γ^{04} are the rates of disease progression of individuals not on ART in stages CD4 > 500, CD4 350 – 500 and CD4 200 – 349 cells/mm^3 respectively and VS_j is the proportion of individuals that are virally suppressed in group j. The reduction in rate of disease progression is represented by $prog_{ART}$. I adapted these equations to model disruptions to viral suppression:

$$\gamma^{32} = \frac{\gamma^{02}}{prog_{ART}} * VS_j * (1-D) + \gamma^{02}\left(1 - VS_j * (1-D)\right) \tag{4.8}$$

$$\gamma^{33} = \frac{\gamma^{03}}{prog_{ART}} * VS_j * (1-D) + \gamma^{03}\left(1 - VS_j * (1-D)\right) \tag{4.9}$$

$$\gamma^{34} = \frac{\gamma^{04}}{prog_{ART}} * VS_j * (1-D) + \gamma^{04}\left(1 - VS_j * (1-D)\right) \tag{4.10}$$

Where D is the factor of disruption resulting from COVID-19.

Finally, AIDS-related mortality occurs for individuals with CD4 350-500, CD4 200-249, CD4 <200 cells/mm^3 ($s = 3,4,5$). AIDS related mortality for those on ART ($r = 3$) is slowed by a factor of $mort_{ART}$, such that:

$$\alpha^{33} = \frac{\alpha^{03}}{mort_{ART}} * VS_j + \alpha^{03}\left(1 - VS_j\right) \tag{4.11}$$

$$\alpha^{34} = \frac{\alpha^{04}}{mort_{ART}} * VS_j + \alpha^{04}\left(1 - VS_j\right) \tag{4.12}$$

$$\alpha^{35} = \frac{\alpha^{05}}{mort_{ART}} * VS_j + \alpha^{05}\left(1 - VS_j\right) \tag{4.13}$$

Where α^{33}, α^{34} and α^{35} are rates of AIDS-related mortality of individuals in stages CD4 350-500, CD4 200-249, CD4 <200 cells/mm^3 respectively when on ART, α^{03}, α^{04} and α^{05} are rates of AIDS-related mortality of individuals in stages CD4 350-500, CD4 200-249, CD4 <200 cells/mm^3 respectively when not on ART, and $mort_{ART}$ is the reduction in rate of AIDS-related mortality.

For each parameter set, $prog_{ART}$ and $mort_{ART}$ are equal. I adapted these equations to model disruptions to viral suppression:

$$\alpha^{33} = \frac{\alpha^{03}}{mort_{ART}} * VS_j * (1 - D) + \alpha^{03}\left(1 - VS_j * (1 - D)\right) \tag{4.14}$$

$$\alpha^{34} = \frac{\alpha^{04}}{mort_{ART}} * VS_j * (1 - D) + \alpha^{04}\left(1 - VS_j * (1 - D)\right) \tag{4.15}$$

$$\alpha^{35} = \frac{\alpha^{05}}{mort_{ART}} * VS_j * (1 - D) + \alpha^{05}\left(1 - VS_j * (1 - D)\right) \tag{4.16}$$

Where D is the factor of disruption resulting from COVID-19.

Sexual partnerships

Chapter 2 showed that partner change rate is calculated separately for commercial (c_i^C) and non-commercial partnerships (c_i^N). Commercial partnerships are only formed between female sex workers (either professional or part-time) and their clients. Non-commercial partnerships can be formed by any sexually active group in the model.

In this analysis, I investigated the impact of a reduction (D) in annual commercial and non-commercial partner change rates, both separately and together.

It is important to note that at every timestep, commercial partnerships are balanced: the total number of commercial partnerships as declared by women ($i = 1,2$) must match the total number of commercial partnerships as declared by men ($i = 5$), such that the following equation is satisfied:

$$c_1^C(1-D)N_1 + c_2^C(1-D)N_2 = c_5^C(1-D)N_5 \tag{4.17}$$

Where c_i^C is the commercial partner change rate of group i; N_i is the total number of individuals in group i; and D is the disruption factor due to COVID-19. The same procedure was implemented for non-commercial partnerships. A more detailed explanation of balancing of partnerships is available in Chapter 2 Section 2.4.2.7.

This balancing procedure is unimportant when the disruption scenario affects the number of partnerships for all groups equally, as the expression $(1-D)$ will factor out. However, there are scenarios where the number of partnerships for a specific group was reduced.

In Scenario D7, I assumed that client demand is maintained, such that partnerships that would have been formed with part-time FSW were instead transferred to professional FSW. This was implemented by setting the partner change rate for part-time FSW to zero, while balancing the partnerships as usual.

In Scenario D6, I assumed that part-time FSW experienced a pause in commercial sex, but that the client sex demand was not transferred to professional FSW. To implement this, I assumed that the number of sex acts per partnerships between part-time FSW ($i = 2$) and clients ($i = 5$) were temporarily set to zero. This is described in the equations below:

$$n_{12}^C = n_{21}^C = 0 \tag{4.18}$$

The effect of this implementation is that the partner change rates and balancing of partnerships are unchanged, but the force of infection experienced by part-time FSW from clients, and vice versa, was temporarily set to zero.

Condom use

In my model structure, the transmission probability of HIV per sex act is modified by the probability that it was protected by a condom, represented by parameters f_{ij}^{C} and f_{ij}^{N} for commercial and non-commercial partnerships respectively, where i and j represent individuals in the sexual partnership. These probabilities varied by group and through time.

I explored various scenarios of reduction in condom use either by particular groups, or by all groups, where the following expressions were implemented, representing frequency of sex acts protected by a condom for commercial and non-commercial partnerships respectively:

$$f_{ij}^{C}(1-D) \tag{4.19}$$

$$f_{ij}^{N}(1-D) \tag{4.20}$$

Where D is the disruption factor due to COVID-19.

4.3.1.4 Estimating impact

The cumulative numbers of incident HIV infections and HIV-related deaths were calculated in the base-case scenario (Scenario A0: no COVID-19), and for all disruption scenarios. These were calculated over 1 and over 5 years separately for professional FSW, clients and for the overall population. For each of the 111 model fits achieved and described in Chapter 3, impact was evaluated as the relative change in new HIV infections or HIV-related deaths compared to the base-case scenario; the equation below describes this calculation:

$$\%\ change\ in\ HIV\ infections = \frac{HIV\ infections\ in\ disruption\ scenario - HIV\ infections\ in\ basecase\ scenario}{HIV\ infections\ in\ basecase\ scenario} \tag{4.21}$$

$$\%\ change\ in\ HIV\ deaths = \frac{HIV\ deaths\ in\ disruption\ scenario - HIV\ deaths\ in\ basecase\ scenario}{HIV\ deaths\ in\ basecase\ scenario} \tag{4.22}$$

Results are presented as the median and 95% quantiles of the 111 fits (95% UI), calculated either over 1 year (2020-2021) or 5 years (2020-2025), and are presented as a percentage. A negative impact represents a reduction in infections or deaths.

It is important to note that while disruptions may be applied to certain groups or all groups, the impact of every scenario, in terms of HIV infections and deaths, was evaluated separately for professional FSW, clients and overall.

4.4 Results

4.4.1 Plausible scenarios

The data-driven plausible scenario, as described earlier, assumed a 3-month disruption including the cessation of ART initiations, and 10% reductions in viral suppression, condom frequency, casual and commercial partnerships. This scenario increased HIV infections by 6.4% (3.3-13.8%) among professional FSW, 7.3% (4.6-11.4%) among clients and 4.7% (2.8-7.1%) in the overall population measured over 1 year, and 2.1% (1.2-4.3%), 2.4% (1.5-3.7%) and 1.6% (1.0-2.6%) measured over 5 years. The equivalent impact in terms of HIV-related deaths was an increase of 6.1% (4.6-8.3%) among professional FSW, 2.6% among clients (1.3-5.2%) and 3.0% (1.6-6.0%) overall measured over 1 year, and 2.0% (1.3-2.8%), 0.6% (0.4-1.3%) and 0.7% (0.4-1.4%) measured over 5 years. In terms of change in relative HIV infections, the impact of this scenario was greater among clients than among professional FSW, but the inverse was true in terms of HIV-related deaths. For this scenario, whether measured in terms of HIV infections or deaths, or measured over 1 year or 5 years, the greatest impact was found either among professional FSW or clients, rather than overall. The relative impact among all groups of a 3-month disruption was greater when measured over 1 year than over 5 years.

The policy-driven plausible scenario explored the impact of disruptions to commercial sex of part-time FSW, a possible result from the closing of bars and restaurants in Cotonou, which lasted from 29th March 2020 to late June 2020 (around 3 months) [332]. If part-time FSW stopped commercial sex for 3 months, and clients did not seek it instead from professional FSW, there was an estimated decrease of 0.1% (0.0-0.4%), 0.3% (0.1-1.5%) and 0.3% (0.1-0.8%) in new HIV infections among professional FSW, clients and overall respectively. The impact in terms of deaths was <0.001% for all groups.

4.4.2 Exploratory scenarios

4.4.2.1 Single scenarios

Figure 4.1 shows the impact of single 6-month disruptions on professional FSW, clients and overall in terms of HIV infections and HIV-related deaths, measured over 1 year. Increases in new HIV infections were estimated for reductions in ART initiations, viral suppression and condom use, while decreases in infections followed reductions in sexual partnerships. A larger disruption (e.g. 50% reduction) had a greater impact than a smaller one (e.g. 10% reduction).

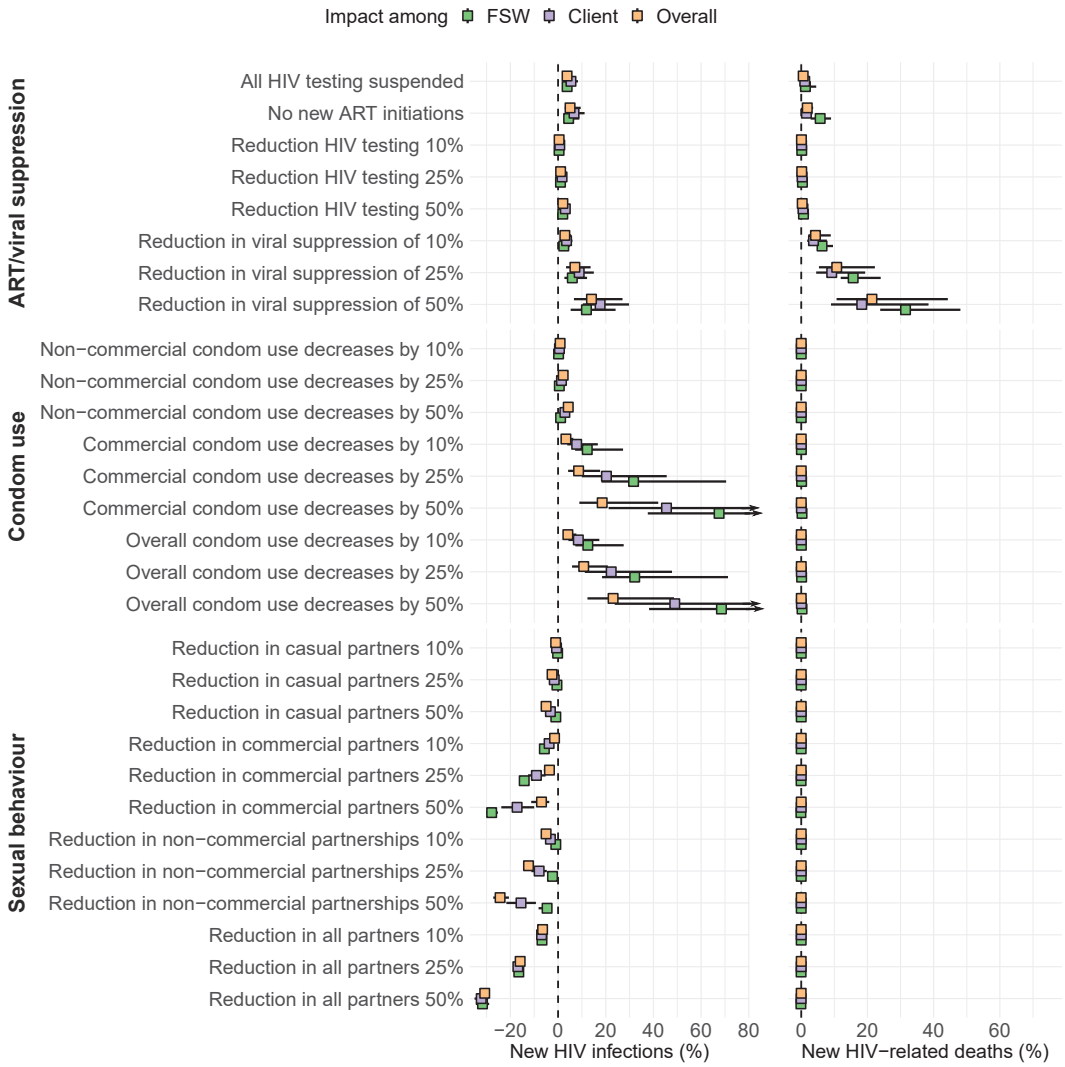


Figure 4.1 Relative change in cumulative number of incident HIV infections and HIV-related deaths over 1 year among professional FSW (green), clients (purple) and overall (orange) under scenarios assuming 6-month single disruptions to all groups, compared to base-case (no COVID-19 nor COVID-19 response) in Cotonou. Squares and lines represent median and 95% percentiles of estimates across 111 model fits.

Broadly, in terms of HIV infections, reductions in condom use led to greater impact than similar reductions in HIV testing or viral suppression; this effect was most pronounced for condom use during commercial sex. Indeed, reducing commercial condom use by 50% for 6 months led to an estimated 67.5% (95% UI: 37.6-158.1%) increase in HIV infections among professional FSW, 45.5% (21.2-110.0%) among clients, and 18.5% (8.9-42.0%) overall. The impact of reducing non-commercial condom use was substantially smaller for all risk groups.

In comparison, reducing viral suppression by 50% was less impactful: leading to a 11.9% (5.3-24.1%), 17.6% (10.3-29.8%) and 14.0% (6.7-27.0%) increase in HIV infections among FSW, clients and overall respectively. Another difference is that while condom use disruptions had a bigger relative impact on professional FSW, disruptions to viral suppression had a bigger impact among clients, because of the increased HIV transmission from professional FSW.

A 50% reduction in commercial partners prevented 27.9% (25.2-30.1%) of new HIV infections among professional FSW, 17.2% (10.0-23.8%) among clients, and 7.0% (3.7-11.3%) overall; the highest impact was among professional FSW and the lowest in the overall population. However, the opposite pattern was observed for non-commercial partnerships: a 50% reduction in non-commercial partnerships prevented 4.6% (2.6-8.2%) of new HIV infections among professional FSW, 15.6% (9.3-21.7%) among clients, and 24.3% (20.6-27.2%) overall. A reduction in all partnerships resulted in a similar impact for all risk groups.

In terms of HIV-related deaths, the greatest impact by far was observed for a 50% reduction in viral suppression, which led to an estimated 31.5% (23.9-48.1%) increase in HIV-related deaths among professional FSW, 18.3% (9.0-38.4%) among clients, and 21.4% (10.7-44.3%) in the overall population. The cessation of all new ART initiations increased HIV-related deaths by 5.7% (2.9-9.0%), 1.6% (0.8-3.4%) and 1.8% (0.9-3.6%) among professional FSW, clients and overall respectively. Disruptions which strongly increased infections, such as those in condom use, did not have time to incur substantial impact in HIV-related deaths when measured over 1 year only. Reducing viral suppression among those already infected, however, caused measurable mortality within 1 year. Other than those related to ART initiation and viral suppression, all single disruptions explored changed HIV-related deaths by <1% over 1 year.

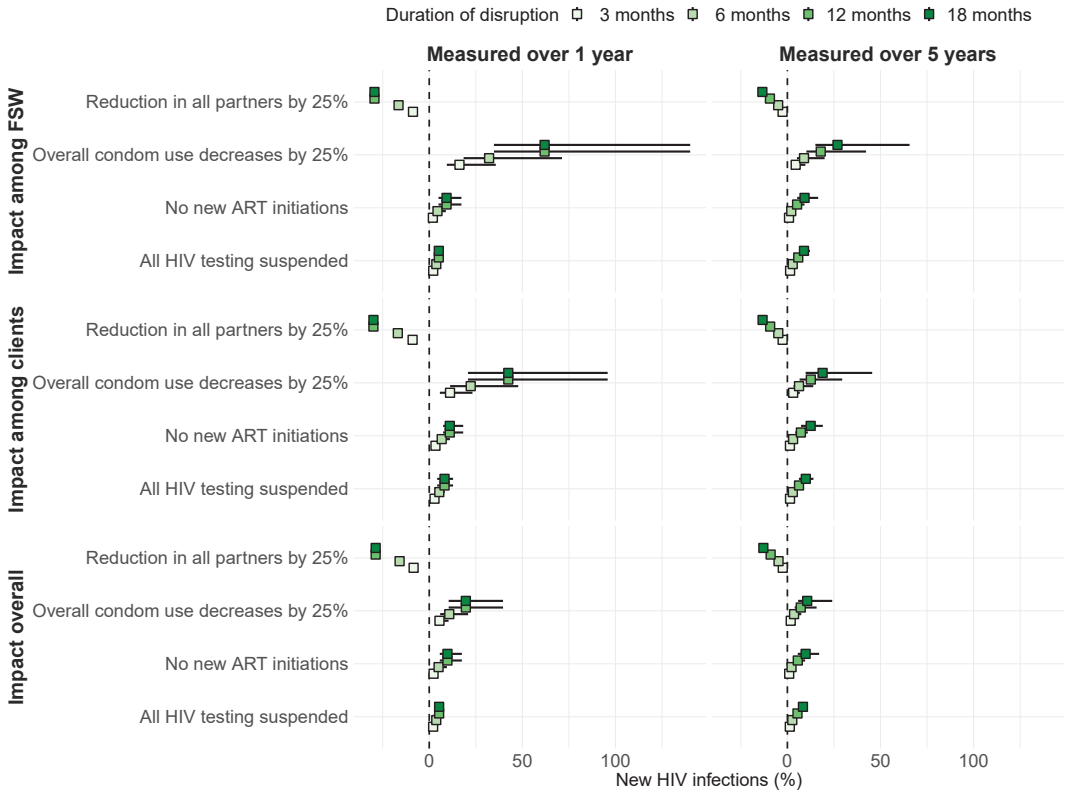


Figure 4.2 Relative change in cumulative number of incident HIV infections over 1 year and 5 years among professional FSW, clients and overall under single scenarios assuming 3 (light green), 6, 12 and 18 (dark green) month disruptions, compared to base-case in Cotonou. Squares and lines represent median and 95% percentiles of estimates across 111 model fits.

Varying the duration of the disruption had implications for the impact. Figure 4.2 shows how disruptions of 3, 6, 12 and 18 months for single scenarios impacted HIV infections among professional FSW, clients and overall. For any duration of disruption, and for all scenarios, the impact was greater in amplitude when measured over 1 year than over 5 years, because a relatively larger proportion of the time under analysis was under disruption. When measuring over 5 years, there was a strongly linear relationship between duration of disruption and new HIV infections. However, when measuring over 1 year, the 12-month and 18-month disruptions had equal impact, as the whole period under study was disrupted in both cases.

Although the relative impact was higher when measuring over 1 year, the absolute impact was higher over 5 years. For example, a 25% reduction in overall condom use for 12-months increased HIV infections among professional FSW by 62% (35-140%) over 1 year, and 18%

(10-42%) over 5 years (Figure 4.2). However, the absolute number of new HIV infections were 8 (4-20%) and 10 (6-24%) over 1 and 5 years respectively (not shown).

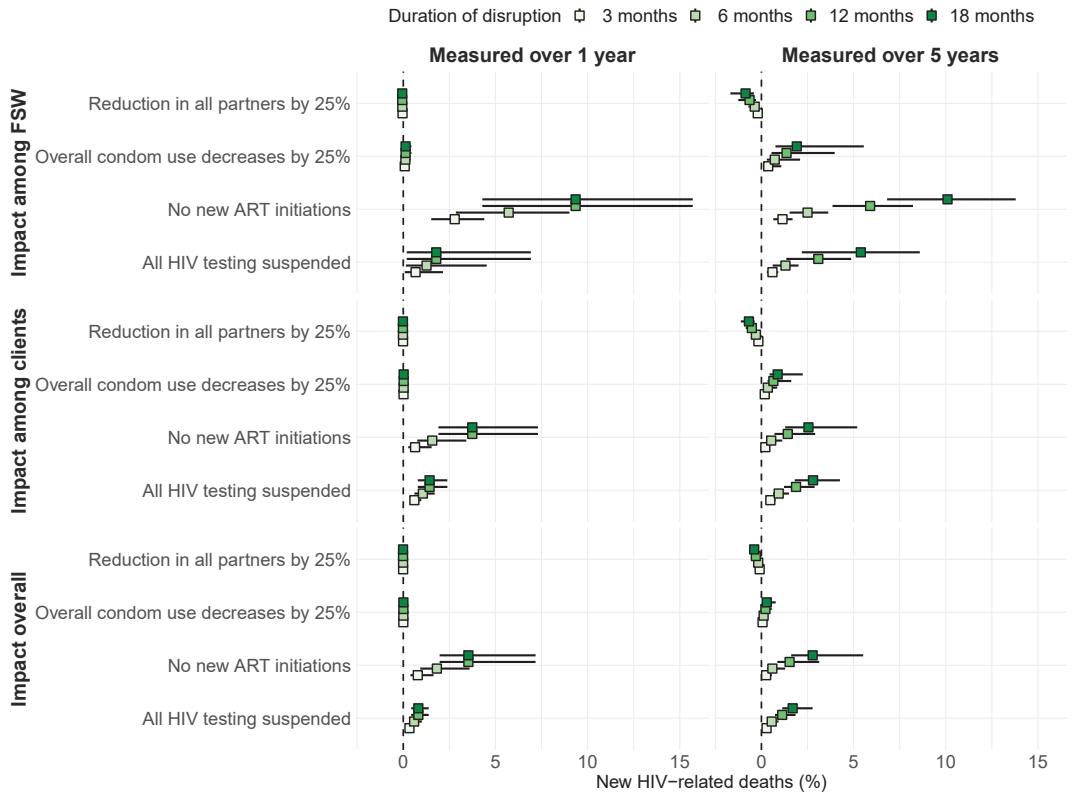


Figure 4.3 Relative change in cumulative number of HIV-related deaths over 1 year and 5 years among professional FSW, clients and overall under single scenarios assuming 3 (light green), 6, 12 and 18 (dark green) month disruptions, compared to base-case scenario (no COVID-19 nor COVID-19 response) in Cotonou. Squares and lines represent median and 95% percentiles of estimates across 111 model fits.

The effects of varying duration of disruption, and the time over which impact is calculated, are shown in Figure 4.3 in terms of HIV-related deaths. Unlike for HIV infections, HIV-related deaths were not always higher when measured over 1 year than 5 years. For example, when reducing overall condom use by 25% for 12 months, the impact was an increase of 0.12% (0.04-0.44%) over 1 year and 1.36% (0.55-3.97%) over 5 years among professional FSW. This was because the deaths incurred by reducing condom use took some years to accumulate. Like with HIV infections, duration of disruption scaled linearly with increases in HIV-related deaths, except for 12 and 18-month disruptions when measured over 1 year which were identical.

To summarise, the single scenario with the worst negative impact depended on the outcome of interest. Disruptions to condom use, particularly for commercial sex, were the most impactful in terms of HIV infections for the time horizons explored. This impact was highest among professional FSW, who rely on high frequencies of condom use to protect themselves and their sexual partners from HIV acquisition. Disruptions to viral suppression also had large impacts on new HIV infections, as the transmission attenuation provided by ART is weakened. The impact of such a disruption was highest among clients, who were more exposed to higher prevalence sexual partners (professional FSW). In terms of HIV-related deaths, disruptions to viral suppression were the most damaging, as this was the only intervention which directly mitigates HIV mortality.

Reductions in sexual partnerships decreased new HIV infections, by reducing the number of opportunities for HIV transmission. The decrease in HIV-related deaths was however very slight for all scenarios reducing sexual partnerships.

Professional FSW and clients tended to experience a disproportionate estimated impact caused by disruptions. This was true not only for reductions in commercial condom use which only affected these groups, but also for more general disruptions such as viral suppression which affected all groups. The relative impact was greater in the overall population compared to key populations only for reductions in casual or non-commercial partnerships, or non-commercial condom use. However, when only looking at infections in the overall population, it is important to note that the impact was still greater with disruptions to commercial condom use than to non-commercial condom use, due to secondary transmissions from key populations into the lower-risk groups.

Longer disruptions had greater impact, and measuring outcomes over 1 year generally returned a greater relative impact than over 5 years, but a smaller absolute impact.

4.4.2.2 Combination scenarios

Figure 4.4 shows the impact of different combination scenarios on HIV infections and HIV-related deaths, calculated separately for professional FSW, clients and the overall population. The combination scenario with the greatest negative impact explored assumed that for 6 months ART initiations stopped, the proportion virally suppressed halved, and the frequency of all condom used halved (scenario C3). This led to an estimated 107% (65-230%) increase in new HIV infections for professional FSW, 94% (52-178%) for clients and 51% (29-90%) overall over 1 year. For both professional FSW and clients, the incremental impact of reduced condom use was far larger than that of both the cessation of ART initiation and reduced viral

suppression. Further, for these two risk groups, the reduction of casual partnerships made little difference compared to the reduction in commercial partnerships.

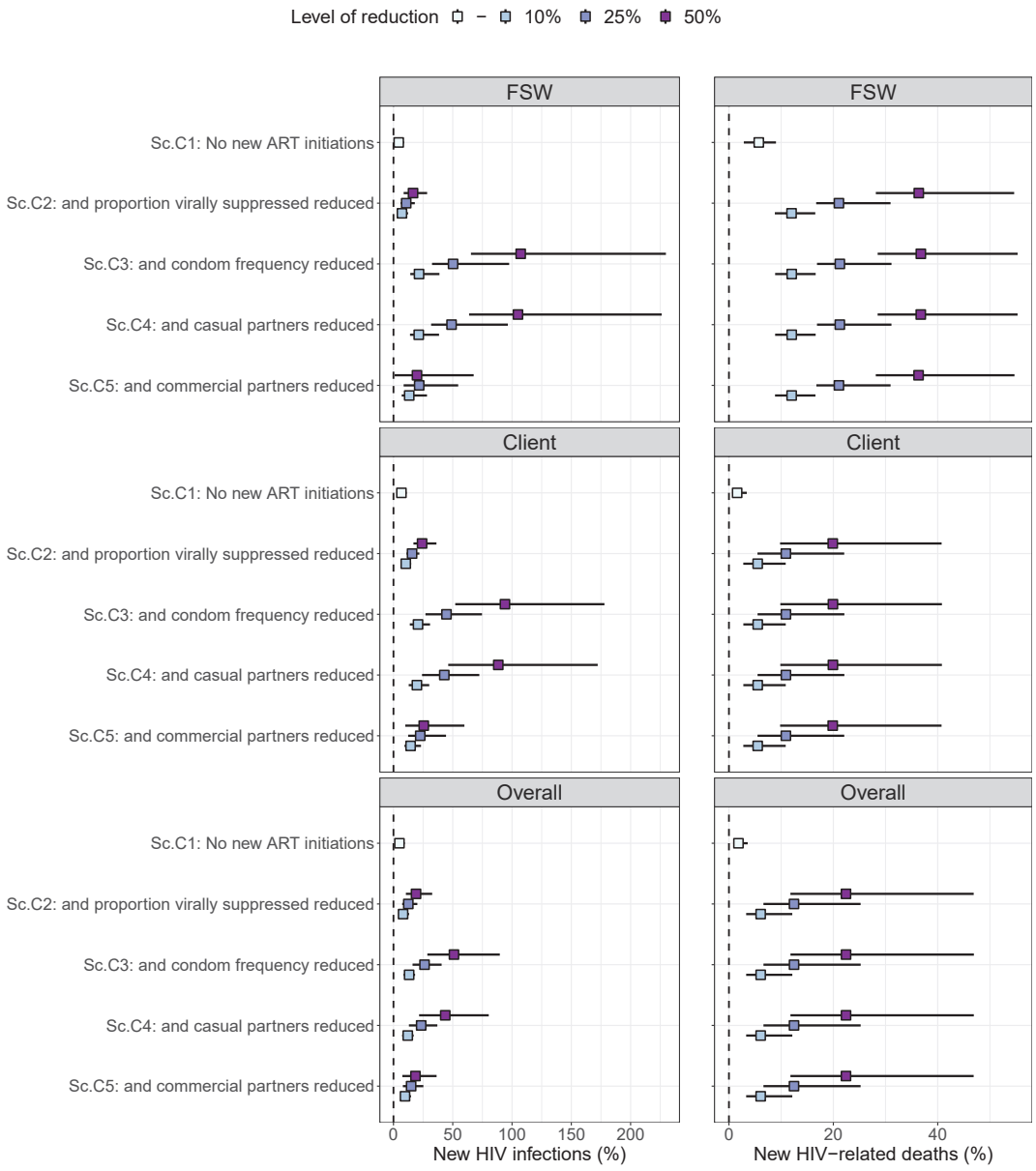


Figure 4.4 Relative change in cumulative number of new HIV infections and HIV-related deaths over 1 year among professional FSW, clients and overall under combination scenarios assuming 6 month disruptions, compared to base-case in Cotonou. For each panel, every new line indicates an additional disruption which occurs on top of the previous one. The level of reduction represented in each disruption is 10% (light blue), 25% or 50% (dark purple). Squares and lines represent median and 95% percentiles of estimates across 111 model fits.

The reduced transmission following a reduction in partnerships did not fully compensate for the increased transmission from other disruptions. With reductions of 50%, the combination of all disruptions (scenario C5) increased HIV infections over one year by 20% (1-68%), 26% (10-60%) and 19% (7-36%) among professional FSW, clients and overall respectively.

The combination of cessation of ART initiation with a 50% reduction in viral suppression (scenario C2) increased HIV-related deaths by 36% (28-55%) among professional FSW, 20% (10-41%) among clients, and 22% (12-47%) overall. The incremental impact of any further disruption was negligible (<0.1%), whether increased from reduced condom use or decreased from reduced partnerships.

The impacts of combined disruptions scaled roughly linearly with the degree of disruption, i.e. the impact when reducing by 50% is around double that when reducing by 25%, and around 5 times that when reducing by 10%.

4.4.2.3 Group-specific scenarios

I also explored various scenarios in which disruptions only affected certain groups; these results are presented in Figures 4.5 and 4.6. In terms of new HIV infections, the scenario with the greatest negative impact was the 6-month 50% reduction in condom use between professional FSW and clients (reduced for both commercial and non-commercial sex), shown in Figure 4.5. This disruption increased new HIV infections by an estimated 68% (38-159%) among professional FSW, 45% (21-110%) among clients and 18% (9-42%) overall, measured over 1 year. The equivalent disruption for lower-risk populations was significantly smaller: halving condom use among these groups increased HIV infections by 1.5% (0.8-2.6%) overall, and less than 0.2% for professional FSW and clients.

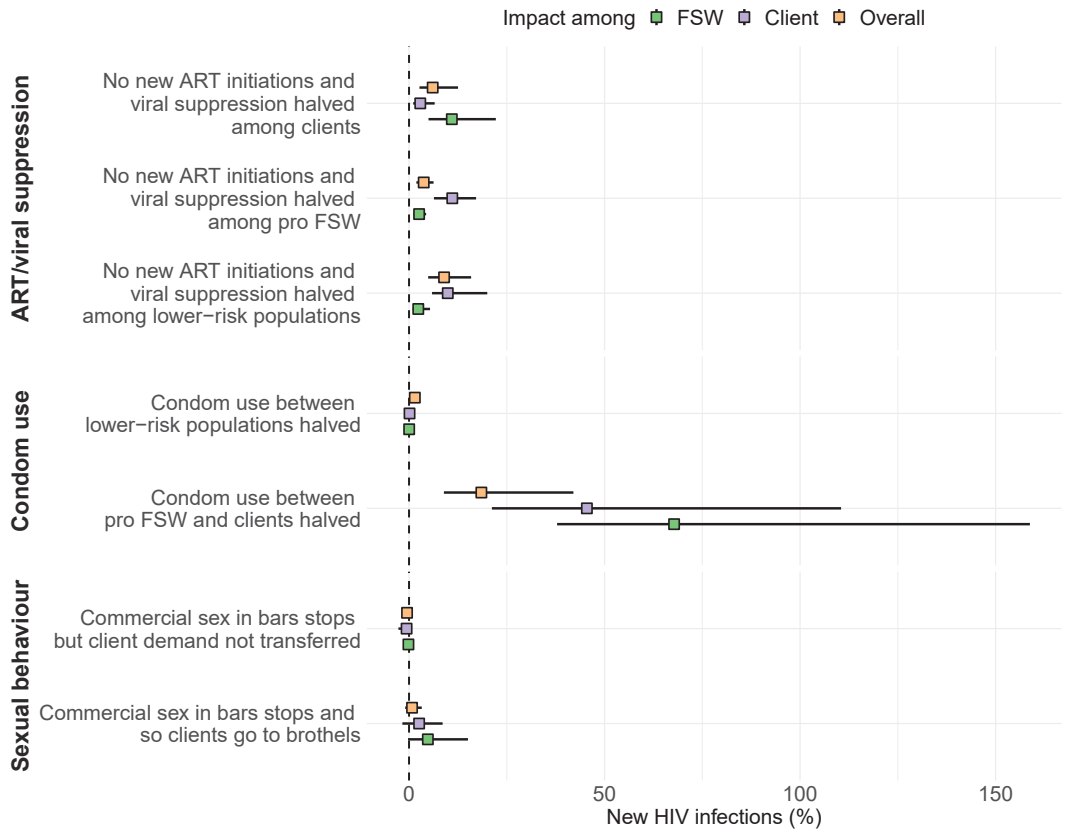


Figure 4.5 Relative change in cumulative number of new HIV infections over 1 year among professional FSW (green), clients (purple) and overall (orange) under scenarios with group-specific 6-month disruptions, compared to base-case in Cotonou. Squares and lines represent median and 95% percentiles of estimates across 111 model fits.

The impacts of disruptions to ART initiation and viral suppression were small over 1 year, and the group most at risk of increased HIV infection and death depended on whose viral suppression was affected. For example, if the interruption of ART initiations and the halving of viral suppression occurred only among clients, the greatest impact was among professional FSW, with an increase of 11% (5-22%) HIV infections, compared to 3% (1-7%) among clients and 6% (3-13%) overall. The equivalent disruption among professional FSW had the most effect on clients, increasing HIV infections by 3% (1-4%) among professional FSW, 11% (6-17%) among clients and 4% (2-6%) overall. The equivalent disruption among lower-risk populations increased infections by 2% (1-5%), 10% (6-20%) and 9% (5-16%) among professional FSW, clients and overall respectively.

Finally I explored the impact of 6-month disruptions to commercial sex of part-time FSW (Figure 4.5). If these FSW stopped commercial sex, and clients did not seek it instead from professional FSW, there was an estimated 0.2% (0.0-0.7%) decrease in new HIV infections among professional FSW, 0.6% (0.1-2.7%) among clients and 0.5% (0.2-1.5%) overall. However, if instead this client demand was transferred to professional FSW, there was an estimated increase in HIV infections of 5% (0-15%) among professional FSW, 3% (0-9%) among clients and 1% (0-3%) overall.

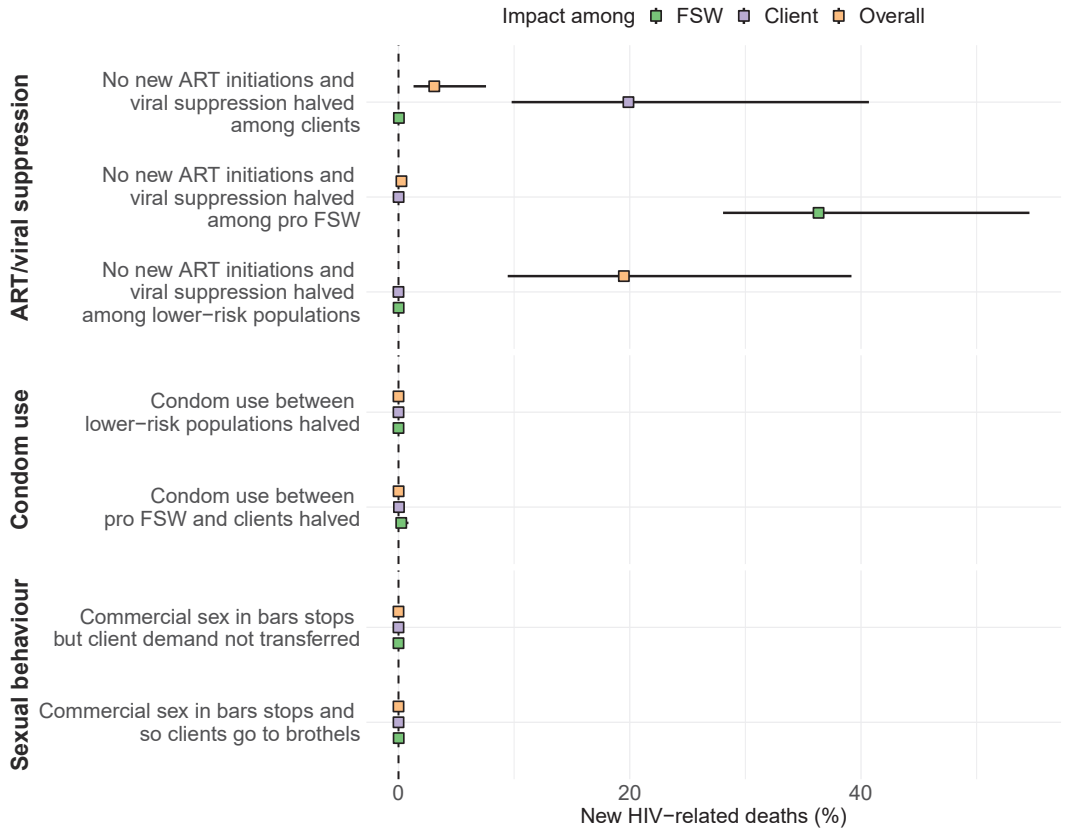


Figure 4.6 Relative change in cumulative number of new HIV-related deaths over 1 year among professional FSW (green), clients (purple) and overall (orange) under scenarios with group-specific 6-month disruptions, compared to base-case in Cotonou. Squares and lines represent median and 95% percentiles of estimates across 111 model fits..

Figure 4.6 shows the impacts of group-specific scenarios in terms of changes in HIV-related deaths. The group who experienced disruptions to ART and viral suppression are those who experienced the greatest increase in HIV-related deaths. For instance, disruptions to ART initiation and viral suppression among clients increased HIV-related deaths by 0% (0-0%)

among professional FSW, 20% (10-41%) among clients and 3% (1-8%) overall. Equivalent disruptions to ART initiation and viral suppression among professional FSW increased HIV-related deaths by 36% (28-55%) among professional FSW, and by less than 0.3% on average among clients and overall. Equivalent disruptions to lower-risk groups increased HIV-related deaths 19% (9-39%) overall, and by less than 0.1% among professional FSW and clients.

All other disruptions, in terms of condom use and sexual partnerships, increased HIV-related deaths by less than 0.1% on average for all groups.

4.5 Discussion

The city of Cotonou in Benin has a long history of successful HIV interventions, including widespread condom promotion and uptake, HIV testing and ART initiation [342, 262, 280, 270, 253]. However, the recent COVID-19 pandemic has caused significant concern regarding the sustained implementation of these interventions across Sub-Saharan Africa [343]. The aim of this chapter was to explore the potential impact of the COVID-19 pandemic on HIV epidemiology in Cotonou. To do this, I used a transmission model already calibrated to the local epidemic (described in Chapter 2), and simulated various scenarios reflecting potential disruptions to already present HIV interventions such as condom use, viral suppression, ART initiations and HIV testing. I also explored scenarios with disruptions to sexual behaviour, whether self-imposed or resulting from physical distancing policies. I calculated impact as the change in new HIV infections and HIV-related deaths following these disruptions, measured over 1 year or 5 years.

4.5.1 COVID-related disruptions to HIV interventions

4.5.1.1 Plausible scenarios

There is, to date, limited data directly measuring the disruptive impact of COVID-19 on HIV interventions. I designed a plausible data-driven scenario, based on a FSW survey in Cotonou, expert opinion, and data from nearby countries. To summarise, this scenario was a 3-month disruption, including a total pause in new ART initiations, as well as 10% reductions in viral suppression, condom frequency, casual and commercial partnerships. When measured over 1 year, these disruptions had a modest impact, increasing HIV infections by 6%, 7% and 5% in professional FSW, clients and overall respectively. Figure 4.1 and Figure 4.4

show that the largest single contributing factor was the disruption to commercial condom use. Indeed, bio-behavioural surveys in 2012 and 2015 showed that condom use between professional FSW and clients exceeded 90% [265, 258, 283, 259], a phenomenon likely to have strongly driven the sharp decrease in incidence among FSW in recent decades [50]. A modest 10% decrease in commercial condom use translates to a near two-fold increase in the number of unprotected sex acts, which risks undoing the successful continued decrease in HIV incidence in Cotonou. These results show that efforts to maintain condom use in these settings are therefore crucial to mitigate the impacts of COVID-19 on new HIV infections, and that even small disruptions can cause significant impacts.

The plausible data-driven scenario led to an estimated increase in HIV-related deaths of 6%, 3% and 3% among professional FSW, clients and overall respectively. The major contributing factor was from disruption to viral suppression, as shown in Figure 4.1 and Figure 4.4. Indeed, professional FSW are the risk group with the highest HIV prevalence, and so they have the greatest risk of mortality from lower viral suppression. Although new HIV infections will lead to mortality if not controlled with ART, the measurement of impact over 1 or 5 years is not long enough for these impacts to be fully measured, given the average duration of an infection without the use of ART is around 10-12 years [344–349].

The plausible data-driven scenario is a conservative disruption, based on the information available at the time of analysis. If it emerges that more severe disruptions occurred during this period, or will occur in future waves of COVID-19, then higher impacts are likely.

4.5.1.2 More recent data

As of 7th September 2021, there is still very little data describing COVID-related disruptions to HIV prevention and treatment services, and sexual behaviour in Cotonou. UNAIDS tracked changes in HIV services in certain Sub-Saharan African countries during the COVID-19 epidemic between March and August 2020 [350]. They found that HIV testing fell in most countries, including in Sierra Leone, where it fell by around 30% compared to before the COVID-19 epidemic. Similar reductions in HIV testing were found by PEPFAR in the West African countries of Ghana and Togo [351]. Although Benin is not a PEPFAR supported country, nor featured in the UNAIDS reports, similar reductions in HIV testing could have occurred.

PEPFAR further found that viral load suppression was maintained during April–June 2020 in all PEPFAR-supported countries [352]. UNAIDS also reported that ART coverage has been relatively unaffected by the COVID-19 or COVID-related restrictions [353]. Local experts

in Cotonou (Michel Alary, MD and Luc Béhanzin, PhD) reported that ART stocks were maintained in Benin, although data on viral suppression and ART dropout were unavailable.

If more data emerges describing any COVID-related disruptions in Cotonou during 2020, this model analysis can provide an estimate of the likely impact on HIV epidemiology.

4.5.2 The disproportionate impact on key populations

While little is known about the impact of COVID-19 and the COVID-19 response on Sub-Saharan African HIV epidemiology in general, even less is known about the impact on key populations in particular. The results from this chapter show that professional FSW and their clients are disproportionately affected by universal disruptions to HIV prevention. The data-driven plausible scenario, which features disruptions applied equally to all groups, has the highest median impact among clients in terms of new HIV infections (7.3% vs. 4.7% overall), and among professional FSW in terms of HIV-related deaths (6.1% vs. 3.0% overall).

The reasons for this can be found by exploring the single scenarios. For example, a reduction of 50% in overall condom use (both commercial and non-commercial) led to a median estimated increase in HIV infections of 69%, 49% and 23% among professional FSW, clients and overall respectively. The first explanation for the high impact among professional FSW is that condoms are used more in commercial sex than in non-commercial sex. If around 90% of commercial partnerships are protected by a condom as observed [265, 258, 283, 259], then 10% are unprotected. A 50% reduction in condom use (to 45% of sex acts protected) would mean unprotected sex acts increase from 10% to 55% of all sex acts, a greater than five-fold increase. Therefore, the originally high condom use by professional FSW means they are more at risk if there are any disruptions. Secondly, due to higher HIV prevalence and number of sexual partnerships among both FSW and clients, they both experience a higher force of infection than lower-risk populations. A reduction in their condom use, therefore, will subject both groups to a higher exposure to HIV acquisition per sex act compared to lower-risk groups.

A similar pattern is observable when considering interruptions to HIV prevention and treatment services. Disruptions to HIV testing, ART initiation and viral suppression all have bigger impacts on key populations than low risk populations. In this case, clients experience the greatest increase in estimated new HIV infections. This is because ART initiation and viral suppression among professional FSW, the risk group with the highest HIV prevalence, decreases the risk of onward infection. Any disruption to the FSW HIV care cascade will

create the biggest HIV risk exposure among clients, with whom FSW have sexual partnerships. Even if the impacts of disruptions seem modest overall, care must be taken that disproportionate impact is not felt by key populations.

On the other hand, the protective effect caused by a reduction in all partnerships due to physical distancing is very similar among key populations and lower-risk populations: a 50% reduction in all partners for 6 months would prevent an estimated median of 32%, 32% and 31% of HIV infections among FSW, clients and overall respectively. Reducing commercial partnerships protects key populations more, while reducing non-commercial partnerships protects lower-risk populations more; reducing all partnerships averages out to similar protection among all groups. Therefore, the specific impact on key populations depends on which partnerships are likely to be disrupted.

Further, the results from this chapter highlight the importance of understanding behavioural responses to disruptions of certain partnerships, particularly in commercial sex. I explored the impact of a disruption to commercial sex from part-time FSW following the closure of bars, clubs and restaurants. The epidemiological impact of this strongly depends on the behaviour of clients. If this client demand disappears, then there is a modest reduction in new HIV infections among all groups: by 0.2%, 0.6% and 0.5% among professional FSW, clients and overall respectively. However, if clients instead buy the same amount of sex from professional FSW in brothels, there is an estimated increase in new HIV infections: a median of 5%, 3% and 1% among professional FSW, clients and overall respectively. Understanding the behaviour of clients in response to disruptions is therefore crucial to estimating the impact of government policy.

In terms of HIV-related deaths, by far the most impactful disruptions were those to viral suppression. As professional FSW are the group with the highest HIV prevalence, any reduction in their levels of viral suppression is the most dangerous. If stocks of ART become limited throughout the city due to COVID-19, then this key population is most at risk of HIV-related death in the short and long term. Efforts should be made to protect key populations from disruptions to their HIV prevention and care.

4.5.3 Core group effect

Section 4.5.2 outlined why key populations may require extra attention in the case of disruptions from COVID-19, as they are disproportionately at risk from HIV infection and mortality. However, my results demonstrate a second reason for focused protection on key populations. Because of high HIV prevalence and incidence among key populations, they

are often seen as the 'drivers' of concentrated epidemics, such as the one in Cotonou [137]. Increased HIV risk among key populations can lead to increased secondary transmissions in lower-risk populations from the extended sexual network in the city. Figure 4.1 shows the estimated increase in HIV infections from single disruptions. A 50% reduction in commercial condom use naturally had more serious consequences for key populations compared to the overall population; meanwhile non-commercial condom disruptions had a bigger impact in the overall population compared to key populations. It is however important to note that for the overall population, commercial condom disruptions are more impactful than non-commercial ones. Although commercial condom use does not directly affect most of the population in Cotonou, the very significant increase in HIV infections among professional and FSW leads to increased secondary infections in lower-risk groups. This clearly demonstrates the effect of the core group theory: protecting key populations will in turn protect lower-risk population too, by preventing onward transmission.

4.5.4 Other patterns in the impacts of disruptions

The results in this chapter describe the potential impact of longer and more drastic disruptions, in the worst cases increasing infections by over 100%. Generally, the length and degree of disruption scaled linearly with impact, both in terms of HIV infections and deaths.

In the scenarios that I explored, disruptions to commercial condom use and viral suppression were the biggest contributors to HIV infections and deaths respectively. If policymakers wish to minimise increases in new HIV infections, priority should be given to prevent disruptions to condom use. However, if policymakers wish to minimise increases in new HIV deaths, priority should be given to prevent disruptions to viral suppression.

The protective effects of physical distancing are also demonstrated. A reduction in sexual partnerships can lead to fewer new HIV infections than predicted without COVID-19. However, reduced sexual partnerships are not enough to compensate for equivalent disruptions to condom use or viral suppression.

Finally, the period over which to measure impact is important to define. A disruption generally has a greater relative impact, but a smaller absolute impact, when measured over a shorter period, rather than a longer one. Although immediate negative impacts must be recognised, it is nevertheless important to also consider these over the long term, as it could take many years for resultant HIV deaths to accumulate.

4.5.5 Comparison to previous studies

The results in this modelling study compare well to those in similar countries. Silhol and colleagues performed a similar analysis on the HIV epidemic in Yaoundé, Cameroon, as a model comparison with this study [317]. They used a transmission model calibrated to historical data in the city. The results from Yaoundé were of similar magnitude as mine from Cotonou for most of the disruptions that were simulated.

For example, a 6-month 50% reduction in viral suppression for all groups increased new HIV infections by 18% (95% uncertainty interval: 16-22%) overall in Yaoundé, compared to 14% (7-27%) in Cotonou. The larger uncertainty interval in Cotonou was due to a greater uncertainty in the posterior proportion of the general population virally suppressed; for example among lower-risk males the 95% UI for this parameter in Yaoundé and Cotonou is 34.2–43.0% and 23.3-63.0% (Table 4.1) respectively [52, 317]. In terms of increased HIV-related deaths, the impact was higher overall in Cotonou compared to Yaoundé: 21% (11–44%) and 13% (9–16%) respectively. A greater impact will occur if a larger ART coverage is disrupted, and indeed Cotonou has higher posterior coverage than Yaoundé among the HIV-infected lower-risk populations in 2020 (75% vs. 55% among women and 68% vs. 49% among men).

A 50% reduction in overall condom use for 6 months gave strikingly similar results between the two cities in terms of increased new HIV infections: 23% (15–31%) and 23% (12–49%) over 1 year in Yaoundé and Cotonou respectively, and 8% (5–10%) and 8% (4–18%) over 5 years. However this similarity obscures differences in commercial and non-commercial condom use. A 50% reduction in commercial condom use increased HIV infections by 4% (2–9%) in Yaoundé and 19% (9–42%) Cotonou. This difference is because Cotonou has an HIV epidemic which is much more concentrated within commercial sex networks than in Yaoundé; although HIV prevalence among these groups is higher in Yaoundé than Cotonou (e.g. among full-time FSW: 13.9% (11.9–16.4%) vs 8.5% (7.1–11.2%) respectively from posterior model distributions in 2020), HIV prevalence among the lower risk populations is relatively higher in Yaoundé than Cotonou. For example, HIV prevalence among lower risk women in Yaoundé is an estimated $\sim$5x higher (4.9% (3.9–5.9%) vs 1.0% (0.8–1.5%)). This comparison indicates that the impact of disruptions depends strongly on the epidemiological characteristics of the locality.

Indeed this effect means that in Cotonou key populations are more severely impacted by similar disruptions. A worst-case scenario simulated in both locations, a 6-month pause in ART initiations combined with 50% reductions in condom use and viral suppression,

increased overall HIV infections by 51% in both cities over 1 year. However, among FSW and clients together the impact was 1.4x higher in Yaoundé, and 1.9x higher in Cotonou. Extra care to protect key populations must be considered in locations where HIV epidemics are more concentrated.

4.5.6 Strengths and limitations

This study has many strengths. First, the model used has been calibrated (and cross-validated) to decades of Cotonou-specific data, gathered from many sources, enabling a close representation of the true HIV epidemic of Cotonou in 2020. Secondly, given that specific risk groups including key populations were explicitly modelled, the effects on each can be measured and compared. Previous studies had not focused so carefully on key populations [306, 337], even though they are usually more vulnerable to HIV incidence and mortality. Thirdly, the transmission model framework reflects chains of transmission, such that the long term impact of disruptions can be understood. Finally, the results were comparable to similar modelling studies in nearby locations, which corroborates the findings and adds reliability.

The main limitation of this study was the lack of empirical data on the COVID-related disruptions being explored. For example, although bars were shut in Cotonou for around 3 months, it has not been shown that part-time FSW stopped commercial sex for this period, nor that clients did not instead transfer their demand to professional FSW. Due to this lack of data, many assumptions on the nature of the disruptions had to be made.

Having said this, however, as I explored different levels of disruption (i.e. reductions between 10% and 50%), one can reasonably assume that if there were any disruptions they fell within this range. A 50% reduction in viral suppression for 6 months increased new HIV infections by an estimated 14% among all groups over one year. Given that such a disruption is very extreme and unlikely to happen, one can therefore be confident that given this model analysis, any disruption to viral suppression is unlikely to increase new HIV infections by more than 14%.

4.5.7 Implications

The implications of this modelling study are numerous. If new data emerge which describe any disruption to sexual behaviour or HIV prevention services in Cotonou, then this model analysis can be co-opted to estimate the potential impact in terms of HIV infections and HIV-related deaths, among key populations and overall. Even if data does not emerge describing

the more chaotic first wave of 2020, where the outbreak was unexpected and difficult to measure, data may yet still be generated in future waves of COVID-19. Epidemiologists argue that this is a likely scenario before vaccine coverage becomes high, or if vaccine escape variants emerge [354].

The results show which disruptions are likely to cause the greatest impacts in terms of changes in new HIV infections or HIV-related deaths. This can help in mitigating the negative consequences of future disruptions. For instance, commercial condom use must be protected from any disruption in order to minimise the impact of COVID-19 on new HIV infections.

This study further highlights the disproportionate negative impact likely incurred by key populations relative to lower-risk ones. Even if a disruption universally affects all groups, professional FSW and their clients can disproportionately feel the impact. A small disruption may cause a modest increase of infections, but most of these cases may be concentrated among key populations. This corroborates previous studies which demonstrate their increased vulnerability, and the need for maintenance of HIV prevention and treatment services among them [355].

The precarious working and living conditions for key populations (such as FSW in Cotonou) make them particularly vulnerable to the impacts of disruptions. Previous studies have found that key populations experience higher levels of stigma and abuse [356], and given the officially illegal status of sex work in Benin [261], FSW may be impacted in unexpected ways and less likely to receive professional aid.

Finally, this exercise underlines once more the effectiveness of HIV prevention interventions. Chapter 3 estimated the impact of adding PrEP and TasP among professional FSW to the already-present suite of HIV prevention tools in Cotonou. This chapter reaffirms the effectiveness of condoms and viral suppression in another fashion. Instead of estimating the positive effects of adding interventions, it estimates the negative effects of disrupting them. This is particularly clear in the case of condom use which, following decades of promotion and education campaigns, has reached very high frequency (90%+) among professional FSW in commercial sex partnerships. This study shows that any backtracking on this progress can lead to devastating results.

Interest in novel pharmaceutical interventions is increasing, particularly for long-acting PrEP [357], or a potential HIV vaccine [358]. While these interventions would undoubtedly be useful if proved effective and feasible in real-life settings, this chapter serves as a strong reminder of the importance of current interventions. Although perhaps less exciting than recent or future inventions, one must not lose sight of the ‘basics’ of HIV prevention:

promotion of condoms, regular testing, and access to ART. The infrastructure to deliver these tried and tested tools must be solidified to ensure resilience to turbulent periods, such as during a fast-paced respiratory pandemic.

Chapter 5

Estimating the impact of non-pharmaceutical interventions for COVID-19 in Weifang

Overview

In this chapter I describe an analysis of 20 SARS-CoV-2 genomic sequences from Weifang, China, using model-based phylodynamic methods. I estimate the burden of SARS-CoV-2 in the city, and various epidemiological and evolutionary parameters. I explore whether the range of restrictions imposed in Weifang were successful in controlling the SARS-CoV-2 outbreak.

This analysis was originally performed in August 2020, and is published with the following citation.

Geidelberg, L., Boyd, O., Jorgensen, D., Siveroni, I., Nascimento, F. F., Johnson, R., ... & Nie, Q. (2021). Genomic epidemiology of a densely sampled COVID-19 outbreak in China. *Virus Evolution*, 7(1), veaa102.

5.1 Introduction

5.1.1 The COVID-19 epidemic in Weifang, China

On the final day of 2019, a cluster of cases of pneumonia was reported by Wuhan Municipal Health Commission, China [359]. A week later, on 7th January 2020, the Chinese Center for Disease Control and Prevention identified the novel coronavirus (then termed '2019-nCoV') from a throat swab in Wuhan, Hubei province [360]. The International Committee on Taxonomy of Viruses (ICTV) later named the virus as SARS-CoV-2, following the discovery of close genetic relatedness to SARS, and because it caused a similarly acute respiratory infection [361]. The illness caused by SARS-CoV-2 was termed COVID-19 on 11th February 2020 [362].

As a result of Chinese New Year celebrations and the approaching Spring Festival, many travellers spread from Wuhan throughout China during December 2019 and January 2020 [363], seeding epidemics in diverse geographical locations. The city of Weifang, Shandong Province, was among the first to report an outbreak cluster of COVID-19 outside of Wuhan. The first case of COVID-19 was identified in Weifang on 23rd January 2020 [363]. As of 10th February 2020, 136 suspected cases and 214 close contacts had been diagnosed by Weifang Center for Disease Control and Prevention; of these, 38 cases were confirmed positive with SARS-CoV-2 [364].

These cases partially overlapped with a range of interventions to control the spread of COVID-19 [365]. Between 23rd-26th January, the local government engaged in public health messaging by establishing telephone hot-lines and encouraging recent travellers from Wuhan to isolate at home. On 24th January, suspected COVID-19 patients were triaged in hospitals, and on 27th January 'fever clinics' were established for consultation and disease diagnosis. Travel restrictions were implemented on 26th January, and school closures were extended on 27th January. As such, unlike the early spread of SARS-CoV-2 in Hubei Province, the majority of community transmissions within Weifang likely occurred after the implementation of these interventions.

On top of being among the first geographically concentrated outbreaks of COVID-19 outside of Wuhan [366], Weifang is of interest as it provided genomic sequencing for a large proportion of its confirmed cases. The whole genome sequences of 20 virus samples from patients were obtained by researchers from the Department of Microbiology, Weifang Center for Disease Control and Prevention, and from the Department of Respiratory Medicine, Weifang. At the date of the last sample, these 20 genomes represented 53% of the 38

confirmed cases in Weifang. There was no deliberate method of sampling to obtain these 20 sequences out of the total confirmed cases, but instead reflected the limited materials available locally given the uncertainty at the time.

Among the 20 patients from whom the virus was sequenced, two suffered severe or critical illness. The patient median age was 36 (range: 6-75). Seven of them had been to Wuhan, as it was their place of work. The thirteen others lived in Weifang, and of these two of them were in contact with a traveller from Wuhan.

5.1.2 Motivation and aims

It is of great interest to academics, policymakers and to the public to understand epidemiological parameters relating to the COVID-19 pandemic (such as reproduction number), and the impact of non-pharmaceutical interventions. However, it can be challenging to understand these for several reasons.

As described in Chapter 1, many COVID-19 patients are asymptomatic, which makes them far less likely to get tested. Case ascertainment can therefore be extremely poor in certain contexts, particularly where there is a high proportion of asymptomatic cases (e.g. in young people), or where there is weak infrastructure for widespread testing. This was especially true in early 2020 for most countries, as tests were not widely available.

As a result of unreliable case counts, fitting mathematical models can prove challenging. Certain models are fitted instead to hospitalisations or deaths, but to reproduce epidemic trajectories of all infections, assumptions must therefore be made on the proportion of cases which are severe or lethal.

Alternatively, genomic sequence alignments are another data set with which to fit models. Indeed, model-based phylodynamic methods have been previously used to analyse sequence data from Wuhan and exported international cases [367].

The objective of this chapter is therefore to estimate the epidemic trajectory of SARS-CoV-2 in Weifang, as well as the impact of non-pharmaceutical interventions on epidemiological parameters. As stated in the aims of this thesis, the key research questions are:

What was the true number of COVID-19 infections in Weifang, China, during January and February 2020?

What was the impact of NPIs on COVID-19 transmission in Weifang?

Secondary questions that I aim to answer are: how does the estimated burden compare to confirmed cases? How did viral transmission vary through time?

To do this, I used an adapted version of a previously published model-based phylodynamic method to be suitable for fitting to the 20 whole SARS-CoV-2 genome sequences in Weifang. I estimated cumulative and daily infections, and through comparison with confirmed cases published by Weifang CDC, I estimated the reporting frequency. I also estimated reproduction number through time, and from this interpreted the impact of NPIs in Weifang.

5.2 Methods

5.2.1 Epidemiological investigation, sampling, and genetic sequencing

Through collaboration with Chinese colleagues, I was able to access whole genome sequences of SARS-CoV-2 from 20 patients in Weifang. These sequences were already cleaned and aligned. In this section I describe the methods for generating these sequences; note that no part of this protocol was performed by me.

Viral RNA was extracted using the Maxwell 16 Viral Total Nucleic Acid Purification Kit (Promega AS1150) with the magnetic bead method, and the RNeasy Mini Kit (QIAGEN 74104) with the column method. Quantitative reverse transcription polymerase chain reaction (RT-qPCR) was carried out using the 2019 novel coronavirus nucleic acid detection kit (BioGerm, Shanghai, China) to confirm the presence of SARS-CoV-2 viral RNA with cycle threshold (Ct) values ranging from 17 to 34, targeting the highly conservative region (ORF1ab/N gene) in the SARS-CoV-2 genome.

Concentration of RNA samples was measured by the Qubit RNA HS Assay Kit (Thermo Fisher Scientific, Waltham, MA, USA). The enzyme DNase was used to remove host DNA. The remaining RNA was used to construct the single-stranded circular DNA library with the MGIEasy RNA Library preparation reagent set (MGI, Shenzhen, China). Purified RNA was then fragmented. Using these short fragments as templates, random hexamers were used to synthesise the first-strand cDNA, then the second strand. Using the short double-strand DNA, a DNA library was constructed through end repair, adaptor ligation, and PCR amplification. PCR products were transformed into a single strand circular DNA library through DNA-denaturation and circularisation. DNA nanoballs (DNBs) were generated with the single-strand circular DNA library by rolling circle replication (RCR). The DNBs were loaded into the flow cell and pair-end 100bp sequencing was performed on DNBSEQ-T7

platform 8 (MGI, Shenzhen, China). Twenty genomes were assembled with length ranging between 26,840 to 29,882 nucleotides.

Total reads were first processed using *Kraken* v0.10.5 (default parameters) with a self-built database of Coronaviridae genomes (including SARS, MERS, and SARS-CoV-2 genome sequences downloaded from GISAID, NCBI, and CNGB) to identify Coronaviridae-like reads. To remove low-quality reads, duplications and adaptor contaminations, *fastp* v0.19.5 (parameters: -q 20 -u 20 -n 1 -l 50) and *SOAPnuke* v1.5.6 (parameters: -l 20 -q 0.2 -E 50 -n 0.02 -5 0 -Q 2 -G -d) were used. The Coronaviridae-like reads of samples with $< 100\times$ average sequencing depth were directly assembled de novo with *SPAdes* v3.14.0 using default settings. The Coronaviridae-like reads of samples with $> 100\times$ average sequencing depth across the SARS-CoV-2 genome were subsampled to achieve $100\times$ sequencing depth before being assembled.

The 20 Weifang sequences had mean 1.1% N content, and are deposited in GISAID (gisaid.org).

5.2.2 Mathematical model

To carry out this analysis, I used the *PhyDyn* package (v1.3.7) implemented in BEAST (v2.6.1), which has previously been published [105]. Briefly, the *PhyDyn* package is a flexible tool for implementing custom compartmental models for phylodynamic inference. As opposed to non-parametric methods, such as *skyline* and *skyride*, *PhyDyn* permits the pre-specification of model structure and parameterisation (e.g. the common SIR model), such that susceptible and infected groups can be explicitly modelled. Further, unlike earlier model-based phylodynamic inference packages, *PhyDyn* allows a more complex set of structured coalescent models. For instance, migration between demes can vary over time, and lineages located in different demes are allowed to coalesce together. Further details are available in Chapter 1.

In order to perform phylodynamic modelling of SARS-CoV-2, the following features were included: i) nonlinear epidemic dynamics in Weifang with a realistic course of infection (incubation and infectious periods), ii) variance in transmission rates which can influence epidemic size estimates, and iii) migration of lineages in and out of Weifang. I justify the inclusion of these features and describe their implementation below.

Nonlinear epidemiological dynamics in Weifang

The maximum number of daily confirmed COVID-19 cases occurred on 5th February, but it is unknown when the maximum prevalence of infection occurred. To capture a nonlinear decrease in cases following epidemic peak, and to account for a realistic distribution of generation times, I used an extension of the susceptible-exposed-infectious-recovered (SEIR) model [368] for epidemic dynamics in Weifang, shown in the set of ordinary differential equations 5.1.

Variance in transmission rates

To estimate total numbers infected, the phylodynamic model must account for epidemiological variables which are known to significantly influence genetic diversity [369]. Foremost among these is the variance in offspring distribution (number of transmissions per primary case). Drawing on evidence based on the previous SARS epidemic, the offspring distribution is highly over-dispersed [369]. Analyses of SARS-CoV-2 sequence data drawn primarily from Wuhan has found that high over-dispersion was required for estimated cases to be consistent with the epidemiological record [367]. High variance of transmission rates will reduce genetic diversity of a sample and failure to account for this factor will lead to highly biased estimates of epidemic size [370]. Models assuming low variance in transmission rates between people would generate estimates of cases that are lower than the known number of confirmed cases.

