The Drivers of Acute Seasonal Infectious Diseases

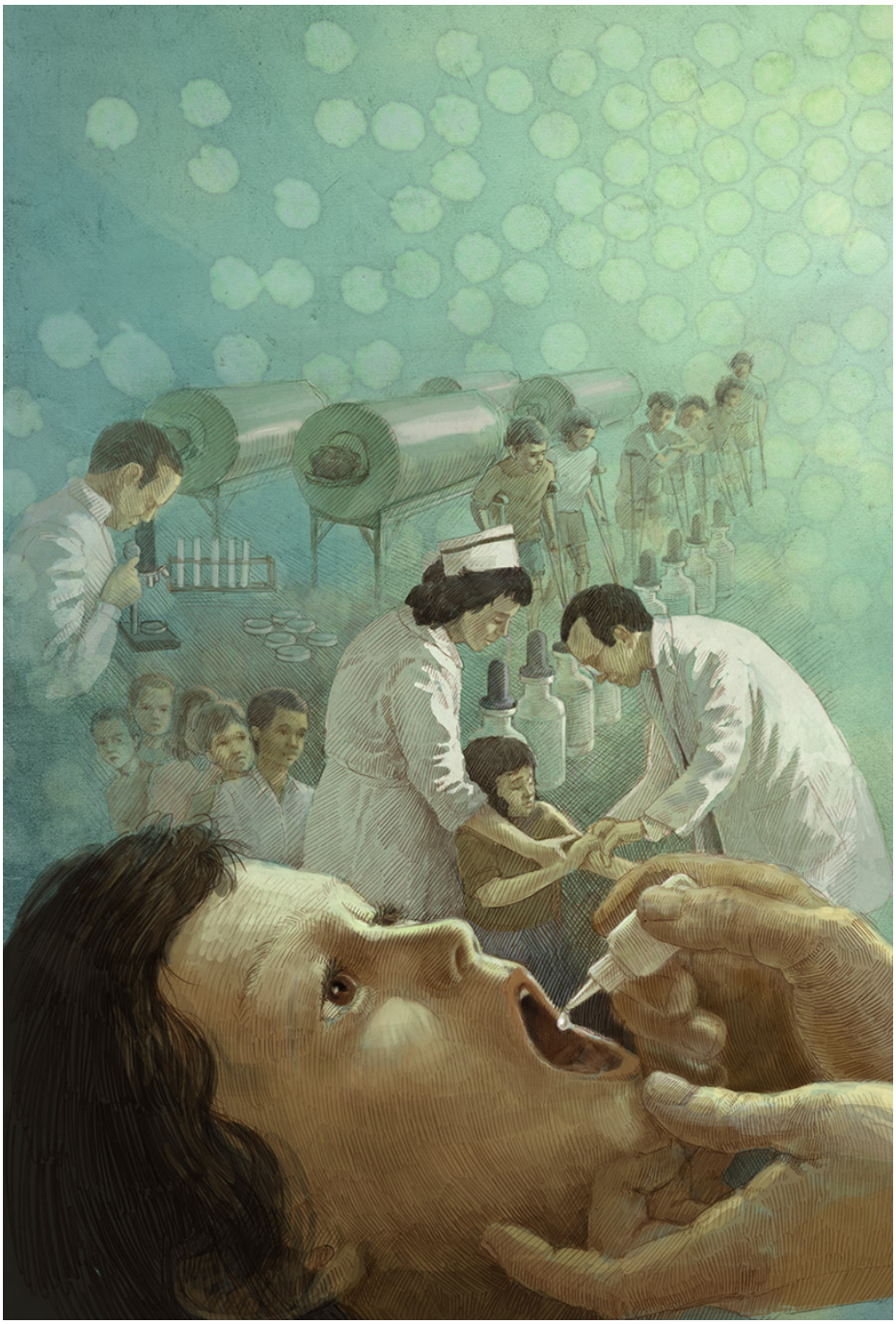
by

Micaela Martinez-Bakker

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Ecology and Evolutionary Biology)
in The University of Michigan
2015

Doctoral Committee:

Professor Pejman Rohani, Co-chair
Professor Aaron A. King, Co-chair
Professor Betsy Foxman
Professor Mercedes Pascual



Polio from the epidemic era, to the Salk and Sabin vaccine eras, and someday global eradication. Art by John Megahan.

ACKNOWLEDGEMENTS

I would like to thank my advisors Pejman Rohani and Aaron A. King. I thank my committee members Mercedes Pascual and Betsy Foxman, and my co-authors Barbara Helm and Kevin M. Bakker.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

Table

LIST OF APPENDICES

Appendix

ABSTRACT

The Drivers of Acute Seasonal Infectious Diseases

by

Micaela Martinez-Bakker

Co-chairs: Pejman Rohani and Aaron A. King

Seasonality is a feature of all ecological systems. Earth's terrestrial and pelagic life has evolved in a highly seasonal abiotic environment with intra-annual variation in photoperiod, temperature, and precipitation, among many other abiotic and biotic factors. Seasonal aspects of mammals and birds include seasonally varying birth rates, seasonal changes in endocrine hormones, and seasonal variation in immunity. One area where seasonal biology is particularly salient is disease ecology. The mechanisms underlying the seasonality of communicable diseases are poorly understood. I propose that much of the unexplained seasonality observed in infectious disease dynamics could be attributed to seasonal biology, including (1) birth seasonality, (2) seasonal variation in immunity, and (3) seasonal cycles in parasite traits and parasite population parameters. In my dissertation, I present work on various aspects of seasonality. In Chapter II, I explored the seasonality of births in human populations and quantified the effects of birth seasonality on measles epidemics. In Chapter III, I reviewed circadian and circannual rhythms in host and parasite populations, and proposed both ecological and evolutionary models for integrating biological rhythms into the study of infectious diseases. In Chapters IV–V, I presented my in-depth

ecological studies of poliovirus, a notoriously seasonal summertime infection. I explored geographical variation in polio's seasonality and tested whether human birth seasonality or transmission seasonality drove epidemics of this disease. In addition to studying polio seasonality, I revealed the connection between (i) polio's emergence and human demography, (ii) the geographical distribution of poliovirus and its persistence, and (iii) polio symptomatology and silent chains of transmission. Lastly, I highlighted the public health implications of seasonal transmission by measuring the efficacy of the two polio vaccines and discussing how seasonality can be utilized for vaccine interventions.

CHAPTER I

Introduction

A unifying feature of infectious diseases is periodicity [1]. Periodicity is observed in infectious disease dynamics at both the population and within-host level. As illustrated by London and Yorke in Fig 1.1, at the population level, infectious disease incidence displays seasonal—and often multiannual—periodicities [2, 3]. At the within-host level, periodicities occur at shorter time scales, such as daily cycles of symptoms and/or parasite abundance [4]. In my dissertation I focused on identifying the drivers of (a) seasonal disease incidence at the population-level, and (b) circadian cycles of infections within-hosts. The systems I studied were polio, measles, and other acute infections. My research on disease seasonality can be categorized into (1) host phenology (i.e., seasonal cycles) and biological rhythms, (2) the ecological role of transmission seasonality, and (3) the public-health implications of transmission seasonality. Given widespread seasonal structuring of host-pathogen interactions, there is much to be gained by exploring the influence of seasonal abiotic and biotic factors on disease dynamics and measuring the relative importance of interconnected seasonal ecological factors. Much attention has been given to seasonal drivers external to host biology, such as temperature, rainfall, humidity, and host aggregation. However, I aimed to elucidate the role of seasonal fluctuations in host biology (i.e. human phenology and biological rhythms) as a driver of disease periodicity.

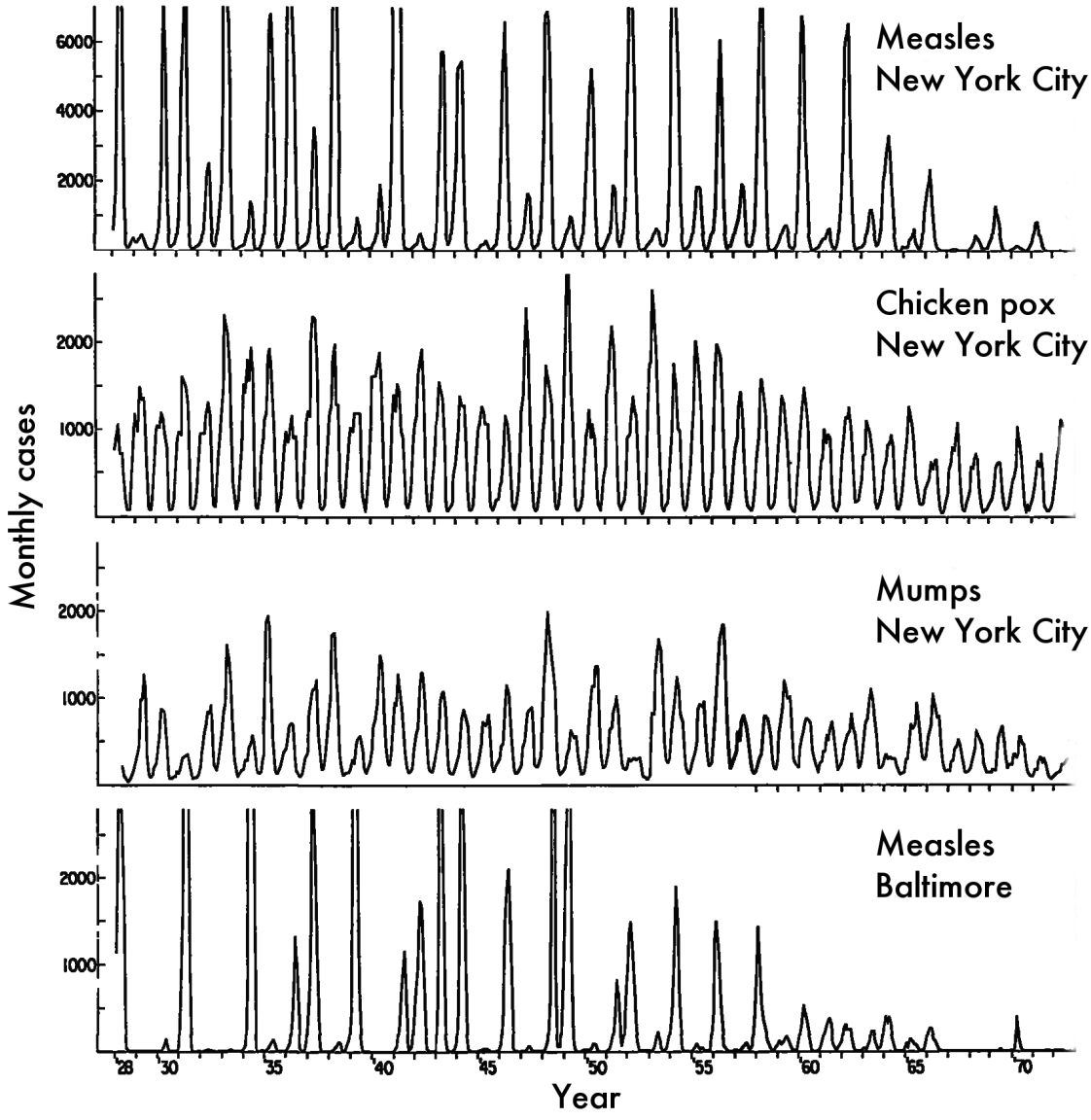


Figure 1.1: **Periodicity of Childhood Diseases.** Time series (top to bottom) of measles, chicken pox, and mumps in New York City, along with measles in Baltimore. Each disease had a seasonal window when epidemics occurred. Epidemics occurred either annually or at multi-annual intervals [5]. Figure reproduced from Wayne P. London, James A. Yorke, Recurrent Outbreaks of Measles, Chickenpox and Mumps: I. Seasonal Variation in Contact Rates, American Journal of Epidemiology, 1973, volume 98, issue 6, 453–468, by permission of Oxford University Press.

1.1 Human Birth Seasonality & Disease

In order to understand the impact of human phenology on seasonal infectious diseases, I studied the influence of human birth seasonality on measles epidemic dynamics. I hypothesized, since fully immunizing childhood diseases, such as measles, rely on births for susceptible recruitment, birth seasonality plays an important role in temporally structuring childhood disease epidemics. Using measles transmission models, I found in Chapter II that high amplitude seasonal fluctuations in births are capable of altering the magnitude and periodicity of measles epidemics. My models predicted that birth seasonality is consequential for measles epidemics in populations with high amplitude birth seasonality, such as populations in Africa. I also found that the phase relationship between the birth pulse and the seasonal peak in transmission is important for determining epidemic size.

Overall, birth seasonality has the effect of shifting the distribution of infections between years. Therefore, birth seasonality does not simply make epidemics "bigger" or "smaller"; rather, birth seasonality can redistribute infections, which has practical implications for epidemic response. In particular, it would be beneficial if public health officials could anticipate the between-year distribution of clinical cases. Models that explicitly account for birth seasonality can facilitate such predictions. Due to (a) the nuanced relationship between birth seasonality and seasonal transmission, and (b) the large amount of regional variation in the timing of birth pulses around the world, I concluded that birth seasonality should be considered for disease control at a regional level. Ideally, birth seasonality would be used to plan vaccination campaigns because birth seasonality is a reliable indicator of the susceptible pool growth rate. Since vaccination initiatives manage the size of the susceptible pool, knowledge of birth seasonality may prove invaluable.

Although my work demonstrated that birth seasonality is relevant for disease dynamics, it is unknown why human births are seasonal. Seasonal reproduction "is

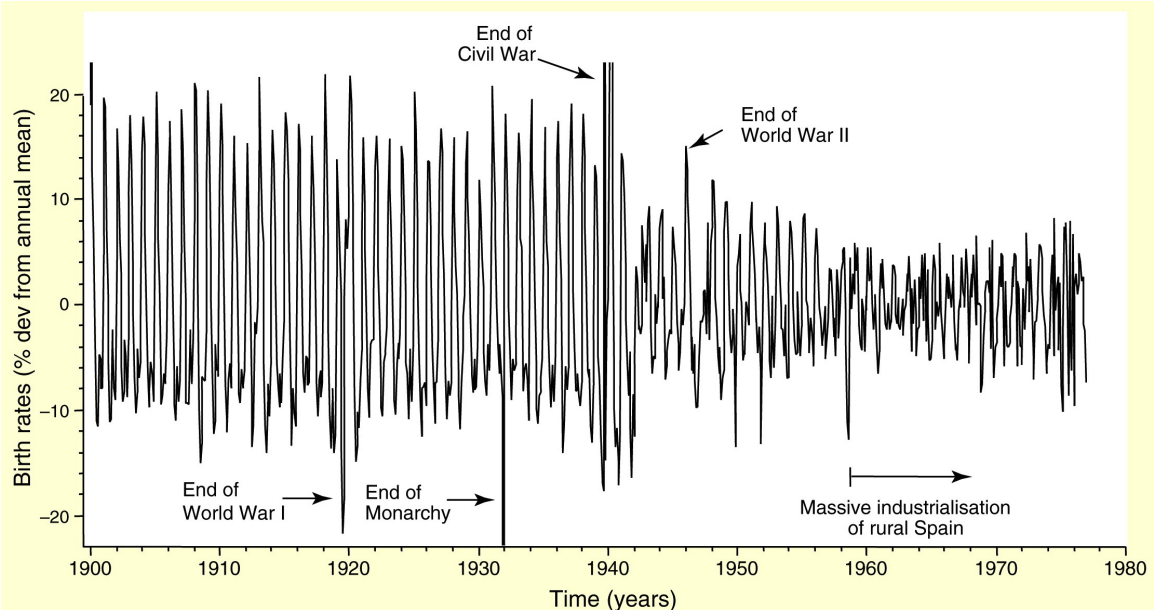


Figure 1.2: **Seasonality of Human Births.** Time series of births in Spain showing consistent seasonality relating to a springtime peak in conception. Societal events, such as wars, altered the amplitude of birth seasonality and introduced irregularities; however, birth seasonality persisted [6]. Reprinted from Current Biology, Vol 18 issue 17, Russell G. Foster, Till Roenneberg, Human Responses to the Geophysical Daily, Annual and Lunar Cycles, Pages R784–R794, Copyright (2008), with permission from Elsevier.

the general rule for the majority of mammals" [7], so it is perhaps not surprising that birth seasonality is a feature of all human populations, both historically and in the modern day. The presence of birth seasonality in populations across geographic regions, cultures, economic regimes, and time periods [8], suggests birth seasonality is a life-history trait, rather than an artifact of culturally motivated human behavior. The persistent seasonality of births is illustrated in Fig 1.2, which shows monthly births in Spain spanning decades [6]. Patterns of birth seasonality are somewhat complex, some populations have biannual cycles of births (i.e., two peaks per year) [9], and as I will discuss in depth in Chapter II, there is a large amount of regional variation in the amplitude and timing of birth seasonality. For instance, within a country the seasonal characteristics of births can vary among populations and shift over time [10].

Using data from the US, Canada, UK, Germany, Greece, Japan, New Zealand, Sweden, South Africa, and Switzerland, Moos & Randall [11] found that each country had a temporally-stable annual conception minimum. The month of minimum-conception had a strong correlation with latitude (correlation coefficient of $r = -0.7678$). The latitudinal pattern of births is highly suggestive that environmental conditions play a role in determining the timing of the conception minimum and the birth pulse. There are various highly debated and rarely agreed upon hypotheses regarding the underlying mechanism of seasonal births in humans. The hypotheses can be broadly summarized as attributing birth seasonality to either behavioral or physiological factors linked to environmental conditions.

Environmental drivers of birth seasonality, proposed by demographers and physiologists, often focus on the effects of temperature or photoperiod on the ability to conceive. Temperature is hypothesized to have a direct effect on conception via its impact on reproductive factors such as sperm count and motility [8]. Whereas, photoperiod, is hypothesized to modulate endogenous physiological rhythms–specifically

in ovarian and androgen activity—that are important for conception in other mammals but are not well characterized in humans [10]. Since it is difficult, or perhaps not plausible, to investigate the mechanisms of birth seasonality using laboratory experiments, studies that have attempted to establish mechanism have been correlative and inconclusive. Importantly, despite underlying mechanisms, the annual variation in human reproduction has implications for susceptible recruitment in childhood disease systems. My work has provided some of the first indication of the public-health importance of human birth seasonality, although the mechanisms of birth seasonality remain nebulous.

1.2 Biological Rhythms & Infection

My research on birth seasonality in humans provided indication that seasonal life history plays an important role in the dynamics of acute infections. I wanted to expand on this and investigate host biological cycles in Chapter III, specifically endogenous biological rhythms, and identify which biological cycles contribute to the ecology of disease. I analyzed the ways in which circadian (i.e., daily) and circannual (i.e., seasonal) rhythms could influence disease transmission and affect within-host dynamics of infection and pathology. The literature on rhythms in the immune system is rapidly expanding. Circadian rhythms have now been implicated in within-host infection dynamics [12] and circannual rhythms in pathogen killing ability have been documented in wild species [13]. Using a transmission model and data from Siberian stonechats, a migratory bird with strong seasonal life history, I illustrated that seasonality in reproduction, migration, and immunity can interact to generate seasonality in acute infections. Interestingly, my model demonstrated that multiple non-sinusoidal out-of-phase seasonal drivers that differ in their epidemiological link to infection dynamics could result in "simple" patterns of seasonal disease incidence. This led me to conclude that the complicated seasonal structure of host behavior and

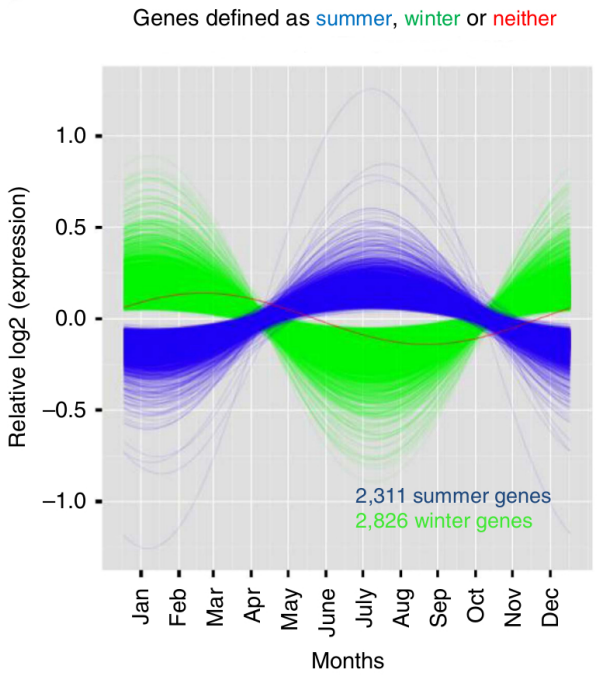


Figure 1.3: **Seasonal Gene Expression in the Human Immune System**. Over 5000 genes ($\sim$ 23%) of the human immune system were identified as seasonal. Plot shows fitted cosine curves for seasonal genes. Seasonal genes were binned into groups of summertime-expressed and wintertime-expressed genes, with the exception of a single gene that didn't fall into either category. Figure reproduced from [14], licensed under a Creative Commons Attribution 4.0 International License.

life history can constrain parasite transmission to particular windows in the annual cycle, which I refer to as the construction of a parasite temporal niche.

Wild species display seasonal life history and circadian clocks occur in all eukaryotes (including eukaryote parasites) and some prokaryotes [15, 16]. The ubiquity of biological rhythms, and the unsolved mystery of rhythmicity in infections, makes the study of biological rhythms an exciting area in disease ecology. In fact, coincident with the publication of my work on biological rhythms, it was discovered that the human immune system displays functional seasonal variation in gene expression [14]. I anticipate as circannual rhythms in humans and other animals continue to be characterized, it will be discovered that biological rhythms are important drivers

of infectious disease periodicity. S.F. Dowell first presented this hypothesis in 2001 [17], and the evidence that has accumulated over the past 14 years supports his idea. In addition to the role of biological rhythms in disease ecology, biological clocks are ancient in origin. In Chapter III, I propose ways in which host and parasite biological rhythms could impose selective pressure upon one another and affect the evolutionary trajectory of rhythms. In order to build scaffolding for future studies of rhythms and infection, I have developed both evolutionary and ecological models addressing the influence of biological rhythms on host-parasite interactions.

1.3 The Ecology of Polio

There are two avenues of inquiry for addressing seasonality in infectious disease systems. First, if seasonal drivers are known, they can be incorporated into transmission models or manipulated in the lab to quantify their effect on disease. Second, if seasonal drivers are unknown, transmission models can be fit to disease data to quantify seasonal epidemiological parameters, but the mechanism(s) will remain unidentified. Quantification of seasonal epidemiological parameters can provide insight into the ecology of infection and disease periodicity, even if seasonal mechanisms are obscure. My work on birth seasonality and biological rhythms emerged from the first avenue of inquiry. I identified birth seasonality and biological rhythms as having the potential to drive disease seasonality. I then incorporated them into transmission models to test whether they were viable mechanisms. In contrast, I followed the second avenue to study the transmission ecology of poliovirus. Polio has a long list of potential seasonal drivers, I therefore investigated the implications of seasonal transmission without knowledge of the mechanism.

Using extensive historical data, I confirmed earlier studies that described polio as a summertime disease [18]. Polio epidemics occur during summer and autumn. I demonstrated that, not only do polio epidemics occur in a fixed seasonal window,

but there is also latitudinal variation in polio's seasonal timing. As shown in Fig 1.4, polio's latitudinal variation was first documented in the 1932 [19], but it was unknown until now whether the latitudinal gradient was a persistent feature of polio's ecology, which it was. The latitudinal gradient suggests that polio seasonality is driven by a mechanism that also varies with latitude. Unfortunately, since many environmental and host factors vary with latitude, I was unable to identify a seasonal driver. However, I characterized the underlying transmission seasonality of poliovirus and the ecological implications of transmission seasonality.

I found that the polio transmission rate undergoes large seasonal shifts. In some regions, seasonal variation in transmission was so extreme that it caused poliovirus extinction during the wintertime low season. Importantly, despite local extinction, due to regional variation in host demography and transmission seasonality, poliovirus was able to persist via sources-sink dynamics. Transmission seasonality, which is a determinant of the reproductive number, is a key parameter for designing vaccination campaigns because it influences the vaccine coverage needed for pathogen extinction [20]. Unfortunately, seasonal fluctuations in transmission are often ignored in campaigns against polio and other vaccine-preventable diseases. Exceptions to this have proven useful; for instance, poliovirus was eradicated in areas of Mexico with relatively low, $<82\%$, oral polio vaccine coverage by launching vaccination campaigns during the low season [21]. Armed with the quantification of seasonal transmission, eradication campaigns could drive pathogens to extinction with relatively low vaccine coverage (low relative to the coverage needed to achieve the same result outside the low season).

With new knowledge of polio transmission, I was able to study the practical implications of transmission seasonality for polio eradication. I collected historical data from the US and the USSR that represented the transitional period during which poliovirus went from being epidemic to rare, and eventually extinct. Using these data,

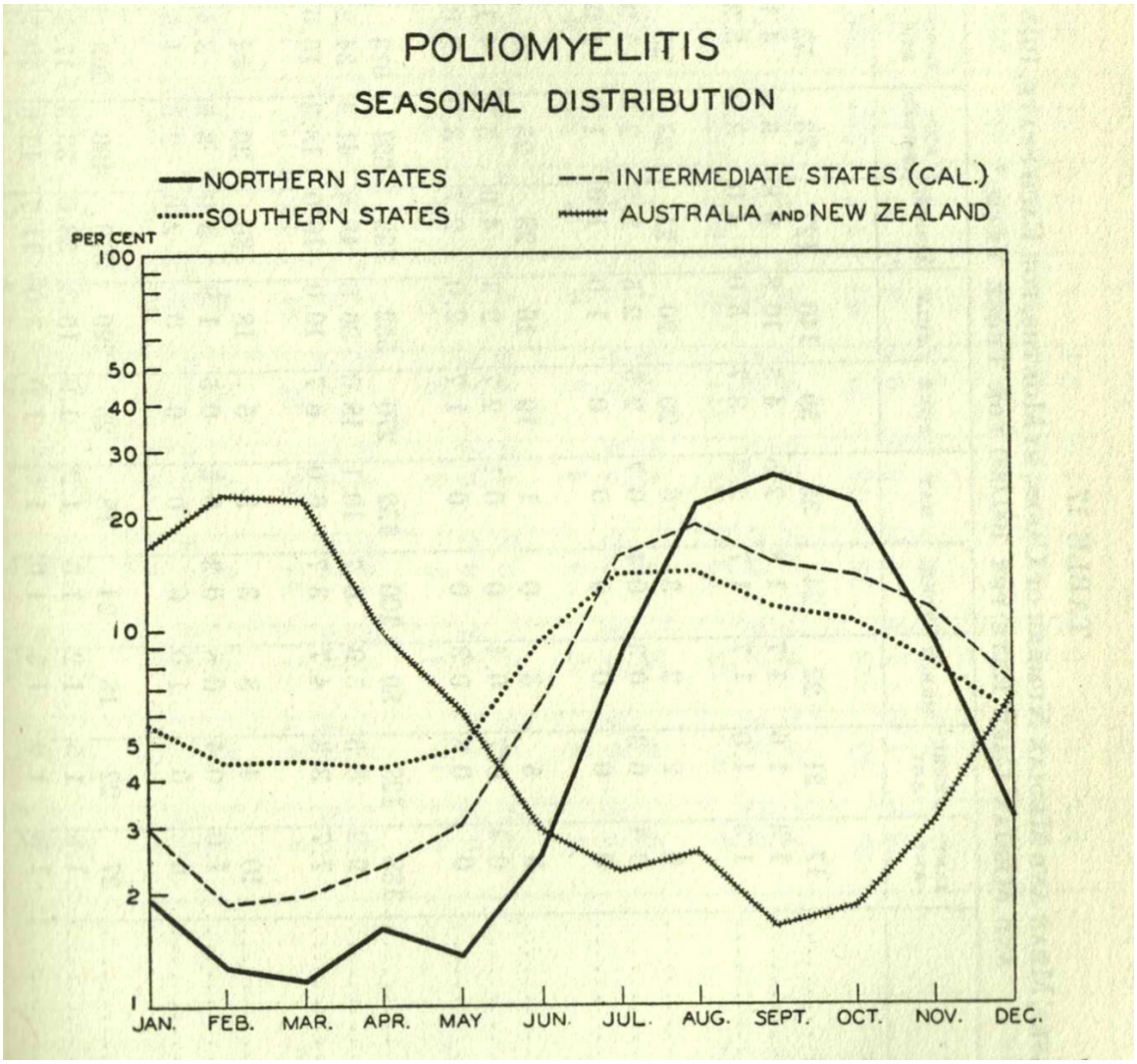


Figure 1.4: **Summertime Polio Epidemics.** The distribution of polio cases in the US, Australia, and New Zealand. Polio epidemics peaked in the summer: Feb–Mar in Australia and New Zealand, and Jun–Oct in the US. US states displayed south-to-north variation in epidemic timing, with southern states having the earliest epidemics, followed by intermediate latitude states. Northern states had the latest epidemics. Figure reproduced from [19].

I estimated the efficacy of the inactive polio vaccine (IPV) and the oral polio vaccine (OPV). I confirmed that OPV is more effective than IPV at reducing poliovirus transmission; a conclusion in line with clinical trials and the experience of the global polio eradication initiative. My key finding was, although IPV is inferior to OPV, IPV effectively reduces poliovirus transmission. I estimated that IPV reduces infectiousness by 69%. This result is a large step forward, not only for polio eradication, but also for the evaluation of vaccines. Vaccine trials are able measure the extent to which vaccines reduce susceptibility to infection. However, trials are logistically incapable of measuring vaccine-induced reductions in infectiousness. For most diseases, it is impractical to trace all secondary infections arising from vaccinated and unvaccinated individuals, which is required to measure realized relative infectiousness.

IPV's 69% reduction in infectiousness is hopeful indication that the future withdrawal of OPV and introduction of IPV is a practical eradication strategy. Although IPV is effective at reducing transmission, my work suggests IPV can only be used to eradicate polio in regions (and in seasons) when the reproductive number is $\leq$ 3. This would require $\geq$ 96% IPV coverage for populations that do not have pre-existing OPV- or naturally-derived immunity. Based on my work on polio in the US and the USSR, polio's reproductive number falls below 3 in many populations during the low season. If this is the case for today's endemic countries, IPV could be used to eradicate polio.

As a collection, my work demonstrates that seasonal ecology is undeniably important for the control and eradication of diseases, including polio. I have exclusively focused on the seasonality of acute infectious diseases. The seasonal factors that influence chronic infections largely overlap with those of acute infections, including birth seasonality and environmental conditions [22]; however, the seasonality of some chronic infections, such as nematode infections, is complicated by environmental life stages of the parasites [23]. In future work, I hope to apply the knowledge gained

in my dissertation to a broader class of infectious diseases and explicitly account for rhythms in host immunity. This will require cross-disciplinary integration of data and experiments from disease ecology, chronobiology (i.e., the study of biological rhythms), immunology, and parasitology.

Bibliography

[1] Soper HE (1929) The Interpretation of Periodicity in Disease Prevalence. Journal of the Royal Statistical Society 92: 34–73.

[2] Altizer S, Dobson A, Hosseini P, Hudson P, Pascual M, et al. (2006) Seasonality and the Dynamics of Infectious Diseases. Ecology Letters 9: 467–484.

[3] Kamo M, Sasaki A (2005) Evolution Toward Multi-year Periodicity in Epidemics. Ecology Letters 8: 378–385.

[4] Hawking F (1975) Circadian and Other Rhythms of Parasites. Advances in Parasitology 13: 123–182.

[5] London WP, Yorke JA (1973) Recurrent Outbreaks of Measles, Chickenpox and Mumps: I. Seasonal Variation in Contact Rates. American Journal of Epidemiology 98: 453–468.

[6] Foster RG, Roenneberg T (2008) Human Responses to the Geophysical Daily, Annual and Lunar Cycles. Current Biology 18: 784–794.

[7] Lincoln GA, Short R (1980) Seasonal Breeding: Natures Contraceptive. Recent Progress in Hormone Research 36: 1–52.

[8] Lam DA, Miron JA (1994) Global Patterns of Seasonal Variation in Human Fertility. Annals of the New York Academy of Sciences 709: 9–28.

[9] James WH (1990) Seasonal Variation in Human Births. Journal of Biosocial Science .

[10] Becker S (1991) Seasonal Patterns of Births and Conception Throughout the World. Advances in Experimental Medicine and Biology 286: 59–72.

[11] Moos WS, Randall W (1995) Patterns of Human Reproduction and Geographic Latitude. International Journal of Biometeorology 38: 84–88.

[12] Bellet MM, Deriu E, Liu JZ, Grimaldi B, Blaschitz C, et al. (2013) Circadian Clock Regulates the Host Response to Salmonella. Proceedings of the National Academy of Sciences of the United States of America 110: 9897–9902.

[13] Versteegh M, Helm B, Kleynhans E, Gwinner E, Tieleman B (2014) Genetic and Phenotypically Flexible Components of Seasonal Variation in Immune Function. The Journal of Experimental Biology 217: 1510–1518.

[14] Dopico XC, Evangelou M, Ferreira RC, Guo H, Pekalski ML, et al. (2015) Widespread Seasonal Gene Expression Reveals Annual Differences in Human Immunity and Physiology. Nature Communications 6: 1–13.

[15] Foster RG, Kreitzman L (2004) Rhythms of Life: The Biological Clocks that Control the Daily Lives of Every Living Thing. London: Profile Books Ltd.

[16] Koukkari WL, Sothern RB (2006) Introducing Biological Rhythms. New York, NY: Springer Science, 92–102 pp.

[17] Dowell SF (2001) Seasonal Variation in Host Susceptibility and Cycles of Certain Infectious Diseases. Emerging Infectious Diseases 7: 369–374.

[18] Sabin AB (1947) The Epidemiology of Poliomyelitis. Journal of the American Medical Association 134: 749–756.

[19] Park W, Blake J, Martin C, Ledingham J, Jordan E, et al. (1932) Chapter VII: Epidemiology. In: Poliomyelitis: A survey made possible by a grant from the Internatinal Committee for the Study of Infantile Paralysis, Baltimore, MD, USA: The Williams & Wilkins Company. pp. 306–478.

[20] Anderson R, May R (1991) Infectious Diseases of Humans: Dynamics and Control. Oxford University Press.

[21] Hull HF, Ward NA, Hull BP, Milstien JB, De Quadros C (1994) Paralytic Poliomyelitis: Seasoned Strategies, Disappearing Disease. Lancet 343: 1331–1337.

[22] Cattadori IM, Boag B, Bjørnstad ON, Cornell SJ, Hudson PJ (2005) Peak Shift and Epidemiology in a Seasonal Host-Nematode System. Proceedings of the Royal Society B: Biological Sciences 272: 1163–1169.

[23] Cornell SJ, Bjørnstad ON, Cattadori IM, Boag B, Hudson PJ (2008) Seasonality, Cohort-Dependence and the Development of Immunity in a Natural Host-Nematode System. Proceedings of the Royal Society B: Biological Sciences 275: 511–518.

CHAPTER II

Human Birth Seasonality: Latitudinal Gradient and Interplay with Childhood Disease Dynamics

Preamble. This chapter is our published manuscript provided verbatim. The citation for this manuscript is:

***Martinez-Bakker, M.**, Bakker, K. M., King, A. A., & Rohani, P. (2014). Human Birth Seasonality: Latitudinal Gradient and Interplay with Childhood Disease Dynamics. Proceedings of the Royal Society B: Biological Sciences, 281(1783), 20132438.*

My (Micaela Martinez-Bakker) contribution to the work was the integration of human birth seasonality into the models of childhood disease. I also worked with my co-authors to design the study and write the manuscript. Specifically, I conducted all the analyses that went into Fig 2.4–2.6 and Appendix A Figures A.7 and A.8. The collection, digitization, and analyses of human birth data were conducted by Kevin M. Bakker (co-author).

Abstract. More than a century of ecological studies have demonstrated the importance of demography in shaping spatial and temporal variation in population dynamics. Surprisingly, the impact of seasonal recruitment on infectious disease systems has received much less attention. Here, we present data encompassing 78 years of monthly natality in the United States, and reveal pronounced seasonality in birth rates, with

geographic and temporal variation in both the peak birth timing and amplitude. The timing of annual birth pulses followed a latitudinal gradient, with northern states exhibiting spring/summer peaks and southern states exhibiting fall peaks, a pattern we also observed throughout the northern hemisphere. Additionally, the amplitude of US birth seasonality was more than two-fold greater in southern states versus those in the north. Next, we examined the dynamical impact of birth seasonality on childhood disease incidence using a mechanistic model of measles. Birth seasonality was found to have the potential to alter the magnitude and periodicity of epidemics, with the effect dependent on both birth peak timing and amplitude. In a simulation study, we fitted an SEIR model to simulated data, and demonstrated that ignoring birth seasonality can bias the estimation of critical epidemiological parameters. Finally, we carried out statistical inference using historical measles incidence data from New York City. Our analyses did not identify the predicted systematic biases in parameter estimates. This may be due to the well-known frequency-locking between measles epidemics and seasonal transmission rates or may arise from substantial uncertainty in multiple model parameters and estimation stochasticity.

The ubiquity of seasonal variation in the incidence of infectious diseases has driven much epidemiological research focused on understanding the responsible underlying mechanisms [1, 2, 3, 4, 5]. Surprisingly, there remains much uncertainty regarding the drivers of seasonal incidence for numerous infections including polio, pertussis, scarlet fever, diphtheria, rotavirus, among others [5, 6, 7, 8]. Early work on diphtheria and measles implicated elevated contact rates among children in school as the driver of pulsed transmission [9, 1], leading to much emphasis on school-term forcing [2, 3, 5, 10, 11]. More recently, additional mechanisms of seasonal transmission have been identified, including climatic drivers of pathogen survival [12], transmission [13, 14], and vector activity [15, 16], seasonal host migration [17], and seasonal fluctuations

in host immunity [18, 19]. Here, we propose that seasonality in host recruitment rates may also shape epidemiology. This is a possiblity that has been appreciated in studies of wildelife diseases [20, 21, 22, 11, 23]. For instance, studying cowpox virus in voles, Begon *et al.* found that susceptible recruitment is seasonal, and higher breeding-season birth rates delayed epidemic peaks [24]. However, despite evidence demonstrating the importance of host demography in recurrent epidemics [25, 26, 27, 28], and the ubiquitous appreciation of seasonal reproduction in broader ecology and evolution [29], we submit that a deep understanding of the dynamical impact of birth seasonality on infectious diseases of humans is currently lacking.

