

UCSF

UC San Francisco Electronic Theses and Dissertations

Title

Modeling the transmission of Ebola and GU Chlamydia in Sub-Saharan African countries under both epidemic and endemic settings

Permalink

<https://escholarship.org/uc/item/0hv370b8>

Author

Wannier, S. Rae

Publication Date

2023

Peer reviewed|Thesis/dissertation

Modeling the transmission of Ebola and GU Chlamydia in Sub-Saharan African countries under both epidemic and endemic settings.

by
S. Rae Wannier

DISSERTATION
Submitted in partial satisfaction of the requirements for degree of
DOCTOR OF PHILOSOPHY

in
Epidemiology and Translational Science

in the
GRADUATE DIVISION
of the
UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

Approved:

DocuSigned by:
Porco, Travis Porco, Travis
1E4E7EDA90F44A6... Chair

DocuSigned by:
Caryn Bern Caryn Bern

DocuSigned by:
James Holland Jones James Holland Jones

DocuSigned by:
Shiboski, Stephen Shiboski, Stephen
013F9EE740B848B...

Committee Members

Copyright 2023

by

S. Rae Wannier

Acknowledgements

I gratefully acknowledge support from the US NIH NIGMS MIDAS program, 1-R01-GM130900.

To my dissertation committee: Dr. Travis Porco, Dr. James H. Jones, Dr. Caryn Bern and Dr. Steve Shiboski thank you for your long support and guidance throughout this process.

To Dr. Lee Worden thank you for always being generous with your time and knowledge, it was always appreciated and I learned so much, and especially for your patience.

To Dr. Dan Kelly thank you for your time and guidance, it has been fun to work together on so many different projects and capacities, and for bringing me into new projects.

To my adviser Dr. Travis Porco, I so appreciate your time, guidance and knowledge throughout my rotation and dissertation. I could not have done this without you, and I am so grateful to have had you as my adviser, supporting me steadily throughout many ups and downs and detours along the way.

And lastly to my family who patiently waited so many years watching me and for mostly being willing not to ask ;) I have finally done it and I look forward to having you with me during my next chapter.

Contributions

The text of the first chapter of this dissertation is a reprint of the material as it appears in *Epidemics* (2019). The senior co-authors listed in this publication listed in this publication directed and supervised the research that forms the basis for the first chapter of the dissertation.

I acknowledge following individuals for their contributions to Chapter 1 of my thesis titled "Estimating the impact of violent events on transmission in Ebola virus disease outbreak, Democratic Republic of the Congo, 2018–2019": Lee Worden, Nicole A. Hoff, Eduardo Amezcua, Bernice Selo, Cyrus Sinai, Mathias Mossoko, Bathe Njoloko, Emile Okitolonda-Wemakoy, Placide Mbala-Kingebeni, Steve Ahuka-Mundeke, Jean Jacques Muyembe-Tamfum, Eugene T. Richardson, George W. Rutherford, James H Jones, Thomas M. Lietman, Anne W. Rimoin, Travis C. Porco, and J. Daniel Kelly.

I acknowledge the following individuals for their contributions to Chapter 2 of my thesis named "Real-time projections of the epidemic curve of the 2020 Mbandaka Ebola outbreak": As readers who looked over the manuscript - Seth Blumberg, Caitlin A. Moe, Bernice Selo, Mathias Mossoko, Placide Mbala, Nicole A. Hoff, Anne W. Rimoin, Jean Jacques Muyembe-Tamfum, Thomas M. Lietman, and J. Daniel Kelly. Author and Travis C. Porco contributed to methodology, formal analysis and visualization. Author, Travis C. Porco and J. Daniel Kelly contributed to conceptualization. Author and Nicole A. Hoff contributed to data curation. Author, and J. Daniel Kelly

writing. Travis C Porco contributed to supervision. RW, CAM, SB, BS, MM, PM, NAH, AWR, JJM, TML, JDK and TCP contributed to review and editing.

I acknowledge the following individuals for their contributions to Chapter 3 of my thesis, titled "Modeling the Impact of Trachoma MDA on GU Chlamydia": L. Worden, T. Lietman, T.C. Porco.

**Modeling the transmission of Ebola and GU Chlamydia in
Sub-Saharan African countries under both epidemic and endemic
settings.**

S. Rae Wannier

Abstract

In this dissertation, I used various methods to model the transmission of two infectious diseases, Ebola in an epidemic setting and GU Chlamydia in an endemic setting within in Sub-Saharan Africa. Since 2015, there have been five outbreaks of Ebola Virus Disease (EVD) in several different countries in Sub-Saharan Africa, one of which became the second largest EVD outbreak in history in the setting of a longstanding conflict zone. It is suspected that after violent events occur, EVD transmission will increase; however, empirical studies to understand the impact of violence on transmission are lacking. In my first chapter, I used spatial and temporal trends of EVD case counts to compare transmission rates between health zones that have versus have not experienced recent violent events during the outbreak. In my second chapter, I also sought to use modeling to make outbreak projections, looking at the 2020 outbreak in The Democratic Republic of Congo. I made short and long-term projections for the outbreak in an effort to assess the potential to provide more accurate forecasting for an ongoing outbreak. I also evaluated how the outbreak's timing and course affected the accuracy of such

forecasts. Lastly in my third chapter, I focused on trachoma endemic areas of Sub-Saharan Africa and modeling the impact of annual Trachoma Mass Drug Administration (MDA) with azithromycin upon the prevalence of genitourinary (GU) chlamydia using a compartmental model. Communities that are especially hard hit with Trachoma are almost exclusively poor communities with poor access to sanitation, screening and antibiotics to treat the infection; conditions that may allow for STDs to maintain a high chain of transmission. The dosing of azithromycin for the Trachoma MDA is consistent with dosing given clinically to treat GU chlamydial (GUC) disease, and recent evidence has suggested it reduces the population prevalence.

In my first chapter investigating the potential impact of violent events upon local instability and increased EVD transmission, I collected daily EVD case counts from DRC Ministry of Health for the 2018 outbreak in the Democratic Republic of Congo (DRC). A time-varying indicator of recent violence in each health zone was derived from events documented in the WHO situation reports. I used the Wallinga-Teunis technique to estimate the reproduction number R for each case by day per zone in the 2018–2019 outbreak. I fit an exponentially decaying curve to estimates of R overall and by health zone, for comparison to past outbreaks.

As of 16 April 2019, the mean overall R for the entire outbreak was 1.11. I found evidence of an increase in the estimated transmission rates in health zones with recently reported violent events versus those without ($p = 0.008$).

The average R was estimated as between 0.61 and 0.86 in regions not affected by recent violent events, and between 1.01 and 1.07 in zones affected by violent events within the previous 21 days, leading to an increase in R between 0.17 and 0.53. Within zones with recent violent events, the mean estimated quenching rate was lower than for all past outbreaks except the 2013–2016 West African outbreak. The difference in the estimated transmission rates between zones affected by recent violent events suggests that violent events contributed to increased transmission and the prolonged nature of the second largest EVD outbreak in history.

In my second chapter performing EVD outbreak projections, several mathematical models were used to predict the final outbreak size and weekly incidence for the 2020 DRC outbreak. Projections were commenced prospectively mid-way through the outbreak, and retrospectively applied for the early outbreak. Short-term forecasts were made using two different models: (i) a particle-filter branching-process model and (ii) a naive auto-regression. Final outbreak size predictions were made using four different models: (i) the particle-filter branching-process model, (ii) Theil-Sen regression, (iii) Gott’s Law and (iv) a novel Bayesian branching process model parameterized using prior outbreak sizes and contingent on the current outbreak size. The Bayesian model examined final size distributions across a range of current outbreak sizes, allowing for an examination of parameter fits.

Overall, there were reasonable amounts of variability in the forecasts cre-

ated by different models. For short-term, auto-regression models showed relatively stable steady-state growth in the outbreak, with somewhat larger confidence intervals while the particle-filter branching model projected an outbreak slowly ending in the same period. Final outbreak size predictions increased overall as the outbreak continued. The median expectation among models increased between 2.5–4.0 fold in September over initial expectations from June as the outbreak grew from 34 to 128 cases. The branching-process model was overall the most stable consistent performer, though the Bayesian model was a close second. Including the West Africa outbreak, easily the largest to date, increased the range of predicted outcomes for the DRC outbreak between 40–50%.

In predicting the 2020 Ebola outbreak, the most consistent performing model was the branching process particle-filter model though the Bayesian model did nearly as well, despite being agnostic to the trajectory of the outbreak. Our short-term models consistently predicted the outbreak would grow, though models disagreed over the slowing pace; it will be important to evaluate the performance of these models in future outbreaks to understand these uncertainties. The growth of the outbreak to well over a hundred cases underscores the real risk EBOV poses to the region and the need for improved understanding of outbreak trajectories even with the presence of three approved vaccines.

In my third chapter, I analyzed the impact of Trachoma MDA upon GU

Chlamydia prevalence using an extended compartmental SIS model, accounting for the natural history of GUC, risk structure, and gender. The model includes slowly developing partial immunity. MDA was modelled as an impulsively forced treatment with varying coverage and efficacy.

My model showed that three years of MDA at current levels reduced the prevalence of GUC in all populations by at least 15%. Between annual MDA, the prevalence partially rebounded to pre-treatment levels. With Coverage x Efficacy ≥ 0.80 , the time between MDA treatments was insufficient to sustain transmission, allowing for GUC burden to be suppressed below 1 in 10,000 after 5 rounds for starting prevalence less than 9.2%. When serial non-compliance is increased from 20% to 80%, this target is achieved for starting prevalences below 4.7%, down from 9.2%. Targeting azithromycin treatment only to high-risk individuals reduces the starting prevalences for which target is reached to 1.8%.

My model suggests that MDA could reduce the prevalence of GUC to less than 1 in 10,000 within 5 years time. This reinforces the suggestions of potential additional health benefits of trachoma MDA and points to potential value of screening and disease treatment even in impoverished areas, and suggests testable hypotheses regarding prevalence in endemic areas under treatment.

Contents

- 1 Estimating the impact of violent events on transmission in Ebola virus disease outbreak, Democratic Republic of the Congo, 2018–2019** **1**
- 1.1 Introduction 1
- 1.2 Methods 2
 - 1.2.1 Data 2
 - 1.2.2 Statistical analysis 3
- 1.3 Results 6
- 1.4 Discussion 14
- 1.5 Supplementary Materials 17

- 2 Real-time projections of the epidemic curve of the 2020 Mbandaka Ebola outbreak** **18**
- 2.1 Introduction 18
- 2.2 Methods 21
 - 2.2.1 Data 21
 - 2.2.2 Analysis 21

2.2.3	Auto-Regression model (AR)	23
2.2.4	Gott's law model (GL)	25
2.2.5	Bayesian model (BM)	25
2.2.6	Scoring	27
2.3	Results	27
2.4	Discussion	34
2.5	Supplementary Materials	39
3	Modeling the Impact of Trachoma MDA on GU Chlamydia	47
3.1	Introduction	47
3.2	Methods	49
3.2.1	Mathematical Model	49
3.2.2	Sensitivity Analysis	51
3.2.3	Running the Model	51
3.3	Results	55
3.4	Discussion	61
4	Bibliography	64

List of Figures

- 1.1 **Map of current Ebola outbreak and violent events.**
The 2018-2019 EVD outbreak in northeastern DRC as of 15 April 2019 with confirmed and probable EVD cases depicted by health zone by color. Violent events are represented as (i) an inner circle: direct events where violence was directed at Ebola relief efforts or (ii) an outer circle: all events where violence either indirectly or directly impacted Ebola relief activities. 8
- 1.2 **Estimated initial reproduction number R_{initial} and quenching rate τ for current and past outbreaks.** R_{initial} and τ were estimated for the current outbreak as an overall summary measure, as well as independently in health zones with over a range of inter-zone transmission mixing parameters ω . a) No transmission between zones: $\omega = 0.0$. b) Low transmission between zones: $\omega = 0.2$. c) Medium transmission between zones: $\omega = 0.5$. d) High transmission between zones: $\omega = 0.8$. e) Full transmission between zones: $\omega = 1.0$ 9

1.3 **Wallinga-Teunis estimated R per day and violent events by health zone for symptom onset and reporting dates,** allowing for mixing between regions. R was estimated over a range of inter-zone transmission mixing parameters ω . Violent events are marked using triangles with colors matching the affected district(s). a,b) Low transmission between zones: $\omega = 0.2$. c,d) Medium transmission between zones: $\omega = 0.5$. e,f) High transmission between zones: $\omega = 0.8$ 11

1.4 **Regression estimates of the effect of violent events within the last 21 days on R_{est}** using case reporting and symptom onset date data over a range of inter-zone transmission mixing parameters ω , where $\omega = 0.0$ no mixing between zones and $\omega = 1.0$ full transmission between zones. 12

1.5 **Wallinga-Teunis estimated R per day by health zone for symptom onset and reporting dates,** Regression estimates of the effect of recent violent events upon R_{est} using case reporting and symptom onset date data over a range of inter-zone transmission mixing parameters ω , where $\omega = 0.0$ no mixing between zones and $\omega = 1.0$ full transmission between zones. a,b) Low transmission between zones: $\omega = 0.2$. c,d) Medium transmission between zones: $\omega = 0.5$. e,f) High transmission between zones: $\omega = 0.8$ 13

1.6	Estimated Shannon entropy of 2018–2019 DRC outbreak.	14
2.1	Fitted reproduction rate R_{initial} and quenching rate τ pairs from historic outbreaks with the fitted probability cloud. R_{initial} and τ pairs were sampled evenly from across the probability distribution to feed the particle-filter branching process (PF-BP) model.	28
2.2	Short-term forecasts for new cases in the recent outbreak for one week to two months out from the last day of each month of the outbreak. Results are shown for the particle-filter branching process (PF-BP) and the auto-regression (AR) models. A red dot indicates the actual observed number of new cases from the outbreak. Forecasts are shown for a) one week, b) two weeks, c) one month and d) two months out.	29

2.3	Final outbreak size probability density distributions for all four models at each forecast date i) Bayesian approach, ii) branching process with particle filtering, iii) Gott’s Law, and iv) regression model. Projections are made based on observed case counts of 34 cases on June 30 th , 72 cases on July 31 st , 109 cases on August 31 st and 128 cases on September 30 th . Red bar indicates the current case count, and the gold bar indicates the final case count for the outbreak. The top panels are zoomed in figures on case counts below 150, and the lower panels are zoomed out to show the tails of the distributions.	33
2.4	One week projections of outbreak size probability density distributions for the BP-PF and AR models at each forecast date i) branching process particle-filter (BP-PF) and ii) auto-regression (AR). Projections are made based on observed case counts of 34 cases on June 30 th , 72 cases on July 31 st , 109 cases on August 31 st and 128 cases on September 30 th	40

2.5	Two week projections of outbreak size probability density distributions for the BP-PF and AR models at each forecast date i) branching process particle-filter (BP-PF) and ii) auto-regression (AR). Projections are made based on observed case counts of 34 cases on June 30 th , 72 cases on July 31 st , 109 cases on August 31 st and 128 cases on September 30 th	41
2.6	One month projections of outbreak size probability density distributions for the BP-PF and AR models at each forecast date i) branching process particle-filter (BP-PF) and ii) auto-regression (AR). Projections are made based on observed case counts of 34 cases on June 30 th , 72 cases on July 31 st , 109 cases on August 31 st and 128 cases on September 30 th	42
2.7	Two month projections of new case probability density distributions for the BP-PF and AR models at each forecast date i) branching process particle-filter (BP-PF) and ii) auto-regression (AR). Projections are made based on observed case counts of 34 cases on June 30 th , 72 cases on July 31 st , 109 cases on August 31 st and 128 cases on September 30 th	43
3.1	Levels of the Compartmental Model and the Full Extended Compartmental Model	50

(a)	Natural history of Chlamydia, with slowly developing immunity	50
(b)	Exposure to the MDA treatment	50
(c)	Gender	50
(d)	Risk Structure for behavior	50
(e)	Full Extended Cartesian model	50
3.2	Percent reduction in prevalence after 1, 3 and 5 years of azithromycin MDA.	56
3.3	Percent reduction in prevalence after 1, 3 and 5 annual rounds of azithromycin targeting only the high-risk group.	57
3.4	Effect of Serial Non-Compliance with the MDA upon the reduction in prevalence after 1, 3 and 5 years of azithromycin MDA.	59
3.5	Uncertainty Analysis for Predicting Reduction in Population Prevalence	60

List of Tables

1.1	Violent Events Reported in WHO Situation Reports and included in our analysis.	17
2.1	Average scores for the short-term predictions for the branching process (PF-BP) model and the auto-regression (AR) model.	31
2.2	Predicted final outbreak size distributions across all four models	34
2.3	Summary of all reported historical Ebola outbreaks	39
2.4	Bayesian predictions of final outbreaks sizes	44
2.5	Sensitivity analysis for Bayesian predictions of final outbreaks sizes including the 2014-15 West Africa outbreak	45
2.6	Predicted short-term new case distributions	46
3.1	Table of Model Parameters and Ranges For Population Dynamics and Sexual Behavior With Sourced References	52

3.2	Table of Model Parameters and Ranges for MDA Treatment and Chlamydia Natural History With Sourced References	53
3.3	Sensitivity Analysis	55

Chapter 1

Estimating the impact of violent events on transmission in Ebola virus disease outbreak, Democratic Republic of the Congo, 2018–2019

1.1 Introduction

Since 1976, 10 of the over 34 reported Ebola virus disease (EVD) outbreaks have been in the Democratic Republic of the Congo (DRC) [1,2]. The current 2018–2019 EVD outbreak in northeastern DRC is, as of April 2019, the second largest EVD outbreak in history and the first to occur in a conflict setting [3]. Although its magnitude substantially trails the 2013–2016 EVD outbreak in West Africa, the current outbreak has surpassed epidemic forecasts, particularly mathematical modelling studies that used historical data from prior outbreaks [4–8]. These studies had projected a final outbreak size of up to 1295 cases as of 25 February

2019 while the current outbreak now has reported 1604 cases [4, 5].

Since the EVD outbreak began, there have been reports of violent events that have ranged from destruction of Ebola care facilities and injured and murdered healthcare workers to events unrelated to Ebola care, linked to elections and local unrest [3, 9–14]. Following many of these violent events, Ebola response activities have been disrupted. In some cases, new EVD case counts have increased in the district affected following these events, despite intensive public health interventions comprising the deployment of rapid diagnostic tests, novel therapeutics, contact tracing, and ring vaccination using a vaccine approved for emergency use with an estimated 97.5% efficacy [3, 15, 16]. There has been a growing sentiment that such events may be contributing to EVD transmission, but quantitative analysis is lacking.

Here, we hypothesized that during the current outbreak to date, there had been higher transmission rates in zones that had recently experienced violent events than in zones that had not experienced such events. Furthermore, we also hypothesized that Shannon entropy computed with respect to space, a measure of the spatial spread of cases and the uniformity of their distribution across districts in this epidemic, has increased over time.

1.2 Methods

1.2.1 Data

A time series of case counts were collected from situation reports presented by the DRC Ministry of Health and were confirmed using situation reports posted by the World Health Organization (WHO) [3]. EVD cases were classified as suspected, probable, or confirmed. Suspected cases underwent diagnostic testing and were subsequently classified either as con-

firmed or not confirmed. Cases reported post-mortem were classified as probable based upon their epidemiological history. The symptom onset date and the reporting date of probable, confirmed, and suspected cases were collected from the beginning of the outbreak on 8 May 2018 through 15 April 2019. Beginning on 8 August 2018, data by health zone were also collected for probable and confirmed cases. As the outbreak has continued, cases were re-assigned by the DRC Ministry of Health from their initially reported health zone to the health zone where epidemiological evidence pointed to EVD acquisition. Symptom onset data were later made available going back to the beginning of the outbreak.

Given the dynamic nature of violence in the region and the relatively short generation time of EVD, we undertook a time-varying, health zone-level analysis of the outbreak comparing transmission in zones that had experienced recent violence to those without recently reported violence. We considered a violent event as one reported within the WHO situation reports until 15 April 2019 [3]. We assigned each health zone as having been exposed to violent events or not. After a violent event, we considered the following three weeks as the exposure period. We modeled the effect of violent events within the previous week, two weeks, three weeks, four weeks, and five weeks to determine the sensitivity to the time period chosen.

1.2.2 Statistical analysis

We used the Wallinga-Teunis technique to estimate the number of secondary cases R for each case in the 2018–2019 outbreak [17]. We defined the serial interval as the interval between disease onset in an index case and disease onset in a person infected by that index case. We used both symptom onset and case report data in our analyses given the limitations of each dataset. Historical outbreaks used case report data, though case report data in this

dataset is subject to case reclassification between zones, reflecting epidemiologic knowledge on the place of transmission rather than place of report, that can show up as transmission. In the current outbreak, case report data were updated nearly daily from the official daily totals reported by the DRC Ministry of Health and revised as needed. Thus, while the symptom onset data may be the fundamental definition of the serial interval, these data were not always revised as more accurate information became available. We employed a gamma distribution with a mean of 14.5 days and a standard deviation of 5 days for the serial interval distribution. This models a serial interval of EVD cases approximately ranging from 3 to 36 days, with mean 14 to 15 days [18–20]. Our application of the Wallinga-Teunis method assumes there are no missing cases.

To compare transmission in the current outbreak to past outbreaks, we applied the same estimation technique to reported cases by date from 18 prior outbreaks [?, 7, 8, 21–29]. We estimated the initial reproduction number R_{initial} and exponential decay (or quenching) rate τ for each outbreak by fitting an exponentially decaying curve $R(d) = R_{\text{initial}}e^{-\tau d}$ to the outbreak's estimates of R by day d . This exponential decay "quenching" parameter approximates the often observed reduction in R over the course of an outbreak that may be due to phenomena such as formal control efforts including case finding and quarantine as well as less formal responses including individual behavioral changes or local depletion of susceptibles. This equation approximates temporal change in transmission rate by a smoothed quenching pattern in which R decreases exponentially over time at a rate defined by the quenching parameter. The estimates R_{initial} and τ obtained by this fit are reported for comparison to historic outbreaks and not used for further modeling.

We repeated the Wallinga-Teunis procedure to estimate the reproduction number R separately for each health zone, creating a time series of estimated reproduction numbers for

each zone. When estimating R for each zone, we considered the possibility of transmission between regions. The probability of an inter-zone transmission relative to intra-zone transmission is denoted ω , representing the reduced probability of a case transmitting outside of its health zone as compared to their probability of transmitting within their own health zone. We compared estimates based on values of the inter-zone transmission-mixing parameter ω across its range from 0 (no transmission allowed between regions) to 1 (all cases in all regions transmit equally to all other cases, regardless of region).

Confidence intervals for the Wallinga-Teunis estimated R by health zone were simulated using the calculated probabilities for each case i transmitting to case j to probabilistically assign a transmission link for every case, based on 5000 simulations [17]. Using the distribution of R per day per zone, we then calculated the negative-binomial confidence interval for each day per zone.

Statistical inference on the resulting estimates of R by day per zone was conducted by regressing estimated R on the presence of a recently reported violent events [30,31]. To adjust for autocorrelation, the standard errors of the estimates were estimated using a time-series bootstrap with blocks of 7 days over 2048 replications.