Further, Endo et al. [371] found that high over-dispersion is required to reconcile estimated reproduction numbers with the observed frequency of international outbreaks. The SEIR model was therefore extended with an additional infectious compartment J which has a higher transmission rate (τ-fold higher) than the I compartment.

The variance of the implied offspring distribution was calibrated to give a similar overdispersion to that of the SARS epidemic.

Equations 5.1 describe the system of ordinary differential equations. Susceptibles become exposed at a force of infection rate proportional to β and the number of infected individuals. Upon leaving the exposed compartment, at a rate which is the inverse of the incubation period, individuals progress to the J compartment with probability p_h, or otherwise to I. Individuals enter the 'recovered' compartment at rate $1/\gamma_1 = 3.8$; however it is important to note that this combines both recovery and mortality.

$$\begin{aligned}
\dot{S}(t) &= -(\beta I(t) + \beta\tau J(t))\frac{S(t)}{S(t)+E(t)+I(t)+J(t)+R(t)} \\
\dot{E}(t) &= (\beta I(t) + \beta\tau J(t))\frac{S(t)}{S(t)+E(t)+I(t)+J(t)+R(t)} - \gamma_0 E(t) \\
\dot{I}(t) &= \gamma_0(1-p_h)E(t) - \gamma_1 I(t) \\
\dot{J}(t) &= \gamma_0 p_h E(t) - \gamma_1 J(t) \\
\dot{R}(t) &= \gamma_1(I(t)+J(t))
\end{aligned} \tag{5.1}$$

The compartments for infected individuals (E, I and J) are considered demes, i.e. they correspond to the state of a lineage. Rates of coalescence may vary between lineages. The compartments S and R are non-demes, and exist in the model simply to facilitate the growth and decline of the infected compartments. They are not to be interpreted as estimates of the true susceptible and recovered populations respectively. I explore this parameterisation in Section 5.4.

The basic reproduction number (R_0) can be calculated for the I and J compartments separately, which when combined as a weighted average, produces the overall estimated R_0:

$$R_0 = (1-p_h) * \beta/\gamma_1 + \tau * p_h * \beta/\gamma_1 \tag{5.2}$$

Importation of lineages from Wuhan

The outbreak in Weifang was likely seeded by multiple lineages imported at various times from the rest of China. The location of sampling is therefore accounted for in the model. Migration is modelled as a bi-directional process with rates proportional to epidemic size in Weifang. The larger reservoir of COVID-19 cases outside of Weifang ($Y(t)$) serves as a source of new infections and is assumed to be growing exponentially (at rate ρ) over this time period.

The equation governing this population is:

$$\dot{Y}(t) = (\rho - \mu)Y(t). \tag{5.3}$$

Migration only depends on the size of variables in the Weifang compartment and thus does not influence epidemic dynamics; it will only influence the inferred probability that a lineage

resides within Weifang. For compartment X (E, I, or J), η is the per-lineage rate of migration out of Weifang, and the total rate of migration in and out of Weifang is ηX.

5.2.3 Model parameterisation

During phylodynamic model fitting, the transition parameters η, β and ρ are estimated, as well as the initial sizes of Y, E, and S. Table 5.1 shows the prior distribution for parameters which were estimated in the phylodynamic analysis.

Table 5.1 Summary of prior distributions for primary epidemiological and evolutionary parameter.

Parameter	**Prior**
Initial infected	Exponential(mean=1)
Initial susceptible	Exponential(mean=500)
Migration rate[1]	Exponential(mean=10)
Transmission rate	Log-normal(mean log=3.21, sd log=0.5)
Reproduction number	Log-normal(mean log=1.03, sd log=0.5)
Molecular clock rate[2]	Uniform(0.0007,0.003)
Transition/transversion	Log-normal(mean log=1,sd log=1.25)
Gamma shape	Exponential(mean=1)

[1]Units: Migrations per lineage per year.

[2]Units: Substitutions per site per year.

Initial values of I, J and R are fixed at 0. Other parameters are fixed based on prior information. I fixed $1/\gamma_0 = 4.1$ days and $1/\gamma_1 = 3.8$ days [367]. I set $p_h = 0.20$ and $\tau = 74$ which yielded a dispersion of the reproduction number that matches a negative binomial distribution with $k = 0.124$ for any value of R_0 between 2 and 5. This dispersion is similar to values estimated for the 2003 SARS epidemic [369].

The phylodynamic model is illustrated in Figure 5.1A as a flow chart. I set the SEIJR model dynamics to begin on 10th January 2020.

5.2.4 Phylogenetic analysis

Using *MAFFT* v7.471 [372], I aligned the 20 Weifang sequences with a previous alignment of 57 non-identical SARS-CoV 2 sequences from outside of Weifang, hereafter known as the 'reference set' [367]. I sampled the reference set from the *GISAID* database [373] downloaded on June 7th 2020, and explicitly included close genetic matches to sequences from Weifang. I placed an upper bound at May 1st on the date of sampling. The distribution

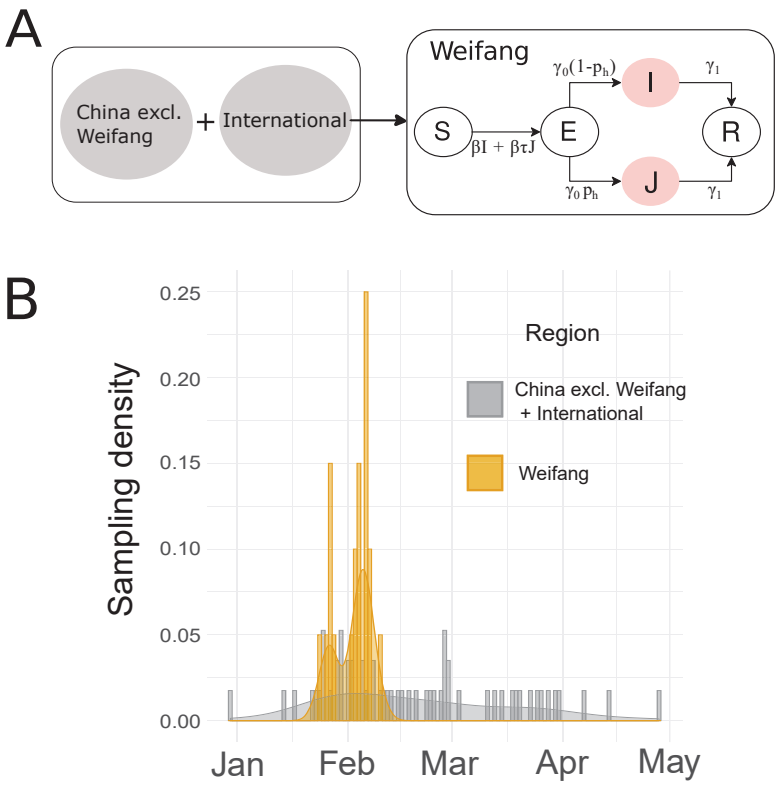


Figure 5.1 Epidemiological model and sample times. A) A diagram representing the structure of the epidemiological SEIR model which was fitted in tandem with the time scaled phylogeny. Infected and infectious individuals may occupy a low (I) or a high (J) transmission rate state to account for high dispersion of the reproduction number. B) Sampling density of sequences from inside (yellow) and outside (grey) of Weifang respectively through time.

of sample dates from inside and outside of Weifang are shown in Figure 5.1B. Of the 57 sequences in the reference set, 20 (35%) were sampled from China.

I carried out a maximum likelihood analysis using *IQTree* v1.6.12 [374] with a HKY+G4 substitution model (parameters -nt 12 -redo -m HKY -s). I used the *getRoot* function in *phangorn* to implement midpoint rooting of the tree. I estimated a time-scaled tree using *treedater* v0.5.0 [113].

Bayesian phylogenetic analysis was carried out using *BEAST* 2.6.1 [375] with a HKY+G4 substitution model and a strict molecular clock. The phylodynamic model was implemented using the *PhyDyn* package v1.3.7 [376] using the QL likelihood approximation and the RK ODE solver. The model was fitted by running 8 MCMC chains of 30 million steps in parallel, and combining chains after removing 50% burn-in.

In order to demonstrate the added utility of the sequence data, I repeated the analysis when sampling only from the prior probability distributions. This was achieved by assuming a constant likelihood, specified in the *BEAST* xml file.

I used the *ggtree* package for all phylogeny visualisations [377].

5.3 Results

5.3.1 Main analysis

5.3.1.1 Model fitting

The posterior distributions of key parameters are summarised in Table 5.2 following phylodynamic model fitting, and in Table 5.3 when sampling from the prior. After fitting the model to sequence data, good mixing was achieved and the posterior parameter distributions were well sampled; the effective sample size (ESS) was above 1,000 for all parameters. When fitting to the sequence data, the parameter posterior distributions are different to when simply sampling from the prior. Notably, mean posterior transmission rate was estimated at 21.5, compared to 28.5 when sampling from the prior. This results in a smaller (and more realistic) estimate of the reproduction number, as shown in the next section.

Table 5.2 Summary of primary parameters for the main analysis, including mean estimated posterior and effective sample size due to autocorrelation.

Statistic	Mean	ESS
Posterior	-41631.97352	1208
Likelihood	-41618.79558	1383
Prior	-13.17793938	1160
Tree likelihood	-41618.79558	1383
Tree height	0.371177175	2947
Molecular clock rate	0.00130656	1784
Gamma shape	0.286880065	2405
Transition/transversion	4.623272977	12172
PhydynSEIR	7.89341053	1130
Initial exposed	4.84479347	2956
Initial susceptible	549.8789703	3182
Transmission rate	21.45699354	4734
Initial exogenous	0.520109449	1486
Exogenous growth rate	20.47238551	1646
Migration rate	1.684142341	9309

Table 5.3 Summary of primary parameters for the main analysis but sampling from the prior, including mean estimated posterior and effective sample size due to autocorrelation.

Statistic	Mean	ESS
Posterior	-41643.01783	10410
Likelihood	-41626.68793	10272
Prior	-16.3299023	18874
Tree likelihood	-41626.68793	10272
Tree height	0.604156711	13422
Molecular clock rate	0.000709584	22985
Gamma shape	0.297625411	3313
Transition/transversion	4.62191852	13931
PhydynSEIR	0	NA
Initial exposed	1.001445038	23576
Initial susceptible	596.9800995	24008
Transmission rate	28.46598941	23490
Initial exogenous	1.00446398	24008
Exogenous growth rate	37.79539512	23614
Migration rate	0.968908965	8496

5.3.1.2 Epidemic trajectory

I estimated cumulative and daily number of infections in Weifang, shown in Figure 5.2A and Figure 5.2B respectively. The peak in daily infections was estimated to occur in late January, preceding the time series of confirmed cases by about a week; this is expected due to delays from infection to appearance of symptoms and delays from symptoms to diagnosis.

The genetic data are strongly informative about timing and size of the epidemic peak: trajectories sampled from the Bayesian prior distribution have a smaller and later epidemic peak (see Figures 5.2 and 5.3) with much less precision.

The epidemic peak in terms of daily new infections was estimated to be a few days after final NPIs in Weifang were implemented.

My central estimate for the cumulative number infected on 10th February is 365 (highest posterior density (HPD): 102-1174), compared to 38 cumulative confirmed cases. Reporting frequency of COVID-19 cases in Weifang (Figure 5.3) was calculated by dividing the number of reported cases by the cumulative mean (and 95% HPD) estimated cases.

I therefore estimated that around 10% of infections were diagnosed; an unknown proportion of infections will be missed by the surveillance system due to very mild, sub-clinical, or asymptomatic infection. This supports the hypothesis that there was a modest (but not large) burden of infection in Weifang over the period that the sequence data were sampled. When fitting to genetic sequence data (compared to sampling from the prior), the reduced uncertainty in estimated cases also provides more precision for the estimated reporting rate (Figure 5.3).

Effective reproduction number over time is shown in Figure 5.2C. I estimate $R_0 = 3.4$ (95% HPD:2.1-5.2) and the initial growth rate in cases was approximately 22% per day. Sampling from the prior yields a much higher estimate for R_0 with an unrealistic HPD upper bound over 10. My central estimate of R_t drops below 1 on the 4th of February, after the implementation of many NPIs in Weifang.

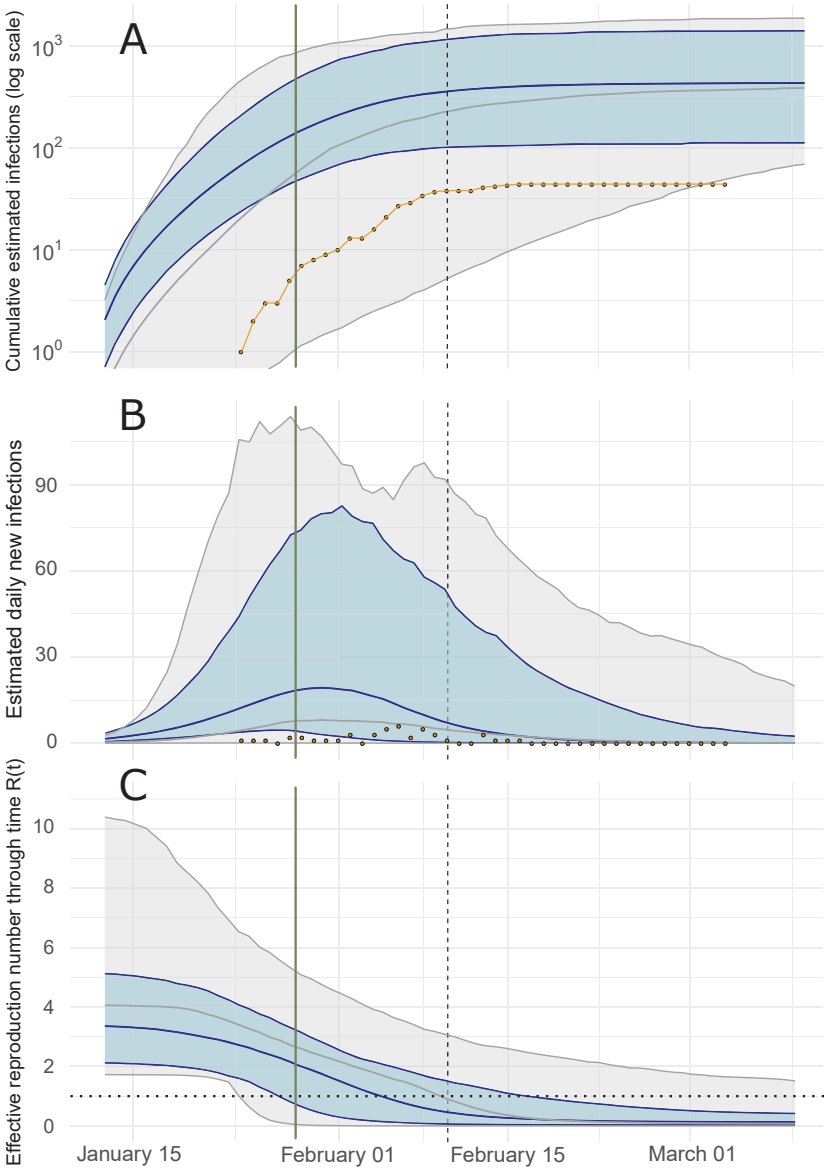


Figure 5.2 Epidemiological trajectory of the Weifang SARS-CoV-2 epidemic in 2020 when fitting the SEIR model to genetic data (blue) and sampling only from prior (grey). Solid lines and shaded area reflect posterior median and 95% HPD. The vertical green solid line represents the date of the final NPIs in Weifang. The vertical dashed line represents the date of the last sequence sampled in Weifang. A) Cumulative estimated infections through time compared to cumulative cases (yellow points) reported by Weifang CDC. B) Daily estimated infections through time compared to daily reported cases (yellow points). C) Effective reproduction number through time R_t. The horizontal dotted line indicates $R_t = 1$.

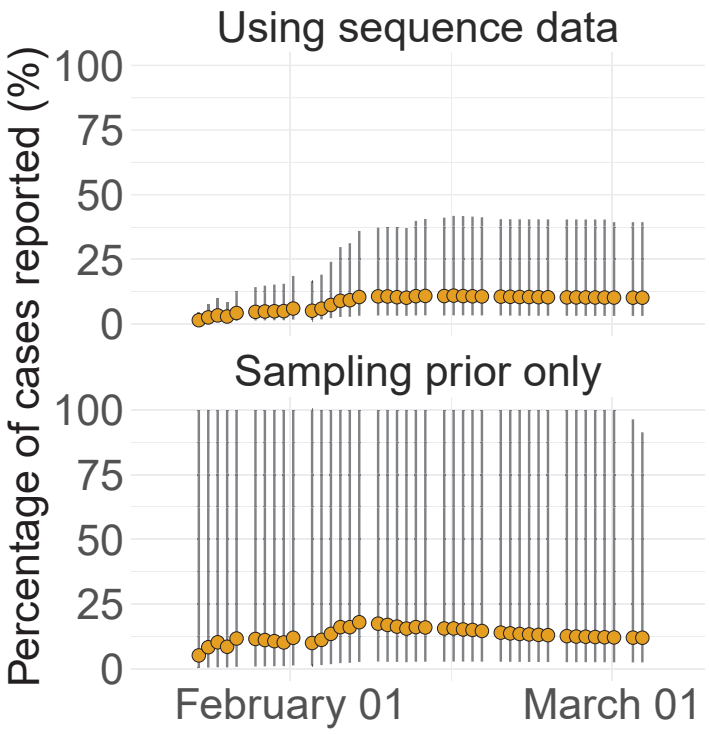


Figure 5.3 Estimated reporting frequency. The yellow points and grey bars reflect the mean and 95% HPD cumulative estimated proportion of cases that were identified through time respectively. Results shown separately for analyses conducted with and without the sequence data.

5.3.1.3 Phylogenetic analysis

There was a mean of three pairwise differences among sequences from Weifang; the corresponding number among the sequences outside of Weifang is eight. Figure 5.4 shows a maximum-likelihood phylogeny of the alignment, including 20 lineages sampled from distinct patients in Weifang (red tips) and 57 genomes sampled from Wuhan and internationally (grey tips).

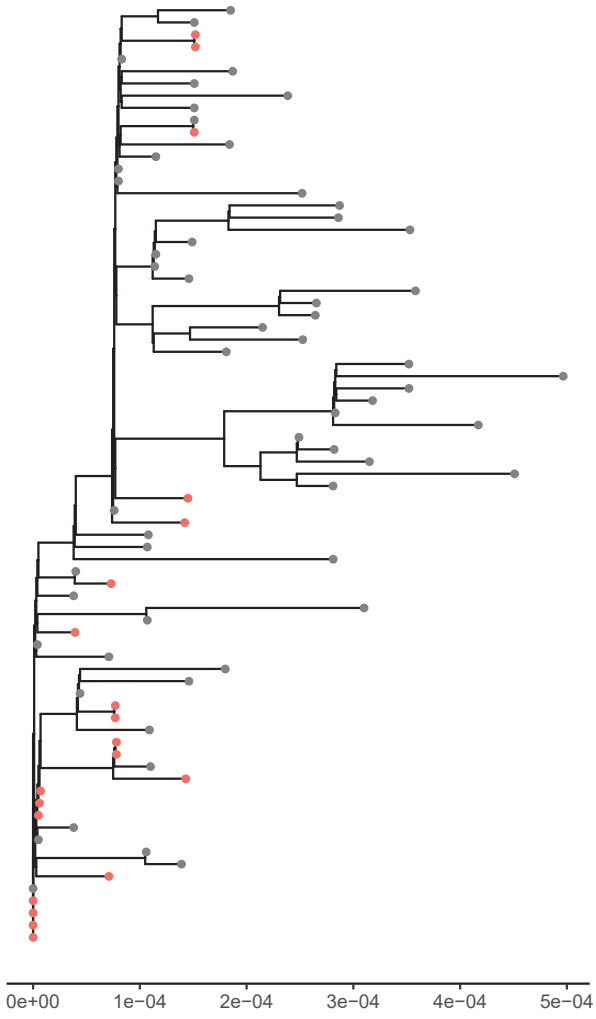


Figure 5.4 A maximum likelihood tree estimated using IQTree using the same data as used for the Bayesian analysis. Red and grey dots represent samples inside and outside Weifang respectively.

Figure 5.5 shows the time-scaled maximum clade credibility (MCC) tree which was co-estimated with the epidemiological parameters in the *BEAST* analysis. The tips from Weifang appear to cluster together, suggesting more community transmission than importations, but there is correspondingly low confidence in tree topology, and only two monophyletic Weifang clades had greater than 50% posterior probability, neither of which are larger than two samples.

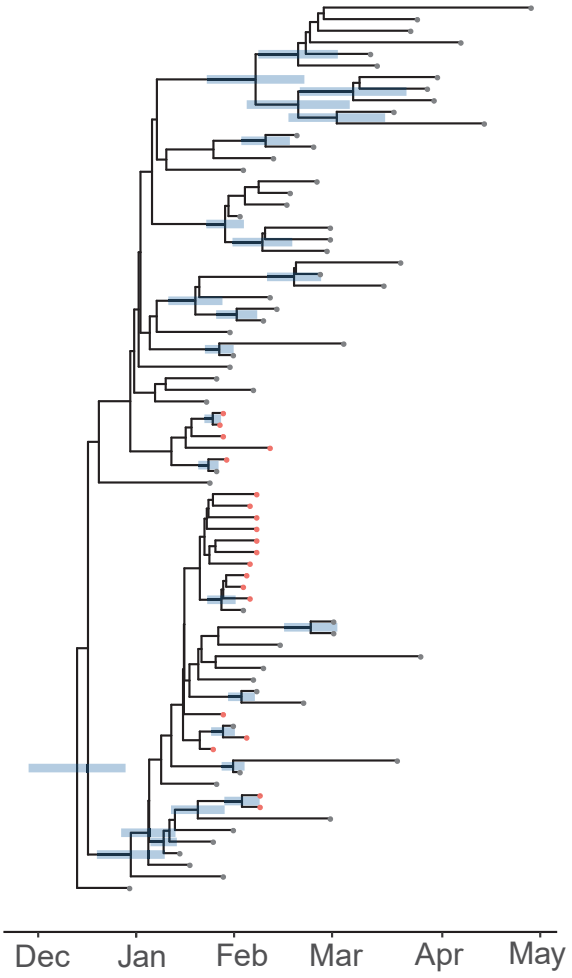


Figure 5.5 A time-scaled phylogeny (maximum clade credibility tree) co-estimated with epidemiological parameters of the SARS-CoV-2 epidemic in 2020. Red and grey tips correspond to samples from inside and outside Weifang, China respectively. The credible interval of time to most recent common ancestor (TMRCA) is shown as a blue bar for all nodes with more than 50% posterior probability support.

Figure 5.6 shows a root-to-tip regression of the fitted tree. Genetic diversity increases approximately linearly with date of sampling, and the estimated mean substitution rate is 1.1×10^{-3} substitutions per site per year. The correlation coefficient was $r^2 = 0.52$; this strong linear relationship suggests that evolution was reasonably reflected by a strict molecular clock.

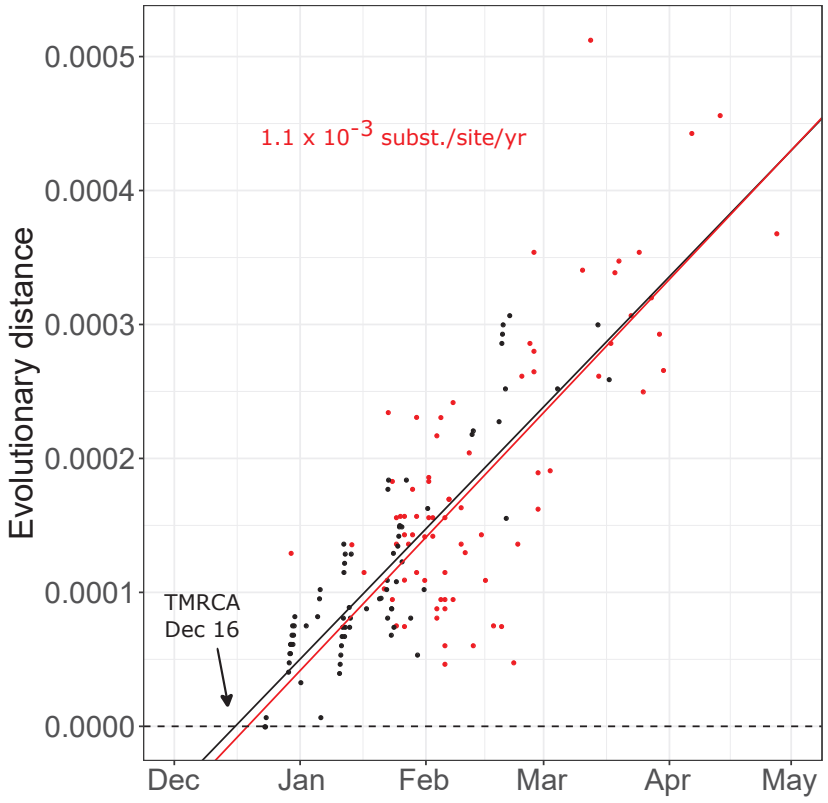


Figure 5.6 A root to tip regression (red and black points indicate sample and internal nodes respectively) showing approximately linear increase in diversity with time of sampling. The red text indicates the mean substitution rate.

5.3.2 Sensitivity analyses

To investigate the robustness of my results to changes in parameterisations, I performed two sensitivity analyses on the initial number of susceptibles and the heterogeneity in offspring distribution.

5.3.2.1 Initial number of susceptibles

In the main analysis, I assumed a prior mean of $S = 500$ for the initial susceptible compartment. I performed a sensitivity analysis on this parameter, changing it from $S = 500$ to $S = 9,086,241$, the latter reflecting the total population of Weifang. Figure 5.7 shows that in this analysis, the reported number of cases is within the HPD of the estimated cumulative infections, and slightly below the central estimate. However, the mean estimated reproduc-

tion number remains constant above 1. There is strong prior belief that the outbreak was controlled (and therefore that R_t must have fallen below 1); I therefore believe these to be unrealistic posterior trajectories, highlighting the impossibility of such a high S prior.

This sensitivity analysis demonstrates that the potential for epidemic decline in our model framework (within a sensible timeframe) is dependent on certain model parameterisations, i.e. a small S prior. It is important to restate that the S compartment is not to be understood as the 'true' number of susceptibles, but is simply included to allow for $R(t)$ to vary, given that the transmission rate parameter (β) is constant.

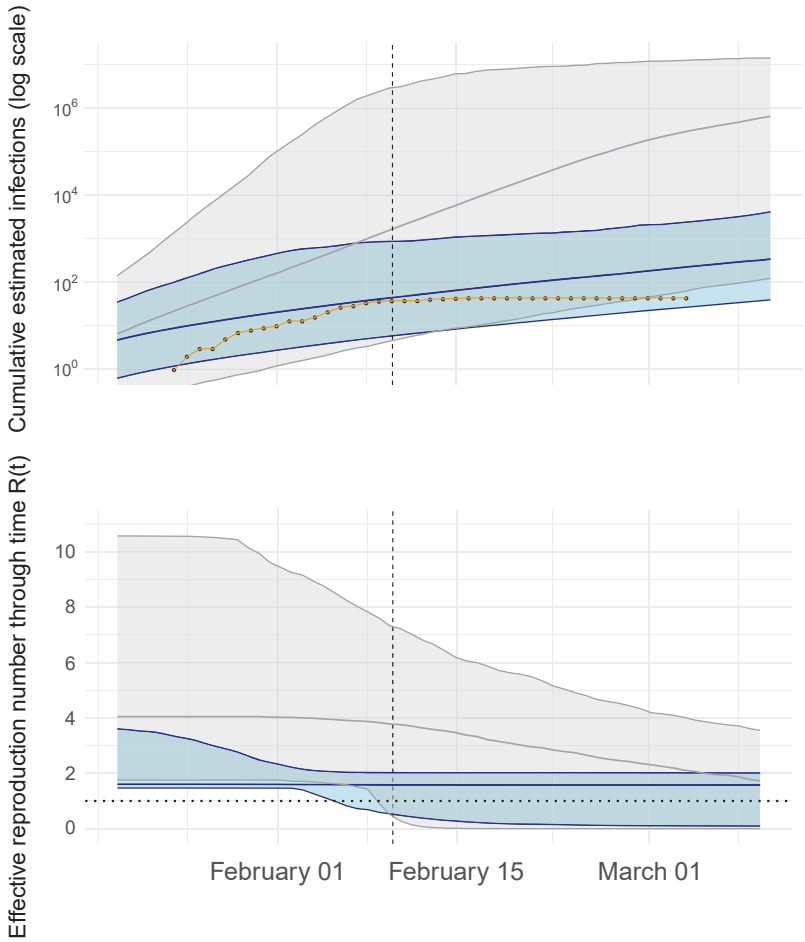


Figure 5.7 Assuming a mean initial susceptible prior of S = 9,086,241, cumulative estimated infections and effective reproduction number through time are shown when fitting the SEIR model to genetic data (blue) and sampling only from prior (grey). Solid lines and shaded area reflect posterior median and 95% HPD. The vertical dashed line represents the date of the last sequence sampled in Weifang; the horizontal dashed line represents $R(t) = 1$. Cumulative cases (yellow points) reported by Weifang CDC.

5.3.2.2 Offspring distribution heterogeneity

I also performed a sensitivity analysis on the heterogeneity of offspring distribution. In the main analysis, the prior distribution of R_0 has mean around 4, which matched a negative binomial distribution with $k = 0.124$, similar to previous modelling from Endo et al 2020 [371]. It is worth noting that in the model, k is largely insensitive to R_0, and very sensitive to τ and p_h. I repeated the analysis while exploring lower offspring heterogeneity by changing from $\tau = 74$ (as in the main analysis) to $\tau = 13$, which resulted in increasing k to 0.2. To maintain a similar mean prior R_0 of around 4 (as R_0 is very sensitive to τ), I increased the prior log-normal distribution mean of β to 4.61 (sd=0.5). Under this parameterisation, the estimated number of infections at last sample (February 10) is 300 (49-1138), compared to 365 (102-1174) in the main analysis. The HPDs are highly overlapping and the conclusions remain the same, but the analysis with lower offspring heterogeneity, as expected, had fewer estimated cases.

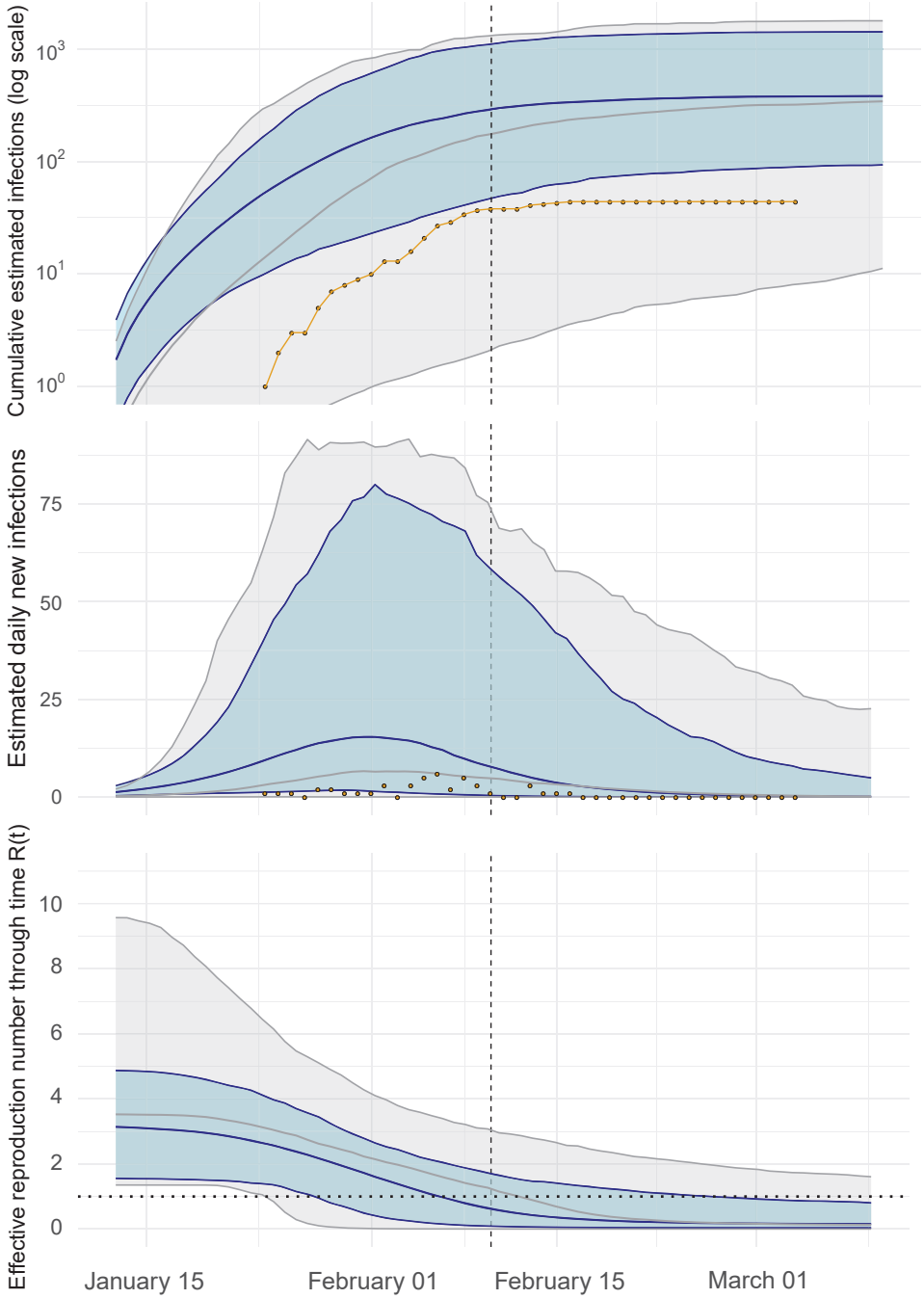


Figure 5.8 Assuming $\tau = 13$ (corresponding to dispersion $k = 0.2$), cumulative and daily estimated infections and effective reproduction number through time are shown when fitting the SEIR model to genetic data (blue) and sampling only from prior (grey). Solid lines and shaded area reflect posterior median and 95% HPD. The vertical dashed line represents the date of the last sequence sampled in Weifang; the horizontal dashed line represents $R(t) = 1$. Cumulative cases (yellow points) reported by Weifang CDC.

5.4 Discussion

In this chapter, I performed a model-based phylodynamic analysis on a small genomic data set of 20 SARS-CoV-2 sequences from the Chinese city of Weifang, as well as 57 sequences sampled from the rest of China and internationally. This modelling framework allows simultaneous estimation of epidemiological and phylogenetic parameters. I estimated the viral epidemic trajectory from January-March 2020, and calculated the percentage of cases that were identified by Weifang CDC. I further estimated the change in reproduction number following the imposition of various NPIs in Weifang. Finally, I produced a posterior time-scaled phylogenetic tree. In this section, I discuss how these results address the key research questions of this chapter.

5.4.1 What was the true number of infections in Weifang?

Despite an initial rapid increase in confirmed cases in Weifang in late January and early February, the number of confirmed cases by Weifang CDC suggest that the outbreak peaked early and the maximum number of daily cases occurred on 5th February 2020. Surveillance of COVID-19 is however rendered difficult by high proportions of illness with mild severity and an unknown proportion of asymptomatic infection [378].

This phylodynamic analysis of 20 SARS-CoV-2 genomes has confirmed independent observations regarding the rate of spread and burden of infection in Weifang, China. The results suggests there was no large hidden burden of infection within Weifang. I estimated that there was a total number of cases around 365 (102-1174) at the date of last sample (10th February 2020), towards the end of the outbreak. This suggests that around 10% of COVID-19 infections were diagnosed in Weifang. This estimate was robust to sensitivity analyses exploring the initial number of susceptibles and the heterogeneity in offspring distribution.

5.4.2 What was the impact of NPIs in Weifang?

Following phylodynamic analysis, I estimated a basic reproduction number of $R_0 = 3.4$, which is equivalent to an initial growth rate of around 22% per day. Fitting the model to genetic data strongly reduces uncertainty relative to simply sampling from the prior. These results are consistent with those estimated in other settings and during the early epidemic in Wuhan [379]. As the epidemic progressed during late January, I estimated a significant

decrease in effective reproduction number ($R(t)$). This coincides with the period of public health interventions, including contact tracing and travel restrictions. I estimated that $R(t)$ fell below 1 on 4th February 2020, consistent with a previous modelling study of Shandong province [380], which showed that R_t fell below 1 on January 29th.

It is widely believed that control efforts reduced epidemic growth rates and contributed to eventual control in China. This phylodynamic analysis supports this hypothesis to be true in Weifang for the following reasons. Given that SARS-CoV-2 has a high R_0, as estimated in this study and elsewhere [379], a decrease in $R(t)$ can be achieved only by either a reduction in transmission or by reaching herd immunity. There is no evidence of widespread transmission in Weifang according to phylodynamic estimates here or independent diagnoses according to Weifang CDC, suggesting herd immunity had been reached. The only reasonable alternative explanation is that NPIs reduced contact rates and transmission, sending reproduction number below one, and thereby prevented the further spread of SARS-CoV-2.

Indeed, one of the interventions implemented was contact tracing, particularly for travellers from Wuhan. As there was widespread recognition of an emerging respiratory infection circulating in Hubei province, other local Chinese authorities implemented preventative interventions to avoid high levels of infections in their provinces. It seems that the strategy of early and intense restrictions were successful at quickly reducing growth rates.

The phylogenetic component to this analysis corroborates this hypothesis. The maximum clade credibility tree suggests that sequences from Weifang are clustered together. If the 20 sequences represented multiple importations into Weifang, this clustering would be 'broken up' by other Chinese samples from the reference set downloaded from GISAID. It is reasonable to therefore conclude that a large degree of the observed cases were due to community transmission, suggesting that the travel restrictions and contact tracing were effective at minimising importations from Wuhan and elsewhere.

The posterior molecular clock rate shown in Table 5.2 is consistent with previous estimates for SARS-CoV-2 [381], adding weight to the robustness of the phylogenetic analysis performed here.

5.4.3 Limitations and strengths

This study has a number of limitations. Firstly, there was a modest number of sequences from Weifang (20). However, this represents a significant proportion of the total number of cases reported; there were 38 confirmed cases at the date of the last genetic sample (10th February), rising no further than 44 from 16th February onwards. Despite relatively few

sequences, our estimated trajectories display uncertainties which are significantly reduced and more realistic, compared to sampling only from the prior.

As there were few sequences used to fit the model, the number of estimated parameters had to be limited in order to prevent overfitting. The sensitivity analyses demonstrated however that the results were robust to altering these fixed parameters to other reasonable values, adding reliability to the study.

It is possible that the outbreak observed in Weifang could be due not to community transmission, but rather multiple importations. However, I explicitly included close genetic matches to the Weifang sequences in generating the reference set. Given that I sampled these from a GISAID database downloaded in June, it is reasonable to assume all close genetic matches would have been chosen. A maximum-likelihood tree of the entire alignment (Figure 5.4) shows that lineages from Weifang have common ancestry with other Chinese lineages at two distinct polytomies and the phylogeny alone gives no information about location of these nodes (Weifang or exogenous). I therefore conclude that the MCC in Figure 5.5, which reflects significant clustering of the Weifang samples, is reasonable.

Community transmission is further supported by the fact that cases were identified via contact tracing. This forms another limitation, as it suggests non-random sampling of cases in Weifang. This could lead to an underestimate of the total number of cases in Weifang. However, as a large proportion of reported cases were included in this analysis, the bias is unlikely to be too significant.

The SEIR model structure also presents some limitations. As β has a constant value, R_t can decrease only as a result of depleting susceptibles. The decrease in R_t is therefore a constraint in the model and occurred even when sampling from the prior. Despite this, the genetic data was informative on the value of β (and therefore R_0), which in turn affects the date at which R_t falls below 1. This analysis demonstrates a reliable mean estimate of R_0, with a narrower uncertainty, compared to sampling from the prior. Although other methods allow for time-varying transmission rate (including other PhyDyn model templates) or models with a piece-wise R_t function [102], this SEIR-type model with constant β required fewer parameters, appropriate for an analysis with only 20 internal sequences.

Further, one of the advantages of a smaller data set is computational speed. This study demonstrates a pipeline for real-time phylodynamic analysis, which could feasibly provide realistic results as a supplement to epidemiological surveillance. However the real-time utility of such methods is dependent on randomised concentrated sampling within localities, coupled with timely sharing of data. Efficient genetic sequencing, processing and data

sharing, coupled with phylodynamic analysis, could prove to be a key tool in the outbreak response toolkit.

Finally, the greatest advantage of this modelling framework is its reliance solely on genomic data, and not on more 'traditional' epidemiological data such as confirmed cases, hospital admissions or deaths. These latter data types are subject to bias for lack of infrastructure for widespread testing, and issues with case ascertainment, case definition and under-reporting. Although fitting models to such data is feasible with various adjustments, there is a danger of uncertainty and bias. Where genomic data is available, model-based phylodynamics is an excellent additional option for estimating epidemiological (and evolutionary) parameters, as well as for investigating the impact of interventions.

5.4.4 Conclusions

This analysis suggests that there was no large hidden burden of COVID-19 in Weifang, China, during the outbreak of January-February 2020. This study corroborates others in the interpretation that NPIs in China were instrumental in controlling the COVID-19 epidemic.

While the value of pathogen genomic analysis is widely recognised for estimating dates of emergence [382] and identifying animal reservoirs [383, 384], analysis of pathogen sequences also has the potential to inform epidemic surveillance and intervention efforts. This is demonstrated clearly in this analysis, drawing on previously developed models and packages for BEAST2 [375, 376], where the results show much narrower uncertainties and more realistic estimates compared to sampling from the prior. Indeed, the added value of fitting to only 20 local sequences in this analysis demonstrates the utility of phylodynamic modelling for outbreaks as compared to traditional epidemiological modelling fitted only to case data.

5.5 Data availability

Genetic sequence data are available from GISAID (gisaid.org).

Accession numbers for sequences from Weifang:
EPI_ISL_413691 EPI_ISL_413693 EPI_ISL_413694
EPI_ISL_413695 EPI_ISL_413696 EPI_ISL_413697
EPI_ISL_413711 EPI_ISL_413729 EPI_ISL_413746
EPI_ISL_413747 EPI_ISL_413748 EPI_ISL_413749

EPI_ISL_413750 EPI_ISL_413751 EPI_ISL_413752
EPI_ISL_413753 EPI_ISL_413761 EPI_ISL_413791
EPI_ISL_413809 EPI_ISL_413692

Accession numbers for sequences from outside of Weifang:
EPI_ISL_414380 EPI_ISL_437621 EPI_ISL_429092
EPI_ISL_418327 EPI_ISL_416335 EPI_ISL_413854
EPI_ISL_402121 EPI_ISL_408480 EPI_ISL_418503
EPI_ISL_450196 EPI_ISL_417030 EPI_ISL_424356
EPI_ISL_451351 EPI_ISL_408010 EPI_ISL_430742
EPI_ISL_416366 EPI_ISL_451343 EPI_ISL_416381
EPI_ISL_407988 EPI_ISL_413882 EPI_ISL_413881
EPI_ISL_413879 EPI_ISL_411954 EPI_ISL_417184
EPI_ISL_418992 EPI_ISL_454935 EPI_ISL_414569
EPI_ISL_416570 EPI_ISL_416600 EPI_ISL_413608
EPI_ISL_451347 EPI_ISL_419242 EPI_ISL_414485
EPI_ISL_414005 EPI_ISL_430847 EPI_ISL_415580
EPI_ISL_413595 EPI_ISL_455376 EPI_ISL_417101
EPI_ISL_417168 EPI_ISL_455410 EPI_ISL_424081
EPI_ISL_440461 EPI_ISL_440433 EPI_ISL_455696
EPI_ISL_444577 EPI_ISL_456208 EPI_ISL_434463
EPI_ISL_437264 EPI_ISL_452673 EPI_ISL_437515
EPI_ISL_437185 EPI_ISL_427257 EPI_ISL_432722
EPI_ISL_437704 EPI_ISL_461275 EPI_ISL_403932

Chapter 6

Estimating the impact of non-pharmaceutical interventions for COVID-19 in England

6.1 Overview

In this chapter, I introduce the COVID-19 epidemic in England, United Kingdom, from January 2020 until 18th March 2021. I describe the available data on SARS-CoV-2 genomes, mobility, and non-pharmaceutical interventions (NPIs). Next, I detail how I sub-sampled and processed these data for use in phylodynamic analyses.

I divided this chapter into four case studies. The first (case study A) investigates the beginning of the epidemic in England before and after the imposition of the first national lockdown. The second (B) studies the city of Leicester during the spring and summer of 2020, where there was a local lockdown. The third (C) explores the second and third waves of COVID-19 in England during the winter period 2020-2021, when there were two national lockdowns. The fourth (D) investigates the same winter period but sub-sampling by lineage, to explore the potential increased transmissibility of variant B.1.1.7 ('alpha'). I conclude this chapter by exploring the association of mobility and NPIs with viral transmission.

Significant portions of this chapter are published or are in peer review. Case study A forms an example of application for the new *mlesky* method, available as a pre-print:

Xavier Didelot, ***Lily Geidelberg****, The COVID-19 Genomics UK (COG-UK) consortium, and Erik M Volz. "Model design for non-parametric phylodynamic inference and applications to*

pathogen surveillance". Pre-print available at:
https://www.biorxiv.org/content/10.1101/2021.01.18.427056v2

Case study D was featured in an article exploring the increased fitness of the B.1.1.7 lineage (e.g. Fig. 1b and 1c, Extended Data Fig. 3).

*Volz, E., Mishra, S., Chand, M., Barrett, J. C., Johnson, R., **Geidelberg, L.**, ... & Ferguson, N. M. (2021). Assessing transmissibility of SARS-CoV-2 lineage B. 1.1. 7 in England. Nature, 593(7858), 266-269.*

Note that an earlier version of this article was published as a report with the following citation:

*Volz, E., Mishra, S., Chand, M., Barrett, J. C., Johnson, R., **Geidelberg, L.**, ... & Ferguson, N. M. (2021). Transmission of SARS-CoV-2 Lineage B. 1.1. 7 in England: Insights from linking epidemiological and genetic data. MedRxiv, 2020-12.*

6.2 Introduction

6.2.1 The COVID-19 epidemic in England

The first cases of COVID-19 in the United Kingdom were confirmed on 31st January 2020 in York [385]. One year later, on 31st January 2021, the number of confirmed cases had risen to over 3.8 million people, with over 100,000 deaths attributed to the disease [386]. The SARS-CoV-2 epidemic was found to be the deadliest respiratory illness in the UK since the Spanish flu 100 years prior [387].

This epidemiological impact was in spite of an unprecedented breadth and depth of public health interventions implemented by the British government. In late January 2020, contact tracing was attempted on all air passengers from Wuhan in the two weeks prior, following advice from the Department of Health and Social Care [388]. In the following weeks, the alert risk level of the disease was raised incrementally, as the number of cases began to rise across England, Scotland, Wales and Northern Ireland. On 12th March 2020, the risk level was raised to "high", and anyone with possible COVID-19 symptoms (a continuous cough or fever) was recommended to self-isolate for seven days [389]. The government further advised on 16th March that people avoid pubs, clubs, theatres, non-essential travel and contacts with others, and that working should be remote where possible [390]. The following day, NHS England announced that, in order to free up 30,000 extra beds in hospitals for COVID-19

patients, all non-urgent operations were to be cancelled [391]. By 20th March schools were closed and public examinations were cancelled [392].

On 23rd March 2020, the UK strategy for mitigating the spread of COVID-19 changed from advisory to compulsory. The Prime Minister announced that all areas of the UK will be subject to a range of restrictions on movement [393]. On top of the closure of all leisure facilities, pubs, restaurants, shops and most businesses, all those residing in Britain were put under a stay-at-home order; leaving home was permitted only for "essential" reasons: food shopping, medical needs and work which could not be performed remotely. Exercise was permitted, outdoors, once a day. With funerals being the only exception, there was a full prohibition on all social gatherings and mixing of households. The mandated stay-at-home order was widely known as "lockdown".

The impetus for lockdown was supported by epidemiological and modelling studies conducted in February and March, which estimated the reproduction number (R) of the virus [39]. As confirmed cases, hospitalisations and deaths seemed to follow an exponential curve, public health officials were concerned that hospitals would soon be overwhelmed by patients with acute COVID-19, having been observed in Northern Italy a few weeks prior [394]; such a scenario was avoidable only if the reproduction number could be brought below one.

A few weeks later, in late April 2020, the number of daily confirmed cases started to decline [386]. The majority of public health officials, epidemiologists and mathematical modellers agreed that the decline could not be attributed to herd immunity [38, 232, 69, 68]. Although a substantial proportion of infections were asymptomatic [190], it was estimated that an insufficient number of individuals had been infected, and that the vast majority remained susceptible. This was further evidenced by the fact that, following a loosening of restrictions in the early summer, cases began to rise again.

On 30th June, following rising cases in Leicester superseding that of the national average [395], the city went into a 'local lockdown', whereby shops and restaurants were closed, and movement into and out of the city was restricted [396]. This local lockdown proceeded for around a month, and was lifted on 1st August. Subsequent local lockdowns were imposed on Greater Manchester, east Lancashire and West Yorkshire [397].

By October 2020, numerous regions were experiencing rising cases; the UK was declared to be suffering its 'second wave'. The government introduced a 'tier system' whereby local regions were ranked according to their case burden, and the stringency of restrictions imposed reflected its tier [398]. Regions with low burden were in tier 1 and received mild restrictions, while those with high burden were in tier 3 and received harsher restrictions. This heterogeneous application of intervention was thought to be a fairer response to the

highly variable observed case counts throughout the country. However, cases continued to climb and a second national lockdown was imposed on 5th November 2020 [399].

Daily case counts fell towards the end of November, so the government lifted lockdown on 2nd December 2020 and re-introduced the tier system [400]. However, this second implementation included four tiers, and more regions were placed into higher tiers compared to October. On 19th December, London and other regions in the South-East of England were placed into Tier 4 [401], the highest tier, as cases continued to rise faster than anticipated. Daily case counts exceeded those of the first wave during the Christmas period, and so the UK government imposed a third lockdown on 5th January 2021, this time also heavily restricting international travel [402]. The winter period of 2020/2021 saw daily COVID-19 related deaths far exceed 1,000 per day [403]. The unforeseen increase in COVID-19 cases was attributed to a novel variant (B.1.1.7) which was estimated to have increased transmissibility compared to previously circulating strains. I discuss the emergence of this strain in Section 6.2.2.1, and describe my analysis of its transmissibility in Section 6.4.

As of 18th March 2021, the cumulative numbers of confirmed cases and COVID-related deaths in the UK exceed 4 million and 125 thousand respectively [386].

6.2.2 SARS-CoV-2 sequence data

On the date of the first national lockdown, March 23rd 2020, the government and UK Chief Scientific Advisor backed the newly-formed COVID-19 Genomics UK Consortium (COG-UK) with an investment of £20 million [404]. The objective of COG-UK was to produce large-scale and rapid sequencing of SARS-CoV-2, the virus causing COVID-19. COG-UK became a large collaboration between the NHS, public health agencies, the Wellcome Sanger Institute, several universities and research institutions.

This enabled high-throughput sequencing in the UK that eclipsed that of other countries, particularly during the first several months of the pandemic. By September 2020, the largest single nation contribution to the global repository of SARS-CoV-2 sequences was from the UK [405], representing about 40% of all sequences. As of 23rd April 2021, 728,344 sequences from the UK had been submitted to COG-UK; the total number of sequences submitted to GISAID was just under three million [179].

The COG-UK central database included sequences acquired from two broad sampling strategies: 'pillar 1' and 'pillar 2'. Pillar 1 includes patients who are unwell with COVID-19 in hospitals. The weekly number of sequences requested from each NHS site by COG-UK is decided prospectively. Sites are asked to provide the first available sequences each day

('surveillance sampling'), without any particular pre-defined selection criteria, so as to make it reasonably random. Local areas within the surveillance of each NHS site should be sampled proportionate to its population size. This surveillance sampling makes up 50% of their allocated daily number requested. The remaining 50% comprises samples which meet local requirements ('targeted sequencing'), to answer particular local research questions. Sites may be requested to prioritise samples from certain risk groups, sample types, locations and viral lineages. Taken together, the pillar 1 database provides detailed data on specific hospital areas and patient groups, and is therefore best suited to answer focused research questions. It is however not ideal for obtaining a representative national picture of the epidemic.

Pillar 2 is based on residual materials from community testing for COVID-19. It is considered to be a more random and representative sample from the UK compared to Pillar 1, as it will not bias for samples taken from hospitals; its composition is similar to designed samples such as REACT-1 and REACT-2 [406]. It aims to provide coverage of England, Scotland, Wales and Northern Ireland, proportionate to the burden of confirmed cases. Where possible, I used pillar 2 sequences for my analyses, as they provided a closer-to-random sample of the country.

The extensive metadata labelling of the COG-UK data enabled efficient sub-sampling by various characteristics, including time and location. Further, as part of the daily pipeline, sequences were aligned within a national and global maximum-likelihood phylogenetic tree, and were thereby labelled by lineage. This enabled researchers to sort by lineage to identify those increasing in frequency, and to estimate elevated transmission and virulence.

6.2.2.1 The variant of concern B.1.1.7

Although SARS-CoV-2 does not evolve at a notably fast pace [407], the high number of transmission events across the world provide ample opportunities for novel mutations to arise and become fixed. Many institutions, including COG-UK, were therefore closely monitoring hospital and community samples for novel variants. Non-synonymous mutations which are likely to cause functional effects in viral infection were of particular interest. An example was B.1.177, a lineage that likely emerged in the summer of 2020 in Spain [408], which quickly rose in frequency across Europe. It was characterised by a substitution on the spike protein (A222V), which initially led to concerns of increased transmissibility (and possibly virulence). Transmission advantage is still debated, but it showed in any case that rapid fixation of novel variants was possible [408].

In December 2020, Public Health England (PHE) announced an increase in monitoring of a novel strain, located within the B.1.1.7 lineage, which became a 'variant under investigation' (VUI – 202012/01) [409]. By 13th December, over one thousand cases within this lineage were identified, mostly in the South East of England, particularly in Kent. There seemed at the time to be a correlation between high frequency of this variant and a large burden of confirmed COVID-19 cases [409]. Retrospective analysis of the COG-UK sequence database revealed a phylogenetic cluster in Kent of 117 samples over the week 10th-18th November 2020 [410]; other samples within the same cluster appeared as far back as 20th September [411]. As investigation continued, the proportion of samples located within this lineage increased, and it became designated a variant of concern on December 18th. By January 1st 2021, B.1.1.7 accounted for around 80% of all samples in the UK, and exceeded 95% by February; it increased to significant frequencies around the world in the following months [412].

Compared to the average accumulation of around 1-2 mutations per month in most SARS-CoV-2 lineages [413], B.1.1.7 is notable for having accrued 14 lineage-specific amino acid replacements prior to its detection [411]; in other words, the high rate of molecular evolution occurred in the phylogenetic branch ancestral to B.1.1.7. In total, B.1.1.7 is defined by 23 mutations [410], eight of which are in the viral spike protein. The mutation N501Y in the spike protein has been associated with increased ACE2 receptor affinity, which has been experimentally shown to increase viral entry into hosts and therefore boost viral infectivity [414]. Another mutation, P681H, is located immediately adjacent to the furin cleavage site, which is also implicated in viral infectivity [415]. A final notable mutation is a deletion at positions 69 and 70 on the spike protein (Δ69-70), which has been independently described as causing diagnostic test failure for common commercial assays [416]. Three of the laboratories which test community cases in the UK use a three-target assay focusing on the nucleoprotein (N), ORF1ab and spike glycoprotein (S) genes [410]. The Δ69-70 mutation causes S-gene target failure (SGTF) in otherwise positive samples. Although possible for other lineages, the vast majority of instances of SGTF were associated with the B.1.1.7 by the end of 2020 in England [417]. Indeed, increasing frequency of SGTF was used as a proxy for measuring increased frequency of B.1.1.7 [410, 101]. Various statistical, mathematical modelling and phylodynamic analyses estimated an increased fitness of B.1.1.7 compared to previously circulating lineages [410, 418]. It has been suggested that the failure to control the virus during the tier system period of December 2020 was due to the emergence of this more transmissible lineage [419].

6.2.3 Estimating the impact of lockdown

6.2.3.1 The challenges

In 2020, the UK economy suffered the greatest decline since 1709 [420], and other medical concerns surrounding cancer, mental health and obesity mounted [301, 421, 422]. There are many who attribute these serious non-COVID related costs to lockdowns. Although polling suggests that the public broadly agreed with imposing lockdowns [423], there has nevertheless been significant political and social conflict.

Proponents of lockdown argue that it has the greatest impact on reducing transmission and COVID-related deaths, a belief supported by many studies [38, 232, 69, 68]. Although costly, its repeated deployment is therefore justified on these grounds. It is nevertheless crucial to scientists, governments, policymakers and the general population to gather as much evidence as possible on the effectiveness of lockdowns and other non-pharmaceutical interventions.

In the weeks following the imposition of the first national lockdown, confirmed COVID-19 cases fell. After the relaxation of restrictions, cases once again began to rise. The second (and third waves) in the UK demonstrate that the population remained largely susceptible, assuming that re-infection is rare [424]. It is therefore reasonable to conclude that herd immunity had not been reached in the spring of 2020, and that the observed reduction in cases could only be attributed to a decreased contact rate. Many have therefore concluded that lockdown succeeded in controlling the epidemic in the UK.

The question, however, is made complex by several factors.

Firstly, lockdowns have always been instigated in parallel to many other interventions. Prior to the first national lockdown on March 23rd 2020, the UK government had implemented several less severe interventions as described earlier. Some of these, for example school closures and public health messaging regarding hand washing, had been previously used to combat other pathogens such as influenza [425]. Economic relief was provided in several packages throughout the year for various reasons, e.g. rough sleepers to obtain accommodation, benefits for those temporarily unemployed due to the virus (the 'furlough' scheme), and for local governments. Disentangling the relative effects of all concomitant interventions is very difficult.

Secondly, news reports from China, Italy and elsewhere had heightened the awareness and fear of many members of the public of COVID-19, and it is likely that many restrictions and behaviour changes were self-imposed [426]. Indeed, using GPS locations from mobile phones, Google released data that suggested decreased movement preceded the enactment

of lockdown [427]. Determining the counterfactual scenario against which to measure the impact of lockdown is, therefore, non-trivial. Indeed some have argued that less stringent interventions, such as those imposed in Sweden, would have been enough to reduce social contacts and control the epidemic [428].

Thirdly, the time horizon over which to measure impact is important. It is possible that COVID-19 becomes endemic in the future [429], and thus all individuals may eventually become exposed [430]. Interventions may therefore not be able to prevent infection, but rather delay it. The likelihood of this scenario is decreased with the introduction of vaccines beginning in December 2020 in the UK [431], but potential vaccine escape variants have already been detected which could plausibly increase the susceptibility of a highly vaccinated population [432]. In any case, measuring the impact of lockdown over 3 months and over 10 years may provide very different estimates.

Fourthly, even if it were established that lockdown decreases transmission, the amplitude of this effect is of equal importance. The justification for lockdown strongly depends on whether reduction is in the order of 5% or 50%.

Finally, even if correlations between lockdown and decreased transmission are demonstrated, it can be difficult to establish the direction of causality, or indeed if there is a causal link at all. It is clear that governments across the world increased restrictions as cases rose, but careful analysis is required to determine that changes in stringency cause a change in cases.

6.2.3.2 Previous studies on the impact of lockdown

Several modelling studies have aimed to estimate the impact of lockdown in the UK. Below I detail the model analyses of groups from Imperial College, the London School of Hygiene and Tropical Medicine (LSHTM) and University of Warwick, whose members contribute to UK COVID-19 response policy through the Scientific Advisory Group for Emergencies (SAGE).

Imperial College published an influential report on 16th March 2020, prior to the national lockdown, estimating the impact of various NPIs on COVID-19 cases and deaths in the UK [38]. Using an individual-based simulation model original developed for pandemic influenza, they modelled the epidemic trajectory under various scenarios. They parameterised SARS-CoV-2 natural history (i.e. severity, generation time, percent infections symptomatic) based on data available at the time from Italy and China. In their scenario assuming an entirely unmitigated epidemic, and a reproduction number of 2.4, they estimated COVID-19 would cause 510,000 deaths in the UK, with significant mortality caused by the overwhelming of

ICU capacity; if all patients could be treated, they estimated around 250,000 deaths in the UK. They concluded that reproduction number could be brought below 1 under a combination of case isolation, social distancing of the entire population and either household quarantine or closing schools and universities.

After the first wave, Imperial College used a Bayesian mechanistic model to assess the impact of NPIs in several European countries including the UK, fitting to observed COVID-related deaths [232]. They estimate that the only intervention to cause a detectable impact on reproduction number was lockdown, by reducing transmission by 81%. The impact of other NPIs, such as closing public events, school closure and isolation of symptomatic cases were estimated to not significantly differ from zero.

A third modelling study from Imperial College sought to evaluate the epidemic in England during the whole year of 2020 [69]. They fitted an age-structured stochastic transmission model to multiple outcomes in surveillance data, including hospital admissions, ICU occupancy, deaths, testing, serological surveys and SARS-CoV-2 prevalence. They conclude that the timing of lockdown had a huge impact on epidemic size and therefore mortality. If lockdown had been imposed a week earlier, they estimate that the first wave death toll would have been reduced by over 50%, and that a week delay would have increased deaths by nearly three-fold. They argue that the Government's *'Eat Out To Help Out'* scheme increased the reproduction number to above 1 in all regions of England, and that the subsequent tiered restrictions reduced transmission but could not bring the reproduction number below 1. They estimate this was only possible with the second lockdown from 5th November to 1st December 2020, in conjunction with population immunity.

The Centre for the Mathematical Modelling of Infectious Diseases at LSHTM also published a study of the impact of NPIs following the first national lockdown [68]. Using a stochastic age-structured SEIR-type transmission model, they estimated that an unmitigated epidemic would have caused 350,000 COVID-related deaths by the end of 2021, without accounting for an increase in CFR if hospital capacity were exceeded. They projected that a strong lockdown would reduce deaths to 50,000, a seven-fold decrease compared to the unmitigated scenario. Like the Imperial College modelling study, they concluded that reproduction number could be brought below 1 only if lockdowns were imposed, and that other NPIs such as school closures and shielding of vulnerable people were not sufficient.

A group at the University of Warwick developed an age-structured transmission model which they fitted to hospitalisations and mortality data from the UK [70]. They estimated that an unmitigated epidemic would have led to 200,000 deaths by the end of 2020, which was reduced to 39,000 deaths if lockdown were imposed throughout the whole year. They

concluded that lockdown was required to control the epidemic, and that other scenarios were not as effective. Further, the release of lockdown, in lieu of a vaccine, would lead to a similar number of deaths as a totally unmitigated scenario, assuming contact mixing patterns return to as before.

In summary, all models from several key modelling institutions have concluded that control of the UK COVID-19 epidemic, represented by a reduction in R to below 1, could be achieved only by full lockdown, in the absence of widespread vaccination.

6.2.3.3 Previous studies on the impact of the tier system

The LSHTM model was also used to evaluate the effect of the tier system introduced in the autumn of 2020 in England. They estimate that for the period of October 2020 to March 2021 inclusive, tier 2 and tier 3 restrictions reduced reproduction number by 2% and 10% respectively; lockdown, including with school closures, reduced reproduction number by an estimated 36%, to a value below 1.

A separate modelling study was conducted by Imperial College to estimate the impact of the tier system between the first and second national lockdowns in the UK [433]. Using a Bayesian hierarchical model fitted to estimated reproduction number by lower tier local authority (LTLA) over time, they estimated that tier 1 had no significant impact on reproduction number, and tiers 2 and 3 reduced it by 6% and 23% respectively. They state in their limitations however that this effect is associative and not causal.

6.2.4 Non-pharmaceutical interventions as a continuous variable

As of March 18th 2021, there have been three announcements of national lockdown: March 23rd 2020, November 5th 2020 and January 5th 2021. These interventions can be understood as binary variables - lockdown or no lockdown - and their impact on epidemiological parameters can be estimated. However, this binary reduction can occlude significant explanatory power of the available data. As discussed, there were several other NPIs which are likely to have contributed to transmission of COVID-19.

Among the most prominent attempts to quantify the extent of varied NPIs is the Oxford COVID-19 Government Response Tracker (OxCGRT) [214]. This project collated information on governmental responses across the world for a number of indicators. These indicators were grouped into three main families: Containment and closure (C), Economic response (E), and Health systems (H). Containment and closure included stay-at-home requirements

(lockdowns), school and workplaces closures, restrictions on travel. Economic responses included income support, international support, debt relief and fiscal measures. Health systems included public information campaigns, testing policies, contact tracing, facemasks and vaccines. Importantly, they converted the presence, absence and duration of these indicators into quantitative measures. Two of the main measures are the health and containment index, and the stringency index, which are select combinations of the above indicators. These indices are downloadable for each nation as a time-series. OxCGRT has been used in numerous studies, from estimating associations between levels of healthcare resources and viral transmission, to estimating impacts of physical distancing on transmission and time-varying reproduction number [434].

Another source of useful data are the Google COVID-19 community mobility reports [435]. Using GPS location from mobile phones, they provide open-access time series data on community mobility for each day since the beginning of the pandemic. Compared to baseline values, they report the changes in daily movement for six categories: retail and recreation, grocery and pharmacy, parks, public transport, workplaces, and residential. These changes can be compared to a "baseline", calculated as the median value for the current week or previous weeks. It is therefore possible to compare changes in mobility following the emergence of COVID-19 in the UK, as well as throughout the course of the epidemic. Google mobility data has been used to estimate the association between movement and COVID-19 transmission [436] and death [437].

Although both the OxCGRT and Google mobility data sets have been widely used, adding support to their reliability, they can be mutually validating; significant associations between government responses and mobility have been directly estimated [434].