To explore this phenomenon, we first characterize the landscape of birth seasonality in modern human populations, and second determine if/how it can impact epidemic dynamics, particularly for immunizing childhood infections. Some precedent has been set in the field of demography, with seasonal variation in human births first documented in the early 1800s [30, 31] and currently recognized as a global characteristic of humans [32, 33, 34, 35, 36]. Early studies of vital statistics in various US regions established a national-level seasonal pattern of births with troughs in the spring and peaks in autumn [35, 37]. Subsequent research has focused on either a few locations over long time periods, or many locations over short time periods. Collectively, these studies showed that northern and southern states have differences in their seasonal birth amplitude [32, 33, 34, 38, 35, 37] and birth/conception minima [38, 39, 32, 40]. Studies of births in Africa and Asia have been sparse, but seasonal peak-trough differences in conception ranging from 11–64% have been documented in Africa and 8–58% in Asia [36]. To date there has been no long term, large scale, spatiotemporal analysis of births in either the US or worldwide.

We have compiled the most extensive spatiotemporal data set on human births to date, and explored the effect of birth seasonality on childhood disease incidence using

simulated and empirical data. Measles was chosen because it is the paradigmatic example of a childhood disease, with two key features: (i) a low mean age of infection during the pre-vaccine era, with infections occurring in the youngest age group, the size of which is tightly linked to the birth rate, and (ii) seasonal transmission, which is a feature of many childhood diseases. We focus on birth seasonality in the presence of seasonal transmission to explore the interplay between these two forces. Our novel demographic data set is comprised of birth records across the globe, consisting of 7.3 $\times$ 10^8 births. Specifically, these data consist of monthly births spanning a 78-year period (1931–2008) for each state in the continental US along with over 200 additional time series from countries spanning the Northern Hemisphere. We have analyzed these data in combination with a transmission model and statistical inference tools to examine the dynamic implications of birth seasonality on childhood infection.

Data

Monthly state-level time series of live births from 1931–2008 were downloaded from US Vital Statistics [41] and digitized. Annual state-level population size data were collected from the US Census Bureau [42] and used to construct monthly time series of birth rates per 1000 individuals per month. Worldwide monthly births were retrieved from the United Nations database [43] and filtered for countries containing at least 5 years of consecutive data.

The US data were split into three eras, to account for the baby boom: (i) Jan 1931–Dec 1945, which we term the Pre-Baby Boom Era, (ii) Jan 1946–Dec 1964, the Baby Boom Era, and (iii) Jan 1965–Dec 2008, the Modern Era. To test for periodicity, a wavelet spectral analysis [44] was performed independently for each US state in each era and for each country in the global data set. The significance of each period was tested by comparing the power of each period against a noise background, using a

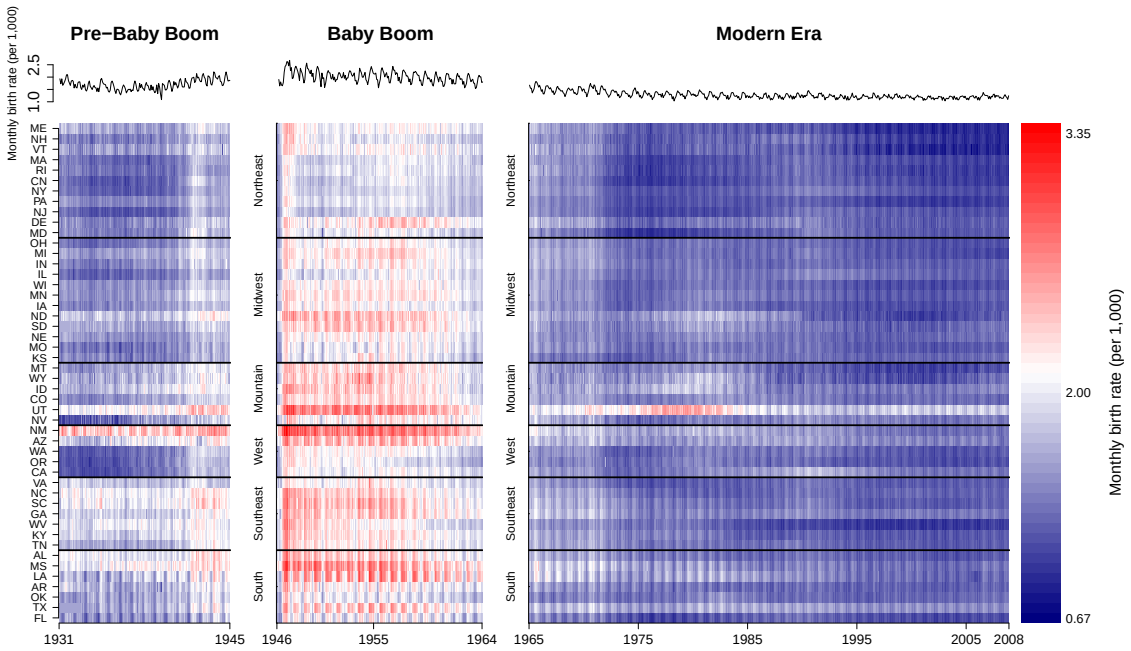


Figure 2.1: **Birth Data from the US**. Temporal patterns of birth rates (per 1000 individuals per month) in the US organized by geographic region, separated into three eras: Pre-Baby Boom (1931–1945), Baby Boom (1946–1965), and Modern Era (1965–2008). The time series for Louisiana is plotted at the top as an example.

lag–1 autocorrelation test. For each data series significant at a 1 year period, phase angle time series were constructed to determine the timing of birth peaks occurring at 11–13 month intervals. Independently, seasonal decomposition was run on all data series to filter out noise, and the seasonal amplitude was calculated by taking one-half the difference between the maxima and minima, measured as a percent of the annual mean (refer to Equations in Appendix A). Inter-annual variation was examined by analyzing the percent change in mean birth rates from one year to the next.

Measles models

We used a discrete-time SEIR model of measles adapted from Earn *et al.* [25], with school-term forcing based on the England & Wales school year (Table S1; [11]).

We incorporated seasonality in births using a sine function with varying amplitude and phase. The equations describing the model and parameter values are provided in Appendix A.

We conducted statistical inference on both simulated and empirical measles incidence data to test for the effect of birth seasonality on epidemic dynamics and parameter estimation. This work aimed to answer the question: how does the omission of birth seasonality affect the precision and bias of estimated parameters? Using a Markovian analogue of our SEIR model (Equations in Appendix A), three time series were generated assuming the following parameterizations: a birth peak day of either 162, 295, or 351 and a 28% birth amplitude (see Table S3 for parameter values). For each simulated time series, our stochastic SEIR model was fit assuming constant births (birth amplitude set to 0%) and an unknown mean transmission rate. All other parameters were assumed known. Thus, the only free parameter was the mean transmission rate, which is directly proportional to R_0. The transmission rate was profiled and the likelihood was calculated using a particle filter (Appendix A Materials & Methods) [45].

In order to test for bias in parameter estimation using real world data, we utilized historical measles case reports from New York City. These data are from the Baby Boom Era, when the birth amplitude was low, approximately 7% for the state of New York. To account for maternal antibodies, we fit models which lagged births 3, 6, or 9 months (see Appendix A for methods). We used both maximum likelihood via iterated filtering [45] and the TSIR methodology of Finkenstadt & Grenfell [46, 47] to quantify the impact of seasonal births on parameter estimates (Fig 5).

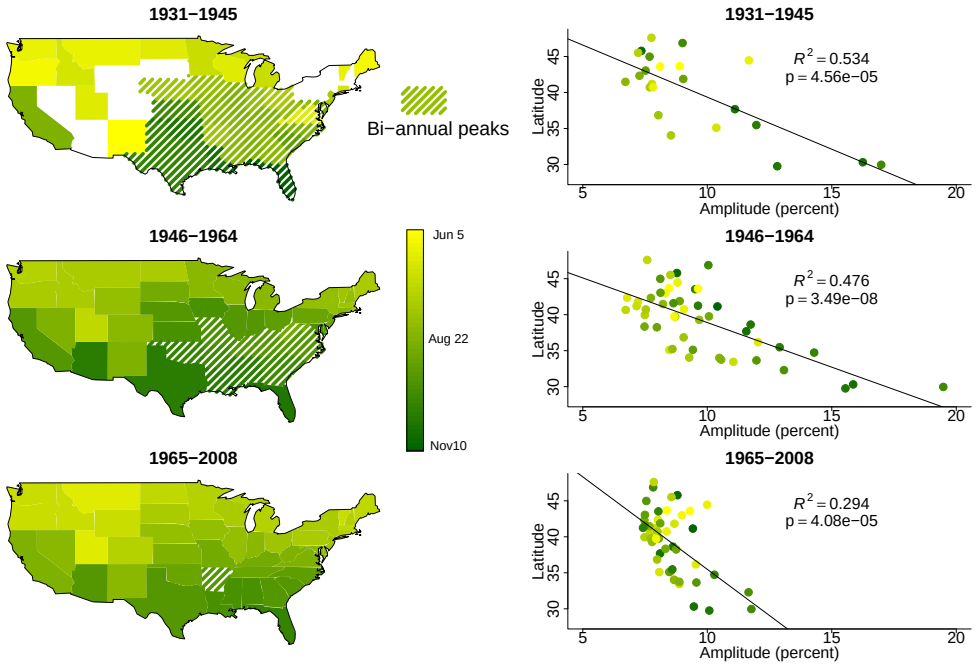


Figure 2.2: **Latitudinal Gradients in US Births**. Spatiotemporal patterns of seasonal birth peak timing and amplitude in the US. (Top panels) Pre-Baby Boom (1931–1945), (Middle panels) Baby Boom (1946–1964), and (Bottom panels) Modern Era (1965–2008). Maps depict the latitudinal gradient in the timing of the birth peak. Colors indicate the mean timing of the birth peak for each state. Hatched regions represent states in which a significant bi-annual peaks were discernible and are color coded based on the timing of their primary annual birth pulse (also see Figs S1 & S2). States shown in white did not exhibit a significant periodicity. Regressions show the latitudinal variation in seasonal amplitude, with the colors representing the peak birth timing for the respective period.

The seasonal timing of births

Figs 1 and S3 provide an overview of birth rates in the US. Most states had significant seasonal (1 year) birth pulses in the Pre-Baby Boom Era, while all states showed significant birth seasonality in the Baby Boom and Modern Eras. Of the 210 time series analyzed outside the US, 132 (63%) had significant birth seasonality. Most of the locations for which seasonality was not significant were short time series (5–7 years) or countries with less than 100 births per month.

We observed a latitudinal gradient in the timing of the birth peak across the US; this same gradient was observed throughout the Northern Hemisphere (Figs 2 & 3). In general, the birth peak occurs earlier in the year in locations further from the equator. For example, in the Pre-Baby Boom Era the birth peak occurred as early as June in the northern states of Oregon and Maine, whereas the peak occurred as late as November in Florida. The variation in birth peak timing was largest during the Pre-Baby Boom Era, when the most out-of-phase states differed by more than 5 months.

During the Baby Boom Era, most states had birth peaks that occurred in August or later. The only peaks which occurred prior to August were in seven northern states and this pattern continued during the Modern Era. The earliest birth peaks always occurred in northern states, followed by mid-latitude states, and the latest peaks occurred in southern states. Across all eras, the latest peak was consistently in Florida, where the peak timing ranged from early October in the Modern Era, to early November in the Pre-Baby Boom Era.

The latitudinal gradient in peak birth timing seen in the US was reflective of a worldwide pattern. The worldwide timing also followed a latitudinal gradient with birth peaks occurring earlier at higher latitudes and later for countries closer to the equator (Fig 3). However, at any given latitude there was a large amount of variation in the timing of the birth peak. In the highest latitude countries ($> 50^\circ$ N), birth peaks occurred between April and July. While there were two outlying mid-latitude countries with birth peaks in March and April (Italy, 1970–1985 and Tajikstan, 1989–1994), typical mid-latitude locations (20–50° N) had peaks between May and November. Countries in the vicinity of the equator (0–20° N) displayed the least amount of variation in timing. The equatorial countries, such as those in the Caribbean, consistently had birth peaks between September and November, with the latest birth peak

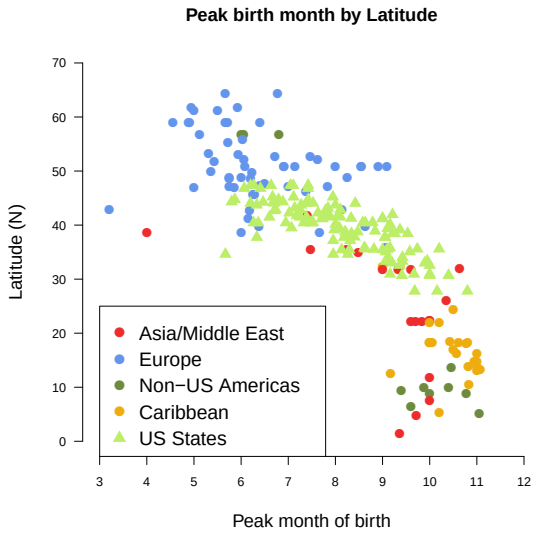


Figure 2.3: **Northern Hemisphere Latitudinal Gradient in Birth Seasonality**. Northern Hemisphere patterns of seasonal birth pulses color coded by region. Birth pulses occurred earlier in the year at northern latitudes. Table S5 provides the details for each country, including the time frame of the data which ranges from the 1960s to 2011.

occurring in Saint Vincent and the Grenadines during the period 1992–2005.

In addition to the annual birth peaks, in the US a significant bi-annual (6 month) birth pulse was detected in 24 states. In the Pre-Baby Boom Era, all states with bi-annual periodicity were clustered in the lower midwest, deep south, and southeast (Figs 2, S1, & S2). In the Baby Boom Era, only 13 states continued to exhibit a bi-annual period. Arkansas is the only state where this bi-annual birth pulse persists in the Modern Era.

The amplitude of seasonal births

Birth amplitude was measured for each time series, each year, as the one-half peak-trough difference with noise removed. Amplitudes are represented as a percent of the mean annual birth rate (Appendix A Materials & Methods). As with the seasonal

timing, in the US the amplitude of birth seasonality displays a latitudinal gradient. Fig 2 depicts the negative relationship between birth amplitude and latitude. We found that 29–53% of the variation in birth amplitude can be explained by latitude ($p < 4.6e\text{-}5$). However, the amount of variation in birth amplitude explained by latitude decreased through time, perhaps due to the decline in birth amplitude throughout southern states during the Modern Era (Figs 2 & S4). We did not observe a latitudinal gradient in birth amplitude outside of the US (Fig S5).

As expected, the mean combined amplitude across all states was found to be comparable to the national level amplitude reported in the literature [48] and was 9.0%, 9.8%, and 8.5% for the Pre-Baby Boom, Baby Boom, and Modern Eras, respectively. Interestingly, due to the geographical variation in birth peak timing, state-level births are out of phase. Thus, aggregated US birth data has a deceptively low amplitude that is not reflective of individual states. Birth amplitudes $> 15\%$ were observed in many southern states throughout the time series (Fig S4).

Prior work has shown that the levels of inter-annual variation in births observed in the US can have a dynamical impact on disease incidence [25]. It would follow logically that variability of this same magnitude over a shorter time period may also be important. Thus, we sought to compare the magnitude of the seasonal variation in births with inter-annual variation. Inter-annual variation was measured for each state as the percent change in mean birth rate from one year to the next. We found that in almost every instance, seasonal variation exceeded inter-annual variation, with seasonal variation in the Modern Era being 2–3 times larger than the variation from year to year (Fig S6).

The effects of birth seasonality on epidemic dynamics

We investigated the impact of birth seasonality on epidemics of childhood disease by employing models of measles transmission. As shown in Fig 4A, birth seasonality can have the effect of amplifying or dampening incidence during epidemic years. Crucially, the impact of birth seasonality depends on the amplitude and phase relationship between susceptible recruitment and transmission seasonality. In our simulation study we did not account for maternal antibodies, thus the peak in susceptible recruitment was equivalent to the birth peak. However, inclusion of maternal antibodies would translate into a lag between the peak in births and the peak in susceptible recruitment. We found that if the peak in susceptible recruitment occurs at the beginning of the year, when children are in school and measles transmission is elevated, the epidemic is amplified due to the availability of susceptibles. In contrast, if the peak in susceptible recruitment occurs at the end of the school year, when children are entering summer break, the epidemic is dampened (Fig 4A).

Independent of the timing of the birth peak, the effect of birth seasonality on measles epidemics depends on the birth amplitude (Figs 4B & S7). The larger the birth amplitude the greater the change in measles incidence. Not only does the amplitude affect incidence, but birth rates with high amplitude fluctuations ($> 40\%$) can alter incidence to such an extent that they can drive dynamical transitions (Fig 4B).

Statistical inference on simulated data led to small biases in the estimate of R_0 for measles (Fig 4C). For the time series in which the birth peak occurred in mid-December, day 351, a time at which susceptible recruitment increases epidemic year incidence, omitting birth seasonality resulted in over-estimating R_0 in order to capture the elevated epidemic year incidence. In contrast, when the birth peak was set to either early June (day 162) or late October (day 295), times at which susceptible

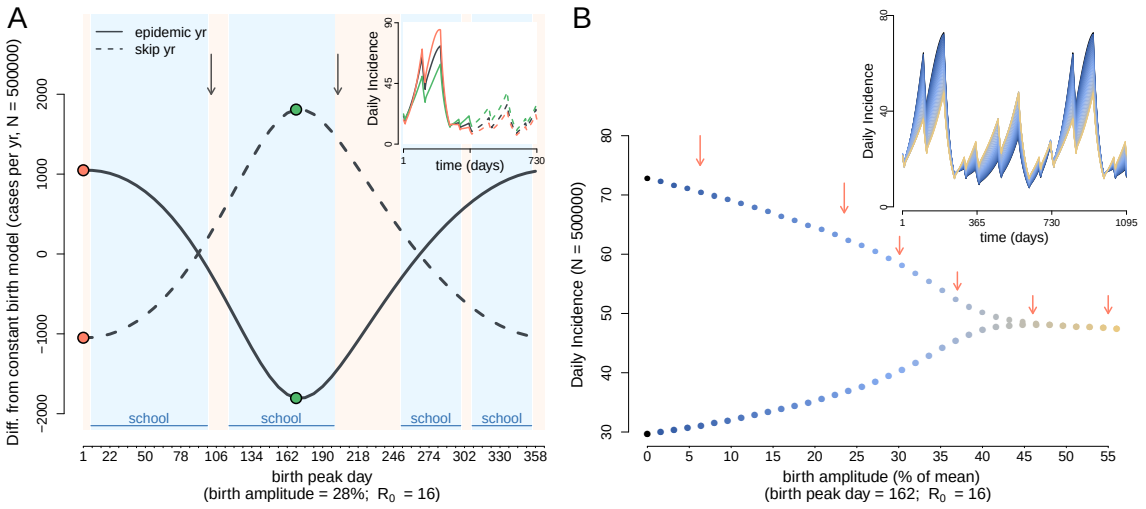


Figure 2.4: **Impact of Birth Seasonality on Childhood Disease**. (A) Epidemic and skip-year incidence varies with birth peak timing along the x-axis. Solid curve shows the change in epidemic year incidence when birth seasonality is added to the measles model. Dashed curve shows the change in the skip-year. The phase relationship between seasonal births and transmission determines whether birth seasonality has an effect on incidence. The greatest increase in epidemic year incidence is when the birth peak occurs after children return from winter holiday (orange points). A decrease in epidemic year incidence occurs when births peak prior to summer vacation (green points). School terms are noted and vertical arrows mark the timing of incidence peaks during the epidemic year. (Inset) Time series from the constant birth model (black), and time series corresponding to the color-matched points on the main graph. (B) Bifurcation diagram showing the change in epidemic and skip-year peak incidence with increasing birth amplitude. In the absence of birth seasonality, epidemics are biennial. As birth amplitude increases, skip-year incidence increases and epidemic year incidence decreases. When birth amplitude reaches ∼40% epidemics become annual. Time series in the inset correspond to the points in the main graph; blue time series are biennial, and golden are annual. Arrows denote the birth amplitude observed in Switzerland, Cuba, Egypt, Nigeria, Guinea, and Sierra Leone, left to right. Amplitudes for Nigeria, Guinea, and Sierra Leone are from [49].

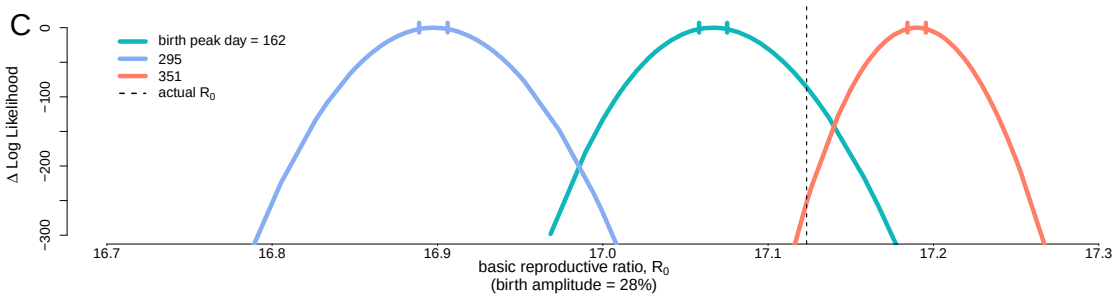


Figure 2.5: **Bias in R_0 estimates due to the exclusion of birth seasonality in a SEIR model.** Time series were generated using a SEIR model with a birth peak day of 162 (turquoise), 295 (blue), or 351 (orange). Each time series was fit to the SEIR model with a birth amplitude of 0%. The actual R_0 value is shown by the dashed line and the likelihood profiles show that maximum likelihood estimate of R_0 is either over- or under-estimated when birth seasonality is excluded from the model. 95% confidence intervals for MLE are indicated on profiles. Note, in the published form of this manuscript, this is panel C, and is combined with the two panels from the previous figure.

recruitment dampens epidemic year incidence and elevates skip-year incidence, we under-estimated R_0. However the bias in R_0 was small, approximately $\leq$0.4–1.3%.

We found that models with seasonal births effectively capture measles dynamics in New York City (Fig 5A). In contrast to our simulation study, however, when multiple unknown parameters were estimated simultaneously, the small predicted bias in R_0 was masked by uncertainty in parameters and Monte Carlo error (Fig 5). Hence, the maximum likelihood parameter estimates (MLEs) for models with and without birth seasonality were nearly identical. The MLEs of the basic reproductive number, R_0, ranged from 19.3–20.3. Thus, the incorporation of birth seasonality into the model did not substantially change parameter estimates, and the dynamics of baby-boom era measles in New York City can be captured by the model without birth seasonality (Fig 5A).

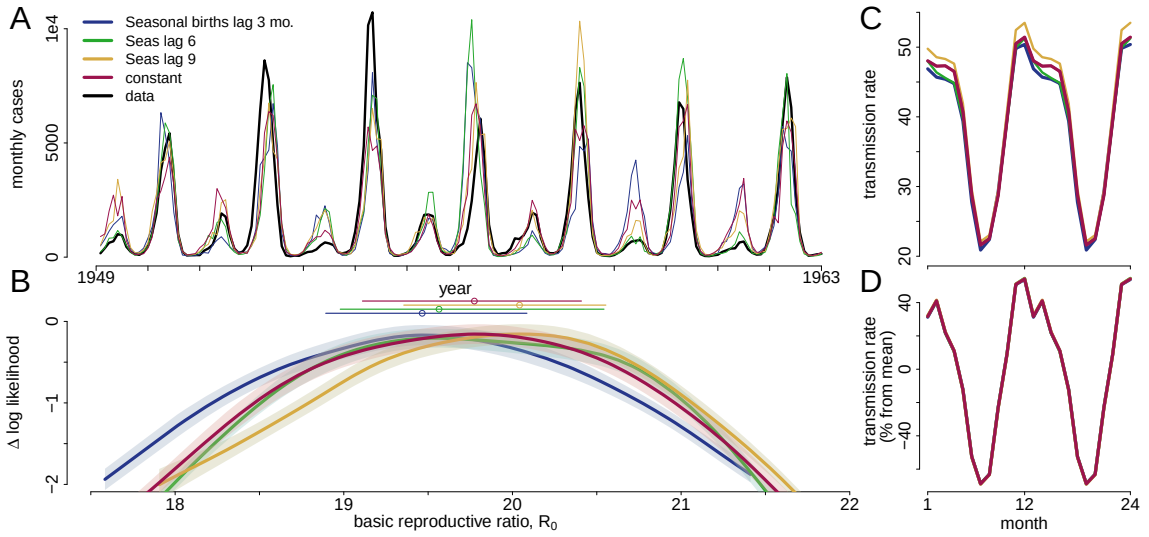


Figure 2.6: **Measles Cases in New York City**. (A) Measles incidence (black) and a stochastic realization using the MLE for each type of birth covariate: seasonal births with a 3 month lag (blue), seasonal births with a 6 month lag (green), seasonal births with a 9 month lag (yellow), and births with no seasonality (maroon). Legend applies to all of Fig 5. (B) The shape of the likelihood surface with respect to R_0. The MLE R_0s are indicated by points and the standard errors are represented by horizontal lines. (C) MLE transmission splines for each model. (D) Transmission splines estimated using TSIR [46, 47] for each type of birth covariate. The MLEs differed with and without birth seasonality, but the differences in the point estimates were overwhelmed by uncertainty in parameter estimates (Figs 5B & 5C). No difference in transmission parameters was observed using the TSIR method.

Discussion

Seasonal fluctuations in human births are observed throughout the world. The timing of the birth peak displayed a marked latitudinal gradient throughout the Northern Hemisphere. The latitudinal gradient in peak birth timing was observed in the US for the entirety of our data, and was reflective of a much broader geographic pattern. National level birth data from Asia, Europe, the Americas, and the Caribbean also exhibited this latitudinal gradient with birth peaks occurring months earlier at locations further from the equator.

Contemporary seasonal birth amplitudes are substantial with a range of 7–12% in the US and 6–35% in other Northern Hemispheric countries. Along with the latitudinal gradient in peak birth timing, in the US we also observed a latitudinal gradient in birth amplitude. States in the southern US have larger seasonal fluctuations in births than northern states. This negative relationship between latitude and amplitude was more pronounced in the Pre-Baby Boom and Baby Boom Eras, relative to the Modern Era. However, this pattern was not observed outside of the US, suggesting this may either be a localized phenomenon or strongly correlated with social, economic, and/or cultural factors in the US.

In addition to the striking geographical variation in timing and amplitude of the annual birth peak, these data displayed additional complexity with the occurrence of bi-annual peaks across the lower midwest, deep south, and southeastern US in the Pre-Baby Boom Era. This bi-annual pulse was lost over time, with only Arkansas exhibiting bi-annual periodicity in the Modern Era. Bi-annual fluctuations in births have been documented in previous studies [38], but our data suggests that bi-annual birth pulses in the US are a relic of the past, lost to societal changes [38, 39], yet may still exist in other countries. Given the robustness of birth seasonality as a global

phenomenon of contemporary human populations, it is surprising that mechanisms driving these patterns remain poorly understood. Demographers have implicated a host of social, environmental, and physiological factors that may interact to drive birth seasonality. While a consensus has yet to be reached, and mechanisms vary geographically, hypothesized drivers include income, culture, race, holidays, rainfall, cold winters, and seasonally variable sperm quality[50, 51, 52, 53, 54, 38, 55, 56]. Although we focused on characterizing the variation in birth seasonality, rather than the mechanisms underlying this variation, it is our hope that the latitudinal gradient in peak birth timing and amplitude observed here will help elucidate the primary drivers of birth seasonality.

Despite our high resolution birth data for the Northern Hemisphere, Southern Hemispheric data proved difficult to obtain. Our analysis focused solely on the US and countries where birth data were readily available. Unfortunately this leaves out many South American and African countries where vaccine preventable childhood diseases are most prevalent. Southern Hemispheric birth data may help us understand the variation observed in the seasonality of childhood infections. For instance, historical work in Africa has shown that measles incidence peaks in April in Uganda, Kenya, and Tanzania, but earlier (November–January) in their southern neighbors Zambia, Zimbabwe, and Malawi [57]. Knowing the seasonal birth peak timing and amplitude in these locations may allow us to better understand this variation. We anticipate the latitudinal gradient in peak birth timing will also be found in the Southern Hemisphere.

The impacts of birth seasonality on epidemic dynamics were explored here in the context of childhood diseases. Our theoretical predictions indicate birth seasonality has the potential to influence the dynamics of fully immunizing infections of childhood – for which susceptible recruitment most heavily relies on births [25, 11]. We

demonstrated that birth amplitude and the timing of the birth peak relative to peak transmission determine whether, and to what extent, birth seasonality affects disease incidence patterns. In our inference study using simulated data, we found that ignoring birth seasonality can bias parameter estimation. As a proof of concept study, we tested for these biases using New York City measles data from the pre-vaccine era. However, we did not detect any systematic biases. There may be a number of reasons for this finding. First, during the time span of these data, the seasonal birth amplitude was low in New York City. Second, the short infectious period of measles is known to lead to pronounced frequency-locking with forcing in transmission [7, 58, 11], which may swamp any dynamical impacts of weakly seasonal susceptible recruitment. Finally, the combination of process- and measurement-noise in the data, combined with uncertainty in parameter estimates and Monte Carlo error may have made it impossible to detect the predicted estimation bias.

Our simulation studies demonstrated that high amplitude birth seasonality, currently observed in many African and Asian countries (Table S5 and [36, 49]), can affect disease periodicity and epidemic magnitude. In these settings, our findings have the potential to explain some of the spatial and temporal variation observed in the periodicity of diseases such as measles, rotavirus, and polio; and present a promising avenue for future research. Indeed, a recent study of birth seasonality across developing countries found that the timing of the birth peak influences epidemic timing, and a high birth rate magnifies the effect of birth seasonality on measles epidemics [59]. Although our study—focused exclusively on measles epidemiology—suggests that high amplitude birth seasonality is required to alter disease incidence, we predict that lower birth amplitudes may have a dynamical effect when coupled with a higher mean birth rate or for childhood diseases with longer infectious periods that may exhibit less frequency-locking with seasonal transmission [7]. Ultimately, our

experience with these systems indicate that the impact of seasonal births on epidemiology will likely be determined by multiple factors, including: the age-distribution of infections, age-specific pattern of contacts, differences in R_0, and the demographic context.

Dynamical consequences of birth seasonality aside, we emphasize that the spatial variation in birth seasonality documented here is pertinent when developing time-specific vaccination campaigns. For example, the World Health Organization implements time-specific vaccination campaigns to supplement routine immunization for the control of measles and polio in Africa, the Eastern Mediterranean, and South-East Asia. Clearly, if these infant immunization campaigns occur prior to the birth pulse, they will be inefficient. Thus, it is our hope that future studies aimed at mitigating childhood diseases will utilize birth seasonality to reduce the burden of disease and tackle some of the unanswered questions in disease ecology.

Acknowledgements

We thank two anonymous reviewers, our editor Vincent Jansen, R. McDaniel, & the Rohani/King labs for their helpful comments on this work. MMB is supported by the NSF Graduate Research Fellowship Program. PR and AAK are supported by the Research and Policy in Infectious Disease Dynamics program of the Science and Technology Directorate, Department of Homeland Security, the Fogarty International Center, National Institutes of Health and by a research grant from the National Institutes of Health (R01AI101155).

Bibliography

[1] Soper, H. E., 1929 The Interpretation of Periodicity in Disease Prevalence. *Journal of the Royal Statistical Society* **92**, 34–73.

[2] London, W. P. & Yorke, J. A., 1973 Recurrent Outbreaks of Measles, Chickenpox and Mumps. I. Seasonal Variation in Contact Rates. *American Journal of Epidemiology* **98**, 453–468.

[3] Fine, P. E. & Clarkson, J. A., 1986 Seasonal Influences on Pertussis. *International Journal of Epidemiology* **15**, 237–247.

[4] Rohani, P., Keeling, M. J., Grenfell, B. T., 2002 The Interplay between Determinism and Stochasticity in Childhood Diseases. *The American Naturalist* **159**, 469-481.

[5] Metcalf, C. J. E., Bjørnstad, O. N., Grenfell, B. T. & Andreasen, V., 2009 Seasonality and Comparative Dynamics of Six Childhood Infections in Pre-vaccination Copenhagen. *Proceedings of the Royal Society B: Biological Sciences* **276**, 4111–4118. (doi:10.1098/rspb.2009.1058).

[6] Gear, H. S., 1949 The Virus of Poliomyelitis: Its Distribution and Methods of Spread. *The Journal of the Royal Society for the Promotion of Health* **69**, 149–153. (doi:10.1177/146642404906900301).

[7] Rohani, P., Earn, D. J., Grenfell, B. T., 1999 Opposite Patterns of Synchrony in Sympatric Disease Metapopulations. *Science* **286**, 968-971.

[8] Török, T., Kilgore, P., Clarke, M., Holman, R., Bresee, J. & Glass, R., 1997 Visualizing Geographic and Temporal Trends in Rotavirus Activity in the United States , 1991 to 1996. *The Pediatric Infectious Disease Journal* **16**, 941–946.

[9] The Lancet, 1899 The Influence of Schools in Accentuating the Spread of Certain Infectious Diseases. *The Lancet* **153**, 184–185, 256, 330–331.

[10] Eames, K. T. D., Tilston, N. L., Brooks-Pollock, E. & Edmunds, W. J., 2012 Measured Dynamic Social Contact Patterns Explain the Spread of H1N1v Influenza. *PLoS Computational Biology* **8**, e1002425.

[11] Keeling, M. J., Rohani, P., 2008 Modeling Infectious Diseases in Humans and Animals. Princeton University Press.

[12] Hemmes, J. H., Winkler, K. C. & Kool, S. M., 1962 Virus Survival as a Seasonal Factor in Influenza and Poliomyelitis. *Antonie van Leeuwenhoek* **28**, 221–233. (doi:10.1007/BF02538737).

[13] Shaman, J., Pitzer, V. E., Viboud, C., Grenfell, B. T. & Lipsitch, M., 2010 Absolute Humidity and the Seasonal Onset of Influenza in the Continental United States. *PLoS Biology* **8**, e1000316. (doi:10.1371/journal.pbio.1000316).

[14] Lowen, A. C., Mubareka, S., Steel, J. & Palese, P., 2007 Influenza Virus Transmission is Dependent on Relative Humidity and Temperature. *PLoS Pathogens* **3**, 1470–1476. (doi:10.1371/journal.ppat.0030151).

[15] Mabaso, M. L. H., Craig, M., Ross, A. & Smith, T., 2007 Environmental Predictors of the Seasonality of Malaria Transmission in Africa: the Challenge. *The American Journal of Tropical Medicine and Hygiene* **76**, 33–38.

[16] Linthicum, K. J., Anyamba, A., Tucker, C. J., Kelley, P. W., Myers, M. F. & Peters, C. J., 1999 Climate and Satellite Indicators to Forecast Rift Valley Fever Epidemics in Kenya. *Science* **285**, 397–400. (doi:10.1126/science.285.5426.397).

[17] Bharti, N., Tatem, a. J., Ferrari, M. J., Grais, R. F., Djibo, A. & Grenfell, B. T., 2011 Explaining Seasonal Fluctuations of Measles in Niger using Nighttime Lights Imagery. *Science* **334**, 1424–1427. (doi:10.1126/science.1210554).

[18] Dowell, S. F., 2001 Seasonal Variation in Host Susceptibility and Cycles of Certain Infectious Diseases. *Emerging Infectious Diseases* **7**, 369–374. (doi: 10.3201/eid0703.010301).

[19] Cannell, J. J., Zasloff, M., Garland, C. F., Scragg, R. & Giovannucci, E., 2008 On the Epidemiology of Influenza. *Virology Journal* **5**, 1–12. (doi:10.1186/1743-422X-5-29).

[20] Bolzoni, L., Dobson, A. P., Gatto, M. & De Leo, G. A., 2008 Allometric Scaling and Seasonality in the Epidemics of Wildlife Diseases. *The American Naturalist* **172**, 818–828. (doi:10.1086/593000).

[21] Altizer, S., Hochachka, W. M. & Dhondt, A. A., 2004 Seasonal Dynamics of Mycoplasmal Conjunctivitis in Eastern North American House Finches. *Journal of Animal Ecology* **73**, 309–322. (doi:10.1111/j.0021-8790.2004.00807.x).

[22] Hosseini, P. R., Dhondt, A. A. & Dobson, A., 2004 Seasonality and Wildlife Disease: How Seasonal Birth, Aggregation and Variation in Immunity Affect the Dynamics of *Mycoplasma gallisepticum* in House Finches. *Proceedings of the Royal Society B: Biological Sciences* **271**, 2569–2577. (doi:10.1098/rspb.2004.2938).

[23] Duke-Sylvester, S. M., Bolzoni, L. & Real, L. A., 2011 Strong Seasonality Produces Spatial Asynchrony in the Outbreak of Infectious Diseases. *Journal of the Royal Society Interface* **8**, 817–825. (doi:10.1098/rsif.2010.0475).

[24] Begon, M., Telfer, S., Smith, M. J., Burthe, S., Paterson, S. & Lambin, X., 2009 Seasonal Host Dynamics Drive the Timing of Recurrent Epidemics in a Wildlife Population. *Proceedings of the Royal Society B: Biological Sciences* **276**, 1603–1610. (doi:10.1098/rspb.2008.1732).

[25] Earn, D. J., Rohani, P., Bolker, B. M. & Grenfell, B. T., 2000 A Simple Model for Complex Dynamical Transitions in Epidemics. *Science* **287**, 667–670.

[26] Pitzer, V. E., Viboud, C., Simonsen, L., Steiner, C., Panozzo, C. A., Alonso, W. J., Miller, M. A., Glass, R. I., Glasser, J. W., Parashar, U. D. *et al.*, 2009 Demographic Variability, Vaccination, and the Spatiotemporal Dynamics of Rotavirus Epidemics. *Science* **325**, 290–294. (doi:10.1126/science.1172330).