The Shannon entropy of the cumulative number of cases was computed based on all health zones and standard errors were computed from the proportion in each geographic region [32]. Time series bootstrap based on a fixed length of 7 was used to compute standard errors of the time trend, based on ordinary least squares linear regression of entropy on days since the first case.

We conducted all analyses in R (v. 3.5 for Macintosh, R Foundation for Statistical Computing, Vienna, Austria).

1.3 Results

As of 15 April 2019, a total of 1,273 EVD cases had been reported. The outbreak is ongoing as of 11 May 2019. Thus far during the outbreak period, a total of 1,044 cases have been reported in the seven health zones that have experienced violent events whereas 229 cases have been reported in zones without violent events. The outbreak has been centered in health zones of Katwa (32%), Beni (20%), Mabalako (8%), Butembo (9%). Butembo and Beni are the health zones with the most violent events (Figure 1.1).

Since the beginning of the outbreak in May 2018, the average reproduction number (mean R_{estimate}) was 1.11. Using reporting data, after August 8, 2018, the average reproduction number was between 1.12 and 1.23 in the 21 days following a violent event in a district, and between 0.81 and 1.08 in all other cases (Figure 1.3). The difference between reproduction numbers following violent events and not was statistically significant and increased as the mixing between zones was assumed increasingly limited (transmission mixing parameter ω decreased). Even at very high levels of inter-zone transmission mixing, corresponding to relatively homogenous transmission within and across health zones, ($\omega = 0.8$, $p=0.016$), the difference was still statistically significant (Figure 1.3).

We compared the initial estimated R and the quenching parameters of past outbreak to the current outbreak and its respective health zones (Figure 1.2). The inter-zone transmission mixing parameter ω was varied over a range from 0 to 1 (Figure 1.2). The log of the estimated quenching parameter was (-6.32) for the current outbreak was closest, though slightly higher, to that estimated for the 2013–2016 outbreak in West Africa (-6.87) . This was paired with a slightly higher estimated $R_{\text{initial}} = 1.69$ than for the West African outbreak where $R_{\text{initial}} = 1.67$. The (R_{initial}) reported here for the West African outbreak is consistent

with the previous literature [33, 34]. These numbers are consistent with the long trajectory and continued transmission of the outbreak as a whole, although the trend of observed R in the current outbreak is rather different from the West African outbreak where the mean overall R was only above 1.0 early in the outbreak before dropping below 1.0 as the outbreak continued, where here it can be seen to fluctuate with three peaks above $R_{\text{estimate}} = 1.5$ and below $R_{\text{estimate}} = 0.5$ (Figure 1.3) [34]. The estimated R_{initial} and quenching parameters clustered around the current outbreak estimates, some more extreme (lower quenching) and other more like smaller outbreaks of the past. Each health zone was also estimated. While the health zones at the center of the outbreak on average appeared to have lower quenching with Mabalako, Butembo and Katwa often having weaker quenching than the West African outbreak, however, this was not consistent because Beni consistently reported a higher level of quenching from the estimates, indicating the complex geographic distribution of transmission among health zones.

In evaluating a probable mixing parameter to use for evaluating the results from this outbreak, we can look at the evidence from the time series of R_{estimate} by health zone and examine their behavior (Figure 1.3). Both extremes of no mixing ($\omega = 0.0$) or full mixing ($\omega = 1.0$) are unrealistic and lead to inconsistent results. When no mixing is allowed ($\omega = 0.0$) this causes false spikes in R_{estimate} in health zones with low levels of transmission following an increase in transmission in a neighboring health zone as the neighboring cases are unable to account for the spread of transmission to the low transmission zone, and this creates a false apparent spike in their own health zone to compensate for a lack of inter-zone transmission. Even low levels of mixing ($\omega = 0.2$) are enough to remove false spikes in R_{estimate} that we see when no mixing is allowed ($\omega = 0.0$). When full mixing is allowed ($\omega = 1.0$) this leads to all health zones having identical R_{estimate} at each time point. This herding behavior is

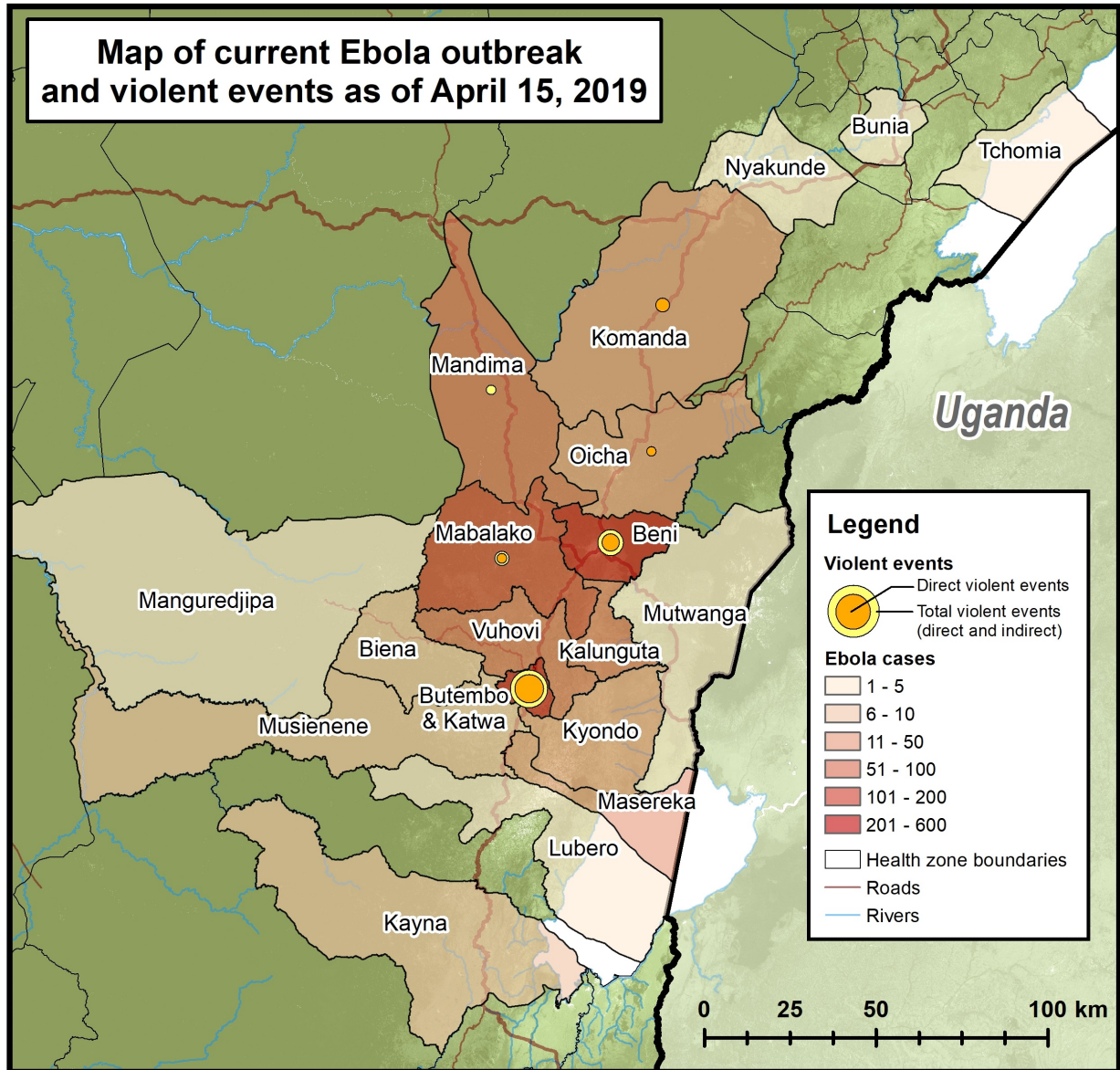


Figure 1.1: Map of current Ebola outbreak and violent events. The 2018-2019 EVD outbreak in northeastern DRC as of 15 April 2019 with confirmed and probable EVD cases depicted by health zone by color. Violent events are represented as (i) an inner circle: direct events where violence was directed at Ebola relief efforts or (ii) an outer circle: all events where violence either indirectly or directly impacted Ebola relief activities.

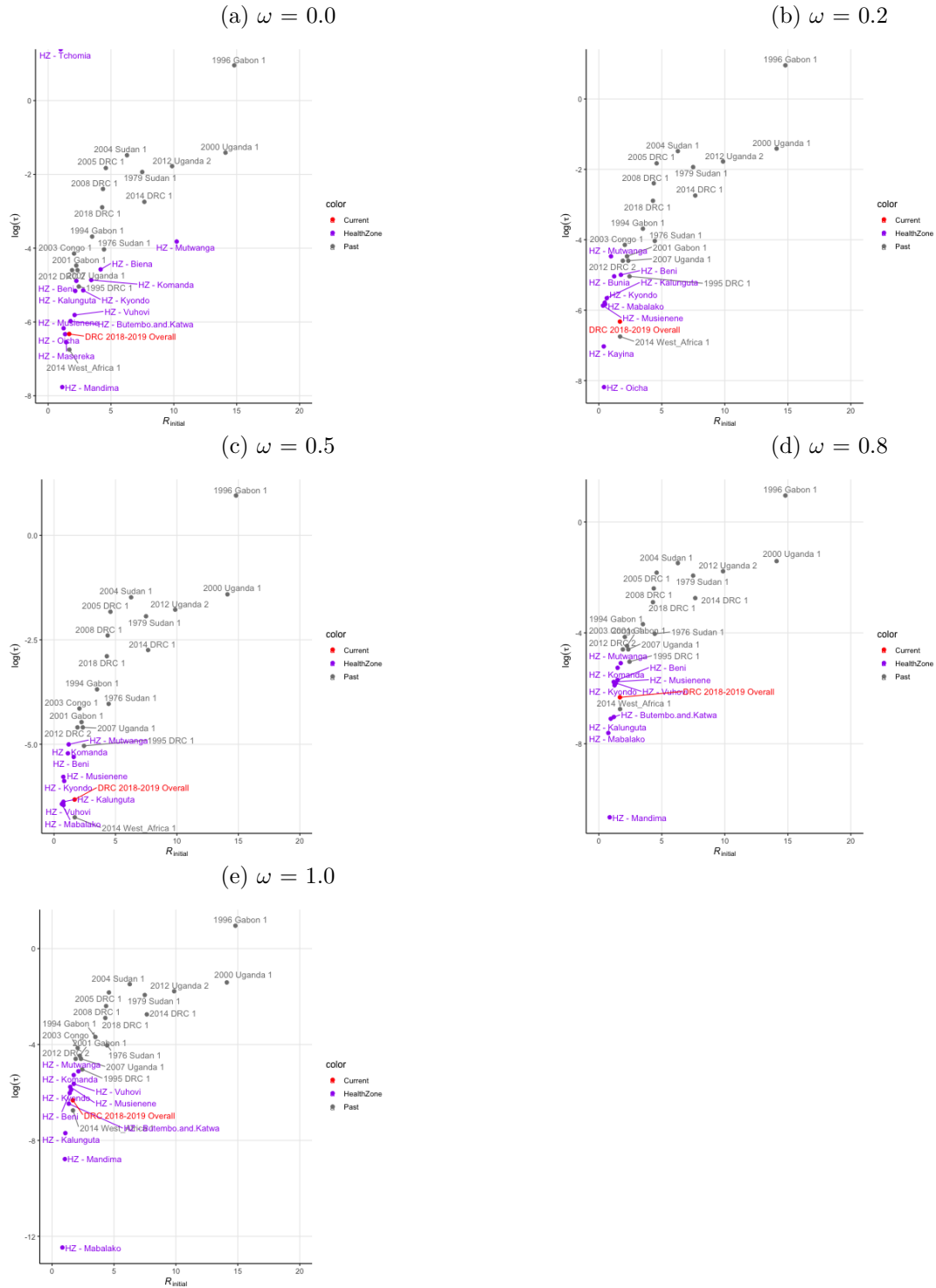


Figure 1.2: **Estimated initial reproduction number R_{initial} and quenching rate τ** for current and past outbreaks. R_{initial} and τ were estimated for the current outbreak as an overall summary measure, as well as independently in health zones with over a range of inter-zone transmission mixing parameters ω . a) No transmission between zones: $\omega = 0.0$. b) Low transmission between zones: $\omega = 0.2$. c) Medium transmission between zones: $\omega = 0.5$. d) High transmission between zones: $\omega = 0.8$. e) Full transmission between zones: $\omega = 1.0$.

strongly seen still when ($\omega = 0.8$) at unrealistically high levels. Even at a middle level of mixing ($\omega = 0.5$) it appears that health zones herd too strongly to the health zone driving transmission, as spikes in transmission at days 200, 235 and 304 cause all health zones reporting cases to spike as well, when it is most probable the increase in transmission is being driven only by a few health zones and the others only responding to the initial change in transmission. Though we do not make an effort to formally estimate a probable mixing parameter, it would be reasonable to consider the estimates taken with $\omega = 0.1$ to $\omega = 0.5$ as these are the probable limits of the range for the mixing parameter in this outbreak.

Using reporting data, at $\omega = 0.1$ to $\omega = 0.8$ we see a difference in the R_{estimate} by violent events (Figure 1.4). Looking only at ω between 0.1 and 0.5, we see an increase in R_{estimate} following violent events of 0.44 (95% CI: 0.28, 0.61, $p < 0.001$) to 0.28 (95% CI: 0.16, 0.40, $p < 0.001$). However, when considering the symptom onset data the strength of the effect of recently reported conflict was reduced, though still significant overall with violent events leading to an increase in R_{estimate} of 0.17 (95% CI: 0.02, 0.32, $p = 0.026$) to 0.20 (95% CI: 0.10, 0.30, $p < 0.001$).

To assess the sensitivity of the lag time chosen after a violent event, we looked at the effect of violent events reported in increments of increasing weeks. Looking at the case report data with $\omega = 0.3$, the overall effect size is fairly constant whether we consider a period of 7 days or 35 days, with an increase in R_{estimate} of 0.37 (95% CI: 0.17, 0.57, $p < 0.001$) and 0.40 (95% CI: 0.28, 0.52, $p < 0.001$), respectively.

When looking at the onset data with $\omega = 0.3$, the strength of the effect of recent violent events is strongest when it is evaluated over a full 21-28 days, as R_{estimate} increases from 0.13 (95% CI: -0.05, 0.31, $p = 0.074$) to 0.21 (95% CI: 0.09, 0.33, $p < 0.001$) as the lag increases 7 to 21 days and then becomes relatively stable.

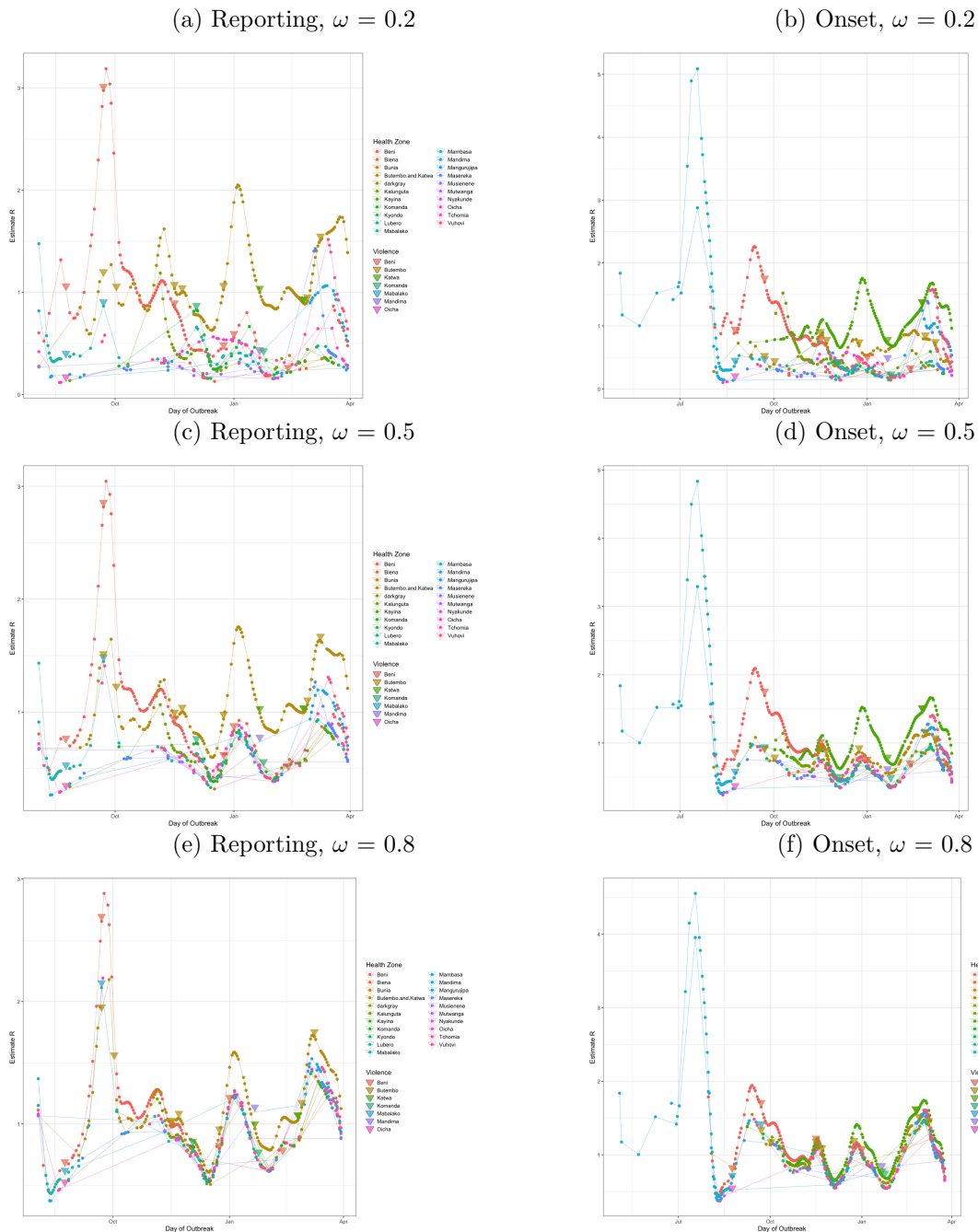


Figure 1.3: Wallinga-Teunis estimated R per day and violent events by health zone for symptom onset and reporting dates, allowing for mixing between regions. R was estimated over a range of inter-zone transmission mixing parameters ω . Violent events are marked using triangles with colors matching the affected district(s). a,b) Low transmission between zones: $\omega = 0.2$. c,d) Medium transmission between zones: $\omega = 0.5$. e,f) High transmission between zones: $\omega = 0.8$.

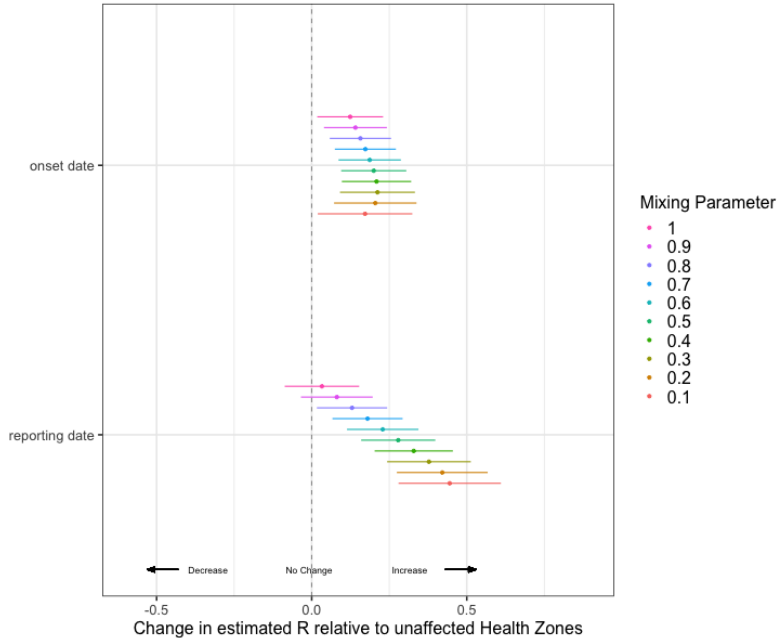


Figure 1.4: **Regression estimates of the effect of violent events within the last 21 days on R_{est}** using case reporting and symptom onset date data over a range of inter-zone transmission mixing parameters ω , where $\omega = 0.0$ no mixing between zones and $\omega = 1.0$ full transmission between zones.

We did not consider event times longer than 35 days apart as many of the affected health zones have events that occur less than one month apart, sometimes occurring as little as one week apart. For these health zones, increasing duration beyond this point does not increase the period of time considered as being impacted by violent events, and thus we lose much of our ability to further distinguish between transmission in zones with recent violent events and just comparing zones affected by violent events to those unaffected.