Further, Koh, Naing and Wong performed a regression analysis across 142 countries on stringency and mobility estimated from OxCGRT and Google respectively, against time-varying reproduction number estimated from the COVID-19 Epiforecasts project [438, 439]. They find that an increase in stringency index by 10 points is correlated with a reduction in R(t) of 6%. They argue that lockdowns should not be considered a binary variable because they vary in stringency, from working from home to complete movement restrictions. They estimate partial and full lockdowns to reduce R(t) by 38% and 32% respectively, but that the effect sizes are not significantly different between them. They instead conclude that earlier lockdowns are more impactful than late ones. They estimate that complete travel bans reduce R(t) by 24%. To address issues concerning causality between policy and viral transmission, they compared R(t) as a median over the 14 days after the date of the explanatory variable (e.g. stringency index).

6.2.5 Aims and justification

As detailed in Section 6.2.3, understanding the impact of lockdown on viral transmission is of critical importance. Few studies have attempted to estimate this using genomic data; none to my knowledge have done so in the English context. The insight gained in doing so would be a useful complement to those analyses working with traditional epidemiological data.

A useful (and easy to interpret) epidemiological parameter obtained from phylodynamic modelling is effective reproduction number $(R(t))$. This statistic has been widely used in policy documents [410], and has heavily featured even in non-scientific media outlets [419, 440]. As described in Section 1.1.2.2, $R(t)$ is calculated from growth rate, which is the rate of change in effective population size (N_e).

Therefore in this chapter, using phylodynamic modelling methods, I estimate the epidemic trajectory of SARS-CoV-2 in England in terms of N_e, and compare it to independent case data published from official sources. Secondly, I aim to estimate $R(t)$ through time, and test its association with the imposition of NPIs including lockdowns, and changes in mobility through time.

The results of these analysis enable me to answer the following research questions, according to the aims of the thesis:

What was the impact of NPIs on COVID-19 transmission in England?

Was the B.1.1.7 variant spreading faster than previous variants in England?

To do this I use high quality publicly available time-series data on mobility and measures of NPI stringency ("covariates"), and compare their associations with my estimates of $R(t)$. Finally, I integrate these covariates in the phylodynamic modelling analysis (see Section 1.1.2.2). After fitting to both genomic and epidemiological co-variate data, growth rates can be re-estimated, and associations between growth rates and covariates can be measured.

Throughout 2020 and the beginning of 2021, there has been significant heterogeneity, both over time and region, in terms of COVID-19 transmission dynamics, government intervention, testing capability, case definition, genomic sampling intensity and viral genetic diversity. It is therefore more practical and useful to focus research questions on smaller time periods or locations.

I therefore present four case studies in this chapter. The first (case study A) considers the beginning of the epidemic in England before and after the imposition of the first national lockdown. I investigate whether the reduction in case burden was attributable to lockdown.

The second (B) focuses on the city of Leicester during the spring and summer of 2020, in which there were both a national and a local lockdown. Here, as well as estimating the impact of lockdowns, I explore whether the high case burden observed in Leicester in the summer months was detectable from the genomic data, even when only using small fractions of the samples. As policy is often tailored to the region depending on local viral activity, it is interesting to see whether local trends can be identified from available genomic data, before the increase in cases and hospitalisations; this case study provides a good setting to do this.

The third case study (C) examines the epidemic in England during the winter months (2020-2021) during which there were interchanging periods of little restriction, tier systems, and two national lockdowns. I aim to evaluate the effectiveness of the observed changing policies on viral transmission.

The fourth case study (D) also focuses on the winter period, but here I accommodate some of the observed heterogeneity in viral lineage that had emerged in this more advanced stage of the epidemic. By estimating the epidemiological trajectories for different lineages, I estimate whether the variant of concern B.1.1.7 had a higher transmission than previously circulating lineages, and investigate whether national efforts to contain the COVID-19 epidemic in England were affected by the emergence of this new variant.

Beyond the importance to scientists, policymakers and the general public, the research questions I undertake are justified methodologically. Firstly, the UK had by far the largest repository of SARS-CoV-2 sequences in the world during the first months of the pandemic [412], and thereby represent a significant opportunity for insights from genomic epidemiology and population genetics, which can complement conclusions derived from more 'traditional' surveillance and epidemiological modelling. However, the nature of the devolved governmental structure of the UK is such that England, Scotland, Wales and Northern Ireland are each able to produce independent COVID-19 response strategies. For this reason I opted to focus on the English epidemic, where the largest genomic data set resides. Further, the likely origin of the B.1.1.7 variant is in South East England, and the close monitoring from its emergence to its near-fixation provides an excellent knowledge base with which to study its impact.

Secondly, as the number of SARS-CoV-2 sequences in the UK approached half a million as of April 13th 2021, the preferred phylodynamic methods are those that can cope with large data sets. Non-parameteric methods such as those described in Section 1.1.2.2 are therefore ideal to maximise the benefits of such a rich genomic data set.

6.3 Materials and Methods

The process for producing epidemic trajectories is summarised as follows. First I summarised the NPI data (including lockdowns) against which I could compare the phylodynamic results. Secondly, I sampled the COG-UK sequence alignment for each case study, described in Section 6.3.3. Next I produced maximum-likelihood trees, then dated them, as described in Section 6.3.4. Finally, I fitted *skygrowth* and *mlesky* phylodynamic models to the dated trees, either with or without NPI data encoded as covariates, as described in 6.3.5.

6.3.1 Summarising NPI covariate data

Binary lockdown variable

There were three UK national lockdowns imposed on March 23rd 2020, November 5th 2020 and January 5th 2021. To encode this information for inclusion in the phylodynamic analyses, I created a model matrix equivalent to $\mathbf{Z}$ as described in Gill et al 2016 [94]. Each lockdown was assumed to be a binary predictor variable ($\mathbf{Z}_j$) of length $t = 412$ elements corresponding to each day between 1st January 2020 and 15th February 2021. For each time point, the value of the predictor was set to 1 if the lockdown was present. For example, for the first lockdown, each value between 23rd March 2020 and 12th May 2020 was set to 1, and the rest set to 0.

Continuous NPI variables

However, as discussed in Section 6.2.4, lockdown is not the only NPI under consideration. In this section I describe how I analysed the data from two main sources: mobility data from Google and the indices of government intervention by OxCGRT. These two data sources provide a richer data set against which to explore associations with SARS-CoV-2 reproduction number.

Figure 6.1 shows the raw data over time, for different categories of movement, during the early months of 2020 (black lines). These data are the covariates against which I compare $R(t)$ in case study A: England before and after the first national lockdown. The raw data was highly variable by day, reflecting the changing mobility throughout the week. To produce intelligible comparisons with reproduction number, the raw data had therefore to be smoothed. I used a local regression smoothing algorithm (*'stats'* R package, *'loess'* function).

The changes in mobility behaviour in these months were very sharp. Therefore, after smoothing, the sheer drop in mobility was only captured when using a small smoothing parameter alpha (e.g. 4%). This means that local regressions are calculated only for the very nearest neighbouring points, allowing for sharp local changes to be respected. In general statistical language this translates to a small bias. However, bias is in trade-off with variance: a smoothing parameter too small will produce a curve with high variance (very "bumpy"), one which is overfitted to the data. Figure 6.1 shows that a parameter value of 4% performs well at balancing these two effects, producing a smooth and accurate recapitulation of the raw data. I therefore proceeded to use this smoothed data (with k=4%) in the analyses described later. In Section 6.5.2.3, I discuss a sensitivity analysis where I instead assume k=20%.

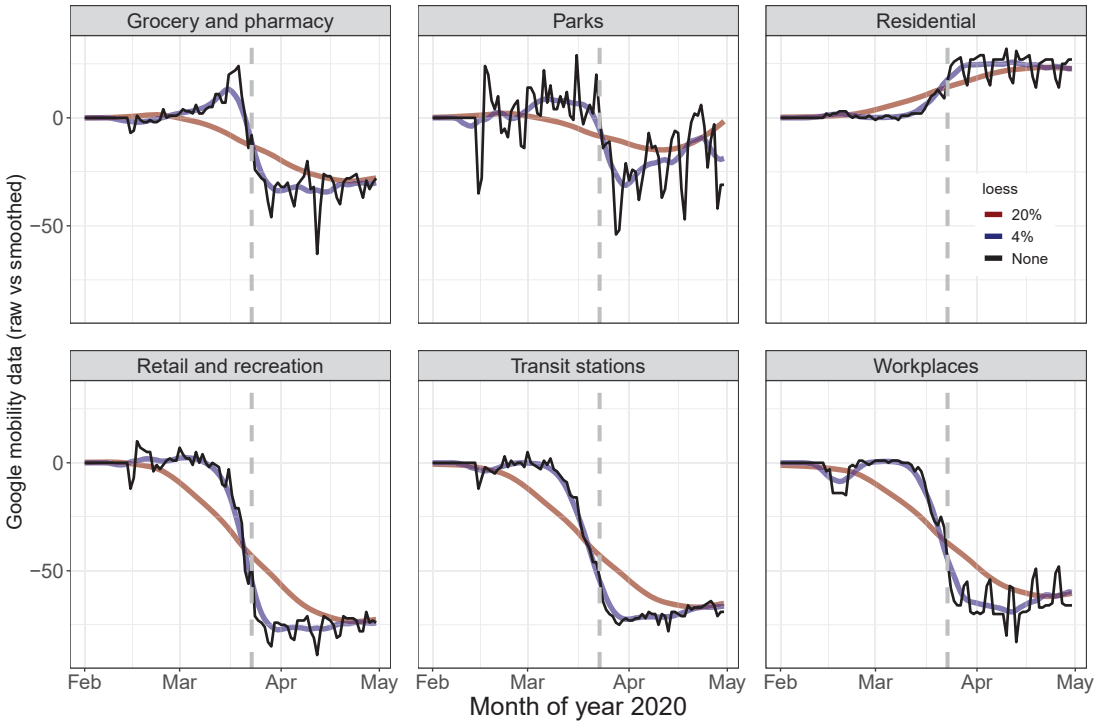


Figure 6.1 National time series data of movement trends for different location types during spring 2020, published by Google. Thin black line represents raw data. Thick red and blue lines represent local regression smoothing with parameters of 20% and 4% respectively. The dashed grey lines indicate date of national lockdown.

I performed a similar exercise for the Oxford COVID-19 Government Response Tracker (OxCGRT) measures of government intervention, summarised in the 'Containment and

Health' and 'Stringency' indices. The 'raw' estimates of these indices for spring 2020 are shown in Figure 6.2 (black lines).

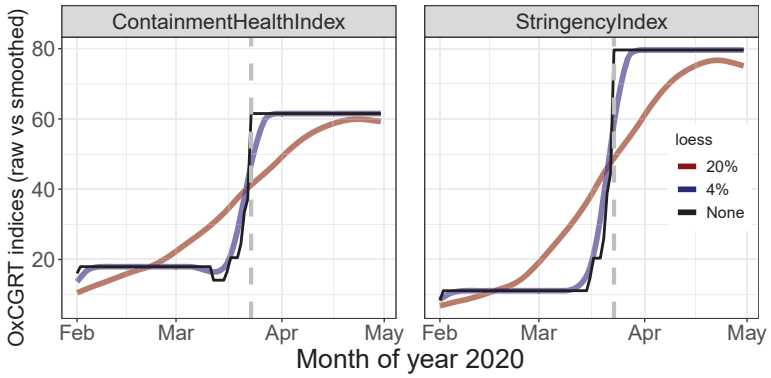


Figure 6.2 Time series composite data of UK government interventions in the early months of 2020, estimated by Oxford COVID-19 Government Response Tracker (OxCGRT). Thin black lines represent raw data. Thick red and blue lines represent local regression smoothing with parameters of 20% and 4% respectively. The dashed grey lines indicate date of national lockdown.

The same trade-off between bias and variance is observable. However, I had a different interpretation of the best degree of smoothing. While mobility provides a direct measure of changing public behaviour, a change in policy does not. Further, the enforcement of policy changes is not immediate. The OxCGRT index shifts occur when changes of policy are announced, and this is represented as a near-instantaneous change. For instance, the announcement of a national lockdown on March 23rd can be seen as an overnight three-fold or five-fold 'jump' in containment and stringency indices respectively. In reality, the manifestation of this policy change is unlikely to be so instantaneous.

Using a 4% smoothing parameter (blue line) value closely imitates this instantaneous jump in policy, while a 20% value (red line), by encompassing a wider set of data points in each local regression, reflects a more incremental change in policy. The bias in such smoothing is clearly seen: the indices 'overestimate' prior to the announcement of lockdown, and 'underestimate' following it. However, I believe that a more incremental curve is a more realistic representation of the implementation of government policy, even if the announcements were themselves instantaneous events. I therefore used the curve produced with a 20% smoothing parameter value in my analyses. I discuss the robustness of this decision in Section 6.5.2.2, by implementing a sensitivity analysis using a 4% smoothing parameter value.

However, it is arguably realistic to assume that all changes in government policy do manifest incrementally, but that this change should strictly follow policy announcements, not precede

them. In other words, the 'slowness' of policy implementation should be interpreted as a rightwards lag. To account for this, I explore the effects of introducing rightward (and leftwards) lag in government policy, also discussed in Section 6.5.2.2.

Another justification for opting for a shallower smoothing curve is the ability to more easily encode first order changes in the data. Reproduction number $(R(t))$ is related to growth rate, which is the first-order rate of change of (N_e); it is therefore most intuitive to compare it to the first-order rate in change of the covariates. In the curve with a 4% smoothing parameter, the flat portions and instantaneous changes of the raw data are mostly preserved. The rate of change would therefore be zero throughout most of the time series, with very extreme values for short periods of time around the date of national lockdown. With such high skew, it would be difficult to compare this rate of change against reproduction number. In the 20% smoothing parameter case however, the more gradual curve enables a greater heterogeneity - and reduced skew - of rates over time, allowing an easier comparison.

The same transformations of the Google mobility and OxCGRT NPI data were applied for covariate data used in case studies B (Leicester during spring and summer 2020) and C and D (England during the winter period 2020-2021). For similar reasons in each case study, I used a loess parameter of 4% for Google mobility and 20% for OxCGRT.

Figures 6.3 and 6.4 display the smoothed mobility and index data used for the Leicester analyses respectively, used in case study B. Google provided mobility data on a regional level; the time series for Leicestershire was extracted to be used in case study B. However, as of March 18th 2021 when this analysis was conducted, there were significant portions of the mobility data missing for the 'parks' variable. The degree of missingness was too great to be able to interpolate between the data points that were reported. I therefore used the national level data for this comparison, and interpretations of results must therefore be made more cautiously. The OxCGRT did not report interventions on the local level, thus I used the national indices. Again, interpretations of these results are therefore subject to more caution.

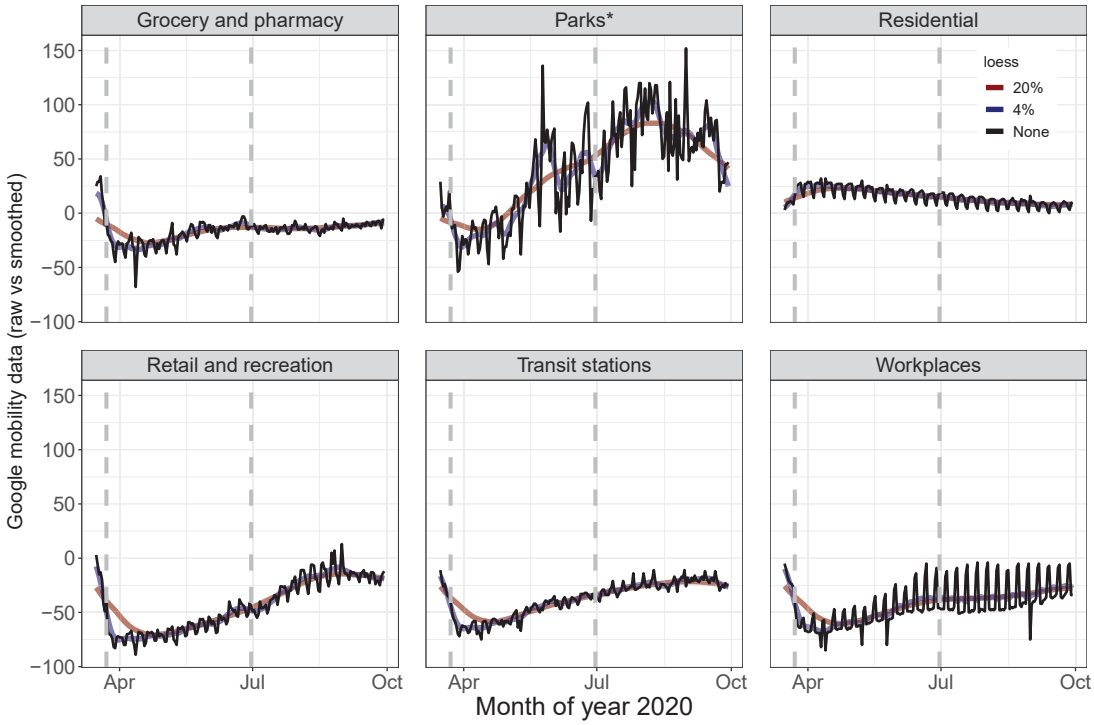


Figure 6.3 Time series data of movement trends for different location types in Leicestershire during spring and summer 2020, published by Google. Thin black line represents raw data. Thick red and blue lines represent local regression smoothing with parameters of 20% and 4% respectively. The dashed grey lines indicate dates of national and local lockdowns.
* Mobility data for Leicestershire as of 18th March 2021 had too much missingness; national data was instead used here

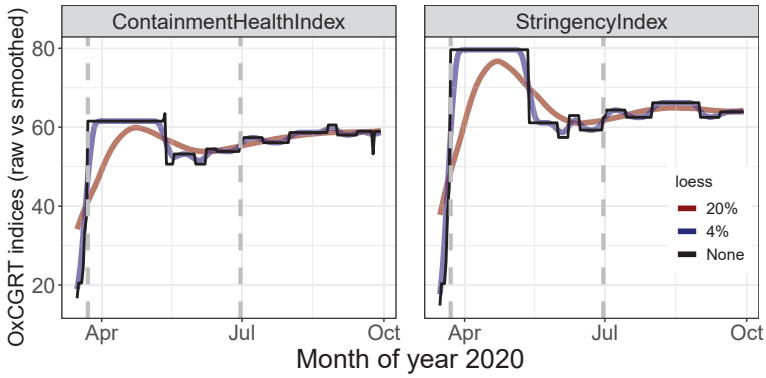


Figure 6.4 Time series composite data of national UK government interventions in the spring and summer months of 2020, estimated by Oxford COVID-19 Government Response Tracker (OxCGRT). Thin black lines represent raw data. Thick red and blue lines represent local regression smoothing with parameters of 20% and 4% respectively. The dashed grey lines indicate dates of national and local lockdowns.

Figures 6.5 and 6.6 show the smoothed covariate data for the winter period of 2020-2021, which is encoded in the analysis for case studies C and D.

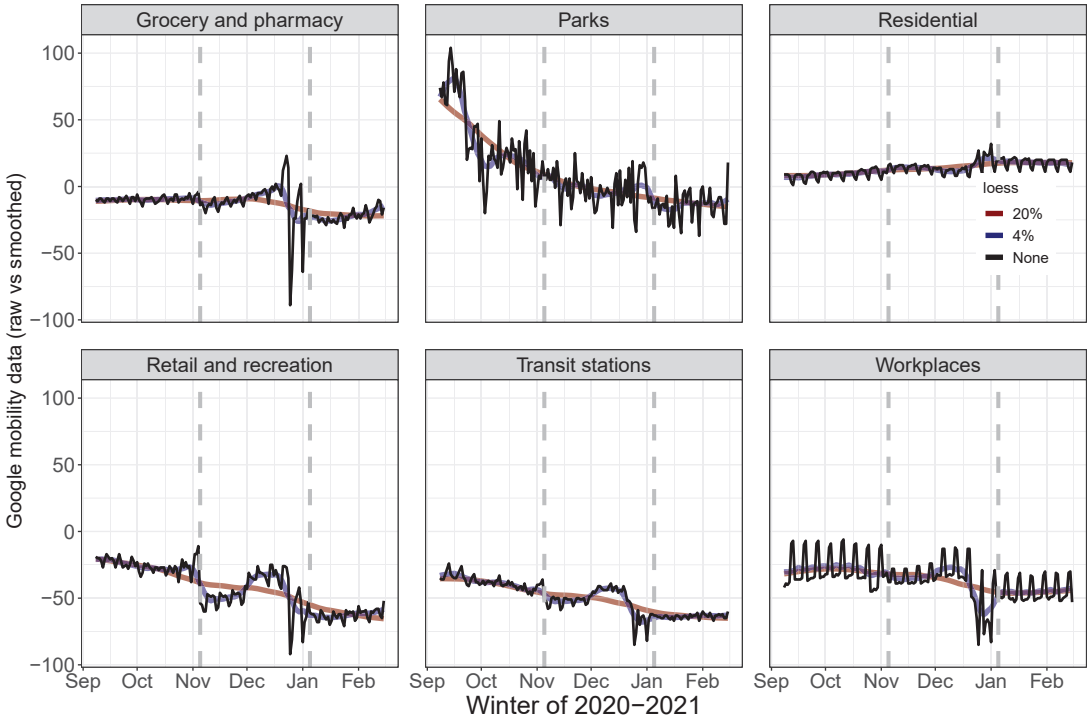


Figure 6.5 Time series data of movement trends for different location types in England during the winter period 2020-2021, published by Google. Thin black line represents raw data. Thick red and blue lines represent local regression smoothing with parameters of 20% and 4% respectively. The dashed grey lines indicate dates of national lockdowns.

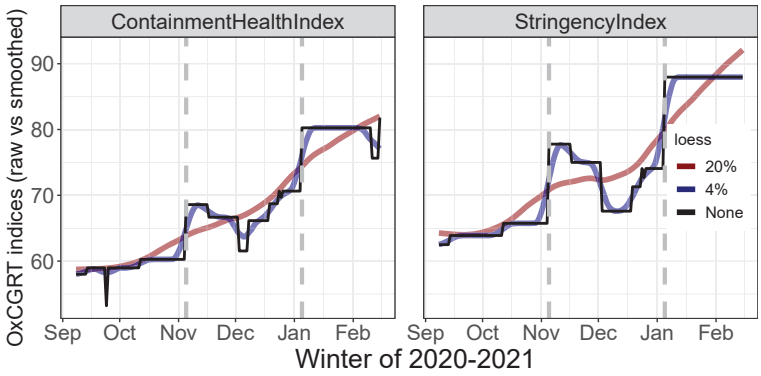


Figure 6.6 Time series composite data of UK government interventions in the winter of 2020-2021, estimated by Oxford COVID-19 Government Response Tracker (OxCGRT). Thin black lines represent raw data. Thick red and blue lines represent local regression smoothing with parameters of 20% and 4% respectively. The dashed grey lines indicate start dates of lockdowns.

Finally for all case studies, after smoothing the covariates, I rescaled them to take positive values within a range of 1-100 and log transformed them to reduce skew.

In summary, there are 3 covariate time series data sets to include in phylodynamic modelling analyses. First is the binary lockdown matrix (separate columns for different lockdowns). Secondly there are the continuous Google mobility time series data. Finally there are the continuous OxCGRT stringency and health containment indices.

6.3.2 S-gene target failure data

As discussed in Section 6.2.2.1, the B.1.1.7 lineage was linked to S-gene target failure (SGTF) in common SARS-CoV-2 diagnostic tests such as the TaqPath assay. However, other lineages (such as B.1.258) with the same functional deletion Δ69-70 have been observed to cause SGTF. Therefore, the reliability of using SGTF frequency as a proxy for B.1.1.7 infection burden varies over time, depending on its changing frequency relative to other Δ69-70 lineages.

Using a spatio-temporal model fitted to counts of genomes with Δ69-70 belonging to VOC and non-VOC lineages, Volz et al [101] estimated the probability that a sample with Δ69-70 belongs to the VOC lineage, as a function of time and place. This probability was termed "true positive proportion" (TPP). They thereby calculated the numbers of TPP-corrected SGTF negative cases attributable to VOC (S^-), and TPP-corrected SGTF positive cases attributable to non-VOC lineages (S^+).

To explore the effectiveness of phylodynamic analysis to estimate viral lineage growth, I compared estimated effective population size against S^- through time. However, the above S^+ and S^- counts reflect only those who were tested, and are underestimates of the true numbers. Therefore for each STP and for each epiweek, I calculated the 'true' number TPP-corrected SGTF negative cases (S_w^{true-}) as the proportion of S^- out the total number of SGTF cases, scaled by the number of cases reported.

The number of confirmed COVID-19 cases was available by epidemiological week ('epiweek') for all 42 STP regions in England. Let C_w represent confirmed cases per epiweek w. Let S_w^- and S_w^+ represent SGTF negative and positive cases per epiweek w respectively. I used the following formula to calculate S_w^{true-} for each STP for each epiweek:

$$S_w^{true-} = C_w \times S_w^- / (S_w^+ + S_w^-)$$

I summed across all STPs to produce an estimate of total S_w^{true-} in England for each epiweek. Where there was more than one data point for any week, I kept only the most recent value. The weekly estimates of S_w^{true-} are shown in Figure 6.22.

6.3.3 Sub-sampling the sequence alignment

The phylodynamic models *skygrowth* and *mlesky*, when implemented in *R* on several computer processing units, can both cope with phylogenetic trees with a few thousand tips to produce epidemic trajectories within a reasonable time frame. However, the COG-UK database for England contains a higher number of sequences by two orders of magnitude. Not all of these sequences can (or should) therefore be used, and appropriate sub-sampling is necessary. The four case studies I describe seek to answer different questions, and the sub-sampling strategies of each are therefore specific and detailed separately in this section.

Case study A: England during the early months of 2020

This aim of this case study was to evaluate the impact of the first national lockdown in England. I therefore sought sequences across the country sampled between 1st January and 30th April 2020. At the beginning of the epidemic in early 2020, the infrastructure for high throughput sequencing of SARS-CoV-2 in England did not exist. Coupled with the lower burden of COVID-19 early in the exponential growth phase, comparatively few sequences were therefore available in the months of January and February. A simple random

draw of sequences sampled between January and April would therefore retrieve little to no representation from the earlier months.

To obtain coverage stretching back into February (or late January), I ordered the sequence alignment by sample date, split it into 5,000 separate bins, and sampled one sequences from each. As a result, there was guaranteed representation from earlier sequences, but the vast majority of samples were from March and April. I then removed duplicate sequences. I replicated this process 10 times to produce 10 alignments with a mean of 4217 sequences each.

Note that during these early months pillar 2 (community sampling) was not available, and thus I sampled from pillar 1, where there were around 23,000 sequences until 30th April 2020.

Case study B: Leicester during the spring and summer of 2020

The aim of this case study was to evaluate the impact of lockdowns on a more local level. Leicester, like the rest of the country, went into lockdown on March 23rd 2020. However, Leicester was unique in being the first British city to enter a local lockdown, due to its disproportionate viral burden during the summer months [395]. Local lockdown was implemented from June 30th to August 1st in Leicester, whereby stay-at-home orders were mandated, and travel from and to the city was prohibited.

As discussed in Section 6.2.5, a motivation for focusing on the city level (as opposed to the national level) was to explore whether routine genomic surveillance is sufficient to provide enough sequences to detect upward trends in infections using phylodynamic modelling, possibly before they are revealed through testing, hospitalisations and deaths. As cases rose in Leicester through the month of June 2020, leading to a sharper focus in this region, a large number of samples were available in June. Such a wealth of data in Leicester during this period is useful, yet it is unusual. It is interesting to investigate whether the rise and fall in viral burden is still detectable with a more 'normal' level of genomic sampling; as such, I performed phylodynamic analyses on various fractions of the whole alignment in Leicester.

I filtered the database for sequences in Leicester (and neighbouring areas) sampled between 1st January and 29th September, in order to obtain a spread of samples before, after and between lockdowns. This resulted in 1,511 sequences between 18th March and 29th September 2020. I took a random sample without replacement of 5% of the Leicester alignment (76 sequences); I performed this 20 times. I repeated the process for sampling fractions of 10%,

15%, 20%, 30%, 40%, 50%, 70%. This generated 8 x 20 alignments on which to perform analyses, with each alignment ranging from 76 to 1,058 sequences.

It is important to reiterate that these fractions do not reflect the proportion of total cases, but rather the proportion of available samples. Only a fraction of cases are sequenced, even under periods of intensive sampling.

Case study C: England during the winter of 2020-2021

The aim of this case study was to examine the epidemic trajectory of the second and third waves of COVID-19 in England, during the winter period of 2020-2021. COVID-19 infections and deaths exceeded those of the first wave, despite having a far deeper understanding of the virus. This period was also notable for heterogeneity in government interventions, with both regional and national restrictions applying at different times.

I filtered the COG-UK database for sequences in England between 7th September 2020 to 31st January 2021, which covered the period before the start of second national lockdown, and after the start of the third. Note however that the third lockdown was still in place at the end of this sample period, and was still in place when this analysis was performed in March 2021.

Pillar 2 sampling, which aims to achieve a more random sample of the community, was in place during this later period of the epidemic. I therefore prioritised these samples and excluded pillar 1 samples.

I ordered the whole England alignment by sample date, split it into 3,000 bins and sampled one sequence from each bin, then deduplicated as before. I repeated this process 100 times to produce 100 alignments with a mean of 2,338 samples each. Note that these alignments contain sequences from any lineage, and is hereafter labelled as "All", so as not to be confused with alignments in case study D.

Case study D: Variant of concern B.1.1.7

In Section 6.4 I present an analysis of the epidemic during the winter months in England, separated by lineages B.1.1.7, B.1.177 and a "control". This work was completed within the Imperial College COVID-19 response team, in collaboration with others. Although I carried out the tree dating and phylodynamic analyses, the sub-sampling (and generation of the maximum-likelihood phylogenetic trees) by lineage were carried out by Dr Erik Volz and Dr Robert Johnson, with the details described in a published article [101].

Briefly, all COG-UK samples were assigned weights according to their local authority and collection date, to account for differences in sampling intensity over space and time. This weight was proportional to the ratio of diagnosed cases to the number of samples in each local authority in each week. Within the B.1.1.7 lineage, 100 alignments with 3,000 sequences were sampled between 15th October 2020 and 1st February 2021. The B.1.1.7 sequences included were then paired with samples from different lineages (B.1.177 and 'control'), matched by upper tier local authority and by sample week, mimicking a case-control study. B.1.177, as described in Section 6.2.2.1, was a lineage that was investigated closely and found to bear no transmission advantage. The "control" group was a sample of any lineage outside of B.1.1.7.

Case study C presents a set of unweighted samples from any lineage, whereby the B.1.1.7 alignments in case study D are weighted samples within the lineage, which are matched by week and local authority to produce the sets B.1.177 and control. Case study C therefore differs from case study D by being unweighted and indiscriminate with respect to lineage. However, both case studies C and D examine the same winter months between 2020 and 2021.

6.3.4 Producing dated phylogenetic trees

As part of the daily COG-UK pipeline, a maximum-likelihood (ML) tree is generated for the whole alignment of sequences from the UK. The documentation for this process is described in a GitHub repository (https://github.com/COG-UK/grapevine) and in this published paper [441]. Briefly, new UK sequences are filtered based on coverage and quality of reads, assigned a lineage using *pangolin* (github.com/cov-lineages/pangolin), then mapped to a reference sequence *Wuhan-Hu-1* using *Minimap* [442]. A maximum-likelihood phylogeny is constructed using *FastTree* version 2.1.11 [443], with *Wuhan/WH04/2020* used as an outgroup. I have not contributed to this pipeline.

Having selected the alignments as described in Section 6.3.3, I retained these tips from the ML tree produced in the COG-UK pipeline, using the *ape* R package [444]. By using the pre-generated tree, I avoided unnecessary and lengthy computational time in its reconstruction. In a sensitivity analysis, I reconstructed the ML tree from the sub-sampled sequences (i.e. not using the COG-UK tree) using *IQTREE* (parameters -nt 12 -redo -m HKY -s). I did this to explore whether avoiding the step of reconstructing the ML tree caused a difference to the results.

I dated the pruned ML trees containing the samples of interest using *treedater* [113], collapsing small branch lengths and randomly resolving polytomies. I sampled clock rate from a normal distribution of mean 5.91×10^{-4} and standard deviation 1.92×10^{-5}; this clock rate distribution was estimated in a Bayesian coalescent phylodynamic analysis performed using BEAST by Professor Andrew Rambaut, described here [101]. Each ML tree was dated 10 times (with a different sampled clock rate) to account for phylogenetic uncertainty. These dated trees were provided as data to which *skygrowth* and *mlesky* were fitted.

A summary of all time-scaled phylogenetic trees is presented in Table 6.1, detailing the number of sequences, sample dates and lineages for all case studies in this Chapter.

Table 6.1 Characteristics of the time-scale phylogenies used in each case study

Case study	Area	Lineages	Earliest sample	Latest sample	Sampling fraction	Samples per alignment	No. alignments	ML* trees per alignment	Time trees per ML tree	No. of time trees total
A	England	All	29/01/2020	30/04/2020	NA	5000	10	1	10	100
B	Leicester	All	18/03/2020	29/08/2020	5%	76	20	1	10	200
B	Leicester	All	18/03/2020	29/08/2020	10%	151	20	1	10	200
B	Leicester	All	18/03/2020	29/08/2020	15%	227	20	1	10	200
B	Leicester	All	18/03/2020	29/08/2020	20%	302	20	1	10	200
B	Leicester	All	18/03/2020	29/08/2020	30%	453	20	1	10	200
B	Leicester	All	18/03/2020	29/08/2020	40%	604	20	1	10	200
B	Leicester	All	18/03/2020	29/08/2020	50%	756	20	1	10	200
B	Leicester	All	18/03/2020	29/08/2020	70%	1058	20	1	10	200
C	England	All	07/09/2020	31/01/2021	NA	3000	100	1	10	1000
D	England	B.1.1.7	15/10/2020	01/02/2021	NA	3000	100	1	10	1000
D	England	B.1.177	11/10/2020	03/02/2021	NA	3000	100	1	10	1000
D	England	control	11/10/2020	03/02/2021	NA	3000	100	1	10	1000

*ML = maximum-likelihood

6.3.5 Phylodynamic modelling

6.3.5.1 Implementation

To estimate the association between SARS-CoV-2 transmission and NPIs, I carried out phylodynamic modelling analyses using *skygrowth* [97] and *mlesky* [99], both implemented in the R programming language.

Using both models, I estimated the epidemic trajectory of SARS-CoV-2 by fitting to the time-scaled phylogenies described in Section 6.3.4. I use the same sets of trees in both model analyses. I perform this for case studies A-C; for case study B, trajectories are estimated for every sampling fraction (see Table 6.1). For case study D, I estimate trajectories for B.1.1.7, B.1.177 and control lineages separately using *mlesky* only. In all case studies, covariate data are included in a *mlesky* analysis.

The default parameters used in the *skygrowth* model analyses are summarised as follows. The precision parameter (τ) had an exponentially distributed prior with scale set to 2.06×10^{-6}, corresponding to a 1% change in growth per week.

The number of grid points (m), i.e. the number of piecewise sections in which $N_e(t)$ is estimated, was set to twice the number of weeks in the period under examination. For instance, case study A simulated the period January 29th 2020 - April 30th 2020 (13 weeks) such that $m = 26$. The prior for R was defined as a lognormal distribution with mean $\log(0.7)$ and standard deviation 0.5. $R(t)$ was calculated according to the following equation [445]:

$$R(t) = r(t)/\gamma + 1$$

where $r(t)$ is the growth rate of effective population size ($N_e(t)$) estimated through time and $(1/\gamma)$ is a constant serial interval of 6.5 days, based on previous studies on transmission pairs [446–448]. This calculation is valid when infectious periods are exponentially distributed [449, 97]. I calculate the 95% highest posterior density (HPD) for $N_e(t)$ and $R(t)$ through time, across all simulations.

The default parameters used in the *mlesky* model analyses are summarised as follows. The precision parameter (τ) was optimised by cross-validation using the *.mlskygrid_oos* function (https://github.com/emvolz-phylodynamics/mlesky/blob/master/R/mlesky.R), bounded between 0.001 and 10. The number of grid points (m) was also set to twice the number of weeks in the period under examination.

The *mlesky* package outputs effective population size ($N_e(t)$) given a time-scaled phylogeny. Let $r(t)$ represent growth rate of effective population size. In the same fashion as *skygrowth*, $R(t)$ is calculated as:

$$R(t) = r(t)/\gamma + 1$$

where $(1/\gamma)$ = 6.5 days.

I calculated the 95% quantiles for $N_e(t)$ and $R(t)$ through time, across all simulations.

When including a covariate time series data set for model fitting in *mlesky*, the *formula_order* parameter is set to 1 (unless stated otherwise), which tests the association between growth rate and the first order rate of change of the covariate. As discussed in Section 6.3.1, growth rate is estimated as the first-order rate of change of $N_e(t)$ and can be related to $R(t)$. The only deviation from this is when including the presence or absence of lockdowns as a covariate, for which I assumed $formula\_order = 2$. This specification compares the association between the 0th order (undifferentiated) covariate with second order changes in $N_e(t)$, i.e. the change in growth rate through time. Although more difficult to interpret, I chose this because the binary matrix, if differentiated to the first order, would be far too sparse to test for associations. Section 6.3.5.2 provides more detailed discussion of the implications of including covariates in *mlesky*.

The steps to complete the time-scaling of the phylogeny and phylodynamic analysis are detailed in scripts in the R programming language, available at: https://github.com/mrc-ide/sarscov2-b.1.1.7/blob/main/phylo/d1_date_trees.R and https://github.com/mrc-ide/sarscov2-b.1.1.7/blob/main/phylo/d2_mlesky.R.

6.3.5.2 Testing association between covariates and viral transmission

The full methodological framework for the inference of $N_e(t)$ using *mlesky* is available at [99]. To explore associations between covariate data and viral transmission (represented by $R(t)$), I use two methods: '*post hoc*' and 'joint'. In this section I summarise the steps of the *mlesky* analysis in order to describe differences in the two methods.

Like in previous phylodynamic frameworks such as *skygrid* [94], a piece-wise function of (log-transformed) effective population size through time ($N_e(t)$) is estimated at different pre-defined gridpoints using a coalescent model fitted to a time-scaled phylogeny. In order to reduce noise and overfitting, the trajectory of $N_e(t)$ is described as a Markov Chain. Given a value of $N_e(t = 0)$ each piece of the chain experiences a 'roughness penalty', where the

second order differences of $log(N_e(t))$ is modelled as a normal distribution with expectation of zero, and a standard deviation dependent on a smoothing parameter τ which is selected by a k-fold cross-validation strategy. This Markov Chain structure ensures that large steps are penalised during the smoothing process.

Post hoc association between covariate and viral transmission

The first method, *post hoc*, was achieved by performing the above, providing a smoothed time series (or 'trajectory') of $N_e(t)$, from which growth rate (and therefore $R(t)$) were calculated. Then, assuming that this reflects the 'true' $R(t)$, I treated it as data against which I tested association with the covariates.

Although rolling correlations between these can be plotted [450], formal testing for significance is not trivial. Measures in a time series, with each element being dependent on the previous time point, violate the assumptions of independent samples for most tests of correlation.

To overcome this, I computed the incremental daily changes in the covariate, a discrete equivalent of the first-order differential, which I then tested for correlation against $R(t)$ (already a first order differential of $N_e(t)$), measuring the effect size and significance of the association. Although imperfect, it is reasonable to assume that each daily change is less dependent on the previous day than the undifferentiated time series.

I tested the same correlations after delaying the $R(t)$ time series forwards or backwards by ten days. If changes in NPI are more strongly correlated with $R(t)$ ten days later, but not ten days earlier, it adds some evidential strength to the direction of causality.

A drawback of this method is that although phylogenetic uncertainty has been explored by resampling sequences and trees, only the central estimate of $R(t)$ across trees is the 'data' against which the covariates are fitted, and thus uncertainty in $R(t)$ is not fully accounted for.

Joint estimation of association between covariate and viral transmission

The second method, 'joint', also involves estimating the maximum likelihood of $N_e(t)$. However, the covariates are now implicated in the smoothing of the trajectory. The steps are summarised here. Note that these steps were programmed by Dr Erik Volz.

An initial fit of $N_e(t)$ against the time-scaled phylogenies is estimated. Daily changes in the covariates, like before, are fitted to daily changes in $N_e(t)$, assuming a linear model. Using

notation described in *skygrid* [94], let k represent the number of observations for which $N_e(t)$ is estimated (pre-defined by the number of grid points), and let p represent the number of covariates, which have observations at the times of grid points. If there are no observations for any covariate at a certain grid point, they are inferred by linear interpolation. In the linear model, the vector of coefficients β_k has length p, and the model matrix $\mathbf{Z}_{kp}$ has dimensions $k \times p$. Let b_k represent the multiplication of these two:

$$b_k = \mathbf{Z}_{kp} \cdot \beta_p$$

The length of b_k is therefore k, the number of observations of $N_e(t)$.

For this analysis, when implementing the roughness penalty in the Markov Chain, the second order differences of $log(N_e(t))$ for each piece i is modelled as a normal distribution with an expectation of b_i. In a final joint optimisation step, the maximum likelihood of $N_e(t)$ and β_k are determined by evaluating the likelihood of the tree given $log(N_e(t))$, combined with the roughness penalty incurred by this trajectory assuming expectations given by b_i.

Each analysis considered a Google mobility type or OxCGRT index as individual univariate predictors; I did not perform any multivariate analyses with these data sets. In other words, β was a scalar with units dR/dx where x is the covariate. However, I did perform a multivariate analysis using 'lockdown2' and 'lockdown3' as joint covariate predictors (see Section 6.3.1). In this case, β had units d^2R/x.

As I performed each *mlesky* analysis across a set of trees, the phylogenetic uncertainty was represented by a distribution of maximum likelihood estimates of β. If 95% of estimated values of beta were greater or less than 0, I concluded there was a positive or negative association respectively between the covariate and viral transmission.

I also explored the impact on β when implementing a delay in the covariate data before performing the *mlesky* analysis. For example, I tested the association between growth rate at time t and the covariate at $t-10$, or at $t+10$. If associations were stronger when the changes in covariate preceded changes in N_e (and not the other way round), then the direction of causality between changes in covariates and transmission could be more strongly stated even if not conclusively proved.

6.4 Results

6.4.1 Case study A: England during the early months of 2020

Figure 6.7 displays the epidemic trajectories of the SARS-CoV-2 epidemic in England during the early months of 2020. There is near-identical accord between both models for the central estimates of $N_e(t)$ and $R(t)$. The measure of uncertainty for *skygrowth* exceeds that of *mlesky*. Epidemic peak is estimated to occur on 1st April for *mlesky*, eight days after the imposition of the first national lockdown, shown in the red box. For *skygrowth*, the peak is broader and more jagged, but also centres around 1st April. These peaks precede the peak in confirmed cases by PHE by several weeks, although widespread testing was not available during this period, and case ascertainment was therefore very poor until April which will downwards bias the case count during February and March (dashed line in Figure 6.7A).

Reproduction number ($R(t)$) also falls below 1 at the beginning of April, although it appears that it was steadily decreasing in the several weeks prior to the start of lockdown. This decline could be due to reduced mobility and behaviour changes due to the pandemic even before lockdown. Alternatively, it could be because the proportion of cases that were due to importation decreased through time, as travel restrictions were imposed and internal transmission increased [451].

For both *skygrowth* and *mlesky*, the first few weeks of the epidemic provided extremely wide uncertainty for reproduction number. I estimated that in late January, $R(t)$ was 1.4 (95% HPD: 0.4-35.0) for *mlesky* and 2.1 (1.1-4.4) for *skygrowth*. In a stochastic transmission model fitted to surveillance data and SARS-CoV-2 prevalence, Knock and colleagues estimated R_0 in England at 2.9 (95% CrI: 2.8-3.1) at the beginning of March [69]. This estimate is higher than the median ones in my analyses, but within the wide uncertainty intervals. The paucity of good quality genomic data at the time in England is likely to lead to unreliable $R(t)$ estimates in my phylodynamic analysis in the early months, reflected in the extremely wide uncertainty intervals.

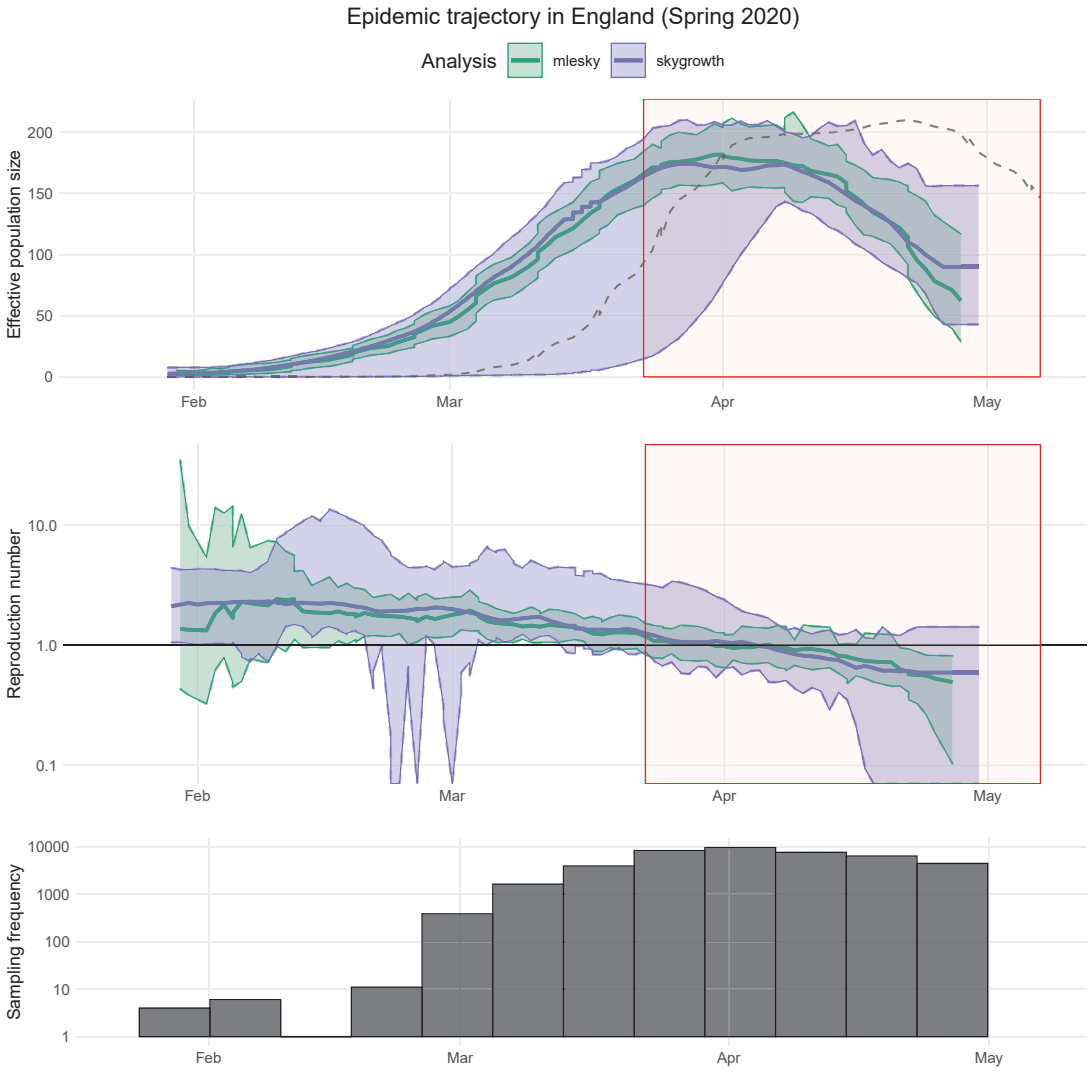


Figure 6.7 The epidemiological trajectory of SARS-CoV-2 in England during spring 2020. Thick solid lines and shaded areas represent median and 95% quantiles respectively, estimated by *mlesky* (green) and *skygrowth* (purple). Red shaded area represents period of first national lockdown in England. Black dotted line represents daily confirmed cases (smoothed and rescaled). Panel A: effective population size ($N_e(t)$) through time. Panel B: reproduction number ($R(t)$) through time. Panel C: frequencies of tips in each sample week in the phylogenies used.

Appendix Figure A.11 shows the daily incremental changes in Google mobility and OxCGRT index data through the spring months of 2020. Superimposed on each plot are the daily incremental changes in median estimate of $R(t)$ using *mlesky*.

Figure 6.8 shows scatterplots of the daily incremental changes in covariates against incremental changes in median $R(t)$ estimated by *mlesky*. This is the '*post hoc*' analysis described in 6.3.5.2. None of the correlations are significant to 5%.

The same correlation was tested after shifting $R(t)$ forwards or backwards by 10 days; still no significant correlations were found. These plots are displayed in the Appendix (Figures A.12 and A.13).

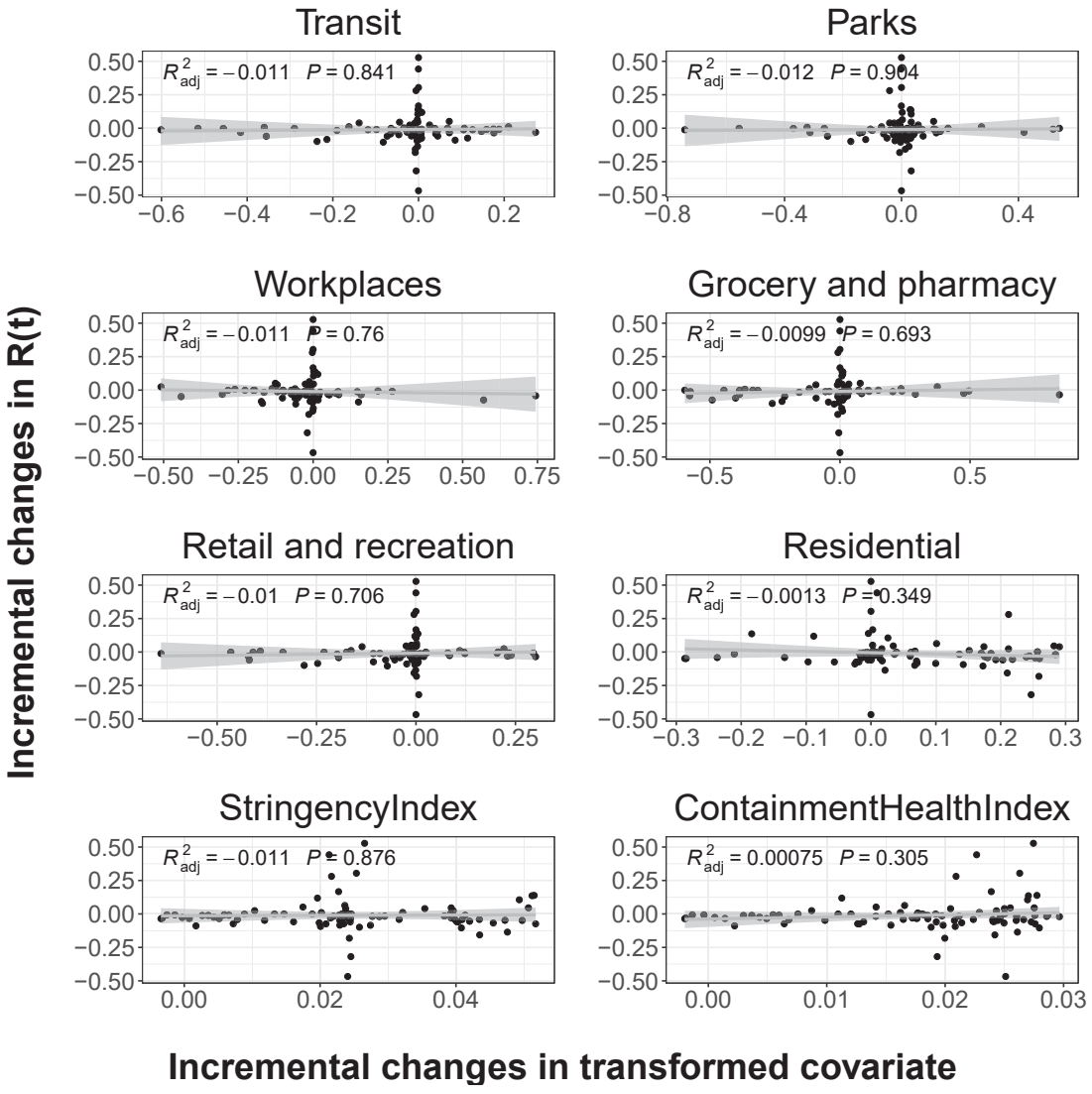


Figure 6.8 Scatterplot of the daily incremental changes in covariates against incremental changes in median $R(t)$ estimated by *mlesky* for England. Panels A-F display Google mobility data; panels G&H display OxCGRT index data. Grey line shows least squares linear regression; coefficients are displayed top left of each panel. No correlation was found to be significant to 5%.

When including the covariate within the phylodynamic analysis ('joint'), several significant associations appear. Figure 6.9 shows that positive daily changes in most Google mobility measures are associated with negative changes in $R(t)$, demonstrated by the negative distribution of coefficients. The exception is residential movement, which has positively association with transmission, although the 95% quantiles overlap with 0.

Government interventions, measured by both containment health and stringency indices, are both negatively associated with $R(t)$ (Figure 6.10).

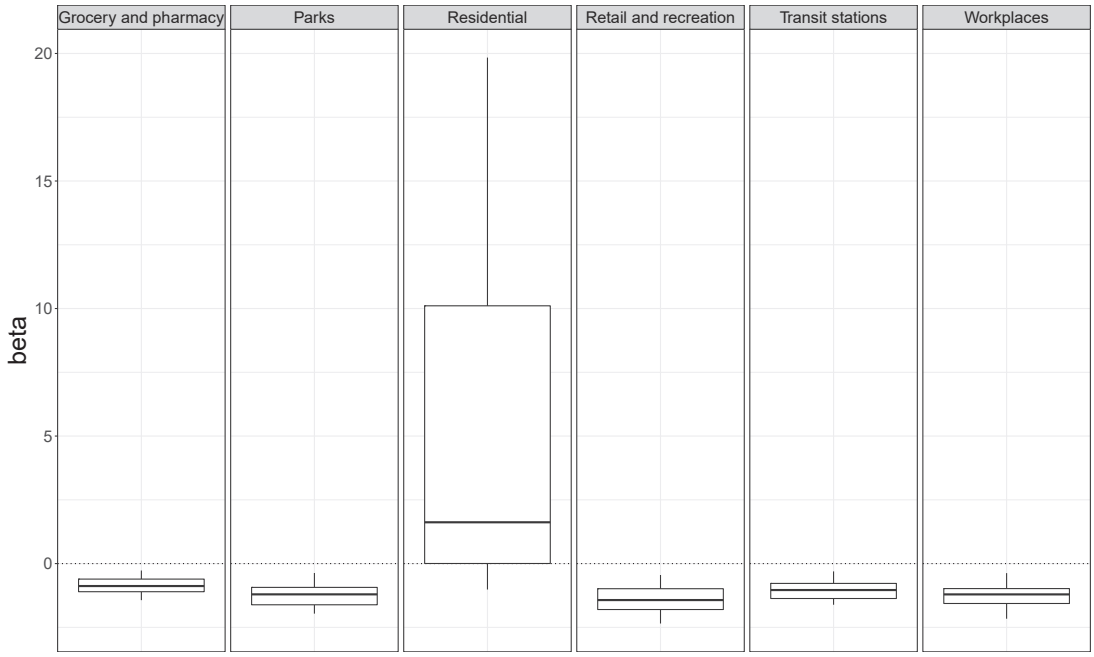


Figure 6.9 Distribution of beta coefficients when testing for univariate association between Google mobility and changes in effective population size estimated in *mlesky*. Boxes represent median and interquartile range; whiskers denote 95% quantiles.

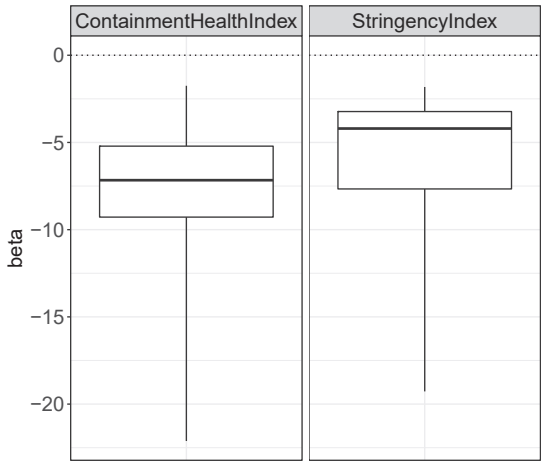


Figure 6.10 Distribution of beta coefficients when testing for univariate association between daily changes in OxCGRT indices and daily changes in effective population size (i.e. growth rate) in England estimated in *mlesky*. Boxes represent median and interquartile range; whiskers denote 95% quantiles.

Figure 6.11 shows the effect of introducing delays in the covariate time series for the OxCGRT Containment and Health index. The effect size of the coefficient has the largest negative value when introducing a time delay of 8 days; that is the association between the index at time t and R at time $t+8$. When the delay is the other way round (i.e. changes in R precede those in NPI by 8 days), the association becomes non-significant.

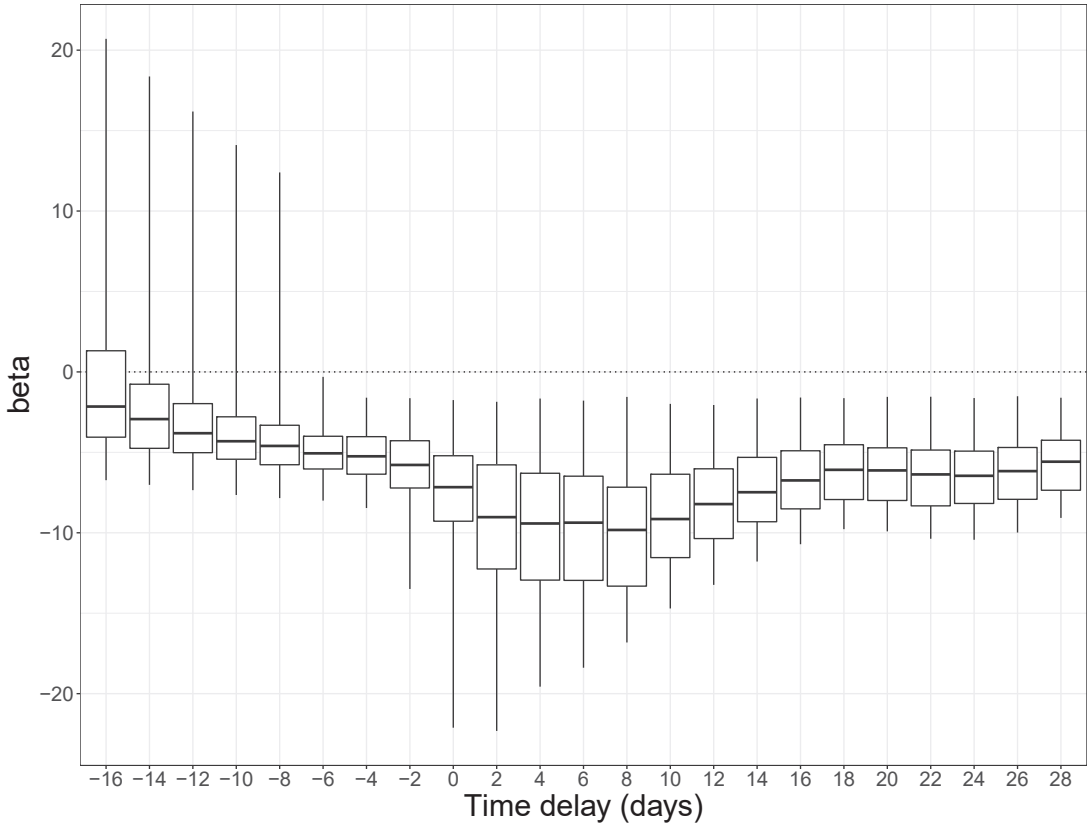


Figure 6.11 Distribution of beta coefficients when testing for univariate association between daily changes in OxCGRT health containment index and daily changes in effective population size (i.e. growth rate) in England estimated in *mlesky*. Boxes represent median and interquartile range; whiskers denote 95% quantiles. A positive delay of 10 represents testing an association between the covariate at time t and R at time $t + 10$ days.

Including covariates with the 'joint' estimation method made little difference to the median estimated epidemic trajectories, but greatly reduced uncertainty in $R(t)$ through time (Appendix Figure A.14).

6.4.2 Case study B: Leicester during the spring and summer of 2020

Figure 6.12 shows the results of phylodynamic analysis with *skygrowth* and *mlesky* without covariates in Leicester, when sampling 70% of sequences during the spring and summer months. There is good concordance in central estimates between the two analyses; *mlesky* has greater uncertainty, particularly in the more recent past. The *skygrowth* and *mlesky* analyses estimate the peak of infections to have occurred on 11th June and 12th June respectively. They both also estimate that $R(t)$ surpasses and falls below 1 repeatedly throughout the period. Note that this plot is cropped in mid August, although simulations continued until the end of August; this is because uncertainty in estimates grew two orders of magnitude in late August and displaying it compressed the axes and obscured the mid-year dynamics.

Both *skygrowth* and *mlesky* estimated a peak in cases in late April which was followed two weeks later by a peak in confirmed cases in mid May (PHE). Both methods later estimated a peak in early June, which was followed around a week later by a peak in confirmed cases. This latter peak is what led to the local lockdown in Leicester.

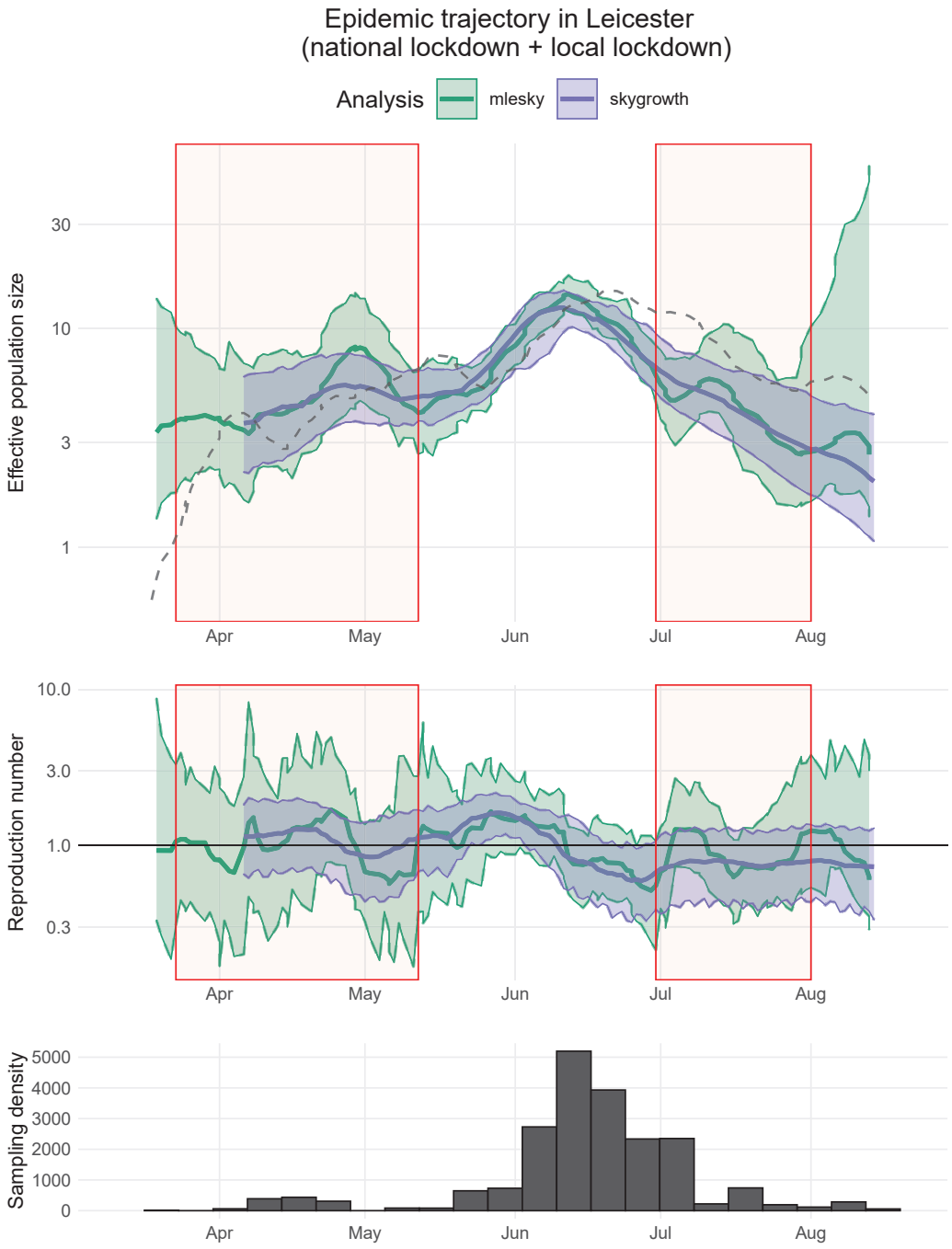


Figure 6.12 The epidemiological trajectory of SARS-CoV-2 in Leicester during the spring and summer months of 2020, using 70% of available sequences. Thick solid lines and shaded areas represent median and 95% quantiles respectively, estimated by *skygrowth* (purple) and *mlesky* (green). The black dotted lines show confirmed cases in Leicester according to PHE (smoothed and rescaled). Red shaded areas represent the first national lockdown in England and local lockdown in Leicester. Panel A: effective population size ($N_e(t)$) through time. Panel B: reproduction number ($R(t)$) through time. Panel C: frequencies of tips in each sample week in the phylogenies used.