[27] Broutin, H., Viboud, C., Grenfell, B. T., Miller, M. A. & Rohani, P., 2010 Impact of Vaccination and Birth Rate on the Epidemiology of Pertussis: a Comparative Study in 64 Countries. *Proceedings of the Royal Society B: Biological Sciences* **277**, 3239–3245. (doi:10.1098/rspb.2010.0994).

[28] Stone, L., Olinky, R. & Huppert, A., 2007 Seasonal Dynamics of Recurrent Epidemics. *Nature* **446**, 533–536. (doi:10.1038/nature05638).

[29] Forrest, J. & Miller-Rushing, A. J., 2010 Toward a Synthetic Understanding of the Role of Phenology in Ecology and Evolution. *Philosophical Transactions of the Royal Society: Biological Sciences* **365**, 3101–3112. (doi:10.1098/rstb.2010.0145).

[30] Villerme, P., 1831 De la Distribution par Mois des Conceptions. *Annales d'Hygiène Publique, Industrielle et Sociale* **5**, 55–155.

[31] Quetelet, A., 1869 Memoire sur les lois des Naissance et de la Mortalitie a Bruxelles. In *Nouveau Memoires de l'Academie Royaledes Sciences et Belles Lettres*

de Bruxelles, pp. 495–512. L'Acadmie Royale de Bruxelles et de l'Université Louvain.

[32] Macfarlane, V., 1970 Seasonality of Conception in Human Populations. *Biometeorology* **4**, 167–182.

[33] Condon, R. G. & Scaglion, R., 1982 The Ecology of Human Birth Seasonality. *Human Ecology* **10**, 495–511.

[34] Lam, D. A. & Miron, J. A., 1994 Global Patterns of Seasonal Variation in Human Fertility. *Annals of the New York Academy of Sciences* **709**, 9–28.

[35] Cowgill, U. M., 1966 The Season of Birth in Man. *Man* **1**, 232–240.

[36] Becker, S., 1991 Seasonal Patterns of Births and Conception Throughout the World. *Advances in Experimental Medicine and Biology* **286**, 59–72.

[37] Rosenberg, H. M., 1966 Seasonal Variation of Births United States, 1933-63. *National Center for Health Statistics* **21**, 1–59.

[38] Bronson, F. H., 1995 Seasonal Variation in Human Reproduction: Environmental Factors. *The Quarterly Review of Biology* **70**, 141–164.

[39] Seiver, D. A., 1985 Trend and Variation in the Seasonality of US Fertility, 1947-1976. *Demography* **22**, 89–100.

[40] Moos, W. S. & Randall, W., 1995 Patterns of Human Reproduction and Geographic Latitude. *International Journal of Biometeorology* **38**, 84–88.

[41] United States Centers for Disease Control and Prevention, 2010. United States Vital Statistics. http://www.cdc.gov/nchs/products/vsus.htm. Accessed: March 1, 2010.

[42] United States Census Bureau, 2010. United States Population Estimates: Historical Estimates. http://www.census.gov/popest/data/historical/index.html. Accessed: March 1, 2010.

[43] United Nations, 2012. Worldwide Births. http://data.un.org. Accessed: May 1, 2012.

[44] Torrence, C. & Compo, G. P., 1998 A Practical Guide to Wavelet Analysis. *Bulletin of the American Meteorological Society* **79**, 61–78.

[45] King, A. A., Ionides, E. L., Bretó, C. M., Ellner, S., Kendall, B., Wearing, H., Ferrari, M. J., Lavine, M. & Reuman, D. C., 2010 *POMP: Statistical Inference for Partially Observed Markov Processes (R package).*

[46] Finkenstadt, B. F. & Grenfell, B. T., 2000 Time Series Modelling of Childhood Diseases: A Dynamical Systems Approach. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* **49**, 187-205.

[47] Bjørnstad, O. N., Finkenstadt, B. F., & Grenfell, B. T., 2002 Dynamics of Measles Epidemics: Estimating Scaling of Transmission Rates using a Time Series SIR Model. *Ecological Monographs* **72**, 169-184.

[48] He, D. & Earn, D. J. D., 2007 Epidemiological Effects of Seasonal Oscillations in Birth Rates. *Theoretical Population Biology* **72**, 274–291. (doi: 10.1016/j.tpb.2007.04.004).

[49] Dorélien, A. M., 2013 A Time to Be Born : Birth Seasonality in Sub-Saharan Africa. *University of Michigan Population Studies Center Research Report Series Report* **13-785**, 1–60.

[50] Lam, D. & Miron, J. A., 1987 The Seasonality of Births in Human Populations. *University of Michigan Population Studies Center Research Report Series* **87-114**, 1–61.

[51] Lam, D. A. & Miron, J. A., 1991 Seasonality of Births in Human Populations. *Social Biology* **38**, 51–78.

[52] Rojansky, N., Brzezinski, A. & Schenker, J. G., 1992 Seasonality in Human Reproduction: an Update. *Human Reproduction* **7**, 735–745.

[53] Levine, R. J., Mathew, R. M., Chenault, C. B., Brown, M. H., Hurtt, M. E., Bentley, K. S., Mohr, K. L. & Working, P. K., 1990 Differences in the Quality of Seman in Outdoor Workers During Summer and Winter. *The New England Journal of Medicine* **323**, 12–16.

[54] Spira, A., 1991 Epidemiologic Aspects of the Relationship Between Temperature and Male Reproduction. *Advances in Experimental Medicine and Biology* **286**, 49–56.

[55] Bobak, M. & Gjonca, A., 2001 The Seasonality of Live Birth is Strongly Influenced by Socio-Demographic Factors. *Human Reproduction* **16**, 1512–1517.

[56] Randall, W., 1987 A Statistical Analysis of the Annual Pattern in Birth in the USA, 1967 to 1976. *Journal of Interdisciplinary Cycle Research* **18**, 179–191.

[57] Morley, D., 1969 Severe Measles in the Tropics. -I. *British Medical Journal* **1**, 297–300.

[58] Bauch, C. T., Earn, D. J. D., 2003 Transients and Attractors in Epidemics. *Proceedings of the Royal Society London B: Biological Sciences* **270**, 1573–1578.

[59] Dorélien, A. M., Ballestros, S., Grenfell, B. T, 2013 Impact of Birth Seasonality on Dynamics of Acute Immunizing Infections in Sub-Saharan Africa. *PLOS ONE* **8**, e75806.

CHAPTER III

The Influence of Biological Rhythms on Host-Parasite Interactions

Preamble. This chapter is our published manuscript provided verbatim. *Reprinted from Publication Trends in Ecology & Evolution, Vol 30 issue 6,* ***Micaela Martinez-Bakker****, Barbara Helm, The influence of biological rhythms on host-parasite interactions, Pages 314–326, Copyright (2015), with permission from Elsevier.*

Abstract. Biological rhythms—from circadian control of cellular processes to annual cycles in life history—are a main structural element of biology. Biological rhythms are considered adaptive because they allow organisms to partition activities to cope with, and take advantage of, predictable fluctuations in environmental conditions. A flourishing area of immunology is uncovering rhythms in the immune system of animals, including humans. Given the temporal structure of immunity, and rhythms in parasite activity and disease incidence, we propose that the intersection of Chronobiology, Disease Ecology and Evolutionary Biology holds the key to understanding host-parasite interactions. We review host-parasite interactions while explicitly considering biological rhythms, and propose that (1) rhythms influence within-host infection dynamics and transmission between hosts, (2) rhythms might account for diel and annual periodicity in host-parasite systems, and (3) rhythms can lead to a host-parasite arms race in the temporal domain.

Biological Timekeeping

Environmental rhythms are a ubiquitous feature of our planet. Many rhythms are caused by geophysical cycles, including diel, tidal, lunar, and annual rhythms. These rhythms are highly predictable and have resulted in the evolution of biological clocks throughout the tree of life [1]. Most biological activities have rhythmic time structure, which scales from gene expression to life history events such as breeding and hibernation. Rhythmic time structure allows organisms to partition and prioritize life history activities—whether they are molecular or behavioral—relative to predictable fluctuations in environmental conditions. For example, for cyanobacteria, which are an ancient lineage, sunlight provides both energy and risk. Cyanobacteria have adapted to this challenge by temporally partitioning photosynthesis from UV-sensitive DNA replication [2]. Likewise, throughout the year, organisms must meet survival needs, while seasonally requiring further resources for reproduction and other activities. This leads to annual cycles of life history, when animals alternate between reproductively active states and inactive states such as dormancy, hibernation, or migration, a retreat to wintering grounds that buffer against resource scarcity [1].

Over evolutionary time, organisms have adapted to environmental fluctuations by an internal representation of time—endogenous biological clocks—that perpetuate biological rhythms even when environmental conditions are kept constant. These rhythms are characteristically innate, evidenced by the observation that individuals who have never experienced environmental fluctuations display rhythmicity [1]. Endogenous biological rhythms oscillate with period lengths that approximate those of geophysical cycles, and are accordingly called circadian, circatidal, circalunar, and circannual. Circadian rhythms, which are most heavily studied, are driven by cell-specific transcription-translation feedback loops that are integrated across the organism. The evolutionary origin of internal clocks is ancient, with circadian clocks being

a unifying feature of eukaryotes and cyanobacteria [3], and a new area of research in other bacterial lineages [4]. The endogenous circannual clock underlying seasonal rhythms is also thought to be evolutionarily conserved, since circannual clocks are found in organisms ranging from dinoflagellates [5] to mammals and birds [6, 7].

An internal representation of time enables the anticipation of favorable environmental conditions, ensuring that activities are initiated in advance to match the opportune time. For example, to rear offspring at the time of maximal food abundance, many species activate their reproductive system and copulate far in advance, potentially under harsh conditions. If individuals initiated breeding activities when food abundance was maximal, offspring would be reared outside the optimal environmental window [6]. Endogenous biological clocks function in concert with the geophysical cycles to which they synchronize [8, 9]. Synchronizing cues (also called *zeitgebers*) include diel and annual changes in light, temperature, and other factors. Jetlag, the overturning of rhythms resulting from changing time zones, and the subsequent re-synchronization of the circadian clock, is familiar to many of us. Species and even populations vary greatly in the way their clocks interact with the environment. They assume different phases, e.g., of activity or reproduction, relative to the environmental cycle, and also differ in the use of synchronizing cues. To varying degrees, organisms retain the ability to adjust their rhythms to respond to current, less predictable, conditions. While some species' rhythms show considerable phenotypic plasticity (e.g., the reproductive rhythm of Great tits, *Parus major*), other species have rigid rhythms that impose fitness costs under rapid environmental change (e.g., the seasonal phenology of Snowshoe hare coat color, *Lepus americanus*) [10, 11, 8]. In addition to phenotypic plasticity, evolutionary malleability of biological rhythms is supported by directional evolution of time adjustments in multiple species, which include heritable shifts in the seasonal timing of life history events such as reproduction, dormancy, and migration

[12, 13].

Biological rhythms are observed across biological processes. In addition to substantial diel and annual fluctuations in activity, reproduction, and metabolism, there is also overwhelming evidence for temporal structuring of immunity. Importantly, such fluctuations cannot be comprehensively characterized as changes in overall immunity; rather, they are a selective re-organization of structural and functional aspects of the immune system [14, 15, 16]. Differentiated temporal structuring of immune defenses can arise from heterogeneous requirements and costs of specific defenses, investment in self-maintenance versus immunity, or the integration of immunity with other aspects of physiology [17, 18]. In light of this, we review biological rhythms pertinent to host-parasite interactions, and propose that rhythms of hosts, parasites, and the environment impose temporal structure on epidemiological and evolutionary dynamics.

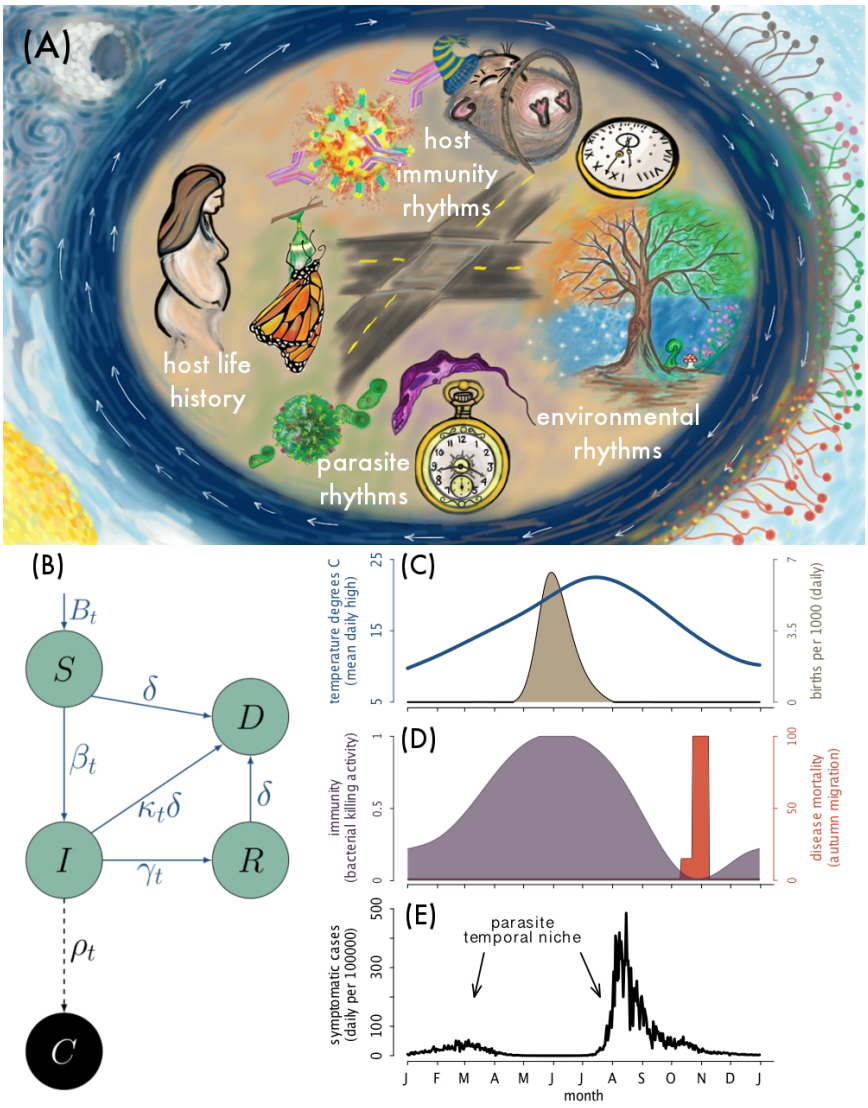


Figure 3.1: **Rhythms and Temporal Niche.** (A) The timing of host and parasite activities falls in the intersection of environmental rhythms, host life history, host immunity rhythms, and parasite rhythms. This intersection is embedded within geophysical rhythms, diel and annual cycles. (B) Biological and environmental rhythms can enter into epidemiological models in multiple ways. The schematic shows a *S*usceptible-*I*nfected-*R*ecovered model, *SIR*, with natural and disease-induced deaths, *D*. The model distinguishes between infections, I, and the subset of infections that are observed as symptomatic cases, C. The model is parameterized using the life history of Siberian stonechats. Four seasonal rhythms enter into the model (births, temperature, immunity, and migration). Host births, B_t, are seasonal. The transmission rate, β_t, is a function of (1) an environmental rhythm (i.e., temperature) that influences parasite transmissibility, and (2) the seasonal immune status of hosts. We assume seasonal immunity also influences the recovery rate, γ_t, and the probability of symptoms, ρ_t. We also assume infected individuals suffer disease-induced mortality, κ_t, associated with the autumn migration (i.e., migratory culling), which multiplies the (here constant) rate of natural mortality δ. (C) Annual fluctuations in temperature and birth seasonality in Siberian stonechats [95]. (D) Annual host immunity is based on bacterial killing activity [53], elevated mortality during autumn migration is inferred from natural migratory timing. (E) Incidence of symptomatic cases assuming: temperature has a positive correlation with transmission, bacterial killing activity reduces transmission, reduces the probability of symptoms, and increases the recovery rate. The four seasonal rhythms act collectively to determine the parasite's temporal niche, the time of year when the parasite is abundant and disease outbreaks occur.

Timekeeping in the Host-Parasite Context

Interactions between hosts and parasites (i.e., microparasites and macroparasites) are embedded within environmental rhythms (Figure 1A). In addition to the environment, host immunity imposes selective pressure on parasites, whilst parasite-driven morbidity and mortality reduces host fitness. These multiple selective forces make optimal timing of allocation of limited resources to survival and reproduction particularly tricky. For hosts, massive investment into parasite resistance, for instance, might only be energetically feasible during a resource pulse (i.e., opportunity) also favorable for reproduction, resulting in an optimization problem for resource allocation to survival versus reproduction [19]. Yet hosts also undoubtedly face the challenge of mitigating the deleterious effects of parasites when resources are scarce, a situation that might favor investment into parasite tolerance versus resistance. For parasites, not only does the host immune response impose risk, additional risks can be introduced by environmental regimes during transmission [20] or environmental life stages [21]; which has led to parasite risk avoidance strategies such as climate-driven arrested development [22]. For both hosts and parasites, therefore, external environmental conditions impose selective pressure by providing fluctuating opportunity for reproduction and risk of mortality. These exogenous factors need not be identical for hosts and parasites, although they co-occur in the same physical environment. For example, we need not expect that rhythms in parasite reproduction, host reproduction, and host immune investment be synchronized. An empirical case of this is the seasonal influence of temperature and humidity on development of the free-living nematode parasite (*Trichostrongylus*) of rabbits, which results in an autumn peak in the force of infection; whereas, the rhythm in host immunocompetence has a peak in the springtime [23].

The temporal structure of host immunity and parasite success suggests that constraints (1) preclude hosts from maintaining high levels of parasite resistance, and (2) prevent parasites from sustaining high reproductive output (fitness). Such constraints would result in trade-offs between investments in opportunity versus risk avoidance [24, 19]. Consequently, we hypothesize that both hosts and parasites time their biological processes with reference to both the external environment and each other, and that therefore in many cases periodic incidence of infectious disease is a consequence of biological rhythms, as has been suggested elsewhere (e.g., [25]). Below, we first lay out empirical evidence for the role of rhythms in host-parasite interactions. In order to inspire quantitative study of biological rhythms in host-parasite systems, we utilize a transmission model to illustrate the epidemiological consequences of rhythms. We then formulate a conceptual evolutionary model for understanding host-parasite dynamics embedded within the rhythmic context in which they are evolving.

Biological Rhythms in Host and Parasite Traits

The incidence of many infectious diseases displays substantial seasonality [26, 27, 28]. Seasonally structured disease incidence can be discussed from the viewpoint of hosts or parasites. From a host's perspective, parasite exposure can be influenced by host behavior, such as seasonal aggregation; a contemporary example being epidemic seasonality of mycoplasmal conjunctivitis in house finches [29]. However, physiological factors influencing host susceptibility to infection and symptomatic disease, such as seasonal changes in immunity, can also drive disease seasonality [28, 23, 30]. Table 1 summarizes some known diel and annual rhythms in host immunity and parasite traits (see [31, 32, 33, 16, 34] for extensive reviews). Although rhythms in immunity are observed across a broad array of taxa, including plants [35, 36] and animals, we focus on mammalian and avian hosts to enhance the link to human health. A very active area of biomedical research is characterizing temporal structure in both innate

and adaptive immunity, and correspondingly, in disease susceptibility [33, 32]. For now, biomedical studies use model organisms; but in the future should include non-model organisms [37], which are either directly relevant for understanding natural host-parasite systems, or enable a broader understanding of within-host dynamics of infection. Longitudinal studies of wild or captive animals compare immune parameters during active versus resting phases and are typically combined with experimental approaches. In the wild these may include repeated immune challenges [38, 39, 40], and in captivity may involve constant conditions, shifted environmental rhythmicity, or biological clock disruption. Studies of rhythms in immunity under natural conditions—wild immunology—are important for understanding how non-model organisms deal with exposure to multiple co-occurring parasites [41]. Studies of wild systems allow us to test, for instance, how seasonal allocation into defense against one parasite can result in enhanced susceptibility to another [42], and whether temporal variation in immune status covaries with other physiological traits and is influenced by nutritional status and parasite exposure [43, 18, 44]. Laboratory studies, in turn, are necessary for distinguishing between endogenous rhythms in immunity versus variation that occurs as a result of patterns of infection or other biotic factors.

By profiling the response to infection across time, much progress has been made in characterizing biological rhythms in immunity. For example, in mice, the circadian rhythmicity of a receptor that recognizes pathogens substantially influences the inflammatory response and survival prospects (for details, see Box 1). The health implications of circadian rhythms in immunity have also been demonstrated using hosts entrained to different light-dark cycles, and mice with genetically modified circadian clocks. Such recent studies have revealed that the immune system is fundamentally circadian in nature [33, 45, 46, 47, 48, 49], which is highlighted by the local circadian clock of macrophages [46], and the feedback between immunity and molecular, cel-

lular, and behavioral rhythms. The emerging picture is that the immune system is an active component of integrated whole-body circadian rhythms in animals [50] and plants [35, 36], lending support to the idea that sophisticated mechanisms of immune defense were also present in their common ancestor [51].

Annual cycles in immunity are not as well characterized as circadian cycles because of the time scale of experimentation [31, 16] but are epidemiologically relevant [28]. Longitudinal studies under controlled captive conditions have revealed substantial annual changes in immune parameters (Table 1). These included a down-regulation of key aspects of immunity during the time of reproductive activation, induced solely by photoperiodic simulation [52, 34, 14, 16]. Such rhythms might have evolved from a trade-off between immune defense and demanding life-cycle stages, and can underlie annual patterns of disease incidence, as suggested, for example, by rhythms of bactericidal capacity of whole blood (Box 1; Figure 1D [53]). It is important to note, however, based on the existing evidence for both circadian and annual immunomodulation, that temporal patterns can differ between innate and adaptive immunity and among traits even within the same immune cell subset [16, 33].

In addition to immunomodulation, several other aspects of host rhythmicity can have population-scale consequences for host-parasite dynamics. Relevant host annual cycles include aggregation [29, 54, 55], sexual contacts (with regard to STDs), habitat use, migration [56, 57, 58], and birth pulses that (1) act to replenish the pool of susceptible individuals, (2) can influence the critical community size required for parasite persistence, and (3) can determine the geographic synchrony of outbreaks [59, 60, 61, 62, 63]. The sweeping effects of annual cycles in host physiology on disease incidence are exemplified by White-Nose Syndrome (WNS), which is drawing many North American bat species near extinction. A new longitudinal study indicated that neither birth pulses nor social behavior affected transmission and intensity of

WNS. Instead, WNS is associated with hibernation [30], which in mammals that have been studied in captivity, is a programmed circannual rhythm [6]. We speculate the link between WNS and hibernation is mediated by hibernation-associated changes in immunity. Evidence thus far suggests that there are large adjustments in immunity in hibernating mammals, including a 90% decrease in circulating white blood cells [64], down-regulation of the acute-phase response to LPS [65], and modifications of intestinal immunity [66, 64].

Migration is another host rhythmicity receiving attention in infectious disease ecology. In monarch butterflies, the protozoan parasite *Ophryocystis elektroscirrha* displays a seasonal pattern of prevalence and a spatial gradient along the monarchs' migratory flyway. Parasite prevalence declines as monarchs migrate, which is likely due to migratory culling [57]. In migratory culling, the coupled energetic demands of migration and fighting infection result in increased mortality of infected individuals during fall and spring migrations. The uninfected are most likely to survive the journey to the breeding or wintering grounds, allowing the destination to be relatively parasite-free. Thus, migratory culling is a direct intersection of host seasonal rhythms and disease prevalence [56], and anthropogenically-driven disruption of this rhythm results in elevated disease burden [58].

Taking the parasite's perspective, rhythmic patterns in parasite dissemination can be influenced by fluctuating abiotic and biotic conditions that affect parasite survival and transmission. Clear examples of abiotic influences are (1) the role of temperature and humidity in transmission of influenza [67], which might be responsible for latitudinal clines observed in influenza incidence [68, 69], and (2) the UV sensitivity of sporulation in *Isospora*, which might have driven the remarkably robust diel pattern of oocyst outputs [21]. In addition to abiotic effects, biotic influences can stem from rhythms of vectors [70] and other parasites [71]. Effects of vector circadian rhyth-

micity have been studied in the malaria vector *Anopheles gambiae*, whose rhythmic gene expression persists under constant conditions. Rhythmically expressed genes include those implicated in the melanization immune response, which encapsulates the Plasmodium parasites, and can thereby affect mosquito to human transmission. Vectors can also temporally structure parasite transmission via their diel patterns of feeding [72] and their phenology [73]. Interspecific influence of parasites on one another's rhythm, to our knowledge, has only been described for *Drosophila* parasitoids, which gain a fitness advantage by temporally segregating circadian rhythms in egg oviposition [71], hypothesized to alleviate competition.

Perhaps then unsurprisingly, parasites—faced with rhythms in their abiotic environment, hosts, and vectors—display what seem to be biological rhythms. Documentation of parasite rhythms dates back over 100 years, long before the discovery of biological clocks. In fact, the early observation that both malaria parasites and microfilariae are abundant in the blood of hosts at night was instrumental to the discovery of mosquitoes as the malaria vector [74]. Experimental studies of parasites report diel and annual rhythms, as measured by fluctuations in parasite burden and infectivity (Table 1), but disentangling the contributions of host and parasite to these rhythms is difficult [75]. To our knowledge, the only described example of a parasite life history event that depends on a host rhythm is reproduction in the ectoparasitic rabbit flea, a vector of myxoma virus. To reproduce, rabbit fleas must undergo maturation on a pregnant or newborn nestling host, and flea maturation is controlled by host hormone cues associated with pregnancy and parturition, thereby synchronizing parasite and host life cycles [76, 77, 78]. There is solid evidence for adjustment of diel parasite rhythms to those of the host, for example from trematodes like *Schistosoma mansoni*—an agent of schistosomiasis in humans—whose emergence from snail hosts is initiated by light [79, 80]. These parasites display diel cycles that shift to

match perturbations in their hosts' circadian rhythm [81]. Similar results were found in rodent-infecting *Trypanosomes*. Rats infected with *T. lewisi* and housed under a normal light-dark cycle (LD 12:12 h) experienced a peak in circulating *T. lewisi* during the early part of night and a trough in the early morning. However, when the host photoperiod was inverted, the parasite rhythm was also reversed [82]. In *Isospora*, the characteristic diel pattern of oocyst output persisted under continuous light in the high Arctic, although feeding and activity rhythmicity of their avian hosts was greatly diminished, suggesting synchronization to subtle host rhythms or possibly self-sustained parasite rhythms [21].

Experimental studies give clear evidence that synchronization to host rhythms impacts parasite fitness. For example, murine host rhythms were experimentally mismatched to that of their malaria parasite (*Plasmodium chabaudi*). This mismatch resulted in a 50 per cent reduction in both parasite replication and production of transmissible life-stages [83]. Follow up experiments have now revealed additional complexities, with the effect of mismatch manifesting differently between parasite life stages, and downstream effects on host disease severity. Mismatch can confer a substantial cost to the parasite, and this cost is experienced at the onset of infection, rather than acquired throughout infection [84, 85]. This suggests there can be intense selective pressure on parasites to maintain a specific phase position relative to their host rhythms, or to vector rhythms, since parasite ability to infect vectors is also time-of-day dependent [72]. Fortunately, the amassing knowledge of biological clocks might help identify host cues used for entrainment of parasite rhythms. For example, the nocturnally-peaking hormone melatonin is a core circadian feature of many vertebrates, and applying this knowledge produced indication that parasites might be using melatonin to synchronize their circadian cell cycle [86, 21].

While we still lack unambiguous evidence for endogenous circadian or circannual rhythms of parasites, recent research suggests the intriguing possibility that parasites can actively manipulate host and vector rhythms to their advantage. For example, various parasites interrupt host diel activity at specific times of day to enhance transmission [87, 88]. The diel timing and synchrony of host behavioral manipulation, along with candidate molecular mechanisms of manipulation, strongly implicate circadian clock pathways [87, 89]. By integrating chronobiology with infectious disease ecology, we might be able to identify, for example, the mechanism by which the trematode *Dicrocoelium* manipulates diel host behavior, inducing suicide, and facilitating trophic transmission [90], and how the notoriously manipulative fungus (*Ophiocordyceps unilateralis s.l.*) seemingly breaks the host circadian clock to perpetuate transmission [87]. Transkingdom cross-regulation between prokaryotic and eukaryotic rhythms is plausible because it has already been documented for other systems (e.g., in bioluminescent squid light organ symbionts and in mammalian gut microbiota) [4, 91].

Despite our knowledge of (1) rhythmic host immunity and physiology, (2) rhythms in parasite reproduction and transmission, and (3) enticing evidence that host rhythms can impact parasite fitness and be exploited by parasites, the effects of biological rhythms on host-parasite dynamical processes remain poorly understood. We surmise that careful consideration of biological rhythms in infectious disease ecology and evolution will provide a better understanding of (1) daily and annual patterns of diseases, (2) within-host parasite dynamics, and (3) parasite transmission.

Models for Investigating Host-Parasite Contributions to Rhythms in Infectious Disease

In order to determine how biological rhythms impose temporal structure on host-parasite dynamical processes, we can integrate empirical data on host and/or parasite rhythms into epidemiological and evolutionary models. Biological rhythms research has great potential for feedback between laboratory studies, field ecology, and dynamical systems modeling. First, rhythms in immunity characterized under laboratory or field conditions can be used in transmission models to make predictions about the epidemiological consequences of those rhythms in nature. Second, observations of diel and annual cycles in infection—characterized via disease incidence, parasite abundance, or host serological markers of infection history—can be used to make predictions regarding rhythms generating such patterns. A new study of the first type [92] explores the effect of annual and biannual rhythms in births (in bats) on the persistence of filoviruses (i.e., Marburgvirus and Ebolavirus). Transmission models predict that filoviruses can persist in species with biannual birth pulses—making them potential reservoirs of infection—and this prediction is supported by serology data showing that species with biannual birth pulses are more likely to be seropositive for filoviruses; demonstrating that explicit consideration of host rhythms can inform targeted surveillance and control of emerging zoonotic diseases. A study of the second type is that of [23], in which long-term field data on nematode infections in European rabbits were used to discriminate among multiple potential seasonal rhythms in the host-parasite system. This led to identification of epidemiologically relevant seasonality in host immunity; the endogenous nature of which can be tested in the lab.

Building upon the examples above, as well as other transmission models that incorporate reproductive rhythms [93, 94, 59, 60, 61], here we provide a Susceptible-

Infected-Recovered (SIR) model of a directly transmitted hypothetical bacterial infection in the Siberian stonechat to illustrate the numerous entry points for biological rhythms into epidemiological processes (Figure 1B). We narrate our model with seasonal rhythms in mind; however, this can be extended to circadian rhythms. The model incorporates ambient temperature as a covariate influencing parasite transmission as well as empirical data on host circannual cycles in reproduction, immunity and migration (Figure 1CD; cf. Box 1) [95, 53]. For migration, the timing is defined empirically, while the model assumes migratory culling of infected individuals only during the autumn migration, when host bacterial killing activity is lowest [53]. Importantly, we propose that circannual cycles in host immunity can influence (1) the transmission rate, (2) the recovery rate, and (3) the pathological consequences of infection, which manifests as symptomatology and enters the model as the report rate. The multiple rhythms: temperature, births, bacterial killing activity, and migratory culling act collectively to shape the observed incidence of disease, which is the model output shown in Figure 1E. We define the resulting seasonal window of elevated disease incidence as the parasite's temporal niche. The seasonal incidence that arises from this model matches the expectations from the underlying data. However, in contrast to most models of seasonal infectious diseases, which only place sinusoidal seasonality in the transmission rate, it contains multiple axes of seasonal forcing. Thus, we provide this model to encourage the inclusion of empirically characterized rhythms into models as covariates. Such models can be used to explore the epidemiological consequences of host and parasite rhythms, although for simplicity the parasite is not explicitly modeled here. Host-parasite systems where modeling parasite rhythms is particularly compelling include nematodes with seasonal arrested development [96, 22], microfilariae which display both circadian and seasonal cycles [97], and *Plasmodium* within-host circadian cycles [75]. Major challenges of incorporating biological rhythms into epidemiological or within-host models, will be (1)

recognizing which host and/or parasite rhythms are epidemiologically relevant, and (2) identifying the functional relationships between rhythms and epidemiological parameters, including: transmission, recovery, and symptomatology.

In addition to the epidemiological consequences of rhythms, we can benefit from understanding the feedback between host and parasite rhythms and the multiple axes that shape their temporal structure. Thus, we provide a conceptual evolutionary model for understanding how hosts and parasites time their biological processes with reference to each other while being embedded in environments with temporally structured risk and opportunity. We pose this model in evolutionary terms, but depending on the varying degrees of plasticity of biological rhythms, individuals may also adjust their rhythms during their life.

Our evolutionary model illustrates three idealized scenarios of how host immune defense varies seasonally, relative to fluctuating environmental conditions (Figure 2A–C). These scenarios are motivated by life history theory and by empirical observations of seasonal immunity in mammals and birds (Table 1). The scenarios assume that immune defense, specifically, parasite resistance, either parallels the availability of resources (A, "resource-driven"), or is reduced when resources are used for reproduction (B, "traded-off"). The third is an extension of the "resource-driven" scenario with modulation related to life history events that can lead to complicated, but potentially important, annual patterns. In our example of migration (C) we assume down-regulation of immune defense during migration (see Figure 1; this could also occur during other vulnerable times such as molting or hibernation [65]), and a shallow trough under favorable conditions in the wintering grounds. We then (D) switch to the parasite's perspective and illustrate how parasites are subject to two axes of seasonal fluctuations: (i) seasonal environmental conditions outside the host and, (ii) seasonal immune defense of the host. We propose that together the two seasonal

axes shape parasite transmission (i.e., parasite fitness; which is captured by the basic reproductive number).

The last and crucial component of our model is the evolutionary feedback between host and parasite rhythms. We propose that due to parasite-induced host morbidity and mortality, selection can drive changes in host seasonal immune defense. Subsequently, since host immune defense is one of the seasonal axes influencing parasites, selection will favor changes in the parasite rhythm. This interplay can continue, driving hosts and parasites to sequentially alter their seasonal rhythms while working within the constraints of environmental conditions. Figure 2E shows these steps.

We suggest that under certain conditions this can escalate into an evolutionary arms race. In this framework, the prerequisite for an arms race is that parasite fitness is sufficiently impacted by the temporal structure of the host immune response, and that the host immune response is predictably rhythmic. To be clear, when considering the temporal structure of immune defense, reference to "low" host immune defense pertains to the parasite in question. However, it must be appreciated that a time of diminished resistance to one parasite (e.g., a helminth) can be a time of high investment into fighting another (e.g., a virus). Furthermore, infection with one parasite can seasonally elevate host susceptibility to another, as is exemplified by concomitant infections of myxoma virus and nematodes [98] and increased susceptibility to bovine TB resulting from helminth coinfection [18]. The arms race itself has two requirements. First, hosts must be able to shift their immune defense to counter exploitation by parasites (host changes in Figure 2E). Changes in host rhythms then translate into a new landscape of time-structured risk and opportunity for parasites. Upon experiencing a new temporal landscape, a dynamic host-parasite arms race can arise only if the second requirement is met: parasites shift their rhythm by changing reproduction within hosts, or release from hosts (parasite changes in Figure 2E). As

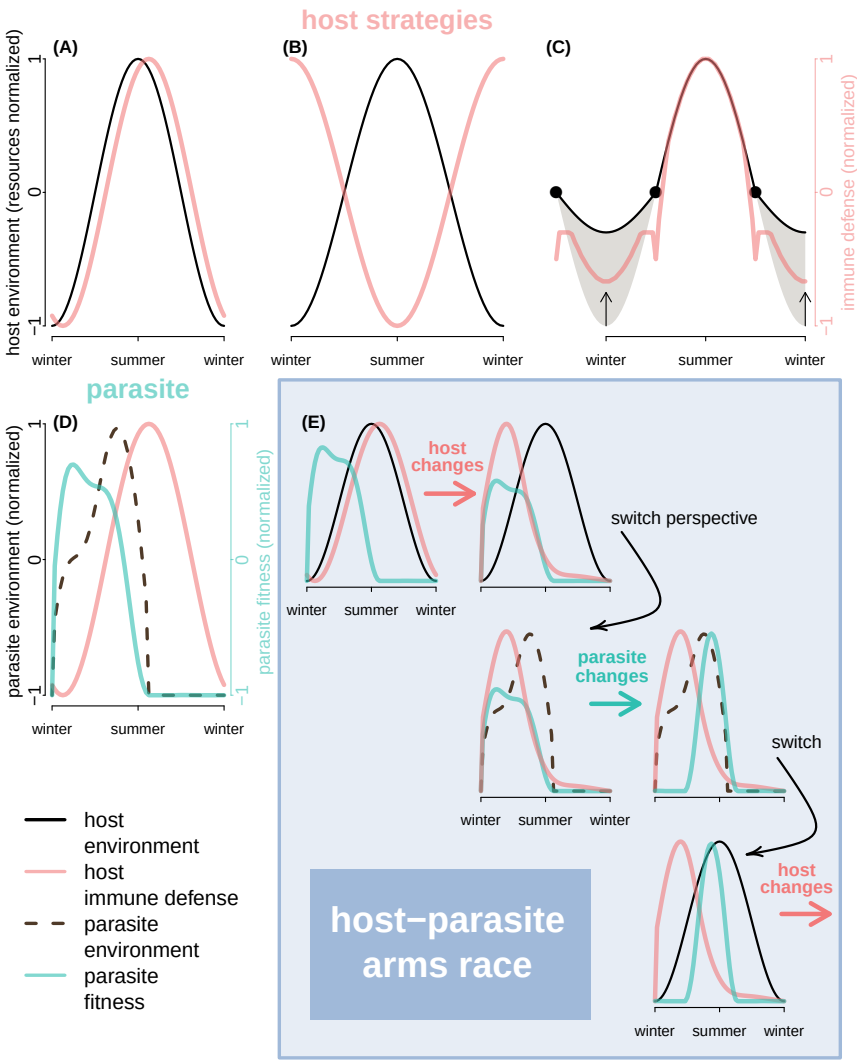


Figure 3.2: **Conceptual model for investigating host-parasite contributions to rhythms in infectious disease.** (A) Host immune defense is resource-driven and tracks the host's environmental conditions (i.e., host resource availability). (B) Host immune defense has an inverse relationship with environmental conditions; this could occur due to a trade-off against investment into reproduction during high resource availability. (C) Resource-driven immune defense in a migrating species that has reduced immune defense during migration. Migrations (indicated by black points) result in shallower environmental troughs since individuals migrate to regions with higher resource availability. For all scenarios, we consider immune defense to be resistance to the parasite in question, although we acknowledge that this simplifies the complexity of the immune system (e.g., independent immunomodulation of innate or adaptive immune parameters). For the resource-driven host immune defense strategy, in (D) we show seasonal parasite fitness shaped by both environmental conditions and seasonal host immune defense. Although host and parasite co-occur in the same physical environment, the environmental rhythms pertinent to the parasite need not be identical to the environmental rhythms pertinent to the host, which is why we distinguish host versus parasite environmental rhythms. (E) Arms race between host and parasite. For illustrative purposes, the arms race is initiated with resource-driven host immune defense and parasite seasonality. The host changes the seasonal timing of peak immune defense to coincide with peak parasite fitness. We then switch to the parasite perspective to consider the parasite's environment. In response to the new host seasonality, the parasite changes its timing of peak reproductive output. These cycles can continue, with both host and parasite seasonality shifting within the bounds of their respective environmental constraints.

with other host-parasite arms races, an arms race in the temporal domain is subject to tradeoffs for both the host and the parasite that might constrain the extent to which their rhythms can be altered. Host tradeoffs can include an immunity-reproduction tradeoff [99]; whereas, tradeoffs for the parasite can include a transmission-virulence or transmission-recovery tradeoff [100, 99, 101]. Also, due to the rapid generation time of parasites, relative to hosts, evolution of host rhythm shifts might be slow relative to the evolution of parasite rhythms, but this would not preclude an arms race from occurring.