Figure 1.6 shows the estimated Shannon entropy over the course of the epidemic. The estimated entropy was 0.99 ± 0.17 on 17 August 2018, rising to 1.88 ± 0.04 by 15 April 2019. We find evidence of an increasing trend ($p < 0.01$, time series bootstrap), showing less concentration of cases over time. Some increase in entropy is expected early in an epidemic, as cases begin to appear in new health zones. Later in an epidemic, a decrease in entropy could occur if there were an increased concentration of cases in a few of the affected health

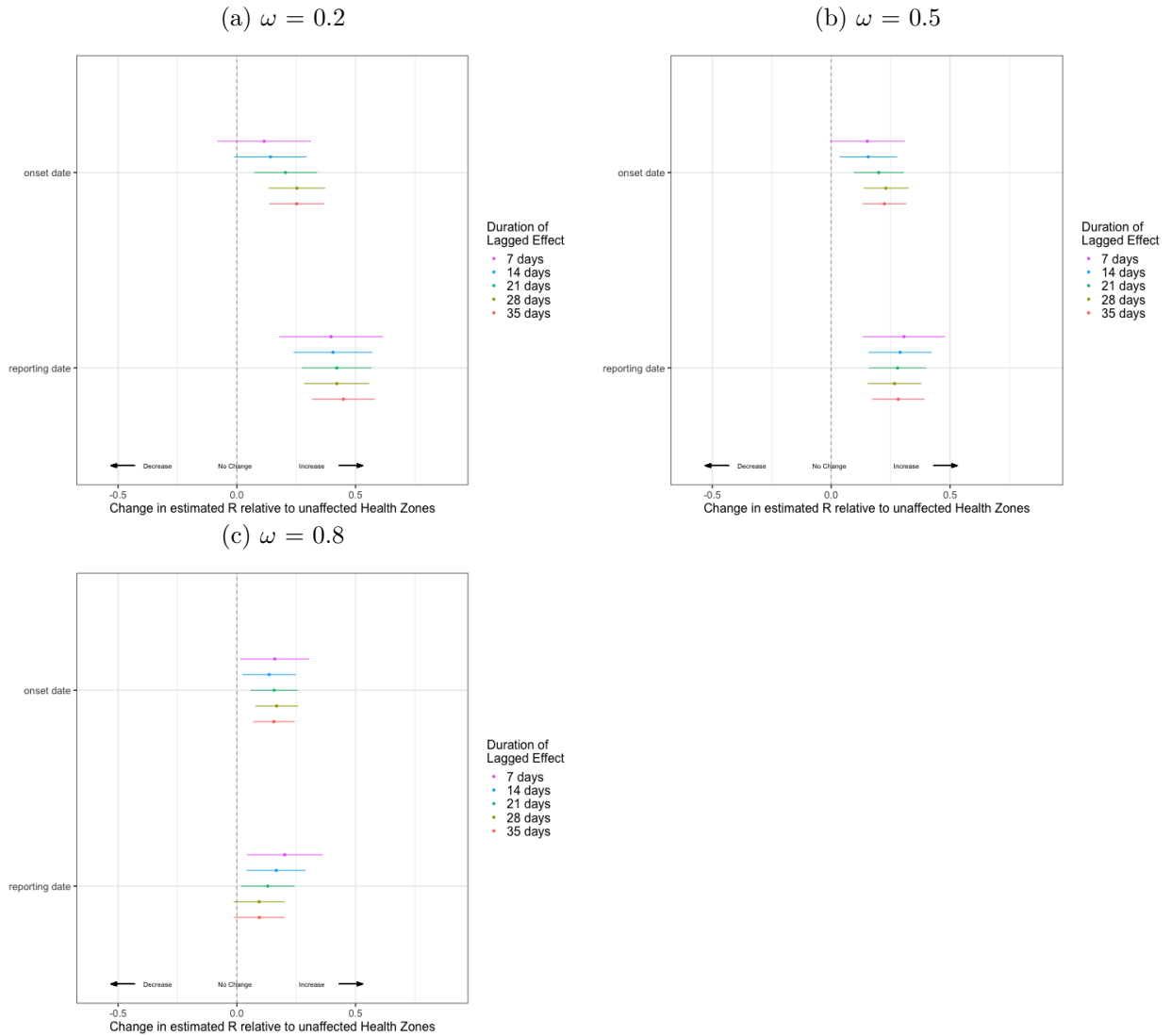


Figure 1.5: **Wallinga-Teunis estimated R per day by health zone for symptom onset and reporting dates**, Regression estimates of the effect of recent violent events upon R_{est} using case reporting and symptom onset date data over a range of inter-zone transmission mixing parameters ω , where $\omega = 0.0$ no mixing between zones and $\omega = 1.0$ full transmission between zones. a,b) Low transmission between zones: $\omega = 0.2$. c,d) Medium transmission between zones: $\omega = 0.5$. e,f) High transmission between zones: $\omega = 0.8$.

zones. It is possible that the continued spatial spread of the current outbreak, which is second only to the West African outbreak, is contributing to the difficulty in controlling this outbreak.

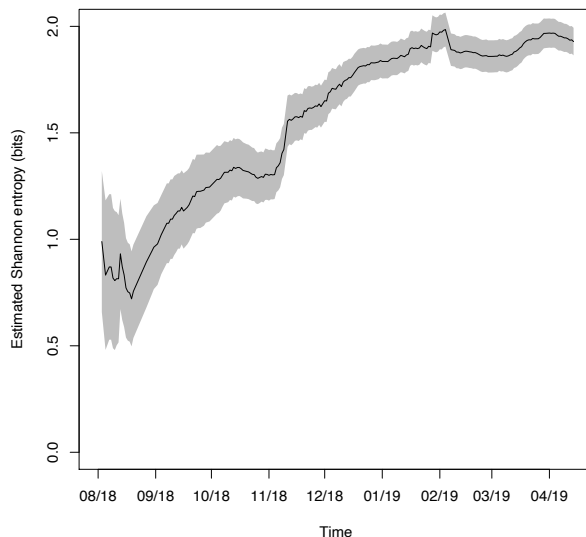


Figure 1.6: **Estimated Shannon entropy of 2018–2019 DRC outbreak.**

1.4 Discussion

Among health zones situated within the EVD outbreak in DRC, we found that the EVD transmission rate (reproduction number R) was higher following violent events. The outbreak was subcritical ($R < 1.0$, non-sustaining transmission) in zones without violent events reported by WHO [3], while it was supercritical (with estimated $R > 1.0$, continued transmission) in zones with reported violent events, suggesting that ongoing violence is likely perpetuating an otherwise declining outbreak.

Our findings suggest that violent events increased transmission in the weeks following a violent event, and that this effect may be sustained for many more weeks. After the

destruction of the Ebola care facility in Katwa, for example, over one month was needed before Ebola virus disease relief efforts were fully resumed. Our time series regression found an effect of violent events on estimated R across both symptom onset data and case report data, across all plausible levels of inter-region transmission (Figure 1.3), and lagged follow-up periods of 14 or more days. The consistency of the effect of recent violent events across data sets strongly supports the idea that violence is indeed contributing to the increased transmission and ongoing nature of this outbreak [35–39]. More research is needed to confirm and further understand how the frequency and intensity of events may affect the contributions of violence to EVD transmission.

There are several limitations in our analysis. Cases may have escaped detection and reporting [40,41]. If cases are missing in a biased way that systematically unreports cases in certain areas or at certain times it could lead to a biased estimate of the effect, though the direction of the bias is unclear. Additionally, if there is heterogeneity in the reporting delays, this could bias the effect estimates made using the case report data, though this objection would not apply to the symptom onset data. The nature and causes of violent events can be quite different. Moreover, such events can affect different numbers of people, can vary in geographic scope as well as the duration of their impact. Other non-violent events not considered in this analysis, such as strikes, may also be contributing to ongoing transmission. Unmeasured causes or determinants of violence in the region may also drive variation in transmission or reporting in an unmeasured way. The epidemic curve in each health zone describes relatively few cases, making interpretations of specific features of their epidemic curves unwise. Our analysis relied on WHO reporting of violent events, and more accurate quantification of these events may be possible. Note that the assumption of exponential decay (quenching) of R_{initial} may not accurately characterize this epidemic; however, this

was only used for comparisons to past outbreaks, and the estimates of the effect of violent events are not affected by this. While we have shown that the ongoing violence has likely hampered control and contributed to this becoming the second largest epidemic, our efforts to quantify the role of violence should be interpreted with caution. These considerations also suggest that meaningful interventions must consider the social and political determinants of armed conflict in the DRC, including the legacies of colonialism, and other forms of historical and ongoing violence [35–39, 42–44].

On 16 April 2019, the Democratic Republic of the Congo’s Minister of Health addressed the worsening outbreak and indicated that with “the difficult security situation, this epidemic had gone beyond public health.” [45] While this may have been the first time an EVD outbreak occurred in an active conflict zone, it is unlikely to be the last time that violent conflict contributes to the prevention of epidemic decline. As EVD surges in DRC, a polio outbreak surges in a conflict area of Nigeria. Ebola virus and other infectious disease outbreaks have the potential to become neglected crises in conflict settings if left unchecked. Despite immense security challenges, Ebola responders and people of the affected area work tirelessly under dangerous conditions and deserve great respect, gratitude, and protection for their ongoing work to contain this public health disaster.

1.5 Supplementary Materials

Table 1.1: **Violent Events Reported in WHO Situation Reports** and included in our analysis.

Date	Setting	Health Zones	Description
24 Aug. 2018	Community	Mabalako Oicha Beni	Community resistance
22 Sep. 2018	Community	Mabalako Butembo Beni	Armed conflict unrelated to EVD response 'ville morte'
2 Oct. 2018	Community	Butembo	Attack on health care workers
16 Nov. 2018	Community	Beni Butembo	Militia attacks, vaccination activities paused after security incidents
22 Nov. 2018	Community	Butembo	Two huts burned due to Ebola suspect
3 Dec. 2018	Community	Komanda	Aggression to SDB teams with 3 of 4 burials missed due to community resistance
24 Dec. 2018	Community Facility	Beni Butembo	Security issues, points of entry burned, SBDs delayed activities
1 Jan. 2019	Community	Beni	Community resistance and 'ville morte' in Beni around SBDs
21 Jan. 2019	Community	Katwa Mandima	Political violence unrelated to the EVD response
24 Jan. 2019	Community	Komanda	Community resistance
12 Feb. 2019	Community	Beni	Surveillance activities disrupted to security incidence
24 Feb. 2019	Community Facility	Katwa	MSF ETC is burned, lab and field activities disrupted
27 Feb. 2019	Facility	Butembo	MSF ETC attacked
9 Mar. 2019	Facility	Butembo	Shooting at ETC after re-opening

Chapter 2

Real-time projections of the epidemic curve of the 2020 Mbandaka Ebola outbreak

2.1 Introduction

A recent outbreak of Ebola virus disease was reported by the WHO on 1 June 2020, in the Democratic Republic of the Congo in the Équateur region [46], which has caused 130 cases and 55 deaths across 15 health zones by the time it was declared officially over on November 18, 2020 [47]. This outbreak was subject to a series of exacerbating and mitigating factors making it hard to predict, and was impeded by competing priorities and insufficient resources. Coming at the same time as the global COVID-19 pandemic, which had caused at least 11,000 confirmed cases in the DRC before the start of the outbreak [48], as well as a large measles outbreak [49]; the outbreak response suffered some from a strain in attention and resources from the central public health department. The outbreak response

also faced challenges of community acceptance of public health interventions, including the ring vaccination program and refusal of care [50]. A greater challenge was simply logistical, with this outbreak covering a broad, largely rural, geographic area, often without good transportation and access for medical and health workers [46]. All of these exacerbating and extenuating circumstances are challenging to model, making it difficult to predict an outbreak's size and response; and yet doing so can help inform the deployment of resources towards control early on. Though most historic outbreaks have been small or mid-size (fewer than a thousand cases), two outbreaks in particular have grown to be extremely large, the recent 2018 outbreak in the DRC and the 2014–16 West African outbreak [7, 8, 21–29, 51–53]. The dynamics controlling final outbreak sizes are poorly understood, and may not be limited exclusively to the regional instability and inexperience with Ebola which are believed to have allowed these outbreaks to flourish [54]. Other factors, such as population density, funerary customs, transport patterns and medical access and trust have all been identified as playing historical roles in Ebola transmission. Predictions of final outbreak sizes can help public health response teams evaluate readiness in the event that an outbreak becomes uncontrolled and prolonged and take early steps to improve their capacity to respond, while also using short-term forecasts to evaluate the scope of operation for immediate responses, including hospital readiness and capacity.

Past outbreaks may provide insight into expected outbreak sizes, despite the changing context of Ebola viral disease (EVD) outbreaks over time, as knowledge of the disease and tools to fight EVD outbreaks have progressed; most particularly the emergence of highly effective vaccines [55]. Each outbreak in the past decade has taken place in a different setting— notably control in the recent 2018–20 outbreak in the DRC was hampered by widespread conflict in the outbreak zone [54], while vaccination efforts in this current outbreak were

hampered both by the remote location and far flung nature of the outbreak. Despite improvements in treatments and vaccination, the four most recent outbreaks have included the two largest outbreaks to date, and underscore the need to consider the ongoing threat that Ebola poses, and the need for updated models that consider these larger outbreaks in making projections for newly emerging outbreaks. Underestimating outbreak trajectories is likely to prove more harmful than overestimations, potentially leaving a local public health team underprepared and under-resourced to adequately respond to a growing outbreak, though overly pessimistic projections are not desirable, as they may unnecessarily divert resources.

Highly accurate forecasts of small, noisy outbreaks may prove elusive [56, 57], and sufficient data are not available to assess how novel events (such as the COVID-19 pandemic) may affect ongoing control and transmission. However, projections derived from past outbreaks may prove a helpful benchmark. In particular, short-term forecasting may prove especially valuable in providing additional feedback to a local disease response in the effectiveness of their responses while avoiding overconfidence in apparent successes. In this paper, we extend our prior work to provide projections [58] for this recent outbreak, seeking to assess the potential to provide more accurate long-term and short-term forecasting for ongoing outbreaks. We also develop a novel Bayesian branching process model using past outbreak trajectories to project new ebola outbreaks. Previous studies have indicated that short-term outbreak predictions are often more accurate and reliable, particularly in low information settings, than final or long-term projections but that, especially medium and long-term model projections, improved with increased data availability and model complexity [59]. Here we evaluate the performance of outbreak forecasts based on past outbreak data, to understand how timing and the outbreak course affect the accuracy of such forecasts and assess whether certain models may perform better under different conditions.

2.2 Methods

2.2.1 Data

Data for the recently ended pandemic were collected from the Ministry of Health of the Democratic Republic of the Congo. Time series for past outbreaks were used, as previously compiled [58] and their inclusion in the prediction models is summarized in the supplementary materials (Table 2.3). A total of twenty historical outbreaks were compiled with time series information for inclusion into the prediction modeling. Forecasts were made prospectively at the end of each month starting in August while the outbreak was ongoing and were retrospectively performed for the months of June and July. In an effort to limit subconscious bias, model development was mostly performed using only the initial severely truncated time series of cases from the first few weeks during development in June and July, and was fully updated in August as we started making prospective predictions.

2.2.2 Analysis

For the short-term forecasts, forecasts looking forward up to two months, we used two models: (1) a branching process model we used for previous Ebola virus disease outbreaks [58, 60] and (2) an autoregression.

We use four methods to estimate the final outbreak size: (1) a branching process model (same as above), (2) Thiel-Sen Regression [58], (3) Gott’s law [61], and (4) a parametric Bayesian method projecting the final size conditional on the current outbreak size and parameterized using prior outbreak sizes.

For all simulation models, outbreaks were terminated once they hit a ceiling of 40,000 cases to improve computational efficiency, and in order to facilitate inter-model comparisons,

the distributions for the regression models were truncated at 40,000 with all remaining tail probability given to an outbreak of precisely 40,000 cases.

Particle-Filter Branching Process Model (PF-BP)

In brief, using a stochastic branching process model previously written [60], we modeled the transmission of Ebola virus (EBOV) using estimated transmission parameters fit to historical EVD outbreak and a particle filtration step for agreement with the current case-count trajectory [58, 60].

To begin, the distribution for the reproduction number and the outbreak trajectory was estimated from fitting an observed relationship to historical outbreaks of size 10 or greater [7, 8, 21–29, 51–53]. The first historical outbreak in each country that met our inclusion criteria (i.e. the first outbreak that was at least ten cases) was excluded from the distribution, to acknowledge the impact that public health experience and community sensitization play in affecting an outbreaks trajectory, as the DRC is now in its eleventh outbreak.

To estimate the reproduction number for the past and current outbreaks we used the Wallinga-Teunis technique to estimate R for each case [17], using a gamma serial interval distribution with a mean of 14.5 days and a standard deviation of 5 days consistent with a serial interval that ranges from 3 to 28 with a mean between 14 to 15 days [18–20]. The R_{initial} and the quenching rate τ were estimated using an exponentially quenched curve given by $R_t = (R_{\text{initial}})e^{-t*\tau}$ fit to each historic outbreak’s time series of R . This models the assumption that transmission rates gradually decline over the course of an outbreak due to a wide range of effects including public health response measures, changes in behavior and potential local depletion of susceptibles.

To project cases in the current outbreak, the stochastic branching process used a nega-

tive binomial secondary case distribution with a mean equal to the simulated reproduction number R as a function of day of the outbreak, and its variance controlled by a dispersion parameter k [62,63]. We sampled across dispersion parameter $k = 0.3, 0.5, \text{ and } 0.7$, to reflect the heterogeneity of transmission observed in past EBOV outbreaks [64–66]. The $(R_{\text{initial}}, \tau)$ pairs sampled uniformly from across the fitted distribution shown in Figure 2.1 informed by the past outbreaks.

Each simulation was initiated with a single case. The simulation was run multiple times with each instance producing a unique epidemic trajectory, generated by the above branching process with the given parameters R_{initial}, τ , and k . These instances were then filtered for trajectories with outcomes matching the cumulative case counts for the current 2020 outbreak trajectory, i.e. simulated trajectories were filtered at two timepoints (i) two weeks before the current date of projection and (ii) at the current date of projection for case counts matching, or within 5 cases of matching, the outbreak size at the respective time points. The purpose of this matching is to select for simulated outbreaks that are both the same size as the current outbreak being predicted, and also one that has recently been growing at a similar pace, while not overly restricting the trajectory fit so that it is computationally practicable. All instances were terminated after 40,000 cases, avoiding run-away epidemics and allowing for efficient computation. As the beginning date of the outbreak is unknown, the first value was allowed to match on any day, reflecting the uncertainty in the precise date the outbreak began.

2.2.3 Auto-Regression model (AR)

An auto-regression model was also fit to the Ebola outbreak time series. A negative binomial autoregressive model was chosen through a validation process (described below)

to forecast additional new case counts at time points one week, two weeks, one month, and two months from the date of forecast. To adjust for disparities in the frequency of case reporting in historic outbreaks (some outbreaks have daily, weekly or monthly data), the data were weighted by the inverse square root of the number of observations contributing to the model. Only historical outbreaks greater than size 10 were considered. Models considered included a two and four-week autoregressive model, as well as parameters considering the historic cumulative case counts (probable and confirmed) at different time points, and logs of historic case counts. When historic case counts for specific dates were missing, each missing case count was interpolated from the two nearest case counts using Stineman’s interpolation function [67], which preserves monotonically increasing case counts while allowing for geometric changes in growth when the rate of reported new cases between subsequent points appears to increase or decrease. All models were fit and the final model was chosen to be the one with the smallest mean-squared error after ensuring residual independence from the predicted case counts. After model fitting and validation, the final model chosen was a negative binomial auto-regression with two- and four-week auto-regressive processes, with a trend adjustment for the ratio of cases from the previous two weeks to four weeks which helps project a downward or upward trend in cases.

Theil-Sen Regression (TS)

We conducted a simple regression forecast based solely on outbreaks of size 10 or greater, based on the time series of prior outbreaks [7, 8, 21–29, 51–53]. Due to the fact that the beginning of the time series were often poorly characterized, all time series were aligned on the day they reached 10 cases, designated as day 0 for the purposes of this regression. Nonparametric Theil-Sen regression (R package `mblm`) was used to project the final outbreak

size based on values of the outbreak size at an earlier time point of choice and the elapsed time from that point to the day the outbreak first reported 10 cases. Stineman’s interpolation was again used between reporting dates to obtain the number of cases reported on the day of forecast, and we reported the median and 95% central coverage intervals for the prediction distribution conditional on the value being no smaller than the observed value on the day of forecast. Finally, we reported the median and 95% central coverage intervals for the prediction distribution. More details can be found elsewhere [58].

2.2.4 Gott’s law model (GL)

Gott’s law approaches prediction with the assumption that, all things being equal, we observe an epidemic at a random point during the epidemic. Thus a specific observation is equally likely to represent 10%, 50% or any other percent of the total cases and final outbreak size. This model serves as a reasonable neutral benchmark for projections, as Ebola outbreaks are often first identified with only a small number of cases, most with fewer than ten cases, and none with more than thirty at first report [7,8,21–29,51–53]. With Gott’s Law, we assume we have no special knowledge of our position on the epidemic curve [61]. If we assume a non-informative uniform prior for the fraction α ($0 < \alpha \leq 1$) of the epidemic included in the last available report with current case count Y_0 , the corresponding probability density function for the final size $Y = Y_0/\alpha$ is $P(Y = y) = Y_0/y^2$, $Y_0 \leq y$. We constructed a probability mass function by integrating the probability density for each individual day.

2.2.5 Bayesian model (BM)

We developed a novel Bayesian forecast of the final size using the same branching process transmission model above. We calculate the conditional distribution of the final size given the

current size as follows. We draw values of R , τ , and k from a prior distribution. Specifically, we chose R from a uniform distribution on the range 0.75 to 3, and we chose $\log(\tau)$ from a uniform distribution on the range $\log(0.01 \text{ days}^{-1})$ to 0 days^{-1} . Finally, we chose k uniformly from the range 0.1 to 1, reflecting substantial over-dispersion as seen in previous studies [64–66,68]. A total of 50,000 different parameter sets were drawn from this prior. Given each parameter set, 8,192 simulations were conducted in an adaptive manner running an initial set of 32 simulations for each parameter set, and running the additional simulations if the initial runs yielded at least one outbreak that was larger than two cases or less than an initial ceiling of 10,000. A secondary adaptive phase was performed for each parameter set running 256 simulations and continuing to the last full phase if there was at least one simulation greater than twenty cases (lower than any of the prediction forecast points) or at least 5% of simulations yielded an outbreak lower than the ceiling of 40,000. The final simulation brought each parameter set to 8,192 with a case ceiling of 40,000, yielding a collection of simulated final sizes. This adaptive simulation was done to improve computational efficiency by tossing parameter sets early on where outbreaks reliably either failed to materialize or nearly always escaped which were fairly certain not to prove a high likelihood simulation. The branching process terminated after 40,000 cases to cover historic outbreaks and improve computational efficiency by preventing runaway epidemics. We then fit the simulated outputs to a Gaussian distribution, and used this to estimate the log-likelihood of the observed outbreaks with a penalty for simulated outbreaks escaping beyond our stoppage criterion. The West Africa outbreak of 2014–16 was omitted from the principal predictions due to difficulties finding and fitting triples for the over-dispersed distribution but included it as a sensitivity analysis (Table 2.5). For consistency we also eliminated the West Africa outbreak from the other predictions models informed using past outbreak distributions. Finally, we

used the sampling/importance resampling algorithm to create a sample from the posterior distribution of outbreak sizes [69]. To establish stability in our prediction of final outbreak sizes, we performed the above simulation four times over (Supplemental Materials)

2.2.6 Scoring

In order to compare model performance, projections were scored using the natural logarithm of the probability assigned to the the subsequently reported case counts for the short-term projections or the final outbreak size for the final size projections. Short-term projections were scored and compared at multiple timepoints: one-week, two-weeks, one-month and two-months forward.