Figure 6.13 shows the results of phylodynamic analysis with *skygrowth* without covariates in Leicester, using different fractions of sampling. Most sequences were sampled in June, due to the increased local testing following an outbreak in the city. A peak in viral population size is clearly detected for any sampling fraction above 5%, however increased sampling fractions were strongly associated with higher estimates of N_e and higher variance in reproduction number.

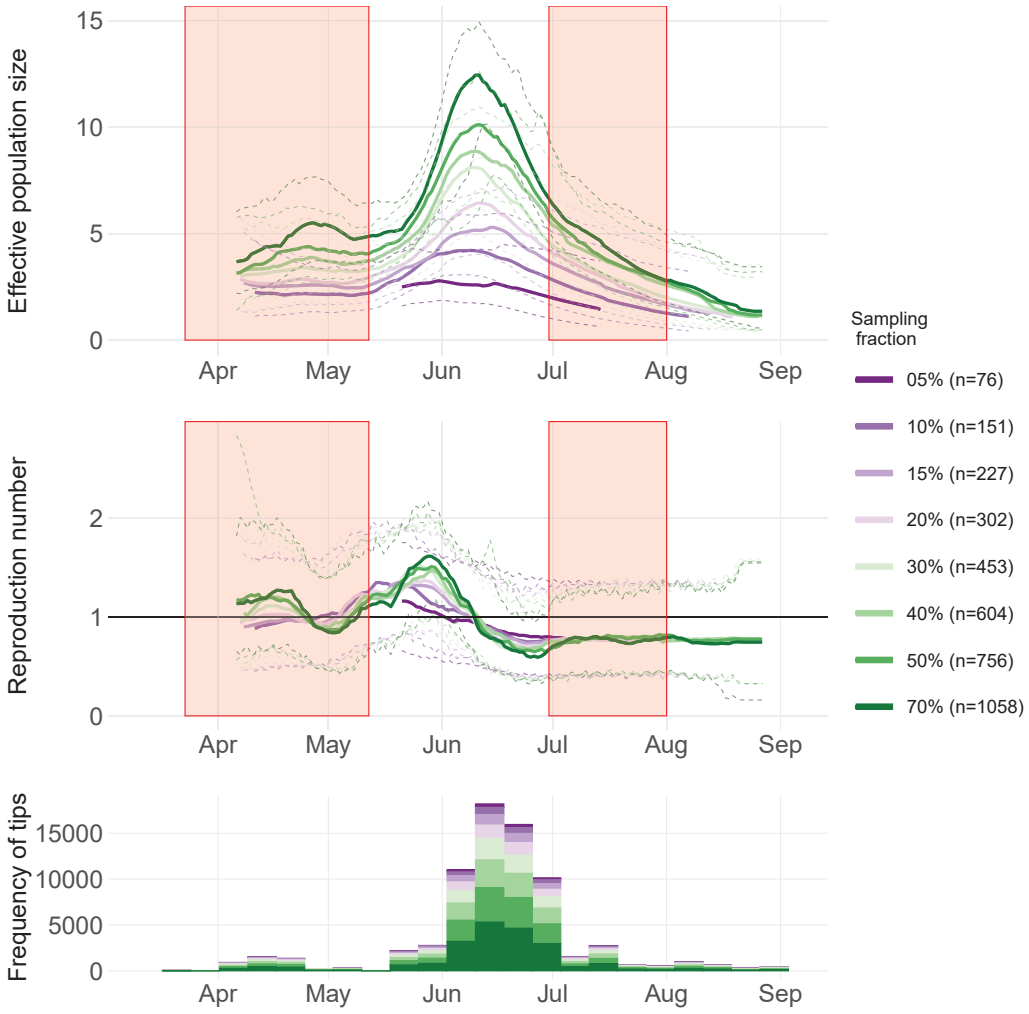


Figure 6.13 The epidemiological trajectory of SARS-CoV-2 in Leicester during the spring and summer months of 2020. Thick solid lines and thin dashed lines represent median and 95% quantiles respectively, estimated by *skygrowth*. Colour represents the fraction of samples analysed, ranging from 5% at the lowest (purple) to 70% at the highest (green). Legend displays sampling fraction and number of tips per alignment and tree. Red shaded areas represent the first national lockdown in England and local lockdown in Leicester. Panel A: effective population size ($N_e(t)$) through time. Panel B: reproduction number ($R(t)$) through time. Panel C: frequencies of tips in each sample week in the phylogenies used.

Figure 6.14 shows the equivalent plot for *mlesky*, but cropped to the beginning of July, and only for sampling fractions of 30% and above. The full plot is available in the Appendix (Figure A.15); the very large uncertainty intervals for the recent past and smaller sampling fractions obscure the patterns in viral population size dynamics. Similar to *skygrowth*, increased sampling fractions were associated with higher estimates of N_e. The peak in viral population size is observable at all sampling fractions (Figure A.16).

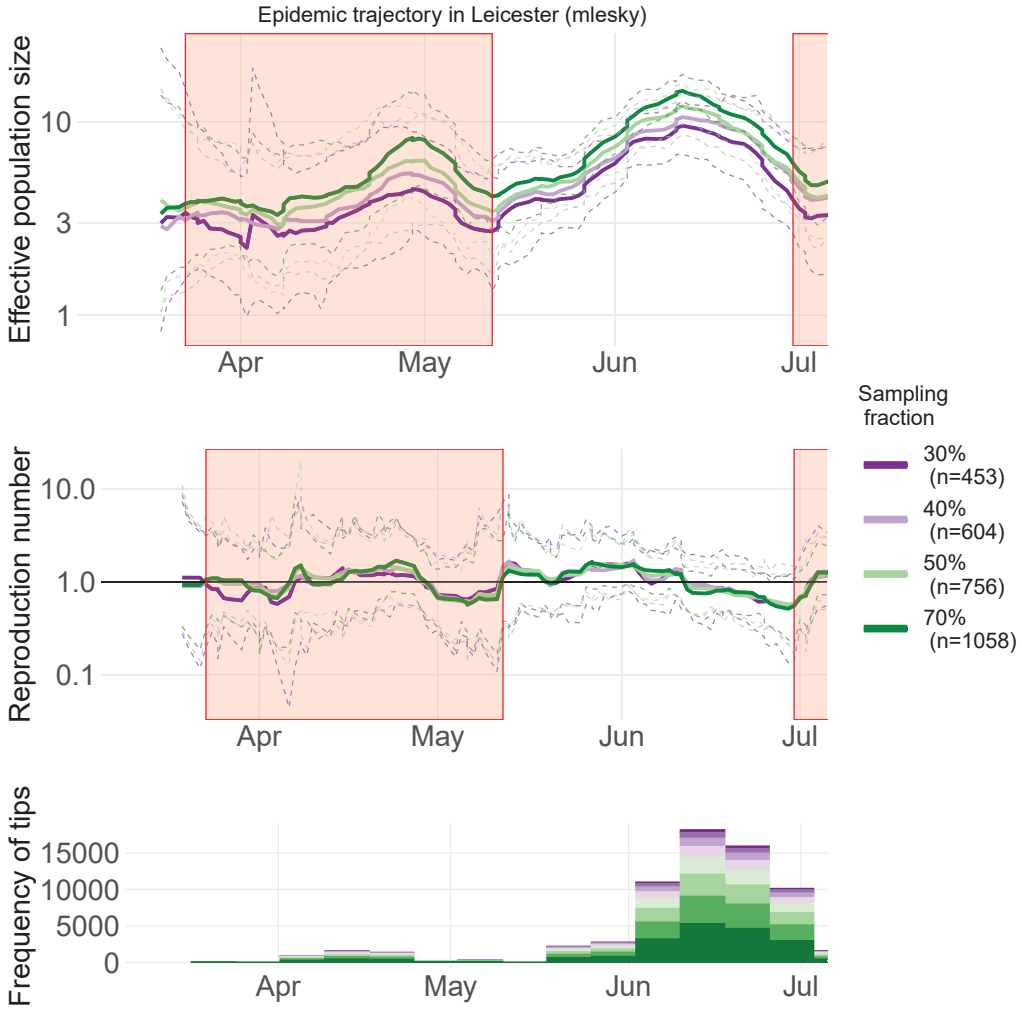


Figure 6.14 The epidemiological trajectory of SARS-CoV-2 in Leicester during the spring and summer months of 2020, cropped to the beginning of July, and only for sampling fractions of 30% and above. Thick solid lines and thin dashed lines represent median and 95% quantiles respectively, estimated by *mlesky*. Colour represents the fraction of samples analysed, ranging from 30% at the lowest (purple) to 70% at the highest (green). Legend displays sampling fraction and number of tips per alignment and tree. Red shaded areas represent the first national lockdown in England and local lockdown in Leicester. Panel A: effective population size ($N_e(t)$) through time. Panel B: reproduction number ($R(t)$) through time. Panel C: frequencies of tips in each sample week in the phylogenies used.

Figure 6.15 shows the tests for correlation between daily incremental changes in covariates and in $R(t)$. There are no significant correlations to 5%.

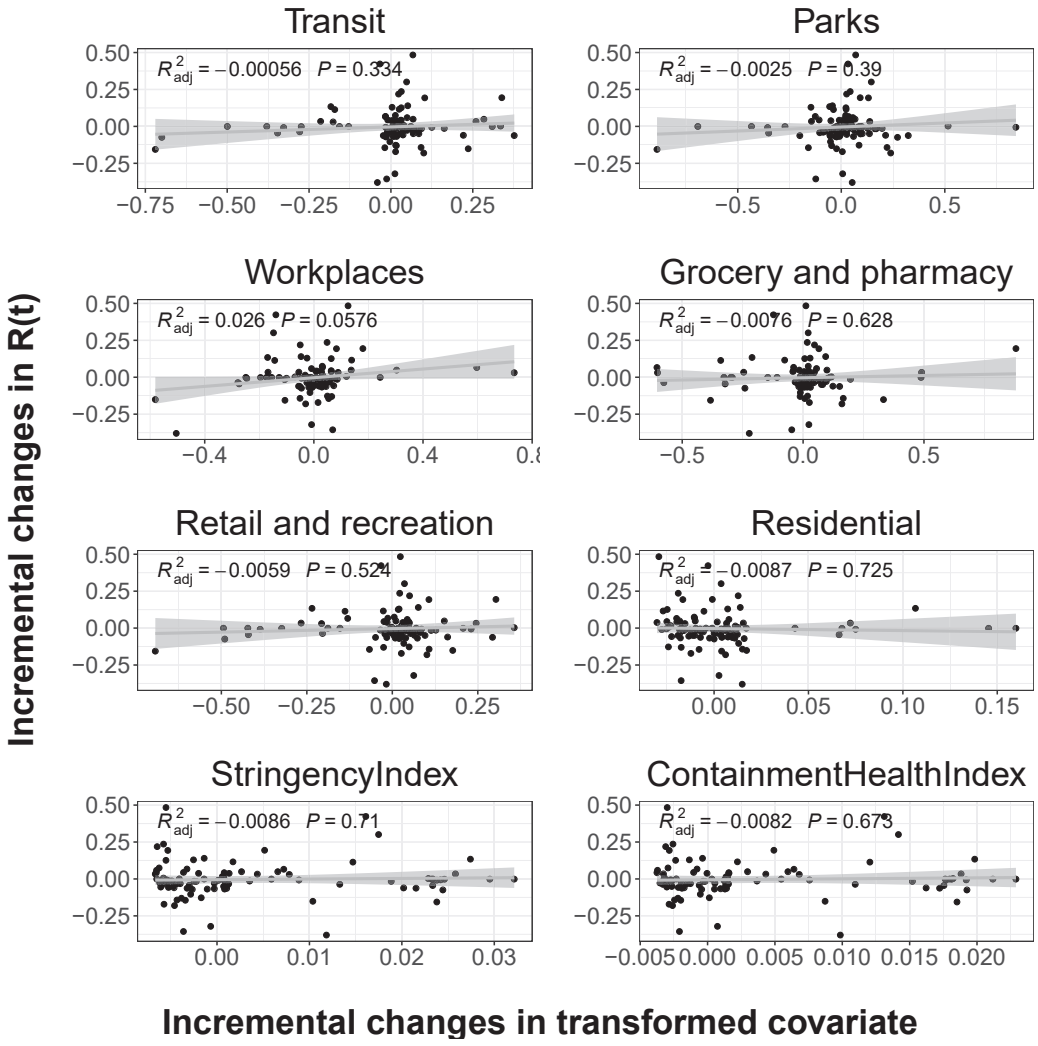


Figure 6.15 Scatterplot of the daily incremental changes in covariates in Leicester against incremental changes in median $R(t)$ estimated by *mlesky* at 70% sampling. Panels A-F display Google mobility data; panels G&H display OxCGRT index data. Grey line shows least squares linear regression; coefficients are displayed top left of each panel. No correlation was found to be significant to 5%.

Figures 6.16 and 6.17 show no significant associations when including the covariate within the phylodynamic analysis ('joint') with *mlesky*, for either Google mobility or OxCGRT indices.

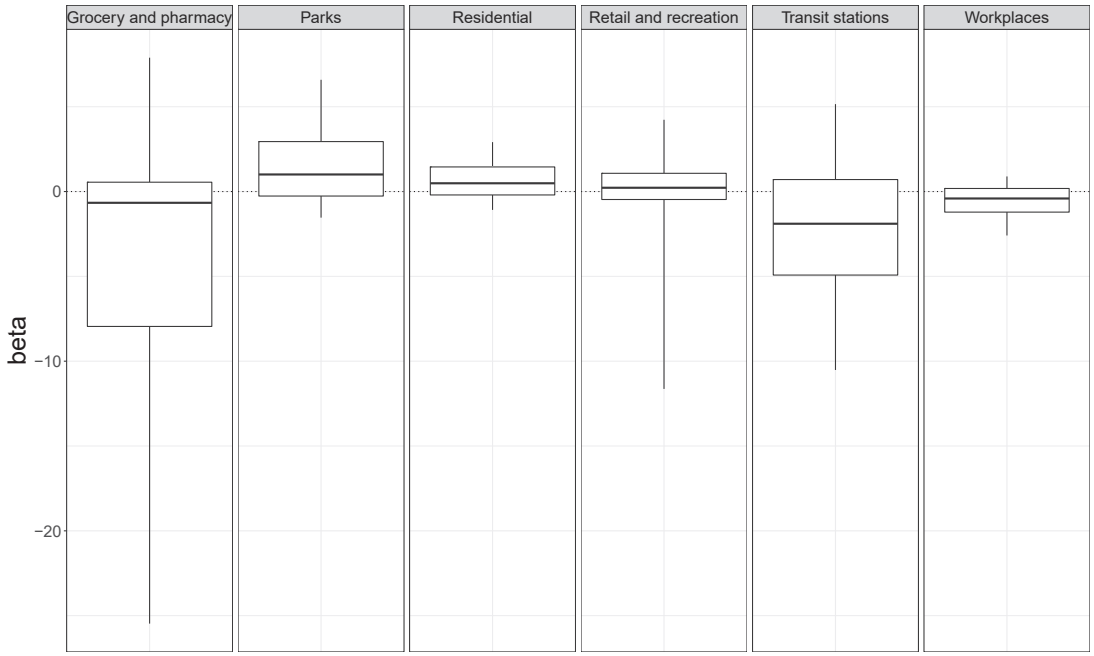


Figure 6.16 Distribution of beta coefficients when testing for univariate association between Google mobility and growth rate in Leicester (70% sampling), estimated with *mlesky*. Boxes represent median and interquartile range; whiskers denote 95% quantiles.

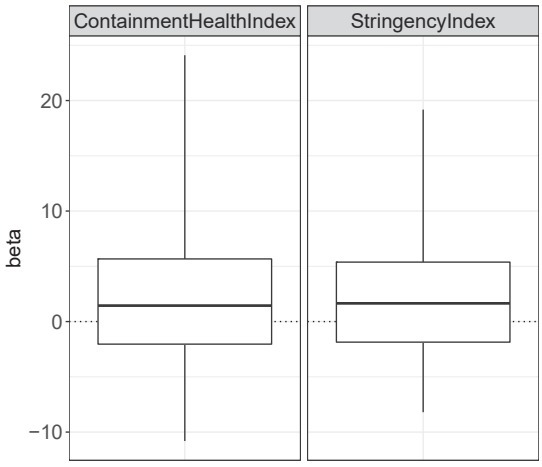


Figure 6.17 Distribution of beta coefficients when testing for univariate association between daily changes in OxCGRT indices and daily changes in effective population size (i.e. growth rate) in Leicester (70% sampling), estimated with *mlesky*. Boxes represent median and interquartile range; whiskers denote 95% quantiles.

In summary, for case study B (Leicester), no associations were found with either method between movement or NPIs against viral transmission. Including covariates with the 'joint' estimation method slightly changed the estimated epidemic trajectory but made no qualitative difference.

The lack of significance in this study is perhaps not unexpected. As there were very few sequences in Leicester in March 2020, this analysis mostly covered the months of April to June 2020 inclusive. This analysis therefore begins with strong restrictions already in place. Testing associations between changing stringency and viral transmission is therefore expected to generate a less significant result compared to case study A, which spanned a period with great variance in stringency (late January to late April).

6.4.3 Case study C: England during the winter of 2020-2021

The estimated epidemic trajectories are shown in Figure 6.18. Both models have strong agreement, estimating a small peak on October 23rd 2020, and a larger peak on 27th December 2020 (*skygrowth*) and 30th December 2020 (*mlesky*). Daily confirmed cases peaked on 7th November and January 4th, after smoothing data from PHE. The peaks in N_e therefore precede confirmed cases by 15 days for the first peak, and 5 - 8 days for the second peak.

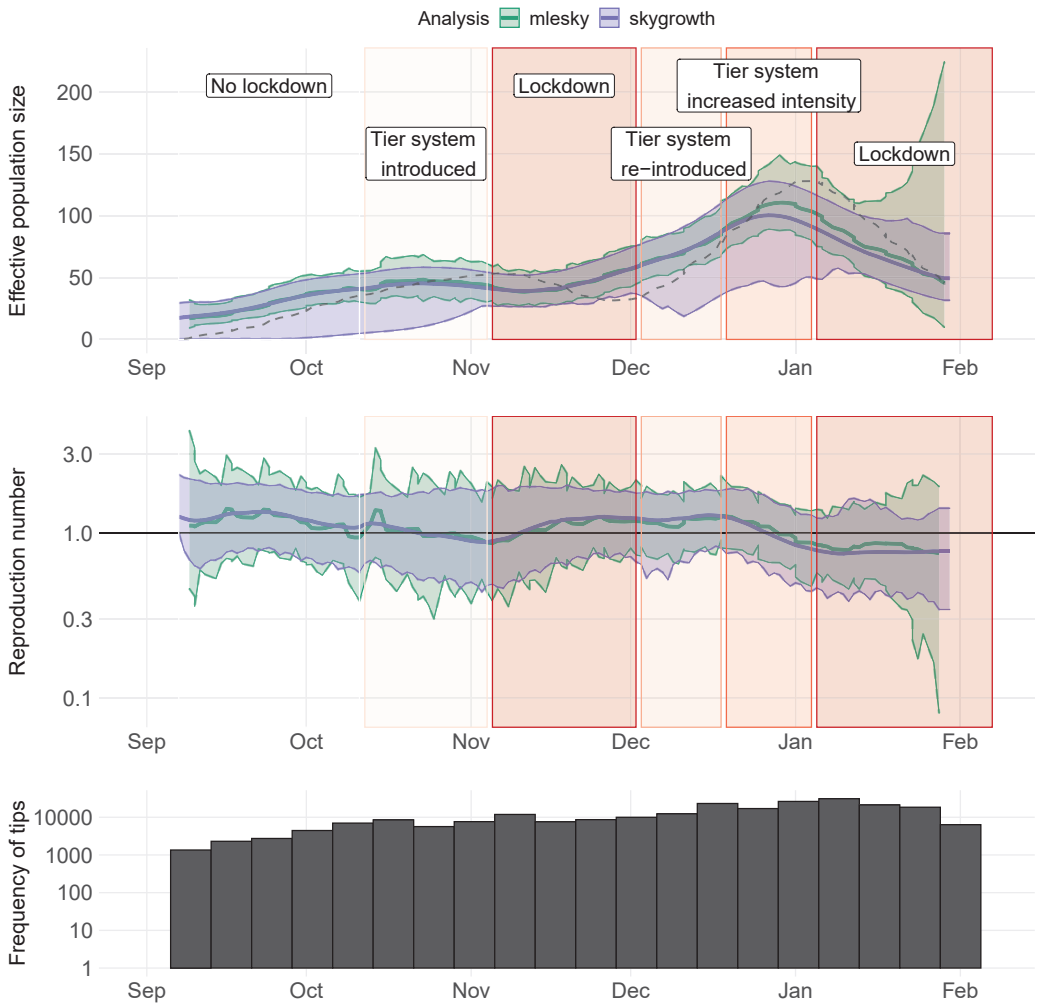


Figure 6.18 The epidemiological trajectory of SARS-CoV-2 in England during the 2020-2021 winter period. Thick solid lines and shaded areas represent median and 95% quantiles respectively, estimated by *mlesky* (green) and *skygrowth* (purple). The black dashed line represents daily confirmed cases from PHE (smoothed and rescaled). Red shaded areas represent periods of lockdowns and tier systems; a darker box reflects more strict restrictions. Panel A: effective population size ($N_e(t)$) through time. Panel B: reproduction number ($R(t)$) through time. Panel C: frequencies of tips in each sample week in the phylogenies used.

Figure 6.19 shows the scatterplot of daily incremental changes in covariates and $R(t)$ estimated by *mlesky*. There was a significant negative correlation for the residential mobility covariate, suggesting an association between changes in residential mobility and changes in $R(t)$.

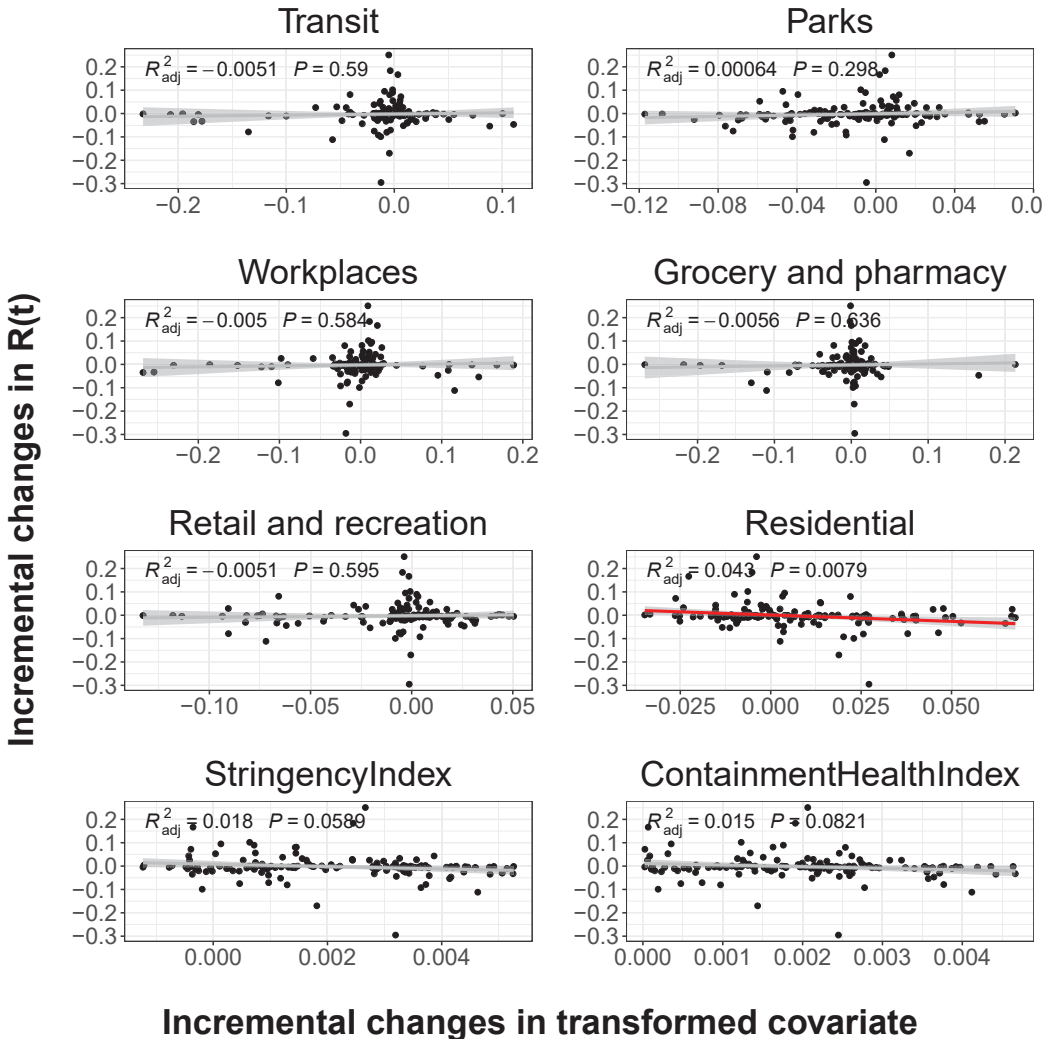


Figure 6.19 Scatterplot of the daily incremental changes in covariates against incremental changes in median $R(t)$ estimated by *mlesky* for England during the winter period. Panels A-F display Google mobility data; panels G&H display OxCGRT index data. The lines show least squares linear regression; coefficients are displayed top left of each panel. The red line shows a significant correlation for residential mobility.

Figure 6.20 shows the association between the presence of lockdowns and second order changes in N_e (i.e. rate of change of growth rate). Although the 95% confidence interval for both lockdowns overlapped with 0 suggesting a non-significant effect, the central estimate was more negative for lockdown 3.

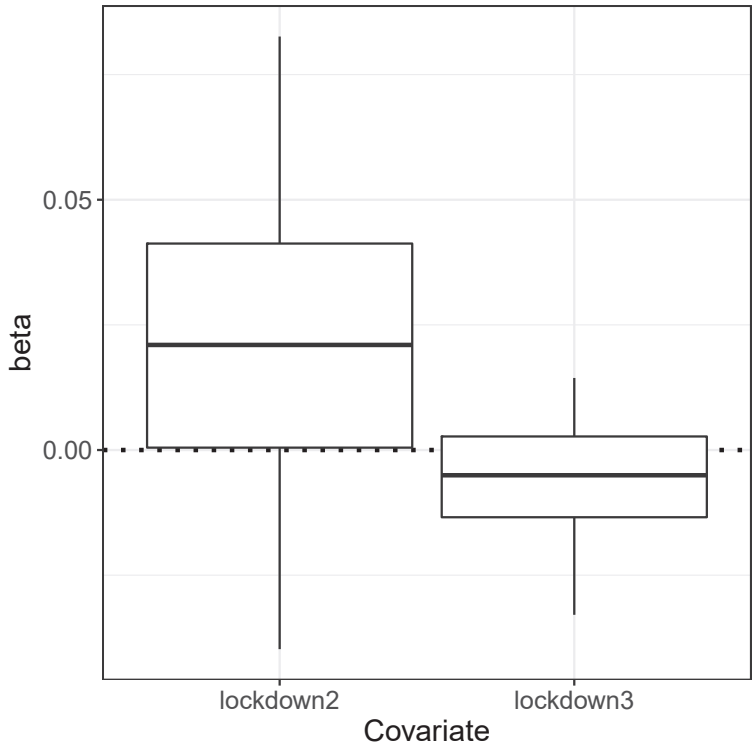


Figure 6.20 Distribution of beta coefficients when testing for multivariate association between presence of lockdowns 2 and 3 against second order changes in effective population size estimated in *mlesky*. Boxes represent median and interquartile range; whiskers denote 95% quantiles.

Including covariates with the 'joint' estimation method slightly changed the estimated epidemic trajectory but made no qualitative difference.

6.4.4 Case study D: Variant of concern B.1.1.7

Figure 6.21 shows the trajectories of lineage B.1.1.7 compared to samples from B.1.177 and other lineages (control), matched by location and week. Frequency of B.1.1.7 grows rapidly from the end of October 2020 to become the dominant strain by the end of the year. The trajectories of reproduction number for each lineage are very similar, but is consistently higher for B.1.1.7 throughout the period studied: between 5th November 2020 and 23rd January 2021, $R(t)$ for B.1.1.7 was 39.8% (daily median, 95% CI: 5.4-70.3) higher than that of the control group. Reproduction number for B.1.1.7 only dips below 1 during the third lockdown.

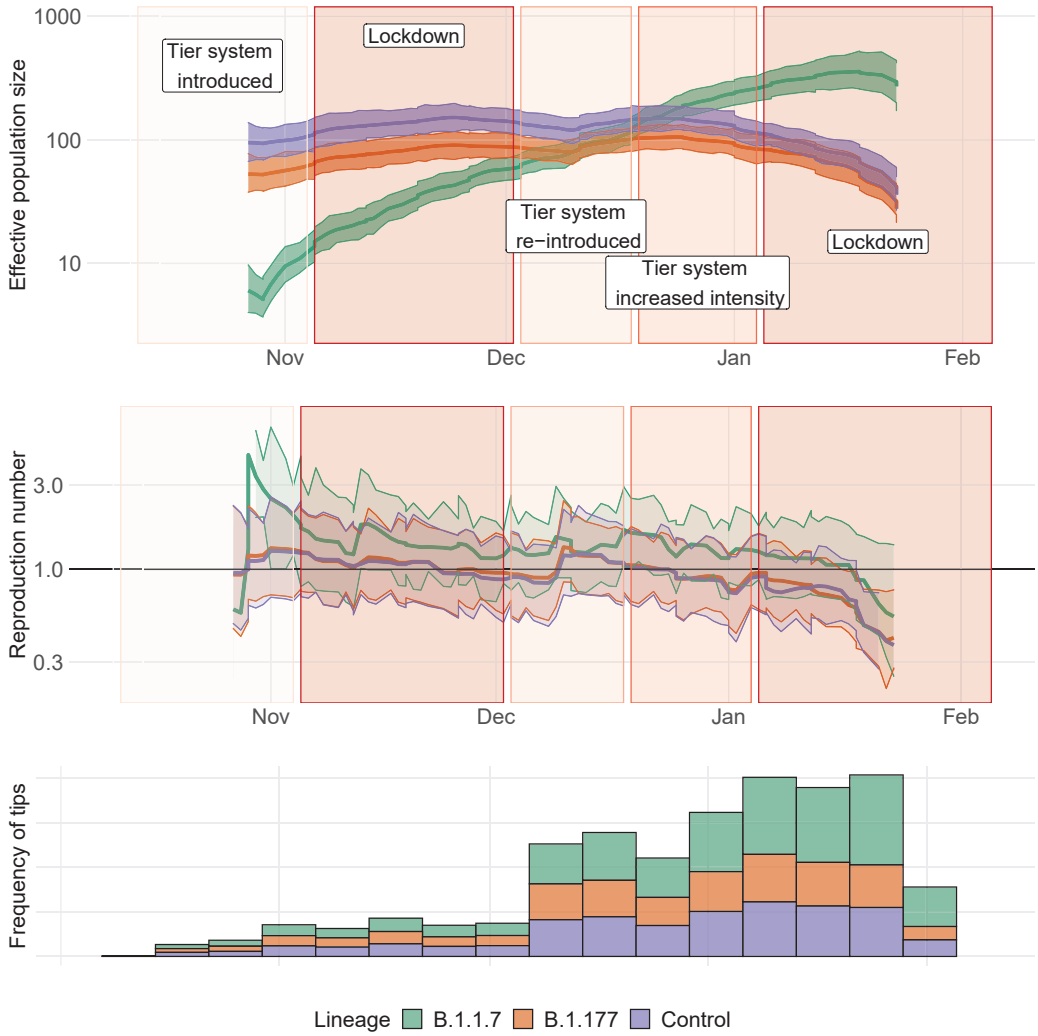


Figure 6.21 The epidemiological trajectories of 2 lineages (B.1.1.7 and B.1.177) of SARS-CoV-2, as well as a control group including samples from any lineage outside of B.1.1.7, in England during the 2020-2021 winter period. Thick solid lines and shaded areas represent median and 95% quantiles respectively, estimated by *mlesky*. B.1.1.7, B.1.177 and control group are shown in green, orange and purple respectively. Red shaded areas represent periods of lockdowns and tier systems; a darker box reflects more strict restrictions. Panel A: effective population size ($N_e(t)$) through time. Panel B: reproduction number ($R(t)$) through time. Panel C: frequencies of tips in each sample week in the phylogenies used.

Figure 6.22 shows that effective population size of the B.1.1.7 lineage, as estimated from genomic data with *mlesky*, is associated with independent estimates of SGTF, producing a robust cross-validation of the phylodynamic results. Figures A.18 and A.19 in the Appendix show the equivalent plots with B.1.177 and control lineages respectively; the effective population size of neither of these two show any association with SGTF.

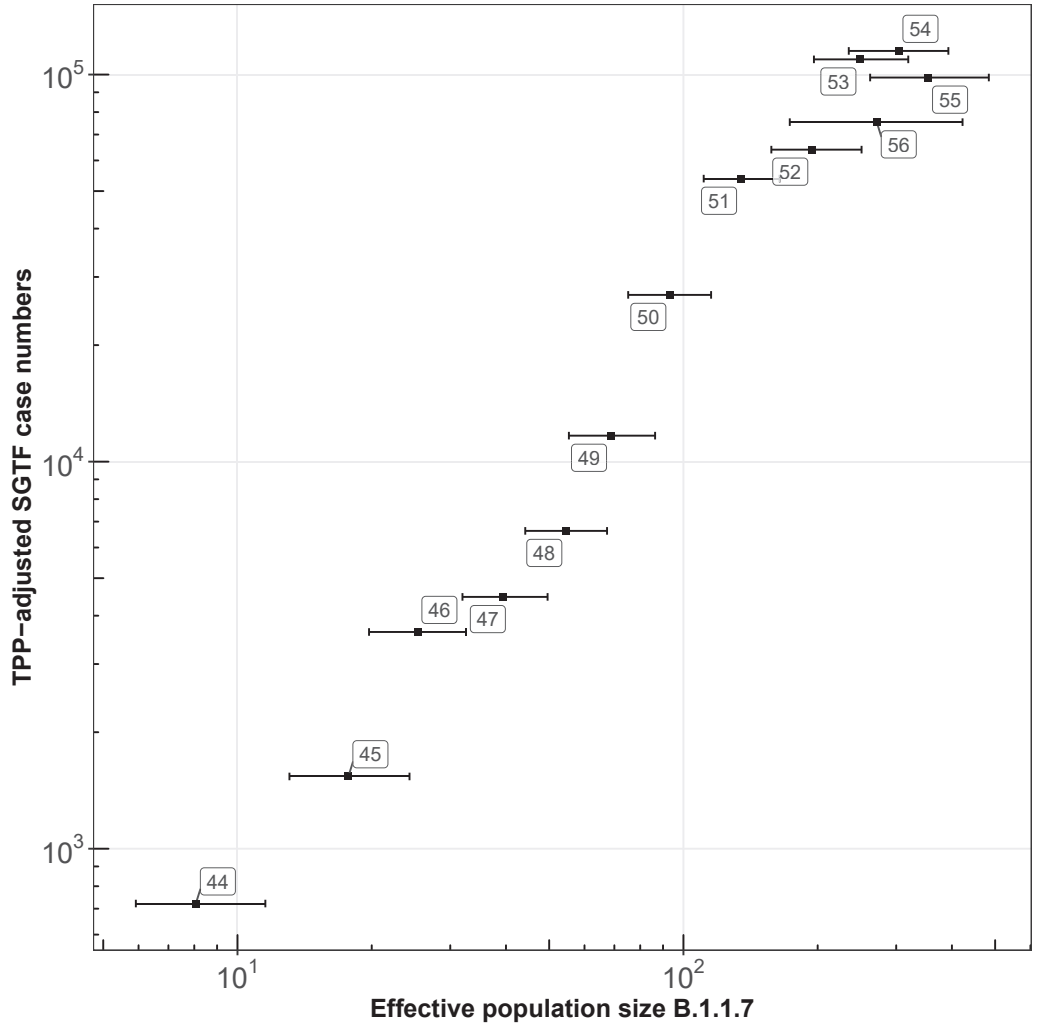


Figure 6.22 Effective population size of B.1.1.7 estimated by *mlesky* compared to TPP-corrected estimates of SGTF by epiweek.

Figure 6.23 shows the association between changes in mobility estimated by Google and reproduction number estimated through *mlesky*. Most associations have a significant negative coefficient, suggesting that mobility is associated with a reduction in transmission. The exception is for workplace mobility, which is positively associated with transmission. The sign of the coefficients is the same for each lineage, although B.1.1.7 always has a larger absolute central estimate.

It is important to note that mobility data was smoothed using loess with a parameter of 4%. As shown in Figure 6.5 with the blue lines, this closely resembles the trajectory of the noisy untransformed data (black lines). I performed a sensitivity analysis on this

assumption, by repeating the analysis with a smoothing parameter of 20%. This produces highly contrasting results, shown in Appendix Figure A.20, where the sign of the coefficients for almost all associations are reversed (with the exception of "Grocery and pharmacy"). It is therefore difficult to draw strong conclusions on the association between mobility and viral transmission.

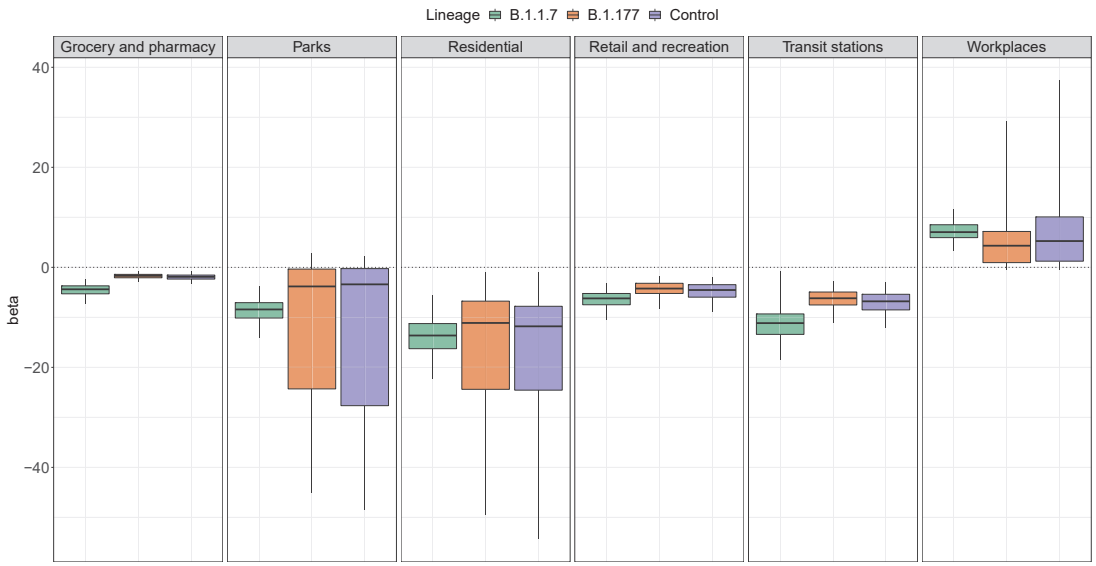


Figure 6.23 Distribution of beta coefficients when testing for univariate association between changes in mobility estimated by Google and reproduction number estimated in *mlesky*. Boxes represent median and interquartile range; whiskers denote 95% quantiles. Lineages B.1.1.7, B.1.177 and control are shown in green, orange and purple boxes respectively.

Figure 6.24 shows associations between presence of the winter lockdowns and the second order changes in effective population size. The coefficients are all negative for every association, suggesting lockdown is associated with a reduction in growth rate every day. The median coefficient value is lower for lockdown 3 than for lockdown 2 for B.1.177 and control lineages, the same pattern observed when investigating all lineages (see Figure 6.20). Further, for lockdown 3, the 95% CI do not overlap with zero for these two groups, forming a significant association. For these groups therefore, lockdown 3 has a more significant association with decreased transmission than lockdown 2. This pattern is not repeated for B.1.1.7, where neither lockdown forms a significant association with changes in growth rates. The median coefficient is more negative for lockdown 2, but B.1.1.7 was still at low frequency during much of this period (see Figure 6.21), thus rendering growth rate estimates

more uncertain. Unlike the two other groups, the B.1.1.7 95% CI overlaps with zero for lockdown 3, suggesting that this lockdown was not as effective in slowing B.1.1.7 as the others. Indeed it is probable that the effect of B.1.1.7 could have prevented a significant association between lockdown 3 and the daily change of growth rates of "All" lineages in the previous case study (Figure 6.20).

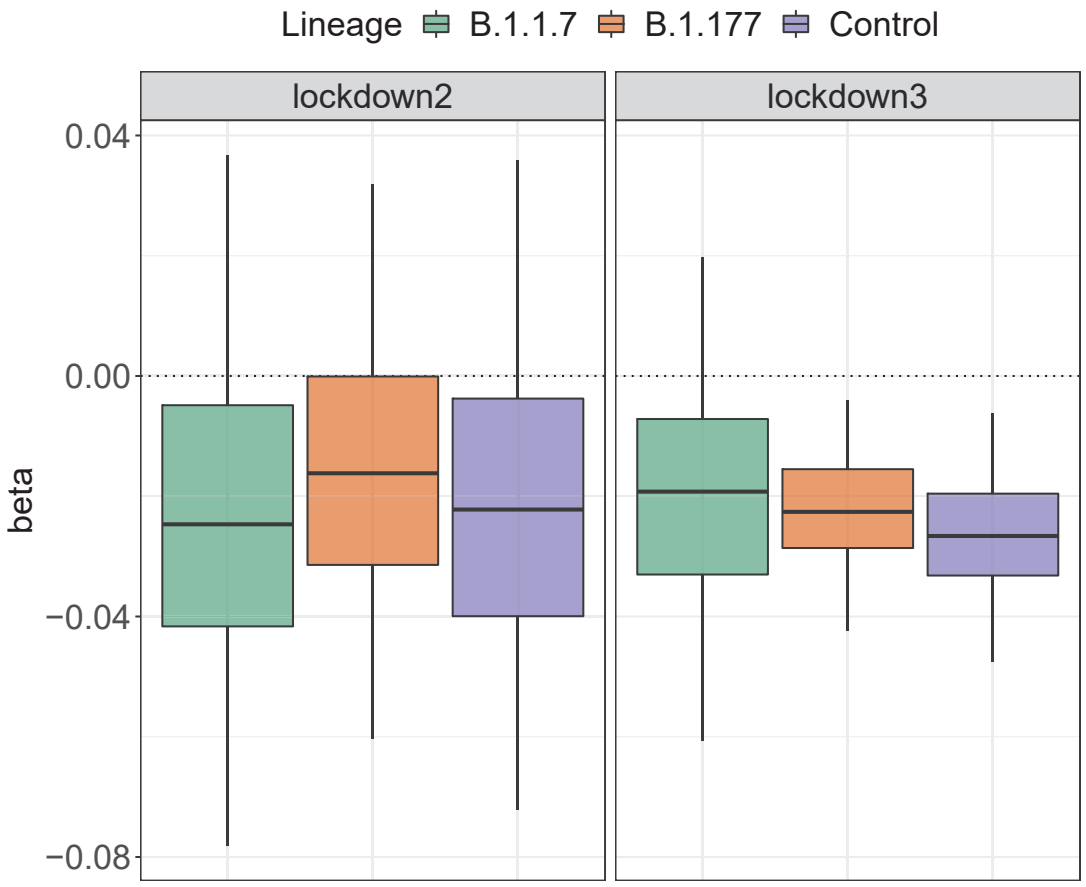


Figure 6.24 Distribution of beta coefficients when testing for multivariate association between presence of lockdowns 2 and 3 against second order changes in effective population size estimated in *mlesky*. Boxes represent median and interquartile range; whiskers denote 95% quantiles. Lineages B.1.1.7, B.1.177 and control are shown in green, orange and purple boxes respectively.

Figure 6.25 shows the associations of daily changes in OxCGRT indices against growth rate. For B.1.177 and control linages, the association is negative and significant for both indices,

with no overlap of the 95% CI and 0. For B.1.1.7, the associations are negative, but the effect size is smaller and overlaps with 0.

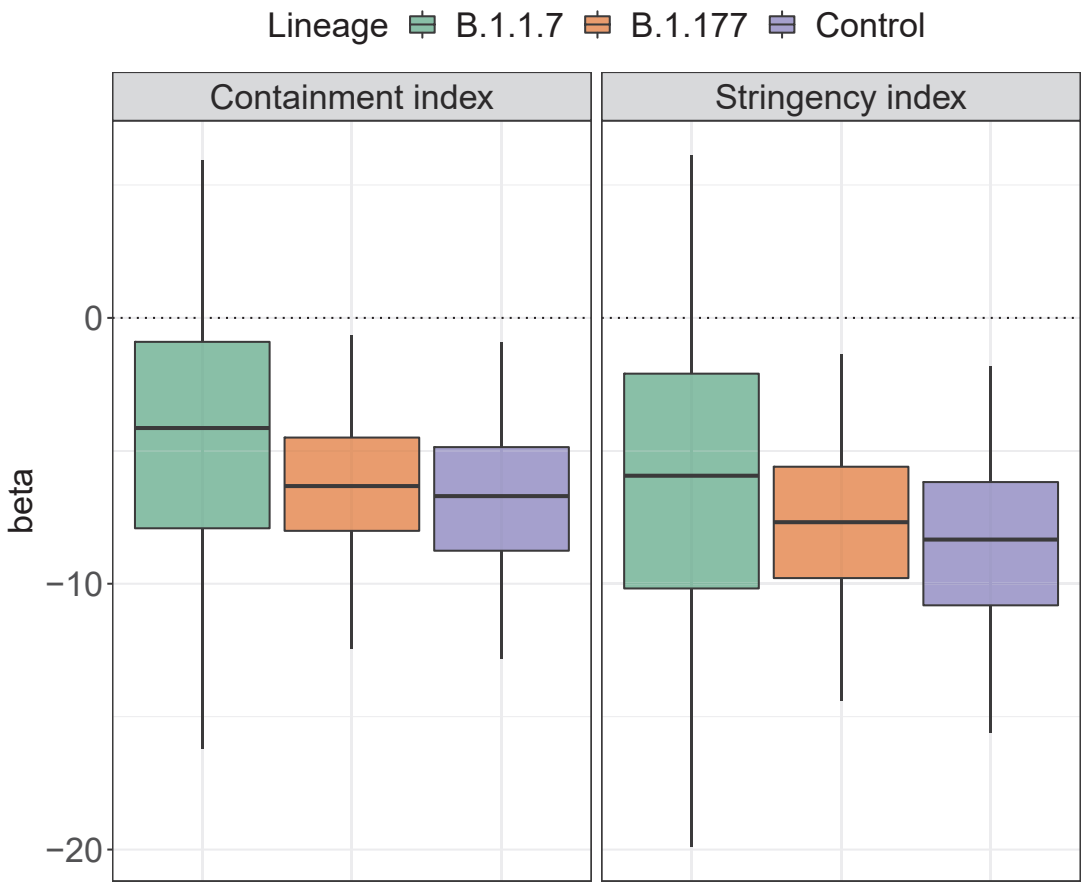


Figure 6.25 Distribution of beta coefficients when testing for univariate association between daily changes in OxCGRT indices and daily changes in effective population size (i.e. growth rate) estimated in *mlesky*. Boxes represent median and interquartile range; whiskers denote 95% quantiles. Lineages B.1.1.7, B.1.177 and control are shown in green, orange and purple boxes respectively.

Figure 6.26 shows the effect of introducing delays in the covariate time series for the OxCGRT Containment and Health index. For B.1.1.7 the effect size of the coefficient has the largest negative value with a time delay of 10-12 days; that is the association between the index at time t and R at time $t+10$ or $t+12$. When the delay is the other way round (i.e. changes in R precede those in NPI by 8 days), the association becomes non-significant. For the control lineage, the greatest effect size is found with a delay between 8-22 days.

Figure A.21 shows a sensitivity analysis using a loess parameter of 4% (as opposed to 20% in the main analysis). This introduces a lot more noise in the coefficient estimates (due to more temporal variance in the covariate), but the results are not qualitatively different: for both B.1.1.7 and control, the greatest significant negative effect sizes are found at 14-16 days delay.

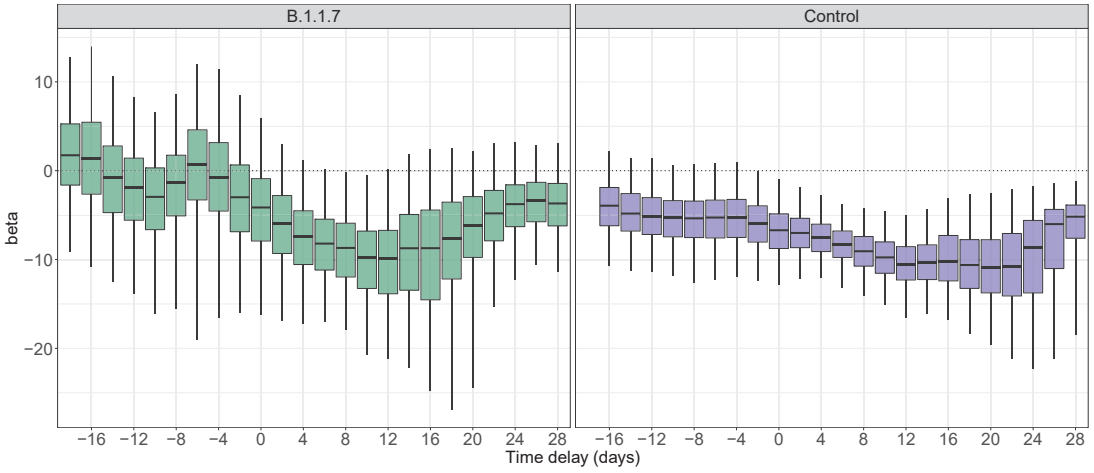


Figure 6.26 Distribution of beta coefficients when testing for univariate association between daily changes in OxCGRT health containment index and daily changes in effective population size (i.e. growth rate) in England estimated in *mlesky*. Boxes represent median and interquartile range; whiskers denote 95% quantiles. A positive time delay of 10 days represents testing an association between the covariate at time t and R at time $t + 10$ days. Lineages B.1.1.7 and control are shown in green and purple boxes respectively.

Including covariates with the 'joint' estimation method slightly changed the estimated epidemic trajectory but made no qualitative difference.

6.5 Discussion

6.5.1 The reliability of phylodynamic modelling for SARS-CoV-2

6.5.1.1 Robustness and corroboration from testing

In this chapter, I aimed to evaluate the epidemic trajectories of SARS-CoV-2 in England at different time periods and geographical levels, using non-parametric phylodynamic modelling. Further, I aimed to explore associations between viral transmission and non-pharmaceutical interventions, summarised as various covariates.

I used two phylodynamic models to estimate trajectories: *mlesky* and *skygrowth*. Where both models were used on the same phylogenies, the central estimates of $N_e(t)$ and $R(t)$ were remarkably similar, although the extent of uncertainty differed according to the particular analysis. This concordance added confidence in the reliability of the results and methods.

When fitting phylodynamic models purely to genomic data, epidemic peaks were always clearly defined and observable. Indeed both models predicted epidemic peaks within days of each other for every case study, sometimes estimating them to have occurred on the same day. Further, these estimated peaks preceded the peak in confirmed cases as reported by PHE by several days; this is consistent with delays in symptom onset and reporting. In Leicester, epidemic peaks became less clearly observable with smaller fractions of sequences included in the analysis; where there was less than 10% sampling the peak was unobservable. Historical reconstruction of SARS-CoV-2 epidemic trajectories using these phylodynamic models is therefore reliable on both a national and regional level, maintaining consistency with surveillance case data, but only when sufficient numbers of sequences are included.

6.5.1.2 Added evidence from analyses by lineage

Fundamental to the theoretical framework of phylodynamics is the incorporation of genetic information for the inference of epidemiological parameters. As such, phylodynamics is very well placed to explore these parameters for distinct viral lineages. By fitting phylodynamic models to phylogenies composed of different lineages, I could compare growth rates between them and against independent epidemiological data.

I found that B.1.1.7 grew significantly faster than other lineages, with an increased fitness of around 40%, during the period between 15th October 2021 and 1st February 2021. This was corroborated by other phylodynamic and statistical analyses [101], conducted by colleagues in parallel to my analysis. A renewal equation based approach estimated that B.1.1.7 had an increased fitness of 89% (95% CrI 43–165), but it was time dependent, decreasing to 54% (95% CrI 34–82) by mid January 2021. Fitting a semi-mechanistic model to TPP-adjusted SGTF frequencies, they estimated an increased fitness of 79% (95% CI 22-149), declining to around 50% by mid January 2021. My estimate for the added fitness of B.1.1.7 is consistent with the estimates from these two other analyses in later periods (i.e. from mid January 2021 onwards). The unique value of phylodynamic modelling is here evident, having fitted models solely to genomic data.

Further evidence that B.1.1.7 had faster transmission was the observed dominance of the lineage several months after the completion of this analysis in Europe, the US and elsewhere [412].

Even more strikingly, my estimated effective population sizes were reliable to a high temporal precision. Figure 6.22 is a plot of effective population size of B.1.1.7 against S-gene target failure, an independent proxy for B.1.1.7 frequency. The very strong relationship between the two measures is clear, and even as N_e begins to decline in week 56 so too does SGTF. The equivalent plots for the B.1.177 and control lineages (Figures A.18 and A.19) show no relationship at all. Phylodynamic modelling is therefore highly effective at estimating viral frequencies by lineage through time, to a precision of at least a week, when benefiting from high throughput and approximately random genomic sampling from the community.

6.5.1.3 Phylodynamic reconstruction is robust to using global ML tree

As described in Section 6.3.4, global maximum-likelihood trees were routinely produced in the daily pipeline when each new set of new samples was added to the database. To obtain trees for phylodynamic analysis, I used this global tree and simply kept the tips of interest while dropping the rest. This maintained genetic distances between samples while avoiding the computationally expensive task of re-estimating thousands of trees. To check whether results are affected by this step, I performed a sensitivity analysis where I re-computed the maximum-likelihood tree from scratch using *IQTREE*. Figure A.17 shows that the epidemiology trajectories are nearly indistinguishable as a result, and therefore that the economising of computational resources was justified. More importantly, it suggests that if results from future phylodynamic modelling are urgent, for example regarding an emergent variant of concern, the time-consuming step of ML tree reconstruction can be avoided if the tips are already represented in a global tree.

6.5.2 Impact of NPIs

A secondary aim of this chapter was to evaluate the impact of lockdown and other NPIs on SARS-CoV-2 transmission. Before exploring any causal effects, first correlations must be established. As described in Section 6.3.5.2, I used two approaches to estimate the association between NPI covariates and viral transmission.

6.5.2.1 How do we best evaluate association between NPI and viral transmission?

The *post hoc* approach involved first estimating the viral epidemic trajectory fitting solely to phylogenies. This trajectory was then assumed to be the 'true' data against which covariates were fitted in a linear model. The effect size and significance of the coefficient were estimated to explore whether there was any association.

Broadly, the *post hoc* method generated very few clear associations. That is, the absolute values of the coefficients (and r-squared values) were close to zero, and only one of any covariate in any case study was found to be significant to 5%. Although this significance threshold is arbitrary, the majority of p-values evaluated exceeded 30%, indicating that associations can be rejected with reasonable confidence. Indeed, I present 24 individual analyses using the *post hoc* method (3 case studies with 8 covariates); there is therefore an expectation of one association emerging at random, assuming a 5% probability. The *post hoc* method therefore did not succeed in discovering any consistent associations between NPI and viral transmission.

However, in the analyses presented with the *post hoc* method, I had fitted a linear regression against daily incremental changes in covariate and $R(t)$. It is possible that such short time increments created a lot of noise, and genetic data will not allow detection of short term changes in growth rates. The signal may be easier to detect within the noise if the time increments were increased, although the time increment should be chosen algorithmically. This will feature as future work described in Chapter 7.

The joint estimation method, however, produced many significant associations. For the *post hoc* method, the coefficient of linear regression describes the slope of fitted covariates against median estimated $R(t)$ across all trees. Phylogenetic uncertainty is therefore unaccounted for when using only the median trajectory. The β in the joint method represents the fitted estimate of beta which produces the most likely expectation in the Markov Chain when smoothing the trajectory of effective population sizes. This fitting process also optimises for different trajectories of $N_e(t)$, and therefore $R(t)$, although I found that these were only slightly different from trajectories estimated without covariates. Crucially, this joint optimisation step was carried out for each analysis (each tree) individually, such that the distribution of β values captures phylogenetic uncertainty.

A further benefit of the joint method is that binary variables can be included as covariates, and their associations with transmission can be therefore explored. The joint method is therefore, on balance, better at discovering associations of covariates with epidemic trajectories, as well as enabling the evaluation of phylogenetic uncertainty.

6.5.2.2 Did stricter restrictions control the virus better?

As described in Section 6.2.4, the OxCGRT was the most prominent attempt to quantify governmental response to COVID-19, and provided an excellent continuous covariate to implement into phylodynamic analysis, including all kinds of non-pharmaceutical interventions. My analyses found that, on a national scale, increased restriction stringency was strongly associated with decreased viral transmission. For example, during the first wave, both the 'Containment Health index' and 'Stringency index' [214] produced a distribution of β coefficients of negative values, for which the 95% quantiles did not overlap with zero (Figure 6.10).

The causal relationship is however very difficult to determine: although it is intuitive that viral transmission decreased because of enforced reduction in contacts, it is also clear that UK government stringency increased as a response to rising cases and hospitalisations. Causality is therefore likely to be bi-directional (or mutually re-enforcing). Nevertheless, introducing delays into the analyses can increase confidence in the claim that restrictions decreased transmission. For the first wave in England, the association was strongest when I tested changes in restrictions against reproduction number 8 days later (Figure 6.11). If several days are required for government policy to be implemented and for the public to respond, then this makes intuitive sense. Conversely, when I tested association between changes in restriction and reproduction number 8 days before, the 95% quantiles of the β distribution overlapped with zero. This result strongly adds weight to the argument that restrictions decrease viral transmission after several days.

When analysing the winter period (lockdowns 2 and 3), a similar pattern emerges (Figure 6.25). There was a negative association between both OxCGRT indices and viral transmission. The only lineage for which the β coefficient distribution overlapped with zero was for B.1.1.7. It is possible that this effect is due to the higher transmissibility of this lineage [101]. However, once again, the introduction of time delays produced a stronger negative association, even for B.1.1.7 (Figure 6.26). The association is most negative when changes in restriction are associated with growth rate 12 days later, and overlap with zero when association with growth rate 12 days earlier. This result is highly encouraging, and adds to the intuitive conclusion that interventions can reduce viral transmission.

The sensitivity analysis shown in Figure A.21 repeats the analysis investigating time delays in health containment index and growth rates of B.1.1.7 and control, except with a smaller smoothing parameter value (4% instead of 20%) for the covariate. The results are qualitatively identical, adding robustness to the conclusions.

For the local analysis in Leicester, however, no significant associations were found. It is unclear whether the brief spike in Leicester was really due to increased transmission, or rather due to imported cases from other regions. It is possible therefore that the local outbreak could have run its course without the need of a local lockdown. Alternatively, the lack of association could be due to the data I used. As discussed in Section 6.3.1, the OxCGRT data was only available on a national level. Indeed it was the unique local lockdown in Leicester that motivated my exploration of this region, and the local restrictions were not reflected in the national OxCGRT covariate data. As such, no strong conclusions regarding the efficacy of local lockdowns can be made in this case.

6.5.2.3 Did movement correlate with transmission?

During the first wave in England, mobility was generally negatively associated with transmission, with the exception of 'residential' (Figure 6.9). However, the early months reflected very sparse non-random sampling, mostly from hospital samples. As a result, phylodynamic reconstruction of the epidemic beginning is perhaps less reliable than other time periods. Nevertheless, residential movement may indeed have been associated with more transmission. A prospective case ascertained study estimated a secondary attack rate of 37% (95% CI 31–43%) in UK households [452], which increased in smaller households; globally, a systematic review and meta analysis estimated secondary attack rates in households to be 16.6% [453]. Prior to the widespread messaging of self-isolation and recognition of COVID-19 related symptoms, household transmission may have represented an important portion of early transmission.

Regarding the other types of mobility, the negative associations of other mobility with transmission may simply be a confounding of the positive association of residential mobility. For example, increased park mobility may not be itself associated with reduced viral transmission, but rather it is associated with reduced household mobility (itself associated with transmission), as people cannot be in two places at once. Indeed, Figure 6.1 clearly shows that as other types of mobility decrease around the time of the first national lockdown, residential movement follows the opposite pattern, as people spend more time near their homes.

During the winter period in England, most mobility types are negatively associated with transmission. The exception here is workplace mobility which was positively associated. This pattern was identical for all lineages explored (Figure 6.23), including for B.1.1.7 whose trajectory was entirely different to B.1.177 and control (Figure 6.21). By this period of the epidemic, after many months of messaging and lockdowns, movement had massively

decreased overall [454]. Further, non-essential retail, pubs, restaurants were all closed during these months thereby limiting opportunity for infection in those venues. It is plausible therefore that most transmissions took place in workplaces. Again, the negative associations with other forms of mobility may simply be confounded by less movement in the workplace.

On the other hand, secondary attack rate for B.1.1.7 is estimated to be higher in households [418]. Therefore, a greater association between residential movement and viral transmission was expected during the winter period, which was not borne out by my analysis.

On top of issues with confounding variables, another complication is that the degree of smoothing applied to the highly noisy mobility data has a highly significant impact on the results. Figure A.20 shows that by assuming a different smoothing parameter, central tendency of the β coefficient switches sign for most forms of mobility, qualitatively changing the interpretation. This sensitivity makes drawing strong conclusions on the association of mobility and transmission highly problematic, and more in-depth tuning of the smoothing would be required to produce results with more confidence. It is important to re-iterate that as discussed in the previous subsection, varying smoothing of the OxCGRT data did not qualitatively change the results.

For the local analysis in Leicester, no form of mobility was found to be associated with viral transmission. Indeed, no associations of any kind were found between any covariate and growth rate in Leicester. One reason could be that due to small data sets, *mlesky* produced reasonably wide uncertainty intervals. Even at 70% sampling, there were only 1,058 sequences, whereas for the national analyses there were 5,000 and 3,000 for the spring and winter waves respectively (see Table 6.1). Furthermore, despite smaller data sets, the time period analysed was up to twice as long as other case studies, e.g. around 6 months in Leicester vs. around 3 months for the first wave in England. This has the effect of "spreading thin" the data and increasing uncertainty. Indeed to produce readable plots (e.g. Figure 6.12), the effective population size is plotted on a log scale, which makes this uncertainty less obvious. As a result of this longer time period, variance in mobility is smaller over time in Leicester compared to other case studies, particularly after smoothing. This could also serve to reduce any signal when testing for associations. Indeed the largest change in mobility in Leicester occurred in late March and early April following the first national lockdown (Figure 6.3), a period which coincides with sparse sampling and large uncertainty bounds in the phylodynamic analysis (Figure 6.12). As a result, the signal which would most likely lead to association with viral transmission is occluded by this uncertainty.

In summary, the associations between mobility and transmission are highly variable throughout the year. The interpretation of these associations must be conducted with caution, due

to the nature of confounding of other variables, sparse sampling causing large uncertainty, and because the results are highly sensitive to the degree of smoothing of the noisy covariate data.

6.5.2.4 Was lockdown 3 more effective than lockdown 2?

Finally I tested the association of lockdowns on transmission when coded as binary variables. This analysis was conducted for lockdown 2 and 3 together as a multivariate analysis. When analysing all lineages during the winter period, the 95% quantiles of the distribution of β coefficients overlapped with zero for both lockdowns, although the central tendency was positive for lockdown 2 and negative for lockdown 3. When exploring the differences by lineage, there was no overlap with zero for lockdown 3, except for B.1.1.7.

As B.1.1.7 is very likely to be more transmissible [101], this result makes intuitive sense: this lineage was less impacted by the strict third lockdown than others. However, the interpretation of these results is not so simple. Because of the binary nature of covariate, the association measured in this analysis is between the presence/absence of lockdown and the second-order changes in effective population size, which is equivalent to the rate of change of the growth rate (see Section 6.3.5). It is not trivial to interpret the meaning of such an association, and the aforementioned analyses which test associations with the growth rate itself are more illuminating. Despite this, the direction of the result is intuitive and encouraging, as it corroborates the widespread understanding that lockdown 3 was more stringent than lockdown 2 [455].

6.5.3 Limitations

Beyond the limitations particular to each individual analysis, some broader limitations emerge. Generally, imperfect availability and sampling strategies of genomic sequences are issues which incur the greatest limitations.

The COVID-19 epidemic in England, as elsewhere, is defined by high heterogeneity in the number of weekly cases. On the December 29th 2020, at the height of the second wave, there were around 72,500 cases reported in England; on 12th April 2021 after months of lockdown and widespread vaccinations, the daily case count was below 1,000 [386]. Although COG-UK's mission is to maintain high sequencing throughput all year round, sampling intensity through time may depend on viral population size: during periods of low burden there are few cases from which to achieve the desired number of samples. As the precision of estimates

of effective population size inference can depend strongly on sampling intensity [456], our results could be impacted by this effect. Pillar 2 sampling has however been found to agree strongly with independent, designed surveys of prevalence, i.e. from REACT-1, REACT-2 and the Office for National Statistics [457, 458]. This provides confidence that pillar 2 sampling is approximately random.

However, early on in the epidemic when widespread sequencing was not available, samples tended to be sparse and derived from hospitals rather than from the community. These samples are likely to be biased towards severe cases and older patients. As a result, phylodynamic reconstruction of these periods can be less reliable, reflected in wider uncertainty intervals, as was the case in Leicester in March. Further, testing associations of covariates with viral transmission during periods of sparse sampling is likely to be difficult.