Conclusion

There is enticing evidence that biological rhythms are structuring elements of host-parasite interactions, both in within-host processes and in epidemiological dynamics. Host circadian rhythms in the immune system influence the progression of infection and parasite burden, and annual rhythms might have similar effects [23, 30]. The existing evidence leads us to conclude that the effects of biological rhythms on the perpetuation of parasites, and on host reactions during infection, can generate population-level rhythms in infectious disease incidence [25], which we here define as the parasite's temporal niche. To formalize temporal niches across parasite taxa and life history strategies, we will need novel integration of epidemiological, immunological, and life history data of both hosts and parasites.

Importantly, circadian rhythms in immunity have direct implications for transmission, and practical application for (i) timing of antibiotic, antiviral, and anthelmintic treatment, and (ii) managing immunopathology, such as cytokine storms. The rhythms of parasites themselves can drive patterns of exposure and illness, as is evident in malaria and filarial infections. Similarly, rhythms in parasitemia and parasite release from hosts can impose temporal structure on transmission, which can

be leveraged for interventions such as deworming campaigns.

We believe that a multi-disciplinary approach at the intersection of Chronobiology, Disease Ecology and Evolutionary Biology holds the key to understanding how biological rhythms influence host-parasite interactions. We have outlined open questions that will bring us closer to understanding the underlying biological interactions in the temporal domain (refer to Outstanding Questions Box). We hope that this Opinion will generate discussion on how to leverage rhythms for translational medicine, for instance to counter the evolution of resistance; and we hope the insights provided here inspire new avenues for interrogating transmission models with host-parasite data from the laboratory and the field, ultimately, to better understand the forces structuring disease incidence and the immunology of non-model organisms.

Finally, although it is beyond the scope of this Opinion, the importance of accounting for biological rhythms is accentuated by the accumulating data on anthropogenically-driven disruptions and mismatches of biological rhythms that are occurring across taxa. Circadian disruption due to light-at-night [102] and altered environmental seasonality due to climate change [13] are challenging the plasticity of rhythms and modifying the fitness advantages of their endogenous basis. For example, the adverse effects of circadian disruption have already been seen in human health and gut microbiota [91]. Given the pervasiveness of rhythms in host immunity, vectors and parasites, we might soon be faced with palpable effects of rhythm disruptions on infectious diseases [103, 104, 73].

Table 3.1: Rhythms in Hosts and Parasites. Diel (circadian) and annual (circannual/seasonal) rhythms in host immunity, parasite reproduction, and parasite release.

Type	**Period**	**Host Trait or Parasite Description**	**Organism/ Species**	**Rhythm**	**Citation**
Immunity	Diel	Macrophages (detection and restriction of parasite invasion)	Mice	8% of transcripts are circadian; autonomous macrophage circadian clock controls rhythm	[46]
Immunity	Diel	Natural Killer Cells (early defense against viruses and intracellular bacteria)	Humans	Circadian trafficking between the blood and organ compartments; inverse trafficking compared to T-lymphocytes	[105]
Immunity	Diel	Toll-like receptor 9 (evolutionarily conserved receptor that recognizes bacteria and viruses)	Mice	Expression and function controlled by circadian clock	[47]
Immunity	Diel	T-lymphocytes (surveillance for infected cells)	Humans	Cytokine production	[106, 107]
Immunity	Diel	Leukocytes	Humans	Abundance in the blood follows a circadian rhythm for neutrophils, lymphocytes, monocytes, and eosinophils	[108]
Immunity	Diel	Whole blood response to LPS stimulation	Humans	Cytokine and chemokine production is circadian in an environment free of time cues	[109]
Immunity	Diel	*Salmonella* colonization and host cytokine response to infection	Mice	Mice have a different immunological response to infection depending on whether infection challenge occurs at day or night; infection during the day results in more inflammation; this effect is due to clock-controlled gene expression	[49]
Immunity	Diel	Clock genes and pro-inflammatory cytokines in spleen; inflammatory response	Birds (captive)	mRNA of cytokines and clock genes are rhythmic under LD cycles and constant conditions; inflammation is rhythmic under LD cycles	[110]
Immunity	Diel	Cellular (PHA) and humoral immune response	Birds (captive)	PHA response and antibody production depend on time of challenge, but their peaks are phase-inversed	[17]

Continued on next page

Type	Period	Host Trait or Parasite Description	Organism/ Species	Rhythm	Citation
Immunity	Annual	Bacterial killing activity	Birds (captive), turtles (wild), humans	Lower bactericidal activity during migration, especially in autumn (birds); Higher bactericidal activity during breeding season (turtles); Higher bacterial killing by neutrophils in summer (humans)	[53, 111, 112, 113]
Immunity	Annual	Leukocytes	Birds (captive)	Annual cycles in several immune traits, no effect of Coccidia on annual cycle of immune measures	[114]
Immunity	Annual	Lysis	Birds (wild)	Lower ability of plasma to lyse foreign cells during migration and winter	[38]
Immunity	Annual	Sickness behavior in response to LPS	Birds (captive and wild), hamsters	Repression of sickness behavior during reproduction in summer (birds); repression of sickness behavior during winter (hamsters)	[39, 115]
Immunity	Annual	Acquired Immunity	Rabbits (wild)	Resistance against nematodes	[23]
Immunity	Annual	Spleen size (spleen is important for both innate and adaptive immunity)	Birds (wild)	Regression of spleen during migration	[116]
Immunity	Annual	Cytokine production stimulated by bacterial endotoxin	Humans, rats, hamsters	Seasonal differences in pro- and anti-inflammatory cytokine production (humans); Summertime photoperiod increases production of pro-inflammatory cytokine TNF-α and extends (rats) or elevates (hamsters) disease symptoms	[117, 118, 119]
Immunity	Annual	Vaccine response	Humans	Seasonal variation in symptoms following live influenza vaccine	[120]
Immunity	Annual	Intestinal immunity	Ground squirrels (captive)	Increase in intestinal leukocytes, pro- and anti-inflammatory cytokines during hibernation	[66]

Continued on next page

Type	Period	Host Trait or Parasite Description	Organism/ Species	Rhythm	Citation
Susceptibility	Diel	Bacterial burden, pathogenesis, and/or virulence of infection	Mice	Timing of infection can affect: (a) bacterial burden due to circadian variation in monocyte trafficking and/or gene expression at site of infection, (b) disease severity from sepsis due to circadian TLR9 expression, and (c) virulence	[45, 49, 47, 121, 122]
Susceptibility	Annual	Susceptibility to fungal growth	Bats (wild)	Hibernating bats have temperatures that match that of hibernacula, allowing explosive growth of WNS fungal pathogen *Pseudogymnoascus destructans*	[30]
Parasite reproduction	Diel	*Plasmodium* species (Malaria parasite) asexual reproduction	Various mammalian hosts	Parasite cohorts of millions of individuals synchronously burst from red blood cells at a particular time in LD cycle	[75]
Parasite reproduction	Annual	Microfilariae (heartworm)	Dogs	Strong seasonal rhythm in microfilaria abundance in dogs infected in the lab	[97]
Parasite development	Annual	Nematodes (parasitic roundworms)	Domestic mammals	Parasites engage in seasonal hypobiosis, arrested development within hosts that allows for persistence when environmental conditions are unfavorable for transmission between hosts	[96]
Parasite discharge	Diel	Coccidia	Birds (wild and captive)	Oocyte release is strictly circadian; although parasites are vulnerable to sun exposure the rhythmic pattern persists under continuous light during the Arctic summer	[123, 21]
Parasite discharge	Diel	*Echinostoma* (parasitic flatworms), Pinworms, *Schistosoma*	*Echinostoma* in mice, Pinworms and Schistosomes in humans	The release of *Echinostoma* eggs by hosts occurs during night when mice are active. Human pinworms migrate out of anus during the night to lay eggs. In contrast, *Schistosoma* eggs are discharged in urine during the day.	[124, 74]
Parasite discharge	Annual	Nematodes (roundworms)	Dall's Sheep (wild)	Seasonal variation in intensity of parasite larvae shed in feces	[125]

Continued on next page

Type	Period	Host Trait or Parasite Description	Organism/ Species	Rhythm	Citation
Parasite manipulation of host behavior	Diel	*Dicrocoelium* trematode	Ants (intermediate host)	In the evening infected ants affix themselves to the top of blades of grass and enter torpor until the next morning, allowing them to be eaten by grazing mammals (definitive hosts)	[90]
Parasite manipulation of host behavior	Diel	Manipulating fungus *Ophiocordyceps*	Ants	*Ophiocordyceps* manipulates host behavior and causes host death at characteristic times of day	[87]

Box 1. Circadian and Seasonal (circannual) Modulation of Host Immune Defense.

Circadian Immune Cycles. In order for hosts to mount an immune response against an infecting pathogen, the immune system must first detect the presence of the pathogen. One way that animals detect pathogens is by immune surveillance for pathogen-associated molecular patterns (PAMPs), which are shared across groups of pathogens. Host cells express pattern recognition receptors (PRRs) that recognize PAMPs. Toll-Like Receptor 9 (TLR9) is an important PRR that can recognize both viruses and bacteria, and is highly evolutionarily conserved. In 2012 Silver et al. discovered that the expression of TLR9 by macrophages and B cells follows a circadian rhythm. Importantly, the circadian rhythm of TLR9 has a significant effect on the immune response and disease severity because the rhythm of TLR9 also produces a rhythm in inflammatory cytokines. The implications were experimentally demonstrated by inducing sepsis. Sepsis can occur during bacterial infections when a severe inflammatory immune response causes damage to the host. Bacterial infection was induced in laboratory mice using a puncture that allowed commensal bacteria to enter

the body cavity from the intestine. Mice were entrained to a light-dark cycle (LD 12:12 h), and infection was induced either at the midpoint of the light period or of the dark period. Mice that were infected during the night, when the TLR9 inflammatory response was elevated, had higher bacterial burdens, earlier mortality, and worse disease scores, hypothermia, and tissue damage than mice that were infected during the day. This study demonstrated that the functional response of the immune system varies according to a circadian rhythm, and this variation is biologically relevant because it can have a significant effect on the dynamics of infection [47].

Annual Immune Cycles. Faced with stark annual fluctuations in environmental conditions and resources, many avian and mammalian species partition life history events such as reproduction, growth, and hibernation into distinct times of year, and their immune system also undergoes seasonal changes. Versteegh et al. 2014 set out to investigate whether annual variation in immunity is due to seasonal adjustments directly driven by environmental or physiological conditions, or originates from a genetically-based circannual rhythm that allows organisms to prepare for changes in the environment. They looked at 5 different immune measures, including bactericidal competence of whole blood as a proxy for functional implications.

To determine whether seasonal immunity is a genetically encoded circannual rhythm, genetically distinct subgroups of a widespread songbird, the stonechat (Saxicola torquata), were bred and raised in a common garden experiment. The subgroups chosen for this experiment differ in their seasonal life history and traits. They included a (i) long-distance migrant, (ii) short-distance migrant, and (iii) a non-migrant, along with hybrids. The prediction was that if seasonal immunity is a direct response to seasonal environmental conditions and energy demands, then by raising birds in an environment where (a) they have ample food, (b) they are not allowed to migrate or breed, and (c) the only fluctuation to which they are exposed is changing day length,

their annual rhythms in immunity would be lost.

The authors found that not only did the annual rhythm in immunity persist under these controlled conditions but also that the subgroups and hybrids of the birds showed specific patterns. The long-distance migrants displayed seasonality in 4 immunity parameters, which included bacterial killing ability (Figure 1D). The short-distance migrants displayed seasonality in only 3 immunity parameters, and the non-migrants displayed seasonality in only 2 parameters. Furthermore, the amplitude of the annual fluctuation was greatest in the long-distance migrants. The inheritance of the rhythm in hemolysis (the ability of antibodies and their compliment system to lyse foreign cells) was also quite striking. Both the long- and short-distance migrants showed reduced hemolysis during the time of the natural autumn migration. The reduction in hemolysis in the long-distance migrants was much more extreme than that of the short-distance migrant, and intermediate in F1- hybrids. Together, this work demonstrates that an inherited, biological clock controls seasonal immunity in stonechats, and these rhythms vary across groups that differ in their seasonal life history [53]. Related avian studies generally confirm annual cycles in immune parameters, and although species differ, there is a common tendency for greater seasonal immunomodulation with increasing migratory lifestyle.

Outstanding Questions.

1. Are rhythms in the immune system adaptive for fighting infection?
2. Are parasite rhythms adaptive for dealing with host rhythms or environmental conditions?
3. Are observed parasite rhythms truly endogenous?
4. If parasite rhythms are endogenous, are they entrained by host rhythms?

5. Are host circannual rhythms in immunity an adaptive response to (a) seasonal parasite exposure or (b) resource limitations and life history trade-offs?

Glossary.

Adaptive immune system - cells of the adaptive immune system include B cells and T cells. The initiation of the adaptive immune response occurs after the initiation of the innate immune response. Receptors of adaptive immune cells require genetic recombination and alteration to generate, resulting in antigen specificity and immunological memory [126, 127].

Annual cycle - cycle with a length of approximately 1 year, with phases occurring consistently at a particular time every year, on an annual/seasonal basis

Circadian - cycle that occurs with an approximate 24-hour period, used in reference to endogenous rhythms

Circannual - cycle that occurs with a period of approximately 1 year, used in reference to endogenous rhythms

Diel cycle - cycle with a length of approximately 1 day, with phases occurring consistently at a particular time during the day-night cycle

Immune defense - immune defense includes (1) resistance against the establishment of infection and the reproduction of parasites, and (2) parasite tolerance, in which the host mitigates the pathological consequences of infection, but tolerates infection

Innate immune system - cells of the innate immune system include macrophages and neutrophils. Innate immune cells are immediate responders to infection. A fundamental distinction between innate and adaptive immune cells is that

innate immune cell receptors responsible for immune recognition are encoded in the germline; whereas, receptors of adaptive immune cells require genetic recombination and alteration to be generated. It was previously thought that the innate immune response is not parasite-specific and lacks memory, but that characterization is now considered incorrect [127].

LD-cycle - cycle in which light and darkness alternate, and each last for a given duration, for example in LD 12:12 h, light and darkness both last for 12 hours

LPS - Lipopolysaccharide is a component of the outer membrane of Gram-negative bacteria that is used in experiments to elicit an anti-bacterial immune response

Macroparasites - parasites that are large and typically metazoans (e.g., helminths)

Microparasites - parasites that are small and often unicellular (e.g., pathogenic viruses, bacteria, and fungi)

Macrophage - phagocytes, often referred to as "big eaters" because they engulf invading bacteria and are responsible for clearance of dead (apoptotic) cells. Macrophages are one of the cells responsible for detection and restriction of parasite invasion.

Parasite Resistance - The ability of the host's immune response to prevent infection from establishing or limit parasite replication. Parasite resistance has a negative impact on parasite fitness [128].

Parasite Tolerance - The ability of the host to mitigate the pathological consequences of infection, rather than mitigate infection itself. Parasite tolerance does not necessarily have a negative impact on parasite fitness [128].

Acknowledgements

We would like to thank our editor Paul Craze and two anonymous reviewers for helping us with this work. We would like to acknowledge Simon Babayan, Isabella Cattadori, Jason Matthiopoulos, Jeff Shi, and Charissa de Bekker for their thoughtful comments on the manuscript, and Lewis Lanier and Alexandra McCubbrey for their insights on innate and adaptive immunity, and macrophages, respectively. Dan Haydon and Barbara Mable for inspiring in the great outdoors; as well as colleagues at the University of Glasgow Institute of BAH&CM and University of Michigan EEB for lively discussion.

Bibliography

[1] Foster, R.G. and Kreitzman, L. (2004) *Rhythms of Life: The Biological Clocks that Control the Daily Lives of Every Living Thing.* Profile Books Ltd, London

[2] Hut, R. and Beersma, D. (2011) Evolution of Time-Keeping Mechanisms: Early Emergence and Adaptation to Photoperiod. *Philosophical Transactions of the Royal Society of London. Series B, Biological sciences* 366, 2141–2154

[3] Dvornyk, V. *et al.* (2003) Origin and Evolution of Circadian Clock Genes in Prokaryotes. *Proceedings of the National Academy of Sciences of the United States of America* 100, 2495–2500

[4] Heath-Heckman, E.A.C. *et al.* (2013) Bacterial Bioluminescence Regulates Expression of a Host Cryptochrome Gene in the Squid-Vibrio Symbiosis. *mBio* 4, 1–10

[5] Anderson, D. and Keafer, B. (1987) An Endogenous Annual Clock in the Toxic Marine Dinoflagellate *Gonyaulax tamarensis*. *Nature* 325, 616–617

[6] Gwinner, E. (1986) *Circannual Rhythms: Endogenous Annual Clocks in the Organization of Seasonal Processes.* Springer-Verlag

[7] Helm, B. *et al.* (2013) Annual Rhythms that Underlie Phenology: Biological Time-Keeping Meets Environmental Change. *Proceedings of the Royal Society B: Biological Sciences* 280

[8] Visser, M.E. *et al.* (2010) Phenology, Seasonal Timing and Circannual Rhythms: Towards a Unified Framework. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 365, 3113–3127

[9] Dunlap, J.C. *et al.*, eds. (2004) *Chronobiology: Biological Timekeeping*. Sinauer Associates, Inc.

[10] Charmantier, A. *et al.* (2008) Adaptive Phenotypic Plasticity in Response to Climate Change in a Wild Bird Population. *Science (New York, N.Y.)* 320, 800–803

[11] Zimova, M. *et al.* (2014) Snowshoe Hares Display Limited Phenotypic Plasticity to Mismatch in Seasonal Camouflage. *Proceedings of the Royal Society B: Biological Sciences* 281, 20140029

[12] Tauber, E. *et al.* (2007) Natural Selection Favors a Newly Derived Timeless Allele in *Drosophila melanogaster*. *Science (New York, N.Y.)* 316, 1895–18988

[13] Bradshaw, W.E. and Holzapfel, C.M. (2006) Evolutionary Response to Rapid Climate Change. *Science* 312, 1477–1478

[14] Lee, K.A. (2006) Linking Immune Defenses and Life History at the Levels of the Individual and the Species. *Integrative and Comparative Biology* 46, 1000–1015

[15] Lee, K.A. and Klasing, K.C. (2004) A Role for Immunology in Invasion Biology. *Trends in Ecology and Evolution* 19, 523–529

[16] Stevenson, T.J. and Prendergast, B.J. (2014) Photoperiodic Time Measurement and Seasonal Immunological Plasticity. *Frontiers in Neuroendocrinology* In press

[17] Siopes, T. and Underwood, H. (2008) Diurnal Variation in the Cellular and Humoral Immune Responses of Japanese Quail: Role of Melatonin. *General and Comparative Endocrinology* 158, 245–249

[18] Ezenwa, V.O. and Jolles, A.E. (2011) From Host Immunity to Pathogen Invasion: The Effects of Helminth Coinfection on the Dynamics of Microparasites. *Integrative and Comparative Biology* 51, 540–551

[19] Sheldon, B. and Verhulst, S. (1996) Ecological Immunology: Costly Parasite Defences and Trade-offs in Evolutionary Ecology. *Trends in Ecology and Evolution* 11, 317–321

[20] Grassly, N.C. and Fraser, C. (2006) Seasonal Infectious Disease Epidemiology. *Proceedings of the Royal Society B: Biological Sciences* 273, 2541–2550

[21] Dolnik, O.V. *et al.* (2011) Keeping the Clock Set Under the Midnight Sun: Diurnal Periodicity and Synchrony of Avian Isospora Parasites Cycle in the High Arctic. *Parasitology* 138, 1077–1081

[22] Van Kuren, A.T. *et al.* (2013) Variability in the Intensity of Nematode Larvae from Gastrointestinal Tissues of a Natural Herbivore. *Parasitology* 140, 632–640

[23] Cornell, S.J. *et al.* (2008) Seasonality, Cohort-Dependence and the Development of Immunity in a Natural Host-Nematode System. *Proceedings of the Royal Society B: Biological Sciences* 275, 511–518

[24] Tschirren, B. and Richner, H. (2006) Parasites Shape the Optimal Investment in Immunity. *Proceedings of the Royal Society B: Biological Sciences* 273, 1773–1777

[25] Dowell, S.F. (2012) Seasonality - Still Confusing. *Epidemiology and Infection* 140, 87–90

[26] Altizer, S. *et al.* (2006) Seasonality and the Dynamics of Infectious Diseases. *Ecology Letters* 9, 467–484

[27] Fisman, D.N. (2007) Seasonality of Infectious Diseases. *Annual Review of Public Health* 28, 127–143

[28] Dowell, S.F. (2001) Seasonal Variation in Host Susceptibility and Cycles of Certain Infectious Diseases. *Emerging Infectious Diseases* 7, 369–374

[29] Hosseini, P.R. *et al.* (2004) Seasonality and Wildlife Disease: How Seasonal Birth, Aggregation and Variation in Immunity Affect the Dynamics of *Mycoplasma gallisepticum* in House Finches. *Proceedings of the Royal Society London Biological Sciences* 271, 2569–2577

[30] Langwig, K.E. *et al.* (2015) Host and Pathogen Ecology Drive the Seasonal Dynamics of a Fungal Disease, White-Nose Syndrome. *Proceedings of the Royal Society B: Biological Sciences* 282, 20142335

[31] Martin, L.B. *et al.* (2008) Seasonal Changes in Vertebrate Immune Activity: Mediation by Physiological Trade-offs. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 363, 321–339

[32] Scheiermann, C. *et al.* (2013) Circadian Control of the Immune System. *Nature Reviews Immunology* 3, 190–198

[33] Arjona, A. *et al.* (2012) Immunity's Fourth Dimension: Approaching the Circadian-Immune Connection. *Trends in Immunology* 33, 607–612

[34] Nelson, R.J. *et al.* (2002) *Seasonal Patterns of Stress, Immune Function, and Disease.* Cambridge University Press

[35] Wang, W. *et al.* (2011) Timing of Plant Immune Responses by a Central Circadian Regulator. *Nature* 470, 110–114

[36] Hua, J. (2013) Modulation of Plant Immunity by Light, Circadian Rhythm, and Temperature. *Current Opinion in Plant Biology* 16, 406–413

[37] Bolker, J. (2012) There's More to Life than Rats and Flies. *Nature* 491, 31–33

[38] Hegemann, A. *et al.* (2012) Immune Function in a Free-living Bird Varies Over the Annual Cycle, but Seasonal Patterns Differ Between Years. *Oecologia* 170, 605–618

[39] Owen-Ashley, N.T. and Wingfield, J.C. (2007) Acute Phase Responses of Passerine Birds: Characterization and Seasonal Variation. *Journal of Ornithology* 148, S583–S591

[40] Buehler, D.M. *et al.* (2008) Captive and Free-living Red Knots *Calidris canutus* Exhibit Differences in Non-induced Immunity that Suggest Different Immune Strategies in Different Environments. *Journal of Avian Biology* 39, 560–566

[41] Pedersen, A.B. and Babayan, S.A. (2011) Wild Immunology. *Molecular Ecology* 20, 872–880

[42] Cizauskas, C.A. *et al.* (2014) Gastrointestinal Helminths Affect Host Susceptibility to Anthrax through Seasonal Immune Trade-offs. *BMC Ecology* 14, 1–15

[43] Cizauskas, C.A. *et al.* (2015) Seasonal Patterns of Hormones, Macroparasites, and Microparasites in Wild African Ungulates: The Interplay Between Stress, Reproduction, and Disease. *PLOS ONE* In press

[44] Sepp, T. *et al.* (2010) Hematological Condition Indexes in Greenfinches: Effects of Captivity and Diurnal Variation. *Physiological and Biochemical Zoology* 83, 276–282

[45] Nguyen, K.D. *et al.* (2013) Circadian Gene Bmal1 Regulates Diurnal Oscillations of Ly6Chi Inflammatory Monocytes. *Science* 341, 1483–1488

[46] Keller, M. *et al.* (2009) A Circadian Clock in Macrophages Controls Inflammatory Immune Responses. *Proceedings of the National Academy of Sciences of the United States of America* 106, 21407–21412

[47] Silver, A.C. *et al.* (2012) The Circadian Clock Controls Toll-like Receptor 9-mediated Innate and Adaptive Immunity. *Immunity* 36, 251–261

[48] Casanova-Acebes, M. *et al.* (2013) Rhythmic Modulation of the Hematopoietic Niche through Neutrophil Clearance. *Cell* 153, 1025–1035

[49] Bellet, M.M. *et al.* (2013) Circadian Clock Regulates the Host Response to *Salmonella*. *Proceedings of the National Academy of Sciences of the United States of America* 110, 9897–9902

[50] Cermakian, N. *et al.* (2013) Crosstalk Between the Circadian Clock Circuitry and the Immune System. *Chronobiology International* 30, 870–888

[51] Janeway, C. *et al.* (2001) Evolution of the Innate Immune System. In *Immunobiology: The Immune System in Health and Disease*. Garland Science, New York, 5th edn.

[52] Nelson, R. and Demas, G.E. (1996) Seasonal Changes in Immune Function. *The Quartly Review of Biology* 71, 511–548

[53] Versteegh, M. *et al.* (2014) Genetic and Phenotypically Flexible Components of Seasonal Variation in Immune Function. *The Journal of Experimental Biology* 217, 1510–1518

[54] Ferrari, M.J. *et al.* (2008) The Dynamics of Measles in Sub-Saharan Africa. *Nature* 451, 679–684

[55] Fine, P.E. and Clarkson, J.a. (1986) Seasonal Influences on Pertussis. *International Journal of Epidemiology* 15, 237–47

[56] Altizer, S. *et al.* (2011) Animal Migration and Infectious Disease Risk. *Science* 331, 296–302

[57] Bartel, R.A. *et al.* (2011) Monarch Butterfly Migration and Parasite Transmission in Eastern North America. *Ecology* 92, 342–351

[58] Satterfield, D.A. *et al.* (2015) Loss of Migratory Behaviour Increases Infection Risk for a Butterfly Host. *Proceedings of the Royal Society B: Biological Sciences* 282, 20141734

[59] Begon, M. *et al.* (2009) Seasonal Host Dynamics Drive the Timing of Recurrent Epidemics in a Wildlife Population. *Proceedings of the Royal Society* 276, 1603–1610

[60] Martinez-Bakker, M. *et al.* (2014) Human Birth Seasonality: Latitudinal Gradient and Interplay with Childhood Disease Dynamics. *Proceedings of the Royal Society B: Biological Sciences* 281

[61] Peel, A.J. *et al.* (2014) The Effect of Seasonal Birth Pulses on Pathogen Persistence in Wild Mammal Populations. *Proceedings of the Royal Society B: Biological Sciences* 281

[62] Duke-Sylvester, S.M. *et al.* (2011) Strong Seasonality Produces Spatial Asynchrony in the Outbreak of Infectious Diseases. *Journal of the Royal Society, Interface* 8, 817–25

[63] Dorlien, A. *et al.* (2013) Impact of Birth Seasonality on Dynamics of Acute Immunizing Infections in Sub-Saharan Africa. *PLOS ONE* 8, e75806

[64] Bouma, H.R. *et al.* (2010) Hibernation: The Immune System at Rest? *Journal of Leukocyte Biology* 88, 619–624

[65] Prendergast, B.J. *et al.* (2002) Periodic Arousal from Hibernation is Necessary for Initiation of Immune Responses in Ground Squirrels. *American Journal of Physiology Regulatory, Integrative and Comparitive Physiology* 282, R1054–R1062

[66] Kurtz, C.C. and Carey, H.V. (2007) Seasonal Changes in the Intestinal Immune System of Hibernating Ground Squirrels. *Developmental and Comparative Immunology* 31, 415–428

[67] Shaman, J. *et al.* (2010) Absolute Humidity and the Seasonal Onset of Influenza in the Continental United States. *PLoS Biology* 8, e1000316

[68] Yu, H. *et al.* (2013) Characterization of Regional Influenza Seasonality Patterns in China and Implications for Vaccination Strategies: Spatio-Temporal Modeling of Surveillance Data. *PLoS Medicine* 10, e1001552

[69] Prendergast, B.J. (2011) Can Photoperiod Predict Mortality in the 1918-1920 Influenza Pandemic? *Journal of Biological Rhythms* 26, 345–52

[70] Meireles-Filho, A.C.A. and Kyriacou, C.P. (2013) Circadian Rhythms in Insect Disease Vectors. *Memórias do Instituto Oswaldo Cruz* 108 Suppl, 48–58

[71] Fleury, F. *et al.* (2000) Adaptive Significance of a Circadian Clock: Temporal Segregation of Activities Reduces Intrinsic Competative Inferiority in *Drosophila* Parasitoids. *Proceedings of the Royal Society London Biological Sciences* 267, 1005–1010

[72] Rund, S.S.C. *et al.* (2011) Genome-Wide Profiling of Diel and Circadian Gene Expression in the Malaria Vector *Anopheles gambiae*. *Proceedings of the National Academy of Sciences of the United States of America* 108, E421–E430

[73] Levi, T. *et al.* (2015) Accelerated Phenology of Blacklegged Ticks Under Climate Warming. *Philosophical Transactions of the Royal Society B: Biological Sciences* 370, 20130556

[74] Hawking, F. (1975) Circadian and Other Rhythms of Parasites. *Advances in Parasitology* 13, 123–182

[75] Mideo, N. *et al.* (2013) The Cinderella Syndrome: Why do Malaria-Infected Cells Burst at Midnight? *Trends in Parasitology* 29, 10–16

[76] Rothschild, M. and Ford, B. (1964) Breeding of the Rabbit Flea (Spilopsyllus cuniculi (Dale)) Controlled by the Reproductive Hormones of the Host. *Nature* 201, 103–104

[77] Mead-Briggs, A. and Rudge, A. (1960) Breeding of the Rabbit Flea, *Spilopsyllus cuniculi* (Dale): Requirement of a 'Factor' from a Pregnant Rabbit for Ovarian Maturation. *Nature* 187, 1136–1137

[78] Rothschild, M. and Ford, B. (1966) Hormones of the Vertebrate Host Controlling Ovarian Regression and Copulation of the Rabbit Flea. *Nature* 211, 261–266

[79] Glaudel, R. and Etges, F. (1973) The Effect of Photoperiod Inversion upon *Schistosoma mansoni* Cercarial Emergence from *Biomphalaria glabrata*. *International Journal for Parasitology* 3, 619–622

[80] Asch, H. (1972) Rhythmic Emergence of *Schistosoma mansoni* Cercariae from *Biomphalaria glabrata*: Control by Illumination. *Experimental Parasitology* 31, 350–355

[81] Toledo, R. *et al.* (1999) Production and Chronobiology of Emergence of the Cercariae of *Euparyphium albuferensis* (Trematoda: Echinostomatidae). *The Journal of Parasitology* 85, 263–267

[82] Cornford, E. *et al.* (1976) Physiological Relationships and Circadian Periodicities in Rodent Trypanosomes. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 70, 238–243

[83] O'Donnell, A.J. *et al.* (2011) Fitness Costs of Disrupting Circadian Rhythms in Malaria Parasites. *Proceedings of the Royal Society B: Biological Sciences* 278, 2429–2436

[84] O'Donnell, A.J. *et al.* (2013) Disrupting Rhythms in *Plasmodium chabaudi*: Costs Accrue Quickly and Independently of how Infections are Initiated. *Malaria Journal* 12, 372

[85] O'Donnell, A.J. *et al.* (2014) Correction: Disrupting Rhythms in *Plasmodium chabaudi*: Costs Accrue Quickly and Independently of how Infections are Initiated. *Malaria Journal* 13, 503

[86] Hotta, C.T. *et al.* (2000) Calcium-Dependent Modulation by Melatonin of the Circadian Rhythm in Malarial Parasites. *Nature Cell Biology* 2, 466–468

[87] De Bekker, C. *et al.* (2014) From Behavior to Mechanisms: An Integrative Approach to the Manipulation by a Parasitic Fungus (*Ophiocordyceps unilateralis s.l.*) of its Host Ants (*Camponotus* spp.). *Integrative and Comparative Biology* 54, 166–176

[88] Adamo, S.A. (2013) Parasites: Evolution's Neurobiologists. *The Journal of Experimental Biology* 216, 3–10

[89] Van Houte, S. *et al.* (2013) Walking with Insects: Molecular Mechanisms Behind Parasitic Manipulation of Host Behaviour. *Molecular Ecology* 22, 3458–3475

[90] Smith Trail, D. (1980) Behavioral Interactions Between Parasites and Hosts: Host Suicide and the Evolution of Complex Life Cycles. *The American Naturalist* 116, 77–91

[91] Thaiss, C.A. *et al.* (2014) Transkingdom Control of Microbiota Diurnal Oscillations Promotes Metabolic Homeostasis. *Cell* 159, 1–16

[92] Hayman, D.T.S. (2015) Biannual Birth Pulses Allow Filoviruses to Persist in Bat Populations. *Proceedings of the Royal Society B: Biological Sciences* 282, 20142591

[93] Ireland, J.M. *et al.* (2004) The Effect of Seasonal Host Birth Rates on Population Dynamics: The Importance of Resonance. *Journal of Theoretical Biology* 231, 229–38

[94] Ireland, J.M. *et al.* (2007) The Effect of Seasonal Host Birth Rates on Disease Persistence. *Mathematical Biosciences* 206, 31–45

[95] Helm, B. (2009) Geographically Distinct Reproductive Schedules in a Changing World: Costly Implications in Captive Stonechats. *Integrative and Comparative Biology* 49, 563–579

[96] Gibbs, H.C. (1973) Hypobiosis in Parasitic Nematodes—An Update. *Advances in Parasitology* 25, 129–174

[97] Hawking, F. (1967) The 24-Hour Periodicity of Microfilariae: Biological Mechanisms Responsible for Its Production and Control. *Proceedings of the Royal Society London Biological Sciences* 169, 59–76

[98] Cattadori, I.M. *et al.* (2007) Variation in Host Susceptibility and Infectiousness Generated by Co-Infection: The Myxoma-*Trichostrongylus retortaeformis* Case in Wild Rabbits. *Journal of the Royal Society Interface Interface* 4, 831–840

[99] Best, A. *et al.* (2009) The Implications of Coevolutionary Dynamics to Host-Parasite Interactions. *The American Naturalist* 173, 779–791

[100] Bull, J.J. and Lauring, A.S. (2014) Theory and Empiricism in Virulence Evolution. *PLoS Pathogens* 10, e1004387

[101] Alizon, S. *et al.* (2009) Virulence Evolution and the Trade-off Hypothesis: History, Current State of affairs and the Future. *Journal of Evolutionary Biology* 22, 245–259

[102] Navara, K.J. and Nelson, R.J. (2007) The Dark Side of Light at Night: Physiological, Epidemiological, and Ecological Consequences. *Journal of Pineal Research* 43, 215–24

[103] Garamszegi, L.Z. (2011) Climate Change Increases the Risk of Malaria in Birds. *Global Change Biology* 17, 1751–1759

[104] Paull, S.H. and Johnson, P.T.J. (2014) Experimental Warming Drives a Seasonal Shift in the Timing of Host-Parasite Dynamics with Consequences for Disease Risk. *Ecology Letters* 17, 445–453

[105] Trifonova, S.T. *et al.* (2013) Diurnal Redistribution of Human Lymphocytes and their Temporal Associations with Salivary Cortisol. *Chronobiology International* 30, 669–681

[106] Kirsch, S. *et al.* (2012) T-Cell Numbers and Antigen-Specific T-cell Function Follow Different Circadian Rhythms. *Journal of Clinical Immunology* 32, 1381–1389

[107] Bollinger, T. *et al.* (2010) The Influence of Regulatory T Cells and Diurnal Hormone Rhythms on T Helper Cell Activity. *Immunology* 131, 488–500

[108] Haus, E. and Smolensky, M.H. (1999) Biological Rhythms in the Immne System. *Chronobiology International* 16, 581–622

[109] Rahman, S.a. *et al.* (2015) Endogenous Circadian Regulation of Pro-Inflammatory Cytokines and Chemokines in the Presence of Bacterial Lipopolysaccharide in Humans. *Brain, Behavior, and Immunity* In press

[110] Naidu, K.S. *et al.* (2010) Inflammation in the Avian Spleen: Timing is Everything. *BMC Molecular Biology* 11, 1–13

[111] Zimmerman, L.M. *et al.* (2010) Variation in the Seasonal Patterns of Innate and Adaptive Immunity in the Red-eared Slider (*Trachemys scripta*). *The Journal of Experimental Biology* 213, 1477–1483

[112] Klink, M. *et al.* (2012) Seasonal Changes in Activities of Human Neutrophils In vitro. *Inflammation Research* 61, 11–16

[113] Buehler, D.M. *et al.* (2008) Seasonal Redistribution of Immune Function in a Migrant Shorebird: Annual-Cycle Effects Override Adjustments to Thermal Regime. *The American Naturalist* 172, 783–796

[114] Pap, P.L. *et al.* (2014) No Evidence for Parasitism-Linked Changes in Immune Function or Oxidative Physiology over the Annual Cycle of an Avian Species. *Physiological and Biochemical Zoology* 87, 729–739

[115] Baillie, S.R. and Prendergast, B.J. (2008) Photoperiodic Regulation of Behavioral Responses to Bacterial and Viral Mimetics: A Test of the Winter Immunoenhancement Hypothesis. *Journal of Biological Rhythms* 23, 81–90

[116] Deerenberg, C. *et al.* (2002) Spleen Size Variation During Long-Distance Migration in the Garden Warbler *Sylvia borin*. *Avian Science* 2, 217–226

[117] Myrianthefs, P. *et al.* (2003) Seasonal Variation in Whole Blood Cytokine Production After LPS Stimulation in Normal Individuals. *Cytokine* 24, 286–292

[118] Prendergast, B.J. *et al.* (2007) Winter Day Lengths Enhance T Lymphocyte Phenotypes, Inhibit Cytokine Responses, and Attenuate Behavioral Symptoms of Infection in Laboratory Rats. *Brain, Behavior, and Immunity* 21, 1096–1108

[119] Pyter, L.M. *et al.* (2005) Photoperiod Alters Hypothalamic Cytokine Gene Expression and Sickness Responses Following Immune Challenge in Female Siberian Hamsters (*Phodopus sungorus*). *Neuroscience* 131, 779–784

[120] Paynter, S. *et al.* (2015) Seasonal Immune Modulation in Humans: Observed Patterns and Potential Environmental Drivers. *The Journal of Infection* 70, 1–10

[121] Feigin, R.D. *et al.* (1969) Daily Periodicity of Susceptibility of Mice to Pneumococcal Infection. *Nature* 224, 379–380

[122] Shackelford, P.G. and Feigin, R.D. (1973) Periodicity of Susceptibility to Pneumococcal Infection: Influence of Light and Adrenocortical Secretions. *Science (New York, N.Y.)* 182, 285–287

[123] Martinaud, G. *et al.* (2009) Circadian Variation in Shedding of the Oocysts of *Isospora turdi* (Apicomplexa) in Blackbirds (*Turdus merula*): An Adaptative

Trait Against Desiccation and Ultraviolet Radiation. *International Journal for Parasitology* 39, 735–739

[124] Platt, T.R. *et al.* (2013) Circadian Egg Production by Echinostoma Caproni (Digenea: Echinostomatidae) in ICR Mice. *The Journal of Parasitology* 99, 179–182

[125] Jenkins, E. *et al.* (2006) Climate Change and the Epidemiology of Protostrongylid Nematodes in Northern Ecosystems: *Parelaphostrongylus odocoilei* and *Protostrongylus stilesi* in Dall's sheep (*Ovis d. dalli*). *Parasitology* 132, 387–401

[126] Janeway, C. *et al.* (2001) Principles of Innate and Adaptive Immunity. In *Immunobiology: The Immune System in Health and Disease*. Garland Science, New York, 5th edn.