2.3 Results

The particle-filter branching process (PF-BP) model and the auto-regression (AR) models were used to perform short-term forecasts for the current outbreak (Figure 2.2, Table 2.1). Both of these models drew R_{initial} and τ pairs informed from the fit of pairs from all historic outbreaks (Figure 2.1). Nearly all of the historical outbreaks have R_{initial} estimated between 0.5-8, with a few outliers between 15-20. These estimates likely reflect a selection only for outbreaks with more than at least ten cases in a disease with high heterogeneity [64–66], while the outlying estimates describe outbreaks that were characterized by extremely steep increase in cases at the onset due to a superspreading event, such as a large funeral for the index case or the repeated sharing of contaminated needles in a hospital following the index case seeking care, with slower transmission following afterwards [8, 62, 70]. The AR model shows a relatively stable steady state growth in the outbreak, projecting on June 30:

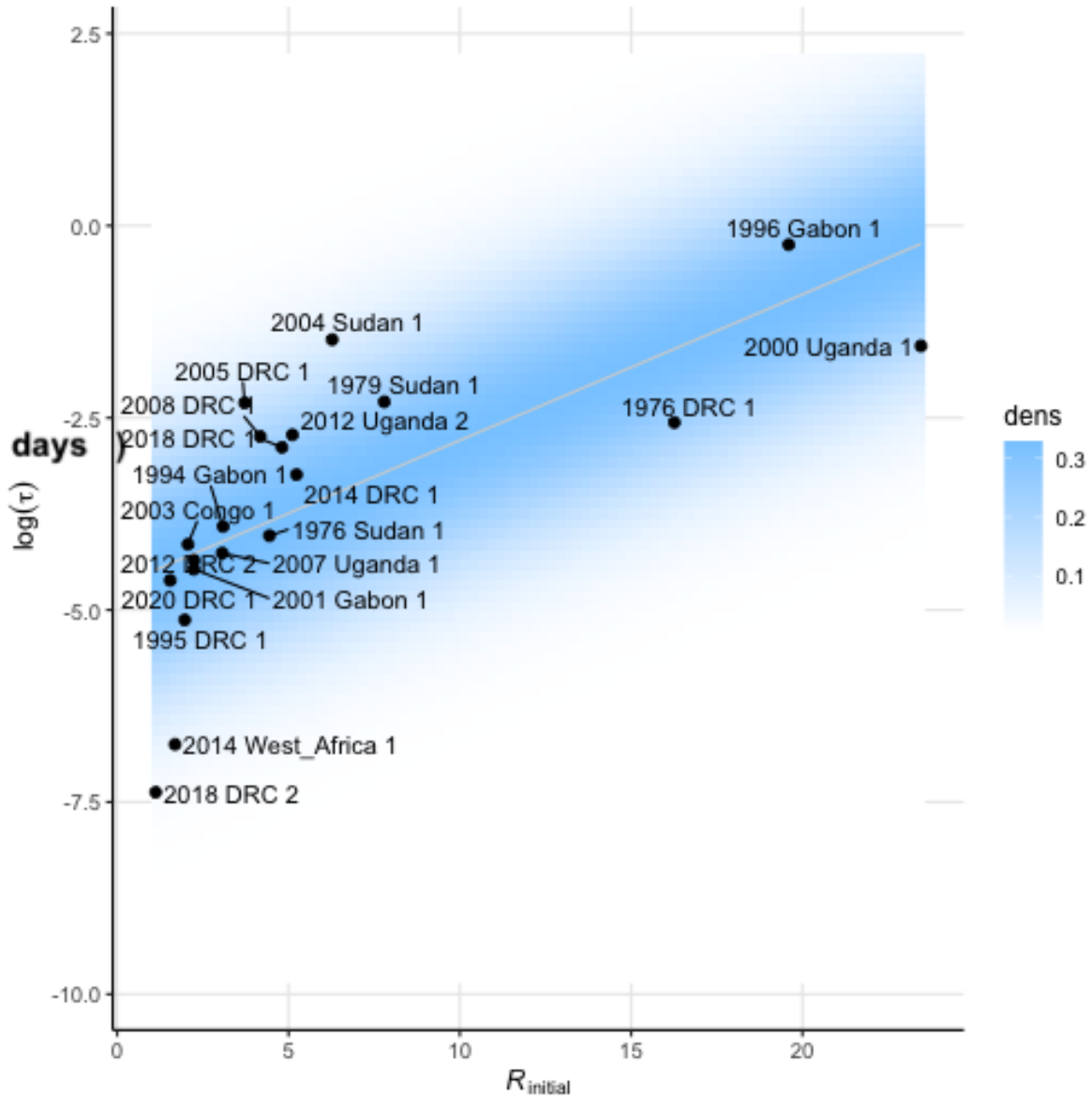


Figure 2.1: **Fitted reproduction rate R_{initial} and quenching rate τ pairs from historic outbreaks with the fitted probability cloud.** R_{initial} and τ pairs were sampled evenly from across the probability distribution to feed the particle-filter branching process (PF-BP) model.

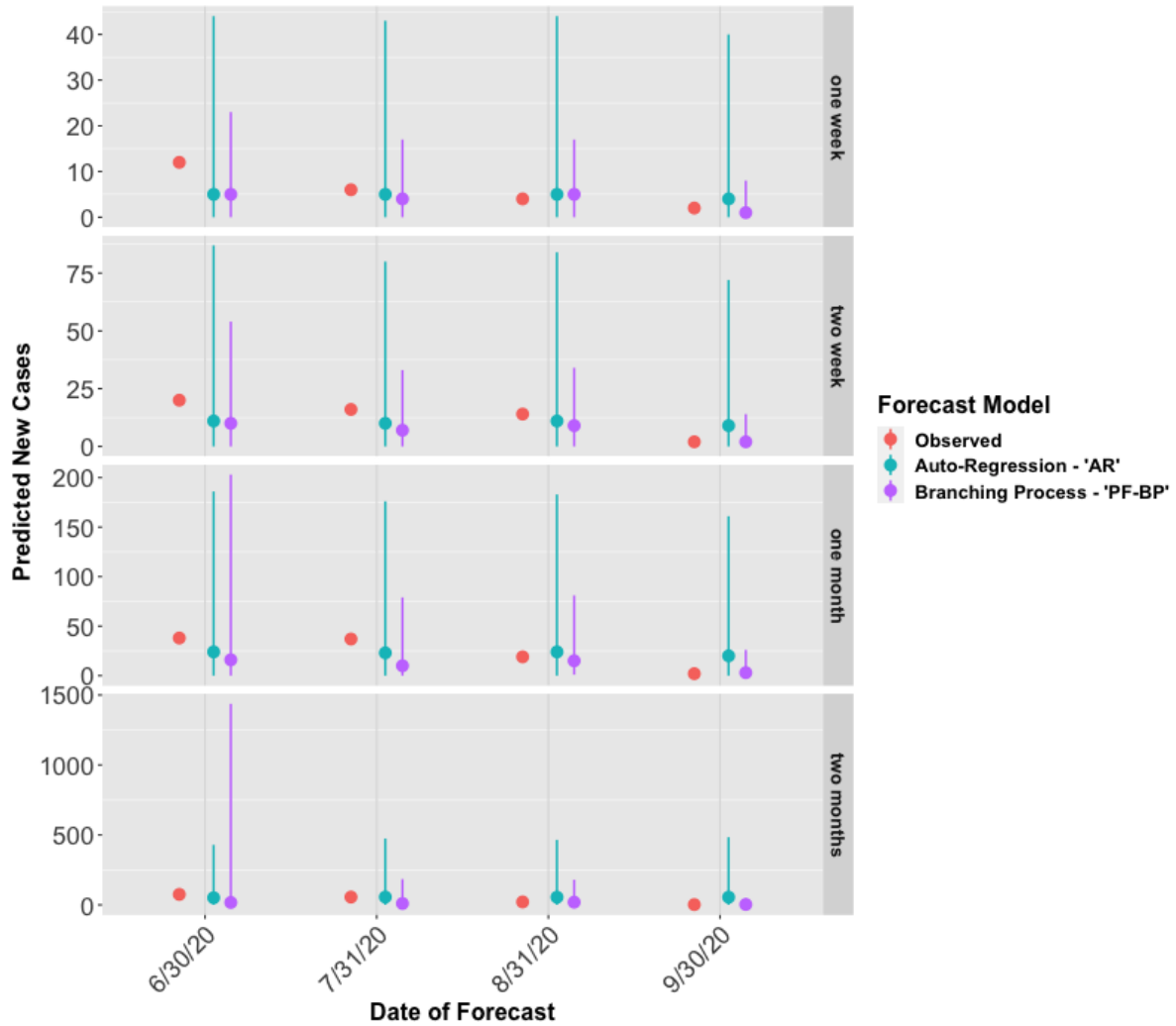


Figure 2.2: **Short-term forecasts for new cases in the recent outbreak for one week to two months out from the last day of each month of the outbreak.** Results are shown for the particle-filter branching process (PF-BP) and the auto-regression (AR) models. A red dot indicates the actual observed number of new cases from the outbreak. Forecasts are shown for a) one week, b) two weeks, c) one month and d) two months out.

5 new cases (95% CI: 0–44), 11 new cases (95% CI: 0–87), 24 new cases (95% CI: 0–186) and 51 new cases (95% CI: 0–429) at one week, two weeks, one month and two months out respectively. Additional detail on short-term predictions can be found in the supplement (Table 2.6, Figures 2.4, 2.5, 2.6, 2.7)). The PF-BP model generally projected an outbreak slowly coming to an end and a smaller spread in probable outcomes than the regression except for the beginning of the outbreak when its projections were far less certain. Early in the outbreak the PF-BP projections were rather broad, projecting 5 new cases (95% CI: 0–23), 10 new cases (95% CI: 0–54), 15 new cases (95% CI: 0–202) and 17 new cases (95% CI: 0–1,427) at one week, two weeks, one month and two months out respectively. However in late in the outbreak on September 30 the projections covered an incredibly narrow interval, while the median expectation of new cases was lower than earlier in the outbreak projecting an outbreak swiftly coming to a close with 1 new cases (95% CI: 0–7), 2 new cases (95% CI: 0–13), 3 new cases (95% CI: 0–26) and 3 new cases (95% CI: 0–46) at one week, two weeks, one month and two months out respectively.

Four prediction models were also used to estimate final outbreak size, including our novel Bayesian model. Overall, final size prediction models showed a larger final outbreak size as the outbreak continued, increasing from a median final size estimate of 45–108 on June 30th with 34 reported cases to a median final outbreak size between 133–417 cases on September 30th with 128 reported cases. The branching process model (PF-BP) was overall the most stable consistent performer (Table 2.2), though the Bayesian model (BM) did nearly as well across the whole outbreak. The Theil-Sen Regression (TS) and the Gott’s Law (GL) both performed quite well at the end of the outbreak, but performed rather poorly at the onset.

The newly developed Bayesian model (BM) was run over 400,000,000 simulations across 50,000 parameter sets. The parameter sets contributing most to the final outbreak size

Table 2.1: **Average scores for the short-term predictions for the branching process (PF-BP) model and the auto-regression (AR) model.**

Averages are calculated both by the length of the projection at one week, two weeks, one month and two months as well as by date of forecast across all projections. Scores are calculated by taking the natural logarithm of the predicted probability of the observed case count for each projection at each month. Smaller negative scores indicate a better performing model.

Projection Type		Auto-Regression (AR)	Branching Process (PF-BP)
Forecast Date	June	-4.61	-4.57
	July	-4.29	-4.37
	August	-3.79	-3.31
	September	-1.98	-2.33
Projection	one week	-2.62	-2.35
	two weeks	-3.46	-3.23
	one month	-4.06	-4.25
	two months	-4.53	-4.75

distribution generally featured high heterogeneity in secondary case distribution ($k= 0.3$) and an R_0 between 1.05–3.40 with a wide-range of decay rates in transmission (Figure 2.1). Supplementary Table 2.4 shows the predicted conditional distribution for final outbreak size across a range of current outbreak sizes, when the unprecedentedly large West Africa outbreak is excluded. This distribution is conditional only on the currently observed case count, and generally had a median predicted final outbreak size between 3–3.5 fold greater than the currently observed case count, IQR (1.7–1.8 to 8–9 times greater than the currently observed case count). Including the West Africa outbreak, by far the largest to date, increased the range of predicted outcomes between 40–50% (Table 2.5).

The prediction of final outbreak size from the BM model increased from 71 cases (95% CI 35 – 1892 cases) on June 30, to 417 cases (95% CI 134 – >40,000 cases) on September 30th, at the tail end of the outbreak (Table 2.2, Figure 2.3). It performed consistently across the duration of this outbreak, but did best at the beginning, but had final scores from -5.32

– -5.81.

The PF-BP predicted increasing final outbreak sizes over the course of the outbreak, increasing from 51 cases (95% CI 34–>40,000 cases) on June 30, to 131 cases (95% CI 128–183 cases) on September 30th. The initial prediction when the outbreak was growing quickly and was early on had the greatest uncertainty in the continued rate of growth of the outbreak and considered a broad range of possible trajectories. As the outbreak stabilized, the anticipated growth and trajectory stabilized the predictions to a narrower range. Overall, the branching process model performed better by score for the shortest term forecasts of one to two weeks while the auto-regression performed slightly better for longer timepoints. The PF-BP otherwise performed relatively similarly to the AR model across the outbreak, though they both performed best at the end when the trajectory was more stable (Table 2.2).

From the TS model, the final outbreak size increased from 45 cases (95% CI 34–80 cases) on June 30, to 137 cases (95% CI 128–163 cases) on September 30th, just as the outbreak was ending. It performed well at the end of the outbreak, with the smallest negative overall score from any model at any point, but did quite poorly at the beginning, due largely to a narrow prediction with a short tail. The GL model performed similarly, with the best predictions at the end of the outbreak. The predictions for GL increased from 67 cases (95% CI 34–1359 cases) on June 30, to 255 cases (95% CI 130–5119 cases) on September 30th.

Interestingly, though in the past GL has served as a relatively pessimistic agnostic model with a long-tail distribution [60], the inclusion of the recent 2018–20 DRC outbreak (final size >500 cases), has brought our other models’ predictions in closer agreement with the GL model. The BM had the heaviest and longest tails.

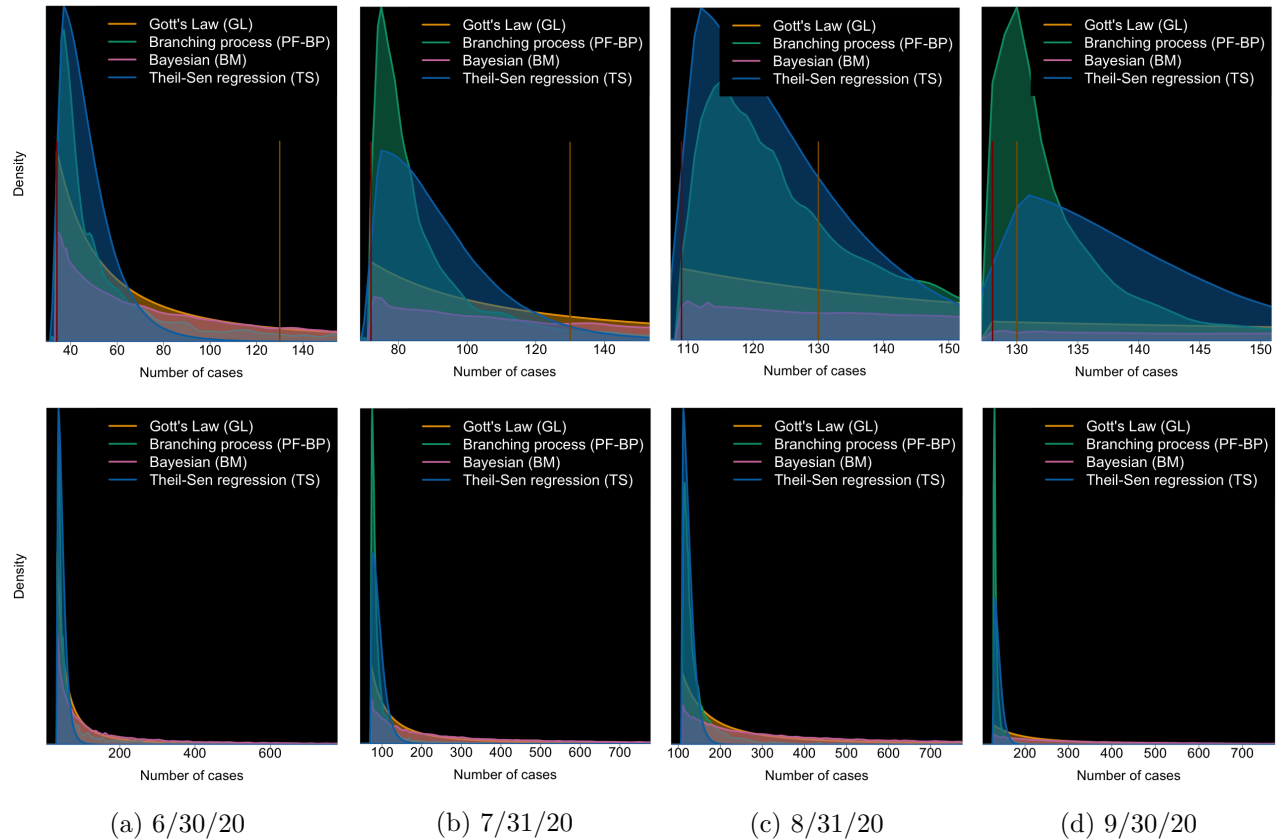


Figure 2.3: **Final outbreak size probability density distributions for all four models at each forecast date** i) Bayesian approach, ii) branching process with particle filtering, iii) Gott’s Law, and iv) regression model. Projections are made based on observed case counts of 34 cases on June 30th, 72 cases on July 31st, 109 cases on August 31st and 128 cases on September 30th. Red bar indicates the current case count, and the gold bar indicates the final case count for the outbreak. The top panels are zoomed in figures on case counts below 150, and the lower panels are zoomed out to show the tails of the distributions.

Table 2.2: **Predicted final outbreak size distributions across all four models**

at the end of each month during the outbreak, including the Bayesian branching process model (PF-BP), the particle filtration branching process model (PF-BP), Theil-Sen Regression model (TS) and Gott’s Law (GL). Performance scores for predicting the final outbreak size of 128 cases from the end of month case counts are included. Scores are calculated by taking the natural logarithm of the predicted probability of the final observed outbreak size for each model at each month. Smaller negative scores indicate a better performing model.

Forecast Date	model	Mean	2.5%	25%	50%	75%	97.5%	score
6/30/20	Bayesian	305	35	55	108	296	1848	-6.2
7/31/20	Bayesian	481	75	122	238	550	2352	-5.81
8/31/20	Bayesian	616	114	184	343	721	2740	-5.67
9/30/20	Bayesian	2596	134	223	417	907	40000*	-5.71
6/30/20	Branching Process	2345	34	39	51	117	40000*	-6.79
7/31/20	Branching Process	356	72	76	82	98	527	-6.4
8/31/20	Branching Process	552	110	118	129	160	812	-4.19
9/30/20	Branching Process	142	128	129	131	136	183	-1.93
6/30/20	Theil-Sen Regression		34	39	45	53	80	-9.97
7/31/20	Theil-Sen Regression		73	80	88	101	138	-5.99
8/31/20	Theil-Sen Regression		110	115	122	132	162	-3.88
9/30/20	Theil-Sen Regression		128	132	137	144	163	-2.93
6/30/20	Gott’s Law	339	34	44	67	135	1359	-6.22
7/31/20	Gott’s Law	665	73	95	143	287	2879	-5.47
8/31/20	Gott’s Law	961	111	144	217	435	4359	-5.05
9/30/20	Gott’s Law	1108	130	170	255	511	5119	-4.89

2.4 Discussion

The recently ended outbreak in the DRC had a final outbreak size of 130 cases, which was always included in the 95% confidence interval for the PF-BP, BM and GL models. The predictions generally increased as the outbreak continued to grow, with both the TS and the PF-BP models showing a narrowing of their predictions as the outbreak continued. The (PF-BP) was overall the most stable consistent performer for final outbreak size during this 2020 outbreak, becoming increasingly confident of outbreak trajectory as time went on. The TS, BM and GL all appeared to be sensitive to the position of the outbreak for their

performance with the BM performing best early in the outbreak, but doing very poorly by the end as it was agnostic to the slowing trajectory of the outbreak. This lack of response to the trajectory may make it less useful to immediate outbreak response but can be important to assess overall readiness of the public health agency early on in the outbreak.

Throughout the outbreak, final outbreak size projections showed fairly high levels of uncertainty, save for the regression, with a sharp disparity between the median and the mode for the outbreak distributions. This uncertainty may simply be a realistic and necessary feature of attempting to predict Ebola outbreaks. The breadth in these predictions can be attributed to two things: i) the over-dispersion in secondary case distribution leading to over-dispersion in historical outbreak sizes which informs the branching process and Bayesian models in particular, a feature of Ebola transmission that is well documented [62–66], and (ii) the inclusion of 2018 DRC outbreak which has affected model fits more generally. These uncertainties are vitally important to highlight as they provide a guidance through the median of the most likely outbreak course. These uncertainties also highlight the need to consider public health readiness in the event that the outbreak should continue on a trajectory more similar to the recent 2018 DRC and West Africa outbreaks.

However, the short-term outbreak predictions continue to be the most immediately applicable for immediate public health responses and are generally considered to be more reliable. Our short-term prediction models consistently predicted that the outbreak would continue to grow, though the models disagreed over whether the course of the outbreak was likely to continue as it had been or slowly come to an end. Dynamic short-term forecasts can provide an important additional feedback tool to local outbreak response teams gaining perspective on changes and successes in outbreak control while also avoiding overconfidence in the ultimate trajectory of the outbreak.

The predictions at all time points for all models were generally larger than previous models and previous outbreaks [58, 60]. The most recent, and second largest outbreak in the DRC from 2018–20 with over 3,000 cases, is only the second outbreak to exceed 500 cases. Previously the West Africa outbreak, the largest to date with nearly 30,000 cases, was often treated as an outlier in previous studies and models, whose inclusion would often more than double the predicted outbreak growth [58, 60]. This more recent outbreak has updated our understanding of the potential for EBOV to lead to large and relatively uncontrolled outbreaks, though the additional inclusion of the West Africa outbreak still pulls the distribution towards larger outbreak sizes. As a consequence, our predictions for the final size of the current outbreak are now heavier tailed and have a higher median. It will be necessary to further evaluate the models’ performances during outbreaks with differing trajectories to assess their responsiveness to changes in growth and trajectory.

Our mathematical models do not consider the unique circumstances that may have likely contributed to the two historically large outbreaks with the recently ended outbreak in the eastern DRC, and previously in the 2014–2016 West Africa outbreak, including war, unrest, and lack of experience with Ebola; circumstances which were not replicated in the recent outbreak and may have a heavy influence upon any predictions presented in this paper. Initial outlooks for the outbreak were optimistic, but subsequent spread to remote regions of the country tempered expectations for the early and swift control of the outbreak by timely treatment, contact tracing and vaccination [71]. Despite ring vaccination having been in use since the end of the West Africa basis, initially as compassionate use, recent outbreaks have continued to be larger than earlier historical outbreaks due to a variety of ongoing challenges. It is likely that the value of prior outbreaks in making predictions becomes less valuable as more knowledge and understanding of the current outbreak becomes apparent,

and the ability of the branching process to respond to the current outbreak's trajectory likely aided its strong performance later in the outbreak.

Conversely, as we observe Ebola outbreaks trending towards larger final outbreak sizes, we may be facing a reality that earlier outbreaks which reflect living conditions from nearly half a century earlier may not be as relevant to predicting outbreak trajectories of today. There have been substantial changes in interconnectedness, faster transport, and larger social networks that make it more likely for outbreaks to escape their local region and make the jump into larger populations which make the outbreaks more difficult to control.

Our models have several limitations. Vaccination is not explicitly included in models, and the recent outbreak was only the third EBOV outbreak to use ring vaccination [50]. Vaccination rollout was slower than anticipated in this outbreak, and the continued reporting of deaths within the community setting, indicated that there were containment failures despite the ring vaccination strategy [50]. Additionally, the limited sample size of past outbreaks greater in size than ten cases, does statistically limit the precision of these analyses, though this also reinforces the necessity of studies such as this one hoping to better understand accurately model the transmission dynamics.