Related to this, the number of sequences over time may constitute another potential sampling bias and therefore a limitation. It is very likely that sampling intensity varied through time; in particular early in the epidemic sampling intensity was likely to be lower due to a weaker sequencing infrastructure. This could have been accounted for by using novel epoch sampling methods [456], or by including number of sequences per grid point as a covariate in the analysis. The number of sequences broadly increased through time in the early months of 2020 (case study A), in similar fashion to the intensity of NPIs as summarised in the OxCGRT index (Figure 6.2). Implementing number of sequences per grid point as a covariate is therefore likely to lead to similar associations as those observed when implementing OxCRT indices as covariates.

Evolutionary rate may pose another limitation. If evolution occurs very quickly, then changes in N_e may not be easily detectable and may be smoothed out.

The associations between NPIs and viral transmission depend not only on genomic sequence availability, but also the quality of the NPI data. Although Google provided high resolution information on movements of different types by region, associations of these with viral transmission were highly sensitive to the degree of smoothing applied to the data. Very careful sensitivity analyses must be explored before affirming strongly any kinds of association. On the other hand, applying data which are too sparse, such as a binary variable describing the presence or absence of lockdowns, brings about other problems in the interpretation of results. A data set which balances these two concerns provides the best covariate with which to test associations.

6.5.4 Conclusions

This chapter has shown phylodynamic modelling to be an extremely powerful method for producing epidemiological information from viral genomic data. The analyses described in this chapter confirm that the high throughput sequencing of SARS-CoV-2 in England throughout the epidemic has permitted high resolution and reliable phylodynamic analysis. When comparing estimated trajectories against independent data, such as confirmed cases or S-gene target failure, each analysis produces good insight into historical epidemiological processes, including those of the recent past. Further, provided that appropriate care is applied in interpretation, associations between governmental interventions and viral transmission can be explored.

Phylodynamic modelling can produce estimates of growth rates with a quick turnaround, particularly if the phylogenetic trees are already constructed. For real time COVID-19 response, however, the limiting factor tends to be the availability of recent samples; by the time that samples have been taken, sequenced and processed, the added value of phylodynamic modelling for urgent response may have already expired. Efforts to reduce delays in the production of available sequence data are highly encouraged to best exploit these methods.

Chapter 7

Discussion

7.1 Overview

In this thesis, I aimed to explore the impact of interventions on viral epidemics using epidemiological and evolutionary models. To do this, I applied various modelling frameworks to investigate the viral epidemics of HIV and SARS-CoV-2 in different contexts. Each framework enables a different method for estimating the impact of pharmaceutical and non-pharmaceutical interventions.

In the first section of this chapter, I summarise the contents of the previous chapters. I present the main conclusions regarding the impact of interventions in each context, to demonstrate how I met my research aims. I also mention significant publications and presentations of my work.

Next, I explore the implications of my research in the context of the wider literature concerning HIV and COVID-19 control respectively. In this section I also consider potential future work.

In the following section, I extract and discuss key themes that run throughout all chapters, allowing me to consider the strengths and limitations of each analysis, and their applicability in each context.

Finally, I submit my final conclusions of this thesis.

7.2 Summary of this work

Chapter 1 introduced background material for understanding the current state of research prior to undertaking each analysis. I presented the global epidemiology and natural history of the two viruses studied in this thesis (HIV and SARS-CoV-2), as well as interventions that are commonly used to advance their control and minimise burden. Finally, I discussed the various methods previously employed to assess the impact of these interventions.

Chapter 2 built the foundation for the transmission model analysis of HIV in Cotonou, Benin. After providing the background of the local HIV epidemic, I summarised the available data used to parameterise and fit the transmission model, whose development I described in the following section. The fitted model provided the framework for the analyses in Chapter 3 and Chapter 4.

This chapter provided the first comprehensive summary of available demographic, epidemiological, sexual behaviour and intervention data between 1993-2017 in Cotonou.

Significant elements of this chapter were featured in the supplementary materials in an article with the following citation:

***Geidelberg, L.**, Mitchell, K. M., Alary, M., Mboup, A., Béhanzin, L., Guédou, F., ... & Boily, M. C. (2021). Mathematical Model Impact Analysis of a Real-Life Pre-exposure Prophylaxis and Treatment-As-Prevention Study Among Female Sex Workers in Cotonou, Benin. Journal of Acquired Immune Deficiency Syndromes, 86(2), e28.*

Chapter 3 used the fitted model described in Chapter 2 to estimate the epidemic trajectory of HIV in Cotonou, projected between 1986-2035. This trajectory suggested that the pharmaceutical and non-pharmaceutical interventions present before 2015 (e.g. ART and condoms, respectively) were instrumental in controlling the HIV epidemic and diminishing incidence. I also estimated the impact of the recent two-year PrEP/TasP demonstration study among FSW, as well as long-term PrEP and TasP scale-up scenarios with varying adherence and coverage. I concluded that while PrEP had greater impact among the FSW in the study, TasP had a greater impact overall, mostly due to better observed uptake, adherence and retention.

I presented elements of this model analysis at CROI 2019, Seattle, Washington, United States. A webcast of this talk can be found here:

Wednesday, 11:45am 6th March 2019, Talk 99 (Theme: TESTING, VIRAL SUPPRESSION, AND IMPACT) http://www.croiwebcasts.org/console/player/41202?mediaType=audio&

The full model analysis was published in the main text of the research article already mentioned for Chapter 2 (Geidelberg, L. et al. 2021).

Chapter 4 explored the impact of COVID-19 on the HIV epidemic in Cotonou. I first described the COVID-19 epidemic in Benin. Next, I detailed scenarios of potential disruptions to already-present HIV interventions. Two of these scenarios were based on the available data, and the others were exploratory. Using the fitted model from Chapter 2, I estimated the impact of these scenarios on HIV transmission and deaths. I concluded that the most likely real-life scenarios had modest impacts on HIV outcomes, but that key populations were disproportionately affected. The exploratory scenarios highlighted the effectiveness of condoms which, if interrupted, could cause significant increases in HIV incidence.

This model analysis formed half of a research article comparing the potential impact of COVID-19 on HIV epidemiology in Yaoundé and Cotonou. The published article has the following citation:

*Silhol, R., **Geidelberg, L.**, Mitchell, K. M., Mishra, S., Dimitrov, D., Bowring, A., ... & Boily, M. C. (2021). Assessing the potential impact of disruptions due to COVID-19 on HIV among key and lower-risk populations in the largest cities of Cameroon and Benin. Journal of Acquired Immune Deficiency Syndromes, 87(3), 899.*

Chapter 5 evaluated the impact of non-pharmaceutical interventions (NPIs) in response to COVID-19 in Weifang, China. I estimated the epidemic trajectory using the model-based phylodynamic method *PhyDyn*, which indicated that disease control (represented by $R(t)$ falling below 1) was achieved a few days after the imposition of strict NPIs. I also found evidence from phylogenetic analysis that travel restrictions and contact tracing were effective at limiting the number of importations of COVID-19 into Weifang.

This phylodynamic analysis was published with the following citation:

***Geidelberg, L.**., Boyd, O., Jorgensen, D., Siveroni, I., Nascimento, F. F., Johnson, R., ... & Nie, Q. (2021). Genomic epidemiology of a densely sampled COVID-19 outbreak in China. Virus evolution, 7(1), veaa102.*

Chapter 6 evaluated the impact of NPIs in response to COVID-19 in England. This chapter was divided into four case studies focused on different time periods (spring 2020, summer 2020, winter 2021), regions (whole of England, Leicester), and genetic variants (B.1.1.7, B.1.177 and control). I applied two non-parametric phylodynamic models (*skygrowth* and *mlesky*) to estimate the epidemic trajectory of COVID-19 in each case study. I used three proxies for summarising NPIs in England: lockdown as a binary variable, mobility as a measure of travel restrictions (Google), and indices of stringency and containment (OxCGRT). I used *mlesky* to explore the associations of these covariates with viral growth rates, and concluded that in most cases, NPIs were associated with epidemic control. I found that the variant of concern B.1.1.7 had 40% higher transmission than previously circulating lineages, and was less impacted by NPIs.

Significant elements of this chapter are published or are in press; the specific details are as follows.

Case study A investigated the epidemic trajectory and impact of NPIs during the first COVID-19 wave in England in spring 2020. This analysis formed a significant real-world example of application for the new *mlesky* method in a research article, currently under review. The pre-print is available at the following link:

Xavier Didelot, ***Lily Geidelberg****, The COVID-19 Genomics UK (COG-UK) consortium, and Erik M Volz. "Model design for non-parametric phylodynamic inference and applications to pathogen surveillance". Pre-print available at:*
https://www.biorxiv.org/content/10.1101/2021.01.18.427056v2

Case study D investigated the epidemic trajectory and impact of NPIs for COVID-19 in England during the winter 2020-2021, and explored how this is affected by genetic lineage (B.1.1.7, B.1.177 and control). Significant elements of this analysis were presented in an article exploring the increased fitness of the B.1.1.7 lineage (e.g. Fig. 1b and 1c, Extended Data Fig. 3).

Volz, E., Mishra, S., Chand, M., Barrett, J. C., Johnson, R., ***Geidelberg, L.****, ... & Ferguson, N. M. (2021). Assessing transmissibility of SARS-CoV-2 lineage B. 1.1. 7 in England. Nature, 593(7858), 266-269.*

7.3 Implications for controlling HIV epidemics

Chapters 2, 3 and 4 of this thesis concern the HIV epidemic of Cotonou, Benin. In this section I discuss the implications of these results in light of the wider literature on HIV epidemic control.

7.3.1 Pharmaceutical interventions among FSW

The United Nations, together with UNAIDS and other partners, aimed to end AIDS as a public health threat, as part of the Sustainable Development Goals [459]. They defined this as a reduction of 90% in both new HIV infections and HIV-related deaths by 2030 compared to 2010 [287]. It is widely understood that controlling the HIV epidemic will require a combination of prevention tools, including voluntary male circumcision, changes to sexual behaviour, needle exchange, condom use, PrEP and antiretroviral treatment as prevention, among others [460]. Indeed, the pharmaceutical interventions considered in Chapter 3, PrEP and TasP, are two potential options which can be introduced within a broader context of HIV prevention tools [156].

I estimated that during the 2-year period of the PrEP/TasP study, PrEP prevented more infections among professional FSW than TasP. However, when measured over 20 years, the impact of TasP was higher among FSW than PrEP. Further, when considering impact in the whole population of Cotonou, TasP was more impactful than PrEP both over 2 years and 20 years. The main reasons for this include better uptake, adherence and retention, as observed in the study. Moreover, given that professional FSW have the highest HIV prevalence among all groups in Cotonou, reducing their viral load with treatment confers significant protection to their male partners which, over long time periods, will in turn protect other FSW and low-risk women. Further to the study, in long term projections I estimated that 20-year TasP interventions are also more impactful than PrEP ones. This model analysis in Cotonou therefore suggests that TasP is preferable to PrEP for maximising impact among all groups.

This result is well replicated in previous studies. The impact of TasP, or of increased ART eligibility for the purposes of prevention, has been estimated to have large impact on HIV infections averted throughout Sub-Saharan Africa, including in South Africa [461, 462], Nigeria [18], Kenya [463] and Zambia [464]. Presently, in many countries, all HIV-infected individuals are eligible for ART [465], including in Benin [169]. TasP is therefore already a current policy de facto. The impact results for TasP in my analyses support this decision, and are consistent with the wider literature.

Whether to include daily oral PrEP as part of the HIV prevention package in Cotonou, however, remains a difficult question. As of 2021, PrEP is not widely available in any form in Benin (expert opinion: Michel Alary, MD). Nevertheless, in 2015 the WHO recommended the provision of PrEP to all at-risk groups [466], a policy which has been adopted into national guidelines of most countries (including Benin [467]). Since then, the number of users of daily oral PrEP has increased every year in all global regions, with the greatest recent expansions in Africa [468]. With the advent of newer forms of PrEP, including the long-acting injectable cabotegravir (CAB LA) and dapivirine vaginal ring (DPV-VR) [469], the adoption of PrEP in some form is likely to continue and be considered in Sub-Saharan African countries like Benin, with priority for key populations.

The impact analysis for PrEP among FSW in Cotonou therefore has important implications. One of the key reasons for its modest estimated impact was that the observed adherence and retention to daily oral PrEP of FSW in the study were poor [169]. These data were more realistic than those observed for women in the clinical trials [284, 470, 161, 471]. Firstly, the PrEP/TasP study was in a real-life setting for FSW, who were recruited by community workers familiar with the population. FSW are at high risk of HIV acquisition, an example of a key population who should be prioritised for PrEP according to WHO recommendations, as compared to the relatively low-risk women recruited to the clinical trials. Secondly, the measure of adherence was biological, using blood plasma concentrations of the active drug tenofovir. Pill counts and self-reporting are more subject to biases and underreporting of non-adherence, which was indeed an observation among the FSW in Cotonou [169, 275]. It is reasonable to consider the observed data in Cotonou as close to a realistic possibility when expanding PrEP to key populations in the community.

An important explanation for poor adherence and retention was the mobility of the women in the study [169]. It has been long documented that the FSW of Cotonou have a diverse background, with the majority being migrants from neighbouring countries in West Africa [254]. They come to Cotonou where there is more custom and less severe sanctions on sex work. However, they periodically return to their home nations either temporarily or permanently. Moreover, many retire from sex work after only a few years. It is therefore very difficult to maintain good adherence and retention to a daily oral PrEP programme for such a mobile population. Unless indicators of long-term retention can be ascertained prior to enrolment, the per capita impact is unlikely to be large.

One potential solution is the introduction of CAB LA, once it receives regulatory approval. It has been demonstrated in a clinical trial (HPTN-084) in seven Sub-Saharan African countries to be 89% (95% CI 68-96%) more effective at preventing HIV infection than daily oral PrEP

among HIV-uninfected women [172], a difference attributed to the difficulty of adhering to a daily oral regimen [472]. Indeed, as I discussed in Chapter 3, the 'perfect adherence' scenarios that I estimated can be interpreted as the impact of CAB LA. The impact of these are around double those estimated for daily oral PrEP among FSW and overall, underscoring a very important difference. If CAB LA is a solution to poor PrEP adherence among women, then my analysis suggests that it would be a better option for FSW in Cotonou and in similar settings.

There are however some important issues to consider. Although acceptability for long-acting injectable PrEP seems high (>75%) in low-risk women from Zimbabwe and South Africa (HPTN-076) [473], it is not certain that similar results would occur for high-risk women such as FSW in Benin. Further, mobility may still pose problems. The half-life of CAB LA is between 21-50 days [474], meaning repeat doses are required every 8-12 weeks. If FSW in Cotonou are sporadically leaving the city, then returning to re-engage in sex work, continued retention to the PrEP programme at the clinic may still be problematic. Even with sure protection in the short to medium term, long term protection is still uncertain. It would require data from real-world implementation studies to obtain a better picture of the likely adherence and retention, and therefore of the likely impact of long-acting PrEP in this context.

7.3.2 Continued focus on key populations

As discussed in Chapter 1, key populations are considered a good choice for HIV prevention interventions, as they generally suffer a disproportionate burden of HIV [355]. Further, resulting from patterns in sexual contact networks, HIV prevention among key populations can prevent secondary transmission events, leading to more impactful interventions [475].

The results from the model analysis in Chapter 3 corroborated this effect. It is important to remember that PrEP and TasP were only provided to FSW; in terms of HIV acquisition, PrEP directly protects the individual users, while TasP only directly protects their partners. As discussed earlier, when considering the long-term picture, TasP prevents more infections than PrEP not only among clients and overall, but also among FSW themselves. This is a striking result, which underscores the importance of secondary transmissions and the wider sexual network.

There is further evidence from my analysis of the impact of COVID-19 on HIV epidemiology, in Chapter 4. I estimated that the likely COVID-19 related disruption and impact was modest, given the available data. However, this impact was not distributed evenly, and key populations

such as FSW and clients were disproportionately affected by disruptions, particularly those to condom use. These results suggest that priority should be made to maintain good condom use and ART coverage among these key populations, to avoid the worst negative consequences of current and future waves of COVID-19 in Benin. The decision to continue to focus on key populations by UNAIDS and WHO is therefore strongly supported by this thesis.

7.3.3 Pharmaceutical vs. non-pharmaceutical interventions

Pharmaceutical interventions are an undoubtedly essential part of the HIV prevention toolkit. It is nevertheless vital to remember that non-pharmaceutical interventions are also indispensable; in some cases they are even more impactful. Following the international AIDS conference in 2002, the recent advent of antiretroviral therapy was celebrated, and scientists encouraged that it be expanded to the millions of HIV-infected individuals worldwide [476]. Despite this clamour, an editorial was published in Sexually Transmitted Infections in 2003, warning about the diverted attention from primary prevention, and specifically about the effectiveness of condoms [477]. They argued that although treatment is essential and should be expanded, the international community must also focus on improved promotion of condom use, and respond to obstacles to condom use, such as poor access, negative attitudes and limited supply.

In the past 20 years, improvements in both condom use and ART coverage have been observed. In Cotonou, condom use by FSW with their clients rose from around 65% of sex acts in 1998, to 98% in 2015 [255, 259]. Condom use in low-risk populations increased from around 4% of sex acts in 1998 to around 21% in 2011 [280, 478]. More recent estimates are unavailable, although a descriptive cross-sectional analytical study in 2017 in Natitingou, Benin, estimated condom use of high school students was around 40% [479], suggesting that with the inflow of younger generations into sexually mixing populations, condom use is still on the rise. Since its introduction in 2000, ART coverage has risen to a median estimate of 83% and 79% among FSW and overall respectively in 2017 [259, 342].

The epidemic trajectory presented in Chapter 3 suggests that, if the PrEP/TasP study had not happened, HIV prevalence would continue to decline with the interventions already present, decreasing by 82% from 2010 to 2030, approaching the UNAIDS goals for 2030. Although the effects of ART are undeniable, so too is the impact of condom use; indeed a previous study estimated that condom use prevented an estimated 62% (52–71%) and 33% (20–46%) of HIV infections in Cotonou among FSWs and overall respectively, between 1993 and 2008 [50]. Moreover, my analysis in Chapter 4 demonstrated that any disruptions to condom use,

particularly for commercial sex, would lead to greater increases in HIV infections compared to any other disruption, including to viral suppression. There are large proportions of sex acts still unprotected among many populations in Cotonou, and closing this gap would be a very safe and inexpensive means to further control HIV. All of this evidence points towards the fact that even twenty years later, the message published in Sexually Transmitted Infections remains true: condom use must continue to be promoted, and access widened, even during the excitement surrounding novel pharmaceutical interventions such as long acting PrEP or a potential HIV vaccine.

Finally, in the last few years, there has been some discussion in the literature surrounding structural interventions. A well-studied example was the Avahan initiative in India, which focused on community and social mobilisation as a form of HIV prevention [480]. This included activities such as crisis response, legal advice and representation, and the development of groups for marginalised populations such as sex workers. Data from the initiative showed that sexual violence and certain engagements with the police incurred risk of HIV infection among FSW in India [481]. There are similar findings in Sub-Saharan Africa: a modelling study of structural determinants for HIV risk estimated that eliminating police violence could prevent 17% of infections in Kenya in the overall population [482]. The authors further estimated that decriminalisation of sex work could prevent 33-46% of infections worldwide over 10 years, by reducing incentive for police aggression towards illegal activities. The COVID-19 pandemic period has seen exacerbation for the problems of violence [483] and homelessness [484] in vulnerable communities. Although structural interventions are difficult to measure, they should also be considered along with other non-pharmaceutical interventions in the suite of HIV prevention tools.

7.3.4 Future work

There are several avenues of future research which could extend the implications of this thesis.

Health policy decisions are not only concerned with impact, but also with the economic implications of new interventions, particularly in resource-limited settings such as Sub-Saharan Africa. A cost-effectiveness analysis is one way to evaluate interventions against health opportunity costs. The dynamic transmission model is very well suited to be extended to a cost-effectiveness analysis, as impact can be calculated in terms of infections averted or DALYs averted, which are common outcomes in cost-effectiveness calculations.

Although I found that broadly TasP had a higher impact than PrEP in the study, as well as in long-term scenarios, this did not assess whether TasP represents the most cost-effective use of available resources. Participation in the PrEP and TasP arms of the study incurs costs both to the providers from diagnostics, drugs and healthcare worker salaries, and to the patients for travel and the hours away from work. In collaboration with a colleague who estimated these costs, including initiation and ongoing annual costs, the cost-effectiveness of the study is currently being analysed. The results from this analysis could contribute further to any policy implications, particularly if cost-effectiveness for long-acting PrEP and other more recent interventions are estimated.

There is also an ongoing implementation study for PrEP among men who have sex with men (MSM) in Cotonou [485], who were recruited in the second half of 2020. My model described only the heterosexual population, but it could be extended to include MSM to estimate the impact of this study or the scale-up of PrEP among this group of high-risk males. Previous clinical trials have found much higher efficacy for oral PrEP among MSM compared to women [486, 487], due to better adherence, and higher concentration of active tenofovir in rectal tissues compared to vaginal tissues [488]. The results of the implementation study in Benin, when available, could parameterise the model for likely PrEP coverage and adherence for MSM, with which I could estimate its impact.

Finally, the same model-based phylodynamic analysis described in Chapter 5 could be applied to the HIV epidemic of Cotonou. Before enrolling onto the TasP arm of the study in Cotonou, the HIV-infected FSW were sequenced to check for any antiretroviral resistance [169]. This yielded around 80 sequences among this key population. There are further sequences publicly available from Cotonou on the Los Alamos National Laboratory HIV sequence database from the mid-1990s until 2020 [489]. An analysis of the HIV epidemic in Cotonou could be performed, providing an alternative estimate of the epidemic trajectory. Comparing this with the results from Chapter 3 would be excellent independent corroboration of the HIV epidemiology in Cotonou, and could be used to compare the applicability of phylodynamics and dynamic transmission modelling in the same context.

7.4 Implications for controlling COVID-19 epidemics

7.4.1 Impact of non-pharmaceutical interventions

Chapter 5 and Chapter 6 estimated the impact of non-pharmaceutical interventions (NPIs) in Weifang and England respectively. Both analyses broadly suggested that the restrictions imposed in each setting reduced viral transmission.

In Weifang, the majority of the confirmed COVID-19 cases identified (and the sequences) were sampled after the imposition of a raft of restrictions, which included public health messaging, establishing phone hotlines, encouraging home isolation for recent visitors from Wuhan, optimising triage of suspected cases in hospitals, travel restrictions, extending school closures, and establishing fever clinics for consultation and diagnosis [365]. All of these restrictions came into effect between 23-27th January 2020; the last genomic sample was taken on 10th February 2020.

Although a range of estimates of R_0 for SARS-CoV-2 have been made from Chinese data alone, with some estimating it to be below 2 [490] and some above 6 [491], one meta-analysis pooled estimate was 3.32 (95% CI: 2.81-3.82) [379]. It is important to note that although I set a prior R_0 of around 4, sampling with a constant likelihood ended up with an unrealistic posterior upper bound (over 10) for R_0. When fitting to genomic data, my phylodynamic analysis estimated that R_0 was 3.4 (95% HPD: 2.1-5.2). This estimate therefore agrees well with others in the literature for the Chinese epidemic.

I estimated that $R(t)$ fell below one on 4th February 2020 in Weifang, soon after the restrictions came into force. By mid-February, no more confirmed cases appeared in Weifang according to the local CDC. I further estimate that the true number of cases is only around 10 times more than what was confirmed by the health authorities, meaning there was no significant unseen burden. Given that R_0 was significantly above 1, and that there was no evidence of depletion of susceptible people, it is reasonable to therefore conclude that COVID-19 was fully controlled by these NPIs. The simultaneous applications of different restrictions mean that their individual impacts are impossible to disentangle. It is therefore unclear from this analysis which intervention, if any, could be removed while still maintaining control.

In England, there was already significant spread of SARS-CoV-2 in the community prior to the imposition of severe restrictions: by the time lockdown began on 23rd March 2020, 5,585 cases and 331 deaths had been recorded in England [386], and they continued to rise for several weeks. As of August 2021, there has never been a day of zero COVID-19 cases in England since the beginning of the epidemic; it is therefore impossible to conclude that the

NPIs eliminated viral transmission to the same extent as in Weifang. However, the results from the four case studies of Chapter 6 do have implications for COVID-19 control.

Case studies A and C investigated spring 2020 and winter 2020-2021 in England respectively. In both cases, there was a negative association between the stringency of restrictions (OxCGRT) and viral growth rates, i.e. viral transmission was slower when stronger NPIs were in place. As discussed in Chapter 6, as well as later in this chapter (section 'Broader themes'), an association does not demonstrate any patterns of causality. However, when introducing delays in the OxCGRT index time series, the strongest associations appear when viral transmission is compared to NPIs several days earlier. There is therefore more evidence that increasing stringency in NPI will lead to slower viral transmission, and less evidence for the opposite causal link. This has considerable implications for COVID-19 policy, given that severe restrictions can be justified only if there is evidence that they successfully reduce transmission.

One previous study estimated the impact of the tier system in England on transmission between the first and second lockdowns [433]. They estimated $R(t)$ at each Lower Tier Local Authority (LTLA) level, before and after the implementation of the restrictions, by fitting a semi-mechanistic Bayesian model to observed cases, deaths and serology data [492]. They observed a decline in $R(t)$ from mean 1.3 to 1.1, but did not attribute all reduction to the tier system. Using a hierarchical Bayesian model, they found that areas with more stringent restrictions saw a greater reduction in viral transmission. The findings from my analysis, when testing associations with the OxCGRT indices, are therefore consistent with this study.

Several other studies concluded that reproduction number could not be brought below 1 in England by any restriction other than full lockdown [70, 68, 69, 232, 38]. It is difficult to compare those results with those from my analysis. Firstly, my analysis cannot distinguish the effects of individual interventions when investigating the OxCGRT index, as the fitting process evaluates the association between changing stringency and growth rates across all timepoints, and summarises the association as a scalar. In case study C, I did explore lockdowns 2 and 3 as covariates, but as explained in Chapter 6 Section 2.3.5, given constraints in the data, I was forced to compare association between the 0th order (undifferentiated) covariate with second order changes in N_e (i.e. rate of change of growth rate). This is an association which is very difficult to interpret, although the analysis did conclude that lockdown 3 was more effective than lockdown 2, which is consistent with the common understanding that it was more stringent [455].

Finally, the results in my thesis regarding mobility were generally inconclusive. Although there were instances where a significant result appeared, the Google mobility data set had

problems in interpretation. There were issues with auto-correlation between variables: more movement in one area entails less movement in another. As a result, if movement of a certain type was found to be negatively associated with growth rates, it is possible that this only reflects a hidden positive association with movement of another type. For example, increased park mobility may not be itself associated with reduced viral transmission, but instead associated with reduced household mobility (itself associated with transmission), as people cannot be in two places at once. Teasing apart these effects is very problematic. Another issue is that the degree of smoothing I applied to the noisy mobility data had a drastic effect on the results, sometimes changing the direction of the associations. For these reasons, the implications of the mobility analyses are modest. It is important to re-iterate that the same sensitivity analyses were applied to the OxCGRT indices, but made no qualitative difference to the results, adding reliability to the association between stringency and reduced viral transmission.

7.4.2 Real-time surveillance and variants of concern

In the UK, the COG-UK consortium enabled high-throughput sequencing of SARS-CoV-2 sequences. By the winter period of 2020, approximately random community sampling was in place, minimising risk of sampling biases. Further, the time it took to sample patients, transport to sequencing hubs, sequence, clean and align the data had declined to 12 days. On the modelling side, improvements to the programming, as well as high-powered computing, enabled very quick time-scaling of trees and fitting the *mlesky* model. As a result, the difference in fitness of lineages could be estimated in a few days, using thousands of sequences.

This rapid turnaround of data and analyses enabled phylodynamic analysis to contribute to real-time surveillance. Indeed, the analysis of growth rates by variant (case study D) was an example of this. When the first version of this analysis was performed in December 2020, published as an Imperial College report [493], lineage B.1.1.7 had emerged from Kent, was quickly expanding in South East and East of England, and was appearing in samples throughout the country. I estimated N_e of B.1.1.7 through time for Kent, Essex, Norfolk and Greater London, all of which were consistent with rapid exponential growth. This analysis, along with several others featured in the Imperial College report and PHE reports [410], contributed to the decision by the UK government to increase restrictions in South East England on 20th December, and later to impose a third lockdown on 5th January 2021 [418].

A few weeks later, after receiving matched samples of other lineages by location and week, I performed a comparative analysis which estimated a substantial fitness increase for B.1.1.7 compared to previously circulating variants, strongly suggesting its rapid dominance was selective. Other modelling and statistical analyses corroborated this conclusion [101], which superseded the null hypothesis that this lineage was more commonly associated with groups that had more social interactions but was not itself fitter (e.g. the founder effect). After estimating this increased fitness, it became the dominant lineage throughout Europe in March 2021 and the United States in April 2021 [412].

At the same time, I compared median estimated N_e of B.1.1.7 against the frequency of S-gene target failure (SGTF) at each epidemiological week. I found the association to be striking; the peak in estimated N_e and peak in SGTF occurred in the same week. SGTF had been used globally as a proxy for frequency of B.1.1.7, and this analysis provides strong support for such practise during the final months of 2020 and first months of 2021. It also represents an excellent demonstration of the utility and reliability of phylodynamic modelling, with estimates accurate to a precision of at least a week, corroborated by independent surveillance data.

In conclusion, efficient analytical pipelines, combined with timely and randomly sampled data, made phylodynamic analysis a good choice for fast-paced epidemic surveillance, although it should be considered as one tool among other epidemiological methods. Further reductions in genomic data turnaround could provide phylodynamics with even more improved timeliness.

7.4.2.1 Recommendations for sequencing efforts

The analyses in this thesis, as well as several others [101, 107, 494, 495], have shown the power and importance of leveraging genomic data for monitoring infectious disease epidemics such as SARS-CoV-2. As a result, several recommendations have been announced to support the collection of further sequence data [496], in lower and middle income countries and in high income countries. In April 2021, the US government announced 1.7 billion dollars to fight COVID-19 variants [497]. The Gates Foundation called for increased funding for Global South countries to aid their sequencing efforts [498]. The European Commission is aiming for 5-10% sequencing coverage of positive tests [499].

The extensive analyses possible on SARS-CoV-2 sequence data, including those in this thesis, demonstrate the power of this data set. As such, the random samples from the community, with coverage proportionate to disease burden, as discussed in Section 6.2.2, are

recommended protocols for any increased SARS-CoV-2 sequence monitoring going forward for nations of any income level. For real-time monitoring of growth rates and variants of concern, extra care must be exercised to minimise the delay between sample collection and sharing with collaborators. The analysis performed in Leicester in this thesis demonstrates that moderate to high coverage of cases in local outbreaks is more likely to yield detectable peaks in COVID-19 case counts.

7.4.3 Future work

The model-based phylodynamic analysis (*PhyDyn*) applied to Weifang could be extended. The same modelling framework can be used in many locations, as there are whole genome sequences of SARS-CoV-2 available from across the world [179]. Further, application of this model to a location which also has seroprevalence data could strengthen the analysis. Estimated COVID-19 trajectories when fitted solely to genomic data could be compared to seroprevalence through time, indicating the reliability of using solely the genomic data. Next, combining the likelihood from the structured coalescent with the likelihood from seroprevalence would enable using both data sets, refining the estimates of reproduction number and epidemic trajectory.

There may be complexities in any future analysis incorporating seroprevalence data in the likelihood, if examining a period with a known variant of concern (VOC) replacing a co-circulating one due to a significantly higher transmission rate. Such a situation may require splitting the analysis by lineage, such as in my analysis in case study D in Chapter 6 (lineages B.1.1.7, B.1.177). This would however require that seroprevalence data are differentiated by lineage, which would be a very rare data set. Therefore, incorporating seroprevalence data into the *PhyDyn* framework would likely work best for the first few months of the COVID-19 epidemic, where little evidence of a VOC with a significantly faster transmission rate was identified.

Alternatively, the *mlesky* analysis performed on SARS-CoV-2 in England could be generalised to any emerging variants of concern. Since discovering the increased fitness of B.1.1.7 compared to previously circulating variants, it became the dominant lineage in England, throughout Europe and elsewhere [412]. As with any virus, however, its evolution continued. Another variant of concern emerged from India, named the “Delta variant” or B.617.2, which has an estimated increased transmissibility of around 50% compared to B.1.1.7 [500–502]. By July 2021, this variant became dominant in India, the United Kingdom, and various countries around the world [412]. It is clear that novel variants will continue to emerge, and

that global genomic surveillance may identify them as they do. The *mlesky* analysis could be used on any emerging lineage under investigation, to explore transmissibility relative to the dominant lineages in the area. If this process is automated, it could provide useful early warning signs of higher-than-expected burden, and policy could be adjusted if necessary.

The impact of pharmaceutical interventions could also be explored within the same phylodynamic framework. The analyses described in Chapter 6 described periods where vaccine coverage was extremely low or non-existent. Since then, vaccine uptake in the United Kingdom has been very successful; as of 4th August 2021, over 46 million people had received one dose of a COVID-19 vaccine, and over 38 million had received two [431]. By age, this translates to a coverage (of at least one dose) above 60% for adults 18-24 years old, and around 95% of adults above 70 years. Between 1st February 2021 and 1st June 2021, a typical day saw 400,000-500,000 doses administered. This rise in vaccine coverage could be modelled as a covariate and implemented in *mlesky*, either in a univariate analysis or in combination with an OxCGRT index as multivariate. Such an analysis could estimate any association between vaccine coverage and viral growth rates. The alignment could be restricted to B.1.1.7 to mitigate complications in transmission rates with the emergence B.1.617.2. A reduction in transmission associated with vaccine uptake would support the effectiveness of the strategy of mass vaccination in the United Kingdom.

Finally, other modelling techniques could be employed to further investigate the SARS-CoV-2 epidemic. One is example is the family of agent-based models which, as discussed in Chapter 1, can easily introduce stochasticity into the system. This may be of great importance when modelling emergent variants of concern: even a variant with a large transmission advantage may go extinct due to random stochastic factors. Agent-based models can also contain contact networks, where rules-based individuals ('agents') can interact between each other to simulate events such as viral transmission [503, 504]. An important difference between such models and compartmental ones, is that individuals of the same class can randomly draw traits and behavioural parameters from distributions, and can therefore all behave slightly differently. The transmission of SARS-CoV-2 can be simulated between individuals, parameterised with various sources such as the OxCGRT and Google mobility data sets [214, 427], or even more detailed data on individual behaviour in real-life settings such as contact patterns [505], face mask usage [506], and hand washing [507]. By leveraging behavioural, epidemiological and viral evolution data in one system, agent-based models could maximise the abundance of data available in certain localities, and make further important inferences on viral epidemiology, and thereby implications for viral control.

7.5 Broader themes

In this section I bring together all analyses in the thesis to consider broader themes. I discuss the strengths and limitations of each modelling framework, and their applicability to certain epidemiological contexts. Finally, I explore the implications of this work for approaching modelling questions.

7.5.1 Taking advantage of the available data

As with every modelling analysis, the quality of the output is limited by the quality of the data used for parameterisation and fitting. The analyses carried out in this thesis exploit very different data sets, varying in type (e.g., sexual behaviour questionnaire vs. whole genomes), abundance (around 25 sexual behavioural surveys vs 600,000 sequences), sampling strategy (contact tracing vs. random), demographic (FSW vs. low-risk groups), country (England vs. Benin), setting (hospital vs. community), stratification (individual vs. aggregated), and study design (cross-sectional vs. longitudinal). The availability of data in the context of the research question must inform the choice of model.

Indeed, the models I employed to answer the specific research questions in each chapter were well chosen to maximise the utility of the data. For instance, after summarising the available demographic, sexual behaviour, epidemiological and intervention data in Cotonou (Chapter 2), it was clear that a dynamic transmission model was one technique that could combine all of them. As an example, from integrated bio-behavioural surveys (IBBS), I extracted estimates of annual partnerships and condom use. Combining these with biological parameters taken from the literature, I calculated per-capita force of infection at each time step. Simultaneously, the model incorporated migration, natural history and engagement with HIV care which were parameterised from other sources. Finally, the model was calibrated by comparing epidemic trajectories with demographic information (from censuses and other studies), HIV prevalence estimates (also extracted from the IBBS) and ART coverage from government databases. The effect is a synergistic combination of numerous data sources, providing a thoroughly reliable representation of the HIV epidemic in Cotonou. Moreover, these data sources are independent and are therefore unlikely to be subject to the same biases in sampling.

For the non-parametric phylodynamic analyses, despite wholly different contexts and data sources, the principle remains the same. The non-parametric phylodynamic analysis of the SARS-CoV-2 epidemic in England, presented in Chapter 6, used thousands of whole

genome sequences to generate phylogenetic trees for model fitting [404]. The epidemic trajectories were cross validated against confirmed cases retrieved from widespread testing [386], allowing for these two diverse data sets to work together. Further, the effective population size of lineage B.1.1.7 was compared to S-gene target failure counts (Chapter 6 case study D) [101], a wholly different proxy for lineage frequency. There is great reliability in results from a model fitted to genomic data being corroborated by independent surveillance of PCR test failures. Lineage frequency through time was estimated with excellent precision (at least a week), demonstrating the power of phylodynamic inference when exploiting a high-quality genomic data set, and should be recommended when such data becomes available.

On the other hand, the analysis in Weifang (Chapter 5) did not benefit from a large data set. Although the model was parameterised using natural history data (e.g. generation time), it was fitted to only a handful of whole genome sequences. Official case counts were compared to model estimates, but because high-throughput COVID-19 testing was unavailable, and case ascertainment was poor so early in the pandemic, confirmed case counts were not a reliable measure of true burden. In fact, one of the motivations of the phylodynamic analysis was to evaluate the degree of underestimation in these case counts. As such, this analysis did not exploit varied data for fitting and cross validation.

In principle however, given that I used an explicit parametric mechanistic model, other data sources could be used for fitting. Indeed, an extension to *PhyDyn* has been developed (unpublished) which incorporates other information into the likelihood function, such as seroprevalence data at multiple timepoints. At the time of analysis however, such options were unavailable, and no useful independent data sources exist for Weifang. Nonetheless, the phylodynamic model analysis was ideal for capitalising on the only reliable data source at the time, and sampling from the prior indicated that the 20 sequences were very informative.

It is worth noting that the modelling frameworks chosen for each specific context in this thesis were not the only ones possible. Indeed, phylodynamic analysis has been applied to HIV epidemics [508–510], and compartmental models have been applied to outbreaks of SARS-CoV-2 [68]. However, the analyses in this thesis were justified for capitalising on the highest quality data sets available in each context, from the long-term cross-sectional HIV survey data in Cotonou, to the vast quantities of community sampled SARS-CoV-2 whole genome sequences in England.

A related implication of this work is the influence modelling can have for effective data collection. In Chapter 4, I estimated that disruptions to commercial condom use had the biggest impact in terms of increased HIV infections in Cotonou, Benin. I concluded that it

is therefore important to concentrate efforts to ensure that such disruptions to condom use do not happen. A broader implication can also be drawn: my model analysis elucidated important areas for urgent data collection. It is recommended that condom use, particularly among pFSW with their clients, is monitored during the COVID-19 period and during other periods of crisis. If reductions in condom use are observed, whether due to lack of supply or economic hardship, then the appropriate response to aid the issue can be identified, and an unwanted increased in HIV infections can be prevented. More generally, modelling can be helpful for identifying important parameters for viral epidemics, and can therefore be leveraged to focus efforts on effective data collection.

7.5.2 Overfitting and bias

A trade-off can however emerge in attempting to capitalise on large amounts of varied data. As mentioned, the dynamic transmission model in Cotonou was able to incorporate demographic, natural history, sexual behaviour and intervention data from numerous sources. Several risk groups were therefore modelled, each representing multiple stages of infection and states of care, amounting to a total of 207 individual compartments. The number of parameters, consequently, climbed into another order of magnitude. Although many were fixed or correlated to each other, just under 100 parameters were estimated, while fitting to far fewer (31) independent observations. The estimated epidemic trajectory is likely to be highly affected by each observation used in fitting. There could therefore be greatly different impacts estimated if new observations are added, or if the ones already used are erroneous. This risk of overfitting was lessened by cross validation and sensitivity analyses, but regardless remains a limitation of such a large model. In summary, there is a trade-off between the inclusion of as many data types as possible, and expanding the model's dimensionality and thereby the risk of overfitting.

A similar limitation exists for the Weifang analysis, which relied on fitting to only 20 whole genome samples . Although many parameters were fixed and the number of compartments was limited, the small data set makes the analysis prone to overfitting. This would mean that estimated trajectories may be significantly different if a different group of 20 patients were sampled in Weifang. Fortunately, the results were robust to sensitivity analysis, and uncertainty was significantly reduced than when sampling from the prior distributions. The predictive power of the model, however, would have been much improved if more genomic samples were available.

The non-parametric analyses in England did not suffer from a similar problem. Firstly, by definition, being non-parametric meant that the model function was not defined a priori, and so the number of parameters cannot be compared to the number of observations. Secondly, tens of thousands of sequences were available for most of the case studies. As such, it became less a question of finding enough data to fit the model, but rather one of choosing the right data via proper sub-sampling. Indeed, as discussed in Chapter 6, the estimated trajectories early in the epidemic (e.g. January/February 2020 in England) relied on 'pillar 1' samples which were not randomly sampled from the community, but taken from cases in hospital. Such trajectories are likely therefore to be biased, forming a limitation of the data and of the model. Community-based random sampling, as performed later in the epidemic, is likely to provide less biased and more reliable epidemic trajectories from phylodynamic analysis.

7.5.3 Discerning complex trajectories

In this thesis, I have mostly compared two classes of analysis: transmission modelling vs. phylodynamic modelling. It can however be useful to consider an alternative classification: parametric vs. non-parametric modelling.

Parametric modelling, by definition, requires the a priori definition of the model function. For example, the Weifang analysis assumed a mechanistic model comprising S, E, I, J, R compartments, and the parameters describing rates of movement between them were already pre-specified. This results in constraints in the shape of epidemic trajectory. Once an epidemic has been seeded, there will be a rise and subsequent fall in cases from the depletion of susceptibles, assuming R_0 is above one. Further increases in cases are impossible with this structure, regardless of parameter values. Such a model analysis is therefore only useful or applicable to a single wave of COVID-19; a second wave can only be captured by changing the model, for instance including further compartments, assuming waning immunity (transition from R to S compartment) or a time-varying transmission rate. It is important to consider the mechanistic constraints of this simple compartmental model when faced with a research question. For the purposes of Chapter 5, where I focused only on the initial spread in Weifang, the model was appropriate.

The dynamic transmission model in Cotonou does not face the same limitations, because it is vastly more complex. A second wave of HIV could be captured by the model in many ways. Indeed, the analysis exploring the impact of COVID on HIV epidemiology provides one such example (Chapter 4). It is possible to model a scenario whereby disruptions were so severe that condom use and viral suppression drop to the point where R_0 increases above one, and

the historical downward trend in HIV prevalence reverses. Having such a complex structure, in order to achieve complex trajectories, was instrumental for the answering the aims of Chapter 4. However, as discussed in the previous section, increased model dimensionality is not without costs. For mechanistic models, the ability to discern complex trajectories is therefore in a trade-off with the degree of overfitting.

The situation for non-parametric phylodynamic models is very different. As mentioned earlier, the model function is not defined a priori, and so the epidemic trajectory is not limited by these mechanistic constraints. Indeed, the shape of the effective population size (N_e) trajectory can be extremely noisy, as observed for the first estimates using *skyline* [91]. Because this is also seen as evidence of overfitting, reflecting noise instead of signal, various methods have been since developed to smooth these trajectories (see Chapter 1 for more details). These include *mlesky* and *skygrowth*, the two non-parametric methods used in this thesis. Even after smoothing, complex epidemic trajectories can still be discerned.

For example, in Leicester (Chapter 6 case study B), there were peaks in confirmed COVID-19 cases in mid-May and mid-June. The *mlesky* model, after fitting to sequence alignments from Leicester, estimated a peak in N_e 1-2 weeks before each peak in confirmed cases. This is consistent with delays in incubation and diagnostic testing. Similarly, in England in winter 2020-2021 (case study C), there were peaks in confirmed cases in early November 2020 and early January 2021. Once again, the *mlesky* model estimated peaks in N_e in late October and late December 2020, a short while before the peak in cases. The COVID-19 epidemic is fast moving and complex, very sensitive to sharp changes in public behaviours, producing a complex epidemic trajectory. The non-parametric design is therefore flexible to take many functional forms, capturing the true peaks and troughs in incidence, without incurring the same costs in overfitting.

7.5.4 Considering natural history

7.5.4.1 Speed of outbreak

The speed of an epidemic is very important when considering how to model it. SARS-CoV-2 is a rapidly spreading virus: early estimates of generation time had a mean of 3.5 days assuming only positive serial intervals, and 2.57 days when including negative ones [186]. As a result, the speed of everything else is accordingly fast, from emerging data on confirmed cases [386, 175], sequences [404, 179], hospitalisations and deaths [196], to the urgent implementation of strict policy decisions [214].

Therefore, model analyses for the purposes of surveillance and informing policy must likewise be fast yet accurate, two important factors which are themselves in trade-off; waiting for more reliable data may improve accuracy, but may render the results no longer useful in a rapidly changing epidemic. As discussed in a previous section, the collection of approximately random community samples by COG-UK provided an excellent data set to perform fast phylodynamic analysis, which proved useful in contributing to real-time epidemic surveillance and urgent policy decisions.

On the other hand, HIV moves at a much slower pace. The high-quality data used to parameterise and fit the dynamic transmission model, from HIV prevalence to sexual behaviour and interventions, spanned several decades. Indeed, even a single estimate of HIV prevalence, calculated from integrated bio-behavioural surveys, may take months or years to finalise. The research questions relevant to policy are accordingly less urgent; consistent condom use can take years for widespread adoption, and the scale-up of pharmaceutical HIV prevention interventions may take years to implement. In such a context, accuracy takes preference over speed; the slower pace enables fitting a complex dynamic transmission model to high quality data through time, which can estimate intervention impact in useful and countable outcomes for policy, such as infections averted and life-years gained.

7.5.4.2 Virus evolution

Related to the speed of outbreak is the rate of molecular evolution. SARS-CoV-2 is not known for a fast mutation rate [407], but its short generation time and widespread transmission leads to the rapid appearance of polymorphisms and lineages. This genetic variation is ultimately what phylodynamic models are fitted to, and is therefore necessary to estimate epidemic trajectories. If polymorphisms did not appear on the epidemiological timescale, i.e. if the majority of samples were genetically identical, then phylodynamic modelling would be a poor candidate for analysis. As it was, genetic variation was plentiful, and so phylodynamics became powerful.

The appearance of novel variants further supports the contribution of phylodynamic modelling of SARS-CoV-2. Variants are themselves determined through phylogenetic analysis; given that the tree is a requirement for phylodynamic analysis, it becomes a good candidate for exploring transmission separately by lineage. Indeed, Chapter 6 provides an excellent demonstration of the added fitness of the B.1.1.7 lineage compared to previously circulating ones. Although it is possible, in principle, to fit mechanistic models to case and death data

labelled by variant, it requires the linking of data sets which incurs further risks to timeliness and bias.

7.5.5 Estimating the impact of interventions depends on context

7.5.5.1 When the intervention is easy to describe

The overall aim of the thesis was to explore the impact of interventions. However, the interventions used to control viral epidemics are highly diverse not only in kind (antiretroviral drugs vs. lockdown), but also in the available data describing them. The PrEP/TasP study offered pharmaceutical items to real individuals, whose participation in the study was monitored over the two years; the actual numbers of initiations and dropouts were therefore available through time [169]. Further, the study benefitted from unique biological PrEP adherence data. Given that PrEP effectiveness is mixed among heterosexual women [284, 470, 161, 471], partially due to poor adherence and retention, it was vital that the model analysis captured this information to properly understand the impact of PrEP. Only certain analyses, such as the dynamic transmission model employed, could exploit such high resolution of detail from this data set.

By design, I had included various compartments to capture numbers of individuals on PrEP and ART. There were four PrEP compartments: high, intermediate and low adherence, and PrEP dropout. This model structure therefore could use the available high-resolution data describing the intervention itself.

The dynamic transmission model framework offers further benefits. Once the model was closely fitted to the PrEP/TasP study data, it represented the real-life scenario in Cotonou. To estimate impact, I compared outcomes from this scenario to another one, a counterfactual scenario, in which the PrEP/TasP study had never happened. The difference in outcomes between these two scenarios constituted the impact of the study, much like comparing the treatment and control arms in a clinical trial. Further, the model design was flexible, such that different scenarios could be simulated by changing few parameters, while holding others constant. The ability to compare scenarios for estimating impact is among the key strengths of this model framework.

It should be noted, however, that this is not without limitations. Notably, it relies on simulating a scenario which never happened, and thus could not be fitted to any data. This naturally adds some uncertainty: what would have happened if the PrEP/TasP study had not run? For example, after the study ended in 2017, ART eligibility was expanded to all HIV-infected

individuals [169], rather than only those with CD4 T-cell counts below 500 cells/mm^3. It is possible that a similar policy change would have occurred anyway, due to the popularity of test-and-treat [511]. It is however also possible that the good uptake, adherence and retention of women on TasP observed in the study provided the impetus for policymakers to expand eligibility in Benin.

Indeed, because of this uncertainty, I carried out a sensitivity analysis. On top of the main counterfactual scenario, where eligibility remained at <500 cells/mm^3, I simulated an alternative counterfactual, where it expanded to all HIV-infected individuals. There was a 30% decrease in the impact of TasP, which reflects the impact of the ART eligibility component (leading to more initiations); around 70% was therefore attributable to increased testing and general increased ART offering. There is therefore some uncertainty in the counterfactual scenarios and in the results, but sensitivity analyses can be performed to account for it.

Furthermore, as compartments represented groups of actual individuals, fitted to demographic data and HIV prevalence, the impact of interventions was measured in this analysis with easily interpretable outcomes, such as infections averted or life-years gained. Such outcomes facilitate comparisons of PrEP and TasP against other interventions, as well as providing the foundation for a cost-effectiveness analysis. These factors are useful for policymakers to decide upon the best interventions with limited resources. The dynamic transmission model, in capturing the well-described high resolution data of the PrEP/TasP study, and able to calculate outcomes useful for policy, was therefore well suited to estimate the impact of this particular intervention, and to contribute to the aims of this thesis.

7.5.5.2 When the intervention is hard to describe

In contrast, other interventions I aimed to analyse were more abstruse. For instance, most interventions for COVID-19 have been, to date, attempts to limit social contacts, using quarantines, travel restrictions, border closures, lockdowns, among many others. Data that describe these interventions are usually much harder to interpret than the PrEP/TasP study. The Oxford COVID-19 Government Response Tracker (OxCGRT) was one attempt to quantify the stringency of government policy through time [214]. However, such indices do not measure what actual people do, and an intuitive public response does not necessarily follow a policy decision, and if it does, it may arrive with a delay. Other proxies of behaviour are available, such as Google mobility data [435], but assuming that mobile phone data will perfectly reflect social contact networks has its limitations [512, 513], so implementing

this data in a force of infection function for an SEIR-like transmission model is arguably unreliable. Moreover, changing mobility is not itself an intervention, but rather a potential consequence of one.

NPIs are therefore much harder to explicitly model compared to the PrEP/TasP study. In such cases, one must perform analyses which assess statistical associations of NPIs with transmission, such as those performed in Chapter 6 of this thesis.

In previous studies of phylodynamic analysis, N_e has been shown to correlate poorly to the number of infected patients with a simple linear function [104]. Clearly defined outcomes for measuring intervention impact, such as infections averted or life-years gained, are therefore very difficult to measure. However, the rate of change of N_e through time (first order differential), i.e. its growth rate, can be easily calculated, and tested for association with NPIs. Although the direction of causality cannot be easily determined, as discussed in the relevant chapter, I found the strongest associations when comparing growth rates several days after policy changes, but not vice versa. This adds weight to the argument that restrictions reduced viral transmission.

It is important to note that this model framework is unlike the dynamic transmission model, where the PrEP/TasP study scenario could be explicitly modelled, and compared to a counterfactual scenario. Instead, the non-parametric phylodynamic model framework could only test statistical associations between growth rates and NPIs for one scenario: the one that really happened. This is indeed less evidentially powerful than an analysis comparing two scenarios. However, given the imprecise measures available for NPIs, they could not be explicitly represented in terms of compartments. Further, given the issues in testing and case ascertainment, epidemic trajectory is arguably better estimated using the abundance of high-quality genomic data instead of cases or deaths.

Moreover, the key research aim was to understand the impact of interventions on viral transmission. Although this analysis cannot provide impact in terms of total numbers of infections, it can provide evidence as to whether NPIs affect transmission, and if so, in which direction. Within the context of a chaotic epidemiological and political situation, where decisions must be made quickly, such evidence is of vital importance, and non-parametric phylodynamic modelling proved to be an excellent and appropriate tool for answering such a difficult research question.

Finally, phylodynamic modelling can give unique insights for answering the question in different ways. The Weifang analysis, for example, involved co-estimation of epidemiological parameters and the phylogenetic tree. While the trajectory of reproduction number through time shed light on NPIs implemented in the city, the phylogeny suggested that most of

the transmission was internal, rather than imported. Travel restrictions and isolation of individuals from Wuhan was therefore likely to be quite effective at limiting the number of importations to Weifang.

7.6 Final conclusions

The principal objective of this thesis was to estimate the impact of interventions on viral epidemics using epidemiological and evolutionary models. The various chapters have answered this in different locations, for different diseases and with different models. Each analysis contributed to the academic literature and had important implications for local and international health policy.

Transmission modelling of HIV in Cotonou, Benin, suggested that implementing TasP among female sex workers was more impactful than daily oral PrEP, due to observed adherence and retention among these women. The impact of PrEP may be improved if adherence could be improved, or if long-acting PrEP was available. However, the mobility of FSW in this context may cause issues with long-term retention to any PrEP programmes.

When exploring COVID-19 related disruptions to already-present HIV interventions, maintaining condom use in commercial sex was found to be the most important for preventing increases in HIV infections, with impact disproportionately affecting key populations. This analysis supports the widespread policy of focusing HIV interventions on key populations, but also underscores the importance of continued maintenance and promotion of condom use.

Various phylodynamic modelling analyses in this thesis suggest that non-pharmaceutical interventions have contributed to control of the SARS-CoV-2 epidemic. I concluded that in Weifang, China, the restrictions were successful in limiting viral transmission and introductions into the city. In England, I found that the increased stringency of interventions was most associated with reduced transmission several days after they are implemented. I also estimated an increased fitness of the then-emerging B.1.1.7 lineage, which later became dominant in the UK and other countries. The analyses in this thesis demonstrate the utility of phylodynamic modelling for real-time surveillance where the appropriate data is available.

Common themes emerge through the thesis. Each analysis was tailored to the context: it consisted in fitting models to the best available data, leading to a careful representation of the underlying viral epidemic, and the estimation of key epidemiological parameters. Each

analysis had to balance several trade-offs, notably between the benefits of complexity and the costs of overfitting and speed.

The work in this thesis shows that the nature of the intervention under investigation is an important consideration. Where an intervention is well described with high-resolution data, it can be explicitly modelled, and its epidemiological impact can be evaluated by comparing scenarios. However, interventions may be pronounced yet very poorly described, or the real-life consequences are impossible to measure. In such cases, proxies for the intervention must be used, and their associations with epidemiological parameters can only be estimated statistically. The nature of the intervention in question therefore determines the most suitable method for evaluating its impact.

In conclusion, on top of the important policy implications of each individual analysis, this thesis explores more general principles regarding the strengths of different models, the quality of data used, and the interpretation of impact. It is paramount that any results are understood in the context of these factors, particularly with the increasing prominence of mathematical models for informing viral epidemic control interventions.

Bibliography

[1] William Heaton Hamer. *Epidemic disease in England: the evidence of variability and of persistency of type*. Bedford Press, 1906.

[2] Matt J Keeling and Pejman Rohani. *Modeling infectious diseases in humans and animals*. Princeton university press, 2011.

[3] Philip S Brachman. Infectious diseases—past, present, and future, 2003.

[4] Theo Vos, Stephen S Lim, Cristiana Abbafati, Kaja M Abbas, Mohammad Abbasi, Mitra Abbasifard, Mohsen Abbasi-Kangevari, Hedayat Abbastabar, Foad Abd-Allah, Ahmed Abdelalim, et al. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *The Lancet*, 396(10258):1204–1222, 2020.

[5] Lauren M Childs, Nadia N Abuelezam, Christopher Dye, Sunetra Gupta, Megan B Murray, Brian G Williams, and Caroline O Buckee. Modelling challenges in context: lessons from malaria, HIV, and tuberculosis. *Epidemics*, 10:102–107, 2015.

[6] Emma S McBryde, Michael T Meehan, Oyelola A Adegboye, Adeshina I Adekunle, Jamie M Caldwell, Anton Pak, Diana P Rojas, Bridget M Williams, and James M Trauer. Role of modelling in COVID-19 policy development. *Paediatric respiratory reviews*, 35:57–60, 2020.

[7] OECD Opportunity. OECD Science, Technology and Innovation Outlook 2021, 2021.

[8] Jatin Machhi, Jonathan Herskovitz, Ahmed M Senan, Debashis Dutta, Barnali Nath, Maxim D Oleynikov, Wilson R Blomberg, Douglas D Meigs, Mahmudul Hasan, Milankumar Patel, et al. The natural history, pathobiology, and clinical manifestations of SARS-CoV-2 infections. *Journal of Neuroimmune Pharmacology*, pages 1–28, 2020.

[9] Hassan ElHawary, Ali Salimi, Nermin Diab, and Lee Smith. Bibliometric analysis of early COVID-19 research: the top 50 cited papers. *Infectious Diseases: Research and Treatment*, 13:1178633720962935, 2020.

[10] Shir Aviv-Reuven and Ariel Rosenfeld. Publication patterns' changes due to the COVID-19 pandemic: A longitudinal and short-term scientometric analysis. *Scientometrics*, pages 1–24, 2021.

[11] gov.uk, 2021. URL https://www.gov.uk/government/collections/scientific-evidence-supporting-the-government-response-to-coronavirus-covid-19. Accessed: 2021-08-11.

[12] Caitlin Rivers, Simon Pollett, and Cecile Viboud. The opportunities and challenges of an Ebola modeling research coordination group. *PLoS neglected tropical diseases*, 14 (7):e0008158, 2020.

[13] Caitlin Rivers, Jean-Paul Chretien, Steven Riley, Julie A Pavlin, Alexandra Woodward, David Brett-Major, Irina Maljkovic Berry, Lindsay Morton, Richard G Jarman, Matthew Biggerstaff, et al. Using "outbreak science" to strengthen the use of models during epidemics. *Nature communications*, 10(1):1–3, 2019.

[14] Robert M Christley, Maggie Mort, Brian Wynne, Jonathan M Wastling, A Louise Heathwaite, Roger Pickup, Zoë Austin, and Sophia M Latham. "Wrong, but useful": negotiating uncertainty in infectious disease modelling. *PLoS One*, 8(10):e76277, 2013.

[15] Maria D Van Kerkhove and Neil M Ferguson. Epidemic and intervention modelling: a scientific rationale for policy decisions? Lessons from the 2009 influenza pandemic. *Bulletin of the World Health Organization*, 90:306–310, 2012.

[16] Stephan Eubank, I Eckstrand, B Lewis, Srinivasan Venkatramanan, Madhav Marathe, and CL Barrett. Commentary on Ferguson, et al.,"Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand". *Bulletin of mathematical biology*, 82(4):1–7, 2020.

[17] Herbert W Hethcote. The mathematics of infectious diseases. *SIAM review*, 42(4): 599–653, 2000.

[18] Kate M Mitchell, Aurélia Lépine, Fern Terris-Prestholt, Kwasi Torpey, Hadiza Khamofu, Morenike O Folayan, Jonah Musa, James Anenih, Atiene S Sagay, Emmanuel Alhassan, et al. Modelling the impact and cost-effectiveness of combination prevention amongst HIV serodiscordant couples in Nigeria. *AIDS (London, England)*, 29(15):2035, 2015.

[19] Ide Cremin, Ramzi Alsallaq, Mark Dybul, Peter Piot, Geoffrey Garnett, and Timothy B Hallett. The new role of antiretrovirals in combination HIV prevention: a mathematical modelling analysis. *AIDS*, 27(3):447–458, 2013.

[20] Jack J Olney, Paula Braitstein, Jeffrey W Eaton, Edwin Sang, Monicah Nyambura, Sylvester Kimaiyo, Ellen McRobie, Joseph W Hogan, and Timothy B Hallett. Evaluating strategies to improve HIV care outcomes in Kenya: a modelling study. *The Lancet HIV*, 3(12):e592–e600, 2016.

[21] Rebecca F Baggaley, Michael A Irvine, Werner Leber, Valentina Cambiano, Jose Figueroa, Heather McMullen, Jane Anderson, Andreia C Santos, Fern Terris-Prestholt, Alec Miners, et al. Cost-effectiveness of screening for HIV in primary care: a health economics modelling analysis. *The Lancet HIV*, 4(10):e465–e474, 2017.

[22] Jessica B McGillen, Sarah-Jane Anderson, Mark R Dybul, and Timothy B Hallett. Optimum resource allocation to reduce HIV incidence across sub-Saharan Africa: a mathematical modelling study. *The Lancet HIV*, 3(9):e441–e448, 2016.

[23] Reuben M Granich, Charles F Gilks, Christopher Dye, Kevin M De Cock, and Brian G Williams. Universal voluntary HIV testing with immediate antiretroviral therapy as a strategy for elimination of HIV transmission: a mathematical model. *The Lancet*, 373 (9657):48–57, 2009.

[24] Mirjam Kretzschmar. Disease modeling for public health: added value, challenges, and institutional constraints. *Journal of public health policy*, 41(1):39–51, 2020.

[25] Reuben Granich, Brian Williams, Julio Montaner, and José M Zuniga. 90-90-90 and ending AIDS: necessary and feasible. *The Lancet*, 390(10092):341–343, 2017.

[26] Dame Deirdre Hine, DBE FFPH FRCP. The 2009 Influenza Pandemic: An independent review of the UK response to the 2009 influenza pandemic, 2021. Accessed: 2021-08-11.

[27] parliament.uk. Memorandum submitted by the Health Protection Agency (SAGE 28), 2021. URL https://publications.parliament.uk/pa/cm201011/cmselect/cmsctech/498/498we11.htm. Accessed: 2021-08-11.

[28] Martin I Meltzer, Charisma Y Atkins, Scott Santibanez, Barbara Knust, Brett W Petersen, Elizabeth D Ervin, Stuart T Nichol, Inger K Damon, and Michael L Washington. Estimating the future number of cases in the Ebola epidemic–Liberia and Sierra Leone, 2014–2015, 2014.

[29] Anton Camacho, Adam Kucharski, Yvonne Aki-Sawyerr, Mark A White, Stefan Flasche, Marc Baguelin, Timothy Pollington, Julia R Carney, Rebecca Glover, Elizabeth Smout, et al. Temporal changes in Ebola transmission in Sierra Leone and implications for control requirements: a real-time modelling study. *PLoS currents*, 7, 2015.

[30] Natsuko Imai, Kalipso Chalkidou, Christl Donnelly, Steven Riley, Peter White, and Neil M Ferguson. Strengthening Epidemiological Research and Modelling for Epidemic Preparedness, 2021. URL https://figshare.com/articles/preprint/Strengthening_Epidemiological_Research_and_Modelling_for_Epidemic_Preparedness/14135633. Accessed: 2021-04-09.

[31] A Camacho, RM Eggo, N Goeyvaerts, A Vandebosch, R Mogg, S Funk, AJ Kucharski, CH Watson, T Vangeneugden, and WJ Edmunds. Real-time dynamic modelling for the design of a cluster-randomized phase 3 Ebola vaccine trial in Sierra Leone. *Vaccine*, 35(4):544–551, 2017.

[32] Adam J Kucharski, Rosalind M Eggo, Conall H Watson, Anton Camacho, Sebastian Funk, and W John Edmunds. Effectiveness of ring vaccination as control strategy for Ebola virus disease. *Emerging infectious diseases*, 22(1):105, 2016.

[33] Imai Natsuko, Dorigatti Ilaria, Cori Anne, R Steven, and M Neil. Report 1: Estimating the potential total number of novel Coronavirus cases in Wuhan City, China. *Imperial College London, London*, 2020.

[34] Natsuko Imai, Anne Cori, Ilaria Dorigatti, Marc Baguelin, Christl A Donnelly, Steven Riley, and Neil M Ferguson. Report 3: transmissibility of 2019-nCoV. *Imperial College London*, pages 1–6, 2020.

[35] Ilaria Dorigatti, Lucy Okell, Anne Cori, Natsuko Imai, Marc Baguelin, Sangeeta Bhatia, Adhiratha Boonyasiri, Zulma Cucunubá, Gina Cuomo-Dannenburg, Rich FitzJohn, et al. Report 4: severity of 2019-novel coronavirus (nCoV). *Imperial College London, London*, 2020.

[36] Erik Volz, Marc Baguelin, Sangeeta Bhatia, Adhiratha Boonyasiri, Anne Cori, Zulma Cucunubá, Gina Cuomo-Dannenburg, Christl A Donnelly, Ilaria Dorigatti, Rich FitzJohn, et al. Report 5: phylogenetic analysis of SARS-CoV-2. *Imperial College London*, 2020.

[37] gov.uk. Addendum to sixteenth SAGE meeting on Covid-19, 16th March 2020, 2021. URL https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/888784/S0384_Sixteenth_SAGE_meeting_on_Wuhan_Coronavirus__Covid-19__.pdf. Accessed: 2021-08-11.

[38] Neil Ferguson, Daniel Laydon, Gemma Nedjati-Gilani, Natsuko Imai, Kylie Ainslie, Marc Baguelin, Sangeeta Bhatia, Adhiratha Boonyasiri, Zulma Cucunubá, Gina Cuomo-Dannenburg, et al. Report 9: Impact of non-pharmaceutical interventions (NPIs) to reduce COVID19 mortality and healthcare demand. *Imperial College London*, 10(77482):491–497, 2020.