[127] Lanier, L.L. (2013) Shades of Grey – The Blurring View of Innate and Adaptive Immunity. *Nature Reviews Immunology* 13, 73–74

[128] Raberg, L. (2014) How to Live with the Enemy: Understanding Tolerance to Parasites. *PLoS Biology* 12, e1001989

CHAPTER IV

Unraveling the Transmission Ecology of Polio

Preamble. This chapter is our published manuscript provided verbatim. The citation for this manuscript is:

***Martinez-Bakker, M.**, King, A. A., & Rohani, P. (2015). Unraveling the Transmission Ecology of Polio. PLoS Biology, Volume 13, Issue 6, June 2015, e1002172.*

Abstract. Sustained and coordinated vaccination efforts have brought polio eradication within reach. Anticipating the eradication of wild poliovirus (WPV), and the subsequent challenges in preventing its re-emergence, we look to the past to identify why polio rose to epidemic levels in the mid-20th century and how WPV persisted over large geographic scales. We analyzed an extensive epidemiological dataset, spanning the 1930s to the 1950s and spatially replicated across each US state, to glean insights into the drivers of polio's historical expansion and the ecological mode of its persistence prior to vaccine introduction. We document a latitudinal gradient in polio's seasonality. Additionally, we fitted and validated mechanistic transmission models to data from each state independently. The fitted models revealed that: (1) polio persistence was the product of a dynamic mosaic of source and sink popula-

tions; (2) geographic heterogeneity of seasonal transmission conditions account for the latitudinal structure of polio epidemics; (3) contrary to the prevailing "disease of development" hypothesis, our analyses demonstrate that polio's historical expansion was straightforwardly explained by demographic trends rather than improvements in sanitation and hygiene; and (4) the absence of clinical disease is not a reliable indicator of polio transmission because widespread polio transmission was likely in the multi-year absence of clinical disease. As the world edges closer to global polio eradication and continues the strategic withdrawal of the Oral Polio Vaccine (OPV), the regular identification of, and rapid response to, these silent chains of transmission is of the utmost importance.

Introduction

Poliovirus, like other members of *Picornaviridae*, usually generates mildly symptomatic infection. However, the clinical manifestation of polio, Acute Flaccid Paralysis (AFP), can result when the virus invades the central nervous system [1]. WPV is transmitted fecal-orally, and in the Northern Hemisphere exhibits seasonal epidemics in late summer and autumn [1, 2, 3]. Polio outbreaks continue today within this narrow seasonal window in Pakistan and Afghanistan [4, 5], but the seasonal transmission structure of polio remains unexplored.

Propelled by public support, the race for the polio vaccine during the post-World War II era led to the development of the Inactive Polio Vaccine (IPV) and the Oral Polio Vaccine (OPV) which reduced the global incidence to less than 0.1% of pre-vaccine levels [6]. Missing the 2014 goal of globally stopping WPV transmission has left eradication elusive, primarily due to political and social obstacles for effective vaccine distribution, including vaccine hesitancy and mistrust. In light of this—and the call for innovative solutions [7]—an understanding of polio's ecology can help guide alternative strategies. Looking toward eradication and beyond, a polio-free

world requires an understanding of the mode by which polio originally emerged and historically persisted. We contend that a retrospective study of the ecology of WPV in the absence of vaccine interventions can inform future planning, and may pinpoint vulnerabilities in WPV's epidemiology that could be leveraged for eradication.

Ironically, because of the success of polio vaccination, critical features of WPV transmission remain obscure. The low global incidence of polio (due to high vaccine coverage), in combination with the relative rarity of symptomatic infections, limits the amount of epidemiological data with which to study transmission. Furthermore, data limitations regarding vaccine coverage in developing countries confound transmission studies, making it difficult to disentangle the effects of the vaccines, demography, and transmission. Therefore, we took advantage of a dataset of unprecedented size and resolution in both space and time to gain insights into the drivers of polio's historical expansion and the ecological mode of its persistence in the pre-vaccine period.

We present analyses of spatially-replicated incidence reports from the pre-vaccine era in the US, and built mechanistic transmission models which incorporate these data to reconstruct the unobservable infection dynamics. Our analyses allow us to dissect three axes of polio epidemiology: (i) geographical and seasonal variation in transmission, (ii) the role of demography in determining incidence, and (iii) the mode by which polio persists.

Methods

Data.

We examined monthly polio case reports (January 1931–December 1954) from the US Public Health Service Morbidity and Mortality Weekly Reports as compiled by [8] and the CDC for each of the 48 contiguous US states and the District of Columbia (Fig 4.1A–B). Prior to 1945, paralytic polio cases were primarily reported

in the US [3, 1]; however, during the later period of this study, non-paralytic cases comprised $\geq$ 40% of reported cases in populous cities such as New York, Detroit, Kansas City, and Sacramento [9]. In addition to polio case data, we obtained state-level births from 1931 onward from US Vital Statistics, and state population sizes from the Population Distribution Branch of the US Census Bureau. Data from US Vital Statistics are housed in the CDC online repository: National Center for Health Statistics, Products, Vital Statistics. The US Census Bureau data were obtained from their Population Estimates Repository, historical data Pre-1980. The polio data set—with cases detailed weekly—has now been independently digitized and is freely available and maintained online through the University of Pittsburgh Project TYCHO. Birth data were not available for Texas and South Dakota beginning in 1931, but began in 1932 and 1933, respectively. For exploratory analyses, we quantified the relationship between disease fadeouts and population size. A threshold of 3 months without a reported infection was chosen to define a fadeout [10]. The portion of fadeout months was taken as the ratio of fadeout months to total months in Fig 4.1C. To estimate spatial synchrony, we used the nonparametric spatial correlation function [11, 12]. To measure the relative timing of polio epidemic peaks for each state and each year, the 1 yr wavelet band phase angle was computed [13] and used to rank states earliest to latest based on their epidemic peak timing.

Models.

We constructed a dynamic stochastic model with components incorporating polio transmission, immunity, seasonality, and symptomatology along with empirical population sizes and birth rates. Birth rates displayed prominent seasonal, secular, and geographical trends (Appendix B)[14]. We utilized Partially Observed Markov Process (POMP) models which are suited for dealing with epidemiological data where the state variables (susceptible, infected, and recovered individuals) were not observed

in the data; rather the infected individuals were partially observed through clinical case reports. For our process models, we used seasonally-forced stochastic monthly discrete-time SIR models where transitions followed a Poisson process. The infectious period was fixed at 1 month, because multiple studies have found the duration of shedding to be 3–4 weeks [15]. Infection-derived immunity was assumed to be life-long [16, 17]. The models contained 6 classes of infants susceptible (S_i^B) to infection. These infant classes contained 0–1 month olds, 1–2 month olds, etc. up to 6 month olds. Models had a single infected class for infants (I^B). The older age class which contained individuals > 6 months of age, had its own susceptible (S^O) and infected class (I^O). The onset of polio symptoms ranges from 5–35 days post-exposure, with a mean of 12 days [18]; therefore, we assumed reporting of symptomatic infections occurred within the 1 month infectious period. We modeled polio reporting explicitly and consistent with clinical evidence, assumed that maternal antibodies protected from severe disease and resulted in unreported infant infections[19, 20, 21, 22, 23]. Thus, we assumed that infections in individuals under 6 months of age were asymptomatic, and only individuals over 6 months of age could be symptomatic and reported as a clinical case. See model schematic in Fig 4.2A. The force of infection was modeled as,

$$\lambda_t = \left(\beta_t \frac{I_t^O + I_t^B}{N_t} + \psi \right) \epsilon_t. \tag{4.1}$$

The first term of the force of infection, $\beta_t \frac{I_t^O + I_t^B}{N_t}$, represents transmission that occurred locally by individuals infected in the state at time t. Whereas, the second term, ψ, encompasses WPV that arose in the population from external sources that were divorced from the local infection dynamics. ψ placed a lower bound on the force of infection, allowing WPV to rebound in the face of local extinction. We interpret ψ as indicating virus imported from other geographic regions; however, it could also be interpreted as representing a small number of individuals in the population that shed WPV for an extended period, or environmental sources helping WPV persist

over the winter. The transmission parameter, β_t, was parameterized using a B-spline, providing it the flexibility to have a constant or seasonal transmission rate. There was seasonality, but no interannual variation, in the transmission rate,

$$\beta_t = \exp \sum_{i=1}^{6} \beta_i \xi_{i_t}. \tag{4.2}$$

Here, each ξ_{i_t} is a periodic B-spline basis with a 1 year period. The process noise, ϵ_t, was gamma distributed with mean 1 and variance that scaled to account for both environmental and demographic stochasticity; refer to Appendix B for further details). We assumed cases were drawn from a rounded, left censored normal distribution with a mean report rate of ρ_t, and dispersion parameter τ,

$$\text{cases}_t = \text{round}(x_t), \qquad x_t \sim \text{normal}(\rho_t I_t^O, \tau I_t^O). \tag{4.3}$$

For calculating the likelihood, we used a binned-normal probability density. Full model details are found in Appendix.

We fitted SIR models (one for each US state) to data independently using Maximization by Iterated particle Filtering (MIF) in the R package pomp [24, 25, 26]. For each US state, we estimated 14–15 parameters. The parameters estimated were: 6 seasonal transmission parameters (β_i), 3 parameters accounting for process and measurement noise, 3 initial conditions for the older age class, the external contribution to the force of infection (ψ), and 1–2 report rates (ρ_t). MIF is a simulation-based likelihood method for parameter estimation. The basis of MIF is particle filtering, which integrates state variables of a stochastic system and estimates the likelihood for fixed parameters. Instead of fixing parameters, MIF varies them throughout the filtering process and selectively propagates particles (i.e., parameter sets) that have the highest likelihoods. By initializing MIF throughout parameter space we estimated the shape of the likelihood surface for each US state and identified the Maximum

Likelihood parameter Estimates (MLEs). MIF was initialized from 1 million parameter sets for a global search, followed by additional phases of increasingly localized searches, which included profiling. In total, for each US state, MIF was initialized from > 10000 locations in parameter space to estimate the shape of the likelihood surface and identify the MLEs.

Prior to 1945, non-paralytic polio cases were rarely included in our data, but the reporting of non-paralytic polio became increasingly common [3, 1]. Thus, we tested an optional parameter to account for increased representation of non-paralytic polio in clinical cases data. We estimated two report rates, one for the pre-baby boom era and another for the baby boom era, and discriminated between models with and without time-varying reporting using Akaike Information Criterion (AIC). Profiles were constructed for the two versions of the model, one in which the report rate was constant through the entire time-period, and one in which the report rate increased during the baby boom era. For each state AIC was used to discriminate between constant and time-varying reporting, and the maximum likelihood estimates (MLEs) were drawn from the appropriate two dimensional profile. Inference was performed using the data from May 1932 to January 1953, with the exception of South Dakota and Texas, whose inference began in May 1933 and 1934, respectively. Inference was initiated in May of the year following the first full year of available data to allow us to construct the infant initial conditions directly from birth data. For model validation, the last two epidemic years were set aside for forecasting. Full details are provided in Appendix B. Likelihood profiles were constructed for each state (example in Fig 4.2B, all others in Appendix B).

To quantify model-data agreement, we evaluated the accuracy of one-step-ahead predictions for all 49 states, both for data used in model parameterization (Fig 4.3B) and for out-of-fit data (Fig 4.3C). Because of correlations between states (which vary significantly in size) and mean incidence, simple linear regression is not appropriate

for accessing model-data agreement; therefore, generalized R^2 was calculated to quantify the proportion of the variance explained by the model relative to that explained by state alone. We calculated the generalized R^2 for the one-step-ahead predictions and out-of-fit predictions (See Appendix B for details). For Fig 4.4A–B, infections were reconstructed using particle filtering means, and the reconstruction was limited to data beginning in Jan 1935, because 1935 is the first full year for which we have the models parameterized for all US states. Following model validation and infection reconstruction, the fitted models were used as simulation tools to explore polio infection dynamics. In Fig 4.4D–E, we used 500 simulations for each state from 1935 through 1954. In Fig 4.4D, we present the state-specific probability of extinction by examining 500 realizations of the fitted models. Specifically, we calculated the annual probability of polio extirpation during the off-season (Dec-May), and averaged across years. Similarly, in Fig 4.4E the minimum number of infections during each off-season was based on 500 simulations. For each simulation, the annual minimum number of infections was identified, the median was taken across the 500 simulations, and averaged across years. In order to identify the covariates and epidemiological parameters that influenced the the number of trough infections—a measure of WPV persistence—we regressed trough infections with various covariates and parameters; results shown in Fig 4.5. In Fig 4.6B–D, distributions were generated by characterizing observations across 500 simulations per state. All simulations and data used for producing the figures in this manuscript are available in Appendix B.

Results

Polio's Seasonality & Latitudinal Gradient.

In the mid-20th century, polio outbreaks in the US were strongly seasonal. Epidemic peaks typically occurred between August–October (Appendix B); but the mag-

nitude was highly variable among states. In the transition from the pre-baby boom era (1931–1945) to the baby boom (1946–1954), epidemics increased in size and became more regular (Fig 4.1A–B). Winter troughs were frequently marked by consecutive months without reported cases. During the baby boom, the frequency of these local fadeouts diminished (Fig 4.1C) while epidemics became more tightly synchronized (Fig 4.1D). There was a striking latitudinal gradient in the timing of epidemics across the entire US (Fig 4.1E–F, Appendix B). Two broad classes of mechanisms can give rise to such a pattern. Seasonal movement of pathogen from southern populations can generate a traveling wave, which has previously been observed in measles [27], dengue [28], influenza [29, 30], and pertussis [31]. Alternatively, the pattern may indicate latitudinal gradients in demographic (e.g., birth rates [14, 32]) and/or environmental factors associated with transmission.

Model fit.

Our extensive search of parameter space resulted in the MLEs for each parameter. To quantify the shape of likelihood surface along two parameter dimensions we identified as important (i.e., the report rate, ρ_t, and the external contribution to the force of infection, ψ), we constructed two dimensional likelihood profiles for each state. Two-dimensional profiles by definition have fixed parameter values along two dimensions of parameter space, while the likelihood is maximized along all other parameter dimensions. There were 12 states that had constant reporting (i.e., the same report rate during the pre-baby boom and baby boom era). Fig 4.3A illustrates that the fitted models generate epidemic trajectories that display (1) the seasonal characteristics of polio, and (2) the large amount of interannual variation in epidemic size. Importantly, the fitted models faithfully reproduce observed dynamics. In particular, the seasonality, epidemic shape, interannual variability in epidemic magnitude, and the increase in incidence during the baby boom are captured by the models. Model

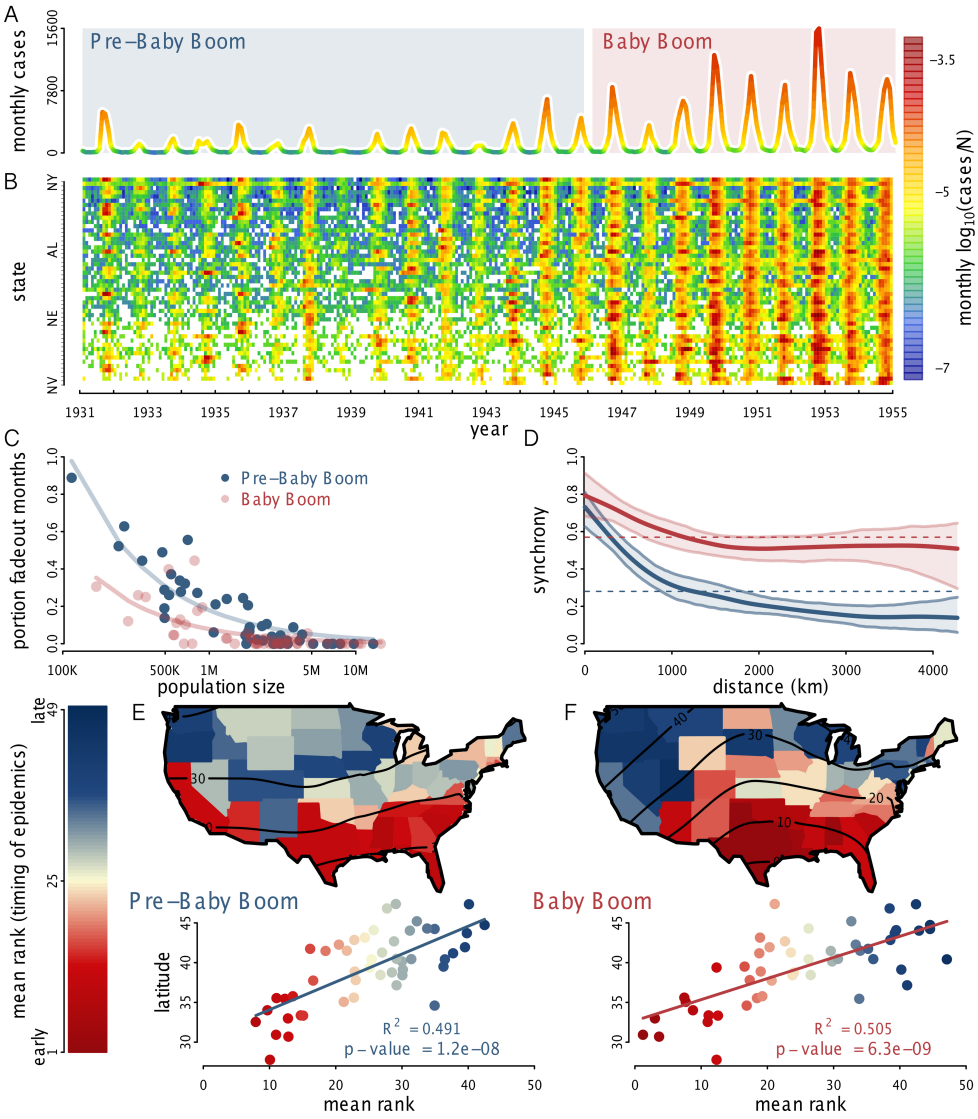


Figure 4.1: **Spatiotemporal patterns in US polio incidence.** (A) Total monthly case reports, 1931–1954, color-coded by per-capita incidence. (B) Log-transformed per-capita incidence by state, ranked top-to-bottom by population size. (C) Disease fadeout frequency as a function of state population size, during the pre-baby boom and the baby boom. The lines represent fitted negative exponential curves, which tended toward zero. (D) Pairwise epidemic synchrony between states during the pre-baby boom and the baby boom. Mean and 95% bootstrap confidence envelope shown. (E, F) Relative timing of polio epidemic peaks during the (E) pre-baby boom and (F) baby boom eras. Color indicates mean rank of each state across years; lower rank indicates earlier epidemic peak. Below each map, relative timing is regressed on latitude. Lower latitude states had significantly earlier epidemic peaks.

fit was formally validated using one-step-ahead predictions for all 49 states (Fig 4.3B) and out-of-fit predictions (Fig 4.3C), which indicate good agreement between models and data. Furthermore, geographical structure in the timing of observed epidemics is captured by the fitted models (Fig 4.3D). State-specific examples of one-step-ahead predictions and out-of-fit predictions are shown in Fig S2–S3.

Explaining the Latitudinal Gradient.

We hypothesized that the latitudinal gradient in epidemic timing was driven by either: (1) geographic variation in transmission due to environmental factors that modulated transmission, (2) the geographic trend in birth seasonality in the US (detailed in [14]), or (3) the movement of pathogen from south to north.

In support of hypothesis 1 (i.e., environmental factors), we identified a spatial pattern in the phase of seasonal transmission (Fig 4.3E–G, Appendix B). States with earlier epidemics had an earlier peak in the seasonal transmission rate in the fitted models. Interestingly, due to polio's long infectious period, peaks in transmission preceded incidence peaks by 1–2 months. States not only varied geographically in the timing of the transmission peak, but also in the wintertime transmission trough depth and trough duration (Fig 4.3G).

Epidemiological theory indicates that birth seasonality can have important dynamical consequences for childhood diseases [33, 14, 34]. To test hypothesis 2 (i.e., birth seasonality), we carried out a comparison of the fitted models with and without birth seasonality. Simulations of both models expressed the latitudinal gradient (Fig S5). Therefore, birth seasonality is not necessary to explain the polio gradient because geographic variation in transmission is sufficient. We attribute the negligible effect of birth seasonality on polio incidence to the low amplitude of birth seasonality, which was $\sim 10\%$ in the US at this time.

We suggest that hypothesis 3 (i.e., pathogen movement) is an unlikely explanation

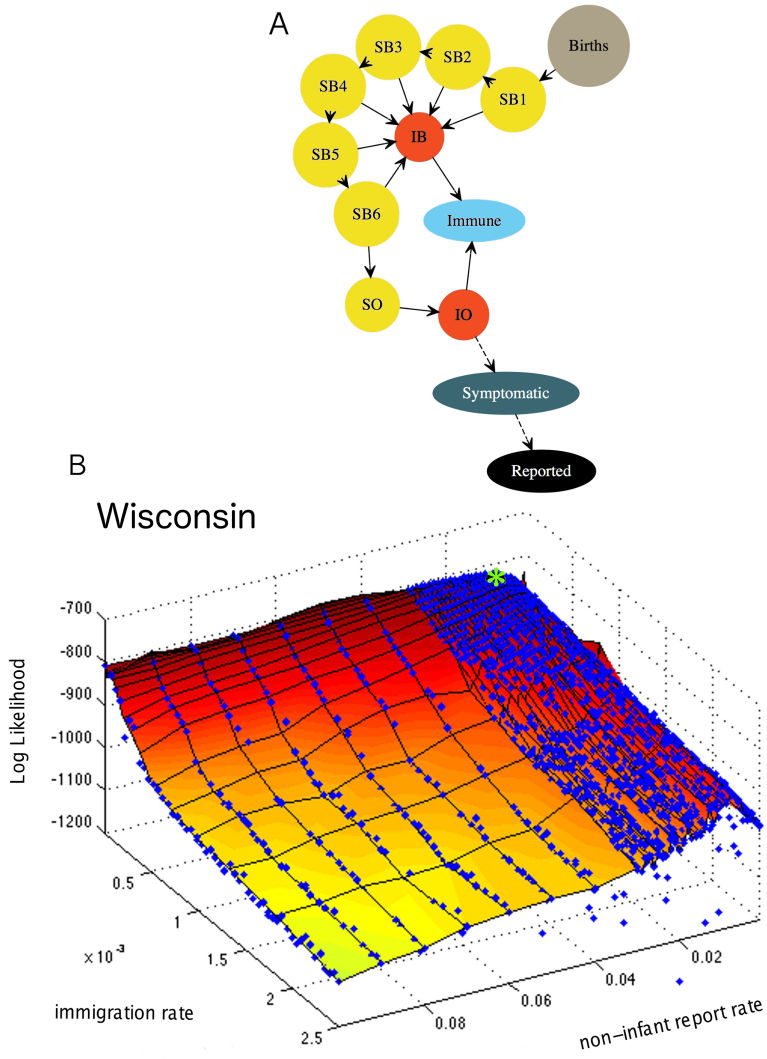


Figure 4.2: **Model schematic and example likelihood profile**. (A) Births enter the first susceptible infant class, S_1^B. Susceptible infants of age 0–6 months, S_{1-6}^B, are susceptible to infection, but are protected from symptomatic disease by maternal antibodies. Susceptible individuals over 6 months of age are in the S^O class. Infected infants and non-infant infections are in I^B and I^O, respectively. Infected individuals over 6 month of age, I^O, can have symptomatic illness and subsequently be reported as a clinical case with mean probability ρ_t. ρ_t is a composite parameter that represents the probability of symptoms and reporting. (B) Likelihood profile for the report rate, ρ_t, of non-infant infections and the immigration rate, ψ, for the state of Wisconsin. MLE indicated by green asterisk. The report rate for Wisconsin was constant through time.

of the latitudinal gradient. If the latitudinal gradient were a wave of pathogen movement, it would require a high wave speed, which we see as incompatible with transport

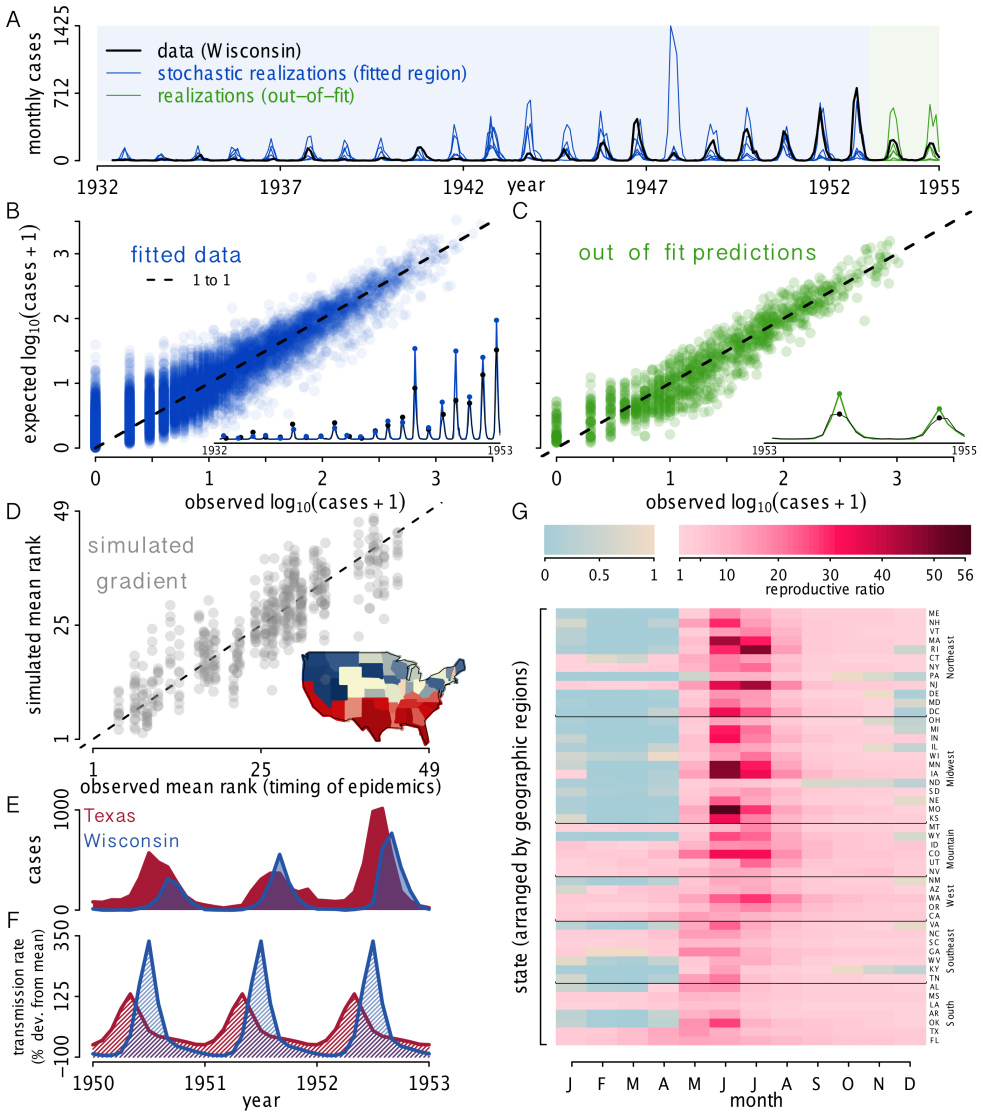


Figure 4.3: **Fitted Model and Seasonality.** (A) Observed data (black, shown for Wisconsin) and three stochastic simulations from the MLE (blue/green) highlight that both observed and simulated polio epidemics have a large amount of interannual variation in size, but a narrow seasonal window. Fitted and out-of-fit data regions are indicated by blue and green, respectively. (B) Model validation showing observed $\log_{10}$(cases) versus expected $\log_{10}$(cases) for fitted data and (C) out-out-fit predictions for all 49 states. Expected cases are one-step-ahead predictions from the fitted models. Insets show observed cases (black) and expected cases (blue/green) for Wisconsin. Fitted data include May 1932–Jan 1953 for all states except South Dakota and Texas, whose covariate data limited our inference to begin in May 1933 and 1934, respectively; out-of-fit data spanned Jan 1953–Dec 1954. The generalized $R^2 = 0.76$ for the fitted data and $R^2 = 0.61$ for out-of-fit data, calculated on the natural scale, while data are plotted on a log-scale for visualization. (D) Observed vs. simulated mean rank of epidemic timing based on 10 realizations of the fitted models. Inset shows the latitudinal gradient from one simulation; colors match Fig 4.1E–F. (E) Monthly polio cases in Texas and Wisconsin, and (F) the MLE transmission rates. Epidemics occured earlier in southern states than northern states because the seasonal peak in transmission occured earlier at lower latitudes. (G) MLEs of the seasonal transmission rate for each state organized by geographic region; in our models this represents the reproductive ratio. The reproductive ratio varies both seasonally and geographically, with some states having a reproductive ratio < 1 during the wintertime off-season.

of pathogen across the landscape. The pattern in Fig 4.1E–F corresponds to a wave traveling approximately 1200 km/mo. For comparison, waves in pertussis have been estimated to travel 110–320 km/mo [31]; waves in dengue appear to move 150 km/mo [28]; and the measles wave-speed in the UK was estimated at 20 km/mo [27]. A polio wave that is 10-fold faster than pertussis in the US is difficult to justify, and unnecessary because our fitted models support hypothesis 1. Thus, we have identified that polio's latitudinal gradient is driven by geographic variation in transmission, and we are left with an unidentified seasonal driver that modulates transmission.

While geographical variation in birth seasonality was insufficient to explain the latitudinal gradient seen in epidemic timing, birth seasonality had a small but observable effect on the simulated incidence of infant infections. To quantify the influence of birth seasonality on infant infections, we compared simulations of the fitted models to simulations for which seasonal fluctuations in births were removed. In the presence of birth seasonality, infant infection incidence was often higher (Fig S6); however, this did not affect the incidence of disease directly, and no indirect effect was observed.

Symptomatology.

It is well known that AFP incidence represents a small fraction of true WPV prevalence [35, 36]. Reassuringly, our independent estimates from the incidence data agree: our MLEs indicate that typically $< 1\%$ of polio infections were reported. We assumed that infected infants under 6 months of age were asymptomatic, due to protection by polio-specific maternal antibodies. The report rate for individuals not maternally-protected was 0.75% (averaged across states) in the pre-baby boom era, and rose to 1.4% in the baby-boom era, with considerable variation across states (Appendix B). Overall, we estimate that there were often 1+ million annual infections in the US; though only 2000–57000 cases were reported every year (Fig 4.4A). Our results are in line with a 1948 serology-based study in North Carolina, which estimated

62–175 subclinical polio infections per paralytic case [37].

Spatiotemporal Heterogeneity in the Reproductive Ratio.

The fitted models revealed vast seasonal and spatial heterogeneity in WPV's reproductive ratio. Fig 4.3G shows large seasonal fluctuations in the reproductive ratio within each US state. Several states maintained a reproductive ratio > 1 throughout the year. In contrast, 28 states had reproductive ratios that fell < 1 for 4–5 months from Dec–Apr.

States in the Northeast and Midwest had extreme seasonal variation in their reproductive ratio. Deep winter troughs in transmission in the Northeast and Midwest often had several consecutive months with a reproductive ratio < 1. In contrast, at the peak of transmission in June and July, these same states had a reproductive ratio > 20. Interestingly, each geographic region other than the Midwest, had at least one state that maintained a reproductive number > 1 throughout the year. Southern states typically maintained an intermediate transmission rate throughout the year.

Epidemic Emergence.

Our analyses provide a new perspective on polio's historical emergence. Commonly described as a "disease of development", polio's emergence has been ascribed to improved hygiene that reduced transmission and pushed the burden of infection onto children more susceptible to paralytic polio. This explanation requires that reduced transmission raised the mean age of infection and therefore the risk of AFP [1]. Our results suggest the marked increase in polio incidence from the 1930s to the 1950s was a straightforward consequence of increased birth rates (Fig 4.4B–C) and that hygiene effects on transmission are not required to explain polio's rise to epidemic levels. Since polio's epidemic emergence was captured in the models due to the changing birth rate, we did not explicitly test reductions in the transmission

rate as an additional contributor to epidemic size and we cannot completely rule out trends in transmission as a contributing factor. While the disease-of-development explanation has also been questioned on other grounds [22], changes in hygiene and sanitation could have contributed to the initial emergence of polio, which occurred from the late 1800s to the early 20th century.

WPV Persistence.

Polio cases were consistently observed throughout the US during the period of this study. We hypothesized: (a) WPV persisted locally in each state, or alternatively, (b) WPV regularly went locally extinct and re-invaded from elsewhere. Due to polio's high asymptomatic infection ratio, distinguishing between these two mechanisms of persistence cannot be done using reported cases alone, since WPV may be present during the off-season even in the absence of clinical cases. In order to determine which of these two persistence mechanisms was the likely explanation of continued infection, we simulated the fitted models and characterized the dynamics of the process models (i.e., the unobserved infection dynamics rather than the observable disease dynamics). We focused on determining whether infections persisted during the wintertime off-season, or if extinction and re-invasion occurred. In particular, we assessed (i) the average annual probability of an extinction event in each state, which results from diminished local transmission and (ii) the annual minimum number of infections. Fig 4.4D–E depicts the geographic variation in these quantities. Some states experienced frequent local extinction during the off-season, followed by re-colonization: we consider these "sink" populations. In contrast to sink states, a few states maintained infections year-round: these we define as "source" populations. The majority of states, however, were neither consistently sources nor sinks, because even sink states had frequent overwintering of WPV. The fitted models suggest that WPV underwent extinction and re-colonization in the classic metapopulation sense.