The particle-filter branching process model and the Theil-Sen regression models are missing two historical outbreaks greater than ten cases in size due to data availability and lack of access to their epidemic curves. Since the simple regression projections and Bayesian predictions are based entirely on past outbreaks of Ebola virus disease (measured and reported in different ways), they cannot account for improvements in control measures and vaccination, nor demographic or geographic differences in the way that a mechanistic model does. The particle-filter branching process model does partly adjust for these changes by including a trajectory in the model of the outbreak across two timepoints. Our models also assume

that the size of prior outbreaks can be modeled by a single distribution. This assumption could be problematic if some of these underlying differences in the outbreak circumstances have an outsized influence upon the trajectory of an outbreak, or if the trend towards larger outbreaks in recent years reflects an intrinsic change in the patterns of EBOV transmission that are not yet fully understood.

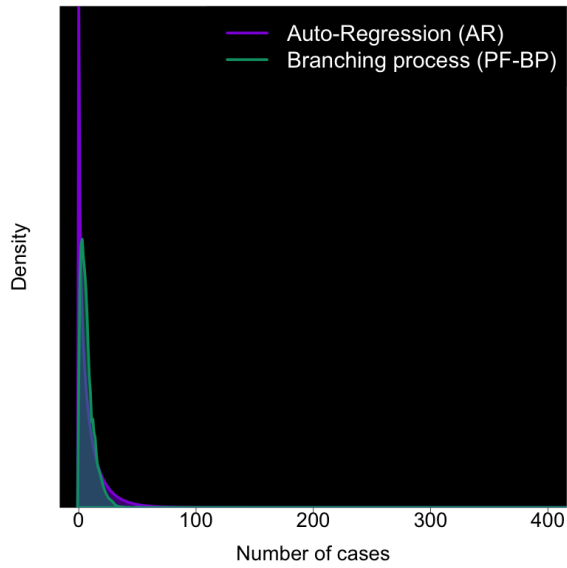
The greatest strength in our approach is the use of multiple methods to estimate the outbreak size. A range of different statistical and methodological approaches is likely to yield a greater understanding of the role of stochasticity plays in determining the final outbreak size as well as providing more robust assessment of an ongoing outbreak's short-term trajectory. This was the first outbreak to occur following the second outbreak to exceed 500 cases, which greatly shifted the modeled case distributions as the West Africa outbreak is no longer a clear outlier. This makes this 2020 outbreak an important first test case of these updated model parameterizations, and underscores the need to continue to evaluate model performance in future outbreaks with diverse trajectories. EBOV remains a serious outbreak disease, despite recent development of several vaccines, implementation of ring vaccination in outbreak response and improvement in treatment options, with the identification of two top performing treatments coming from a clinical trial undertaken during the 2018–20 Beni outbreak [50]. Thus, prediction of the final outbreak size remains an important element of the public health response.

2.5 Supplementary Materials

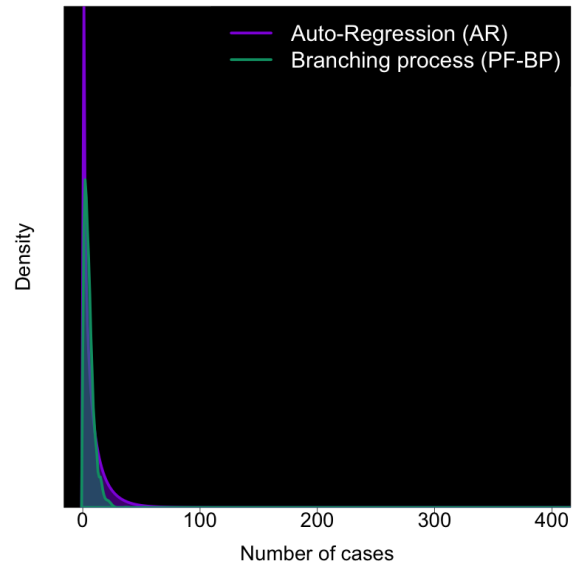
Table 2.3: **Summary of all reported historical Ebola outbreaks**

of greater than ten cases in size, their cases counts, and their inclusion in the various prediction models. TS = Theil-Sen Regression, AR = Auto-regression, BM = Bayesian Model, PF-BP = Particle-Filter Branching Process Model and GL = Gott’s Law.

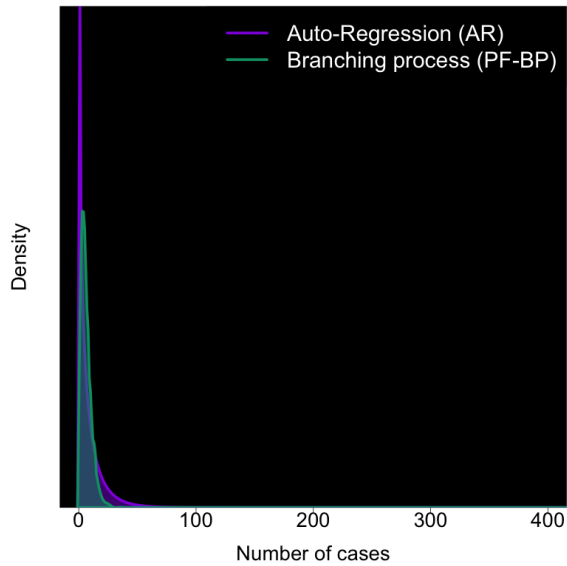
Time Period	Country	Reported Cases	Time Series Cases	TS	AR	BM	PF-BP	GL
1976	DRC*	318	262	Yes	Yes	Yes	No	No
1976	Sudan	284	284	Yes	Yes	Yes	No	No
1979	Sudan	34	34	Yes	Yes	Yes	Yes	No
1994–95	Gabon	52	49	Yes	Yes	Yes	No	No
1995	DRC	315	317	Yes	Yes	Yes	Yes	No
1996	Gabon	37	29	Yes	Yes	Yes	Yes	No
1996–97	Gabon	60	–	No	No	No	No	No
2000–01	Uganda	425	436	Yes	Yes	Yes	No	No
2001–02	Gabon, Rep. of the Congo	124	124	Yes	Yes	Yes	Yes	No
2002–03	Rep. of the Congo	143	–	No	No	No	No	No
2003	Rep. of the Congo	35	35	Yes	Yes	Yes	Yes	No
2004	Sudan	17	17	Yes	Yes	Yes	Yes	No
2005	DRC	12	12	Yes	Yes	Yes	Yes	No
2007	DRC	264	264	Yes	Yes	Yes	Yes	No
2007–08	Uganda	131	127	Yes	Yes	Yes	Yes	No
2008–09	DRC	32	32	Yes	Yes	Yes	Yes	No
2012	Uganda	24	24	Yes	Yes	Yes	Yes	No
2012	DRC	52	52	Yes	Yes	Yes	Yes	No
2014	DRC	66	62	Yes	Yes	Yes	Yes	No
2014	Nigeria-offshoot of West Afr. outbreak Guinea,	20	–	No	No	No	No	No
2014–016	Liberia, Sierra Leone	28,616	21,422	No*	No*	No*	No	No
2018	DRC	54	54	Yes	Yes	Yes	Yes	No
2018–20	DRC	3463	3463	Yes	Yes	Yes	Yes	No



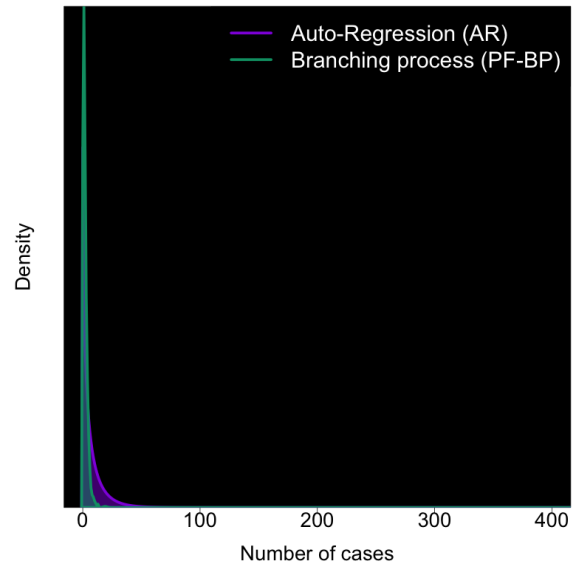
(a) 6/30/20



(b) 7/31/20

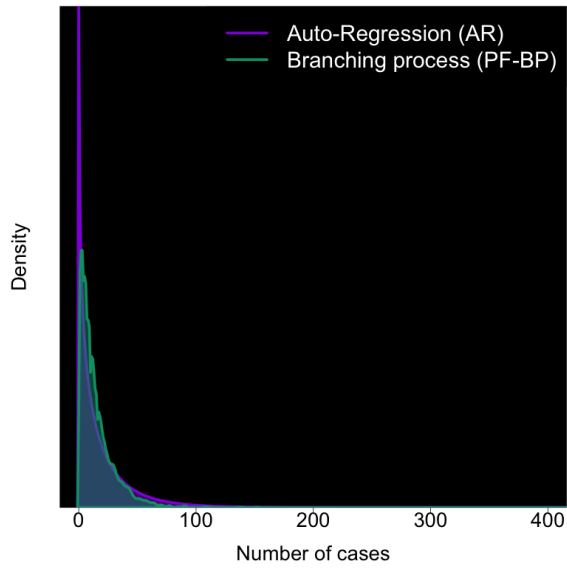


(c) 8/31/20

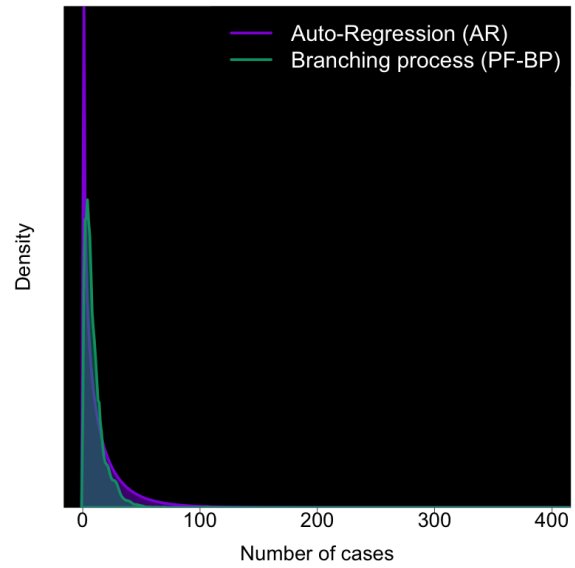


(d) 9/30/20

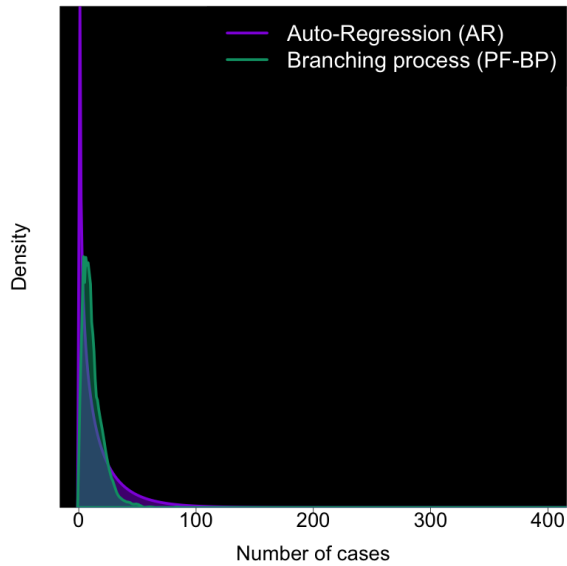
Figure 2.4: **One week projections of outbreak size probability density distributions for the BP-PF and AR models at each forecast date** i) branching process particle-filter (BP-PF) and ii) auto-regression (AR). Projections are made based on observed case counts of 34 cases on June 30th, 72 cases on July 31st, 109 cases on August 31st and 128 cases on September 30th.



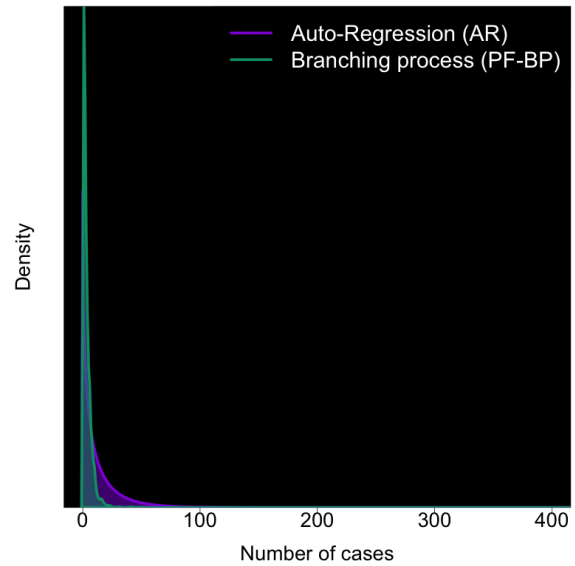
(a) 6/30/20



(b) 7/31/20

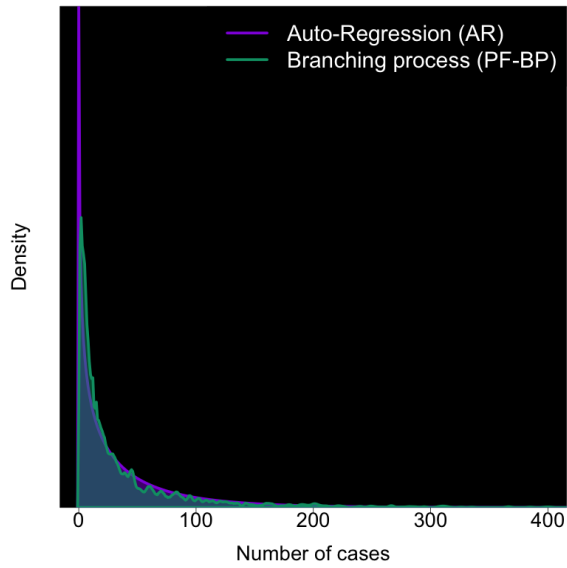


(c) 8/31/20

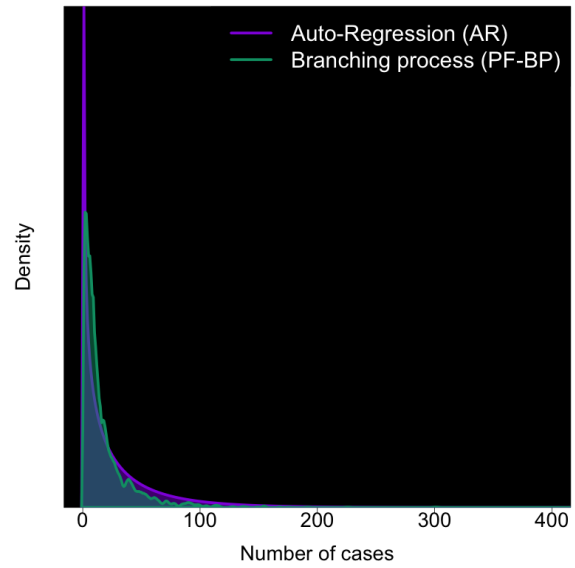


(d) 9/30/20

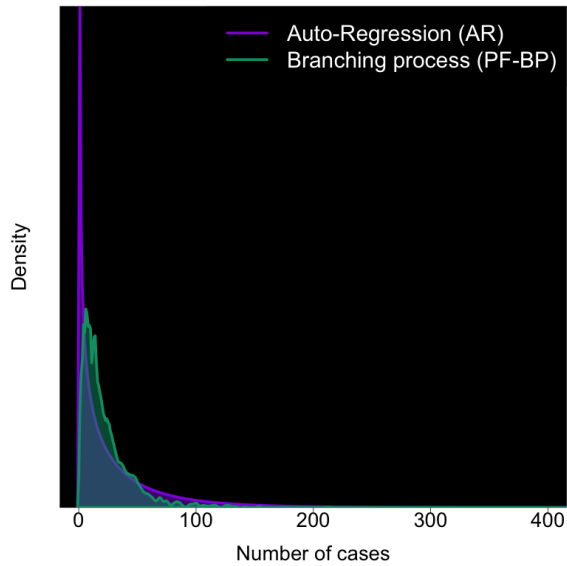
Figure 2.5: **Two week projections of outbreak size probability density distributions for the BP-PF and AR models at each forecast date** i) branching process particle-filter (BP-PF) and ii) auto-regression (AR). Projections are made based on observed case counts of 34 cases on June 30th, 72 cases on July 31st, 109 cases on August 31st and 128 cases on September 30th.



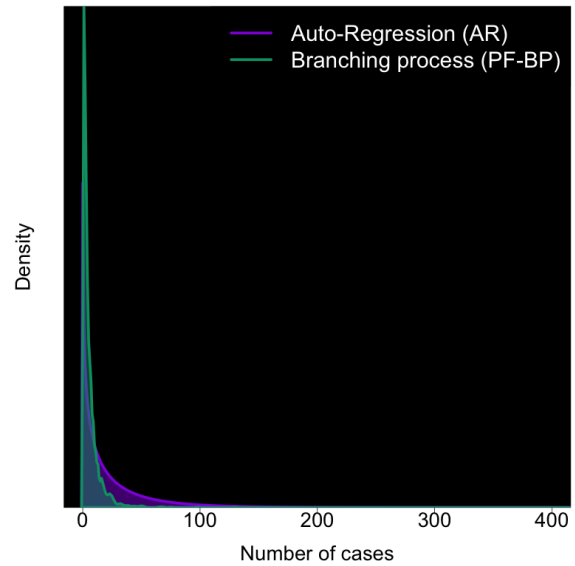
(a) 6/30/20



(b) 7/31/20

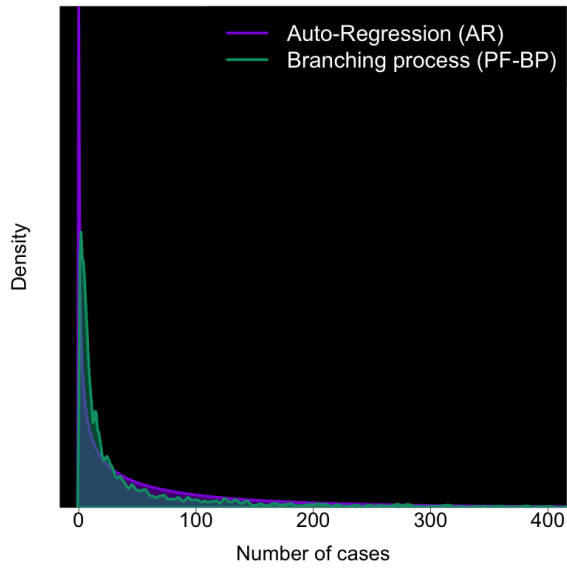


(c) 8/31/20

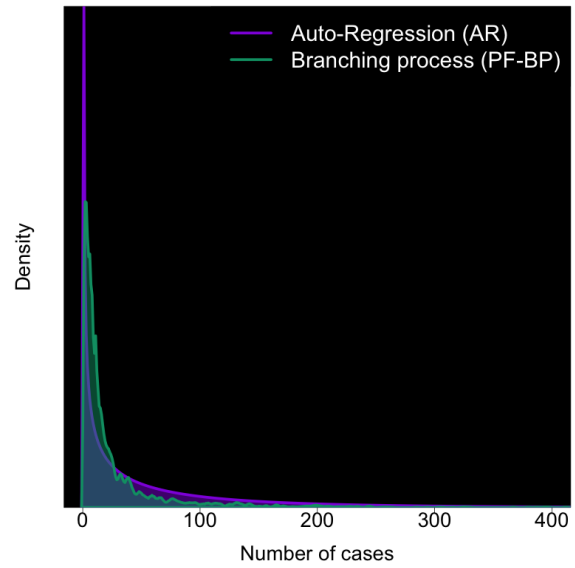


(d) 9/30/20

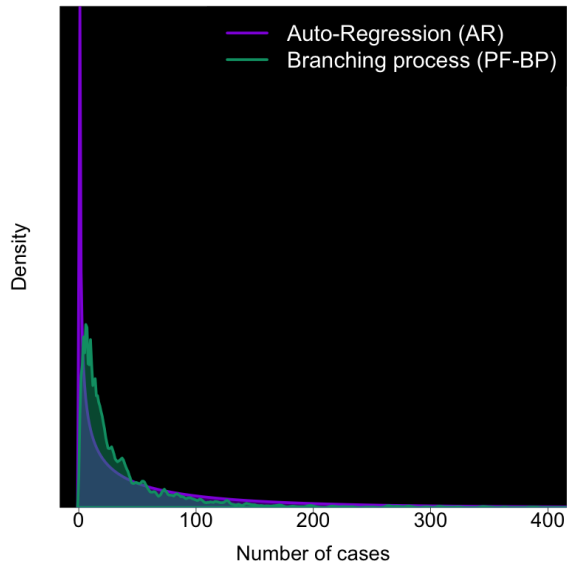
Figure 2.6: **One month projections of outbreak size probability density distributions for the BP-PF and AR models at each forecast date i) branching process particle-filter (BP-PF) and ii) auto-regression (AR).** Projections are made based on observed case counts of 34 cases on June 30th, 72 cases on July 31st, 109 cases on August 31st and 128 cases on September 30th.



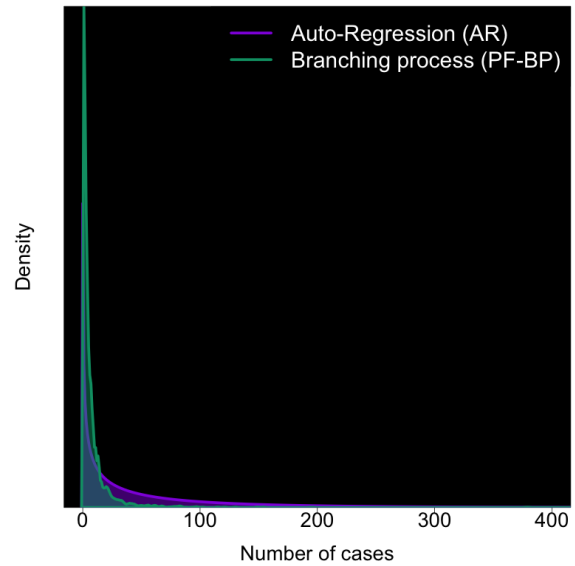
(a) 6/30/20



(b) 7/31/20



(c) 8/31/20



(d) 9/30/20

Figure 2.7: **Two month projections of new case probability density distributions for the BP-PF and AR models at each forecast date i) branching process particle-filter (BP-PF) and ii) auto-regression (AR).** Projections are made based on observed case counts of 34 cases on June 30th, 72 cases on July 31st, 109 cases on August 31st and 128 cases on September 30th.