[39] gov.uk. Scientific evidence supporting the government response to coronavirus (COVID-19), 2021. URL https://www.gov.uk/government/collections/scientific-evidence-supporting-the-government-response-to-coronavirus-covid-19.

[40] gov.uk. Joint Committee on Vaccination and Immunisation, 2021. URL https://www.gov.uk/government/groups/joint-committee-on-vaccination-and-immunisation. Accessed: 2021-08-11.

[41] gov.uk. New and Emerging Respiratory Virus Threats Advisory Group, 2021. URL https://www.gov.uk/government/groups/new-and-emerging-respiratory-virus-threats-advisory-group. Accessed: 2021-08-11.

[42] New and Emerging Respiratory Virus Threats Advisory Group. NERVTAG meeting on SARS-CoV-2 variant under investigation VUI-202012/01, 2021. URL https://app.box.com/s/3lkcbxepqixkg4mv640dpvvg978ixjtf/file/756963730457. Accessed: 2021-08-11.

[43] Richard Pitman, David Fisman, Gregory S Zaric, Maarten Postma, Mirjam Kretzschmar, John Edmunds, Marc Brisson, ISPOR-SMDM Modeling Good Research Practices Task Force, et al. Dynamic transmission modeling: a report of the ISPOR-SMDM modeling good research practices task force-5. *Value in health*, 15(6):828–834, 2012.

[44] MJ Keeling and L Danon. Mathematical modelling of infectious diseases. *British medical bulletin*, 92(1), 2009.

[45] P Van den Driessche and James Watmough. Further notes on the basic reproduction number. In *Mathematical epidemiology*, pages 159–178. Springer, 2008.

[46] Samuel Alizon and Carsten Magnus. Modelling the course of an HIV infection: insights from ecology and evolution. *Viruses*, 4(10):1984–2013, 2012.

[47] Narat Punyacharoensin, William John Edmunds, Daniela De Angelis, Valerie Delpech, Graham Hart, Jonathan Elford, Alison Brown, Noel Gill, and Richard G White. Modelling the HIV epidemic among MSM in the United Kingdom: quantifying the contributions to HIV transmission to better inform prevention initiatives. *AIDS*, 29(3): 339–349, 2015.

[48] Mathieu Maheu-Giroux, Juan F Vesga, Souleymane Diabaté, Michel Alary, Stefan Baral, Daouda Diouf, Kouamé Abo, and Marie-Claude Boily. Changing dynamics of HIV transmission in Côte d'Ivoire: modeling who acquired and transmitted infections and estimating the impact of past HIV interventions (1976-2015). *Journal of acquired immune deficiency syndromes*, 75(5):517, 2017.

[49] Kamaldeen Okuneye and Abba B Gumel. Analysis of a temperature-and rainfall-dependent model for malaria transmission dynamics. *Mathematical biosciences*, 287: 72–92, 2017.

[50] John R Williams, Michel Alary, Catherine M Lowndes, Luc Béhanzin, Annie-Claude Labbé, Séverin Anagonou, Marguerite Ndour, Isaac Minani, Clément Ahoussinou, Djimon Marcel Zannou, et al. Positive impact of increases in condom use among female sex workers and clients in a medium HIV prevalence epidemic: modelling results from Project SIDA1/2/3 in Cotonou, Benin. *PLoS One*, 9(7):e102643, 2014.

[51] Kate M Mitchell, Brooke Hoots, Dobromir Dimitrov, Danielle German, Colin Flynn, Jason E Farley, Marcy Gelman, James P Hughes, Deborah Donnell, Adeola Adeyeye, et al. Improvements in the HIV care continuum needed to meaningfully reduce HIV incidence among men who have sex with men in Baltimore, US: a modelling study for HPTN 078. *Journal of the International AIDS Society*, 22(3):e25246, 2019.

[52] Romain Silhol, Stefan Baral, Anna L Bowring, Christinah Mukandavire, Iliassou M Njindam, Amrita Rao, Sheree Schwartz, Ubald Tamoufe, Serge C Billong, Oudou Njoya, et al. Quantifying the evolving contribution of HIV interventions and key populations to the HIV epidemic in Yaounde, Cameroon. *JAIDS Journal of Acquired Immune Deficiency Syndromes*, 86(4):396–405, 2021.

[53] Mingwang Shen, Yanni Xiao, and Libin Rong. Global stability of an infection-age structured HIV-1 model linking within-host and between-host dynamics. *Mathematical biosciences*, 263:37–50, 2015.

[54] Lythgoe, Katrina A and Fraser, Christophe. New insights into the evolutionary rate of HIV-1 at the within-host and epidemiological levels. *Proceedings of the Royal Society B: Biological Sciences*, 279(1741):3367–3375, 2012.

[55] Mideo, Nicole and Alizon, Samuel and Day, Troy. Linking within-and between-host dynamics in the evolutionary epidemiology of infectious diseases. *Trends in ecology & evolution*, 23(9):511–517, 2008.

[56] Britton, Tom and House, Thomas and Lloyd, Alun L and Mollison, Denis and Riley, Steven and Trapman, Pieter. Five challenges for stochastic epidemic models involving global transmission. *Epidemics*, 10:54–57, 2015.

[57] Lawrence F Shampine. *Numerical solution of ordinary differential equations*. Routledge, 2018.

[58] Eleanor Gouws, Peter J White, John Stover, and Tim Brown. Short term estimates of adult HIV incidence by mode of transmission: Kenya and Thailand as examples. *Sexually transmitted infections*, 82(suppl 3):iii51–iii55, 2006.

[59] Sharmistha Mishra, Michael Pickles, James F Blanchard, Stephen Moses, Zara Shubber, and Marie-Claude Boily. Validation of the modes of transmission model as a tool to prioritize HIV prevention targets: a comparative modelling analysis. *PLoS One*, 9 (7):e101690, 2014.

[60] Yoann Madec, Faroudy Boufassa, Kholoud Porter, Maria Prins, Caroline Sabin, Antonella d'Arminio Monforte, Pauli Amornkul, Barbara Bartmeyer, Mette Sannes, Alain Venet, et al. Natural history of HIV-control since seroconversion. *AIDS*, 27(15): 2451–2460, 2013.

[61] Sarah MacCarthy, Michael Hoffmann, Laura Ferguson, Amy Nunn, Risha Irvin, David Bangsberg, Sofia Gruskin, and Ines Dourado. The HIV care cascade: models, measures and moving forward. *Journal of the International AIDS Society*, 18(1): 19395, 2015.

[62] F Baryarama, JYT Mugisha, and LS Luboobi. A mathematical model for the dynamics of HIV/AIDS with gradual behaviour change. *Computational and Mathematical Methods in Medicine*, 7(1):15–26, 2006.

[63] Geoffrey P Garnett and Roy M Anderson. Balancing sexual partnership in an age and activity stratified model of HIV transmission in heterosexual populations. *Mathematical Medicine and Biology: A Journal of the IMA*, 11(3):161–192, 1994.

[64] Mathieu Maheu-Giroux, Juan F Vesga, Souleymane Diabaté, Michel Alary, Stefan Baral, Daouda Diouf, Kouamé Abo, and Marie-Claude Boily. Population-level impact of an accelerated HIV response plan to reach the UNAIDS 90-90-90 target in Côte d'Ivoire: Insights from mathematical modeling. *PLoS medicine*, 14(6):e1002321, 2017.

[65] Marie-Claude Boily, Rebecca F Baggaley, Lei Wang, Benoit Masse, Richard G White, Richard J Hayes, and Michel Alary. Heterosexual risk of HIV-1 infection per sexual act: systematic review and meta-analysis of observational studies. *The Lancet infectious diseases*, 9(2):118–129, 2009.

[66] James W Haefner. *Modeling Biological Systems:: Principles and Applications*. Springer Science & Business Media, 2005.

[67] Maia Martcheva. *An introduction to mathematical epidemiology*, volume 61. Springer, 2015.

[68] Nicholas G Davies, Adam J Kucharski, Rosalind M Eggo, Amy Gimma, W John Edmunds, Thibaut Jombart, Kathleen O'Reilly, Akira Endo, Joel Hellewell, Emily S Nightingale, et al. Effects of non-pharmaceutical interventions on COVID-19 cases, deaths, and demand for hospital services in the UK: a modelling study. *The Lancet Public Health*, 5(7):e375–e385, 2020.

[69] Edward S Knock, Lilith K Whittles, John A Lees, Pablo N Perez-Guzman, Robert Verity, Richard G FitzJohn, Katy AM Gaythorpe, Natsuko Imai, Wes Hinsley, Lucy C Okell, et al. The 2020 SARS-CoV-2 epidemic in England: key epidemiological drivers and impact of interventions. *medRxiv*, 2021.

[70] Matt J Keeling, Edward M Hill, Erin E Gorsich, Bridget Penman, Glen Guyver-Fletcher, Alex Holmes, Trystan Leng, Hector McKimm, Massimiliano Tamborrino, Louise Dyson, et al. Predictions of COVID-19 dynamics in the UK: short-term forecasting and analysis of potential exit strategies. *PLoS computational biology*, 17 (1):e1008619, 2021.

[71] Jon C Emery, Timothy W Russell, Yang Liu, Joel Hellewell, Carl AB Pearson, Gwenan M Knight, Rosalind M Eggo, Adam J Kucharski, Sebastian Funk, Stefan Flasche, et al. The contribution of asymptomatic SARS-CoV-2 infections to transmission on the Diamond Princess cruise ship. *Elife*, 9:e58699, 2020.

[72] Theresa Ryckman, Stephen Luby, Douglas K Owens, Eran Bendavid, and Jeremy D Goldhaber-Fiebert. Methods for Model Calibration under High Uncertainty: Modeling Cholera in Bangladesh. *Medical Decision Making*, 40(5):693–709, 2020.

[73] Adrian E Raftery and Le Bao. Estimating and projecting trends in hiv/aids generalized epidemics using incremental mixture importance sampling. *Biometrics*, 66(4):1162–1173, 2010.

[74] Martin Lange and Hans-Hermann Thulke. Elucidating transmission parameters of african swine fever through wild boar carcasses by combining spatio-temporal notification data and agent-based modelling. *Stochastic environmental research and risk assessment*, 31(2):379–391, 2017.

[75] Bärbel F Finkenstädt, Ottar N Bjørnstad, and Bryan T Grenfell. A stochastic model for extinction and recurrence of epidemics: estimation and inference for measles outbreaks. *Biostatistics*, 3(4):493–510, 2002.

[76] Hazel R Parry and Andrew J Evans. A comparative analysis of parallel processing and super-individual methods for improving the computational performance of a large individual-based model. *Ecological Modelling*, 214(2-4):141–152, 2008.

[77] Sarah P Otto and Troy Day. *A biologist's guide to mathematical modeling in ecology and evolution*. Princeton University Press, 2011.

[78] Benoît Dayrat. The roots of phylogeny: how did Haeckel build his trees? *Systematic Biology*, 52(4):515–527, 2003.

[79] Charles Semple, Mike Steel, et al. *Phylogenetics*, volume 24. Oxford University Press on Demand, 2003.

[80] Michael SY Lee and Alessandro Palci. Morphological phylogenetics in the genomic age. *Current Biology*, 25(19):R922–R929, 2015.

[81] Masami Hasegawa, Hirohisa Kishino, and Naruya Saitou. On the maximum likelihood method in molecular phylogenetics. *Journal of molecular evolution*, 32(5):443–445, 1991.

[82] James Franklin Crow, Motoo Kimura, et al. An introduction to population genetics theory. *An introduction to population genetics theory.*, 1970.

[83] Daniel L Hartl, Andrew G Clark, and Andrew G Clark. *Principles of population genetics*, volume 116. Sinauer associates Sunderland, MA, 1997.

[84] John FC Kingman. Origins of the coalescent: 1974-1982. *Genetics*, 156(4):1461–1463, 2000.

[85] John Frank Charles Kingman. The coalescent. *Stochastic processes and their applications*, 13(3):235–248, 1982.

[86] John Wakeley and Ori Sargsyan. Extensions of the coalescent effective population size. *Genetics*, 181(1):341–345, 2009.

[87] Oliver G Pybus and Andrew Rambaut. Evolutionary analysis of the dynamics of viral infectious disease. *Nature Reviews Genetics*, 10(8):540–550, 2009.

[88] Bryan T Grenfell, Oliver G Pybus, Julia R Gog, James LN Wood, Janet M Daly, Jenny A Mumford, and Edward C Holmes. Unifying the epidemiological and evolutionary dynamics of pathogens. *science*, 303(5656):327–332, 2004.

[89] Erik M Volz, Sergei L Kosakovsky Pond, Melissa J Ward, Andrew J Leigh Brown, and Simon DW Frost. Phylodynamics of infectious disease epidemics. *Genetics*, 183 (4):1421–1430, 2009.

[90] Erik M Volz, Katia Koelle, and Trevor Bedford. Viral phylodynamics. *PLoS computational biology*, 9(3):e1002947, 2013.

[91] Oliver G Pybus, Andrew Rambaut, and Paul H Harvey. An integrated framework for the inference of viral population history from reconstructed genealogies. *Genetics*, 155(3):1429–1437, 2000.

[92] Simon YW Ho and Beth Shapiro. Skyline-plot methods for estimating demographic history from nucleotide sequences. *Molecular ecology resources*, 11(3):423–434, 2011.

[93] Vladimir N Minin, Erik W Bloomquist, and Marc A Suchard. Smooth skyride through a rough skyline: Bayesian coalescent-based inference of population dynamics. *Molecular biology and evolution*, 25(7):1459–1471, 2008.

[94] Mandev S Gill, Philippe Lemey, Shannon N Bennett, Roman Biek, and Marc A Suchard. Understanding past population dynamics: Bayesian coalescent-based modeling with covariates. *Systematic biology*, 65(6):1041–1056, 2016.

[95] Benjamin Roche, John M Drake, Justin Brown, David E Stallknecht, Trevor Bedford, and Pejman Rohani. Adaptive evolution and environmental durability jointly structure phylodynamic patterns in avian influenza viruses. *PLoS Biol*, 12(8):e1001931, 2014.

[96] Verity Hill and Guy Baele. Bayesian estimation of past population dynamics in BEAST 1.10 using the Skygrid coalescent model. *Molecular biology and evolution*, 36(11):2620–2628, 2019.

[97] Erik M Volz and Xavier Didelot. Modeling the growth and decline of pathogen effective population size provides insight into epidemic dynamics and drivers of antimicrobial resistance. *Systematic Biology*, 67(4):719–728, 2018.

[98] Eric de Silva, Neil M Ferguson, and Christophe Fraser. Inferring pandemic growth rates from sequence data. *Journal of The Royal Society Interface*, 9(73):1797–1808, 2012.

[99] Xavier Didelot, Lily Geidelberg, Erik M Volz, et al. Model design for non-parametric phylodynamic inference and applications to pathogen surveillance. *bioRxiv: the preprint server for biology*, 2021.

[100] Matthew T Maurano, Sitharam Ramaswami, Paul Zappile, Dacia Dimartino, Ludovic Boytard, André M Ribeiro-dos Santos, Nicholas A Vulpescu, Gael Westby, Guomiao Shen, Xiaojun Feng, et al. Sequencing identifies multiple early introductions of SARS-CoV-2 to the New York City region. *Genome research*, 30(12):1781–1788, 2020.

[101] Erik Volz, Swapnil Mishra, Meera Chand, Jeffrey C Barrett, Robert Johnson, Lily Geidelberg, Wes R Hinsley, Daniel J Laydon, Gavin Dabrera, Áine O'Toole, et al. Assessing transmissibility of SARS-CoV-2 lineage B.1.1.7 in England. *Nature*, pages 1–17, 2021.

[102] Simon DW Frost and Erik M Volz. Viral phylodynamics and the search for an 'effective number of infections'. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1548):1879–1890, 2010.

[103] Katia Koelle and David A Rasmussen. Rates of coalescence for common epidemiological models at equilibrium. *Journal of the Royal Society Interface*, 9(70):997–1007, 2012.

[104] Erik M Volz. Complex population dynamics and the coalescent under neutrality. *Genetics*, 190(1):187–201, 2012.

[105] Erik M Volz and Igor Siveroni. Bayesian phylodynamic inference with complex models. *PLoS computational biology*, 14(11):e1006546, 2018.

[106] Fabrícia F Nascimento, Stefan Baral, Lily Geidelberg, Christinah Mukandavire, Sheree R Schwartz, Gnilane Turpin, Nguissali Turpin, Daouda Diouf, Nafissatou Leye Diouf, Karleen Coly, et al. Phylodynamic analysis of HIV-1 subtypes B, C and CRF 02_AG in Senegal. *Epidemics*, 30:100376, 2020.

[107] Manon Ragonnet-Cronin, Olivia Boyd, Lily Geidelberg, David Jorgensen, Fabricia F Nascimento, Igor Siveroni, Robert A Johnson, Marc Baguelin, Zulma M Cucunubá, Elita Jauneikaite, et al. Genetic evidence for the association between COVID-19 epidemic severity and timing of non-pharmaceutical interventions. *Nature communications*, 12(1):1–7, 2021.

[108] Danielle Miller, Michael A Martin, Noam Harel, Omer Tirosh, Talia Kustin, Moran Meir, Nadav Sorek, Shiraz Gefen-Halevi, Sharon Amit, Olesya Vorontsov, et al. Full genome viral sequences inform patterns of SARS-CoV-2 spread into and within Israel. *Nature communications*, 11(1):1–10, 2020.

[109] Hidetoshi Shimodaira and Masami Hasegawa. CONSEL: for assessing the confidence of phylogenetic tree selection. *Bioinformatics*, 17(12):1246–1247, 2001.

[110] Anna Smielewska, Edward Emmott, Kyriaki Ranellou, Ashley Popay, Ian Goodfellow, and Hamid Jalal. UK circulating strains of human parainfluenza 3: an amplicon based next generation sequencing method and phylogenetic analysis. *Wellcome Open Research*, 3, 2018.

[111] AC Ghani, CA Donnelly, and GP Garnett. Sampling biases and missing data in explorations of sexual partner networks for the spread of sexually transmitted diseases. *Statistics in medicine*, 17(18):2079–2097, 1998.

[112] Eamon B O'Dea and Claus O Wilke. Contact heterogeneity and phylodynamics: how contact networks shape parasite evolutionary trees. *Interdisciplinary perspectives on infectious diseases*, 2011, 2011.

[113] E M Volz and S D W Frost. "Scalable relaxed clock phylogenetic dating". *Virus Evol*, 3(2), July 2017.

[114] Simon Dellicour, Guy Baele, Gytis Dudas, Nuno R Faria, Oliver G Pybus, Marc A Suchard, Andrew Rambaut, and Philippe Lemey. Phylodynamic assessment of intervention strategies for the West African Ebola virus outbreak. *Nature communications*, 9(1):1–9, 2018.

[115] Mario Stevenson. HIV-1 pathogenesis. *Nature medicine*, 9(7):853–860, 2003.

[116] Kathleen Case. Nomenclature: human immunodeficiency virus, 1986.

[117] Thierry Huet, Rémi Cheynier, Andreas Meyerhans, Georges Roelants, and Simon Wain-Hobson. Genetic organization of a chimpanzee lentivirus related to HIV-1. *Nature*, 345(6273):356–359, 1990.

[118] Sabrina Locatelli and Martine Peeters. Cross-species transmission of simian retroviruses: how and why they could lead to the emergence of new diseases in the human population. *AIDS*, 26(6):659–673, 2012.

[119] Joseph M Watts, Kristen K Dang, Robert J Gorelick, Christopher W Leonard, Julian W Bess Jr, Ronald Swanstrom, Christina L Burch, and Kevin M Weeks. Architecture and secondary structure of an entire HIV-1 RNA genome. *Nature*, 460(7256):711–716, 2009.

[120] Simon Wain-Hobson, Pierre Sonigo, Olivier Danos, Stewart Cole, and Marc Alizon. Nucleotide sequence of the AIDS virus, LAV. *Cell*, 40(1):9–17, 1985.

[121] J Votteler and U Schubert. Human immunodeficiency viruses: molecular biology. 2008.

[122] George M Shaw and Eric Hunter. HIV transmission. *Cold Spring Harbor perspectives in medicine*, 2(11):a006965, 2012.

[123] Laura Pattacini, Pamela M Murnane, Erin M Kahle, Michael J Bolton, Jeffrey J Delrow, Jairam R Lingappa, Elly Katabira, Deborah Donnell, M Juliana McElrath, Jared M Baeten, et al. Differential regulatory T cell activity in HIV type 1-exposed seronegative individuals. *AIDS research and human retroviruses*, 29(10):1321–1329, 2013.

[124] Nathan W Cummins and Andrew D Badley. Making sense of how HIV kills infected CD4 T cells: implications for HIV cure. *Molecular and cellular therapies*, 2(1):1–7, 2014.

[125] Donald S Burke. Recombination in HIV: an important viral evolutionary strategy. *Emerging infectious diseases*, 3(3):253, 1997.

[126] Michael T Bretscher, Christian L Althaus, Viktor Müller, and Sebastian Bonhoeffer. Recombination in HIV and the evolution of drug resistance: for better or for worse? *Bioessays*, 26(2):180–188, 2004.

[127] Laurence Vergne, Coumba Touré Kane, Christian Laurent, Ndella Diakhaté, Ndeye Fatou Ngom Gueye, Pape Mandoumbé Gueye, Papa Salif Sow, Mame Awa Faye, Florian Liégeois, Adama Ndir, et al. Low rate of genotypic HIV-1 drug-resistant strains in the Senegalese government initiative of access to antiretroviral therapy. *AIDS*, 17:S31–S38, 2003.

[128] Abhishek Pandey and Alison P Galvani. The global burden of HIV and prospects for control. *The Lancet HIV*, 6(12):e809–e811, 2019.

[129] UNAIDS. FACT SHEET 2021, 2021. URL https://www.unaids.org/sites/default/files/media_asset/UNAIDS_FactSheet_en.pdf. Accessed: 2021-08-11.

[130] World Health Organisation. Key facts HIV, 2021. URL https://cdn.who.int/media/docs/default-source/hq-hiv-hepatitis-and-stis-library/key-facts-hiv-2020.pdf?sfvrsn=582c3f6e_13. Accessed: 2021-08-11.

[131] Yibeltal Assefa and Charles F Gilks. Ending the epidemic of HIV/AIDS by 2030: Will there be an endgame to HIV, or an endemic HIV requiring an integrated health systems response in many countries? *International Journal of Infectious Diseases*, 100:273–277, 2020.

[132] UNAIDS. UNAIDS TERMINOLOGY GUIDELINES, 2021. URL https://www.unaids.org/sites/default/files/media_asset/2015_terminology_guidelines_en.pdf. Accessed: 2021-08-11.

[133] World Health Organisation. HIV PREVENTION, DIAGNOSIS, TREATMENT AND CARE FOR KEY POPULATIONS, 2021. URL http://apps.who.int/iris/bitstream/handle/10665/246200/9789241511124. Accessed: 2021-08-11.

[134] Sharmistha Mishra, Romain Silhol, Jesse Knight, Refilwe Phaswana-Mafuya, Daouda Diouf, Linwei Wang, Sheree Schwartz, Marie-Claude Boily, and Stefan Baral. Estimating the epidemic consequences of hiv prevention gaps among key populations. *Journal of the International AIDS Society*, 24:e25739, 2021.

[135] UNAIDS. GLOBAL AIDS UPDATE, 2021. URL https://www.unaids.org/sites/default/files/media_asset/global-AIDS-update-2016_en.pdf. Accessed: 2021-08-11.

[136] David Wilson and Daniel T Halperin. "Know your epidemic, know your response": a useful approach, if we get it right. *The Lancet*, 372(9637):423–426, 2008.

[137] Marie-Claude Boily, C Lowndes, and M Alary. The impact of HIV epidemic phases on the effectiveness of core group interventions: insights from mathematical models. *Sexually transmitted infections*, 78(suppl 1):i78–i90, 2002.

[138] Judith Vandepitte, Rob Lyerla, Gina Dallabetta, Francois Crabbé, Michel Alary, and Anne Buvé. Estimates of the number of female sex workers in different regions of the world. *Sexually transmitted infections*, 82(suppl 3):iii18–iii25, 2006.

[139] Christine Harcourt and Basil Donovan. The many faces of sex work. *Sexually transmitted infections*, 81(3):201–206, 2005.

[140] Kate Shannon and Joanne Csete. Violence, condom negotiation, and HIV/STI risk among sex workers. *Jama*, 304(5):573–574, 2010.

[141] Stefan Baral, Chris Beyrer, Kathryn Muessig, Tonia Poteat, Andrea L Wirtz, Michele R Decker, Susan G Sherman, and Deanna Kerrigan. Burden of HIV among female sex workers in low-income and middle-income countries: a systematic review and meta-analysis. *The Lancet infectious diseases*, 12(7):538–549, 2012.

[142] Mohammad Ali Mansournia and Douglas G Altman. Population attributable fraction. *Bmj*, 360, 2018.

[143] Christinah Mukandavire, Josephine Walker, Sheree Schwartz, Marie-Claude Boily, Leon Danon, Carrie Lyons, Daouda Diouf, Ben Liestman, Nafissatou Leye Diouf, Fatou Drame, et al. Estimating the contribution of key populations towards the spread of HIV in Dakar, Senegal. *Journal of the International AIDS Society*, 21:e25126, 2018.

[144] Marie-Claude Boily, Michael Pickles, Michel Alary, Stefan Baral, James Blanchard, Stephen Moses, Peter Vickerman, and Sharmistha Mishra. What really is a concentrated HIV epidemic and what does it mean for West and Central Africa? Insights from mathematical modeling. *JAIDS Journal of Acquired Immune Deficiency Syndromes*, 68:S74–S82, 2015.

[145] Peter Vickerman, Anna M Foss, Michael Pickles, Kathleen Deering, Supriya Verma, Eric Demers, Catherine M Lowndes, Stephen Moses, Michel Alary, and Marie-Claude Boily. To what extent is the HIV epidemic in southern India driven by commercial sex? A modelling analysis. *AIDS*, 24(16):2563–2572, 2010.

[146] Windy MV Wariki, Erika Ota, Rintaro Mori, Ai Koyanagi, Narumi Hori, and Kenji Shibuya. Behavioral interventions to reduce the transmission of HIV infection among sex workers and their clients in low-and middle-income countries. *Cochrane Database of Systematic Reviews*, (2), 2012.

[147] Myron S Cohen, Ying Q Chen, Marybeth McCauley, Theresa Gamble, Mina C Hosseinipour, Nagalingeswaran Kumarasamy, James G Hakim, Johnstone Kumwenda, Beatriz Grinsztejn, Jose HS Pilotto, et al. Prevention of HIV-1 infection with early antiretroviral therapy. *New England journal of medicine*, 365(6):493–505, 2011.

[148] Marie A Brault, Donna Spiegelman, James Hargreaves, Denis Nash, and Sten H Vermund. Treatment as prevention: concepts and challenges for reducing HIV incidence. *Journal of acquired immune deficiency syndromes*, 82(2):S104, 2019.

[149] Collins C Iwuji, Joanna Orne-Gliemann, Frank Tanser, Sylvie Boyer, Richard J Lessells, France Lert, John Imrie, Till Bärnighausen, Claire Rekacewicz, Brigitte Bazin, et al. Evaluation of the impact of immediate versus WHO recommendations-guided antiretroviral therapy initiation on HIV incidence: the ANRS 12249 TasP (Treatment as Prevention) trial in Hlabisa sub-district, KwaZulu-Natal, South Africa: study protocol for a cluster randomised controlled trial. *Trials*, 14(1):1–15, 2013.

[150] Mélanie Plazy, Kamal El Farouki, Collins Iwuji, Nonhlanhla Okesola, Joanna Orne-Gliemann, Joseph Larmarange, France Lert, Marie-Louise Newell, François Dabis, Rosemary Dray-Spira, et al. Access to HIV care in the context of universal test and treat: challenges within the ANRS 12249 TasP cluster-randomized trial in rural South Africa. *Journal of the International AIDS Society*, 19(1):20913, 2016.

[151] Richard Hayes, Helen Ayles, Nulda Beyers, Kalpana Sabapathy, Sian Floyd, Kwame Shanaube, P Bock, S Griffith, A Moore, D Watson-Jones, et al. HPTN 071 (PopART) Study Team. HPTN 071 (PopART): rationale and design of a cluster-randomised trial of the population impact of an HIV combination prevention intervention including universal testing and treatment—a study protocol for a cluster randomised trial. *Trials*, 15(1):57, 2014.

[152] Richard J Hayes, Deborah Donnell, Sian Floyd, Nomtha Mandla, Justin Bwalya, Kalpana Sabapathy, Blia Yang, Mwelwa Phiri, Ab Schaap, Susan H Eshleman, et al. Effect of universal testing and treatment on HIV incidence—HPTN 071 (PopART). *New England Journal of Medicine*, 381(3):207–218, 2019.

[153] Diane V Havlir, Laura B Balzer, Edwin D Charlebois, Tamara D Clark, Dalsone Kwarisiima, James Ayieko, Jane Kabami, Norton Sang, Teri Liegler, Gabriel Chamie, et al. HIV testing and treatment with the use of a community health approach in rural Africa. *New England Journal of Medicine*, 381(3):219–229, 2019.

[154] Joseph Makhema, Kathleen E Wirth, Molly Pretorius Holme, Tendani Gaolathe, Mompati Mmalane, Etienne Kadima, Unoda Chakalisa, Kara Bennett, Jean Leidner,

Kutlo Manyake, et al. Universal testing, expanded treatment, and incidence of HIV infection in Botswana. *New England Journal of Medicine*, 381(3):230–242, 2019.

[155] Sheena McCormack, David T Dunn, Monica Desai, David I Dolling, Mitzy Gafos, Richard Gilson, Ann K Sullivan, Amanda Clarke, Iain Reeves, Gabriel Schembri, et al. Pre-exposure prophylaxis to prevent the acquisition of HIV-1 infection (PROUD): effectiveness results from the pilot phase of a pragmatic open-label randomised trial. *The Lancet*, 387(10013):53–60, 2016.

[156] World Health Organization et al. *Guideline on when to start antiretroviral therapy and on pre-exposure prophylaxis for HIV*. World Health Organization, 2015.

[157] Peter L Anderson, David V Glidden, Albert Liu, Susan Buchbinder, Javier R Lama, Juan Vicente Guanira, Vanessa McMahan, Lane R Bushman, Martín Casapía, Orlando Montoya-Herrera, et al. Emtricitabine-tenofovir concentrations and pre-exposure prophylaxis efficacy in men who have sex with men. *Science translational medicine*, 4(151):151ra125–151ra125, 2012.

[158] Jean-Michel Molina, Catherine Capitant, Bruno Spire, Gilles Pialoux, Laurent Cotte, Isabelle Charreau, Cecile Tremblay, Jean-Marie Le Gall, Eric Cua, Armelle Pasquet, et al. On-demand preexposure prophylaxis in men at high risk for HIV-1 infection. *N Engl J Med*, 373:2237–2246, 2015.

[159] Robert M Grant, Peter L Anderson, Vanessa McMahan, Albert Liu, K Rivet Amico, Megha Mehrotra, Sybil Hosek, Carlos Mosquera, Martin Casapia, Orlando Montoya, et al. Uptake of pre-exposure prophylaxis, sexual practices, and HIV incidence in men and transgender women who have sex with men: a cohort study. *The Lancet infectious diseases*, 14(9):820–829, 2014.

[160] Lut Van Damme, Amy Corneli, Khatija Ahmed, Kawango Agot, Johan Lombaard, Saidi Kapiga, Mookho Malahleha, Fredrick Owino, Rachel Manongi, Jacob Onyango, et al. Preexposure prophylaxis for HIV infection among African women. *New England Journal of Medicine*, 367(5):411–422, 2012.

[161] Jeanne M Marrazzo, Gita Ramjee, Barbra A Richardson, Kailazarid Gomez, Nyaradzo Mgodi, Gonasagrie Nair, Thesla Palanee, Clemensia Nakabiito, Ariane Van Der Straten, Lisa Noguchi, et al. Tenofovir-based preexposure prophylaxis for HIV infection among African women. *New England Journal of Medicine*, 372(6): 509–518, 2015.

[162] Jared M Baeten, Deborah Donnell, Patrick Ndase, Nelly R Mugo, James D Campbell, Jonathan Wangisi, Jordan W Tappero, Elizabeth A Bukusi, Craig R Cohen, Elly Katabira, et al. Antiretroviral prophylaxis for HIV prevention in heterosexual men and women. *New England Journal of Medicine*, 367(5):399–410, 2012.

[163] Mackenzie L Cottrell, Kuo H Yang, Heather MA Prince, Craig Sykes, Nicole White, Stephanie Malone, Evan S Dellon, Ryan D Madanick, Nicholas J Shaheen, Michael G Hudgens, et al. A translational pharmacology approach to predicting outcomes of preexposure prophylaxis against HIV in men and women using tenofovir disoproxil fumarate with or without emtricitabine. *The Journal of infectious diseases*, 214(1): 55–64, 2016.

[164] JK Kyongo, M Kiragu, R Karuga, C Ochieng, A Ngunjiri, C Wachihi, H Musyoki, L Digolo, L Otiso, L Gelmon, et al. How long will they take it? Oral pre-exposure prophylaxis (PrEP) retention for female sex workers, men who have sex with men and young women in a demonstration project in Kenya. In *Journal of the International AIDS Society*, volume 21, pages 54–55. JOHN WILEY & SONS LTD THE ATRIUM, SOUTHERN GATE, CHICHESTER PO19 8SQ, W . . . , 2018.

[165] Robyn Eakle, Gabriela B Gomez, Niven Naicker, Rutendo Bothma, Judie Mbogua, Maria A Cabrera Escobar, Elaine Saayman, Michelle Moorhouse, WD Francois Venter, Helen Rees, et al. HIV pre-exposure prophylaxis and early antiretroviral treatment among female sex workers in South Africa: results from a prospective observational demonstration project. *PLoS medicine*, 14(11):e1002444, 2017.

[166] Joseph Kagaayi, Larry W Chang, Victor Ssempijja, M Kate Grabowski, Robert Ssekubugu, Gertrude Nakigozi, Godfrey Kigozi, David M Serwadda, Ronald H Gray, Fred Nalugoda, et al. Impact of combination HIV interventions on HIV incidence in hyperendemic fishing communities in Uganda: a prospective cohort study. *The Lancet HIV*, 6(10):e680–e687, 2019.

[167] Joseph Kagaayi, James Batte, Hadijja Nakawooya, Boniface Kigozi, Gertrude Nakigozi, Susanne Strömdahl, Anna Mia Ekström, Larry W Chang, Ron Gray, Steven J Reynolds, et al. Uptake and retention on HIV pre-exposure prophylaxis among key and priority populations in South-Central Uganda. *Journal of the International AIDS Society*, 23(8):e25588, 2020.

[168] Moussa Sarr, Daouda Gueye, Aminata Mboup, Ousmane Diouf, Mame D Bousso Bao, Anna Julienne Ndiaye, Birahim P Ndiaye, Stephen E Hawes, Eric Tousset, Abdoulaye Diallo, et al. Uptake, retention, and outcomes in a demonstration project of pre-exposure prophylaxis among female sex workers in public health centers in Senegal. *International journal of STD & AIDS*, 31(11):1063–1072, 2020.

[169] Aminata Mboup, Luc Béhanzin, Fernand A Guédou, Nassirou Geraldo, Ella Goma-Matsétsé, Katia Giguère, Marlène Aza-Gnandji, Léon Kessou, Mamadou Diallo, René K Kêkê, et al. Early antiretroviral therapy and daily pre-exposure prophylaxis for HIV prevention among female sex workers in Cotonou, Benin: a prospective observational demonstration study. *Journal of the International AIDS Society*, 21(11): e25208, 2018.

[170] Tony Kirby. Cabotegravir, a new option for PrEP. *The Lancet Infectious Diseases*, 20 (7):781, 2020.

[171] Christine Tagliaferri Rael, Javier Lopez-Ríos, Stacey A McKenna, Doyel Das, Curtis Dolezal, Elena Abascal, Alex Carballo-Diéguez, Rebecca Schnall, Thomas J Hope, José Bauermeister, et al. Transgender Women's Barriers, Facilitators, and Preferences on Tailored Injection Delivery Strategies to Administer Long-Acting Injectable Cabotegravir (CAB-LA) for HIV Pre-exposure Prophylaxis (PrEP). *AIDS and Behavior*, pages 1–13, 2021.

[172] S Delany Moretlwe, J Hughes, P Bock, S Gurrion, P Hunidzarira, D Kalonji, N Kayange, J Makhema, P Mandima, C Mathew, et al. Long acting injectable

cabotegravir is safe and effective in preventing HIV infection in cisgender women: interim results from HPTN 084. *Journal of the International AIDS Society*, 24(S1): 8–9, 2021.

[173] World Health Organisation. Trial results reveal that long-acting injectable cabotegravir as PrEP is highly effective in preventing HIV acquisition in women, 2021. URL https://www.who.int/news/item/09-11-2020-trial-results-reveal-that-long-acting-injectable-cabotegravir-as-prep-is-highly-effective-in-preventing-hiv-acquisition-in-women. Accessed: 2021-08-11.

[174] John Thornhill and Chloe Orkin. Long-acting injectable HIV therapies: the next frontier. *Current Opinion in Infectious Diseases*, 34(1):8–15, 2021.

[175] Max Roser, Hannah Ritchie, Esteban Ortiz-Ospina, and Joe Hasell. Coronavirus pandemic (COVID-19), 2021. Accessed: 2021-08-11.

[176] Peng Zhou, Xing-Lou Yang, Xian-Guang Wang, Ben Hu, Lei Zhang, Wei Zhang, Hao-Rui Si, Yan Zhu, Bei Li, Chao-Lin Huang, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*, 579(7798):270–273, 2020.

[177] Hong Zhou, Jingkai Ji, Xing Chen, Yuhai Bi, Juan Li, Qihui Wang, Tao Hu, Hao Song, Runchu Zhao, Yanhua Chen, et al. Identification of novel bat coronaviruses sheds light on the evolutionary origins of SARS-CoV-2 and related viruses. *Cell*, 2021.

[178] Berend Jan Bosch, Ruurd Van der Zee, Cornelis AM De Haan, and Peter JM Rottier. The coronavirus spike protein is a class I virus fusion protein: structural and functional characterization of the fusion core complex. *Journal of virology*, 77(16):8801–8811, 2003.

[179] Yuelong Shu and John McCauley. GISAID: Global initiative on sharing all influenza data–from vision to reality. *Eurosurveillance*, 22(13):30494, 2017.

[180] Guoping Li, Xiang He, Lei Zhang, Qin Ran, Junyi Wang, Anying Xiong, Dehong Wu, Feng Chen, Jinlyu Sun, and Christopher Chang. Assessing ACE2 expression patterns in lung tissues in the pathogenesis of COVID-19. *Journal of autoimmunity*, 112:102463, 2020.

[181] Zharko Daniloski, Tristan X Jordan, Hans-Hermann Wessels, Daisy A Hoagland, Silva Kasela, Mateusz Legut, Silas Maniatis, Eleni P Mimitou, Lu Lu, Evan Geller, et al. Identification of required host factors for SARS-CoV-2 infection in human cells. *Cell*, 184(1):92–105, 2021.

[182] Philip V'kovski, Annika Kratzel, Silvio Steiner, Hanspeter Stalder, and Volker Thiel. Coronavirus biology and replication: implications for SARS-CoV-2. *Nature Reviews Microbiology*, 19(3):155–170, 2021.

[183] Silvia Angeletti, Domenico Benvenuto, Martina Bianchi, Marta Giovanetti, Stefano Pascarella, and Massimo Ciccozzi. COVID-2019: the role of the nsp2 and nsp3 in its pathogenesis. *Journal of medical virology*, 92(6):584–588, 2020.

[184] Souheil Zayet, Vincent Gendrin, and Timothee Klopfenstein. Natural history of COVID-19: back to basics. *New Microbes and New Infections*, 38:100815, 2020.

[185] Christel Faes, Steven Abrams, Dominique Van Beckhoven, Geert Meyfroidt, Erika Vlieghe, Niel Hens, et al. Time between symptom onset, hospitalisation and recovery or death: statistical analysis of Belgian COVID-19 patients. *International journal of environmental research and public health*, 17(20):7560, 2020.

[186] Tapiwa Ganyani, Cecile Kremer, Dongxuan Chen, Andrea Torneri, Christel Faes, Jacco Wallinga, and Niel Hens. Estimating the generation interval for coronavirus disease (COVID-19) based on symptom onset data, March 2020. *Eurosurveillance*, 25(17):2000257, 2020.

[187] Conor McAloon, Áine Collins, Kevin Hunt, Ann Barber, Andrew W Byrne, Francis Butler, Miriam Casey, John Griffin, Elizabeth Lane, David McEvoy, et al. Incubation period of covid-19: a rapid systematic review and meta-analysis of observational research. *BMJ open*, 10(8):e039652, 2020.

[188] Robert Verity, Lucy C Okell, Ilaria Dorigatti, Peter Winskill, Charles Whittaker, Natsuko Imai, Gina Cuomo-Dannenburg, Hayley Thompson, Patrick GT Walker, Han Fu, et al. Estimates of the severity of coronavirus disease 2019: a model-based analysis. *The Lancet infectious diseases*, 20(6):669–677, 2020.

[189] Audrey F Pennington, Lyudmyla Kompaniyets, April D Summers, Melissa L Danielson, Alyson B Goodman, Jennifer R Chevinsky, Leigh Ellyn Preston, Lyna Z Schieber, Gonza Namulanda, Joseph Courtney, et al. Risk of clinical severity by age and race/ethnicity among adults hospitalized for COVID-19—United States, March–September 2020. In *Open forum infectious diseases*, volume 8, page ofaa638. Oxford University Press US, 2021.

[190] Oyungerel Byambasuren, Magnolia Cardona, Katy Bell, Justin Clark, Mary-Louise McLaws, and Paul Glasziou. Estimating the extent of asymptomatic COVID-19 and its potential for community transmission: systematic review and meta-analysis. *Official Journal of the Association of Medical Microbiology and Infectious Disease Canada*, 5 (4):223–234, 2020.

[191] Andrew T Levin, William P Hanage, Nana Owusu-Boaitey, Kensington B Cochran, Seamus P Walsh, and Gideon Meyerowitz-Katz. Assessing the age specificity of infection fatality rates for COVID-19: systematic review, meta-analysis, and public policy implications. *European journal of epidemiology*, pages 1–16, 2020.

[192] Hannah Peckham, Nina M de Gruijter, Charles Raine, Anna Radziszewska, Coziana Ciurtin, Lucy R Wedderburn, Elizabeth C Rosser, Kate Webb, and Claire T Deakin. Male sex identified by global COVID-19 meta-analysis as a risk factor for death and ITU admission. *Nature communications*, 11(1):1–10, 2020.

[193] Matthew A Raifman and Julia R Raifman. Disparities in the population at risk of severe illness from COVID-19 by race/ethnicity and income. *American journal of preventive medicine*, 59(1):137–139, 2020.

[194] Jing Yang, Ya Zheng, Xi Gou, Ke Pu, Zhaofeng Chen, Qinghong Guo, Rui Ji, Haojia Wang, Yuping Wang, and Yongning Zhou. Prevalence of comorbidities in the novel Wuhan coronavirus (COVID-19) infection: a systematic review and meta-analysis. *Int J Infect Dis*, 10(10.1016), 2020.

[195] Hagai Rossman, Tomer Meir, Jonathan Somer, Smadar Shilo, Rom Gutman, Asaf Ben Arie, Eran Segal, Uri Shalit, and Malka Gorfine. Hospital load and increased COVID-19 related mortality in Israel. *Nature communications*, 12(1):1–7, 2021.

[196] Johns Hopkins Coronavirus Resource Center. Mortality analyses. *Johns Hopkins Coronavirus Resource Center*, 2020.

[197] World Health Organisation. Coronavirus disease (COVID-19) advice for the public: Mythbusters, 2021. URL https://www.who.int/emergencies/diseases/novel-coronavirus-2019/advice-for-public/myth-busters. Accessed: 2021-08-11.

[198] Peter C Taylor, Andrew C Adams, Matthew M Hufford, Inmaculada de la Torre, Kevin Winthrop, and Robert L Gottlieb. Neutralizing monoclonal antibodies for treatment of covid-19. *Nature Reviews Immunology*, pages 1–12, 2021.

[199] RECOVERY Collaborative Group. Dexamethasone in hospitalized patients with Covid-19. *New England Journal of Medicine*, 384(8):693–704, 2021.

[200] Mark G Thompson, Jefferey L Burgess, Allison L Naleway, Harmony L Tyner, Sarang K Yoon, Jennifer Meece, Lauren EW Olsho, Alberto J Caban-Martinez, Ashley Fowlkes, Karen Lutrick, et al. Interim estimates of vaccine effectiveness of BNT162b2 and mRNA-1273 COVID-19 vaccines in preventing SARS-CoV-2 infection among health care personnel, first responders, and other essential and frontline workers—eight US locations, December 2020–March 2021. *Morbidity and Mortality Weekly Report*, 70(13):495, 2021.

[201] Laith J Abu-Raddad, Hiam Chemaitelly, and Adeel A Butt. Effectiveness of the BNT162b2 Covid-19 Vaccine against the B. 1.1. 7 and B. 1.351 Variants. *New England Journal of Medicine*, 2021.

[202] Eric J Haas, Frederick J Angulo, John M McLaughlin, Emilia Anis, Shepherd R Singer, Farid Khan, Nati Brooks, Meir Smaja, Gabriel Mircus, Kaijie Pan, et al. Impact and effectiveness of mRNA BNT162b2 vaccine against SARS-CoV-2 infections and COVID-19 cases, hospitalisations, and deaths following a nationwide vaccination campaign in Israel: an observational study using national surveillance data. *The Lancet*, 397(10287):1819–1829, 2021.

[203] Mahesh Jayaweera, Hasini Perera, Buddhika Gunawardana, and Jagath Manatunge. Transmission of COVID-19 virus by droplets and aerosols: A critical review on the unresolved dichotomy. *Environmental research*, 188:109819, 2020.

[204] Emanuel Goldman. Exaggerated risk of transmission of COVID-19 by fomites. *The Lancet Infectious Diseases*, 20(8):892–893, 2020.

[205] Lauren C Tindale, Jessica E Stockdale, Michelle Coombe, Emma S Garlock, Wing Yin Venus Lau, Manu Saraswat, Louxin Zhang, Dongxuan Chen, Jacco Wallinga, and Caroline Colijn. Evidence for transmission of covid-19 prior to symptom onset. *Elife*, 9:e57149, 2020.

[206] Xinghua Chang, Maoxing Liu, Zhen Jin, and Jianrong Wang. Studying on the impact of media coverage on the spread of COVID-19 in Hubei Province, China. *Math. Biosci. Eng*, 17(4):3147–3159, 2020.

[207] Xiaoyan Mu, Anthony Gar-On Yeh, and Xiaohu Zhang. The interplay of spatial spread of COVID-19 and human mobility in the urban system of China during the Chinese New Year. *Environment and Planning B: Urban Analytics and City Science*, page 2399808320954211, 2020.

[208] Simiao Chen, Juntao Yang, Weizhong Yang, Chen Wang, and Till Bärnighausen. COVID-19 control in China during mass population movements at New Year. *The Lancet*, 395(10226):764–766, 2020.

[209] World Health Organisation. Novel Coronavirus (2019-nCoV) SITUATION REPORT 4 - 24 JANUARY 2020, 2020. URL https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200124-sitrep-4-2019-ncov.pdf.

[210] World Health Organisation. Novel Coronavirus (2019-nCoV) SITUATION REPORT 51 - 11 MARCH 2020, 2020. URL https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200311-sitrep-51-covid-19.pdf.

[211] Tim K Tsang, Peng Wu, Yun Lin, Eric HY Lau, Gabriel M Leung, and Benjamin J Cowling. Effect of changing case definitions for COVID-19 on the epidemic curve and transmission parameters in mainland China: a modelling study. *The Lancet Public Health*, 5(5):e289–e296, 2020.

[212] Karthikeyan Ravichandran, Subbaiyan Anbazhagan, H Agri, RN Rupner, VKO Rajendran, K Dhama, and BR Singh. Global status of COVID-19 diagnosis: an overview. *J Pure Appl Microbiol*, 14(suppl 1):879–892, 2020.

[213] Patrizio Vanella, Christian Wiessner, Anja Holz, Gérard Krause, Annika Möhl, Sarah Wiegel, Berit Lange, and Heiko Becher. The Role of Age Distribution, Time Lag Between Reporting and Death and Healthcare System Capacity in Case Fatality Estimates of COVID-19. 2020.

[214] Blavatnik School of Government. COVID-19 Government Response Tracker, 2021. URL ttps://www.bsg.ox.ac.uk/research/research-projects/covid-19-government-response-tracker. Accessed: 2021-04-09.

[215] Chinenyenwa Ohia, Adeleye S Bakarey, and Tauseef Ahmad. COVID-19 and Nigeria: putting the realities in context. *International Journal of Infectious Diseases*, 95: 279–281, 2020.

[216] Lawrence Mwananyanda, Christopher J Gill, William MacLeod, Geoffrey Kwenda, Rachel Pieciak, Zachariah Mupila, Rotem Lapidot, Francis Mupeta, Leah Forman, Luunga Ziko, et al. Covid-19 deaths in Africa: prospective systematic postmortem surveillance study. *BMJ*, 372, 2021.

[217] Fabrice Mougeni, Ance Mangaboula, and Bertrand Lell. The potential effect of the African population age structure on COVID-19 mortality. *Medrxiv*, 2020.

[218] Jun Zheng. SARS-CoV-2: an emerging coronavirus that causes a global threat. *International journal of biological sciences*, 16(10):1678, 2020.

[219] John-Sebastian Eden. Genome sequencing and its use in public health responses to COVID-19. *Microbiology Australia*, 42(1):44–46, 2021.

[220] Xingguang Li, Junjie Zai, Qiang Zhao, Qing Nie, Yi Li, Brian T Foley, and Antoine Chaillon. Evolutionary history, potential intermediate animal host, and cross-species analyses of SARS-CoV-2. *Journal of medical virology*, 92(6):602–611, 2020.

[221] Shu-Miaw Chaw, Jui-Hung Tai, Shi-Lun Chen, Chia-Hung Hsieh, Sui-Yuan Chang, Shiou-Hwei Yeh, Wei-Shiung Yang, Pei-Jer Chen, and Hurng-Yi Wang. The origin and underlying driving forces of the SARS-CoV-2 outbreak. *Journal of biomedical science*, 27:1–12, 2020.

[222] Mahan Ghafari, Louis du Plessis, Oliver Pybus and Aris Katzourakis. "Virological". http://virological.org/, 2020. Accessed: 2021-08-21.

[223] Zhongming Zhao, Haipeng Li, Xiaozhuang Wu, Yixi Zhong, Keqin Zhang, Ya-Ping Zhang, Eric Boerwinkle, and Yun-Xin Fu. Moderate mutation rate in the SARS coronavirus genome and its implications. *BMC evolutionary biology*, 4(1):1–9, 2004.

[224] Maciej F Boni, Philippe Lemey, Xiaowei Jiang, Tommy Tsan-Yuk Lam, Blair W Perry, Todd A Castoe, Andrew Rambaut, and David L Robertson. Evolutionary origins of the SARS-CoV-2 sarbecovirus lineage responsible for the COVID-19 pandemic. *Nature microbiology*, 5(11):1408–1417, 2020.

[225] Andrew Rambaut. "Phylodynamic Analysis | 46 genomes | 31 Jan 2020". http://virological.org/t/phylodynamic-analysis-46-genomes-31-jan-2020/356, January 2020. Accessed: 2020-2-2.

[226] Andrew Rambaut, Edward C Holmes, Áine O'Toole, Verity Hill, John T McCrone, Christopher Ruis, Louis du Plessis, and Oliver G Pybus. Addendum: A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. *Nature Microbiology*, 6(3):415–415, 2021.

[227] Andrew Rambaut, Edward C Holmes, Áine O'Toole, Verity Hill, John T McCrone, Christopher Ruis, Louis du Plessis, and Oliver G Pybus. A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. *Nature microbiology*, 5(11):1403–1407, 2020.

[228] Public Health England. SARS-CoV-2 variants of concern and variants under investigation in England - Technical briefing 8, 2021. URL https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/975742/Variants_of_Concern_VOC_Technical_Briefing_8_England.pdf.

[229] Shengjie Lai, Nick W Ruktanonchai, Liangcai Zhou, Olivia Prosper, Wei Luo, Jessica R Floyd, Amy Wesolowski, Mauricio Santillana, Chi Zhang, Xiangjun Du, et al. Effect of non-pharmaceutical interventions to contain COVID-19 in China. *nature*, 585(7825):410–413, 2020.

[230] Wei Chen, Qing Wang, YQ Li, HL Yu, YY Xia, ML Zhang, Ying Qin, Ting Zhang, ZB Peng, RC Zhang, et al. Early containment strategies and core measures for prevention and control of novel coronavirus pneumonia in China. *Zhonghua yu fang yi xue za zhi [Chinese journal of preventive medicine]*, 54(3):1–6, 2020.

[231] Mark Zanin, Cheng Xiao, Tingting Liang, Shiman Ling, Fengming Zhao, Zhenting Huang, Fangmei Lin, Xia Lin, Zhanpeng Jiang, and Sook-San Wong. The public health response to the COVID-19 outbreak in mainland China: a narrative review. *Journal of thoracic disease*, 12(8):4434, 2020.

[232] Seth Flaxman, Swapnil Mishra, Axel Gandy, H Juliette T Unwin, Thomas A Mellan, Helen Coupland, Charles Whittaker, Harrison Zhu, Tresnia Berah, Jeffrey W Eaton, et al. Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature*, 584(7820):257–261, 2020.

[233] Ralf Reintjes. Lessons in contact tracing from Germany, 2020.

[234] france24, 2021. URL http://web.archive.org/web/20200425104041/https://www.france24.com/en/20200327-french-pm-extends-coronavirus-lockdown-by-two-weeks-until-april-15. Accessed: 2021-08-11.

[235] euractiv, 2021. URL http://web.archive.org/web/20200421180927/https://www.euractiv.com/section/coronavirus/news/spain-goes-under-lockdown-to-fight-coronavirus/. Accessed: 2021-08-11.

[236] BBC. Coronavirus: Italy extends emergency measures nationwide, 2021. URL https://www.bbc.co.uk/news/world-europe-51810673. Accessed: 2021-08-11.

[237] Financial Times, 2021. URL https://web.archive.org/web/20200523152613/https://www.ft.com/content/31de03b8-6dbc-11ea-89df-41bea055720b. Accessed: 2021-08-11.

[238] Moustapha Mbow, Bertrand Lell, Simon P Jochems, Badara Cisse, Souleymane Mboup, Benjamin G Dewals, Assan Jaye, Alioune Dieye, and Maria Yazdanbakhsh. COVID-19 in Africa: Dampening the storm? *Science*, 369(6504):624–626, 2020.

[239] Christopher Adam, Mark Henstridge, and Stevan Lee. After the lockdown: macroeconomic adjustment to the COVID-19 pandemic in sub-Saharan Africa. *Oxford Review of Economic Policy*, 36(Supplement_1):S338–S358, 2020.

[240] The World Bank. WORLD DEVELOPMENT INDICATORS, 2021. URL https://datatopics.worldbank.org/world-development-indicators/.

[241] United Nations. 2019 Revision of World Population Prospects, 2019. URL https://population.un.org/wpp/.

[242] Institut National de la statistique et d'analyse économique du Bénin (INSAE). Recensement General de la Population et de l'Habitation 1er Edition. Report, 1979.

[243] Institut National de la statistique et d'analyse économique du Bénin (INSAE). Recensement General de la Population et de l'Habitation 2eme Edition. Report, 1992.

[244] Institut National de la statistique et d'analyse économique du Bénin (INSAE). Quatrième Recensement Général de la Population et de l'Habitation . Report, 2013.

[245] Institut National de la statistique et d'analyse économique du Bénin (INSAE). Recensement General de la Population et de l'Habitation 3eme Edition. Report, 2002.

[246] Anignikin V Padonou, Afolabi RI Biaou Olayé, and Benjamin Allagbé. Analysis of the Competitiveness Factors of the Port of Cotonou in the Economic Growth of Benin. 2019.

[247] Mohamed Nasser Baco. Ethno-ecological Variability in the Consumption of Leafy Green Plants in the Republic of Benin. *Asian Journal of Agricultural Extension, Economics & Sociology*, pages 1–15, 2019.

[248] Karim Okanla. Competitive squall at the pier, 2019. URL https://www.dandc.eu/en/article/autonomous-port-cotonou-benins-economic-engine-facing-major-new-challenge-neighbouring. Accessed: 2021-08-11.

[249] J Joost Beuving. Lebanese traders in Cotonou: a socio-cultural analysis of economic mobility and capital accumulation. *Africa*, 76(3):324–351, 2006.

[250] International Monetary Fund. World Economic Outlook Database, 2021. URL https://www.imf.org/en/Publications/WEO/weo-database/2021/April. Accessed: 2021-08-11.

[251] M. De Bruyere Bigot A., I. Zohoun. *Premiers cas de seropositivite anti-HIV-1 au Benin*, volume 16 of *La Press Medicale*. 1987.

[252] Way P Boyle Torrey B, Mulligan M. *Blood donors and AIDS in Africa: The gift relationship revisited*. US Bureau of the Census, 1990.

[253] UNAIDS; Ministry of Health Benin; CNLS. Rapport de suivi de la déclaration de politique sur le VIH/SIDA au Bénin 2016 (Country progress report, Benin). Report, 2016.

[254] M. Alary, L. Mukenge-Tshibaka, F. Bernier, N. Geraldo, C. M. Lowndes, H. Meda, C. A. Gnintoungbe, S. Anagonou, and J. R. Joly. Decline in the prevalence of HIV and sexually transmitted diseases among female sex workers in Cotonou, Benin, 1993-1999. *AIDS*, 16(3):463–70, 2002. ISSN 0269-9370 (Print) 0269-9370 (Linking). URL https://www.ncbi.nlm.nih.gov/pubmed/11834959.

[255] CM Lowndes, M Alary, H Meda, CAB Gnintoungbe, L Mukenge-Tshibaka, C Adjovi, A Buve, L Morison, M Laourou, L Kanhonou, et al. Role of core and bridging groups in the transmission dynamics of HIV and STIs in Cotonou, Benin, West Africa. *Sexually transmitted infections*, 78(suppl 1):i69–i77, 2002.

[256] Ministry of Health Benin. Enquête de surveillance de deuxième génération relative aux IST, VIH et SIDA au Bénin (ESDG-2017): Travailleuses de Sexe, Serveuses de Bar et Restaurants & Clients PS. Report, 2017.

[257] Projet Équité en Santé Sexuelle et Santé. Cartographie des Travailleuses de Sexe des communes de Cotonou, Abomey-Calavi, Sèmè-Podji et Parakou: Rapport Synthèse du Mapping, Benin, 2013-2014. Report.

[258] Ministry of Health Benin. Enquête de surveillance de deuxième génération des IST/VIH/SIDA au Bénin (ESDG-2012): Travailleuses du sexe et Serveuses de bars/restaurants. Report, 2012.

[259] Ministry of Health Benin. Enquête de surveillance de deuxième génération relative aux IST, VIH et SIDA au Bénin (ESDG-2015): Professionnelles de Sexe & Serveuses de Bar et Restaurants. Report, 2015.

[260] Abdullahi Garba Kangiwa. The socio–economic factors and effects of prostitution in Nigeria. *European Journal of Research in Social Sciences Vol*, 3(5), 2015.

[261] Maria Huberta Kroone. Prostitution in Benin, Benin (West-Africa). URL http://www.aktiebenin.nl/userfiles/file/Papers/prostitution%20in%20Benin.pdf. Accessed: 2021-08-11.

[262] L. Behanzin, S. Diabate, I. Minani, M. C. Boily, A. C. Labbe, C. Ahoussinou, S. Anagonou, D. M. Zannou, C. M. Lowndes, and M. Alary. Decline in the prevalence of HIV and sexually transmitted infections among female sex workers in Benin over 15 years of targeted interventions. *J Acquir Immune Defic Syndr*, 63 (1):126–34, 2013. ISSN 1944-7884 (Electronic) 1525-4135 (Linking). doi: 10.1097/QAI.0b013e318286b9d4. URL https://www.ncbi.nlm.nih.gov/pubmed/23337368.

[263] M Alary, L Béhanzin, F Guédou, E Goma Matsétsé, M Olodo, M Aza-Gnandji, A Dossouvo, F Kintin, A Akpaka, E Chagas, et al. O13. 4 Prevalence of gonorrhea, chlamydia and syphilis among men who have sex with men initiating HIV pre-exposure prophylaxis in Cotonou, Benin, 2021.

[264] Kyle T Bernstein, Julia L Marcus, Giuliano Nieri, Susan S Philip, and Jeffrey D Klausner. Rectal gonorrhea and chlamydia reinfection is associated with increased risk of HIV seroconversion. *JAIDS Journal of Acquired Immune Deficiency Syndromes*, 53(4):537–543, 2010.

[265] Ministry of Health Benin. Enquête de surveillance de deuxième génération des IST/VIH/SIDA au Bénin (ESDG-2012): Camionneurs et Clients des Travailleuses du sexe. Report, 2012.

[266] Michel Alary and Catherine M Lowndes. The central role of clients of female sex workers in the dynamics of heterosexual HIV transmission in sub-Saharan Africa. *AIDS*, 18(6):945–947, 2004.

[267] Bertran Auvert, Anne Buve, Emmanuel Lagarde, M Kahindo, J Chege, N Rutenberg, R Musonda, M Laourou, E Akam, HA Weiss, et al. Male circumcision and HIV infection in four cities in sub-Saharan Africa. *AIDS*, 15:S31–S40, 2001.

[268] UNAIDS; PNLS. Rapport de suivi de la déclaration de politique sue le VIH/SIDA au Bénin 2010 (Country progress report, Benin). Report, 2010.

[269] United States Agency for International Development. Assessment of HIV testing policy and practices in Benin. Report, 2005. URL https://pdf.usaid.gov/pdf_docs/Pnade810.pdf.

[270] PNLS Ministry of Health Benin. Politique, normes et procedures pour la prise en charge des personnes vivant avec le VIH. Report, 2012. URL http://www.grandir.sidaction.org/wp-content/uploads/2015/06/Normes-et-Procedures-VIH-Benin_2012.pdf.

[271] Dovonou Albert Comlan, Alassani Adébayo Cossi, Attinsounon Cossi Angelo, Saké Kadidjatou, Adè Serge, Ahoui Séraphin, Degla Jivather, Tchégnonsi Tognon Francis, Zannou Djimon Marcel, Adè Gabriel, and Houngbé Fabien. Clinical and immunological characteristics in HIV-infected patients at the treatment initiation at the University Hospital of Parakou (Benin). *Open Journal of Immunology*, 7:51–58, 2017.

[272] Myron S Cohen, M Kumi Smith, Kathryn E Muessig, Timothy B Hallett, Kimberly A Powers, and Angela D Kashuba. Antiretroviral treatment of HIV-1 prevents transmission of HIV-1: where do we go from here? *The Lancet*, 382(9903):1515–1524, 2013.

[273] M. S. Cohen, Y. Q. Chen, M. McCauley, T. Gamble, M. C. Hosseinipour, N. Kumarasamy, J. G. Hakim, J. Kumwenda, B. Grinsztejn, J. H. Pilotto, S. V. Godbole, S. Chariyalertsak, B. R. Santos, K. H. Mayer, I. F. Hoffman, S. H. Eshleman, E. Piwowar-Manning, L. Cottle, X. C. Zhang, J. Makhema, L. A. Mills, R. Panchia, S. Faesen, J. Eron, J. Gallant, D. Havlir, S. Swindells, V. Elharrar, D. Burns, T. E. Taha, K. Nielsen-Saines, D. D. Celentano, M. Essex, S. E. Hudelson, A. D. Redd, T. R. Fleming, and Hptn Study Team. Antiretroviral therapy for the prevention of HIV-1 transmission. *N Engl J Med*, 375(9):830–9, 2016. ISSN 1533-4406 (Electronic) 0028-4793 (Linking). doi: 10.1056/NEJMoa1600693. URL https://www.ncbi.nlm.nih.gov/pubmed/27424812.

[274] Aminata Mboup, Luc Béhanzin, Fernand Guédou, Katia Giguère, Nassirou Geraldo, Djimon Marcel Zannou, René K Kêkê, Moussa Bachabi, Flore Gangbo, Dissou Affolabi, et al. Comparison of adherence measurement tools used in a pre-exposure prophylaxis demonstration study among female sex workers in Benin. *Medicine*, 99 (21), 2020.

[275] Aminata Mboup, Souleymane Diabaté, Luc Béhanzin, Fernand Aimé Guédou, Djimon Marcel Zannou, René K Kêkê, Moussa Bachabi, Flore Gangbo, Mark A Marzinke, Craig Hendrix, et al. Determinants of HIV Preexposure Prophylaxis Adherence Among Female Sex Workers in a Demonstration Study in Cotonou, Benin: A Study of Behavioral and Demographic Factors. *Sexually Transmitted Diseases*, 48(8):565, 2021.

[276] Lily Geidelberg, Kate M Mitchell, Michel Alary, Aminata Mboup, Luc Béhanzin, Fernand Guédou, Nassirou Geraldo, Ella Goma-Matsétsé, Katia Giguère, Marlène Aza-Gnandji, et al. Mathematical Model Impact Analysis of a Real-Life Pre-exposure Prophylaxis and Treatment-As-Prevention Study Among Female Sex Workers in Cotonou, Benin. *Journal of Acquired Immune Deficiency Syndromes*, 86(2):e28, 2021.

[277] UNAIDS. "Spectrum model", 2019. URL https://aidsinfo.unaids.org/.