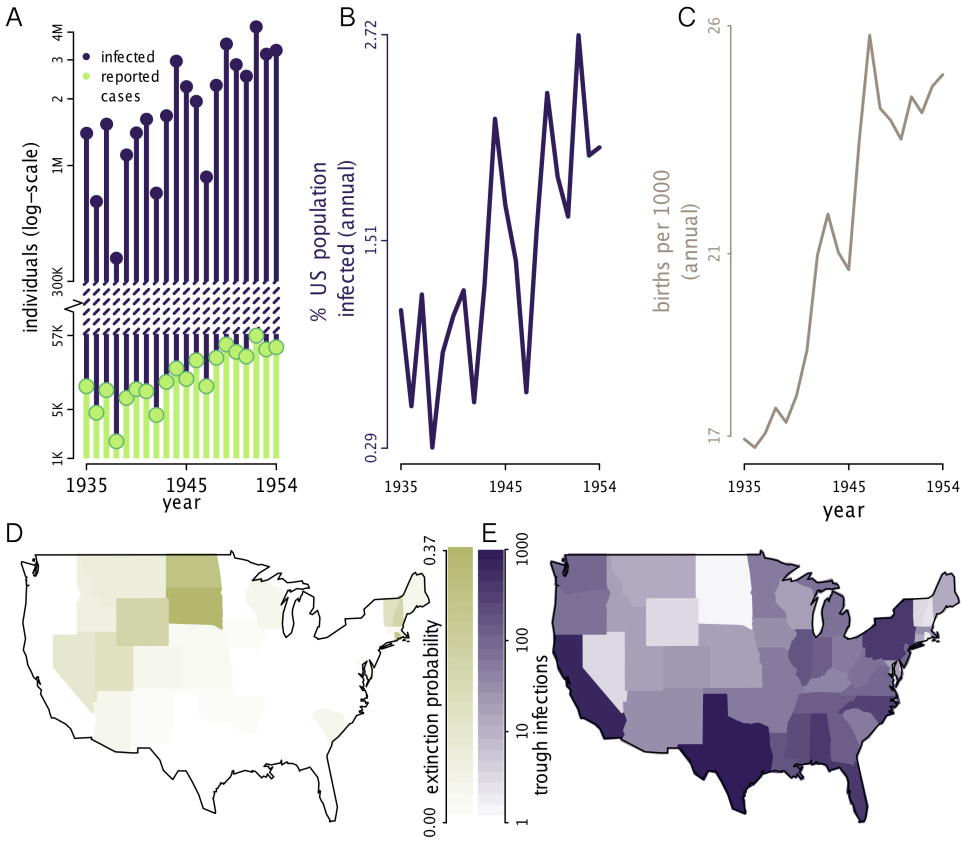


Figure 4.4: **Epidemic Emergence and Source-Sink Dynamics.** (A) Annual number of infected individuals in the US in contrast to the small number of reported cases. Annual infections were reconstructed for the US using particle filtering means. The particle filtering mean is the expected value at time t, given the data up to time t: $\mathbb{E}(\mathbf{X}_t \mid \text{cases}_t)$. (B) Annual infections in the US represented as the percent of the population. Reconstructed infections show an increase in infection incidence that accompanies (C) the increase in the birth rate. (D) Simulated WPV extinction probability. The probability of extinction measured as the mean annual probability of observing an extinction during the off-season (Dec–May). "sink" populations are those states with a high extinction probability. (E) Simulated trough infections. Trough infections indicate the minimum number of infections during off-seasons. For each state, the median was taken across simulations and averaged across years. "source" populations are those that maintain a high number of infections. Panels D–E were constructed using the 500 stochastic simulations for each state.

Source-Sink Population Predictors.

We explored characteristics that contributed to states having been WPV sources versus sinks. We used simulated trough infections, shown in Fig 4.4E, as the indicator of a source versus a sink. States that maintained a high number of trough

infections enabled WPV to persist through the off-season; whereas, states with a low number of trough infections were likely to have experienced regular WPV extinction. State population size accounted for 65% of the variation in the number of trough infections, Fig 4.5A. We used multiple regression models to determine whether the (i) mean birth rate, (ii) amplitude of birth seasonality, (iii) immigration rate, (iv) seasonal minimum reproductive ratio, and/or (v) seasonal amplitude of the reproductive ratio explained the residual variation in trough infections, after controlling for population size. The mean birth rate and amplitude of birth seasonality had a negligible impact on the residual variation in trough infections; therefore, they were removed from the multiple regression model. A multiple regression model with the immigration rate, seasonal minimum reproductive ratio, and the seasonal amplitude of the reproductive ratio explained 56% of the residual variation in trough infections, Fig 4.5B. Interestingly, even though there were no clear geographic patters of source-versus-sink localization, Fig 4.4D–E, there was strong geographic clustering in the minimum reproductive ratio, Fig 4.5C, demonstrating that even though source-sink predictors display geographic clustering, the combination of predictors can generate a source-sink mosaic. We found that after accounting for population size, states with a higher immigration rate had more trough infections, Fig 4.5D–E. States with a higher transmission amplitude, however, had fewer trough infections; we interpret this as being due to susceptible depletion followed by deep infection troughs in states with a high transmission amplitude, Fig 4.5D. The minimum reproductive ratio had a positive relationship with tough infections; states that maintained a reproductive ratio > 1 during the off-season tended to have more trough infections during the off-season, Fig 4.5E.

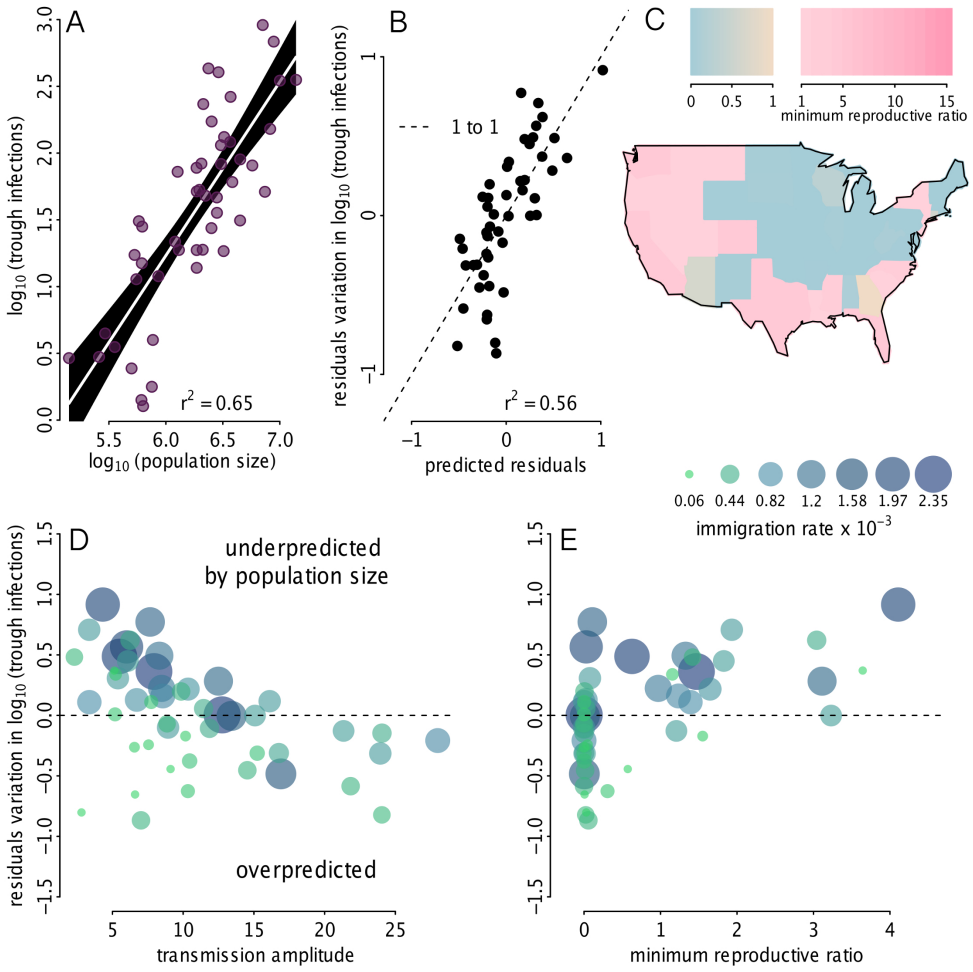


Figure 4.5: **Source-Sink Population Predictors.** (A) Linear regression of state population size versus simulated trough infections, both on a $\log_{10}$ scale. Trough infections are those shown in Fig 4.4E. (B) Residuals from the regression of population size versus trough infections were used as the dependent variable in the multiple regression model, where the predictors were: the state's seasonal minimum reproductive ratio, the immigration rate, and the seasonal amplitude of the reproductive ratio, measured as half the peak-tough difference in the reproductive ratio. Plot shows on the y-axis, the residuals, r_i, from panel A, along with the prediction of the residuals based on the multiple regression $r_i = b_0 + b_1 \min(R_t^i) + b_2 \frac{\max(R_t^i) - \min(R_t^i)}{2} + b_3 \psi_i$, where R_t^i is the reproductive ratio, ψ_i is the immigration rate, and i indicates the state. Taken together, panels A and B demonstrate that the predictors of a source versus sink are: the population size, the minimum reproductive ratio, the amplitude of the reproductive ratio, and the immigration rate. (C) Map of the seasonal minimum reproductive ratio showing geographic clustering. (D) The residuals, r_i, versus the seasonal amplitude of the reproductive ratio (i.e., the transmission amplitude), point size and color indicate the immigration rate, ψ_i. (E) The residuals, r_i, versus the seasonal minimum reproductive ratio, point size and color indicate the immigration rate, ψ_i.

Silent Infections.

Disease eradication programs face the significant challenge of verifying success in the presence of asymptomatic infections. Typically, a criterion for success is the absence of disease for an extended period; however, the utility of this criterion is questioned when the symptomatic cases reported are only the tip of the iceberg in terms of infection. Using our fitted models, we explored the reliability of absence-of-disease as an indicator of WPV extinction. Due to widespread subclinical infections, there was a stark contrast between the simulated number of polio infections and clinical cases (Fig 4.4A). This contrast (i.e., the disconnect between infections and clinical cases), can lead to epidemiological scenarios where absence-of-disease is uninformative. In our models, WPV persistence was achieved by one of two mechanisms (1) local unbroken chains of transmission, or (2) local extinction followed by rapid reintroduction. For each of these two mechanisms, we found that clinical case data can be misleading, as outlined in Table 4.1. For instance, if WPV circulated at low levels of infection, extended absence of clinical cases could lead to the conclusion that WPV was locally eradicated. Similarly, if local extinction of WPV occurred, and was quickly followed by reintroduction and clinical cases, local extinction could go unrecognized, potentially misdirecting targets for control (e.g., to focus on sink populations rather than source populations).

By simulating our fitted models we identified extended periods absent of disease and used these periods to quantify the number of silent infections (Fig 4.6B–C). We observed that if infections were maintained at relatively low numbers (i.e., under 100 infections per month), then WPV could circulate silently for > 30 months (Fig 4.6B). The silent circulation of WPV can result in thousands of infections before a single reported case is observed (Fig 4.6C). Our models assumed homogeneous mixing within each state, and it is important to recognize that different mixing patterns could increase or decrease the duration of silent infections. Due to the silent circulation of

polio, it is difficult—and perhaps indefensible—to use clinical case data (i.e., without fitted models) to evaluate WPV persistence. We simulated the fitted models to quantify the distribution of cases observed during periods with WPV extinction (Fig 4.6D). The distribution of cases surrounding WPV extinctions is fairly symmetric due to the reintroduction of WPV following extinction. Therefore, we conclude that, in the face of rapid reintroduction following WPV extinction, case data cannot be used to identify extinction events. Though it is desirable to use fitted models to identify signals of extinction, and apply this knowledge to case data, it would require extensive evaluation of silent circulation.

	Local Persistence of WPV	**Local Extinction & Reintroduction**
Extended absence of disease	Disease data are uninformative, and potentially misleading, because WPV is circulating silently via subclinical infections	Disease data reflect that WPV goes extinct and is reintroduced
Disease observed regularly	Disease data reflect that WPV persists and transmission is ongoing	Disease data are uninformative because they mask that WPV goes extinct and is reintroduced

Table 4.1: **Four scenarios for the relationship between WPV infections and clinical disease.** Local persistence of polio—within a state, region, or country—occurs when WPV overwinters during the off-season and the transmission chain is unbroken year-round. In contrast, local extinction and reintroduction occurs when WPV goes extinct during the off-season, breaking the chain of transmission; a new transmission chain begins when WPV is reintroduced from elsewhere. Discriminating among these scenarios is necessary for planning eradication strategies in endemic regions.

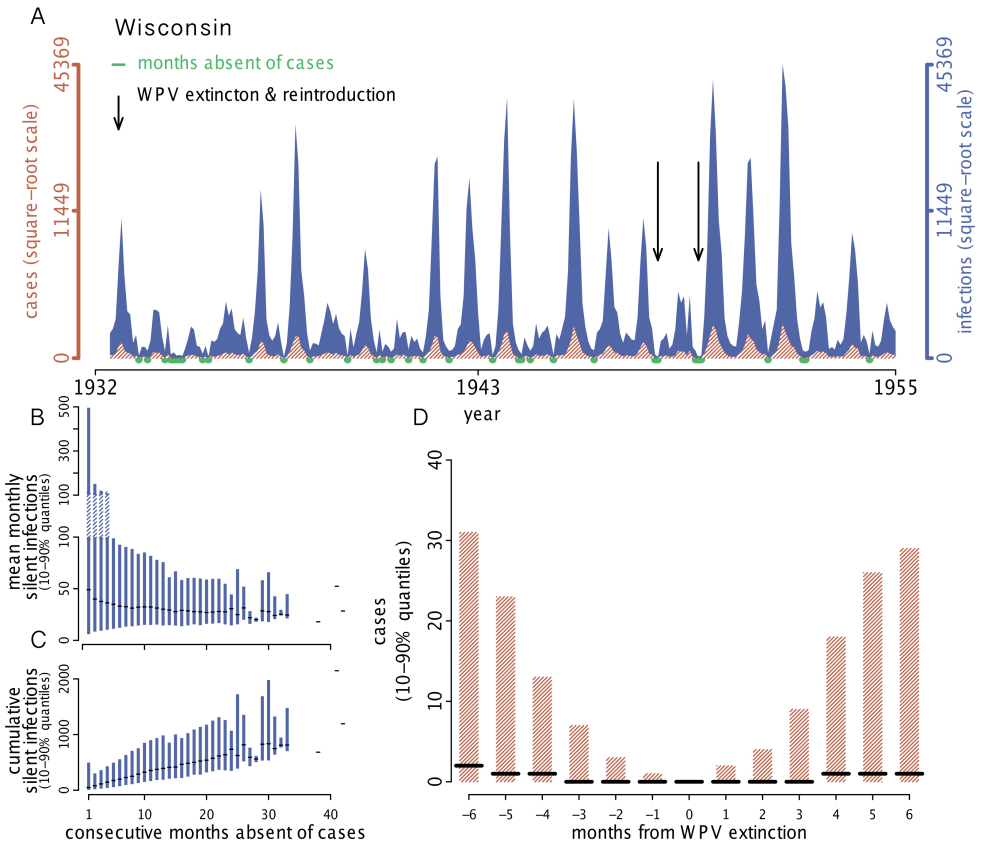


Figure 4.6: **Persistence Mechanisms.** (A) Example of simulated infections and cases for Wisconsin. Months absent of reported cases are indicated in green. During periods when the disease is absent, WPV infections are often silently transmitted in the population. In this simulation, there were two instances (indicated by arrows) when the local chain of transmission was broken and WPV went locally extinct, but quickly rebounded due to reintroduction. This example illustrates the two polio persistence mechanisms observed throughout the US, which are (i) local WPV persistence via unbroken chains of transmission, and (ii) WPV extinction and reintroduction. (B) Distributions of mean monthly silent infections during periods absent of reported disease. (C) Distributions of cumulative silent infections during periods absent of disease. Distributions in B and C are 10–90% quantiles and the median based on 500 simulations per US state. Silent infections are those that occur in the absence of reported cases, and highlight the unobservable dynamics of polio. (D) Simulated cases surrounding WPV extinction events. Distributions show 10–90% quantiles and the median number of cases observed up to 6 months preceding and 6 months following an extinction event. Generally, fewer than 5 cases per month are reported two months to either side of an extinction event. However, it is unclear whether 5 months with < 5 cases is a reliable signal of extinction.

Discussion

This work sheds light on the fundamental ecology of WPV. Latitudinal gradients have been identified in several acute viral infections: including influenza, RSV, rotavirus, and now polio [39, 40]. Our results indicate that the observed latitudinal gradient in the timing of polio epidemics is driven by a latitudinal gradient in demographic and/or environmental factors associated with transmission. Determining which mechanism is responsible has implications for control and surveillance efforts. Specifically, knowledge of the seasonal driver could allow for regionally-timed national immunization campaigns or the ability to forecast changes in epidemic seasonality.

Our identification of birth rate as a driver of polio's epidemic emergence during the baby boom of the 1940s and 1950s, is yet another demonstration [41, 42] of the need for full integration of demography into the study of childhood infectious disease epidemiology. The rate of susceptible recruitment has long been known to control the magnitude and frequency of epidemics of fully-immunizing childhood diseases [41, 43]. Today, in an era of human population expansion and emerging infectious diseases, we are reminded of the importance of characterizing changes in host population ecology. Due to limits of our demographic data, we were unable to address the early emergence phase of polio from the late 1800s through the 1920s. Rather, we focused on the later phase of emergence in the US, as the disease transitioned from small epidemics in the 1930s and early 1940s to large epidemics during the baby boom era. Though there were increases in the report rate, which contributed to the trend in observed cases, we also discovered an increase in the incidence of infection. Importantly, the increase in infection incidence closely tracked birth rates in the mid-1900s.

To the extent that our results bear on contemporary polio ecology, the identification of source-sink dynamics in the US suggests that successful local elimination of polio in a sink population is inconsequential in the presence of a source population. This prediction has unfortunately been repeatedly borne out in current

epidemics. Regional elimination of polio has been followed by reintroduction from endemic countries, such as the 2013 outbreak in Somalia, Ethiopia, and Kenya with WPV introduced from Nigeria, and repeat reinfection of Afghanistan from Pakistan [38]. Moreover, the metapopulation structure of WPV demonstrates that preventing emigration of WPV from source populations—which may be highly localized—is a requirement for efficient control.

We estimate that $> 99\%$ of infections were subclinical, with the reporting of total infections regularly below 1%. Importantly, subclinical infections are likely more common today than in our fitted models. This is because, first, both non-paralytic and AFP cases were reported in the US, whilst only AFP cases are currently reported. Second, our models were fit to data during the vaccine-free period of polio endemicity; therefore, infection incidence was elevated each summer, allowing the number of infections to grow sufficiently large to result in a high probability of clinical infections. In contrast, today, as polio's reproductive number approaches $R_t = 1$ in highly vaccinated endemic countries, WPV can circulate at levels below that needed for likely clinical observation. The recovery of environmental WPV isolates in Israel in the complete absence of AFP cases supports this expectation [38]. Furthermore, Fig 4.6B demonstrates that polio may circulate silently for extended periods (i.e., > 3 years) if the number of infections remains below the threshold for likely detection. Two years of silent WPV circulation has been confirmed. The outbreak in Central Africa detected in October 2013 was traced back to WPV circulation in Chad during 2011 [44]. Populations expected to have a small number of monthly infections in the presence of WPV—due to their demography or because they are highly vaccinated—would therefore be desirable targets for intense environmental surveillance. In terms of information gained, environmental surveillance is a powerful tool for identifying silent transmission in locations where polio would otherwise go undetected. In Pakistan the level of environmental surveillance has increased since 2011, and WPV has

consistently been detected, even in the absence of AFP cases [45].

In the absence of validated transmission models, case data are relied upon to determine whether a pathogen has gone locally extinct and estimate the critical community size required for pathogen persistence. In light of polio's propensity for silent circulation, we conclude that AFP data can be misleading; this conclusion extends to any communicable disease in which clinical cases represent a small fraction of infections. Extended periods absent of reported cases can mask infections circulating at levels below the threshold for likely reporting. We therefore advocate fitting transmission models to contemporary data to draw inference regarding extinction. Since infection can persist even in the extended absence of reported cases, knowledge of the local infection dynamics could reveal invaluable epidemiological information. Transmission models fit to endemic countries (i.e., Pakistan, Afghanistan, and Nigeria) could be used to identify how demographic and environmental factors interact with vaccine coverage to determine regional WPV persistence. In addition to coupling case data with transmission models for endemic countries, another useful extension would be to combine genetic data from WPV isolates with transmission models to further distinguish between sustained local transmission and imported infection. Genetic studies have found reductions in WPV genetic diversity in Afghanistan, suggesting local extinction of some WPV strains [6].

Vaccination campaigns might take further advantage of the seasonality and geographic clustering of WPV's reproductive ratio. Low transmission season vaccination campaigns have been utilized by the GPEI [6]. We found that the "low season" reproductive ratio can have geographic clusters where the reproductive ratio is > 1, which, if identified in the contemporary setting, might be useful targets for intense low season vaccination campaigns. Additionally, if the (1) seasonal reproductive ratio, (2) birth seasonality, and (3) vaccine coverage are quantified for endemic countries, vaccination campaigns could use this information to determine the regionally-optimal timing for

national vaccination days. These three quantities could be used to estimate the seasonal effective reproductive number, and evaluate alternative vaccination strategies. For instance, one strategy might be to extend the duration of the wintertime trough (i.e., by generating or extending the window during which the effective reproductive number is < 1), which may push WPV to extinction. Alternative strategies might be to vaccinate in the months prior to the seasonal peak in transmission or six months following the peak in births. In the past, mass OPV campaigns held during the low transmission season were deemed "most effective" [46], but it is unclear to what extent this strategy is used today.

Historical data, particularly in pre-vaccine periods, offer a unique glimpse into the ecology of infection, in the relative absence of human intervention. Historical data offer several advantages. First, reporting rates from historical eras are informative because they are reflective of (a) the symptomatology of infection, and (b) clinical diagnosis of symptomatic infection. Second, it can be difficult to infer unobserved infection dynamics using data for diseases that are near their eradication or elimination threshold. This is because the parameterization of transmission models with data containing few cases—and lacking recurrent epidemics—can result in ambiguous parameter estimates. The recurrent nature of historical epidemics gives us the unique opportunity to unravel disease-specific transmission ecology. Once the baseline transmission ecology is known, it can be coupled with data from contemporary periods to test hypotheses regarding modern-day epidemics and their geographic coupling.

Our analyses demonstrate the power of an approach focused on coupling mechanistic transmission models with long-term, spatially replicated longitudinal incidence data. Specifically, we document intriguing continental-scale gradients in polio seasonality, which we suggest are explained by latitudinal gradients in local transmission rates. We also show that the historical emergence of epidemic polio was largely a consequence of demographic trends rather than improvements in hygiene. Impor-

tantly, we demonstrate that historical polio persistence in the US was driven by an ever changing arrangement of source-sink populations. Finally, we found that even protracted AFP-free periods do not reliably indicate WPV extinction. Because of the difficulty in establishing fundamental aspects of WPV transmission in heavily vaccinated populations, it is our hope that these insights will act as a baseline for understanding modern polio transmission and disentangling vaccine effects from the natural ecology of the disease.

Acknowledgments

We would like to thank our editor, Roland Roberts, as well as three anonymous reviewers for helping us shape this manuscript. We would like to thank Kevin Bakker for digitizing the demography data and for feedback on the manuscript. We thank the Rohani lab, King lab, Adam Lauring, Alexandra McCubbrey, and Elizabeth Halloran for their helpful comments on this manuscript. We thank the CDC for the polio incidence records. MMB is supported by the NSF Graduate Research Fellowship Program and the University of Michigan Rackham Merit Fellowship. PR and AAK are supported by the Research and Policy in Infectious Disease Dynamics program of the Science and Technology Directorate, the Department of Homeland Security, the Fogarty International Center, the National Institutes of Health, and by a research grant from the National Institutes of Health (1R01AI101155). This research was supported in part through computational resources and services provided by Advanced Research Computing at the University of Michigan, Ann Arbor; and by UAF Life Science Informatics. UAF Life Science Informatics as a core research resource is supported by Grant Number RR016466 from the National Center for Research Resources, a component of the National Institutes of Health. MMB would also like to thank the NSF supported Ecology and Evolution of Infectious Disease Workshop.

Bibliography

[1] Nathanson N, Kew OM (2010) From Emergence to Eradication: The Epidemiology of Poliomyelitis Deconstructed. American Journal of Epidemiology 172: 1213–29.

[2] Freyche M, Nielsen J (1955) The Incidence of Poliomyelitis since 1920. In: Poliomyelitis: World Health Organization Monograph Series No. 26. pp. 59–104.

[3] Nathanson N, Martin J (1979) The Epidemiology of Poliomyelitis: Enigmas Surrounding its Appearance, Epidemicity, and Disapperance. American Journal of Epidemiology 110: 672–692.

[4] World Health Organization Country Office Islamabad Pakistan (2010) Weekly AFP Surveillance Update Pakistan 2010 Week 49. Technical report.

[5] Afghanistan National Surveillance Office (2011) Afghanistan Polio Eradication Initiative Surveillance Update 2011 Week 21. Technical Report May.

[6] World Health Organization (2013) Polio Eradication & Endgame Strategic Plan 2013-2018. Technical report.

[7] Donaldson L, Mogedal S, El Sayed N, Nduati R, Goldstein S, et al. (2014) Independent Monitoring Board of the Global Polio Eradication Initiative: Tenth Report October 2014. Technical report.

[8] Trevelyan B, Smallman-Raynor M, Cliff AD (2005) The Spatial Dynamics of Poliomyelitis in the United States: From Epidemic Emergence to Vaccine-Induced Retreat, 1910-1971. Annals of the Association of American Geographers 95: 269–293.

[9] Nelson H (1949) Poliomyelitis-Epidemiology and Preventive Measures. Journal of the Royal Sanitary Institute (Great Britain) 69: 160–170.

[10] Bartlett MS (1957) Measles Periodicity and Community Size. Journal of the Royal Statistical Society Series A 120: 48–70.

[11] Bjørnstad ON, Ims RA, Lambin X (1999) Spatial Population Dynamics: Analyzing Patterns and Processes of Population Synchrony. Trends in Ecology and Evolution 14: 427–432.

[12] Bjørnstad ON (2013). Spatial Nonparametric Covariance Function (R package ncf).

[13] Gouhier T (2014) biwavelet: Conduct univariate and bivariate wavelet analyses. URL `http://github.com/tgouhier/biwavelet`. (Version 0.17.4).

[14] Martinez-Bakker M, Bakker K, King AA, Rohani P (2014) Human Birth Seasonality: Latitudinal Gradient and Interplay with Childhood Disease Dynamics. Proceedings of the Royal Society B: Biological Sciences 281.

[15] Alexander JP, Gary HE, Pallansch MA (1997) Duration of Poliovirus Excretion and its Implications for Acute Flaccid Paralysis Surveillance: A Review of the Literature. The Journal of Infectious Diseases 175, Suppl: S176–S182.

[16] Howe HA (1952) Antibodies and Immunity to Poliomyelitis. In: Poliomyelitis: Papers and Discussions Presented at the Second International Poliomyelitis Conference. pp. 295–300.

[17] Paul JR (1955) Epidemiology of Poliomyelitis. In: Poliomyelitis: World Health Organization Monograph Series No. 26. pp. 9–29.

[18] Casey A (1942) The Incubation Period in Epidemic Poliomyelitis. Journal of the American Medical Association 120: 805–807.

[19] Gear JHS (1955) Distribution of Antibodies to Poliomyelitis in the General Population. In: Poliomyelitis: Papers and Discussions Presented at the Third International Poliomyelitis Conference. pp. 137–148.

[20] Strean GJ, Gelfand MM, Pavilanis V, Sternberg J (1957) Maternal- Fetal Relationships: Placental Transmission of Poliomyelitis Antibodies in Newborn. Canadian Medical Association Journal 77: 315–323.

[21] Evans CA (1960) Factors Influencing the Occurrence of Illness During Naturally Acquired Poliomyelitis Virus Infections. Bacteriological Reviews 24: 341–352.

[22] Anderson RM, May RM (1991) Infectious Diseases of Humans: Dynamics and Control. 114–116 pp.

[23] Sabin AB (1951) Paralytic Consequences of Poliomyelitis Infection in Different Parts of the World and in Different Population Groups. American Journal of Public Health 41: 1215–1230.

[24] King AA, Ionides EL, Breto CM, Ellner S, Kendall B, et al. (2010). pomp: Statistical Inference for Partially Observed Markov Processes (R package).

[25] Ionides EL, Bhadra A, Atchadé Y, King A (2011) Iterated Filtering. The Annals of Statistics 39: 1776–1802.

[26] King AA, Nguyen D, Ionides E (in press) Statistical Inference for Partially Observed Markov Processes via the R Package pomp. Journal of Statistical Software.

[27] Grenfell BT, Bjørnstad ON, Kappey J (2001) Travelling Waves and Spatial Hierarchies in Measles Epidemics. Nature 414: 716–23.

[28] Cummings DAT, Irizarry RA, Huang NE, Endy TP, Nisalak A, et al. (2004) Travelling Waves in the Occurrence of Dengue Haemorrhagic Fever in Thailand. Nature 427: 344–347.

[29] Viboud C, Bjørnstad ON, Smith DL, Simonsen L, Miller Ma, et al. (2006) Synchrony, Waves, and Spatial Hierarchies in the Spread of Influenza. Science 312: 447–451.

[30] Alonso WJ, Viboud C, Simonsen L, Hirano EW, Daufenbach LZ, et al. (2007) Seasonality of Influenza in Brazil: A Traveling Wave from the Amazon to the Subtropics. American Journal of Epidemiology 165: 1434–1442.

[31] Choisy M, Rohani P (2012) Changing Spatial Epidemiology of Pertussis in Continental USA. Proceedings of the Royal Society B: Biological Sciences 279: 4574–81.

[32] Pitzer VE, Viboud C, Simonsen L, Steiner C, Panozzo CA, et al. (2009) Demographic Variability, Vaccination, and the Spatiotemporal Dynamics of Rotavirus Epidemics. Science 325: 290–294.

[33] Keeling M, Rohani P (2008) Modeling Infectious Diseases in Humans and Animals. Princeton: Princeton University Press.

[34] Dorélien A, Ballesteros S, Grenfell BT (2013) Impact of Birth Seasonality on Dynamics of Acute Immunizing Infections in Sub-Saharan Africa. PLoS ONE 8: e75806.

[35] Brown GC (1955) Virus Excretion and Antibody Response in Clinical and Subclinical Cases of Poliomyelitis. Annals of the New York Academy of Sciences 61: 989–997.

[36] Atkinson W, Wolfe S, Hamborsky J (2012) Poliomyelitis. In: Epidemiology and Prevention of Vaccine-Preventable Diseases 12th Edition. pp. 249–262.

[37] Melnick JL, Ledinko N (1953) Development of Neutralizing Antibodies Against the Three Types of Poliomyelitis Virus During an Epidemic Period. American Journal of Hygiene 58: 207–222.

[38] Donaldson L, El Sayed N, Goldstein S, Koplan J, Mogedal S, et al. (2013) Independent Monitoring Board of the Global Polio Eradication Initiative Eighth Report - October 2013. Technical Report October.

[39] Bloom-Feshbach K, Alonso WJ, Charu V, Tamerius J, Simonsen L, et al. (2013) Latitudinal Variations in Seasonal Activity of Influenza and Respiratory Syncytial Virus (RSV): A Global Comparative Review. PLoS ONE 8: e54445.

[40] Lebaron CW, Lew J, Glass RI, Weber JM, Ruiz-Palacios GM (1990) Annual Rotavirus Epidemic Patterns in North America: Results of a 5-Year Retrospecitive Survey of 88 Centers in Canada, Mexico, and the United States. Journal of the American Medical Association 264: 983–988.

[41] Earn DJ, Rohani P, Bolker BM, Grenfell BT (2000) A Simple Model for Complex Dynamical Transitions in Epidemics. Science 287: 667–70.

[42] Morris SE, Pitzer VE, Viboud C, Metcalf CJE, Bjørnstad ON, et al. (2015) Demographic Buffering: Titrating the Effects of Birth rate and Imperfect Immunity on Epidemic Dynamics. Journal of the Royal Society Interface 12: 20141245.

[43] Broutin H, Viboud C, Grenfell BT, Miller MA, Rohani P (2010) Impact of Vaccination and Birth Rate on the Epidemiology of Pertussis: a Comparative Study in 64 Countries. Proceedings of the Royal Society Biological Sciences 277: 3239–3245.

[44] Mckenzie A, Wolff C, Tangermann R, Chenoweth P, Tallis G, et al. (2014) Assessing and Mitigating the Risks for Polio Outbreaks in Polio-Free Countries — Africa, 2013–2014. CDC Morbidity and Mortality Weekly Report 63.

[45] Sein C (2013) Evaluating Surveillance Indicators Supporting the Global Polio Eradication Initiative, 2011-2012. Morbidity and Mortality Weekly Report 62: 270–274.

[46] Hull HF, Ward NA, Hull BP, Milstien JB, De Quadros C (1994) Paralytic Poliomyelitis: Seasoned Strategies, Disappearing Disease. Lancet 343: 1331–1337.

CHAPTER V

Both Salk and Sabin Vaccines Effectively Reduce Polio Transmission in Epidemic Settings

Abstract. Wild poliovirus (WPV) is on the brink of eradication worldwide. There are two vaccines against polio, the Salk Inactive Polio vaccine (IPV) and the Sabin live Oral Polio vaccine (OPV). OPV is the primary weapon of the Global Polio Eradication Initiative (GPEI) and is employed throughout the developing world. Unfortunately, OPV has a serious drawback: reversion of the live attenuated virus to a pathological form that can result in circulating vaccine derived poliovirus [1, 2]. In this way, OPV acts as a source of poliovirus, and although OPV has been able to push WPV near the eradication threshold, it cannot be used for complete eradication. For this reason, OPV will be withdrawn and IPV re-introduced. The GPEI is now faced with deciding where and when to withdraw OPV and introduce IPV. However, this decision is being made in the face of great uncertainty regarding the efficacy of IPV at reducing polio transmission [3, 1]. If IPV fails to induce herd immunity in endemic countries, the premature withdrawal of OPV could lead to the resurgence of polio. Here, we use the initial launch of IPV and OPV in the US and the USSR to measure the extent that IPV and OPV reduce WPV transmission. We conclude, relative to unvaccinated individuals, infected OPV and IPV vaccinated individuals transmit WPV 96% and 69% less, respectively; suggesting that IPV could be used to eradicate polio in regions

with an off-season WPV reproductive number ≤ 3.

Polio from Past to Present

WPV is enigmatic for causing paralytic polio in a small fraction of infected individuals. Due to intense global vaccination efforts, WPV is on the brink of eradication worldwide with only three remaining endemic countries (Pakistan, Afghanistan, and Nigeria). There are two polio vaccines that were licensed in the 1950s and 1960s, IPV and OPV, also known as the Salk and Sabin Vaccines, respectively. Most countries have eradicated polio using OPV, with the exception of Scandinavia and the Netherlands, which used IPV [4]. OPV is the primary eradication tool for several reasons: (1) OPV stimulates intestinal immunity by mimicking a natural infection and generating a secretory IgA response that reduces viral shedding, (2) OPV contains attenuated live virus that replicates in vaccinated individuals who then shed attenuated virus and indirectly vaccinate others, and (3) OPV is inexpensive and easily administered as oral drops or candy [3, 5]. IPV, on the other hand, is an injectable vaccine that contains inactive virus killed using formalin [6]. The killed virus in IPV does not establish infection within vaccinated individuals. A recent trial of OPV vaccination followed by an IPV dose demonstrated that IPV can boost mucosal immunity [5]; however, it is unknown whether IPV alone can stimulate enteric mucosal immunity and reduce transmission, or if it merely stimulates humoral immunity and protects from paralytic polio [3, 5]. It is widely recognized that OPV provides superior protection from WPV infection and reduces transmission; however, OPV has a serious drawback. The attenuated virus in OPV can undergo recombination with other enteroviruses or can acquire point mutations and revert back to a pathological form, known as vaccine-derived poliovirus [7, 8]. Estimates of the frequency of vaccine-derived poliovirus vary. Generally, it is estimated that 1–2 cases of paralytic polio arise per million primary OPV immunizations administered [2]. In these rare

cases, the vaccinated individual acquires a paralytic vaccine-derived poliovirus infection that can be transmitted [1]. In this way, OPV acts as a source of poliovirus, and although OPV has been able to push WPV near the global eradication threshold, it cannot be used for complete eradication. For this reason, the 2013–2018 GPEI action plan calls for the complete withdrawal of trivalent OPV, and subsequently, a gradual withdrawal of bivalent OPV and introduction of IPV [9]. The GPEI is faced with deciding where and when to withdraw OPV and introduce IPV while WPV continues to circulate. It is unknown whether IPV has the ability to push WPV to extinction in endemic regions because there is a lack of information on the extent to which IPV reduces WPV transmission. Here, we utilize extensive longitudinal data on poliomyelitis incidence and vaccination status of the population of the US and the USSR in the 1950s and 1960s to estimate the population-level impact of IPV and OPV on transmission of WPV.

Early in the 1900s polio became an epidemic disease in industrialized nations. In the US, the incidence of polio increased steadily throughout the 1900s until the sequential development of IPV and OPV. In the US alone, the unprecedented magnitude of polio epidemics in the 1950s resulted in > 10000 clinical cases and millions of subclinical infections annually [10]. The majority of clinical cases were concentrated in school-aged children [2]. Due to this age-structure, when the polio vaccines were introduced, which first occurred in the US (for IPV) and USSR (for OPV), preference was given to school-aged children and individuals under 20 years old [11, 12, 13]. In order to gauge the safety and success of the massive polio immunization campaigns in the US and USSR, the US launched the Poliomyelitis Surveillance Unit (PSU) at the Centers for Disease Control and Prevention (CDC) [14]; at the same time, the USSR Poliomyelitis Research Laboratory tracked the USSR's intense OPV vaccination campaigns and clinical cases. Detailed data from these surveillance units provided the unique opportunity to study the population-level effects of IPV and OPV on epi-

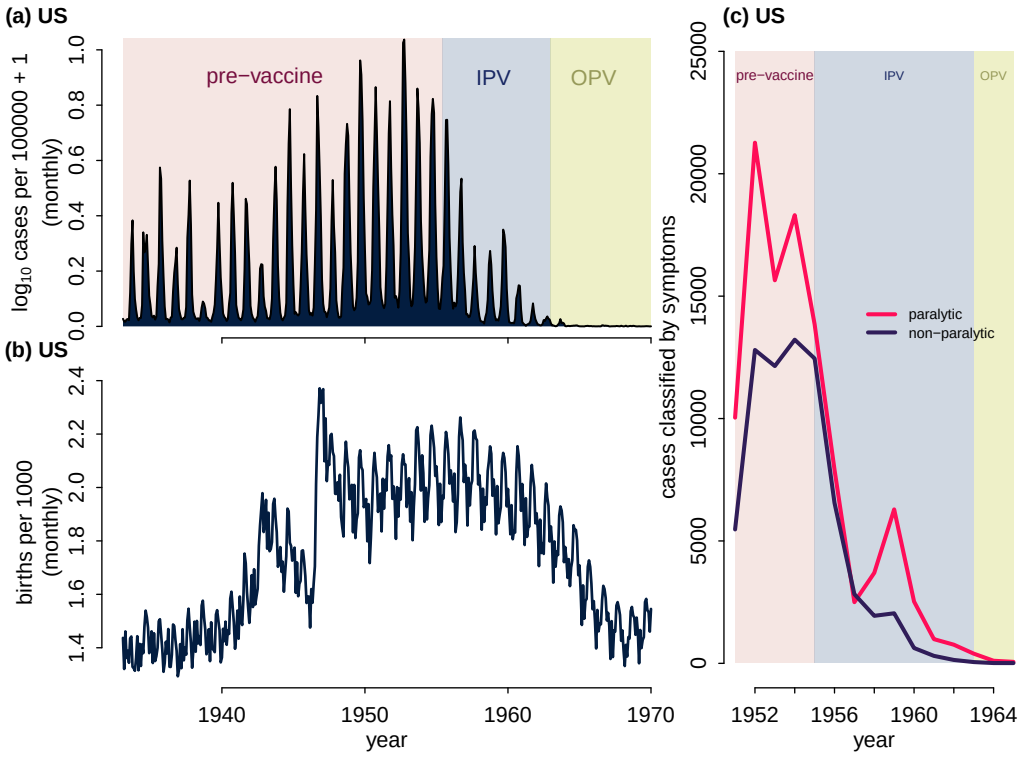


Figure 5.1: **MMWR and NOVS data.** (a) MMWR data, polio cases per 100000 population in the contiguous US from January 1933–December 1969. (b) Births per 1000 population over the same time period. (c) NOVS data, polio cases for which symptomatology was classified as paralytic or non-paralytic; this is a subset of the cases in MMWR, all others were unspecified.

demic polio. Here for the first time, we curated data detailing the roll-out of IPV and OPV in the US and USSR, combined it with clinical case data and mathematical transmission models, and measured the efficacy of IPV and OPV at the population level.