Table 2.4: **Bayesian predictions of final outbreaks sizes**

Bayesian model (BM) predictions of final outbreaks sizes conditional on currently observed cumulative case counts excluding the historical West Africa outbreak. The final size distribution of R , τ , and k triples were fit using the Gaussian kernel density estimator which was used to estimate the log-likelihood of the previously observed outbreaks' final sizes to create a conditional ensemble distribution for the final size of the current outbreak

Current Case Count	Mean	2.5%	25%	50%	75%	97.5%
20	277.0	20	33	66	197	1554
30	303.0	31	50	104	297	1779
40	321.0	41	65	120	314	1897
50	370.0	51	82	158	396	2013
60	438.0	62	101	200	487	2249
70	494.0	73	123	246	571	2385
80	491.0	84	136	258	572	2278
90	490.0	94	148	273	578	2172
100	498.0	104	162	290	590	2133
110	560.0	115	182	328	668	2381
120	565.0	125	196	343	678	2313
140	659.0	147	236	414	800	2622
160	725.0	168	270	469	884	2804

Table 2.5: **Sensitivity analysis for Bayesian predictions of final outbreaks sizes including the 2014-15 West Africa outbreak**

Sensitivity analysis for Bayesian model (BM) predictions of final outbreaks sizes including the 2014-15 West Africa outbreak. Final size predictions are conditional on currently observed cumulative case counts. The final size distribution of R , au , and k triples were fit using the Gaussian kernel density estimator which was used to estimate the log-likelihood of the previously observed outbreaks' final sizes to create a conditional final size distribution.

Current Case Count	West Africa	Mean	2.5%	25%	50%	75%	97.5%
20	no	277	20	33	66	197	1554
30	no	303	31	50	104	297	1779
40	no	321	41	65	120	314	1897
50	no	370	51	82	158	396	2013
60	no	438	62	101	200	487	2249
70	no	494	73	123	246	571	2385
80	no	491	84	136	258	572	2278
90	no	490	94	148	273	578	2172
100	no	498	104	162	290	590	2133
110	no	560	115	182	328	668	2381
120	no	565	125	196	343	678	2313
140	no	659	147	236	414	800	2622
160	no	725	168	270	469	884	2804
20	yes	2006	20	36	98	657	18842
30	yes	1548	31	60	173	1215	12061
40	yes	1791	41	82	244	1596	12240
50	yes	1736	52	98	288	1549	12631
60	yes	1890	63	123	364	1889	12787
70	yes	1992	75	144	449	2087	13014
80	yes	2058	85	172	526	2218	13040
90	yes	2130	95	200	596	2335	13120
100	yes	2168	106	229	663	2456	12898
110	yes	2250	119	287	836	2649	12431
120	yes	2247	130	297	831	2674	12688
140	yes	2596	154	401	1142	3224	13578
160	yes	3072	182	540	1532	3855	15080

Table 2.6: **Predicted short-term new case distributions**

at one week, two weeks, one month and two months out for both the branching process (PF-BP) model and the auto-regression (AR) model.

Model	Forecast Date	Forecast	Mean	2.5%	25%	50%	75%	97.5%
Auto-Regression	6/30/20	1 week	9	0	1	4	12	42
Auto-Regression	6/30/20	2 week	19	0	2	9	25	85
Auto-Regression	6/30/20	1 month	37	0	4	18	49	177
Auto-Regression	6/30/20	2 month	75	0	6	32	97	391
Auto-Regression	7/31/20	1 week	9	1	2	5	11	38
Auto-Regression	7/31/20	2 week	16	1	3	8	21	70
Auto-Regression	7/31/20	1 month	29	1	4	15	38	135
Auto-Regression	7/31/20	2 month	64	1	6	27	82	332
Auto-Regression	8/31/20	1 week	9	1	2	5	12	40
Auto-Regression	8/31/20	2 week	17	1	3	9	22	75
Auto-Regression	8/31/20	1 month	31	1	4	16	41	146
Auto-Regression	8/31/20	2 month	68	1	6	29	87	351
Auto-Regression	9/30/20	1 week	7	0	0	3	9	33
Auto-Regression	9/30/20	2 week	13	0	1	6	17	60
Auto-Regression	9/30/20	1 month	23	0	2	11	30	112
Auto-Regression	9/30/20	2 month	55	0	4	22	70	291
Branching Process	6/30/20	1 week	7	0	2	5	10	23
Branching Process	6/30/20	2 week	15	0	4	10	20	54
Branching Process	6/30/20	1 month	39	0	5	16	43	203
Branching Process	6/30/20	2 month	202	0	5	17	71	1437
Branching Process	7/31/20	1 week	6	0	2	4	7	17
Branching Process	7/31/20	2 week	10	0	3	7	13	33
Branching Process	7/31/20	1 month	18	0	4	10	21	79
Branching Process	7/31/20	2 month	29	0	4	10	26	184
Branching Process	8/31/20	1 week	6	0	3	5	8	17
Branching Process	8/31/20	2 week	12	0	5	9	16	34
Branching Process	8/31/20	1 month	22	1	8	15	29	81
Branching Process	8/31/20	2 month	38	1	9	19	43	180
Branching Process	9/30/20	1 week	2	0	0	1	3	8
Branching Process	9/30/20	2 week	4	0	1	2	5	14
Branching Process	9/30/20	1 month	6	0	1	3	7	26
Branching Process	9/30/20	2 month	9	0	1	3	8	47

Chapter 3

Modeling the Impact of Trachoma

MDA on GU Chlamydia

3.1 Introduction

Despite trachoma being targeted for global elimination through the use of annual and semi-annual Mass Drug Administration (MDA) of azithromycin, the impacts that these interventions will have upon the burden of its sister disease Chlamydia are very poorly understood and documented. Communities that are especially hard hit by trachoma are almost exclusively poor communities with poor access to sanitation, screening and antibiotics to treat the infection [72]. These are the same conditions that allow STDs, such as Chlamydia, to maintain a high, unbroken chain of transmission [73–75].

Genitourinary (GU) Chlamydia is the most common bacterial sexually transmitted infection world-wide [76,77]. The global prevalence of GU infection by *Chlamydia trachomatis*, or Chlamydia, in 2012 was 4.2% (95% CI: 3.7-4.7%) among women ages 15-49 years and 2.7% (95% CI: 2.0-3.6%) among men ages 15-49. However, the burden of Chlamydia is especially

high in resource-limited settings where the overall prevalence has been reported to be as high as 29% (CI: 7-31%) [75, 77–80]. GU Chlamydia can lead to infertility, adverse pregnancy outcomes for both the mother and the neonate, and an increased risk of HIV transmission [76, 77, 81]. However, infections with *C. trachomatis* can often be asymptomatic, and some high-income countries perform routine screening of high-risk groups to detect and treat *C. trachomatis* infection. However, in low-resource settings, such programs are not possible due to the expense and infrastructure necessary. [74]

In response to the continued challenges represented by STD's such as Chlamydia, and findings that the benefits of screening programs were overestimated [82], some have called for the use of MDA or targeted administration of treatment with azithromycin in extremely high prevalence areas, and several studies have used Periodic Presumptive Treatment (PPT) of sex-workers with azithromycin in similar settings [83–88]. Though there is evidence that the near-simultaneous treatment from MDA greatly reduces the probability of re-infection from partners [89] there is very little evidence to help understand what such an approach would mean for the long-term sustainability of decreased prevalence, as well as concerns as to the long-term viability of treatment efficacy, and concerns of developing anti-biotic resistance. Though both diseases are caused by the same pathogen, only trachoma is currently targeted for eradication through MDA and other initiatives. At the same time, recent evidence suggests that trachoma MDA with the antibiotic azithromycin reduces the prevalence of genital Chlamydia. [74] Currently, there are very few studies that have looked at the impact of trachoma MDA upon the burden of genital Chlamydia, and no efforts have been made to model or predict the reduction in genital Chlamydia burden in these populations. The trachoma MDA dosing with 20 mg/kg of azithromycin [90] is clinically consistent with dosing regularly used in the treatment of GU Chlamydia [76], and thus it is highly likely that

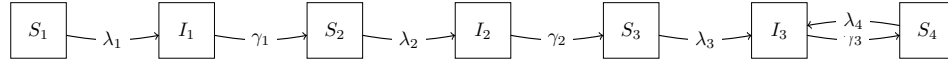
trachoma MDA is treating undiagnosed or unrecognized community GU Chlamydia. This may also mitigate concerns regarding the rise of resistant strains that might arise from partial or insufficient dosing. Here we propose a mathematical disease transmission model to investigate the impact of trachoma MDA upon chlamydia prevalence in these settings.

3.2 Methods

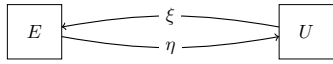
3.2.1 Mathematical Model

We modeled Chlamydia transmission using a standard SIS hidden markov compartmental model adapted for the natural history of a Chlamydia infection [91]. We generated an extended compartmental SIS model (Figure 3.1), using a generalized Cartesian product operation [92] to combine the features of the model including: the natural history of GU Chlamydia, risk structure, gender, MDA exposure and population maturation and aging. We modelled the duration of illness under the assumption of a constant risk of recovery using a standard negative binomial distribution.

We have included in this model four susceptible (S_1, S_2, S_3, S_4) and three infectious classes (I_1, I_2, I_3) (Figure 3.1) to allow for the natural cycle of recovery and re-infection over the course of an individual's sexually active life, and the process of slowly developing immunity after each subsequent infection [93]. During the course of natural infection, as individuals develop immunity, individuals move from I_1 to I_2 to I_3 , which each have distinct infection responses and dynamics. Due to the short time scale of the exposed, but non-infectious, class relative to the disease course, we did not explicitly model this transition state [93]. In the presence of MDA, an individual may either be allowed to move from the infectious class (I_i) to the next susceptible class (S_{i+1}) or back down to the previous



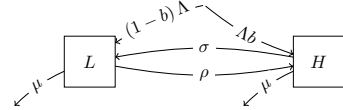
(a) Natural History of Chlamydia



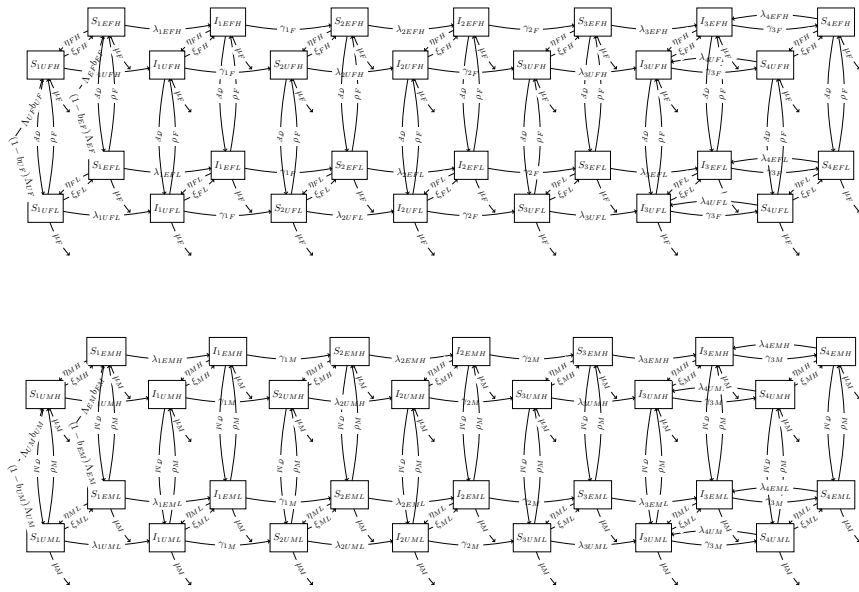
(b) Exposure to MDA Treatment



(c) Gender



(d) Risk Structure



(e) Full Model

Figure 3.1: Levels of the Compartmental Model and the Full Extended Compartmental Model

a. Natural history of Chlamydia, with slowly developing immunity. γ = mean duration of infection, S = Susceptible, I = Infected, $\lambda = \text{force of infection} = \beta \cdot I/N = \text{Transmission Probability Per Contact} \times \text{Infected} / \text{Total Population}$. b. Exposure to the MDA treatment. c. Gender. d. Risk structure for behavior. e. Full extended Cartesian model. (Tables 3.1 and 3.2)

susceptible class (S_i) if the treatment has been given prior to individual mounting an effective immune response to develop partial immunity which takes many months to occur [91]. MDA was modelled as an annual impulsively forced treatment with varying coverage and efficacy.

We added gender and risk groups, to address the sexual nature of Chlamydia transmission

and also allowed for differential susceptibility and transmissibility to GU Chlamydia by gender. For risk structure, we included high risk core groups who had relatively high levels of sexual activity, and low-risk groups with lower rates of sexual contacts made per year [91,93]. Given the nature of the countries and areas targeted by trachoma elimination, these high risk groups were modelled as female sex workers (FSWs) from the regions and those who have contact with these FSW's [94–97].

3.2.2 Sensitivity Analysis

A sensitivity analyses was conducted to determine the sensitivity of the primary outcome (relative reduction in prevalence) model parameter selections to investigate the robustness of model findings. All parameters were sampled across their respective distributions (Table 3.1, Table 3.2) using Latin Hypercube (LH) sampling [98]. Due to the possible range for parameters and the independent nature of their sampling, the resulting simulations were screened to ensure that high risk groups had contact rates at least twice that of the low risk groups. The remaining simulations were then run for 5 years of MDA and the sensitivity of the % reduction in prevalence from baseline after 3 years and 5 years of MDA were assessed using a Partial Rank Correlation Coefficient (PRCC) [98].

3.2.3 Running the Model

We tested the efficacy of trachoma MDA treatment on reducing GU Chlamydia, varying coverage levels and population baseline prevalence using the base parameters shown in Tables 3.1 and 3.2. Our primary outcome of interest was the reduction in prevalence relative to baseline one year after MDA treatment (immediately prior to subsequent MDA treatment), allowing for the expected bounce back in population prevalence in the time following

Table 3.1: Table of Model Parameters and Ranges For Population Dynamics and Sexual Behavior With Sourced References

Model parameters for population dynamics and sexual behavior by compartments are shown along with the literature references. The values used for primary model calculations are shown, as well as the ranges used for the sensitivity analysis and prediction uncertainty.

Parameter (units)	Group	Symbol	Value	Range
Population Dynamics				
Sexual maturation rate (model entry)	female	δ_F	100 ppl./year	
	male	δ_M	100 ppl./year	
Rate of annual aging out of model	female	μ_F	0.05	0.036 - 0.067
	Male	μ_M	0.05	0.036 - 0.067
Sexual Behavior				
Proportion of "maturations" recruited to high risk group	exposed female	b_{EF}	0.05	0.01 - 0.1 [72,99]
	exposed male	b_{EM}	0.1	0.01 - 0.3
	unexposed female	b_{UF}	0.09	0.01 - 0.5 [72,99]
	unexposed male	b_{UM}	0.18	0.01 - 0.5
Proportion high-risk contacts	high risk female	a_{HH}	0.66	0.41 - 0.95 [99,100]
	low risk female	a_{LH}	0.15	0 - 0.41 [101-103]
Average # new partne- ships per year	high risk female	c_{FH}	144	12 - 420 [99, 100, 104]
	low risk female	c_{FL}	0.24	0.1 - 2.4 [101-103]
Rate of movement low to high risk per year	female	ρ_F	0.012	0.001 - 0.18 [99]
	male	ρ_M	0.012	0.001 - 0.18 [101-103]
Rate of movement high to low risk per year	female	σ_F	0.24	0.1 - 0.5 [99]
	males	σ_M	0.12	0.06 - 0.6 [101-103]

Table 3.2: **Table of Model Parameters and Ranges for MDA Treatment and Chlamydia Natural History With Sourced References**

Model parameters for MDA treatment and the natural history of chlamydia by compartments are shown along with the literature references. The values used for primary model calculations are shown, as well as the ranges used for the sensitivity analysis and prediction uncertainty.

Parameter (units)	Group	Symbol	Value	Range
MDA treatment				
Efficacy	-	τ	0.98	0.9 - 1 [76]
Serial non-compliance	high risk female	κ_{FH}	0.5	0 - 1
	low risk female	κ_{FL}	0.5	0 - 1
	high risk male	κ_{MH}	0.5	0 - 1
	low risk male	κ_{ML}	0.5	0 - 1
Proportion of maturation exposed to MDA	female	δ'_{EF}	0.9	0.5 - 1 [105]
	male	δ'_{EM}	0.9	0.5 - 1
Chlamydia Natural History				
Transmission probability per contact during first infection	female to male	β_{FM}	0.17	0.04 - 0.7 [91,101-103]
	male to female	β_{MF}	0.2	0.04 - 0.7
Reduction in transmission prob. in subsequent infections	female to male	β'_{FM}	0.1	0 - 0.5 [91]
	male to female	β'_{MF}	0.1	0 - 0.5
Mean duration of infection (months)	female	γ_F	10	2 - 50 [91,101,103,106,107]
	male	γ_M	5	2 - 50
Reduction in disease duration in subsequent infections	female	γ'_F	0.1	0 - 0.2 [91]
	male	γ'_M	0.1	0 - 0.2
Susceptibility	female	ζ_F	0.9	0.5 - 0.999 [91]
	males	ζ_M	0.9	0.5 - 0.999

treatment. Under this assumption, with enough MDA administrations within the population, the burden of GU Chlamydia would be steadily reduced and eventually drop below the level of sustained transmission within the population. Due to previous studies indicating the efficiency of targeting and treating high-risk individuals [108, 109], we investigated the impact of targeting the azithromycin treatment only to the high risk groups under the same conditions.

One often raised concern with MDA strategies is serial non-compliance with annual treatment [110], that the same individuals are non-compliant year after year, creating a reservoir of disease for continued transmission. To investigate the impact of serial non-compliance with MDA treatment, we varied the rate of annual movement from the exposed to the non-exposed classes from 0 to 1 across a range of population baseline prevalences (the base model is run at $\kappa=0.5$).

Lastly, in order to investigate the potential uncertainty in our predictions, we ran the model 10,000 times across a range of parameter sets (Tables 3.1 and 3.2) randomly sampled from a distribution reflecting our uncertainty in each parameter at each of four levels of MDA coverage at 60%, 70%, 80% and 90%.

Table 3.3: **Sensitivity Analysis**

The Partial Ranked Correlation Coefficients (PRCC) were calculated for all parameters with the percent reduction in prevalence relative to baseline and reported when $PRCC \geq |0.10|$ after 3 years and 5 years . Parameters were varied across the ranges reported in Tables 3.1 and 3.2, and sampled across the range using Latin Hypercube sampling. The model was then run across all parameter sets, and the sensitivity analysis was performed on the resulting output.

Parameter	3 Years		5 Years	
	PRCC	std. error	PRCC	std. error
Recruitment to high-risk group for MDA exposed F (δ'_{EF})	0.10	0.03	0.09	0.03
Proportion high-risk contacts among high-risk F (a_{HH})	0.10	0.03	0.10	0.03
Transmission per contact M to F (β_{MF})	-0.30	0.02	-0.30	0.02
Reduced trans. prob. in subsequent infect. M to F (β'_{MF})	-0.13	0.04	-0.13	0.03
Contact rate high-risk females (c_{FH})	-0.52	0.02	-0.52	0.02
Mean duration of infection - F (γ_F)	-0.65	0.03	-0.64	0.01
Mean duration of infection - M (γ_M)	-0.18	0.03	-0.18	0.01
Mean disease duration in subsequent infections - F (γ'_F)	-0.36	0.02	-0.35	0.02
Movement low to high risk group - F (ρ_F)	-0.25	0.02	-0.26	0.03
Movement high to low risk group - F (σ_F)	0.32	0.02	0.32	0.03
Efficacy (τ)	0.09	0.03	0.09	0.03

3.3 Results

Overall, we found that the MDA treatment was effective at reducing prevalence levels relative to baseline in the overall population. Our model showed that three years of MDA at current levels reduced the prevalence of GU Chlamydia in all populations by at least 10.5% relative to baseline (Figure 3.2). Though the prevalence of GU Chlamydia partially bounced back to pre-treatment levels between annual administrations as untreated or ineffectively treated individuals continued transmitting GU Chlamydia to newly susceptible or naive hosts, the time between MDA was not enough to allow transmission to revert to pre-MDA levels.

Looking at percent (%) reduction in prevalence relative to baseline one year after MDA

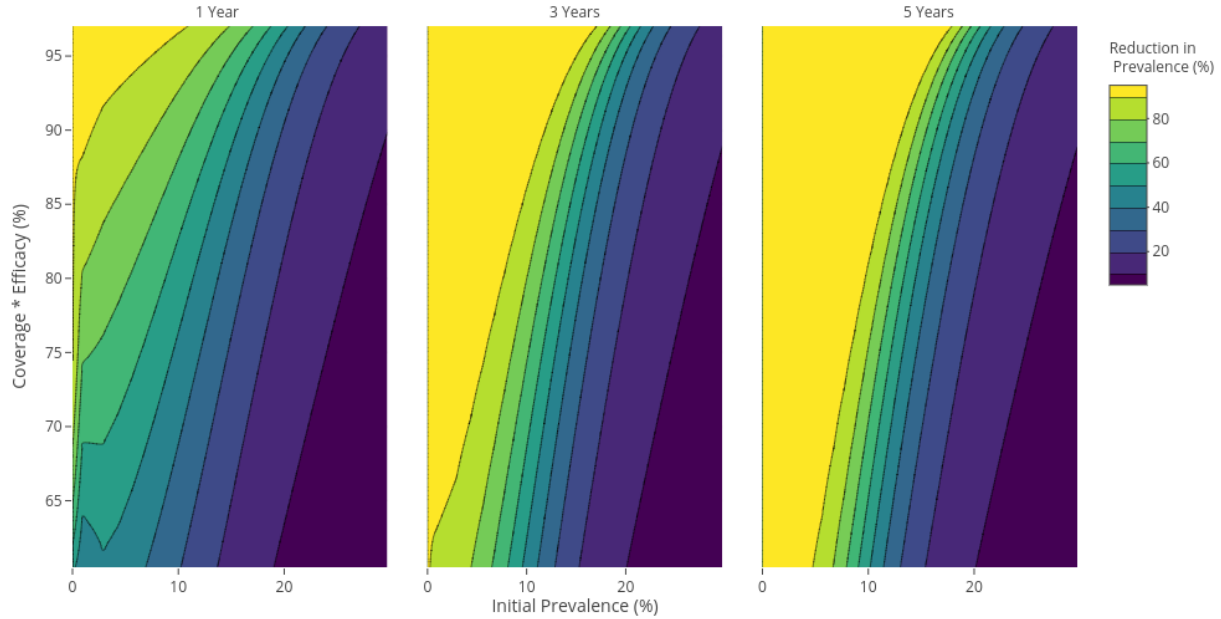


Figure 3.2: **Percent reduction in prevalence after 1, 3 and 5 years of azithromycin MDA.**

The percent (%) reduction in prevalence relative to baseline was calculated as the reduction in prevalence one year after the last annual MDA treatment compared to the baseline equilibrium for each simulation. Results are reported for the same initial conditions following 1, 3 and 5 years of annual azithromycin treatment.

treatment (immediately prior to subsequent MDA treatment) relative to the baseline equilibrium prevalence, it appeared that the burden of GU Chlamydia would be steadily reduced and eventually drop below the level of sustained transmission under all model parameterizations. However, even at a 90% coverage X efficacy (CE) and a baseline prevalence above 15% prevalence, this target would take more than 10 years of MDA administration to achieve (Figure 3.2). With $CE \geq 0.80$, the time between MDA treatments was insufficient to sustain transmission, allowing for GU Chlamydia burden to be suppressed below 1 in 10,000 after 5 rounds of MDA for baseline prevalences less than 2.9%, and with $CE \geq 90\%$, prevalence was reduced below 1 in 10,000 after 5 rounds for baseline prevalences less than 6.5%. Three rounds of MDA at 90% CE was sufficient to reduce prevalence below 1 in 10,000 only for baseline prevalences below 2.9%.