[278] Jacob Levi, Alice Raymond, Anton Pozniak, Pietro Vernazza, Philipp Kohler, and Andrew Hill. Can the UNAIDS 90-90-90 target be achieved? A systematic analysis of national HIV treatment cascades. *BMJ global health*, 1(2):e000010, 2016.

[279] P Vickerman, C Watts, RW Peeling, D Mabey, and M Alary. Modelling the cost effectiveness of rapid point of care diagnostic tests for the control of HIV and other sexually transmitted infections among female sex workers. *Sexually transmitted infections*, 82(5):403–412, 2006.

[280] L. Behanzin, S. Diabate, I. Minani, C. M. Lowndes, M. C. Boily, A. C. Labbe, S. Anagonou, D. M. Zannou, A. Buve, and M. Alary. Decline in HIV prevalence among young men in the general population of Cotonou, Benin, 1998-2008. *PLoS One*, 7(8):e43818, 2012. ISSN 1932-6203 (Electronic) 1932-6203 (Linking). doi: 10.1371/journal.pone.0043818. URL https://www.ncbi.nlm.nih.gov/pubmed/22952773.

[281] Andrea Low, Nicolas Nagot, Issouf Konate, Nicolas Meda, Michel Segondy, Philippe Van de Perre, Philippe Mayaud, and Peter Vickerman. Potential impact of existing interventions and of antiretroviral use in female sex workers on transmission of HIV in Burkina Faso: a modeling study. *JAIDS Journal of Acquired Immune Deficiency Syndromes*, 68:S180–S188, 2015.

[282] Catherine M Lowndes, Michel Alary, Annie-Claude Labbe, Cyriaque Gnintoungbe, Michelyne Belleau, Leonard Mukenge, Honore Meda, Marguerite Ndour, Severin Anagonou, and Alphonse Gbaguidi. Interventions among male clients of female sex workers in Benin, West Africa: an essential component of targeted HIV preventive interventions. *Sexually transmitted infections*, 83(7):577–581, 2007.

[283] Ministry of Health Benin. Enquête de surveillance de deuxième génération relative aux IST, VIH et SIDA au Bénin (ESDG-2015): Clients des TS, Camionneurs et Personnes Privées de Liberté. 2015.

[284] Deborah Donnell, Jared M Baeten, Namandjé N Bumpus, Justin Brantley, David R Bangsberg, Jessica E Haberer, Andrew Mujugira, Nelly Mugo, Patrick Ndase, Craig Hendrix, et al. HIV protective efficacy and correlates of tenofovir blood concentrations in a clinical trial of PrEP for HIV prevention. *Journal of acquired immune deficiency syndromes*, 66(3):340, 2014.

[285] Michael Stein. Large sample properties of simulations using latin hypercube sampling. *Technometrics*, 29(2):143–151, 1987.

[286] Supriya Sarkar, Phaedra Corso, Shideh Ebrahim-Zadeh, Patricia Kim, Sana Charania, and Kristin Wall. Cost-effectiveness of HIV prevention interventions in sub-Saharan Africa: a systematic review. *EClinicalMedicine*, 10:10–31, 2019.

[287] Shannon Hader. COVID-19 shows need for UN General Assembly push to end AIDS. *The Lancet HIV*, 2021.

[288] AVERT. PRE-EXPOSURE PROPHYLAXIS (PREP) FOR HIV PREVENTION, 2021. URL https://www.avert.org/professionals/hiv-programming/prevention/pre-exposure-prophylaxis. Accessed: 2021-08-11.

[289] Connie L Celum, Sinead Delany-Moretlwe, Margaret McConnell, Heidi Van Rooyen, Linda-Gail Bekker, Ann Kurth, Elizabeth Bukusi, Chris Desmond, Jennifer Morton, and Jared M Baeten. Rethinking HIV prevention to prepare for oral PrEP implementation for young African women. *Journal of the International AIDS Society*, 18:20227, 2015.

[290] Carlos F Cáceres, Kevin R O'Reilly, Kenneth H Mayer, and Rachel Baggaley. PrEP implementation: moving from trials to policy and practice, 2015.

[291] Katia Giguère, Luc Béhanzin, Fernand A Guédou, Denis Talbot, François A Leblond, Ella Goma-Matsétsé, Djimon M Zannou, Dissou Affolabi, René K Kêkê, and Flore Gangbo. PrEP use among female sex workers: no evidence for risk compensation. *JAIDS Journal of Acquired Immune Deficiency Syndromes*, 82(3):257–264, 2019. ISSN 1525-4135.

[292] HIV-Causal Collaboration et al. The effect of combined antiretroviral therapy on the overall mortality of HIV-infected individuals. *AIDS (London, England)*, 24(1):123, 2010.

[293] R Landovitz et al. HPTN 083 Interim results: Pre-exposure prophylaxis (PrEP) 376 containing long-acting injectable cabotegravir (CAB-LA) is safe and highly effective for 377 cisgender men and transgender women who have sex with men (MSM, TGW). *New Engl J Med*, 2021.

[294] Linda-Gail Bekker, Leigh Johnson, Frances Cowan, Cheryl Overs, Donela Besada, Sharon Hillier, and Willard Cates Jr. Combination HIV prevention for female sex workers: what is the evidence? *The Lancet*, 385(9962):72–87, 2015.

[295] Richard Cash and Vikram Patel. Has COVID-19 subverted global health? *The Lancet*, 395(10238):1687–1688, 2020.

[296] Piya Majumdar, Ankita Biswas, and Subhashis Sahu. COVID-19 pandemic and lockdown: cause of sleep disruption, depression, somatic pain, and increased screen exposure of office workers and students of India. *Chronobiology international*, 37(8): 1191–1200, 2020.

[297] Stephen X Zhang, Yifei Wang, Andreas Rauch, and Feng Wei. Unprecedented disruption of lives and work: Health, distress and life satisfaction of working adults in China one month into the COVID-19 outbreak. *Psychiatry research*, 288:112958, 2020.

[298] Thanassis Karalis. Planning and evaluation during educational disruption: Lessons learned from Covid-19 pandemic for treatment of emergencies in education. *European Journal of Education Studies*, 2020.

[299] Immanuel Debbarma and Tabitha Durai. Educational disruption: Impact of COVID-19 on students from the Northeast states of India. *Children and youth services review*, 120:105769, 2021.

[300] Christopher M Jones, Ganesh Radhakrishna, Katharine Aitken, John Bridgewater, Pippa Corrie, Martin Eatock, Rebecca Goody, Paula Ghaneh, James Good, Derek Grose, et al. Considerations for the treatment of pancreatic cancer during the COVID-19 pandemic: the UK consensus position. *British journal of cancer*, 123(5):709–713, 2020.

[301] Matthew D Rutter, Matthew Brookes, Thomas J Lee, Peter Rogers, and Linda Sharp. Impact of the COVID-19 pandemic on UK endoscopic activity and cancer detection: a National Endoscopy Database Analysis. *Gut*, 70(3):537–543, 2021.

[302] Keith Alcorn and Gus Cairns. Two UK studies find that HIV infection may be a risk factor for dying from COVID-19, 2020. URL https://www.aidsmap.com/news/aug-2020/two-uk-studies-find-hiv-infection-may-be-risk-factor-dying-covid-19. Accessed: 2021-08-11.

[303] Sandeep Prabhu, Selvamuthu Poongulali, and Nagalingeswaran Kumarasamy. Impact of COVID-19 on people with HIV: a review. *Journal of Virus Eradication*, page 100019, 2020.

[304] Centers for Disease Control, Prevention, et al. Ten ways healthcare systems can operate effectively during the COVID-19 pandemic. 2020. URL https://www.cdc.gov/coronavirus/2019-ncov/hcp/ways-operate-effectively.html. Accessed: 2021-08-11.

[305] Hongbo Jiang, Yi Zhou, and Weiming Tang. Maintaining HIV care during the COVID-19 pandemic. *The Lancet HIV*, 7(5):e308–e309, 2020.

[306] Britta L Jewell, Edinah Mudimu, John Stover, Debra Ten Brink, Andrew N Phillips, Jennifer A Smith, Rowan Martin-Hughes, Yu Teng, Robert Glaubius, Severin Guy Mahiane, et al. Potential effects of disruption to HIV programmes in sub-Saharan Africa caused by COVID-19: results from multiple mathematical models. *The Lancet HIV*, 7(9):e629–e640, 2020.

[307] Gowri Nagendra, Caroline Carnevale, Natalie Neu, Alwyn Cohall, and Jason Zucker. The potential impact and availability of sexual health services during the COVID-19 pandemic. *Sexually transmitted diseases*, 47(7):434, 2020.

[308] Thabani Nyoni and Moses Okumu. COVID-19-compliant strategies for supporting treatment adherence among people living with HIV in sub-Saharan Africa. *AIDS and Behavior*, 24(9):2473–2476, 2020.

[309] Solomon Mukwenha, Tafadzwa Dzinamarira, Owen Mugurungi, and Godfrey Musuka. Maintaining robust HIV and tuberculosis services in the COVID-19 era: A public health dilemma in Zimbabwe. *International Journal of Infectious Diseases*, 100: 394–395, 2020.

[310] KENPHIA. KENPHIA 2018 PRELIMINARY REPORT, 2018. URL https://www.health.go.ke/wp-content/uploads/2020/02/KENPHIA-2018-PREL-REP-2020-HR3-final.pdf.

[311] Francesca Odhiambo, Edwin Mulwa, James Ayieko, Jayne Kulzer, Maurice Aluda, Purba Chatterjee, Raphael Onyango, George Nyanaro, Elizabeth A Bukusi, and Craig R Cohen. Implementation of HIV care in Western Kenya during corona virus disease 2019 response. *AIDS*, 34(8):F1–F2, 2020.

[312] Matthew Ponticiello, Juliet Mwanga-Amumpaire, Patricia Tushemereirwe, Gabriel Nuwagaba, Rachel King, and Radhika Sundararajan. "Everything is a mess": how COVID-19 is impacting engagement with HIV testing services in rural Southwestern Uganda. *AIDS and Behavior*, 24:3006–3009, 2020.

[313] World Health Organisation. Disruption in HIV, Hepatitis and STI services due to COVID-19, 2020. Accessed: 2021-08-11.

[314] Times Live Zambia. Almost 11,000 HIV-positive patients in Gauteng have skipped ARV collection during lockdown, 2020. URL https://www.timeslive.co.za/news/south-africa/2020-05-19-almost-11000-hiv-positive-patients-in-gauteng-have-skipped-arv-collection-during-lockdown/. Accessed: 2021-08-11.

[315] Claire Keene, Erika Mohr-Holland, Tali Cassidy, Vera Scott, Aurelie Nelson, Jennifer Furin, and Laura Triviño-Duran. How COVID-19 could benefit tuberculosis and HIV services in South Africa. *The Lancet Respiratory Medicine*, 8(9):844–846, 2020.

[316] Mohamed A Hammoud, Lisa Maher, Martin Holt, Louisa Degenhardt, Fengyi Jin, Dean Murphy, Benjamin Bavinton, Andrew Grulich, Toby Lea, Bridget Haire, et al. Physical distancing due to COVID-19 disrupts sexual behaviors among gay and bisexual men in Australia: implications for trends in HIV and other sexually transmissible infections. *JAIDS Journal of Acquired Immune Deficiency Syndromes*, 85(3):309–315, 2020.

[317] Romain Silhol, Lily Geidelberg, Kate M Mitchell, Sharmistha Mishra, Dobromir Dimitrov, Anna Bowring, Luc Béhanzin, Fernand Guédou, Souleymane Diabaté, Sheree Schwartz, et al. Assessing the potential impact of disruptions due to COVID-19 on HIV among key and lower-risk populations in the largest cities of Cameroon and Benin. *Journal of Acquired Immune Deficiency Syndromes*, 87(3):899, 2021.

[318] Joshua Kimani, Joyce Adhiambo, Rosemary Kasiba, Peninah Mwangi, Veronica Were, John Mathenge, Pascal Macharia, Francois Cholette, Samantha Moore, Souradet Shaw, et al. The effects of COVID-19 on the health and socio-economic security of sex workers in Nairobi, Kenya: Emerging intersections with HIV. *Global Public Health*, 15(7):1073–1082, 2020.

[319] WQ Toh, C Treibich, S Szawlowski, H Cust, EHA Mbaye, K Gueye, CT Ndour, and A Lépine. How COVID-19 impacted risky sexual behaviours of female sex workers in Senegal and its wider implications: Evidence from a cohort study.

[320] Carrie E Lyons, Ashley Grosso, Fatou M Drame, Sosthenes Ketende, Daouda Diouf, Ibrahima Ba, Kate Shannon, Rebecca Ezouatchi, Amara Bamba, Abo Kouame, et al. Physical and sexual violence affecting female sex workers in Abidjan, Côte d'Ivoire: Prevalence, and the relationship with the work environment, HIV and access to health services. *Journal of acquired immune deficiency syndromes*, 75(1):9, 2017.

[321] Abebe Aemro Selassie. Sub-Saharan Africa's Difficult Road to Recovery. URL https://blogs.imf.org/2020/10/22/sub-saharan-africas-difficult-road-to-recovery/,year=2020. Accessed: 2021-08-11.

[322] J Jeffery Reeves, Hannah M Hollandsworth, Francesca J Torriani, Randy Taplitz, Shira Abeles, Ming Tai-Seale, Marlene Millen, Brian J Clay, and Christopher A Longhurst. Rapid response to COVID-19: health informatics support for outbreak management in an academic health system. *Journal of the American Medical Informatics Association*, 27(6):853–859, 2020.

[323] Peter F Rebeiro, Stephany N Duda, Kara K Wools-Kaloustian, Denis Nash, Keri N Althoff, et al. Implications of COVID-19 for HIV research: data sources, indicators and longitudinal analyses. *Journal of the International AIDS Society*, 23(10), 2020.

[324] Reuters. More African countries confirm first coronavirus cases as Jack Ma pledges aid, 2020. URL https://www.reuters.com/article/us-health-coronavirus-africa/somalia-liberia-benin-and-tanzania-confirm-first-coronavirus-cases-idUSKBN2131IA. Accessed: 2021-08-11.

[325] Carl AB Pearson, Cari Van Schalkwyk, Anna M Foss, Kathleen M O'Reilly, Juliet RC Pulliam, CMMID COVID-19 working group, et al. Projected early spread of COVID-19 in Africa through 1 June 2020. *Eurosurveillance*, 25(18):2000543, 2020.

[326] World Health Organisation. Le Bénin renforce sa riposte à la COVID-19 en augmentant le dépistage, 2020. URL https://www.afro.who.int/fr/news/le-benin-renforce-sa-riposte-la-covid-19-en-augmentant-le-depistage. Accessed: 2021-08-11.

[327] Elijah Paintsil et al. COVID-19 threatens health systems in sub-Saharan Africa: the eye of the crocodile. *The journal of clinical investigation*, 130(6):2741–2744, 2020.

[328] Gouvernment du Bénin. CORONAVIRUS - Communiqué du Ministre de l'intérieur au sujet du Cordon Sanitaire, 2020. URL https://www.gouv.bj/actualite/571/coronavirus---communique-ministre-interieur-sujet-cordon-sanitaire/. Accessed: 2021-08-11.

[329] africaradio. CORONAVIRUS : LE BÉNIN N'A PAS LES "MOYENS" DE CONFINER SA POPULATION, 2020. URL https://www.africaradio.com/news/coronavirus-le-benin-n-a-pas-les-moyens-de-confiner-sa-population-165962. Accessed: 2021-08-11.

[330] Gouvernment du Bénin. Mesures de riposte - Covid-19, 2020. URL https://www.gouv.bj/coronavirus/mesures/. Accessed: 2021-08-11.

[331] Emmanuel Bonnet, Oriane Bodson, Frédéric Le Marcis, Adama Faye, NE Sambieni, Florence Fournet, Florence Boyer, Abdourahmane Coulibaly, Kadidiatou Kadio, FB Diongue, et al. The COVID-19 pandemic in francophone West Africa: from the first cases to responses in seven countries. *BMC public health*, 21(1):1–17, 2021.

[332] UNICEF. Benin: COVID-19 Situation Report – 15, 2020. URL https://reliefweb.int/report/benin/benin-covid-19-situation-report-15. Accessed: 2021-08-11.

[333] European Investment Bank. RIPOSTE SANITAIRE COVID-19 BENIN, 2020. URL https://www.eib.org/en/projects/pipelines/all/20200624. Accessed: 2021-08-11.

[334] Fernand Guédou et al. Rapport sur le volet impact de COVID-19 sur la prévention de VIH : extrait de l'enquête évaluative de l'utilisation de l'autotest VIH salivaire chez les TS à Cotonou et environs, 2020.

[335] Marianne Boisvert Moreau, Frédéric D Kintin, Septime Atchekpe, Georges Batona, Luc Béhanzin, Fernand A Guédou, Marie-Pierre Gagnon, and Michel Alary. HIV Self-Testing Implementation, Distribution and Use Among Female Sex Workers in Cotonou, Benin: A Qualitative Evaluation of Acceptability and Feasibility. 2021.

[336] Luc Béhanzin, Souleymane Diabaté, Isaac Minani, Catherine M Lowndes, Marie-Claude Boily, Annie-Claude Labbé, Séverin Anagonou, Djimon Marcel Zannou, Anne Buvé, and Michel Alary. Assessment of HIV-related risky behaviour: a comparative study of face-to-face interviews and polling booth surveys in the general population of Cotonou, Benin. *Sexually transmitted infections*, 89(7):595–601, 2013.

[337] Alexandra B Hogan, Britta L Jewell, Ellie Sherrard-Smith, Juan F Vesga, Oliver J Watson, Charles Whittaker, Arran Hamlet, Jennifer A Smith, Peter Winskill, Robert Verity, et al. Potential impact of the COVID-19 pandemic on HIV, tuberculosis, and malaria in low-income and middle-income countries: a modelling study. *The Lancet Global Health*, 8(9):e1132–e1141, 2020.

[338] Najmul Haider, Abdinasir Yusuf Osman, Audrey Gadzekpo, George O Akipede, Danny Asogun, Rashid Ansumana, Richard John Lessells, Palwasha Khan, Muzamil Mahdi Abdel Hamid, Dorothy Yeboah-Manu, et al. Lockdown measures in response to COVID-19 in nine sub-Saharan African countries. *BMJ Global health*, 5(10):e003319, 2020.

[339] Issideen Ayinla Osseni. COVID-19 pandemic in sub-Saharan Africa: preparedness, response, and hidden potentials. *Tropical medicine and health*, 48(1):1–3, 2020.

[340] Marguerite Massinga Loembé, Akhona Tshangela, Stephanie J Salyer, Jay K Varma, Ahmed E Ogwell Ouma, and John N Nkengasong. COVID-19 in Africa: the spread and response. *Nature Medicine*, 26(7):999–1003, 2020.

[341] UNAIDS. Sex workers must not be left behind in the response to COVID-19, 2020. URL https://www.unaids.org/en/resources/presscentre/pressreleaseandstatementarchive/2020/april/20200408_sex-workers-covid-19. Accessed: 2021-08-11.

[342] Ministry of Health Benin. Rapport de l'audit de la file active des personnes vivant avec le VIH au Bénin (Audit report). Report, 2017.

[343] Floriano Amimo, Ben Lambert, and Anthony Magit. What does the COVID-19 pandemic mean for HIV, tuberculosis, and malaria control? *Tropical medicine and health*, 48(1):1–4, 2020.

[344] Jim Todd, Judith R. Glynn, Milly Marston, Tom Lutalo, Sam Biraro, Wambura Mwita, Vinai Suriyanon, Ram Rangsin, Kenrad E. Nelson, Pam Sonnenberg, Dan Fitzgerald, Etienne Karita, and Basia Żaba. Time from HIV seroconversion to death: a collaborative analysis of eight studies in six low and middle-income countries before highly active antiretroviral therapy. *AIDS*, 21(Suppl 6):S55–S63, 2007. ISSN 0269-9370. doi: 10.1097/01.aids.0000299411.75269.e8. URL http://dx.doi.org/10.1097/01.aids.0000299411.75269.e8.

[345] Judith R. Glynn, Pam Sonnenberg, Gill Nelson, Andre Bester, Stuart Shearer, and Jill Murray. Survival from HIV-1 seroconversion in Southern Africa: a retrospective cohort study in nearly 2000 gold-miners over 10 years of follow-up. *AIDS*, 21(5): 625–632, 2007. ISSN 0269-9370. doi: 10.1097/qad.0b013e328017f857. URL http://dx.doi.org/10.1097/qad.0b013e328017f857.

[346] L. Lavreys, J. M. Baeten, V. Chohan, R. S. McClelland, W. M. Hassan, B. A. Richardson, K. Mandaliya, J. O. N. Achola, and J. Overbaugh. Higher set point plasma viral load and more-severe acute HIV type 1 (HIV-1) illness predict mortality among high-risk HIV-1-infected African Women. *Clinical Infectious Diseases*, 42 (9):1333–1339, 2006. ISSN 1058-4838 1537-6591. doi: 10.1086/503258. URL http://dx.doi.org/10.1086/503258.

[347] Margaret May, Robin Wood, Landon Myer, Patrick Taffé, Andri Rauch, Manuel Battegay, and Matthias Egger. CD4+T cell count decreases by ethnicity among untreated patients with HIV infection in South Africa and Switzerland. *The Journal of Infectious Diseases*, 200(11):1729–1735, 2009. ISSN 0022-1899 1537-6613. doi: 10.1086/648096. URL http://dx.doi.org/10.1086/648096.

[348] Nikos Pantazis, Charles Morrison, Pauli N. Amornkul, Charlotte Lewden, Robert A. Salata, Albert Minga, Tsungai Chipato, Harold Jaffe, Shabir Lakhi, Etienne Karita, Kholoud Porter, Laurence Meyer, and Giota Touloumi. Differences in HIV natural history among African and non-African seroconverters in Europe and seroconverters in Sub-Saharan Africa. *PLoS ONE*, 7(3):e32369, 2012. ISSN 1932-6203. doi: 10.1371/journal.pone.0032369. URL http://dx.doi.org/10.1371/journal.pone.0032369.

[349] Muhammad Bakari, Willy Urassa, Fred Mhalu, Gunnel Biberfeld, Kisali Pallangyo, and Eric Sandström. Slow progression of HIV-1 infection in a cohort of antiretroviral naïve hotel workers in Dar es Salaam, Tanzania as defined by their CD4 cell slopes. *Scandinavian Journal of Infectious Diseases*, 40(5):407–413, 2008. ISSN 0036-5548 1651-1980. doi: 10.1080/00365540701708285. URL http://dx.doi.org/10.1080/00365540701708285.

[350] Global HIV Prevention Coalition. Preventing HIV infections at the time of a new pandemic, 2021. URL https://www.unaids.org/en/resources/documents/2021/20210701_HIVPrevention_new_pandemic. Accessed: 2021-08-11.

[351] S/GAC Chair, Fatuma Y. Sanneh and PPM, Diana L. Huestis. INFORMATION MEMO FOR AMBASSADOR STEPHANIE SULLIVAN, WEST AFRICAREGION, 2021. URL https://www.state.gov/wp-content/uploads/2021/01/ROP21_PLL_WestAfricaRegion_01.12.2021.pdf. Accessed: 2021-08-11.

[352] Shirley Lee Lecher, Mary Naluguza, Christina Mwangi, Jonathan N'tale, Dianna Edgil, George Alemnji, and Heather Alexander. Notes from the Field: Impact of the COVID-19 Response on Scale-Up of HIV Viral Load Testing—PEPFAR-Supported Countries, January–June 2020. *Morbidity and Mortality Weekly Report*, 70(21):794, 2021.

[353] UNAIDS. COVID-19's impact on HIV treatment less severe than feared, 2020. URL https://www.unaids.org/en/resources/presscentre/featurestories/2020/october/20201016_covid-impact-on-hiv-treatment-less-severe-than-feared. Accessed: 2021-08-11.

[354] Sebastian Contreras and Viola Priesemann. Risking further COVID-19 waves despite vaccination. *The Lancet Infectious Diseases*, 21(6):745–746, 2021.

[355] World Health Organization et al. *Consolidated guidelines on HIV prevention, diagnosis, treatment and care for key populations*. World Health Organization, 2016.

[356] Barbara A Friedland, Laurel Sprague, Laura Nyblade, Stefan D Baral, Julie Pulerwitz, Ann Gottert, Ugo Amanyeiwe, Alison Cheng, Christoforos Mallouris, Florence Anam, et al. Measuring intersecting stigma among key populations living with hiv: implementing the people living with hiv stigma index 2.0. *Journal of the International AIDS Society*, 21(Suppl Suppl 5), 2018.

[357] Meredith E Clement, Ryan Kofron, and Raphael J Landovitz. Long-acting injectable cabotegravir for the prevention of HIV infection. *Current Opinion in HIV and AIDS*, 15(1):19, 2020.

[358] Anthony S Fauci. An HIV vaccine is essential for ending the HIV/AIDS pandemic. *Jama*, 318(16):1535–1536, 2017.

[359] World Health Organization. Archived: WHO Timeline - COVID-19, 2021. URL https://www.who.int/news/item/27-04-2020-who-timeline---covid-19. Accessed: 2021-07-16.

[360] World Health Organisation. Novel Coronavirus(2019-nCoV) Situation Report – 1 , 2020. URL https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200121-sitrep-1-2019-ncov.pdf. Accessed: 2021-08-11.

[361] World Health Organization. Naming the coronavirus disease (COVID-19) and the virus that causes it.

[362] World Health Organisation. Novel Coronavirus(2019-nCoV) Situation Report – 22 , 2020. URL https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200211-sitrep-22-ncov.pdf?sfvrsn=fb6d49b1_2. Accessed: 2021-08-11.

[363] Qingxiu Wang, Shiliang Zheng, Wei Tan, Li Qi, Weiguang Shao, Ming Zhang, and Xiaodong Sun. Epidemiology and clinical characteristics of 43 COVID-19 patients in Weifang, China. *Annals of Palliative Medicine*, 2020.

[364] Lily Geidelberg, Olivia Boyd, David Jorgensen, Igor Siveroni, Fabrícia F Nascimento, Robert Johnson, Manon Ragonnet-Cronin, Han Fu, Haowei Wang, Xiaoyue Xi, et al. Genomic epidemiology of a densely sampled COVID-19 outbreak in China. *Virus evolution*, 7(1):veaa102, 2021.

[365] Hongshan Mao. "Weifang City announces fever clinics". *Weifang News Network*, January 2020.

[366] ZhiXia Gao, Jian Zhou, YanYan Jiang, LuFeng Xu, Lin Sun, ShaoQi Yu, ZiLiang Fan, and JunJie Fan. Analysis of family cluster outbreak of coronavirus disease 2019 caused by nosocomial infection. *China Tropical Medicine*, pages 673–675, 2020.

[367] Erik Volz, Marc Baguelin, Sangeeta Bhatia, Adhiratha Boonyasiri, Anne Cori, Zulma Cucunubá, Gina Cuomo-Dannenburg, Christl A Donnelly, Ilaria Dorigatti, Rich Fitzjohn, Han Fu, Katy Gaythorpe, Azra Ghani, Arran Hamlet, Wes Hinsley, Natsuko Imai, Daniel Laydon, Lucy Nedjati-Gilani, Gemma abd Okell, Steven Riley, Sabine

van Elsland, Haowei Wang, Yuanrong Wang, Xiaoyue Xi, and Neil M Ferguson. "Report 5: Phylogenetic analysis of SARS-CoV-2", 2020.

[368] Matt J Keeling and Pejman Rohani. *"Modeling Infectious Diseases in Humans and Animals"*. Princeton University Press, September 2011.

[369] J O Lloyd-Smith, S J Schreiber, P E Kopp, and W M Getz. "Superspreading and the effect of individual variation on disease emergence". *Nature*, 438(7066):355–359, November 2005.

[370] Lucy M Li, Nicholas C Grassly, and Christophe Fraser. "Quantifying Transmission Heterogeneity Using Both Pathogen Phylogenies and Incidence Time Series". *Mol. Biol. Evol.*, 34(11):2982–2995, November 2017.

[371] Akira Endo, Sam Abbott, Adam J Kucharski, Sebastian Funk, et al. Estimating the overdispersion in COVID-19 transmission using outbreak sizes outside China. *Wellcome Open Research*, 5(67):67, 2020.

[372] Kazutaka Katoh and Daron M Standley. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular biology and evolution*, 30(4):772–780, 2013.

[373] Stefan Elbe and Gemma Buckland-Merrett. Data, disease and diplomacy: GISAID's innovative contribution to global health. *Global Challenges*, 1(1):33–46, 2017.

[374] Bui Quang Minh, Heiko Schmidt, Olga Chernomor, Dominik Schrempf, Michael Woodhams, Arndt von Haeseler, and Robert Lanfear. IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era, November 2019.

[375] Remco Bouckaert, Timothy G Vaughan, Joelle Barido-Sottani, Sebastián Duchêne, Mathieu Fourment, Alexandra Gavryushkina, Joseph Heled, Graham Jones, Denise Kühnert, Nicola De Maio, et al. BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. *PLoS computational biology*, 15(4):e1006650, 2019.

[376] Erik M Volz and Igor Siveroni. "Bayesian phylodynamic inference with complex models". *PLoS Comput. Biol.*, 14(11):e1006546, November 2018.

[377] Guangchuang Yu, David K Smith, Huachen Zhu, Yi Guan, and Tommy Tsan-Yuk Lam. "ggtree : an R package for visualization and annotation of phylogenetic trees with their covariates and other associated data". *Methods Ecol. Evol.*, 8(1):28–36, January 2017.

[378] Wei-Jie Guan, Zheng-Yi Ni, Yu Hu, Wen-Hua Liang, Chun-Quan Ou, Jian-Xing He, Lei Liu, Hong Shan, Chun-Liang Lei, David S C Hui, Bin Du, Lan-Juan Li, Guang Zeng, Kwok-Yung Yuen, Ru-Chong Chen, Chun-Li Tang, Tao Wang, Ping-Yan Chen, Jie Xiang, Shi-Yue Li, Jin-Lin Wang, Zi-Jing Liang, Yi-Xiang Peng, Li Wei, Yong Liu, Ya-Hua Hu, Peng Peng, Jian-Ming Wang, Ji-Yang Liu, Zhong Chen, Gang Li, Zhi-Jian Zheng, Shao-Qin Qiu, Jie Luo, Chang-Jiang Ye, Shao-Yong Zhu, Nan-Shan Zhong, and China Medical Treatment Expert Group for Covid-19. "Clinical Characteristics of Coronavirus Disease 2019 in China". *N. Engl. J. Med.*, February 2020.

[379] Yousef Alimohamadi, Maryam Taghdir, and Mojtaba Sepandi. The estimate of the basic reproduction number for novel coronavirus disease (COVID-19): a systematic review and meta-analysis. *Journal of Preventive Medicine and Public Health*, 2020.

[380] Juanjuan Zhang, Maria Litvinova, Wei Wang, Yan Wang, Xiaowei Deng, Xinghui Chen, Mei Li, Wen Zheng, Lan Yi, Xinhua Chen, et al. Evolving epidemiology and transmission dynamics of coronavirus disease 2019 outside Hubei province, China: a descriptive and modelling study. *The Lancet Infectious Diseases*, 2020.

[381] Qing Nie, Xingguang Li, Wei Chen, Dehui Liu, Yingying Chen, Haitao Li, Dongying Li, Mengmeng Tian, Wei Tan, and Junjie Zai. Phylogenetic and phylodynamic analyses of SARS-CoV-2. *Virus Research*, page 198098, 2020.

[382] Stephen K Gire, Augustine Goba, Kristian G Andersen, Rachel S G Sealfon, Daniel J Park, Lansana Kanneh, Simbirie Jalloh, Mambu Momoh, Mohamed Fullah, Gytis Dudas, Shirlee Wohl, Lina M Moses, Nathan L Yozwiak, Sarah Winnicki, Christian B Matranga, Christine M Malboeuf, James Qu, Adrianne D Gladden, Stephen F Schaffner, Xiao Yang, Pan-Pan Jiang, Mahan Nekoui, Andres Colubri, Moinya Ruth Coomber, Mbalu Fonnie, Alex Moigboi, Michael Gbakie, Fatima K Kamara, Veronica Tucker, Edwin Konuwa, Sidiki Saffa, Josephine Sellu, Abdul Azziz Jalloh, Alice Kovoma, James Koninga, Ibrahim Mustapha, Kandeh Kargbo, Momoh Foday, Mohamed Yillah, Franklyn Kanneh, Willie Robert, James L B Massally, Sinéad B Chapman, James Bochicchio, Cheryl Murphy, Chad Nusbaum, Sarah Young, Bruce W Birren, Donald S Grant, John S Scheiffelin, Eric S Lander, Christian Happi, Sahr M Gevao, Andreas Gnirke, Andrew Rambaut, Robert F Garry, S Humarr Khan, and Pardis C Sabeti. "Genomic surveillance elucidates Ebola virus origin and transmission during the 2014 outbreak". *Science*, 345(6202):1369–1372, September 2014.

[383] Peng Zhou, Xing-Lou Yang, Xian-Guang Wang, Ben Hu, Lei Zhang, Wei Zhang, Hao-Rui Si, Yan Zhu, Bei Li, Chao-Lin Huang, Hui-Dong Chen, Jing Chen, Yun Luo, Hua Guo, Ren-Di Jiang, Mei-Qin Liu, Ying Chen, Xu-Rui Shen, Xi Wang, Xiao-Shuang Zheng, Kai Zhao, Quan-Jiao Chen, Fei Deng, Lin-Lin Liu, Bing Yan, Fa-Xian Zhan, Yan-Yi Wang, Geng-Fu Xiao, and Zheng-Li Shi. "A pneumonia outbreak associated with a new coronavirus of probable bat origin". *Nature*, February 2020.

[384] Gytis Dudas, Luiz Max Carvalho, Andrew Rambaut, and Trevor Bedford. MERS-CoV spillover at the camel-human interface. *Elife*, 7, April 2018.

[385] BBC. Coronavirus: How the UK dealt with its first Covid case, 2021. URL https://www.bbc.co.uk/news/uk-england-55622386. Accessed: 2021-04-09.

[386] gov.uk. GOV.UK Coronavirus (COVID-19) in the UK, 2021. URL https://coronavirus.data.gov.uk/details/cases. Accessed: 2021-04-14.

[387] Office for National Statistics. Coronavirus: a year like no other, 2021. URL https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/articles/coronavirusayearlikenoother/2021-03-15. Accessed: 2021-04-09.

[388] BBC. China coronavirus: UK tracing up to 2,000 Wuhan visitors, 2021. URL https://www.bbc.co.uk/news/uk-51232163. Accessed: 2021-04-09.

[389] gov.uk. Prime Minister's statement on coronavirus (COVID-19): 12 March 2020, 2021. URL https://www.gov.uk/government/speeches/pm-statement-on-coronavirus-12-march-2020. Accessed: 2021-04-09.

[390] gov.uk. Prime Minister's statement on coronavirus (COVID-19): 16 March 2020, 2021. URL https://www.gov.uk/government/speeches/pm-statement-on-coronavirus-16-march-2020. Accessed: 2021-04-09.

[391] Sir Simon Stevens (NHS Chief Executive). NEXT STEPS ON NHS RESPONSE TO COVID-19, 2021. URL https://www.england.nhs.uk/coronavirus/wp-content/uploads/sites/52/2020/03/urgent-next-steps-on-nhs-response-to-covid-19-letter-simon-stevens.pdf.

[392] BBC. Coronavirus: Teachers to estimate grades after exams cancelled, 2021. URL https://www.bbc.co.uk/news/education-51980831. Accessed: 2021-04-09.

[393] gov.uk. Prime Minister's statement on coronavirus (COVID-19): 23 March 2020, 2021. URL https://www.gov.uk/government/speeches/pm-address-to-the-nation-on-coronavirus-23-march-2020. Accessed: 2021-04-09.

[394] BBC. Coronavirus: Northern Italy quarantines 16 million people, 2021. URL https://www.bbc.co.uk/news/world-middle-east-51787238. Accessed: 2021-04-09.

[395] PHE. Preliminary investigation into COVID-19 exceedances in Leicester (June 2020), 2021. URL https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/897128/COVID-19_activity_Leicester_Final-report_010720_v3.pdf. Accessed: 2021-04-09.

[396] BBC. Leicester lockdown: 100 days of extra restrictions in city, 2021. URL https://www.bbc.co.uk/news/uk-england-leicestershire-54423723. Accessed: 2021-04-09.

[397] BBC. Coronavirus: Visiting people at home banned in parts of Northern England, 2020. URL https://www.bbc.co.uk/news/uk-53602362. Accessed: 2021-04-09.

[398] gov.uk. Local COVID alert levels: what you need to know, 2021. URL https://web.archive.org/web/20201014194215/http://gov.uk/guidance/local-covid-alert-levels-what-you-need-to-know. Accessed: 2021-04-09.

[399] BBC. Covid-19: PM announces four-week England lockdown, 2021. URL https://www.bbc.co.uk/news/uk-54763956. Accessed: 2021-04-09.

[400] gov.uk. Prime Minister announces Tier 4: 'Stay At Home' Alert Level in response to new COVID variant, 2021. URL https://www.gov.uk/government/news/prime-minister-announces-tier-4-stay-at-home-alert-level-in-response-to-new-covid-variant. Accessed: 2021-04-09.

[401] LBC. "New COVID-19 tiers: Matt Hancock confirms areas moved to Tier 2 and 3", 2020. URL https://www.lbc.co.uk/news/matt-hancock-announcement-new-covid-19-tiers-areas/. Accessed: 2021-04-09.

[402] BBC. Covid: England's third national lockdown legally comes into force, 2021. URL https://www.bbc.co.uk/news/uk-55554550. Accessed: 2021-04-09.

[403] BBC. Covid-19: UK records more than 1,000 daily virus deaths, 2021. URL https://www.bbc.co.uk/news/uk-55565254. Accessed: 2021-04-09.

[404] COVID-19 Genomics UK Consortium. UK launches whole genome sequence alliance to map spread of coronavirus, 2020. URL https://www.cogconsortium.uk/uk-launches-whole-genome-sequence-alliance-to-map-spread-of-coronavirus/. Accessed: 2021-03-18.

[405] Yuki Furuse. Genomic sequencing effort for SARS-CoV-2 by country during the pandemic. *International Journal of Infectious Diseases*, 103:305–307, 2021.

[406] Steven Riley, Kylie EC Ainslie, Oliver Eales, Caroline E Walters, Haowei Wang, Christina J Atchison, Peter Diggle, Deborah Ashby, Christl A Donnelly, Graham Cooke, et al. Transient dynamics of SARS-CoV-2 as England exited national lockdown. *medRxiv*, 2020.

[407] Troy Day, Sylvain Gandon, Sébastien Lion, and Sarah P Otto. On the evolutionary epidemiology of SARS-CoV-2. *Current Biology*, 30(15):R849–R857, 2020.

[408] Emma B Hodcroft, Moira Zuber, Sarah Nadeau, Katharine HD Crawford, Jesse D Bloom, David Veesler, Timothy G Vaughan, Iñaki Comas, Fernando González Candelas, Tanja Stadler, et al. Emergence and spread of a SARS-CoV-2 variant through Europe in the summer of 2020. *MedRxiv*, 2020.

[409] Public Health England. "PHE investigating a novel variant of COVID-19", 2020. URL https://www.gov.uk/government/news/phe-investigating-a-novel-variant-of-covid-19. Accessed: 2021-04-09.

[410] Public Health England. "Investigation of novel SARS-CoV-2 variant: Variant of Concern 202012/01", 2020. URL https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/959438/Technical_Briefing_VOC_SH_NJL2_SH2.pdf. Accessed: 2021-04-09.

[411] Andrew Rambaut, Nick Loman, Oliver Pybus, Wendy Barclay, Jeff Barrett, Alesandro Carabelli, Tom Connor, Tom Peacock, David L Robertson, and Erik Volz. "Preliminary genomic characterisation of an emergent SARS-CoV-2 lineage in the UK defined by a novel set of spike mutations". https://virological.org/t/preliminary-genomic-characterisation-of-an-emergent-sars-cov-2-lineage-in-the-uk-defined-by-a-novel-set-of-spike-mutations/563, December 2020. Accessed: 2021-04-09.

[412] Áine O'Toole, Emily Scher, Anthony Underwood, Ben Jackson, Verity Hill, JT McCrone, Chris Ruis, Khali Abu-Dahab, Ben Taylor, Corin Yeats, Louis du Plessis, David Aanensen, Eddie Holmes, Oliver Pybus, and Andrew Rambaut. "pangolin: lineage assignment in an emerging pandemic as an epidemiological tool", 2021. URL github.com/cov-lineages/pangolin. Accessed: 2021-04-09.

[413] Sebastian Duchene, Leo Featherstone, Melina Haritopoulou-Sinanidou, Andrew Rambaut, Philippe Lemey, and Guy Baele. Temporal signal and the phylodynamic threshold of SARS-CoV-2. *Virus evolution*, 6(2):veaa061, 2020.

[414] Tyler N Starr, Allison J Greaney, Sarah K Hilton, Daniel Ellis, Katharine HD Crawford, Adam S Dingens, Mary Jane Navarro, John E Bowen, M Alejandra Tortorici, Alexandra C Walls, et al. Deep mutational scanning of SARS-CoV-2 receptor binding domain reveals constraints on folding and ACE2 binding. *Cell*, 182(5):1295–1310, 2020.

[415] Bryan A Johnson, Xuping Xie, Adam L Bailey, Birte Kalveram, Kumari G Lokugamage, Antonio Muruato, Jing Zou, Xianwen Zhang, Terry Juelich, Jennifer K Smith, et al. Loss of furin cleavage site attenuates SARS-CoV-2 pathogenesis. *Nature*, 591 (7849):293–299, 2021.

[416] Antonin Bal, Gregory Destras, Alexandre Gaymard, Karl Stefic, Julien Marlet, Sébastien Eymieux, Hadrien Regue, Quentin Semanas, Constance d'Aubarede, Geneviève Billaud, et al. Two-step strategy for the identification of SARS-CoV-2 variant of concern 202012/01 and other variants with spike deletion H69–V70, France, August to December 2020. *Eurosurveillance*, 26(3):2100008, 2021.

[417] Vítor Borges, Carlos Sousa, Luís Menezes, António Maia Gonçalves, Miguel Picão, José Pedro Almeida, Margarida Vieita, Rafael Santos, Ana Rita Silva, Mariana Costa, et al. Tracking SARS-CoV-2 lineage B. 1.1. 7 dissemination: insights from nationwide spike gene target failure (SGTF) and spike gene late detection (SGTL) data, Portugal, week 49 2020 to week 3 2021. *Eurosurveillance*, 26(10):2100131, 2021.

[418] Nicholas G Davies, Sam Abbott, Rosanna C Barnard, Christopher I Jarvis, Adam J Kucharski, James D Munday, Carl AB Pearson, Timothy W Russell, Damien C Tully, Alex D Washburne, et al. Estimated transmissibility and impact of SARS-CoV-2 lineage B.1.1.7 in England. *Science*, 2021.

[419] The Independent. Coronavirus: return to full lockdown might not be enough to control new variant, SAGE warns, 2020. URL https://www.independent.co.uk/news/health/coronavirus-new-variant-lockdown-sage-r-rate-b1780909.html. Accessed: 2021-04-09.

[420] Richard Partington. UK economy hit by record slump in 2020 but double-dip recession avoided, 2020. URL https://www.theguardian.com/business/2021/feb/12/uk-avoided-double-dip-recession-despite-covid-slump-in-2020-ons-gdp. Accessed: 2021-03-18.

[421] Matthias Pierce, Holly Hope, Tamsin Ford, Stephani Hatch, Matthew Hotopf, Ann John, Evangelos Kontopantelis, Roger Webb, Simon Wessely, Sally McManus, et al. Mental health before and during the COVID-19 pandemic: a longitudinal probability sample survey of the UK population. *The Lancet Psychiatry*, 7(10):883–892, 2020.

[422] Eric Robinson, Emma Boyland, Anna Chisholm, Joanne Harrold, Niamh G Maloney, Lucile Marty, Bethan R Mead, Rob Noonan, and Charlotte A Hardman. Obesity, eating behavior and physical activity during COVID-19 lockdown: A study of UK adults. *Appetite*, 156:104853, 2021.

[423] Jonathan Jackson, Krisztian Posch, Ben Bradford, Zoe Hobson, and Arabella Kyprianides. The lockdown and social norms: Why the UK is complying by consent rather than compulsion. *British Policy and Politics at LSE*, 2020.

[424] Heidi Ledford. COVID-19 reinfection: three questions scientists are asking. *Nature*, 585(7824):168–169, 2020.

[425] Charlotte Jackson, Punam Mangtani, Jeremy Hawker, Babatunde Olowokure, and Emilia Vynnycky. The effects of school closures on influenza outbreaks and pandemics: systematic review of simulation studies. *PloS one*, 9(5):e97297, 2014.

[426] Julian Jamison, Donald Bundy, Dean Jamison, Jacob Spitz, and Stephane Verguet. Comparing the impact on COVID-19 mortality of self-imposed behavior change and of government regulations across 13 countries. *medRxiv*, 2020.

[427] Google. "COVID-19 Community Mobility Report: United Kingdom", 2021. URL https://www.gstatic.com/covid19/mobility/2021-04-04_GB_Mobility_Report_en-GB.pdf. Accessed: 2021-04-09.

[428] Shina CL Kamerlin and Peter M Kasson. Managing COVID-19 spread with voluntary public-health measures: Sweden as a case study for pandemic control. *Clinical Infectious Diseases*, 2020.

[429] Nicky Phillips. The coronavirus is here to stay-here's what that means. *Nature*, 590 (7846):382–384, 2021.

[430] Jennie S Lavine, Ottar N Bjornstad, and Rustom Antia. Immunological characteristics govern the transition of COVID-19 to endemicity. *Science*, 371(6530):741–745, 2021.

[431] gov.uk. People who have received 1st dose vaccinations, by report date, 2021. URL https://coronavirus.data.gov.uk/details/vaccinations. Accessed: 2021-04-09.

[432] Markus Hoffmann, Prerna Arora, Rüdiger Groß, Alina Seidel, Bojan F Hörnich, Alexander S Hahn, Nadine Krüger, Luise Graichen, Heike Hofmann-Winkler, Amy Kempf, et al. SARS-CoV-2 variants B. 1.351 and P. 1 escape from neutralizing antibodies. *Cell*, 2021.

[433] Daniel J Laydon, Swapnil Mishra, Wes R Hinsley, Pantelis Samartsidis, Seth Flaxman, Axel Gandy, Neil M Ferguson, and Samir Bhatt. Impact of the Tier system on SARS-CoV-2 transmission in the UK between the first and second national lockdowns. *medRxiv*, 2021.

[434] Thomas Hale, Noam Angrist, Rafael Goldszmidt, Beatriz Kira, Anna Petherick, Toby Phillips, Samuel Webster, Emily Cameron-Blake, Laura Hallas, Saptarshi Majumdar, et al. A global panel database of pandemic policies (Oxford COVID-19 government response tracker). *Nature Human Behaviour*, pages 1–10, 2021.

[435] Google. "COVID-19 Community Mobility Reports", 2021. URL https://www.google.com/covid19/mobility/?hl=en. Accessed: 2021-04-09.

[436] Pierre Nouvellet, Sangeeta Bhatia, Anne Cori, Kylie EC Ainslie, Marc Baguelin, Samir Bhatt, Adhiratha Boonyasiri, Nicholas F Brazeau, Lorenzo Cattarino, Laura V Cooper, et al. Reduction in mobility and COVID-19 transmission. *Nature communications*, 12 (1):1–9, 2021.

[437] Hakan Yilmazkuday. Stay-at-home works to fight against COVID-19: international evidence from Google mobility data. *Journal of Human Behavior in the Social Environment*, pages 1–11, 2020.

[438] Wee Chian Koh, Lin Naing, and Justin Wong. Estimating the impact of physical distancing measures in containing COVID-19: an empirical analysis. *International Journal of Infectious Diseases*, 100:42–49, 2020.

[439] Sam Abbott, Joel Hellewell, Robin N Thompson, Katharine Sherratt, Hamish P Gibbs, Nikos I Bosse, James D Munday, Sophie Meakin, Emma L Doughty, June Young Chun, et al. Estimating the time-varying reproduction number of SARS-CoV-2 using national and subnational case counts. *Wellcome Open Research*, 5(112):112, 2020.

[440] BBC. Coronavirus: What is the R number and how is it calculated?, 2021. URL https://www.bbc.co.uk/news/health-52473523. Accessed: 2021-08-23.

[441] Ana da Silva Filipe, James G Shepherd, Thomas Williams, Joseph Hughes, Elihu Aranday-Cortes, Patawee Asamaphan, Shirin Ashraf, Carlos Balcazar, Kirstyn Brunker, Alasdair Campbell, et al. Genomic epidemiology reveals multiple introductions of SARS-CoV-2 from mainland Europe into Scotland. *Nature microbiology*, 6 (1):112–122, 2021.

[442] Heng Li. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics*, 34 (18):3094–3100, 2018.

[443] Morgan N Price, Paramvir S Dehal, and Adam P Arkin. FastTree 2–approximately maximum-likelihood trees for large alignments. *PloS one*, 5(3):e9490, 2010.

[444] Emmanuel Paradis, Julien Claude, and Korbinian Strimmer. APE: analyses of phylogenetics and evolution in R language. *Bioinformatics*, 20(2):289–290, 2004.

[445] Jacco Wallinga and Marc Lipsitch. How generation intervals shape the relationship between growth rates and reproductive numbers. *Proceedings of the Royal Society B: Biological Sciences*, 274(1609):599–604, 2007.

[446] Yung-Wai Desmond Chan, Stefan Flasche, Tin-Long Terence Lam, Mei-Hung Joanna Leung, Miu-Ling Wong, Ho-Yeung Lam, and Shuk-Kwan Chuang. Transmission dynamics, serial interval and epidemiology of COVID-19 diseases in Hong Kong under different control measures. *Wellcome Open Research*, 5:91, 2020.

[447] Qifang Bi, Yongsheng Wu, Shujiang Mei, Chenfei Ye, Xuan Zou, Zhen Zhang, Xiaojian Liu, Lan Wei, Shaun A Truelove, Tong Zhang, et al. Epidemiology and transmission of COVID-19 in 391 cases and 1286 of their close contacts in Shenzhen, China: a retrospective cohort study. *The Lancet Infectious Diseases*, 20(8):911–919, 2020.

[448] Joseph T Wu, Kathy Leung, Mary Bushman, Nishant Kishore, Rene Niehus, Pablo M de Salazar, Benjamin J Cowling, Marc Lipsitch, and Gabriel M Leung. Estimating clinical severity of COVID-19 from the transmission dynamics in Wuhan, China. *Nature medicine*, 26(4):506–510, 2020.

[449] J Wallinga and M Lipsitch. How generation intervals shape the relationship between growth rates and reproductive numbers. *Proc. Biol. Sci.*, 274(1609):599–604, 2007. ISSN 0962-8452. doi: 10.1098/rspb.2006.3754.

[450] Kylie EC Ainslie, Caroline E Walters, Han Fu, Sangeeta Bhatia, Haowei Wang, Xiaoyue Xi, Marc Baguelin, Samir Bhatt, Adhiratha Boonyasiri, Olivia Boyd, et al. Evidence of initial success for China exiting COVID-19 social distancing policy after achieving containment. *Wellcome Open Research*, 5, 2020.

[451] Louis du Plessis, John T McCrone, Alexander E Zarebski, Verity Hill, Christopher Ruis, Bernardo Gutierrez, Jayna Raghwani, Jordan Ashworth, Rachel Colquhoun, Thomas R Connor, et al. Establishment and lineage dynamics of the SARS-CoV-2 epidemic in the UK. *Science*, 371(6530):708–712, 2021.

[452] Jamie Lopez Bernal, Nikolaos Panagiotopoulos, Chloe Byers, Tatiana Garcia Vilaplana, Nicola L Boddington, XuSheng Zhang, Andre Charlett, Suzanne Elgohari, Laura Coughlan, Rosie Whillock, et al. Transmission dynamics of COVID-19 in household and community settings in the United Kingdom. *medRxiv*, 2020.

[453] Zachary J Madewell, Yang Yang, Ira M Longini, M Elizabeth Halloran, and Natalie E Dean. Household Transmission of SARS-CoV-2: A Systematic Review and Meta-analysis. *JAMA network open*, 3(12):e2031756–e2031756, 2020.

[454] Christopher I Jarvis, Kevin Van Zandvoort, Amy Gimma, Kiesha Prem, Petra Klepac, G James Rubin, and W John Edmunds. Quantifying the impact of physical distance measures on the transmission of COVID-19 in the UK. *BMC medicine*, 18:1–10, 2020.

[455] The Telegraph. Lockdowns 1, 2 and 3: What are the differences, and what are you unable to do?, 2021. URL https://www.telegraph.co.uk/news/2021/01/07/lockdown-first-second-third-vs-tier-4-what-difference-can-cant-do/. Accessed: 2021-05-19.

[456] Kris V Parag, Louis du Plessis, and Oliver G Pybus. Jointly inferring the dynamics of population size and sampling intensity from molecular sequences. *Molecular biology and evolution*, 37(8):2414–2429, 2020.

[457] Steven Riley, Kylie EC Ainslie, Oliver Eales, Caroline E Walters, Haowei Wang, Christina J Atchison, Claudio Fronterre, Peter J Diggle, Deborah Ashby, Christl A Donnelly, et al. Resurgence of SARS-CoV-2 in England: detection by community antigen surveillance. *medRxiv*, 2020.

[458] Steven Riley, Caroline E Walters, Haowei Wang, Oliver Eales, David Haw, Kylie EC Ainslie, Christina Atchinson, Claudio Fronterre, Peter J Diggle, Deborah Ashby, et al. REACT-1 round 9 interim report: downward trend of SARS-CoV-2 in England in February 2021 but still at high prevalence. *medRxiv*, 2021.

[459] Paul R De Lay, Adèle Benzaken, Quarraisha Abdool Karim, Sani Aliyu, Carolyn Amole, George Ayala, Kalipso Chalkidou, Judy Chang, Michaela Clayton, Aleny Couto, et al. Ending AIDS as a public health threat by 2030: Time to reset targets for 2025. *PLoS medicine*, 18(6):e1003649, 2021.

[460] UNAIDS. HIV prevention, 2021. URL https://www.unaids.org/en/topic/prevention. Accessed: 2021-08-11.

[461] Reuben Granich, James G Kahn, Rod Bennett, Charles B Holmes, Navneet Garg, Celicia Serenata, Miriam Lewis Sabin, Carla Makhlouf-Obermeyer, Christina De Filippo Mack, Phoebe Williams, et al. Expanding ART for treatment and prevention of HIV in South Africa: estimated cost and cost-effectiveness 2011-2050. *PloS one*, 7 (2):e30216, 2012.

[462] Jennifer A Smith, Monisha Sharma, Carol Levin, Jared M Baeten, Heidi van Rooyen, Connie Celum, Timothy B Hallett, and Ruanne V Barnabas. Cost-effectiveness of community-based strategies to strengthen the continuum of HIV care in rural South Africa: a health economic modelling analysis. *The Lancet HIV*, 2(4):e159–e168, 2015.

[463] Ramzi A Alsallaq, Jasmine Buttolph, Charles M Cleland, Timothy Hallett, Irene Inwani, Kawango Agot, and Ann E Kurth. The potential impact and cost of focusing HIV prevention on young women and men: A modeling analysis in western Kenya. *PLoS One*, 12(4):e0175447, 2017.

[464] Ranjeeta Thomas, William JM Probert, Rafael Sauter, Lawrence Mwenge, Surya Singh, Sarah Kanema, Nosivuyile Vanqa, Abigail Harper, Ronelle Burger, Anne Cori, et al. Cost and cost-effectiveness of a universal HIV testing and treatment intervention in Zambia and South Africa: evidence and projections from the HPTN 071 (PopART) trial. *The Lancet Global Health*, 9(5):e668–e680, 2021.

[465] Olga Tymejczyk, Ellen Brazier, Constantin T Yiannoutsos, Michael Vinikoor, Monique van Lettow, Fred Nalugoda, Mark Urassa, Jean d'Amour Sinayobye, Peter F Rebeiro, Kara Wools-Kaloustian, et al. Changes in rapid HIV treatment initiation after national "treat all" policy adoption in 6 sub-Saharan African countries: regression discontinuity analysis. *PLoS medicine*, 16(6):e1002822, 2019.

[466] aidsmap. WHO guidelines recommend universal antiretroviral treatment, PrEP for those at substantial risk, 2015. URL https://www.aidsmap.com/news/sep-2015/who-guidelines-recommend-universal-antiretroviral-treatment-prep-those-substantial. Accessed: 2021-08-11.

[467] World Health Organisation. Global State of PrEP, 2021. URL https://www.who.int/groups/global-prep-network/global-state-of-prep. Accessed: 2021-08-11.

[468] World Health Organisation. Global data shows increasing PrEP use and widespread adoption of WHO PrEP recommendations, 2021. URL https://www.who.int/news-room/feature-stories/detail/global-data-shows-increasing-prep-use-and-widespread-adoption-of-who-prep-recommendations. Accessed: 2021-08-11.

[469] PrEPwatch. About the Dapivirine Ring, 2021. URL https://www.prepwatch.org/about-prep/dapivirine-ring/. Accessed: 2021-08-11.

[470] M. C. Thigpen, P. M. Kebaabetswe, L. A. Paxton, D. K. Smith, C. E. Rose, T. M. Segolodi, F. L. Henderson, S. R. Pathak, F. A. Soud, K. L. Chillag, R. Mutanhaurwa, L. I. Chirwa, M. Kasonde, D. Abebe, E. Buliva, R. J. Gvetadze, S. Johnson, T. Sukalac, V. T. Thomas, C. Hart, J. A. Johnson, C. K. Malotte, C. W. Hendrix, J. T. Brooks, and T. D. F. Study Group. Antiretroviral preexposure prophylaxis for heterosexual

HIV transmission in Botswana. *N Engl J Med*, 367(5):423–34, 2012. ISSN 1533-4406 (Electronic) 0028-4793 (Linking). doi: 10.1056/NEJMoa1110711. URL https://www.ncbi.nlm.nih.gov/pubmed/22784038.

[471] Lut Van Damme, Amy Corneli, Khatija Ahmed, Kawango Agot, Johan Lombaard, Saidi Kapiga, Mookho Malahleha, Fredrick Owino, Rachel Manongi, Jacob Onyango, Lucky Temu, Modie Constance Monedi, Paul Mak'Oketch, Mankalimeng Makanda, Ilse Reblin, Shumani Elsie Makatu, Lisa Saylor, Haddie Kiernan, Stella Kirkendale, Christina Wong, Robert Grant, Angela Kashuba, Kavita Nanda, Justin Mandala, Katrien Fransen, Jennifer Deese, Tania Crucitti, Timothy D. Mastro, and Douglas Taylor. Preexposure prophylaxis for HIV infection among African women. *New England Journal of Medicine*, 367(5):411–422, 2012. ISSN 0028-4793 1533-4406. doi: 10.1056/nejmoa1202614. URL http://dx.doi.org/10.1056/nejmoa1202614.

[472] World Health Organisation. Trial results reveal that long-acting injectable cabotegravir as PrEP is highly effective in preventing HIV acquisition in women, 2020. URL https://www.who.int/news/item/09-11-2020-trial-results-reveal-that-long-acting-injectable-cabotegravir-as-prep-is-highly-effective-in-preventing-hiv-acquisition-in-women. Accessed: 2021-08-11.

[473] Elizabeth E Tolley, Sue Li, Sahar Z Zangeneh, Millicent Atujuna, Petina Musara, Jessica Justman, Subash Pathak, Linda-Gail Bekker, Shobha Swaminathan, Jill Stanton, et al. Acceptability of a long-acting injectable HIV prevention product among US and African women: findings from a phase 2 clinical Trial (HPTN 076). *Journal of the International AIDS Society*, 22(10):e25408, 2019.

[474] Raphael J Landovitz, Ryan Kofron, and Marybeth McCauley. The promise and pitfalls of long acting injectable agents for HIV prevention. *Current Opinion in HIV and AIDS*, 11(1):122, 2016.

[475] Maria N Pyra, Jessica E Haberer, Nina Hasen, Jason Reed, Nelly R Mugo, and Jared M Baeten. Global implementation of PrEP for HIV prevention: setting expectations for impact. *Journal of the International AIDS Society*, 22(8):e25370, 2019.

[476] Philip B Berger. The XIV International AIDS Conference: a call for action... now. *Cmaj*, 167(5):483–484, 2002.

[477] PJ Feldblum, MJ Welsh, and MJ Steiner. Don't overlook condoms for HIV prevention, 2003.

[478] Ministère du Développement. République du Bénin Enquête Démographique et de Santé (Demographic and Health Survey), 2011. URL https://dhsprogram.com/pubs/pdf/FR270/FR270.pdf.

[479] David S Houéto, Eloi B N'Koué N'Da, and Emmanuel N Sambiéni. Factors Associated With Condom Use Among High School Students of Natitingou Agricultural Technical School, Benin, in 2017. *International Quarterly of Community Health Education*, 41 (4):363–368, 2021.

[480] Tisha Wheeler, Usha Kiran, Gina Dallabetta, Matangi Jayaram, Padma Chandrasekaran, Annie Tangri, Hari Menon, Sameer Kumta, Sema Sgaier, Aparajita Ramakrishnan, et al. Learning about scale, measurement and community mobilisation: reflections on the implementation of the Avahan HIV/AIDS initiative in India. *J Epidemiol Community Health*, 66(Suppl 2):ii16–ii25, 2012.

[481] Jennifer Toller Erausquin, Elizabeth Reed, and Kim M Blankenship. Police-related experiences and HIV risk among female sex workers in Andhra Pradesh, India. *Journal of Infectious Diseases*, 204(suppl_5):S1223–S1228, 2011.

[482] Kate Shannon, Steffanie A Strathdee, Shira M Goldenberg, Putu Duff, Peninah Mwangi, Maia Rusakova, Sushena Reza-Paul, Joseph Lau, Kathleen Deering, Michael R Pickles, et al. Global epidemiology of HIV among female sex workers: influence of structural determinants. *The Lancet*, 385(9962):55–71, 2015.

[483] Gautam Gulati and Brendan D Kelly. Domestic violence against women and the covid-19 pandemic: What is the role of psychiatry? *International journal of law and psychiatry*, 71:101594, 2020.

[484] Nádia Nara Rolim Lima, Ricardo Inácio de Souza, Pedro Walisson Gomes Feitosa, Jorge Lucas de Sousa Moreira, Claudio Gleidiston Lima da Silva, and Modesto Leite Rolim Neto. People experiencing homelessness: Their potential exposure to covid-19. *Psychiatry research*, 288:112945, 2020.

[485] Carin Ahouada, Souleymane Diabaté, Ndeye Ndiagna Gning, Septime Hessou, Georges Batona, Fernand A Guédou, Luc Béhanzin, Marcel D Zannou, and Michel Alary. Acceptability of pre-exposure prophylaxis (PrEP) among men who have sex with men (MSM) in Benin: a qualitative study. *AIDS care*, 32(2):242–248, 2020.

[486] J. M. Molina, C. Capitant, B. Spire, G. Pialoux, L. Cotte, I. Charreau, C. Tremblay, J. M. Le Gall, E. Cua, A. Pasquet, F. Raffi, C. Pintado, C. Chidiac, J. Chas, P. Charbonneau, C. Delaugerre, M. Suzan-Monti, B. Loze, J. Fonsart, G. Peytavin, A. Cheret, J. Timsit, G. Girard, N. Lorente, M. Preau, J. F. Rooney, M. A. Wainberg, D. Thompson, W. Rozenbaum, V. Dore, L. Marchand, M. C. Simon, N. Etien, J. P. Aboulker, L. Meyer, J. F. Delfraissy, and Anrs Ipergay Study Group. On-demand preexposure prophylaxis in men at high risk for HIV-1 infection. *N Engl J Med*, 373(23):2237–46, 2015. doi: 10.1056/NEJMoa1506273.

[487] R. M. Grant, J. R. Lama, P. L. Anderson, V. McMahan, A. Y. Liu, L. Vargas, P. Goicochea, M. Casapia, J. V. Guanira-Carranza, M. E. Ramirez-Cardich, O. Montoya-Herrera, T. Fernandez, V. G. Veloso, S. P. Buchbinder, S. Chariyalertsak, M. Schechter, L. G. Bekker, K. H. Mayer, E. G. Kallas, K. R. Amico, K. Mulligan, L. R. Bushman, R. J. Hance, C. Ganoza, P. Defechereux, B. Postle, F. Wang, J. J. McConnell, J. H. Zheng, J. Lee, J. F. Rooney, H. S. Jaffe, A. I. Martinez, D. N. Burns, D. V. Glidden, and Team iPrEx Study. Preexposure chemoprophylaxis for HIV prevention in men who have sex with men. *N Engl J Med*, 363(27):2587–99, 2010. doi: 10.1056/NEJMoa1011205.

[488] Rogers Sekabira, Ian Mcgowan, Krista Yuhas, Rhonda M Brand, Mark A Marzinke, Yukari C Manabe, Ian Frank, Joseph Eron, Raphael J Landovitz, Peter Anton, et al.

Higher colorectal tissue HIV infectivity in cisgender women compared with MSM before and during oral preexposure prophylaxis. *AIDS*, 35(10):1585–1595, 2021.

[489] Carla Kuiken, Bette Korber, and Robert W Shafer. HIV sequence databases. *AIDS reviews*, 5(1):52, 2003.

[490] Zhanwei Du, Lin Wang, Simon Cauchemez, Xiaoke Xu, Xianwen Wang, Benjamin J Cowling, and Lauren Ancel Meyers. Risk for transportation of coronavirus disease from Wuhan to other cities in China. *Emerging infectious diseases*, 26(5):1049, 2020.