The race to develop a polio vaccine resulted in the development of IPV by Jonas Salk and OPV by Albert Sabin. Placebo-controlled IPV trials were run by the University of Michigan and included the vaccination of over 700000 children [15], resulting in IPV licensing in April 1955. Once licensed, IPV was quickly launched throughout the US. Initially IPV was given to 2nd grade children but the program was rapidly expanded to include all children and young-adults [16, 17]. OPV followed at the

heels of IPV. The vaccine trials for OPV were held in the USSR. In September 1958 Albert Sabin provided the USSR with virus seeds for the three attenuated poliovirus serotypes. The USSR used these seeds to culture the viruses in monolayer monkey kidney cells. The USSR immediately cultured enough of each serotype to vaccinated 10 million people. The first cycles of immunization were conducted January–March 1959. The USSR held the first mass immunization campaigns with OPV in Estonia and Lithuania. On December 16, 1959 the USSR approved widespread vaccination with OPV, and vaccinated individuals 2 month olds to 20 years old. This involved vaccinating 35% of total population, 77 million people, between January and June 1960. OPV was administered via Pomadka-Candy, which allowed for easy dissemination [11, 12]. The US sequentially licensed each monovalent OPV and licensed trivalent OPV in 1963. After the World Health Organization reviewed the details and design of the USSR's OPV trials and confirmed their success, OPV became the vaccine of choice in the 1960s [18]. By 1969 OPV was more widely used in the US than IPV. In order to estimate IPV and OPV efficacy, we focused on data detailing the introduction of IPV in the US and the introduction of OPV in three regions of the USSR: Estonia, Lithuania, Ukraine, and the Russian Soviet Federative Socialist

Data

Since the early 1900s, polio cases were regularly recorded in the US Weekly Public Health Reports and the CDC Morbidity and Mortality Weekly Reports. We obtained monthly reported polio cases in the US from January 1933–December 1969 from the US CDC. We refer the reader to [10] for our baseline analyses of polio transmission ecology in the US. We will refer to the US monthly polio data as the "MMWR data" (Fig 5.1a). Beginning in 1951, the National Office of Vital Statistics (NOVS), which collected the MMWR data, allowed symptomatology to be classified in case reports. Cases could be reported as paralytic or non-paralytic polio, but many cases

were unspecified. We obtained the subset of cases for which symptomatology was provided, which we will refer to as the "NOVS data" (Fig 5.1c). The NOVS data are annual from 1951–1965.

In addition to the MMWR and NOVS data, we digitized independent polio case data collected by the CDC PSU and housed at the CDC library. The PSU was a task force commissioned to gather detailed data on polio cases following the launch of IPV in 1955 [14]. PSU data were collected independent of the MMWR and NOVS data; because of this, not all cases represented in MMWR were included in PSU data, and vice versa. The PSU data are annual from 1958–1965 and detail polio cases by age, vaccine status, and symptomatology (Fig 5.2). The NOVS and PSU data both classify cases as either paralytic or non-paralytic. The interpretation of non-paralytic polio cases has always been problematic due to the lack of a consistency in the clinical definition. There are two broad clinical definitions that have been used to classify non-paralytic polio. Some authors define non-paralytic polio as a symptomatic infection that lacks signs of central nervous system (CNS) involvement; others define non-paralytic polio as an infection with CNS involvement that does not result in paralysis. Those that use the latter definition generally place symptomatic infections lacking CNS involvement into another category known as abortive polio [19]. Inconsistency in the definition of non-paralytic polio requires acknowledgement when interpreting polio case reports. We acknowledge that although paralytic polio is a well-defined clinical manifestation of infection, non-paralytic polio is more of a "catch-all" category for clinical WPV infections that do not manifest as acute flaccid paralysis (AFP). All together, we have obtained three polio data sets from the US (i.e., monthly MMWR data, annual NOVS data, and annual PSU data), which we couple with complementary vaccination and demography data.

The PSU not only collected polio case data, but—along with the newly formed National Immunization Survey—also collected data on the distribution of polio vac-

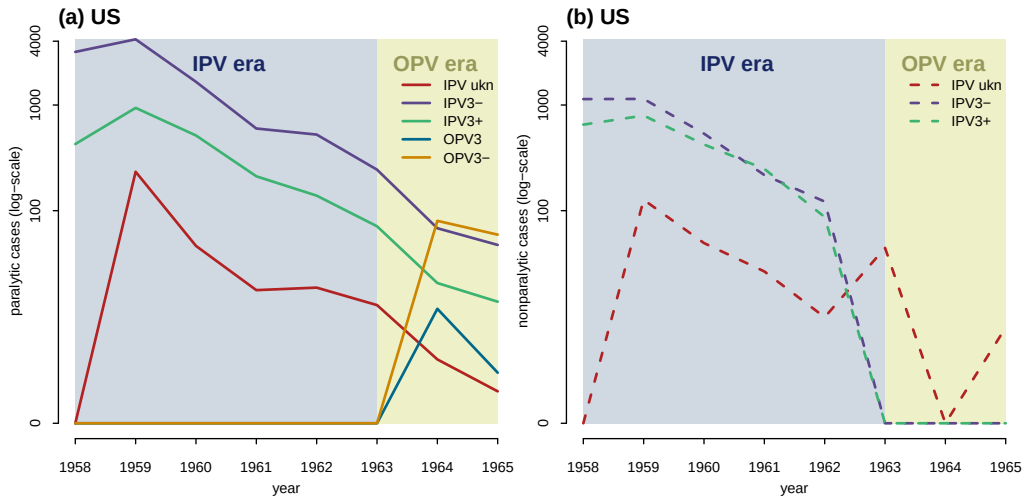


Figure 5.2: **PSU Data.** Cases reported to the Poliomyelitis Surveillance Unit detailed by vaccine status and symptomatology. Individuals with the full 3+ dose series of IPV or OPV are indicated by IPV3+ and OPV3; unvaccinated and under-vaccinated individuals are indicated as IPV3- or OPV3-; "ukn" indicates unknown vaccine status. Symptomatology is divided into two categories: (a) paralytic or (b) non-paralytic polio.

cines and age-structured vaccine coverage with IPV and OPV. We digitized PSU vaccine data which included quarterly shipments of IPV in the US, annual shipments of monovalent and trivalent OPVs, and the annual percent of individuals in five-year age groups that have 0, 1, 2, 3, or 4+ doses of IPV, OPV, or combined IPV and OPV (Figs 5.3 & 5.4). In addition to the disease and vaccine data, we also gathered demography data for the US that included monthly births (Fig 5.1b), which were an important determinant of polio incidence [10], and annual population size estimates from the US Census Bureau.

For the USSR, we digitized monthly polio cases reported in Lithuania and Estonia from January 1955–December 1960. The data were digitized from the Institute for Poliomyelitis Research, Academy of Medical Sciences, Moscow, USSR. Data reported in the Bulletin of the World Health Organization 1961, vol 25, pp 79–91 [13]. Mass immunization with OPV in Lithuania and Estonia occurred in the early part of 1959 and was repeated in the second quarter of 1960. The population sizes were $\sim$

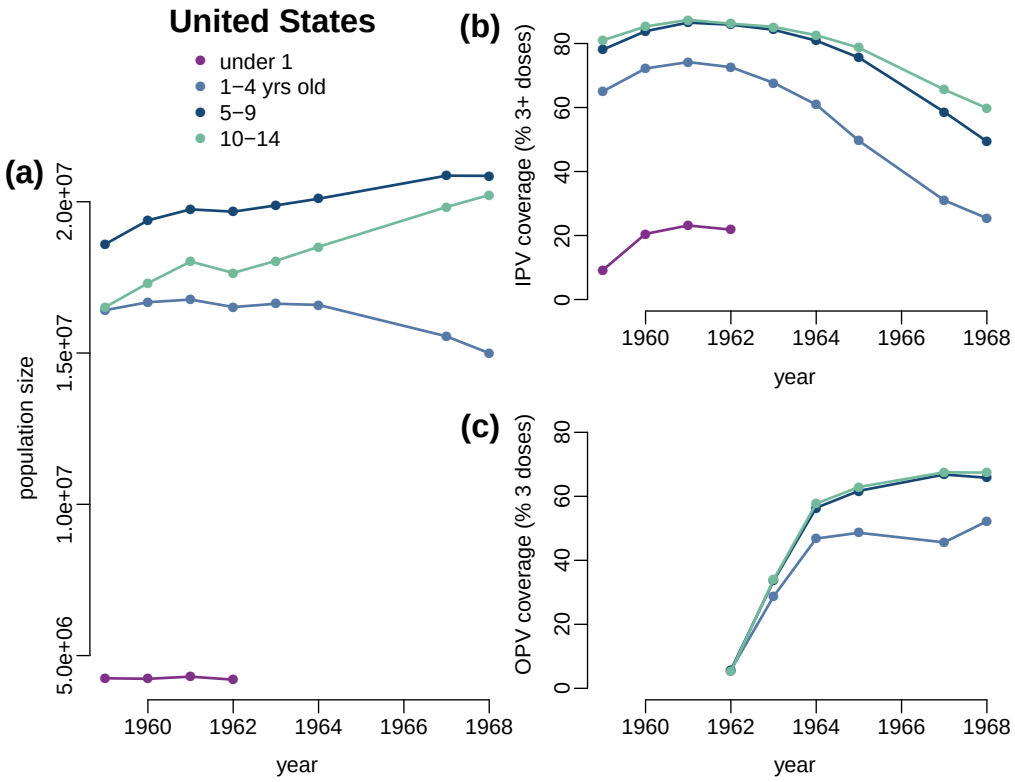


Figure 5.3: **Age structured population size and polio vaccine coverage in the US.** (a) Annual size of childhood age classes. (b) Percent of children in each age group with 3+ doses of IPV. (c) Percent of children with 3 doses of OPV.

2700000 for Lithuania, and ∼ 1200000 for Estonia. By December 1959, the estimated number of persons vaccinated in Lithuania and Estonia was 547000 (20% of population) and 695000 (58%), respectively. By December 1960, the estimated number of persons vaccinated in Lithuania was 1760000 (65%), and Estonia 827240 (69%). Monthly paralytic polio cases reported in the RSFSR and Ukraine were obtained from a shorter time period, January 1957–December 1960. The data were originally reported as incidence per 100000. The population size was ∼ 119603000 for the RSFSR, and ∼ 42606000 for Ukraine. By December 1959, only 3310000 people (2.8% of the population) were vaccinated in the RSFSR and 2378000 (5.6%) in Ukraine. However, by December 1960, the estimated number of persons vaccinated in RSFSR was 42604150 (36%), and 14152778 in Ukraine (33%). In the RSFSR monovalent OPVs

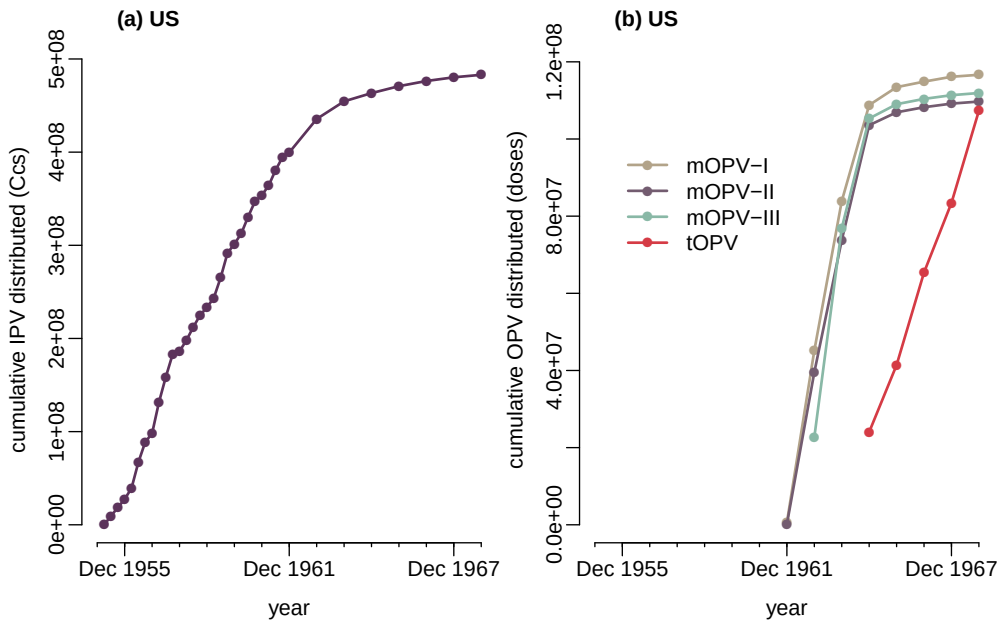


Figure 5.4: **Distribution of IPV and OPV during the roll-out in the US.** (a) Cumulative distribution of IPV in the US. Data are quarterly from 1955–1961, after which they are annual. (b) Cumulative distribution of OPV monovalent type I, II, and II, and trivalent OPV.

were given three times in 1960 and subsequently trivalent OPV was used for revaccination. In some regions of Ukraine, trivalent vaccine was administered twice; and in the remaining regions of Ukraine, monovalent vaccines were administered separately and trivalent was used for revaccination.

Transmission Model

We constructed a stochastic discrete-time Susceptible-Infected-Recovered polio transmission model that explicitly accounted for vaccine status. A pre-vaccine version of this transmission model was previously validated for studying polio transmission [10]. The model includes classes to track individuals with 3+ doses of IPV and/or 3 doses of OPV. In the PSU data we observed polio cases in vaccinated individuals (Fig 5.2); therefore, we allowed for infection of, and transmission by, vaccinated individuals. For vaccinated individuals, we estimated their (i) susceptibility to infection

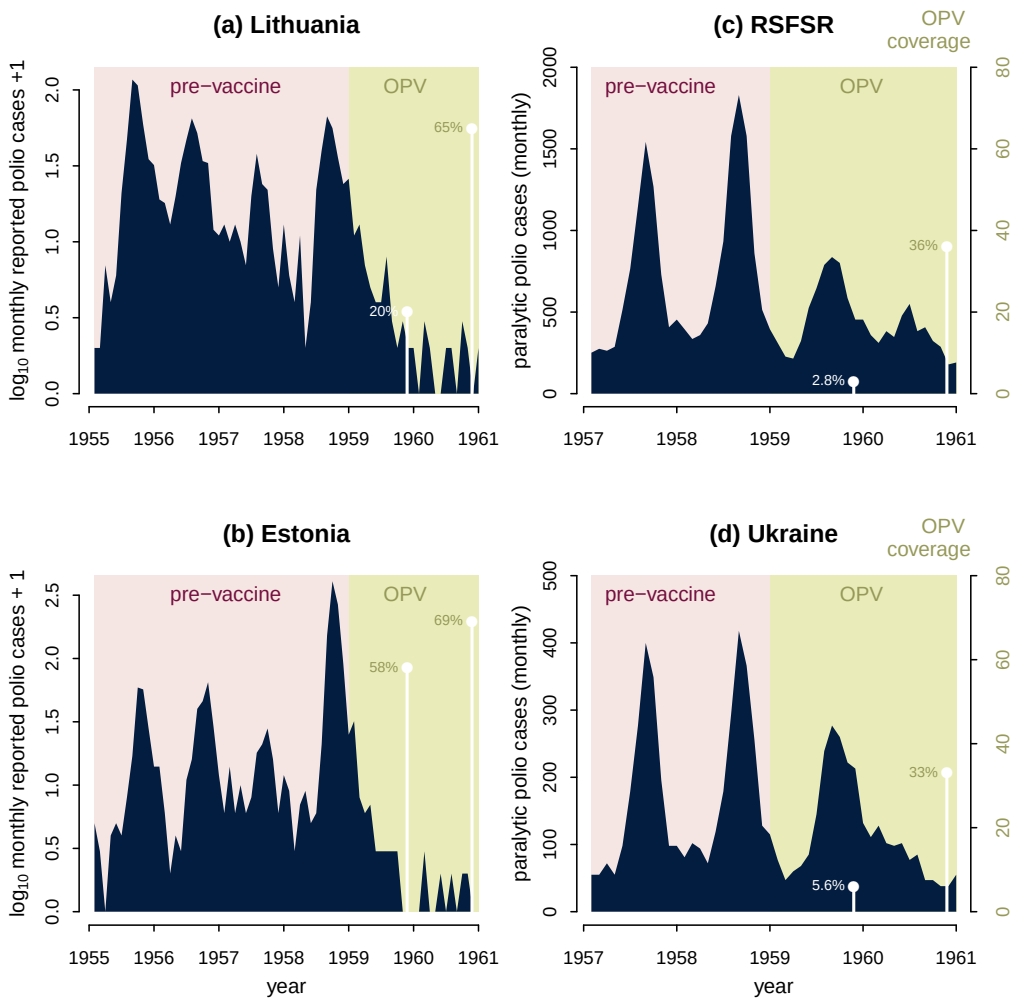


Figure 5.5: **Polio cases in the USSR.** Monthly polio cases reported in (a) Lithuania, (b) Estonia, (c) RSFSR, and (d) Ukraine, along with the percent of the population vaccinated with OPV (white points). OPV was first introduced in Lithuania and Estonia, where vaccine coverage reached $> 60\%$ in the second year of vaccination. As of December 1959, only 2.8% of population was vaccinated in the RSFSR and 5.6% in Ukraine. However, by December 1960, the reported vaccination coverage in the RSFSR was 36% and 33% in Ukraine.

(a measure of vaccine leakiness), and (ii) infectiousness, relative to unvaccinated individuals. The model was constructed in order to infer (a) the mode of vaccine action (i.e., whether each vaccine reduces susceptibility to infection, reduces infectiousness, and/or protects against pathology), and (b) vaccine efficacy (i.e., the total amount that each vaccine reduces transmission). Fig 5.6 shows the structure of the transmis-

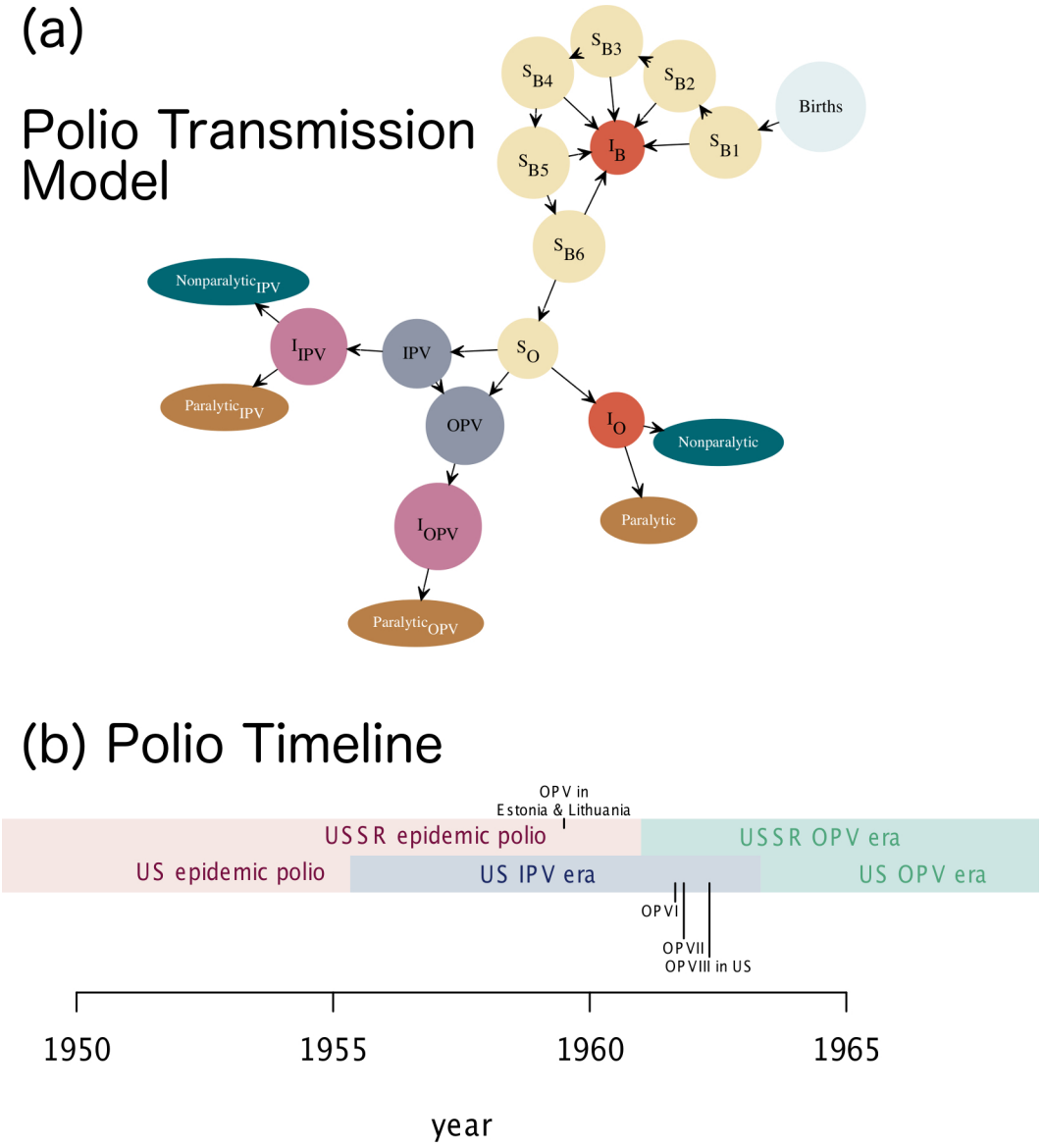


Figure 5.6: **Vaccine era polio transmission model.** (a) Schematic model of polio transmission and vaccination. The model has two vaccinated categories, IPV and OPV, which include individuals with 3+ doses of IPV and 3 doses of OPV, respectively. The model categorizes infections by vaccine status and symptomatology (i.e., paralytic and non-paralytic polio). (b) The timeline of polio in the US and USSR, epidemic periods are indicated along with the roll-out of IPV and OPV in each country. In the US, monovalent OPVs were licensed and used for epidemic response before the switch from IPV to OPV for immunization in April 1963 when trivalent OPV was licensed.

sion model. In the absence of IPV and/or OPV in a country or region, the IPV and OPV classes and their downstream classes are empty. In addition to the epidemiological process model, we also included a parallel model to track the vaccine coverage in the US and each of the four regions in the USSR. This was necessary because $\sim 99\%$ of polio infections are subclinical [10]; therefore, individuals vaccinated may already have naturally-derived immunity, making it tenuous to directly use the susceptible class from the epidemiological model to infer vaccine uptake rates from the vaccine coverage data. The vaccine tracker model for the US tracked entry and exit from the under 15 age group via births, deaths, and aging, and it is in this pure demography model that we account for IPV and OPV vaccination. By fitting the vaccine tracker model to the IPV and OPV coverage data, we obtained per capita vaccination rates for the under 15 year age group, which we then applied to our susceptible class in the epidemiological process model. We used a similar vaccine tracker model for each region of the USSR, but in the USSR we tracked OPV coverage for all age groups.

All statistical inference was done using maximum likelihood by iterated particle filtering (mif) in the R package pomp [20]. Our statistical inference pipeline is discussed in detail in the Appendix. For the US and each region in the USSR, we fit the data independently and sequentially in the following order: we (1) fit polio transmission parameters and report rates to the pre-vaccine era case data, (2) fit the per capita vaccine uptake rates to the vaccine coverage data, and (3) fixed the maximum likelihood estimate (MLE) transmission parameters, report rates, and vaccine uptake rates in the model and fit the vaccine efficacy parameters to the vaccine era data.

The NOVS data from the US allowed us to estimate the probability of paralytic and non-paralytic polio pathology. We estimated the probability of unvaccinated individuals having paralytic or non-paralytic polio upon infection, and we assumed these parameters are determined by host-pathogen biology and hold in both the US and USSR. For each region in the USSR (i.e., Estonia, Lithuania, Ukraine, and the

RSFSR), transmission parameters were fit to the USSR pre-vaccine era data, the paralytic and non-paralytic probabilities were taken from our US estimates, but the report rates of symptomatic polio infections were estimated for each region in the USSR.

We fit IPV efficacy parameters by simultaneously fitting the model to the US MMWR data and the NOVS data from January 1955–December 1962, the end of the IPV era in the US. One difficulty of studying vaccine efficacy in the US is that by the time OPV was fully rolled-out in the US in 1963, the incidence of polio was already low because IPV had been in use for 8 years. Additionally, as the vaccine era progressed in the US, the breakdown of reported cases in the PSU data signaled a reduction in reporting of non-paralytic polio, reducing the reported cases even further. The small number of reported cases during the OPV era in the US, and uncertainty regarding the stationarity of reporting, left us with little data with which to estimate OPV parameters; therefore, we took advantage of the OPV roll-out in the USSR to study OPV efficacy.

We fit OPV efficacy parameters to the paralytic polio data from the RSFSR. The RSFSR was specifically chosen because it reported only paralytic polio cases. The PSU data from the US reported paralytic polio in OPV vaccinated individuals, but never reported a non-paralytic case in an OPV vaccinated individual; we do not know if this is because OPV protects from the non-paralytic polio or if this is due to reporting bias. With no way of knowing whether the former or the latter was the case, we left non-paralytic infections in OPV-vaccinated individuals out of the model and used only paralytic polio cases when fitting OPV efficacy parameters. The observation model for each country/region is schematically depicted in Fig 5.7. The process model state variables that fed into the observation models included paralytic and non-paralytic polio infections in (1) unvaccinated or under-vaccinated individuals, (2) individuals fully vaccinated with 3+ doses of IPV, and (3) individuals fully vaccinated

with 3 doses of monovalent OPVs or trivalent OPV. The observation models linked observable state variables with the data from each country/region. In the US, the full observation model allowed for observations in the MMWR, NOVS, and PSU data as depicted in Fig 5.7. However, due to computational limits on calculating joint likelihoods for 3+ time series simultaneously, a nested observation model for the US that only included the MMWR and NOVS data was used for fitting IPV efficacy.

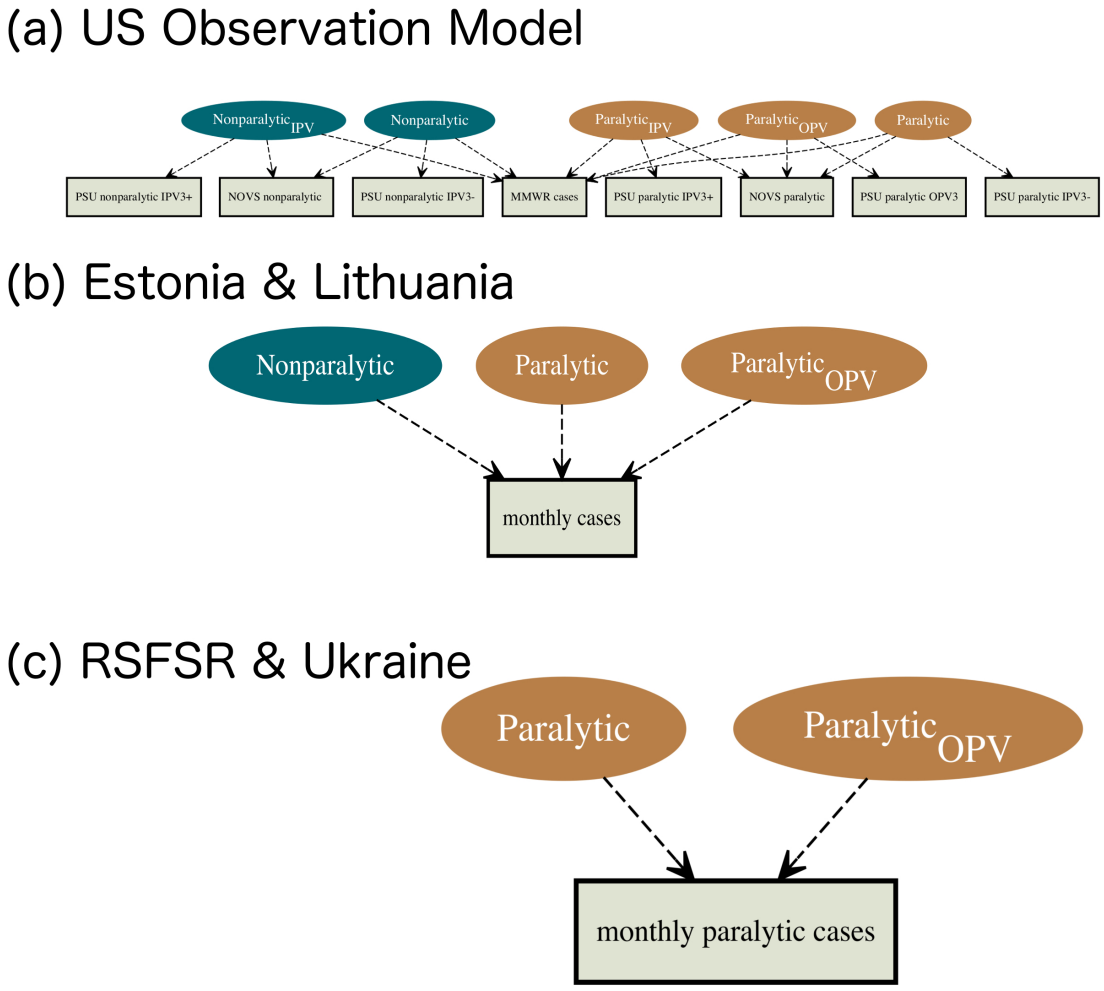


Figure 5.7: **Polio observation models for the US and USSR.** The observation models detail how symptomatic infections from the process model (turquoise and orange ovals) were reported in the (a) US via cases in the MMWR, NOVS, and PSU data sets, (b) Estonia and Lithuania, and (c) the RSFSR and Ukraine. In the US, the process model state variables that fed into the observation model were paralytic and non-paralytic infections broken down by vaccine status, either on a monthly or annual basis. The full US observation model is shown in (a); however, a nested version that includes only the MMWR and NOVS data was used for fitting IPV efficacy parameters. For Lithuania and Estonia, the observable state variables were the monthly number of non-paralytic and paralytic infections broken down by OPV status, assuming OPV protects from non-paralytic polio. For the RSFSR and Ukraine, the observable state variables were the monthly number of paralytic polio infections by OPV status. Subscripts indicate individuals with IPV3+ or OPV3 doses, lack of subscripts indicate unvaccinated/under-vaccinated individuals.

Results

Patterns in the US

The incidence of polio in the US displayed strong seasonality, with epidemics occurring in the late summer and autumn (Fig 5.1). The transmission rate MLEs for the US were 1.1 during the seasonal trough and 7.5 during the peak (Fig 5.8). There is a great deal of geographic variation in transmission seasonality in the US [10]; thus, the MLEs from the national data represent an aggregate estimate. The seasonality persisted during the IPV era, but dissipated after the introduction of OPV (Fig 5.1). Figures 5.1 & 5.2 show the dramatic decline in polio incidence following the introduction of IPV and OPV. In the 8 years preceding the introduction of IPV, April 1947–March 1955, there were 274719 reported polio cases in the MMWR, as compared to the 69608 reported cases during the 8 year IPV era, April 1955–March 1963. Although the incidence of polio had already declined substantially when trivalent OPV was introduced in April 1963, during the first 5 years of the OPV era, the number of reported cases was 645, which was 3.3% of the number reported in the previous 5 years, and 0.3% of those reported the last 5 years of the pre-vaccine period. The NOVS data demonstrate that paralytic polio comprised more of the reported cases than non-paralytic polio, this is likely due to the high report rate of paralytic polio, which other studies estimated to be 60–80% in the US at this time [21]. The breakdown of reported cases by vaccine status in the PSU data revealed that paralytic polio was regularly reported in individuals with 3+ doses of IPV, but unvaccinated and under-vaccinated individuals (i.e., IPV3- and OPV3-) represented the largest number of paralytic polio cases.

Patterns in the USSR

The USSR displayed epidemic seasonality similar to that in the US. Epidemics occurred in the summer and autumn (Fig 5.5). However, with the exception of Estonia, one noticeable difference between the epidemics in the US and those of the USSR is that the off-season trough in cases is shallow in the USSR relative to the US cases. This suggests the USSR maintained WPV infections at a higher level during the wintertime off-season. This is supported by the estimates of the transmission seasonality in the USSR (Fig 5.8). The seasonal minimum reproductive number was estimated to be 21, 3, 3, and 2 for Lithuania, Estonia, the RSFSF, and Ukraine, respectively. The seasonal maximum reproductive numbers were 63, 53, 8, and 7. The seasonal peak reproductive numbers for Lithuania and Estonia > 50 are on the level of that observed in the US Midwest, where explosive epidemic dynamics occurred. However, the shallow trough in Lithuania, with a reproductive number of 21 is far above that estimated in any other location to date [10]. Interestingly, the transmission rates for Ukraine and the RSFSR are similar to US estimates, with the exception that the RSFSR and Ukraine both had a shoulder in the transmission rate at the turn of the year.

Vaccine-induced reduction in polio incidence was observed across the USSR. Fig 5.5 shows monthly cases of polio in the USSR before and after the introduction of OPV. Vaccine coverage in Lithuania and Estonia was >60% within the first two years of OPV introduction. Seasonal epidemics were observed in both Lithuania and Estonia prior to OPV introduction, and these epidemics were lost immediately following the introduction of OPV. In the RSFSR and Ukraine, the widespread introduction of OPV was delayed a year after its roll-out in Estonia and Lithuania. Because OPV was introduced later, vaccine coverage in the RSFSR and Ukraine was lower as of December 1960, 36% and 33%, respectively. Based on the data obtained from the RSFSR and Ukraine, the incidence of paralytic polio declined during the first 2 years

of OPV vaccination, but not as dramatically as in Lithuania and Estonia.

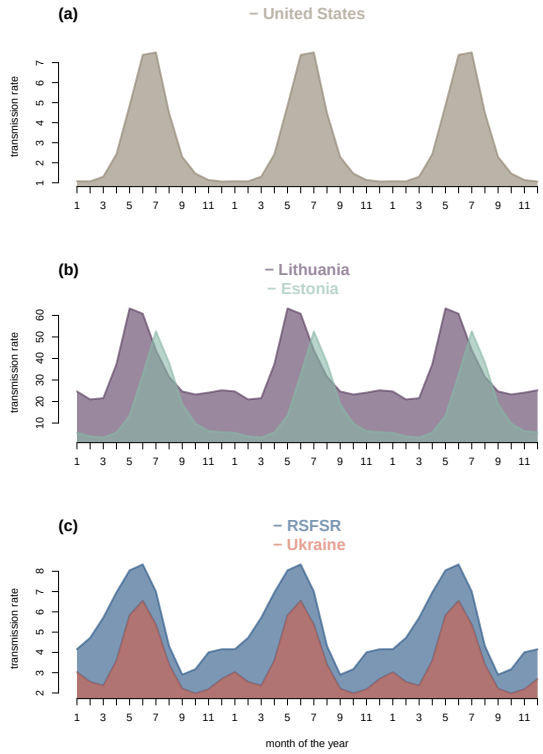


Figure 5.8: **MLEs of the seasonal transmission rate.** Rates for the (a) US, (b) Lithuania and Estonia, and (c) RSFSR and Ukraine. The transmission rate parameters were estimated using the pre-vaccine era data. For the USSR, the transmission rate equates to the reproductive number.

We used one-month-ahead predictions to validate our MLEs of polio transmission parameters. Fig 5.9 shows one-month-ahead predictions for the US, and Fig 5.10 shows the predictions for each region of the USSR. We used data from the pre-vaccine era for fitting and validating transmission parameters. In the US, the combined MMWR and NOVS data were used to estimate the probability of paralytic and non-paralytic polio in unvaccinated infected individuals, as opposed to subclinical symptoms or silent infection. Under the assumption that the report rate of paralytic polio was between 60–80% [21], we estimated that the probability that an unvaccinated infected individual was paralyzed by poliovirus was 0.008 and the probability of non-paralytic polio was 0.05. Taken together with the ambiguity in the clinical

definition of non-paralytic polio, these estimates suggest that $\leq 5.7\%$ of individuals infected with WPV have CNS involvement.