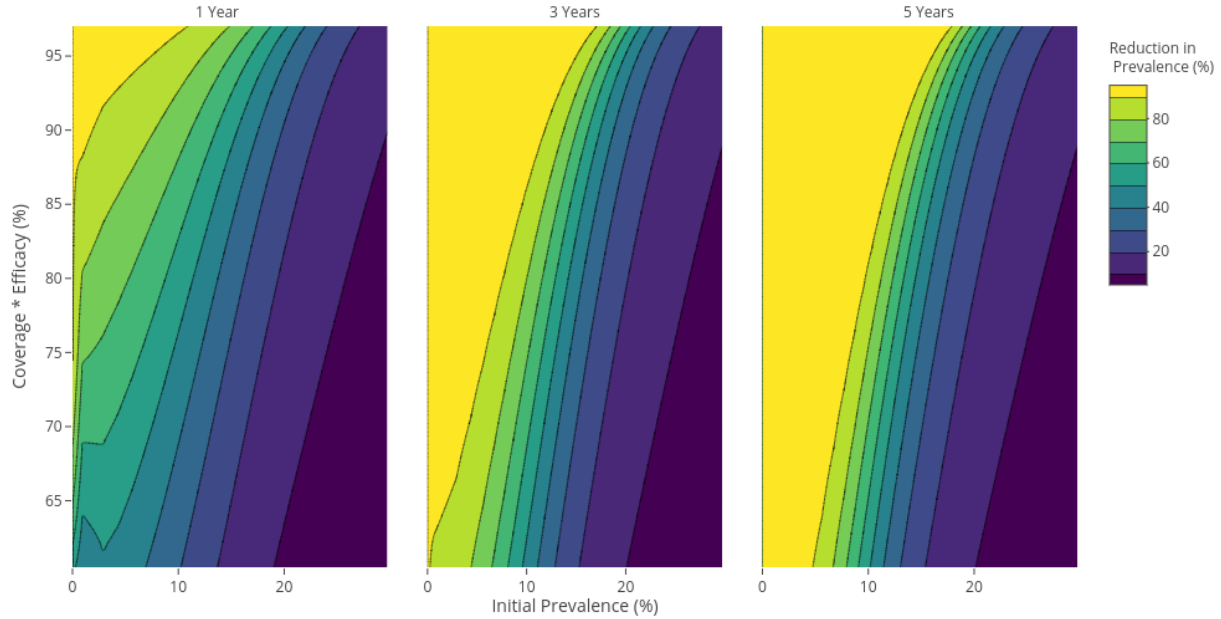


Figure 3.3: Percent reduction in prevalence after 1, 3 and 5 annual rounds of azithromycin targeting only the high-risk group.

The percent (%) reduction in prevalence was calculated as the reduction in total population prevalence one year after the last annual MDA treatment compared to the baseline equilibrium for each simulation. Results are reported for the same initial conditions following 1, 3 and 5 years of annual azithromycin treatment targeting only the high-risk groups.

We also looked at prevalence reductions of 50% and 90% as outcomes relevant to public health interests. At 80% CE, a 50% reduction in prevalence relative to baseline was achieved for starting prevalences of up to 13.7% after 3 rounds of MDA, and starting prevalences up to 14.3% after 5 rounds of MDA. Likewise at 80% CE, a 90% reduction in prevalence relative to baseline was achieved for starting prevalences of up to 7.7% after 3 rounds of MDA, and baseline prevalences up to 9.5%.

Again considering percent reduction in prevalence relative to baseline, we looked at the sensitivity of the model to various model parameters (Table 3.3). Overall, the model appeared more sensitive to changes in female defined parameters than the males. This is likely because the female high-risk group, modelled as they were after female sex workers, was assumed to have higher contact rates than their male high-risk counterparts, and in this

model represent the core risk group facilitating transmission within the population. The model is quite sensitive to changes in group-specific risk structure and duration of infection. Decreasing the duration of infection, lowering the contact rate, increasing the movement from high to low risk groups and decreasing the movement from low to high risk groups (effectively reducing the high-risk population size), and decreasing transmission probability from males to females were all strongly associated with increasing the reduction in prevalence relative to baseline following MDA transmission. Efficacy was relatively weakly correlated with the efficacy of MDA administration (PRCC = 0.09). Though somewhat surprising, this demonstrates that population dynamics are critical in affecting the efficacy of the MDA intervention. More would need to be known about the population being targeted to determine if they are a good candidate for such a program.

Given a reticence to implementation of a population-level intervention and previous studies on targeting treatments for GU Chlamydia [108,109], we looked at the impact of targeting the intervention only to high-risk individuals (Figure 3.3). When targeting only high-risk individuals, absolute prevalence was reduced below 1 in 10,000 when CE \geq 0.8 for baseline prevalences below 0.73% after 5 rounds of MDA, down from achieving the same target from a baseline prevalence of 2.9% using population-wide MDA. This target was not achieved in any simulation after only three rounds. Other targets were similarly affected as at 80% CE, a 50% reduction in prevalence relative to baseline was achieved for starting prevalences up to 9.4% after 3 rounds of MDA, and baseline prevalences up to 10.3% after 5 rounds of MDA, down from achieving these targets at up to 13.7% and 14.3% baseline prevalences respectively using population-wide MDA. A 90% reduction in baseline prevalence at 80% CI was achieved for baseline prevalences up to 0.79% after 3 rounds of MDA, and baseline prevalences up to 4.7% after 5 rounds of MDA, down from 7.7% and 9.5% respectively.

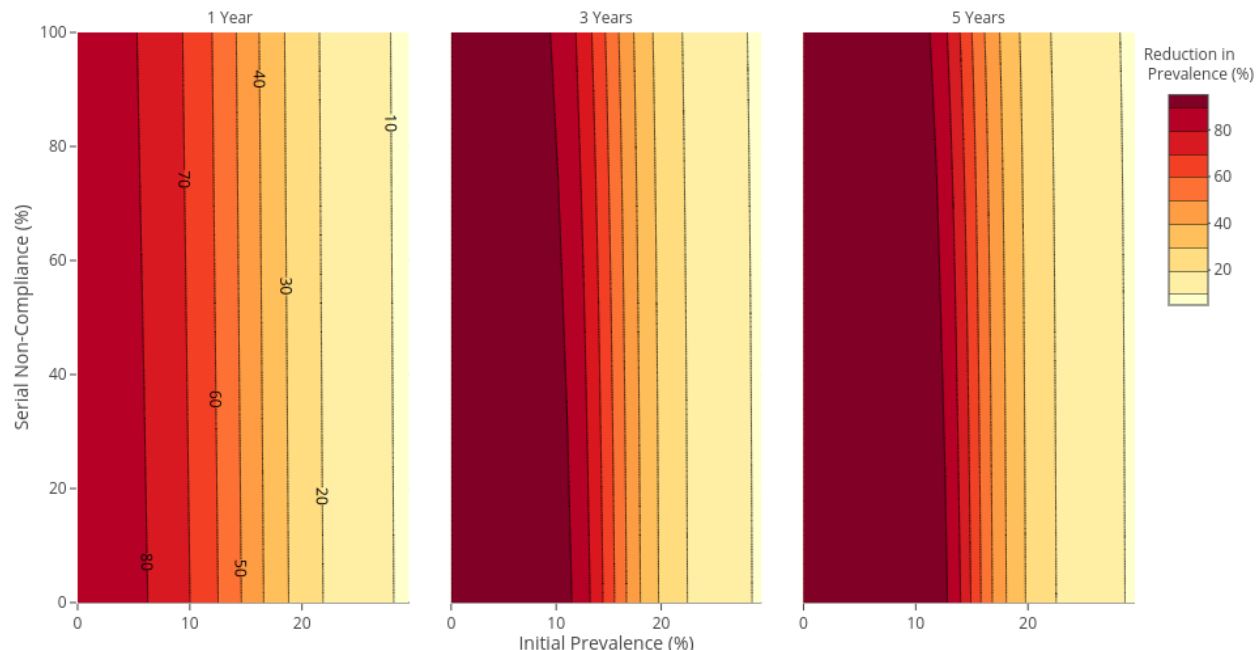


Figure 3.4: Effect of Serial Non-Compliance with the MDA upon the reduction in prevalence after 1, 3 and 5 years of azithromycin MDA.

Serial non-compliance is varied from 0 (exposure to MDA is completely random year-over-year) to 1 (perfect non-compliance where those who are unexposed to MDA stay unexposed) when compliance is set at 90%. The % reduction in prevalence was calculated as the reduction in total population prevalence one year after the last annual MDA treatment compared to the baseline equilibrium for each simulation.

Serial non-compliance with MDA has often been seen to limit the efficacy of such interventions, as a consistent subset of the population is failing to receive MDA treatment and is acting as a vehicle for continued transmission. When serial non-compliance with the azithromycin MDA is increased from 20% to 80%, the baseline prevalence for which population-wide prevalence dropped below 1 in 10,000 was reduced from 5.9% to 2.8% after 5 rounds of MDA (Figure 3.4). Similar results were seen for targets of 50% and 90% reduction in population prevalence relative to baseline where, after 5 rounds of MDA, when serial non-compliance is increased from 20% to 80%, the baseline prevalence for which a 50% reduction is achieved drops from 16.5% to 15.9% with higher serial non-compliance, and starting prevalences achieving a 90% target reduction dropped from 12.5% to 11.0% with

higher serial non-compliance.

From an uncertainty analysis allowing model parameters to be randomly selected from their distribution (Tables 3.1 and 3.2), there appeared to be a reasonable amount of uncertainty in the model predictions. At 80% CE, a 50% reduction is predicted to occur between 6.8% and 20.2% after 3 rounds of MDA, and between 7.2% and 20.3% after 5 rounds of MDA. Similarly a 90% reduction is predicted to occur between 0% and 8.3% after 3 rounds of MDA, and between 0% and 10.5% after 5 rounds of MDA.

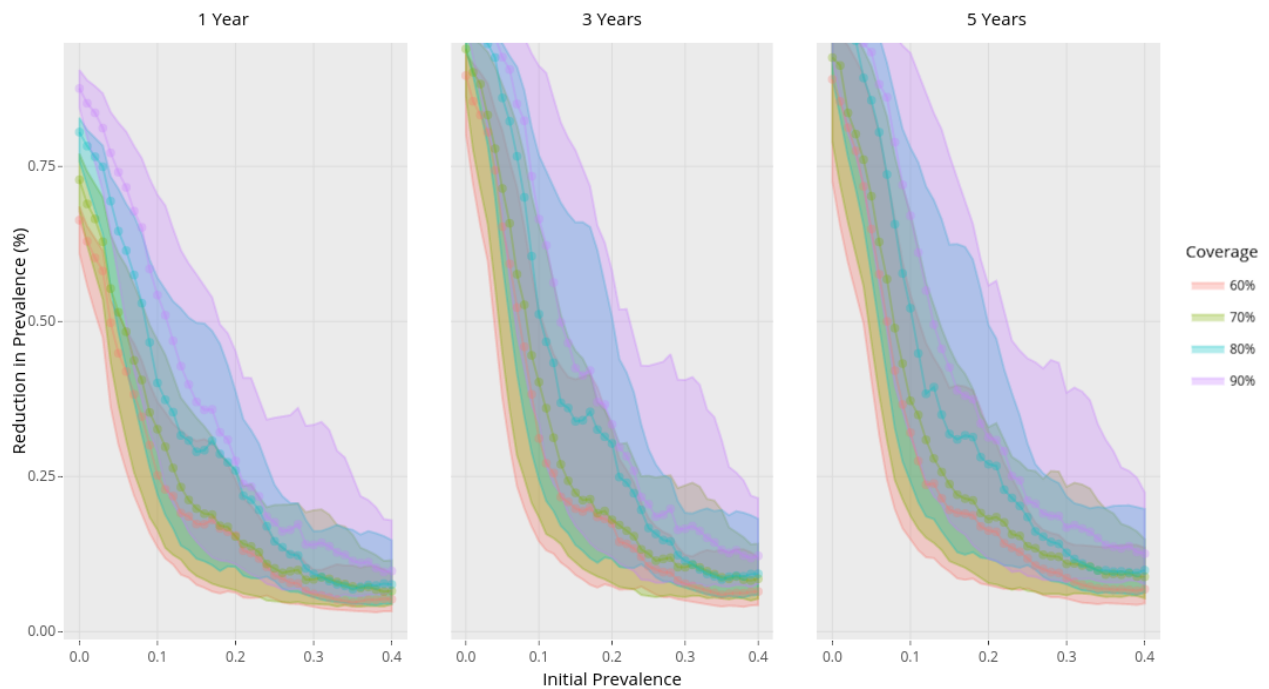


Figure 3.5: Uncertainty Analysis for Predicting Reduction in Population Prevalence

Reduction in % population prevalence relative to baseline across a range of starting prevalences at 60%, 70%, 80% and 90% population coverage, reported after 1, 3, and 5 rounds of MDA. Shown with 90% confidence intervals from 10,000 simulations randomly sampled parameter distributions in Tables 3.1 and 3.2.

3.4 Discussion

One principal advantage the trachoma MDA program enjoys in fighting Chlamydia is a near-simultaneous treatment of all individuals. Reducing the delay between the treatment of an index case and an infected partner has been shown to greatly reduce the probability of reinfection from nearly 20% down to 2.5%. [89]

Our model suggests that MDA could reduce the prevalence of GUC to less than 1 in 10,000 within 5 years time for populations with baseline prevalences up to 6.5%. This reinforces the suggestions of potential additional health benefits of trachoma MDA and may also invite potential exploration of targeting high burden populations in impoverished areas. Efforts that would potentially target only high-risk persons could still have a meaningful impact on public health, halving baseline prevalences of up to 10.3% after 3 rounds of MDA and halving baseline prevalences of up to 13.7% after a full 5 rounds of MDA.

However, if the reduction target is set too low, there is likely be an observed resurgence in prevalence following termination of a program, and costs and benefits should be weighed accordingly. Populations being targeted for trachoma elimination tend to be poor, and with a very low population exposure to antibiotics [72]. Such a program is likely to have a widespread beneficial impact [111]. However, modeling is often prone to overestimate the impact of such interventions.

There are several limitations in our analysis. First, our study does not distinguish between natural recovery from a Chlamydial infection and those seeking treatment with antibiotics outside of MDA, though the manner of recovery is known to have some impact on the level of acquired immunity following infection [91]. Second, though the populations targeted by trachoma have very low reported antibiotic exposure, there is a non-zero, low-level rate

of treatment outside of the MDA, and thus we may have overstated the impact of the trachoma MDA program upon reducing Chlamydia prevalence. Similarly, we have made a closed system that doesn't allow for re-introduction from outside the target population. Real world inter-mixing and migration is likely to have an impact on any predicted program benefits. We also did not model fast versus slow clearing infections which have been shown to act as a natural reservoir for chlamydia to persist in an untreated population, though it is likely this would lead to an underrepresentation of the model impacts as the MDA would treat the slow clearing infections [107].

Similarly, we did not explicitly model sexual behaviors affecting risk, such as condom use, however we did adjust the probability of transmission per partnership to reflect the reported rates of sexually protective behaviors, however the reported rate of protective behaviors, such as condom use was anywhere from 50-80% [99, 104]. Our model was fairly simple, and there are more sophisticated models available, such as an agent-based approach which might do better. It would be insightful to compare the results of micro-agent based simulations to this approach. Lastly, our model depends on observed data only through parameters based upon prior publications. We were not aiming to target to a specific country, but rather a general region affected by trachoma, namely sub-saharan Africa, that shows a large level of heterogeneity, our parameters are only able to roughly target precise sub-regional values and differences. Seemingly small changes in transmission, sexual behaviors of high and low risk groups and especially the behaviors of the FSW may have a large impact on the transportability of results. It would be appropriate to fit this model to a particular country or region's demographics to better understand the impact of trachoma MDA upon local GU Chlamydia burden. However, our model does suggest testable hypotheses regarding prevalence in endemic areas under treatment.

Overall, we find that trachoma MDA efforts are almost certainly having an impact upon Chlamydia prevalence in the targeted population, as the one year lag between treatments was not enough to support prevalence returning to the original levels. With the continued challenges presented by GU Chlamydia in low-resource settings, and findings that the benefits of screening programs were overestimated [82], the use of MDA or targeted administration of treatment with azithromycin (using PPT of sex-workers) in extremely high prevalence areas may be the way forward to addressing the the continued burden of GU Chlamydia [83–88].

Bibliography

- [1] World Health Organization Regional Office for Africa. Health topics: Ebola virus disease. <http://www.afro.who.int/health-topics/ebola-virus-disease>, 2018. Online; accessed 5-November-2018.
- [2] Outbreaks chronology: Ebola virus disease. pages <https://www.cdc.gov/vhf/ebola/outbreaks/history/chronology.html>, archived at <http://www.webcitation.org/6wduMHBrt>.
- [3] World Health Organization Regional Office for Africa. Health topics: Ebola situation reports: Democratic republic of congo. <https://www.who.int/ebola/situation-reports/drc-2018/en/>, 2019. Online; accessed 10-January-2018.
- [4] L. Worden, S. R. Wannier, N.A. Hoff, K. Musene, B. Selo, M. Mossoko, E. Okitolonda-Wemakoy, J. J. Muyembe-Tamfum, G. W. Rutherford, T. M. Lietman, A. W. Rimoin, T. C. Porco, and J. D. Kelly. Real-time projections of epidemic transmission and estimation of vaccination impact during an Ebola virus disease outbreak in the eastern region of the Democratic Republic of Congo. *bioRxiv*, 2018.
- [5] Kelly JD, L. Worden, S. R. Wannier, N.A. Hoff, Mukadi P, Sinai C, Ackley S, Chen X, Gao D, Selo B, Mossoko M, Okitolonda-Wemakoy E, Richardson ET, G. W. Ruther-

- ford, T. M. Lietman, J. J. Muyembe-Tamfum, A. W. Rimoin, and T. C. Porco. Projections of Ebola outbreak size and duration with and without vaccine use in Équateur, Democratic Republic of Congo, as of May 27, 2018. *PLoS One*, 14(3):e0213190, 2019.
- [6] Enanoria WT, Worden L, Liu F, Gao D, Ackley S, Scott J, Deiner M, Mwebaze E, Ip W, Lietman TM, and Porco TC. Evaluating subcriticality during the Ebola epidemic in West Africa. *PLoS One*, 10(10):e0140651, 2015.
- [7] Ebola haemorrhagic fever in sudan, 1976. report of a who/international study team. *Bull World Health Organ*, 56(2):247–70, 1978.
- [8] J. G. Breman, D. L. Heymann, G. Lloyd, J. B. McCormick, M. Miatudila, F. A. Murphy, J. J. Muyembé-Tamfun, P. Piot, J. F. Ruppol, P. Sureau, G. van der Groen, and K. M. Johnson. Discovery and description of Ebola Zaire Virus in 1976 and relevance to the West African epidemic during 2013–2016. *The Journal of infectious diseases*, 214:S93–S101, 2016.
- [9] Christopher Dickey. What’s worse than ebola? fighting it in a war zone. <https://www.thedailybeast.com/whats-worse-than-ebola-fighting-it-in-a-war-zone>, 2018. Online; accessed 20-October-2018.
- [10] Branswell H. Ebola response suffers another setback, as WHO evacuates some staff after attack. *Stat*, 11 2018. Accessed: 2019-05-10. (Archived by WebCiteó at <http://www.webcitation.org/78GtZOWVV>).
- [11] Branswell H. Ebola response teams scrambling to care for patients after attacks set back efforts. *Stat*, 03 2019. Accessed: 2019-05-10. (Archived by WebCite at <http://www.webcitation.org/78Gu5eqcW>).

- [12] Branswell H. 'on a knife edge': Ebola outbreak threatens to escalate as violence rises. *Stat*, 05 2019. Accessed: 2019-05-10. (Archived by WebCite at <http://www.webcitation.org/78GuOL4UO>).
- [13] Branswell H. CDC's Redfield: it could take another year to control Ebola in DRC. *Stat*, 03 2019. Accessed: 2019-05-10. (Archived by WebCite at <http://www.webcitation.org/78GvnBbNX>).
- [14] Branswell H. Ebola response is working, WHO director-general says, amid criticism and violence. *Stat*, 03 2019. Accessed: 2019-05-10. (Archived by WebCite at <http://www.webcitation.org/78GwJ4itx>).
- [15] M. S. ElSherif, C. Brown, D. MacKinnon-Cameron, L. Li, T. Racine, J. Alimonti, T. L. Rudge, C. Sabourin, P. Silvera, J. W. Hooper, S. A. Kwilas, N. Kilgore, C. Badorrek, W. J. Ramsey, D. G. Heppner, T. Kemp, T. P. Monath, T. Nowak, S. A. McNeil, J. M. Langley, S. A. Halperin, and Canadian Immunization Research Network. Assessing the safety and immunogenicity of recombinant vesicular stomatitis virus Ebola vaccine in healthy adults: a randomized clinical trial. *Canadian Medical Association journal*, 189:E819–E827, 2017.
- [16] WHO. Preliminary results on the efficacy of rVSV-ZEBOV-GP Ebola vaccine using the ring vaccination strategy in the control of an Ebola outbreak in the democratic republic of the congo: an example of integration of research into epidemic response. 04 2019.
- [17] J. Wallinga and P. Teunis. Different epidemic curves for severe acute respiratory syndrome reveal similar impacts of control measures. *Am J Epidemiol*, 160(6):509–16,

2004.