[491] Biao Tang, Xia Wang, Qian Li, Nicola Luigi Bragazzi, Sanyi Tang, Yanni Xiao, and Jianhong Wu. Estimation of the transmission risk of the 2019-nCoV and its implication for public health interventions. *Journal of clinical medicine*, 9(2):462, 2020.

[492] Swapnil Mishra, Jamie Scott, Harrison Zhu, Neil M Ferguson, Samir Bhatt, Seth Flaxman, and Axel Gandy. A COVID-19 model for local authorities of the United Kingdom. *Medrxiv*, 2020.

[493] Erik Volz, Swapnil Mishra, Meera Chand, Jeffrey C Barrett, Robert Johnson, Lily Geidelberg, Wes R Hinsley, Daniel J Laydon, Gavin Dabrera, Áine O'Toole, et al. Transmission of SARS-CoV-2 Lineage B. 1.1. 7 in England: Insights from linking epidemiological and genetic data. *MedRxiv*, pages 2020–12, 2021.

[494] Darlan S Candido, Ingra M Claro, Jaqueline G De Jesus, William M Souza, Filipe RR Moreira, Simon Dellicour, Thomas A Mellan, Louis Du Plessis, Rafael HM Pereira, Flavia CS Sales, et al. Evolution and epidemic spread of sars-cov-2 in brazil. *Science*, 369(6508):1255–1260, 2020.

[495] Houriiyah Tegally, Eduan Wilkinson, Richard J Lessells, Jennifer Giandhari, Sureshnee Pillay, Nokukhanya Msomi, Koleka Mlisana, Jinal N Bhiman, Anne von Gottberg, Sibongile Walaza, et al. Sixteen novel lineages of sars-cov-2 in south africa. *Nature Medicine*, 27(3):440–446, 2021.

[496] WHO. Classification of Omicron (B.1.1.529): SARS-CoV-2 Variant of Concerns, 2021. URL https://www.who.int/news/item/26-11-2021-classification-of-omicron-b.1.1.529-sars-cov-2-variant-of-concern. Accessed: 2022-01-25.

[497] The White House. Fact Sheet: Biden Administration Announces $1.7 Billion Investment, 2021. URL https://www.whitehouse.gov/briefing-room/statements-releases/2021/04/16/fact-sheet-biden-administration-announces-1-7-billion-investment-to-fight-covid-19-variants/. Accessed: 2022-01-25.

[498] Gargee Ghosh. Global action to end the COVID-19 crisis, 2021. URL https://www.gatesfoundation.org/ideas/articles/global-response-needed-to-end-covid-19. Accessed: 2022-01-25.

[499] ECDC. Guidance for representative and targeted genomic SARS-CoV-2 monitoring, 2021. URL https://www.ecdc.europa.eu/sites/default/files/documents/Guidance-for-representative-targeted-genomic-SARS-CoV-2-monitoring.pdf. Accessed: 2021-08-11.

[500] Public Health England. SARS-CoV-2 variants of concern and variants under investigation in England - Technical briefing 13, 2021. URL https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/990339/Variants_of_Concern_VOC_Technical_Briefing_13_England.pdf.

[501] gov.uk. Eighty-ninth SAGE meeting on COVID-19, 13 May 2021, 2021. URL https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/988403/S1236_Eighty-nineth_SAGE.pdf.

[502] Finlay Campbell, Brett Archer, Henry Laurenson-Schafer, Yuka Jinnai, Franck Konings, Neale Batra, Boris Pavlin, Katelijn Vandemaele, Maria D Van Kerkhove, Thibaut Jombart, et al. Increased transmissibility and global spread of SARS-CoV-2 variants of concern as at June 2021. *Eurosurveillance*, 26(24):2100509, 2021.

[503] Alun L Lloyd and Steve Valeika. Network models in epidemiology: an overview. *Complex population dynamics: nonlinear modeling in ecology, epidemiology and genetics*, pages 189–214, 2007.

[504] Tae Jin Lee, Masayuki Kakehashi, and Arni SR Srinivasa Rao. Network models in epidemiology. In *Handbook of Statistics*, volume 44, pages 235–256. Elsevier, 2021.

[505] Juanjuan Zhang, Maria Litvinova, Yuxia Liang, Yan Wang, Wei Wang, Shanlu Zhao, Qianhui Wu, Stefano Merler, Cécile Viboud, Alessandro Vespignani, et al. Changes in contact patterns shape the dynamics of the covid-19 outbreak in china. *Science*, 368 (6498):1481–1486, 2020.

[506] Lasse S Liebst, Peter Ejbye-Ernst, Marijn de Bruin, Josephine Thomas, and Marie R Lindegaard. Face-touching behaviour as a possible correlate of mask-wearing: A video observational study of public place incidents during the covid-19 pandemic. *Transboundary and Emerging Diseases*, 2021.

[507] Nicholas Fielmua, Bismarck Yelfogle Guba, and Darius Tuonianuo Mwingyine. Hand hygiene and safety behaviours at shopping centres in covid-19: an observation in wa township in ghana. *Journal of Water, Sanitation and Hygiene for Development*, 11(3): 442–452, 2021.

[508] Nuno R Faria, Marc A Suchard, Ana Abecasis, João D Sousa, Nicaise Ndembi, Idalina Bonfim, Ricardo J Camacho, Anne-Mieke Vandamme, and Philippe Lemey. Phylodynamics of the HIV-1 CRF02_AG clade in Cameroon. *Infection, Genetics and Evolution*, 12(2):453–460, 2012.

[509] Rebecca R Gray, Andrew J Tatem, Susanna Lamers, Wei Hou, Oliver Laeyendecker, David Serwadda, Nelson Sewankambo, Ronald H Gray, Maria Wawer, Thomas C Quinn, et al. Spatial phylodynamics of HIV-1 epidemic emergence in east Africa. *AIDS (London, England)*, 23(14):F9, 2009.

[510] Gareth J Hughes, Esther Fearnhill, David Dunn, Samantha J Lycett, Andrew Rambaut, Andrew J Leigh Brown, and UK HIV Drug Resistance Collaboration. Molecular phylodynamics of the heterosexual HIV epidemic in the United Kingdom. *PLoS pathogens*, 5(9):e1000590, 2009.

[511] AVERT. TREATMENT AS PREVENTION (TASP) FOR HIV, 2021. URL https://www.avert.org/professionals/hiv-programming/prevention/treatment-as-prevention. Accessed: 2021-08-11.

[512] Silvio Romano, Annalisa Fierro, and Antonella Liccardo. Beyond the peak: A deterministic compartment model for exploring the COVID-19 evolution in Italy. *PloS one*, 15(11):e0241951, 2020.

[513] Hamada S Badr and Lauren M Gardner. Limitations of using mobile phone data to model COVID-19 transmission in the USA. *The Lancet Infectious Diseases*, 21(5): e113, 2021.

Appendix A

Supplementary Figures

Chapter 3

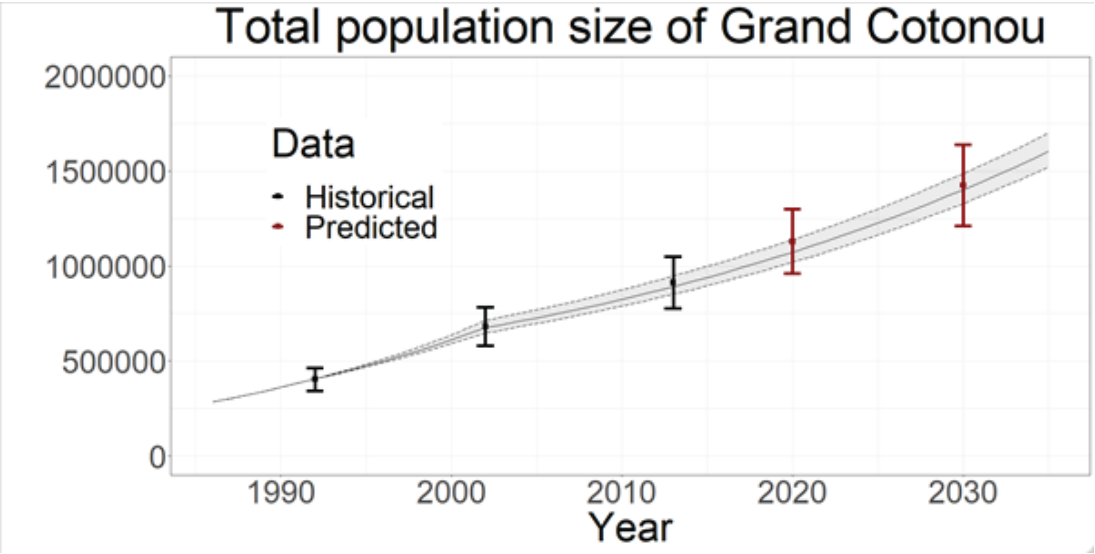


Figure A.1 Posterior model predictions for total population size of Grand Cotonou from the 111 posterior parameter sets, showing median (solid line) and 95%UI across all fits (shaded regions) of counterfactual (no PrEP or TasP) scenario, compared to available fitting historical demographic data shown in thick bars (black) and estimated future predictions assuming constant population growth rate (red).

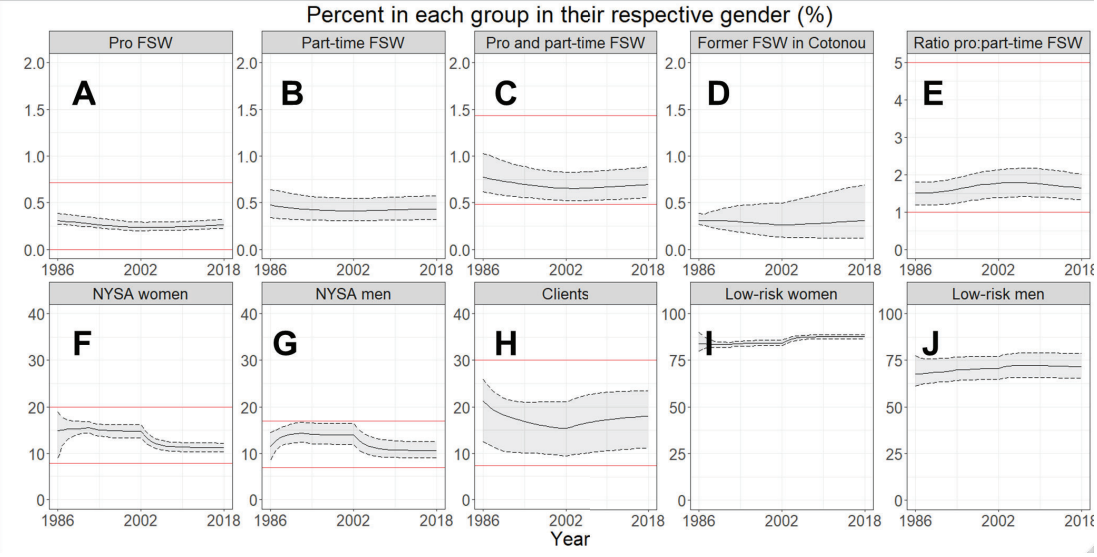


Figure A.2 Posterior model predictions for the percent of each group in their respective gender (e.g. the percent of women who are professional FSW), from the 111 posterior parameter sets, showing median (solid line) and 95%UI across all fits (shaded regions) of counterfactual (no PrEP or TasP) scenario, compared to available fitting historical demographic data shown in red lines.

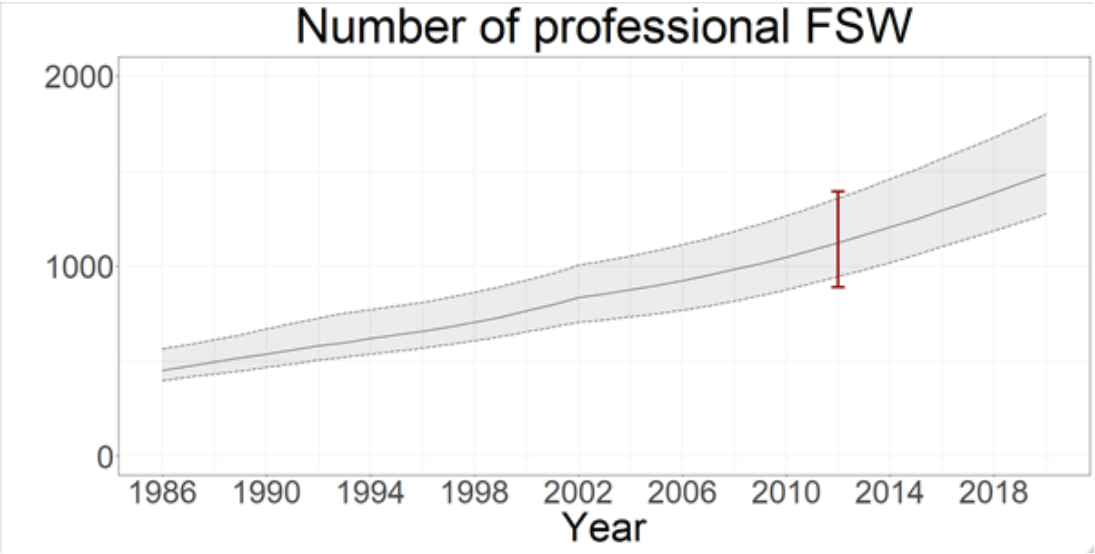


Figure A.3 Posterior model predictions of number of professional FSW from the 111 posterior parameter sets, showing median (solid line) and 95% UI across all fits (shaded regions) of counterfactual (no PrEP or TasP) scenario, compared to available fitting historical demographic data shown with a red bar.

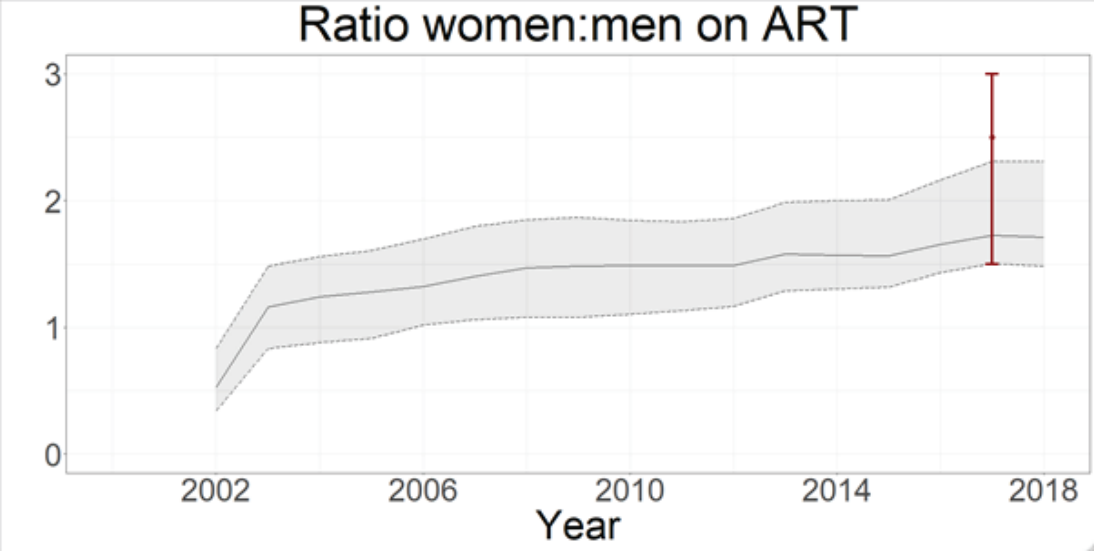


Figure A.4 Posterior model predictions of the ratio of women to men on ART from the 111 posterior parameter sets, showing median (solid line) and 95% UI across all fits (shaded regions) of counterfactual (no PrEP or TasP) scenario, compared to available fitting intervention data shown with a red bar.

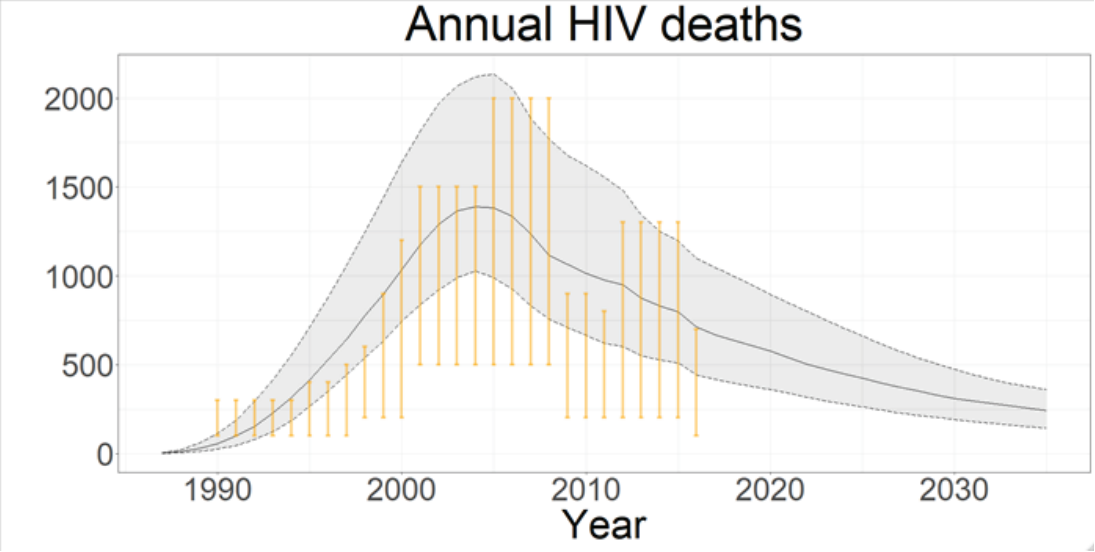


Figure A.5 Posterior model predictions from the 111 posterior parameter sets, showing median (solid line) and 95% UI across all fits (shaded regions) of counterfactual (no PrEP or TasP) scenario, of annual deaths due to HIV infection compared to available cross-validation UNAIDS Reference Group Spectrum model estimates (orange bars) in Cotonou.

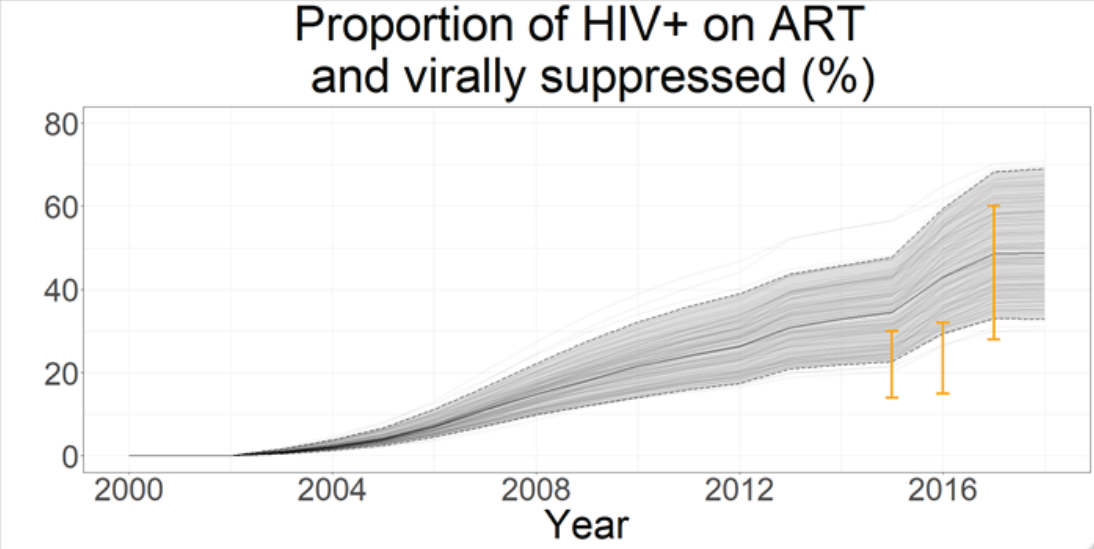


Figure A.6 Posterior model predictions from the 111 posterior parameter sets, showing median (solid line) and 95% UI across all fits (shaded regions) of counterfactual (no PrEP or TasP) scenario, of proportion of HIV-positive people on ART and virally suppressed, compared to available cross-validation UNAIDS Reference Group Spectrum model estimates (orange bars) in Cotonou.

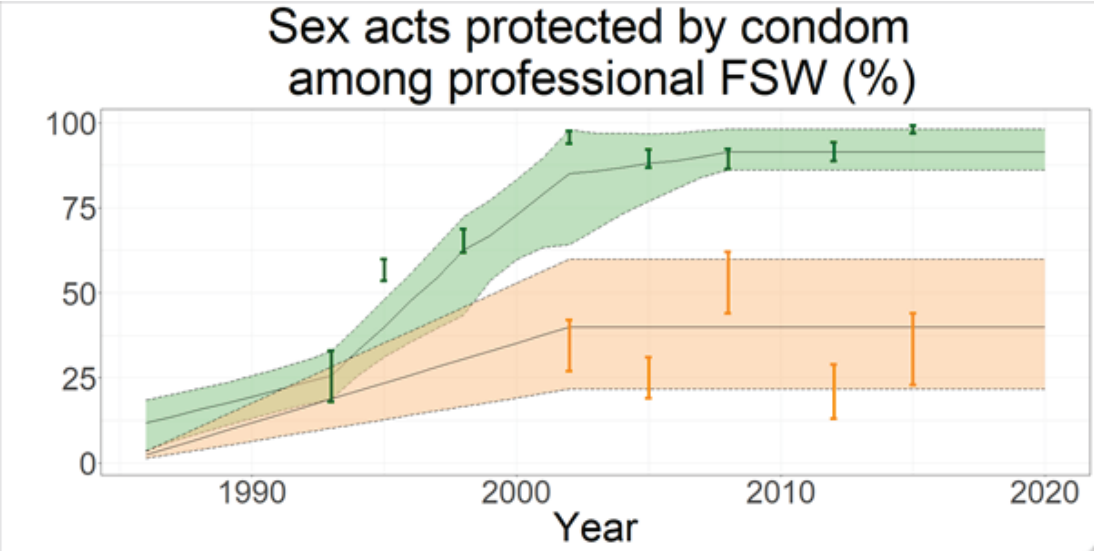


Figure A.7 Posterior percent of commercial and non-commercial partnerships (shown in green and orange respectively) between professional FSW and clients protected by condoms from the 111 posterior parameter sets, showing median (solid line) and 95% UI across all fits (shaded regions) of counterfactual (no PrEP or TasP) scenario, compared to available historical behavioural data in bars.

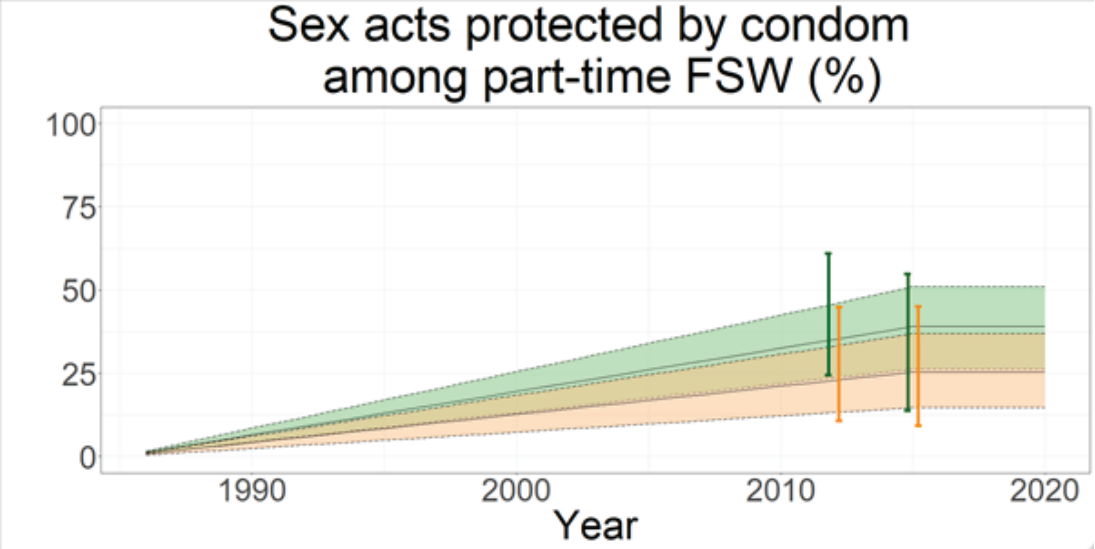


Figure A.8 Posterior percent of commercial and non-commercial partnerships (shown in green and orange respectively) protected by condoms between part-time FSW and clients from the 111 posterior parameter sets, showing median (solid line) and 95% UI across all fits (shaded regions) of counterfactual (no PrEP or TasP) scenario, compared to available historical behavioural data in bars.

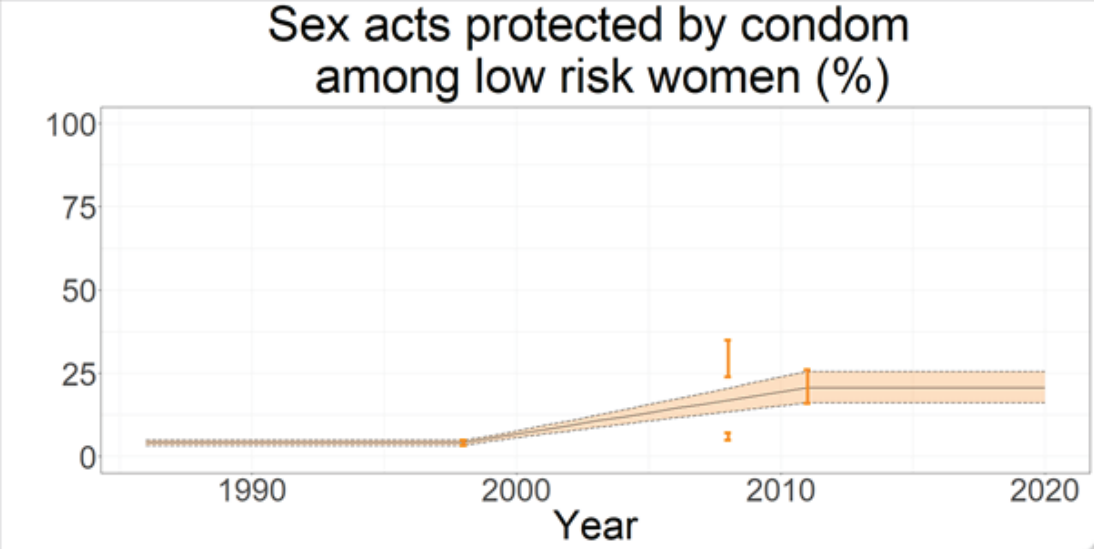


Figure A.9 Posterior percent of non-commercial partnerships between low-risk women and their partners (clients and low-risk men) protected by condoms from the 111 posterior parameter sets, showing median (solid line) and 95% UI across all fits (shaded region) of counterfactual (no PrEP or TasP) scenario, compared to available historical behavioural data in orange bars.

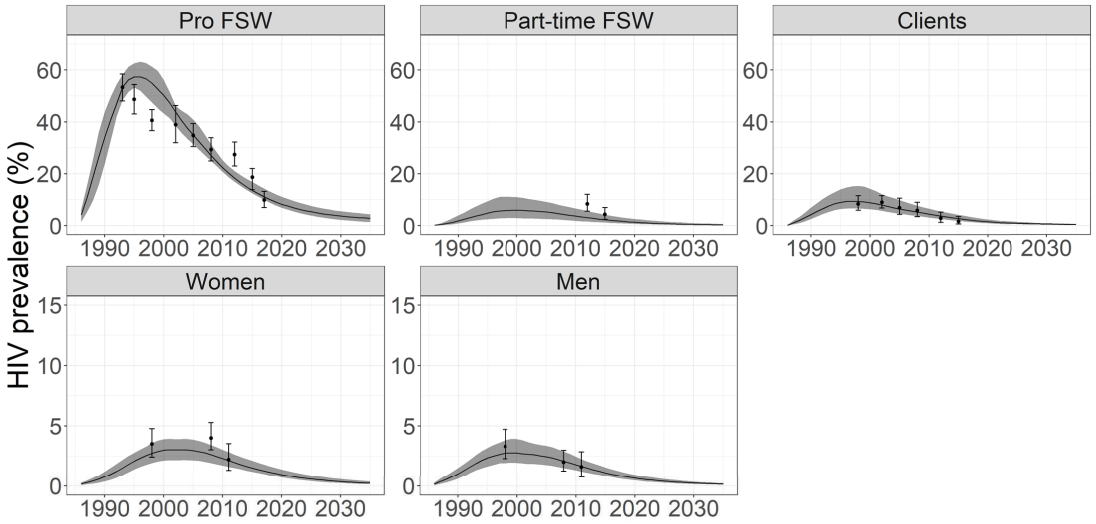


Figure A.10 Model predictions from the 111 posterior parameter sets, in solid lines (median) and shaded regions (95%UI across all fits). Error bars represent cross-sectional HIV prevalence data up to 2017 by risk group. Projections are shown until 2035 in the absence of the PrEP/TasP intervention (counterfactual scenario). The "women" and "men" panels include all women and men in Cotonou, including key populations.

Chapter 6

Case study A

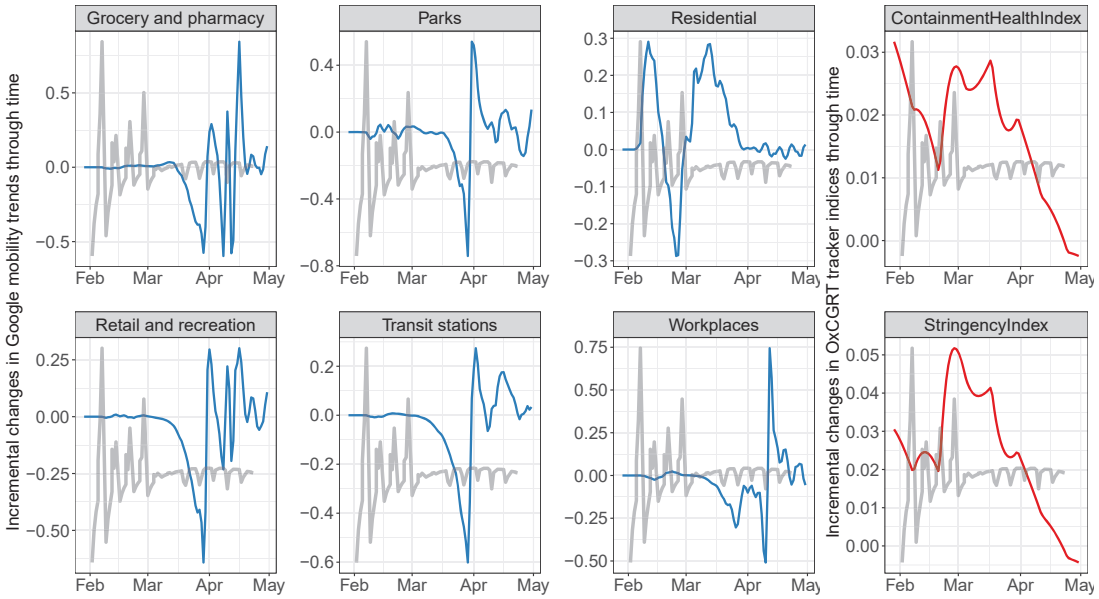


Figure A.11 Daily incremental changes in the transformed time series data from Google mobility (blue lines) and OxCGRT indices (red lines). The re-scaled daily incremental changes in median $R(t)$ estimated with *mlesky* is superimposed on each panel (grey lines).

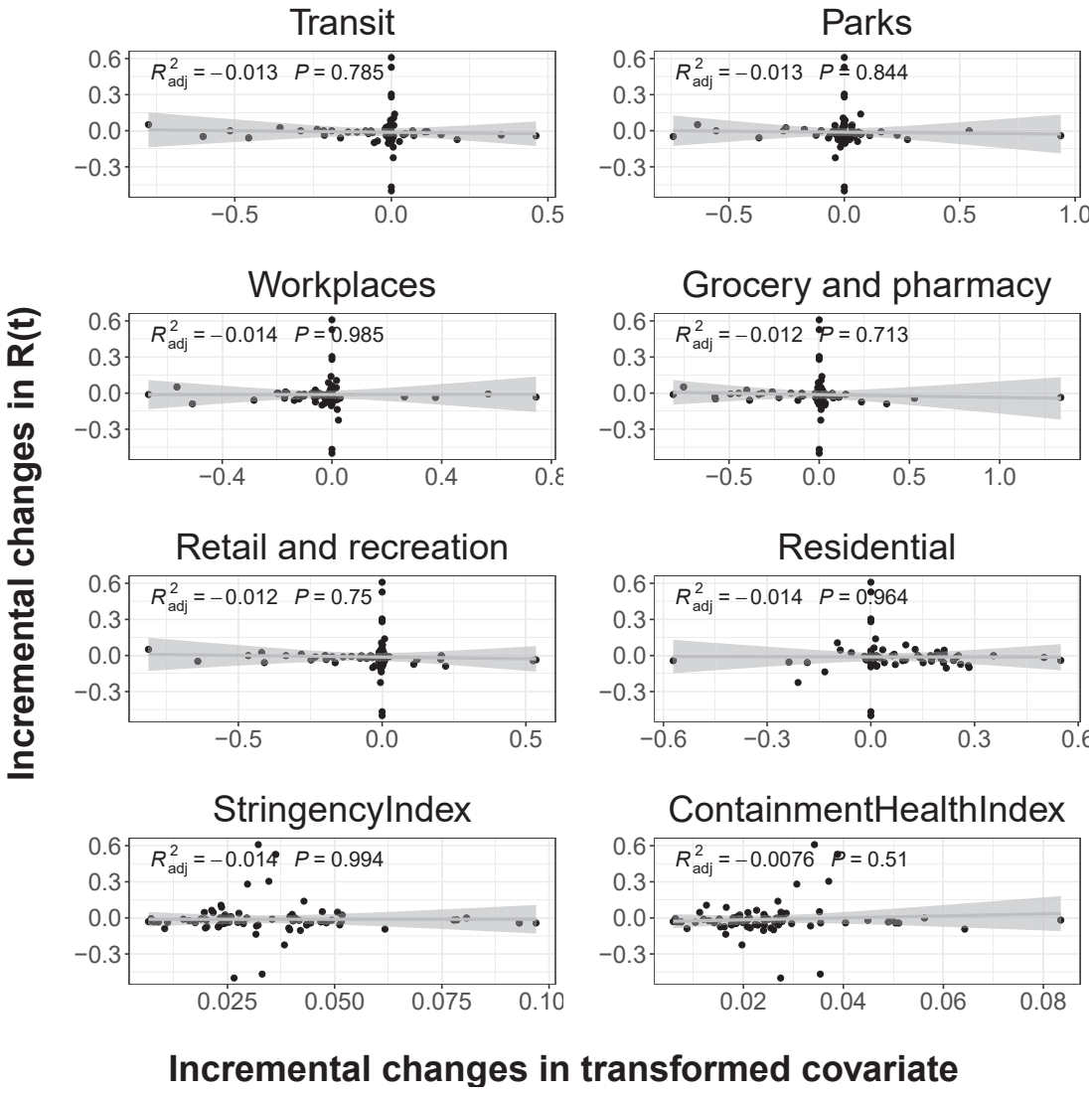


Figure A.12 Scatterplot of the daily incremental changes in transformed covariates against incremental changes in median $R(t)$ shifted 10 days earlier estimated by *mlesky*. Panels A-F display Google mobility data; panels G&H display OxCGRT index data. Grey line shows least squares linear regression; coefficients are displayed top left of each panel. No correlation was found to be significant to 5% significance.

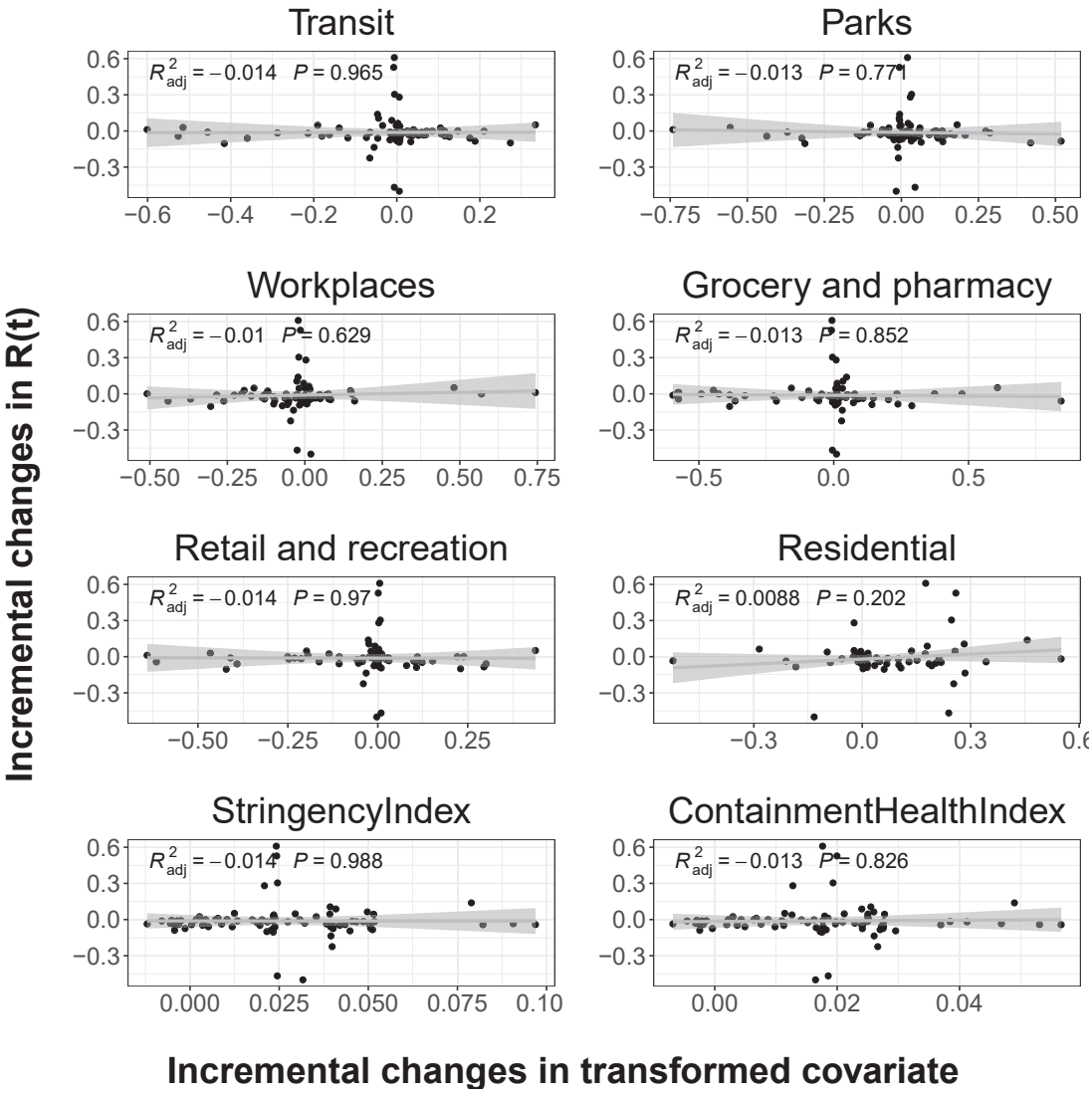


Figure A.13 Scatterplot of the daily incremental changes in transformed covariates against incremental changes in median $R(t)$ shifted 10 days later estimated by *mlesky*. Panels A-F display Google mobility data; panels G&H display OxCGRT index data. Grey line shows least squares linear regression; coefficients are displayed top left of each panel. No correlation was found to be significant to 5% significance.

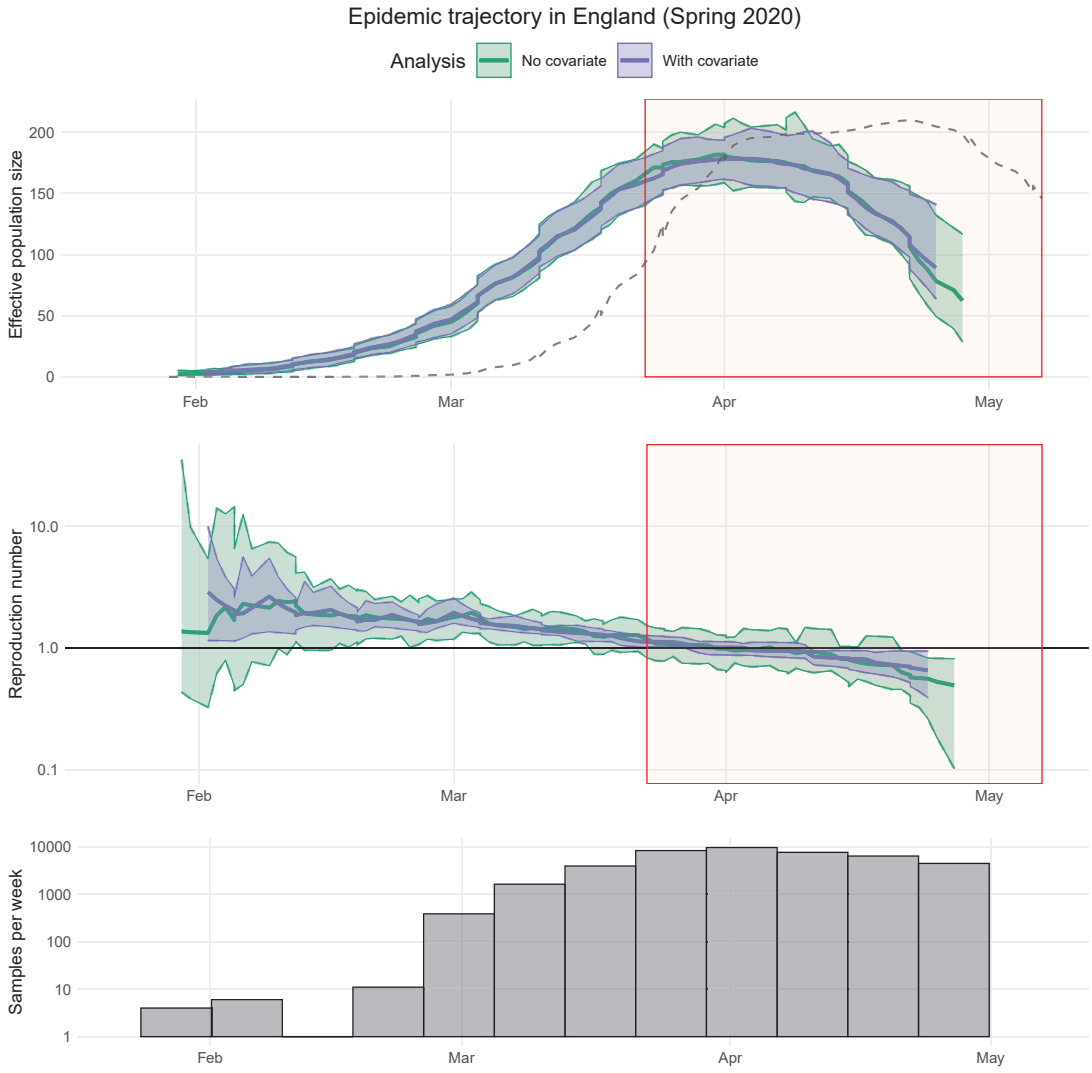


Figure A.14 The epidemiological trajectory of SARS-CoV-2 in England during spring 2020. Thick solid lines and shaded areas represent median and 95% quantiles respectively, estimated by *mlesky* with (purple) and without (green) covariates. Red shaded area represents period of first national lockdown in England. Black dotted line represents daily confirmed cases (smoothed and rescaled). Panel A: effective population size ($N_e(t)$) through time. Panel B: reproduction number ($R(t)$) through time. Panel C: frequencies of tips in each sample week in the phylogenies used.

Case study B

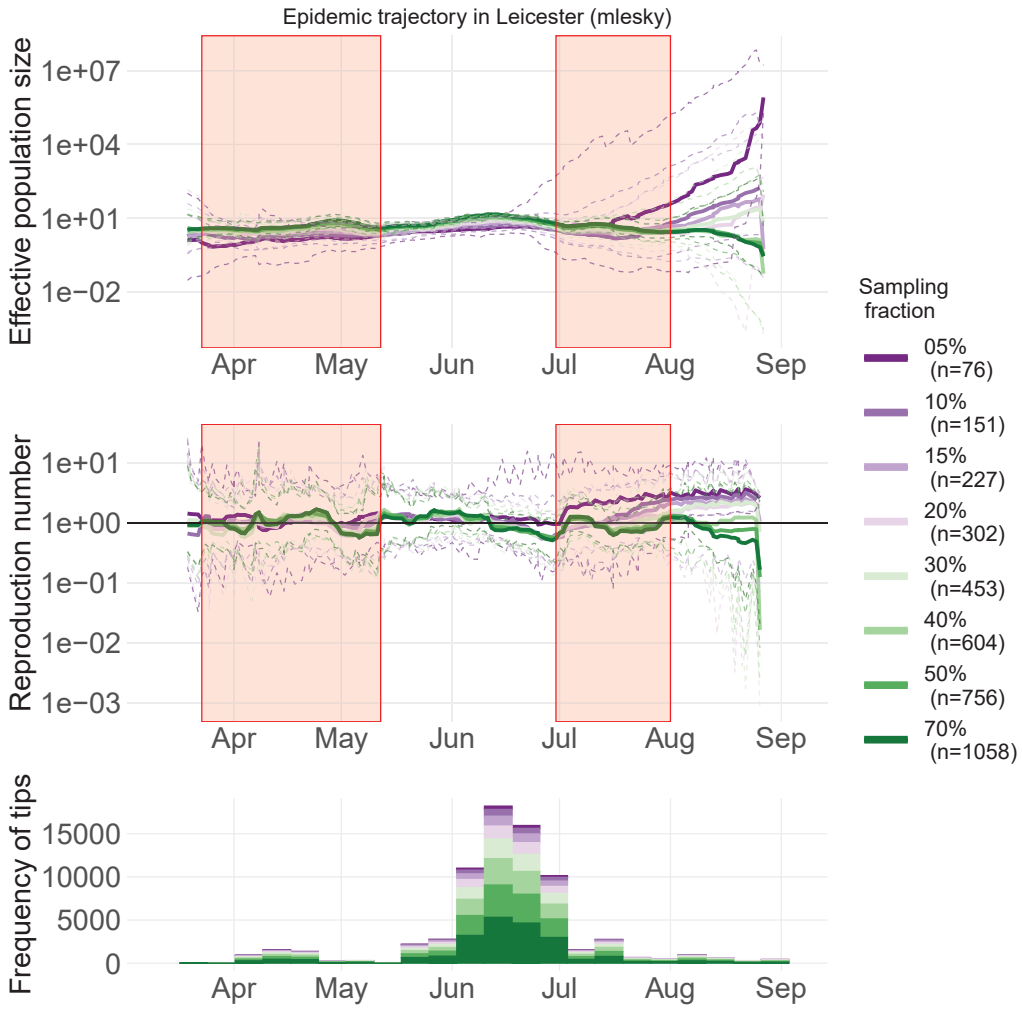


Figure A.15 The epidemiological trajectory of SARS-CoV-2 in Leicester during the spring and summer months of 2020. Thick solid lines and thin dashed lines represent median and 95% quantiles respectively, estimated by *mlesky*. Colour represents the fraction of samples analysed, ranging from 5% at the lowest (purple) to 70% at the highest (green). Legend displays sampling fraction and number of tips per alignment and tree. Red shaded areas represent the first national lockdown in England and local lockdown in Leicester. Panel A: effective population size ($N_e(t)$) through time. Panel B: reproduction number ($R(t)$) through time. Panel C: frequencies of tips in each sample week in the phylogenies used.

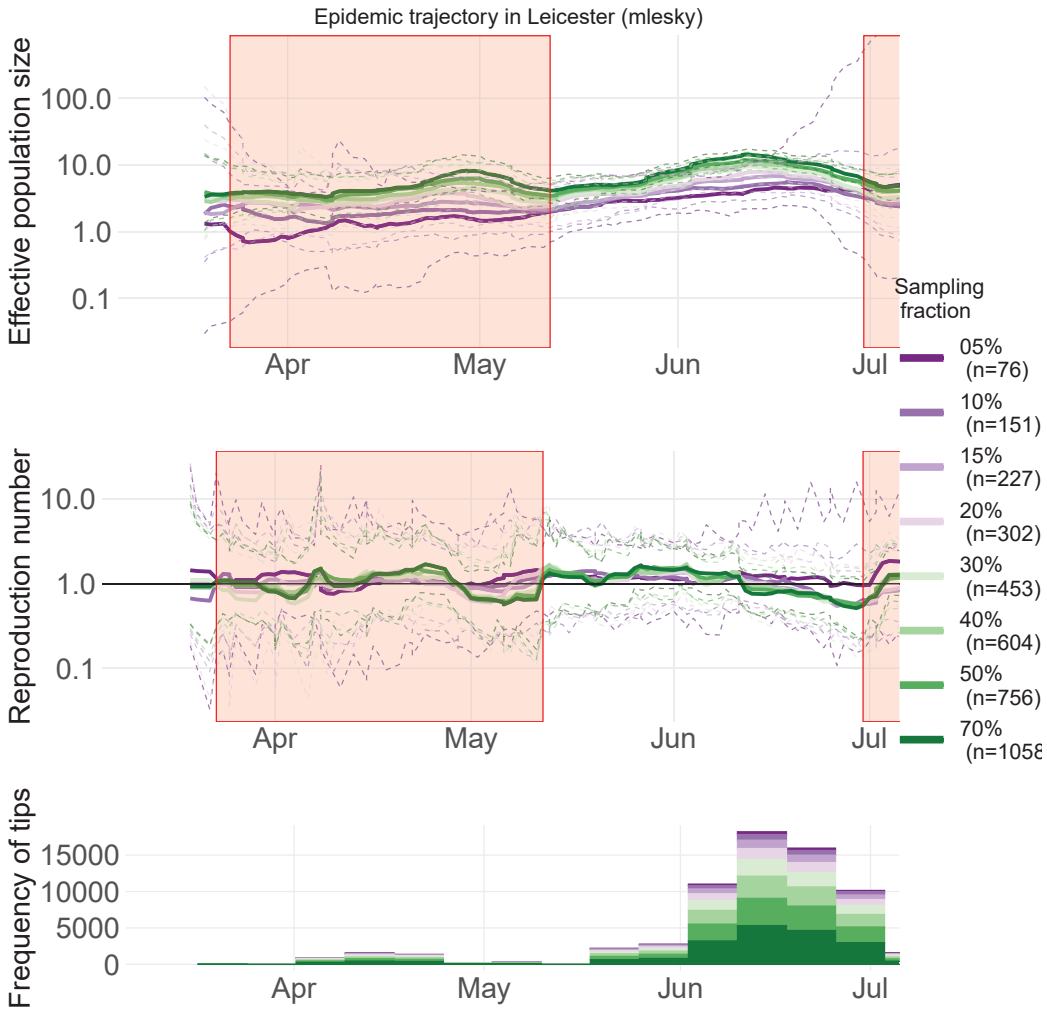


Figure A.16 The epidemiological trajectory of SARS-CoV-2 in Leicester during until July 2020. Thick solid lines and thin dashed lines represent median and 95% quantiles respectively, estimated by *mlesky*. Colour represents the fraction of samples analysed, ranging from 5% at the lowest (purple) to 70% at the highest (green). Legend displays sampling fraction and number of tips per alignment and tree. Red shaded areas represent the first national lockdown in England and local lockdown in Leicester. Panel A: effective population size ($N_e(t)$) through time. Panel B: reproduction number ($R(t)$) through time. Panel C: frequencies of tips in each sample week in the phylogenies used.

Case study D

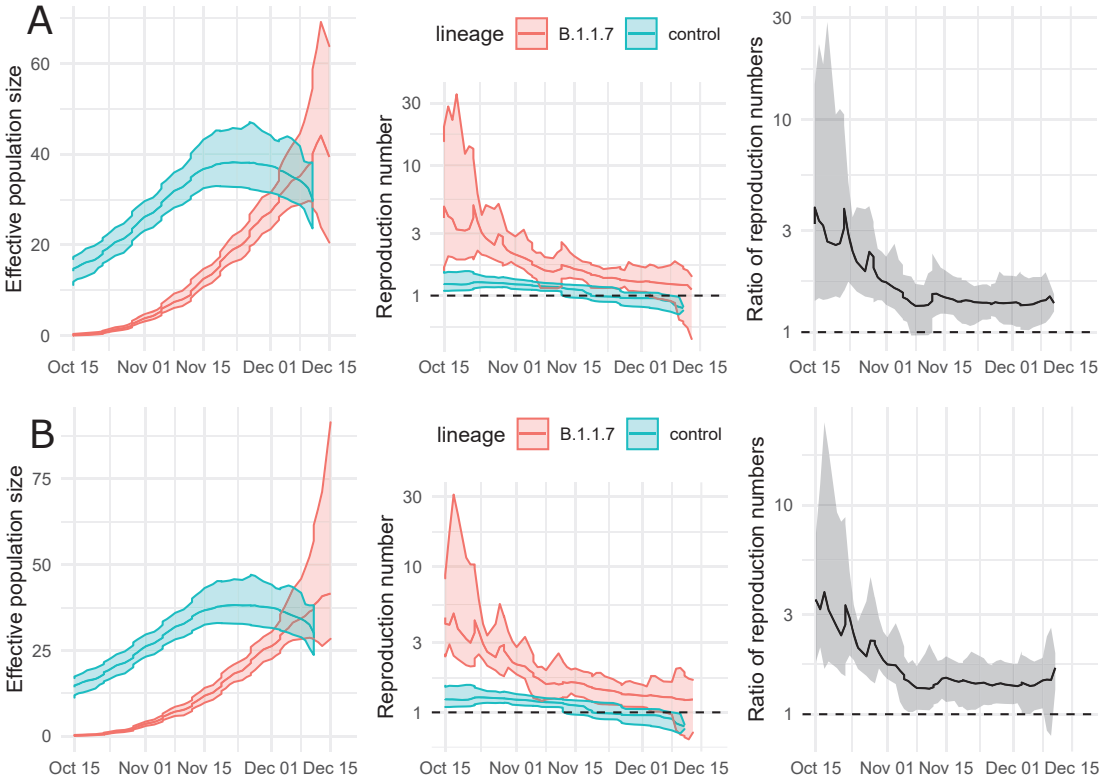


Figure A.17 The epidemiological trajectory of B.1.1.7 (red) and control (blue) lineages in England during the 2020-2021 winter period. Thick solid lines and shaded areas represent median and 95% quantiles respectively, estimated by *mlesky*. Effective population size, reproduction number and the ratio of reproduction numbers (B.1.1.7/control) are shown through time. Panel A: using the global ML tree Panel B: reconstructing an ML tree from the sequences

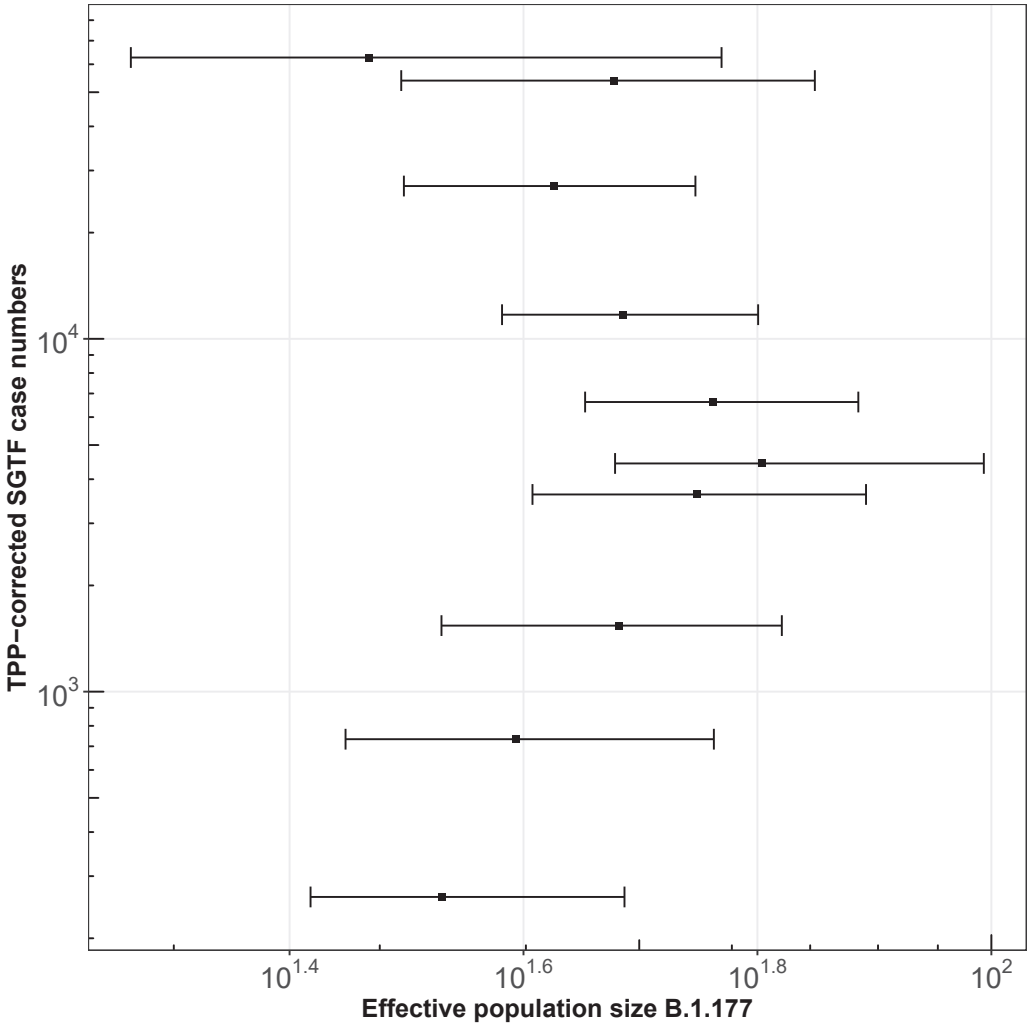


Figure A.18 Comparison of effective population size of B.1.177 lineage against TPP-corrected SGTF case numbers by epiweek.

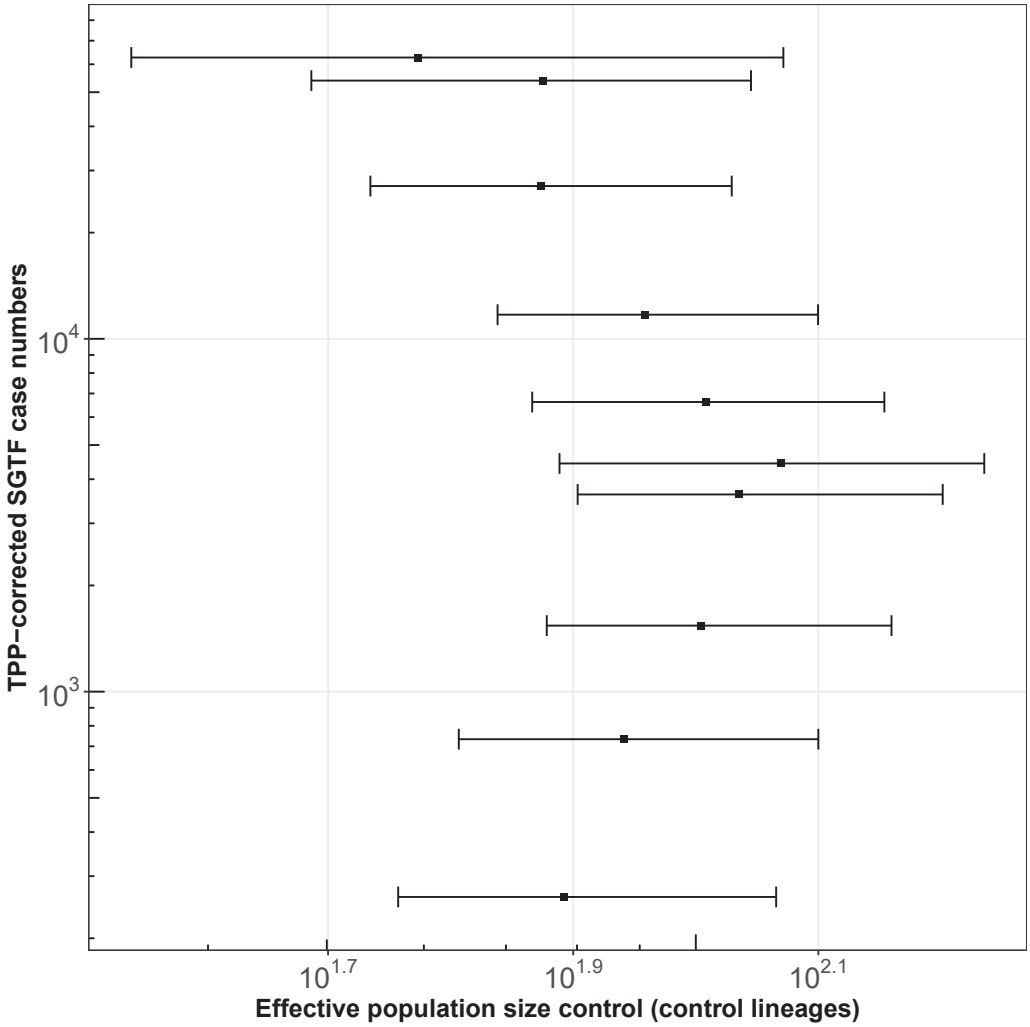


Figure A.19 Comparison of effective population size of control lineages against TPP-corrected SGTF case numbers by epiweek.

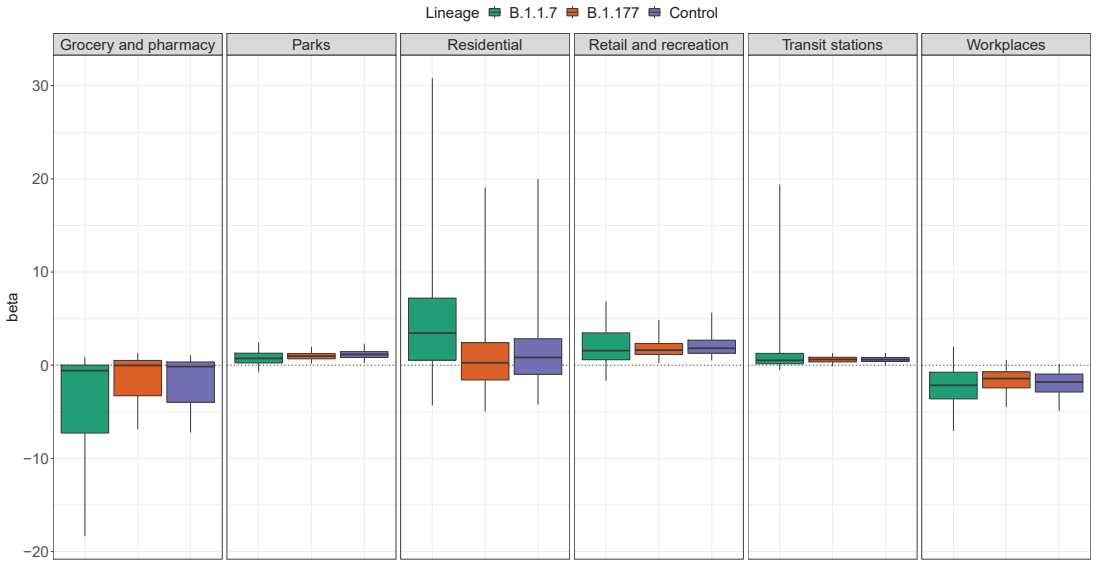


Figure A.20 Distribution of beta coefficients when testing for univariate association between changes in mobility estimated by Google (smoothed with loess parameter 20%) and reproduction number estimated in *mlesky*. Boxes represent median and interquartile range; whiskers denote 95% quartiles. Lineages B.1.1.7, B.1.177 and control are shown in orange, white and grey boxes respectively.

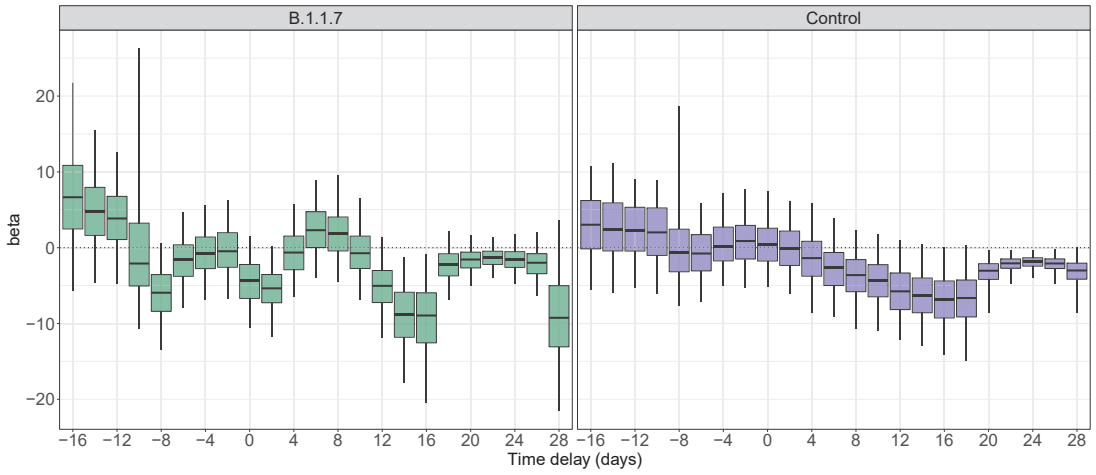


Figure A.21 Distribution of beta coefficients when testing for univariate association between daily changes in OxCGRT (smoothed with loess parameter 4%) health containment index and daily changes in effective population size (i.e. growth rate) in England estimated in *mlesky*. Boxes represent median and interquartile range; whiskers denote 95% quartiles. A positive time delay of 10 days represents testing an association between the covariate at time t and R at time $t + 10$ days. Lineages B.1.1.7 and control are shown in green and purple boxes respectively.