- [18] B Aylward, P Barboza, L Bawo, and et al. The impact of control strategies and behavioural changes on the elimination of Ebola from Lofa County, Liberia. *N Engl J Med*, 371(16):1481–95, 2014.
- [19] ZS Wong, CM Bui, AA Chughtai, and CR Macintyre. A systematic review of early modelling studies of Ebola virus disease in West Africa. *Epidemiol Infect*, 145(6):1069–94, 2017.
- [20] G. Chowell and H. Nishiura. Transmission dynamics and control of Ebola virus disease (EVD): a review. *BMC Medicine*, 12:196, 2014.
- [21] R. C. Baron, J. B. McCormick, and O. A. Zubeir. Ebola virus disease in southern Sudan: hospital dissemination and intrafamilial spread. *Bulletin of the World Health Organization*, 61(6):997–1003, 1983.
- [22] A. J. Georges, E. M. Leroy, A. A. Renaut, C. T. Benissan, R. J. Nabias, M. T. Ngoc, P. I. Obiang, J. P. Lepage, E. J. Bertherat, D. D. Bénoni, E. J. Wickings, J. P. Amblard, J. M. Lansoud-Soukate, J. M. Milleliri, S. Baize, and M. C. Georges-Courbot. Ebola hemorrhagic fever outbreaks in Gabon, 1994–1997: epidemiologic and health control issues. *Journal of Infectious Diseases*, 179 Suppl 1:S65–S75, 1999.
- [23] A. S. Khan, F. K. Tshioko, D. L. Heymann, B. Le Guenno, P. Nabeth, B. Kerstiëns, Y. Fleerackers, P. H. Kilmarx, G. R. Rodier, O. Nkuku, P. E. Rollin, A. Sanchez, S. R. Zaki, R. Swanepoel, O. Tomori, S. T. Nichol, C. J. Peters, J. J. Muyembe-Tamfum, and T. G. Ksiazek. The reemergence of Ebola hemorrhagic fever, Democratic Republic

- of the Congo, 1995. Commission de Lutte contre les Epidémies à Kikwit. 179 Suppl 1:S76–S86, 1999.
- [24] CDC and Ministry of Health: T. Oyok, C Odonga, E Mulwani, J Abur, F Kaducu, M Akech, J Olango, P Onek, J Turyanika, I Mutyaba, HRS Luwaga, G Bisoborwa, A Kaguna, FG Omaswa, S Zaramba, S Okware, A Opio, J Amandua, J Kamugisha, and E Mukoyo. Outbreak of ebola hemorrhagic fever — uganda, august 2000–january 2001. *MMWR: Morbidity and Mortality Weekly Report*, 50(5):73–77, 2001.
- [25] Outbreak(s) of ebola haemorrhagic fever, congo and gabon, october 2001-july 2002. *Wkly Epidemiol Rec*, 78(26):223–8, 2003.
- [26] P Boumandouki, P Formenty, A Epelboin, P Campbell, C Atsangandoko, Y Alarangar, EM Leroy, ML Kone, A Molamou, O Dinga-Longa, A Salemo, RY Kounkou, V Mombouli, JR Ibara, P Gaturuku, S Nkunku, A Lucht, and H Feldmann. Prise en charge des malades et des défunts lors de l'épidémie de fièvre hémorragique due au virus ebola d'octobre à décembre 2003. *Bulletin de la Société de pathologie exotique*, 98(3):218–223, 2005.
- [27] World Health Organization. Weekly epidemiological record. *Abonnement annuel*, 43(80):369–376, 2005.
- [28] D. Nkoghe, M. L. Kone, A. Yada, and E. Leroy. A limited outbreak of Ebola haemorrhagic fever in Etoumbi, Republic of Congo, 2005. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 105(8):466–472, 2011.
- [29] A. Rosello, M. Mossoko, S. Flasche, A. J. Van Hoek, P. Mbala, A. Camacho, S. Funk, A. Kucharski, B. K. Ilunga, W. J. Edmunds, P. Piot, M. Baguelin, and J. J. Tamfum.

- Ebola virus disease in the Democratic Republic of the Congo, 1976–2014. *Elife*, 4, 2015.
- [30] M. S. Bartlett. On the theoretical specification and sampling properties of autocorrelated time-series. *Suppl. J. Roy. Statist. Soc.*, 8:27–41, 1946.
- [31] Dimitris N. Politis. The impact of bootstrap methods on time series analysis. *Statistical Science*, 18(2):219–230, 2003.
- [32] S. G. From. Confidence intervals for Gini’s diversity measure and Shannon’s entropy using adjusted proportions. *Communications in Statistics—Theory and Methods*, 32(5):935–954, 2003.
- [33] M. V. Barbarossa, A. Dénes, G. Kiss, Y. Nakata, G. Röst, and Z. Vizi. Transmission dynamics and final epidemic size of Ebola Virus disease outbreaks with varying interventions. *PLoS One*, 10(7):e0131398, 2015.
- [34] M. S. Y. Lau, B. D. Dalziel, S. Funk, A. McClelland, A. Tiffany, S. Riley, C. J. E. Metcalf, and B. T. Grenfell. Spatial and temporal dynamics of superspreading events in the 2014–2015 West Africa Ebola epidemic. *Proceedings of the National Academy of Sciences of the United States of America*, 114:2337–2342, 2017.
- [35] ET Richardson and MP Fallah. The genesis of the Ebola outbreak in West Africa. *The Lancet Infectious Diseases*, 19(4):348–349, 2019.
- [36] RG Frankfurter, M Kardas-Nelson, A Benton, MB Barrie, Y Dibba, PE Farmer, and ET Richardson. Indirect rule redux: The political economy of diamond mining and its relation to the Ebola outbreak in Kono District, Sierra Leone. *Review of African Political Economy*, 45(158):1–19, 2019.

- [37] ET Richardson, JD Kelly, O Sesay, MD Drasher, IK Desai, R Frankfurter, PE Farmer, and MB. Barrie. The symbolic violence of outbreak: A mixed methods, quasi-experimental impact evaluation of social protection on Ebola survivor wellbeing. *Social Science and Medicine*, 195:77–82, 2017.
- [38] ET Richardson, MB Barrie, CT Nutt, JD Kelly, R Frankfurter, MP Fallah, and PE Farmer. The Ebola suspect’s dilemma. *The Lancet Global Health*, 5(3):e254–e256, 2017.
- [39] ET Richardson, MB Barrie, JD Kelly, Y Dibba, S Koedoyoma, and PE Farmer. Biosocial approaches to the 2013-16 Ebola pandemic. *Health and Human Rights*, 18(1):115–128, 2016.
- [40] ET Richardson, JD Kelly, MB Barrie, AW Mesman, S Karku, K Quiwa, RH Marsh, S Koedoyoma, F Daboh, KP Barron, M Grady, E Tucker, KL Dierberg, GW Rutherford, M Barry, JH Hones, MB Murray, and PE Farmer. Minimally symptomatic infection in an Ebola ‘Hotspot’: A cross-sectional serosurvey. *PLoS Negl Trop Dis.*, 10(11):e0005087, 2016.
- [41] Kelly JD, Barrie MB, Mesman AW, Karku S, Quiwa K, Drasher M, Schlough GW, Dierberg K, Koedoyoma S, Lindan CP, Jones JH, Chamie G, Worden L, Greenhouse B, Weiser SD, Porco TC, Rutherford GW, and Richardson ET. Anatomy of a hotspot: Chain and seroepidemiology of Ebola virus transmission, Sukudu, Sierra Leone, 2015-16. *Journal of Infectious Diseases*, 217(8):1214–1221, 2018.
- [42] J. Sweeney. DRC: Conflict and displacement in Nord Kivu and Ituri, 5 2019. ACAPS Briefing Note: Displacement 14-May 2019.

- [43] Nzongola-Ntalaja G. *The Congo from Leopold to Kabila: A People's History*. Zed Books, London, 2013.
- [44] Council on Foreign Relations. Violence in the democratic republic of congo. *Global Conflict Tracker*, 05 2019.
- [45] Rdc: Ebola doit être vaincu le plus tôt (Tshibuyi). *Actualite*, 2019, April 27.
- [46] New Ebola outbreak detected in northwest Democratic Republic of the Congo; WHO surge team supporting the response, 2019.
- [47] WHO. Ebola virus disease—eleventh outbreak—Democratic Republic of the Congo—2020. *Communicable Disease Threats Report*, Week 28, 5-11 July 2020, 2020.
- [48] WHO. Coronavirus disease (covid-19) situation report 172.
- [49] WHO. Deaths from Democratic Republic of the Congo measles outbreak top 6000.
- [50] Covid-19 virtual press conference - dr. michael ryan. *WHO*, July 2020.
- [51] A. MacNeil, E. C. Farnon, O. W. Morgan, P. Gould, T. K. Boehmer, D. D. Blaney, P. Wiersma, J. W. Tappero, S. T. Nichol, T. G. Ksiazek, and P. E. Rollin. Filovirus outbreak detection and surveillance: lessons from bundibugyo. *J Infect Dis*, 204 Suppl 3:S761–7, 2011.
- [52] Uganda: Ebola situation report. *Bull World Health Organ*, 2013.
- [53] Centers for Disease Control and Prevention. Number of cases and deaths in guinea, liberia, and sierra leone during the 2014-2016 west africa ebola outbreak. 2017. Online; accessed 10-May-2018.

- [54] S. R. Wannier, L. Worden, N. A. Hoff, E. Amezcua, B. Selo, C. Sinai, M. Mossoko, B. Njoloko, E. Okitolonda-Wemakoy, P. Mbala-Kingebeni, S. Ahuka-Mundeke, J. J. Muyembe-Tamfum, E. T. Richardson, G. W. Rutherford, J. H. Jones, T. M. Lietman, A. W. Rimoin, T. C. Porco, and J. D. Kelly. Estimating the impact of violent events on transmission in Ebola virus disease outbreak, Democratic Republic of the Congo, 2018–2019. *Epidemics*, 28:100353, 2019.
- [55] Excitement over use of Ebola vaccine in outbreak tempered by real-world challenges. 23 May 2018.
- [56] D. Champredon, M. Li, B. M. Bolker, and J. Dushoff. Two approaches to forecast Ebola synthetic epidemics. *Epidemics*, 2017.
- [57] M Li, J Dushoff, and BM Bolker. Fitting mechanistic epidemic models to data: A comparison of simple markov chain monte carlo approaches. *Stat Methods Med Res*, 27(7):1956–1967, 2018.
- [58] J. D. Kelly, L. Worden, S. R. Wannier, N. A. Hoff, P. Mukadi, C. Sinai, S. Ackley, X. Chen, D. Gao, B. Selo, M. Mossoko, E. Okitolonda-Wemakoy, E. T. Richardson, G. W. Rutherford, T. M. Lietman, J. J. Muyembe-Tamfum, A. W. Rimoin, and T. C. Porco. Projections of Ebola outbreak size and duration with and without vaccine use in Équateur, Democratic Republic of Congo, as of May 27, 2018. *PLoS One*, 14(3):e0213190, 2019.
- [59] Viboud C, Sun K, Gaffey R, Ajelli M, Fumanelli L, Merler S, Zhang Q, Chowell G, Simonsen L, Vespignani A, and RAPIDD Ebola Forecasting Challenge group. the

- RAPIDD ebola forecasting challenge: Synthesis and lessons learnt. *Epidemics*, 22:13–21, 2018.
- [60] L. Worden, R. Wannier, N. A. Hoff, K. Musene, B. Selo, M. Mossoko, E. Okitolonda-Wemakoy, J. J. Muyembe Tamfum, G. W. Rutherford, T. M. Lietman, A. W. Rimoin, T. C. Porco, and J. D. Kelly. Projections of epidemic transmission and estimation of vaccination impact during an ongoing Ebola virus disease outbreak in northeastern Democratic Republic of Congo, as of Feb. 25, 2019. *PLoS Neglected Tropical Diseases*, 13(8):e0007512, 2019.
- [61] J. R. Gott. Implications of the Copernican principle for our future prospects. *Nature*, 363:315–319, 1993.
- [62] S. Blumberg and J. O. Lloyd-Smith. Comparing methods for estimating r_0 from the size distribution of subcritical transmission chains. *Epidemics*, 5(3):131–45, 2013.
- [63] 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):355–9, 2005.
- [64] Althaus CL. the RAPIDD ebola forecasting challenge: Synthesis and lessons learnt. *Lancet*, 15(5):P507–508, 2015.
- [65] Toth DJA, Gundlapalli AV, Khader K, Pettey WBP, Rubin MA, Adler FR, and Samore MH. Estimates of outbreak risk from new introductions of ebola with immediate and delayed transmission control. *Emerg Infect Dis*, 21(8):1402–1408, 2015.
- [66] Glennon EE, Jephcott FL, Restif O, and Wood JLN. Beyond r_0 : Heterogeneity in secondary infections and probabilistic epidemic forecasting. *PLoS Negl Trop Dis*, 13(6):e0007428, 2019.

- [67] Stineman RW. A consistently well-behaved method of interpolation.
- [68] Höbert-Dufresne L, Althouse BM, Scarpino SV, and Allard A. Beyond r_0 : Heterogeneity in secondary infections and probabilistic epidemic forecasting. *Journal of the Royal Society Interface*, pages 1–8, 2020.
- [69] Smith AFM and Gelfand AE. Bayesian statistics without tears: A sampling-resampling perspective. *The American Statistician*, 46(2):84–88, 1992.
- [70] Okware SI, Omaswa FG, Zaramba S, Opio A, Lutwama JJ, Kamugisha J, Rwaguma EB, Kagwa P, and Lamunu M. An outbreak of ebola in uganda. *Trop Med and Internat Health*, 7(12):1068–1075, 2002.
- [71] Covid-19 virtual press conference - dr. michael ryan. *WHO*, september 2020.
- [72] Lietman TM, Gebre T, Ayele B, Ray KJ, Maher MC, See CW, Emerson PM, and Porco TC. The epidemiological dynamics of infectious trachoma may facilitate elimination. *Epidemics*, 3:119–124, 2011.
- [73] Marks M, Kako H, Butcher R, and et al. Prevalence of sexually transmitted infections in 20 female clinic attendees in Honiara, Solomon Islands. *BMJ Open*, (5:e007276), 2015.
- [74] Marks M, Bottomley C, Tome H, Pitakaka R, Butcher R, Sokana O, Kako H, Solomon AW, and Mabey DC. Mass drug administration of azithromycin for trachoma reduces the prevalence of genital chlamydia trachomatis infection in the Solomon Islands. *BMJ Sex Transm Infect*, 92:261–265, 2016.

- [75] Johnson LF, Dorrington RE, and Bradshaw D. The role of immunity in the epidemiology of gonorrhoea, chlamydial infection and trichomoniasis: insights from a mathematical model. *Epidemiology and Infection*, 139(12):1875–1883, 2011.
- [76] Mullick S, Watson-Jones D, Beksinska M, and et al. Sexually transmitted infections in pregnancy: prevalence, impact on pregnancy outcomes, and approach to treatment in developing countries. *Sex Trasm Infect*, 81.:294–302, 2005.
- [77] Norman J. Epidemiology of female genital chlamydia trachomatis infections. *Best Practice & Research Clinical Obstetrics and Gynaecology*, 16(6):775–787, 2002.
- [78] Newman L, Rowley J, Vander Hoorn S, Wijesooriya NS, Unemo M, Low N, Stevens G, Gottlieb S, Kiarie J, and Temmerman M. Global estimates of the prevalence and incidence of four curable sexually transmitted infections in 2012 based on systematic review and global reporting. *PLoS One*, 10(12):e0143304, 2015.
- [79] Wangnapi RA, Soso S, Unger HW, and et al. Prevalence and risk factors for chlamydia trachomatis, neisseria gonorrhoeae and trichomonas vaginalis infection in pregnant women in Papua New Guinea. 91:194–200, 2015.
- [80] Walsh MS, Hope E, Isaia L, and et al. Prevalence of chlamydia trachomatis infection in Samoan women aged 18 to 29 and assessment of possible risk factors: a community-based study. *Trans R Soc Trop Med Hyg*, 109:245–51, 2015.
- [81] Gravett MG, Nelson HP, DeRouen T, and et al. Independent associations of bacterial vaginosis and chlamydia trachomatis infection with adverse pregnancy outcome. *JAMA*, 256:1899–1903., 1986.

- [82] Ronn MM, Wolf EE, Chesson H, Menzies NA, Galer K, Gorwitz R, Gift T, Hsu K, and Salomon JA. The use of mathematical models of chlamydia transmission to address public health policy questions: A systematic review. *Sexually Transmitted Diseases*, 44(5), 2017.
- [83] O'Farrell N, Oula R, Morison L, and et al. Periodic presumptive treatment for cervical infections in service women in 3 border provinces of Laos. *Sex Transm Dis*, 33:558e64, 2006.
- [84] Steen R, Vuylsteke B, DeCoito T, and et al. Evidence of declining STD prevalence in a South African mining community following a core-group intervention. *Sex Transm Dis*, 27:1e8, 2000.
- [85] Labbe AC, Dzokoto D, Nzambi K, and et al. A randomised placebo-controlled trial of routine monthly antibiotics against gonococcal and chlamydial infections among female sex workers in Ghana and Benin: intention-to-treat analysis. *15th Biennial Congress of the ISSTD*, 27e30, July 2003. Ottawa.
- [86] Vickerman P, Terris-Prestholt F, Delaney S, and et al. Are targeted HIV prevention activities still cost-effective in high prevalence settings? Results from an STI treatment intervention for sex-workers in Hillbrow, South Africa. *Sex Transm Dis*, 33(Suppl):S122e32, 2006.
- [87] Kaul R, Kimani J, Nagelkerke N, and et al. Monthly antibiotic chemoprophylaxis and incidence of sexually transmitted infections and HIV-1 infections in Kenyan sex workers: a randomized controlled trial. *JAMA*, 291:2555e62, 2004.

- [88] Cowan FM, Hargrove JW, Langhaug LF, and et al. The appropriateness of core group interventions using presumptive periodic treatment among rural Zimbabwean women who exchange sex for gifts or money. *J Acquir Immune Defic Syndr*, 38:202e7, 2005.
- [89] Low N, Heijne JCM, Herzon SA, and Althaus CL. Reinfection by untreated partners of people treated for chlamydia trachomatis and neisseria gonorrhoeae: mathematical modelling study. *Sexually Transmitted Infections*, 90:254–256, 2014.
- [90] Harding-Esch EM, Sillah A, Edwards T, Burr SE, Hart JD, Joof H, Laye M, Makalo P, Manjang A, Molina S, Sarr-Sissoho I, Quinn TC, Lietman T, Holland MJ, Mabey D, West SK, Bailey R, and Partnership for Rapid Elimination of Trachoma (PRET) study group. Mass treatment with azithromycin for trachoma: when is one round enough? results from the PRET trial in the gambia. *PLoS Negl Trop Dis.*, 192(7):e21115, 2013.
- [91] Brunham RC, Pourbohloul B, Mak S, White R, and Rekart ML. The unexpected impact of a chlamydia trachomatis infection control program on susceptibility to reinfection. *Journal of Infectious Diseases*, 192:1836–44, 2005.
- [92] Worden L and Porco TC. Products of compartmental models in epidemiology. *Comput Math Methods Med.*, page 8613878, 2017.
- [93] Hethcote HW and Yorke JA. Gonorrhea. transmission dynamics and control. *Biom. J.*, 28(304), 1986.
- [94] Asare K, Osman F, Ngcapu S, Vandormael A, Naicker N, Khanyile M, Mindel A, Karim SSA, Tomita A, and Garrett N. Burden of sexually transmitted infections from acute hiv infection among women in south africa: Evidence from a prospective cohort study. *Ann Epidemiol*, 74:132–139, 2022.

- [95] Olson KM, Geisler WM, Bakshi RK, Gupta K, and Tiwari HK. Predicting the probability of chlamydia reinfection in african american women using immunologic and genetic determinants in a bayesian model. *Sex Transm Dis*, 48(11):813–818, 2021.
- [96] Okiring J, Getahun M, Gutin SA, Lebu S, Lee J, Maeri I, Eyul P, Bukusi EA, Cohen CR, Neilands TB, Ssali S, Charlebois ED, and Camlin CS. Sexual partnership concurrency and age disparities associated with sexually transmitted infection and risk behavior in rural communities in kenya and uganda. *Int J Infect Dis*, 120:158–167, 2022.
- [97] Diabate S, Chamberland A, Geraldo N, Tremblay C, and Alary M. Gonorrhoea, chlamydia and hiv incidence among female sex workers in cotonou, benin: A longitudinal study. *PLoS One*, 13(5):e0197251, 2018.
- [98] J Wu, R Dhingra, M Gamghir, and JV Remais. Sensitivity analysis of infectious disease models: methods, advances and their application. *J R Soc Interface*, 10(86):20121018, 2013.
- [99] Government of the Republic of South Sudan, Ministry of Health. A bio-behavioral HIV survey of female sex workers in South Sudan. The Eagle Survey. A final report. 2016.
- [100] Ngugi EN, Roth E, Mastin T, Nderitu MG, and Yasmin S. Female sex workers in Africa: Epidemiology overview, data gaps, ways forward. *Sahara J*, 9(3):148–153, 2012.

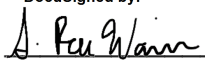
- [101] Looker KJ, Wallace LA, and Turner KME. Impact and cost-effectiveness of chlamydia testing in Scotland: a mathematical modelling study. *Theoretical Biology and Medical Modelling*, 12(2):1–18, 2015.
- [102] Althaus CL, Heijne JCM, and Low N. Towards more robust estimates of the transmissibility of chlamydia trachomatis. *Sexually Transmitted Diseases*, 39(5):402–4, 2012.
- [103] Owusu-Edusei K, Chesson HW, Gift T, Brunham RC, and Bolan G. Cost-effectiveness of chlamydia vaccination programs for young women. *Emerging Infectious Diseases*, 21(6):960–968, 2015.
- [104] Vickerman P, Ndowa F, O’Farrell N, Steen R, Alary M, and Delany-Moretlwe S. Using mathematical modelling to estimate the impact of periodic presumptive treatment on the transmission of sexually transmitted infections and HIV among female sex workers. *Epidemiology Sex Transm Infect*, 86:163–168, 2010.
- [105] Liu F, Porco TC, Mkocha HA, Muñoz B, Ray KJ, Bailey RL, Lietman TM, and West SK. The efficacy of oral azithromycin in clearing ocular chlamydia: Mathematical modeling from a community-randomized trachoma trial. *Epidemics*, 6:10–17, 2014.
- [106] Geisler W. Duration of untreated, uncomplicated chlamydia trachomatis genital infection and factors associated with chlamydia resolution: A review of human studies. *The Journal of Infectious Diseases*, 201(Suppl 2):S104–S113, 2010.
- [107] Lewis J, Price MJ, Horner PJ, and White PJ. Genital chlamydia trachomatis infections clear more slowly in men than women, but are less likely to become established. *Journal of Infectious Diseases*, 216:237–44, 2017.

- [108] Chesson HW, Mayaud P, and Aral SO. *Sexually Transmitted Infections: Impact and Cost-Effectiveness of Prevention*. publisher, 2017.
- [109] Mayaud P and Mabey D. Approaches to the control of sexually transmitted infections in developing countries: old problems and modern challenges. *Sexually Transmitted Infections*, 80:174–182, 2004.
- [110] Dyson L, Stolk WA, Farrell SH, and Hollingsworth TD. Measuring and modelling the effects of systematic non-adherence to mass drug administration. *Epidemics*, 18:56–66, 2017.
- [111] Keenan JK, Bailey RL, West SK, Arzika AM, Hart J, Weaver J, Kalua K, Mrango Z, Ray K, Cook C, Lebas E, O’Brien KS, Emerson PM, Porco TC, and Lietman TM for the MORDOR Study Group. Azithromycin to reduce childhood mortality in sub-saharan africa. *N Engl J Med*, 378:1583–1592, 2018.

Publishing Agreement

It is the policy of the University to encourage open access and broad distribution of all theses, dissertations, and manuscripts. The Graduate Division will facilitate the distribution of UCSF theses, dissertations, and manuscripts to the UCSF Library for open access and distribution. UCSF will make such theses, dissertations, and manuscripts accessible to the public and will take reasonable steps to preserve these works in perpetuity.

I hereby grant the non-exclusive, perpetual right to The Regents of the University of California to reproduce, publicly display, distribute, preserve, and publish copies of my thesis, dissertation, or manuscript in any form or media, now existing or later derived, including access online for teaching, research, and public service purposes.

DocuSigned by:

BE5C8EB8B6FE45F... Author Signature

3/20/2023
Date