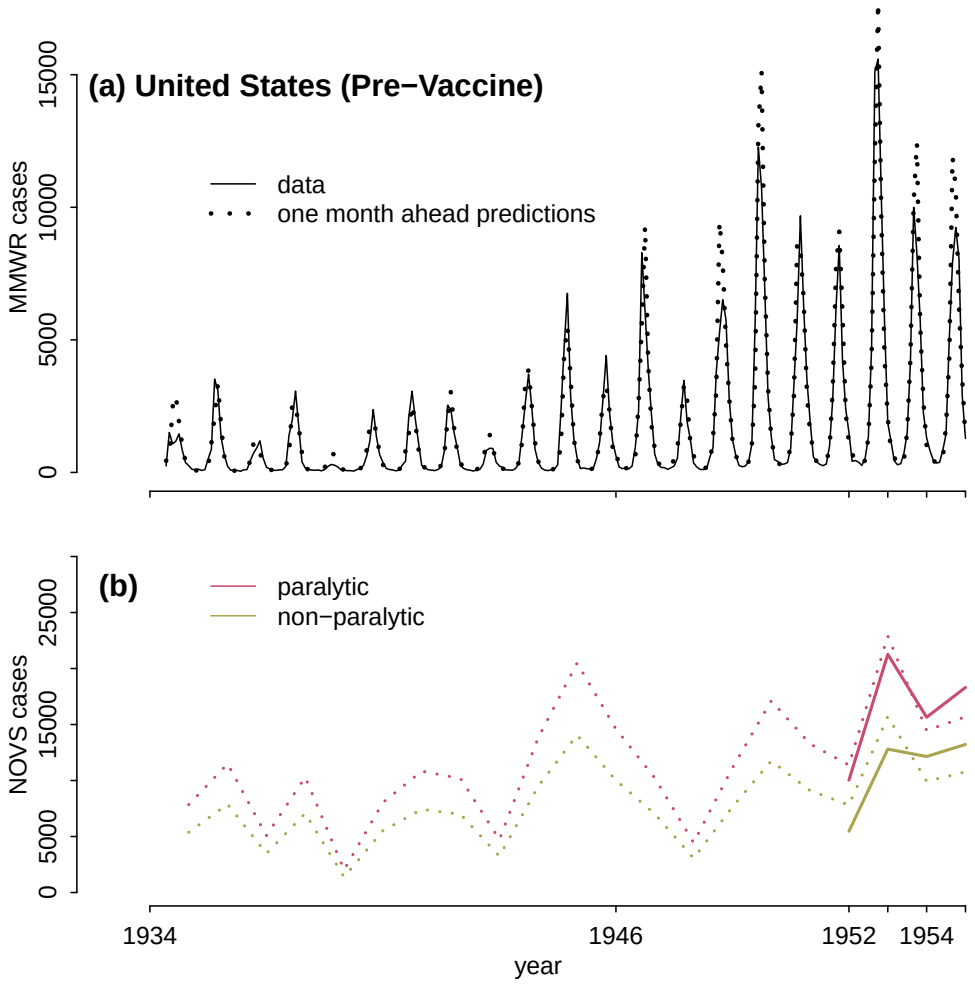


Figure 5.9: **US pre-vaccine data and one-month-ahead predictions from the MLE.** (a) Monthly MMWR polio cases and model predictions. (b) Annual NOVS polio cases that were classified as paralytic or non-paralytic polio, along with model predictions. One-month-ahead predictions are particle filtering prediction means.

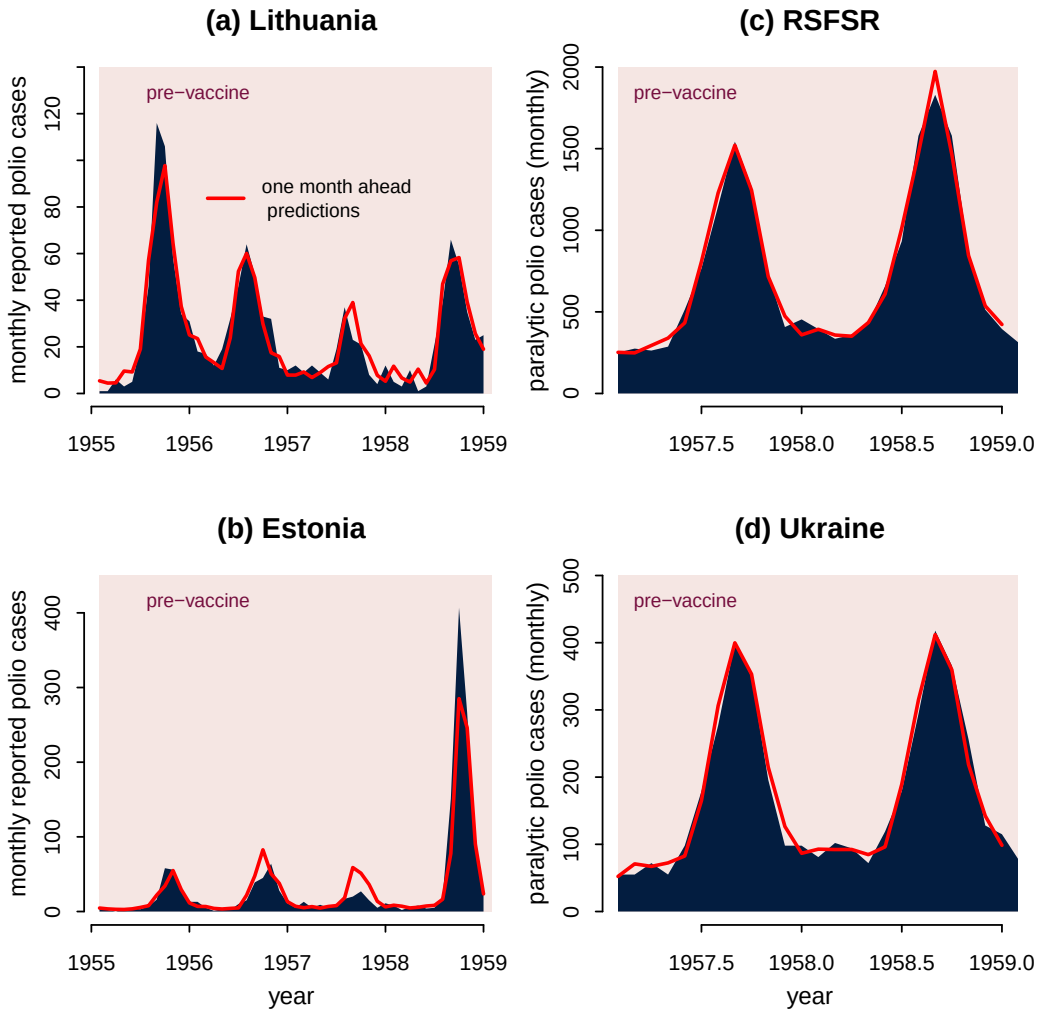


Figure 5.10: **USSR pre-vaccine data and one-month-ahead predictions from the MLE.** Monthly polio cases and one-month-ahead predictions for (a) Lithuania and (b) Estonia. Paralytic polio cases and one-month-ahead predictions for (c) the RSFSR and (d) Ukraine. One-month-ahead predictions are particle filtering prediction means.

IPV Mode of Action & Efficacy

In order to fit IPV efficacy parameters, we first estimated monthly per capita vaccine uptake rates for the US population under 15 years of age. Fig 5.11 shows the MLE vaccine coverage fit simultaneously to four time series of vaccine coverage: IPV3+ coverage, OPV3 coverage, dual IPV3+ and OPV3 coverage, and the percent of the under 15 age group with less than 3 doses of each vaccine. In accordance with the cumulative vaccine distribution data in Fig 5.4, we assumed a linear increase in IPV coverage from 1955–1958, years for which we lacked vaccine coverage data. We expected trends in cumulative vaccine distribution to be a good proxy for trends in vaccine coverage, because vaccine coverage is the cumulative percent of children vaccinated. Using the IPV era data from April 1955–December 1962, the MLE vaccine uptake rates, and the MLE transmission parameters, we estimated the mode of action and efficacy of IPV.

The IPV modes of action we tested were reductions in (1) susceptibility to WPV infection, (2) transmission by infected vaccinated individuals (i.e., infectiousness), (3) the probability of paralytic polio, and (4) the probability of non-paralytic polio. Modes 1 and 2 act by reducing WPV transmission in the population and could generate herd immunity. In contrast, modes 3 and 4 are effects on the pathological consequences of infection, which would protect the individual vaccinated, and reduce the observation of polio as a disease, but would not reduce transmission of WPV. The efficacy of IPV was taken as the overall reduction in WPV transmission by IPV, which we measured as the product of IPV's effect on susceptibility and infectiousness. All IPV parameters were estimated using the IPV era data from the US.

Our results indicate that IPV reduces polio transmission at the population-level by reducing susceptibility and infectiousness, either alone or in combination. Due to tradeoffs between parameter estimates for susceptibility and infectiousness (Fig 5.12a), we were unable to distinguish whether IPV acts primarily by reducing susceptibility

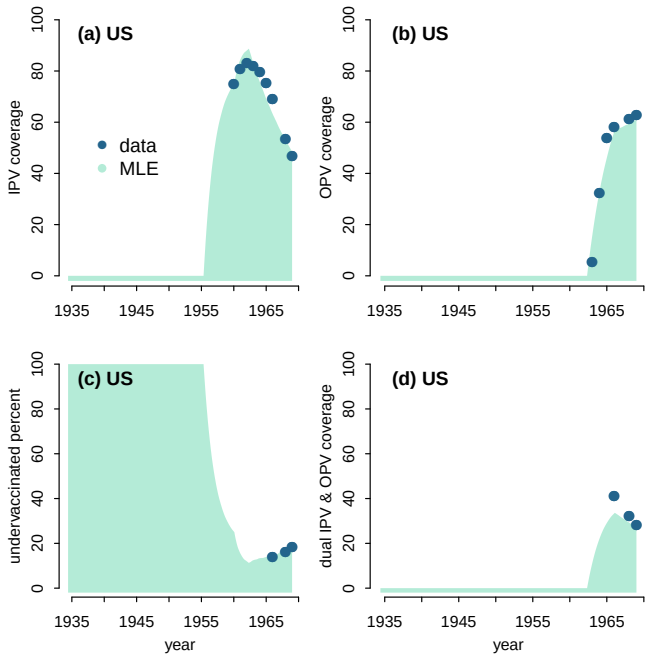


Figure 5.11: **US MLE vaccine coverage**. Vaccine coverage based on MLE vaccine uptake rates fit to the vaccine coverage data for the 1–14 year old age group.

(i.e., by preventing the establishment of infection) or by reducing infectiousness (i.e., by reducing the duration or amount of viral shedding). To deal with this tradeoff, we assumed that IPV effectively acts by reducing infectiousness, rather than susceptibility. This assumption is supported by clinical data which suggest IPV does not reduce the odds of the establishment of infection upon poliovirus challenge, but IPV can reduce the quantity of virus shed and the duration of shedding [3]. Under this assumption, we profiled IPV efficacy and estimated that IPV efficacy is 69% (95% CI 67–74%), meaning that IPV reduces transmission to 31% of the natural level (Fig 5.12b). We also estimated that IPV reduces the probability of paralytic polio by 75% and reduces the probability of non-paralytic polio by 99%. However, we have not measured the level of uncertainty in the latter two parameters. Taken together, these results suggest that IPV acts via multiple modes to reduce both transmission and pathology.

We used hindcasting to validate our estimate of IPV efficacy. Specifically, we used the fitted model to conduct December-to-December predictions of the MMWR and NOVS data. The December-to-December predictions tested the power of the model to predict "the next year's epidemic". The December-to-December predictions (Fig 5.13) demonstrated that the model predicts the decline in polio cases driven by the roll-out of IPV. However, the predictions usually underestimate the number of cases in both the MMWR and NOVS data, suggesting that our estimates are nearby, but are not, the true MLEs. We speculate that the under-predictions are due to overestimating the reduction in the paralytic and non-paralytic probabilities. If this is the case, profiling the reduction in the paralytic and non-paralytic probability will improve model fit.

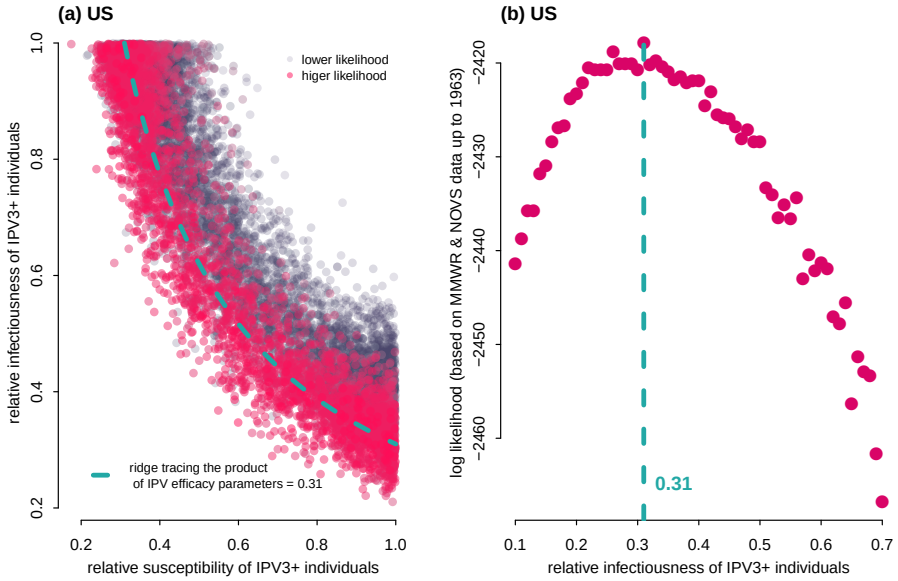


Figure 5.12: **IPV efficacy.** (a) Tradeoff in IPV efficacy parameters observed during the inference process. The parameters measured the susceptibility to infection of IPV3+ individuals, relative to unvaccinated individuals, and the infectiousness of IPV3+ individuals, relative to unvaccinated infected individuals. The tradeoff is such that there is ridge in the likelihood surface and the product of the IPV parameters is well identified. (b) Due to the tradeoff between the IPV efficacy parameters we fixed the susceptibility to infection of IPV3+ individuals to ~ 1 (meaning IPV does not reduce susceptibility) and estimated the infectiousness of IPV3+ individuals. We used the MMWR and NOVS data from the IPV era for profiling because it became computationally infeasible to use the PSU data (see Appendix). We estimated IPV infectiousness to be 0.31. Thus, IPV reduces transmission by 69% (95% CI 67–74%). In (a) we indicate the IPV efficacy ridge tracing the MLE at 0.31 (turquoise dashed line).

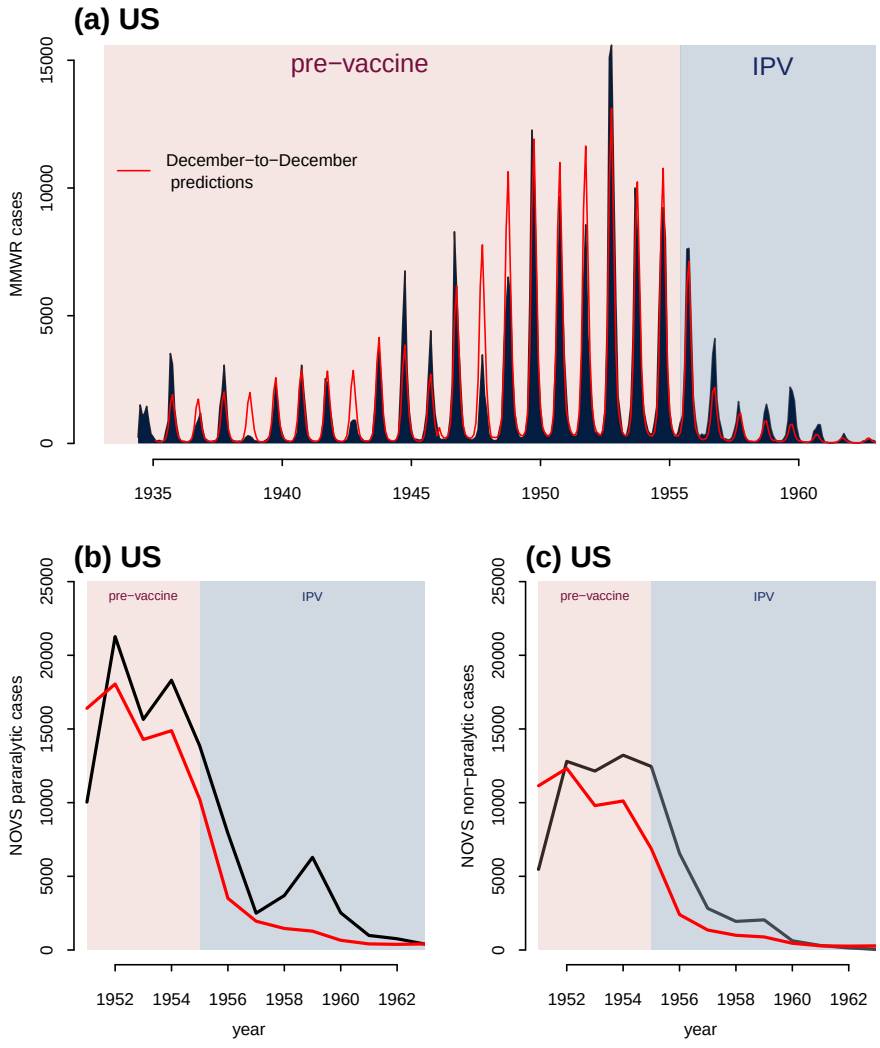


Figure 5.13: **IPV efficacy, one year ahead predictions for the US.** IPV efficacy parameters were fit to the MMWR and NOVS data up to December 1962. Using the MLE for IPV efficacy, a particle filter was run to December of each year and use to simulate forward through the next year to predict epidemic trajectories. The December-to-December predictions were done simultaneously for (a) MMWR data, (b) paralytic polio reported to NOVS, and (c) non-paralytic polio reported to NOVS. December-to-December predictions tested the ability of the model to predict the next year's epidemic. In general, the model under-predicted the number of cases, suggesting that we are nearby, but have not yet reached the true MLE.

OPV Mode of Action & Efficacy

The modes of action tested for OPV were reduction in (1) susceptibility to WPV infection, (2) transmission by infected OPV vaccinated individuals, and (3) the probability of paralytic polio. All OPV parameters were inferred using paralytic polio data from the RSFSR and then were independently validated by predicting the reduction in polio cases in Lithuania, Estonia, and Ukraine following the introduction of OPV. Since we anticipated a tradeoff in estimates of OPV efficacy parameters, as was the case for IPV, we first profiled relative susceptibility, assuming full infectiousness, and vice versa. We then profiled along all three modes of action.

In the RSFSR, we created a two-dimensional likelihood profile of OPV3 relative infectiousness and OPV3 relative paralytic probability (Fig 5.14 top), and a two-dimensional likelihood profile of OPV3 relative susceptibility and OPV3 relative paralytic probability (Fig 5.14 bottom). For both profiles we let mif estimate the OPV vaccine uptake rate. Using the MLE vaccine uptake rate from the two-dimensional profiles, we then produced a 50x50x50 three-dimensional profile of OPV relative susceptibility, infectiousness, and paralytic probability, by calculating likelihoods at all 125000 profile points. We identified the MLE OPV efficacy from the three-dimensional profile (Fig 5.15a). The MLE OPV efficacy was 0.04, indicating that OPV reduces polio transmission by 96%. Interestingly, the MLE for OPV relative susceptibility was 1, and infectiousness was 0.04, indicating that OPV acts by reducing infectiousness, rather than susceptibility. The OPV relative paralytic probability was 0.67, suggesting that OPV does reduce the probability of paralysis. We used one-step-ahead predictions in the RSFSR to validate OPV efficacy (Fig 5.15b). The one-step-ahead predictions captured the reduction in paralytic polio following the introduction of OPV; however, the model tended to overestimate the number of cases. We anticipate that higher resolution profiling will result in a MLE with a reduced OPV paralytic probability and a better fit to the vaccine era data.

We conducted independent validation of OPV efficacy by deploying the OPV efficacy MLE in the models for Lithuania, Estonia, and Ukraine. The independent validation was able to predict the decline in polio incidence in Lithuania, Estonia, and Ukraine (Fig 5.16). Once again the model often over-predicted the number of polio cases in the presence of OPV. We believe this discrepancy can be overcome with higher resolution profiling of OPV efficacy and the reduction in paralytic probability.

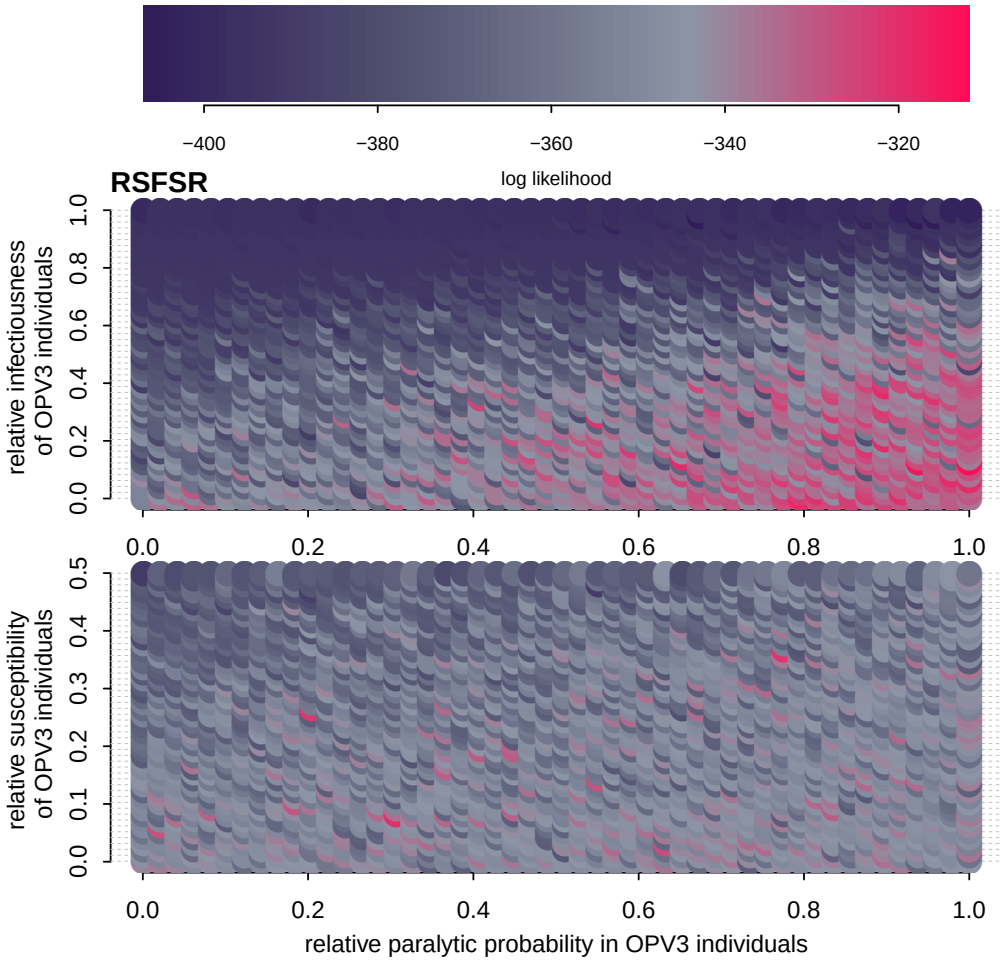


Figure 5.14: **OPV efficacy 2D profiles.** The RSFSR was used to estimate OPV efficacy parameters. Two dimensional likelihood profiles of (top) the reduction in infectiousness caused by the full 3 dose series of OPV, relative to unvaccinated individuals, and the reduction in the paralytic probability. Top profile assumes that OPV does not reduce susceptibility. (bottom) The reduction in susceptibility to infection and the reduction in the paralytic probability. The bottom profile assumes OPV does not reduce infectiousness. The relative infectiousness of OPV individuals seems to be well identified in the top profile, but the bottom profile is relatively flat, and low likelihood, suggesting that we cannot assume OPV does not reduce infectiousness.

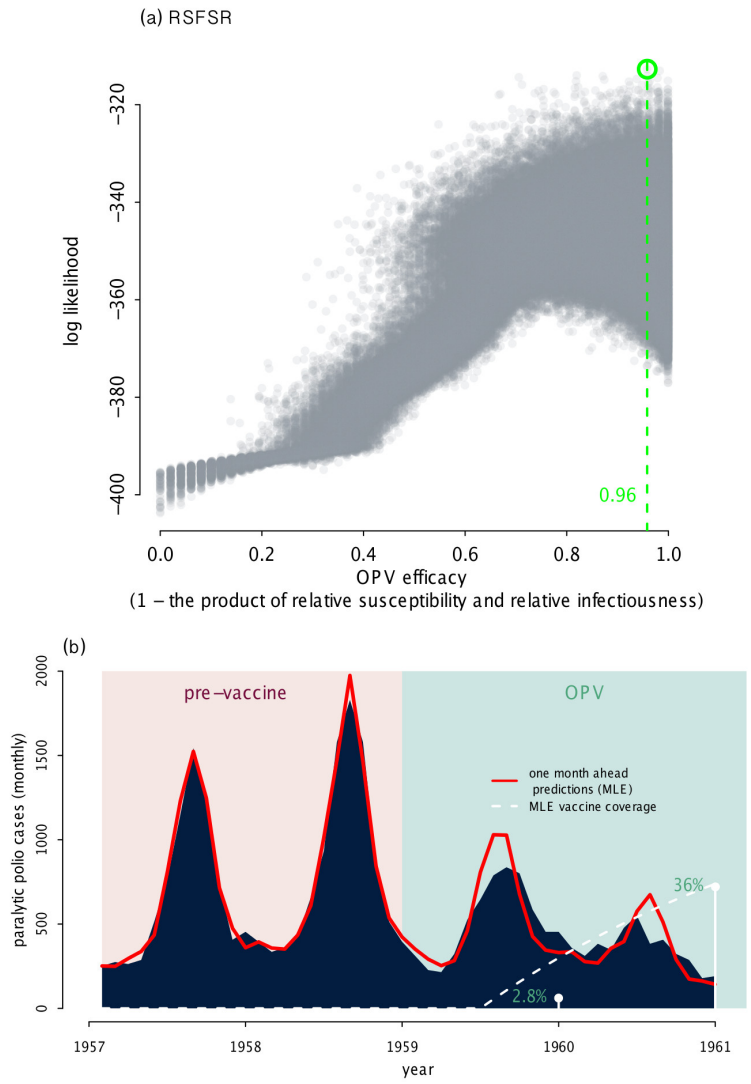


Figure 5.15: **OPV efficacy 3D profile.** The RSFSR was used to estimate OPV efficacy parameters. A three dimensional 50x50x50 likelihood profile of the reduction in infectiousness caused OPV, the reduction in susceptibility, and the reduction in the paralytic probability was created. Plotted in (a) the log likelihood of OPV efficacy across the profiled range. We measured efficacy by taking the combined effect of OPV on susceptibility and infectiousness. The MLE efficacy of 96% is indicated in green. The OPV efficacy MLE was then validated using (b) one-month-ahead predictions. The white line is the vaccine coverage based on the MLE uptake rate.

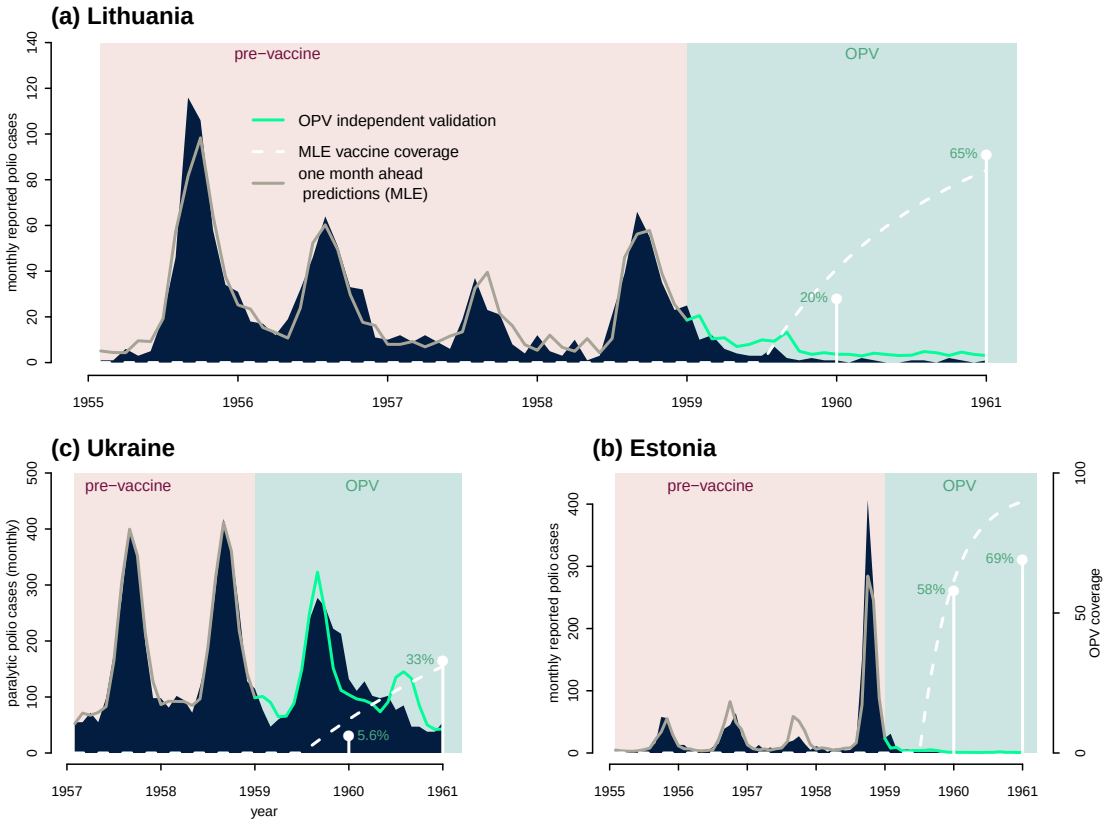


Figure 5.16: **Independent validation of OPV efficacy.** The OPV efficacy MLE from the RSFSR was used to predict out-of-fit, the cases in (a) Lithuania, (b) Ukraine, and (c) Estonia.

Discussion

Using detailed data on polio incidence and the roll-out of IPV and OPV in the US and USSR, we estimated the efficacy of IPV and OPV. In line with clinical trial data and the eradication experience [3, 4], we found that OPV is superior to IPV in reducing polio transmission. Our fitted models indicate the full 3+ dose series of IPV reduces WPV transmission by 69%, relative to unvaccinated individuals. OPV, on the other hand, reduces transmission by 96%. We found that IPV reduces transmission by either reducing susceptibility to infection or by reducing the infectiousness of vaccinated individuals. Due to uncertainty regarding IPV's mode of action, we assumed IPV reduces infectiousness, rather than susceptibility. The basis for this assumption is that multiple clinical studies have found that IPV does not reduce the odds of infection and shedding, but reduces the quantity of virus shed in stool by 63–91%. Our estimated 69% reduction in transmission suggests that viral shedding is an indicator of infectiousness. Clinical data also indicate that IPV decreases the infectious period; however, due to the monthly structure of our model we were unable to test shorter infectious periods.

The infectiousness of OPV vaccinated individuals was 0.04, relative to unvaccinated individuals. Importantly, the 96% reduction in transmission we estimated matches clinical trials that found that the odds of OPV individuals shedding virus after infection challenge was 0.13 [3]. The consistency between our IPV and OPV efficacy estimates and the clinical trial data suggests that the individual-level efficacy trials directly translated to reductions in transmission measured at the population-level. In addition to effects on transmission, we also have an indication that IPV reduces the probability of paralytic polio by 75% and the probability of non-paralytic polio by 99%. Given the lack of reported non-paralytic polio in OPV3 individuals, we assumed the full series of OPV offers complete protection from non-paralytic polio, an assumption that requires further testing. As for paralytic polio, our MLE indi-

cates that OPV reduces the probability of paralysis in infected individuals by 33%. However, this is likely to change with refinement of the MLEs.

An important consideration for vaccine evaluation is how efficacy changes through time. Mucosal immunity provided by OPV is known to wane over time [22]. Since our inference was done using data from the initial roll-out of IPV and OPV, we did not have the longer-term data necessary to evaluate waning vaccine-derived immunity. However, if eradication is the goal, the waning of immunity over the course of several years would be inconsequential if mass immunization raised "recent" vaccine coverage high enough to hit the eradication target, and waning occurred post-eradication. In other words, in the face of imperfect vaccines, globally-synchronized mass immunization can be used to boost age groups, creating a honeymoon period with high levels of immunity that could drive WPV to extinction.

Our profiles of IPV and OPV efficacy indicated that the reduction in transmission provided by each vaccine was well identified. As previously mentioned, for IPV we were unable to distinguish whether the mode of action was reduction in susceptibility to infection, reduction in infectiousness, or both. The next step would be to fix the infectiousness and estimate the susceptibility of vaccinated individuals. Due to the strict tradeoff observed between IPV susceptibility and infectiousness in Fig 5.12, we predict that the MLE for IPV efficacy will be ~ 0.31, regardless of the mode of action.

We believe there is one parameter with the potential to affect our estimate of IPV efficacy, and that is the immigration rate of infection. The immigration rate was a parameter that accounted for introduction of WPV from regions outside the US. Since the majority of WPV infections are silent or subclinical, the immigration rate was included in the model to account for WPV introduction into the population by the movement of people with subclinical infections. Given the large amount of international travel into the US in the post-World War II era, we decided to include immigration into the force of infection. The immigration rate was estimated along

with the other epidemiological parameters in the pre-vaccine era. By profiling the immigration rate (see Appendix), we found that the MLE was at the lower range of our profile. Although immigration had a large impact on local polio epidemics [10], it did not seem to have a substantial impact on the national-level epidemics in the US. If we set immigration to zero, the pre-vaccine trajectories changed very little. However, as WPV circulation decreased after years of vaccination with IPV, we believe the immigration rate had a large effect on infection dynamics. This is because, as the number of locally-derived WPV infections decreased, the relative effect of the immigration rate on the force of infection increased. The immigration rate is important because, if we had estimated a lower immigration rate, we would have a lower baseline transmission rate, and therefore would require less efficacious IPV to see the same reduction in cases. A very important follow up will be to more precisely estimate the immigration rate in the US and then refit IPV efficacy. The confidence intervals for IPV efficacy could also be calculated to account for uncertainty in the immigration rate.

Other improvements to the US model include better estimates of IPV uptake. The model currently captures vaccine coverage well, but the decline in IPV coverage during the transition to the OPV era could be improved. With improved estimates of IPV uptake and efficacy in the US, we believe we will be capable of fitting the full observation model with the PSU data. As previously mentioned, the fitted IPV model underestimated the number of cases in MMWR and NOVS (Fig 5.13). We believe this was due to overestimating the reduction in the IPV paralytic probability and the non-paralytic probability. Simulations from the fitted model had too few IPV paralytic and non-paralytic cases reported to PSU. We suspect that this is why we hit the limits of floating-point number representation when trying to calculate the joint likelihood of the MMWR, NOVS, and PSU data.

As for OPV efficacy, the fitted OPV efficacy in the RSFSR and independent val-

idation in Lithuania, Estonia, and Ukraine make a strong case for the accuracy of our estimate. Although the estimate of OPV efficacy may be accurate, finer scale profiling is warranted. The fitted models over-predicted the cases in the USSR, suggesting that the true MLE has either a higher OPV efficacy (i.e., a greater reduction in infectiousness or susceptibility) or a reduction in the OPV paralytic probability. It was unexpected that the relative OPV susceptibility was 1, meaning OPV does not reduce susceptibility. However, we assume further profiling in the neighborhood of the OPV efficacy MLE will result in a lower estimate of this value and a better fit to the data. An exciting extension to the independent validation of OPV efficacy will be to deploy the MLE OPV efficacy to predict the OPV era data from the US. Similarly, the IPV efficacy estimated in the US could be independently validated using data from another country that relied on IPV, such as the Netherlands.

In conclusion, we found that OPV reduces WPV transmission on the population level by reducing the infectiousness of vaccinated individuals by 96%. More importantly, we demonstrated that IPV acts either by reducing the susceptibility to infection or infectiousness by 69%. Although the MLE indicates IPV efficacy is 69%, the efficacy may be higher because we used data from the 1950s and 1960s. The regulations for vaccine licensure changed in 1968 and required higher potency IPV [23]. Thus, IPV licensed and used today may have higher efficacy than the original rolled-out in the US.

Clinical IPV efficacy studies— that simulated WPV challenge using OPV—suggested that IPV does not protect from the establishment of infection, but reduces the quantity of WPV shed in stool and the duration of shedding [3]. Our results indicate that this individual-level estimate scales up to the population-level and IPV does in fact reduce transmission. This is a hopeful indication that, if maintained at high levels of vaccine coverage, IPV could significantly reduce polio transmission and result in herd immunity. Given an IPV efficacy of $> 69\%$, a "back of the envelope calculation"

suggests that IPV could be used to lower the reproductive number of WPV < 1 in regions with an off-season WPV reproductive number ≤ 3. If this is the case, WPV could be driven to local extinction with high IPV coverage. Based on studies of imperfect vaccines [24], we estimate that with an off-season reproductive number ≤ 3, 96% vaccine coverage with IPV alone, or 69% vaccine coverage with OPV alone, could eradicate polio. Given, that much of the world has OPV-derived immunity, and some older individuals have naturally-acquired immunity, eradication might potentially be achieved with a switch to IPV and coverage $< 96\%$.

Our results eliminate much of the uncertainty that surrounded IPV efficacy, and give a hopeful outlook for the successful withdrawal of OPV and reliance on IPV. However, aside from questions regarding efficacy, another reason that IPV has limited use is due to the difficulty of administering the 3 dose injection of IPV in mass immunization campaigns. Fortunately, this year a new microneedle technology was developed for administering IPV, which can be done cheaply, with limited training, and simple transport [25]. In light of our results, new technology for administering IPV, and the drawbacks of OPV, we suggest that the cautious withdrawal of OPV be fully considered under (1) the explicit knowledge of an IPV efficacy of $\sim 69\%$ and (2) the dramatic seasonal fluctuations in WPV's reproductive number and the limited seasonal "window" for eradication.

Epilogue

Before publication of this work there are several extensions that must be completed. The list below contains some of the primary extensions.

- We will profile the IPV uptake rate to improve the fit to IPV coverage data in the US
- We will test whether infection immigration is necessary in the US transmission

model and we will refit IPV efficacy parameters under a "no immigration" assumption.

- Using the US data, we will profile IPV efficacy along the IPV-efficacy-ridge shown in Fig 5.12. The profiling will be along 3 parameter dimensions: IPV efficacy, IPV relative paralytic probability, and IPV relative non-paralytic probability.

- For the USSR, we will be expanding the data both in the pre-vaccine era and the vaccine era. With the expanded data, we will refit the WPV transmission parameters in each region in the USSR

- Using expanded data from the USSR, we will once again simultaneously profile OPV relative susceptibility, OPV relative infectiousness, and OPV relative paralytic probability.

Bibliography

[1] Mangal TD, Aylward RB, Grassly NC (2013) The Potential Impact of Routine Immunization with Inactivated Poliovirus Vaccine on Wild-type or Vaccine-derived Poliovirus Outbreaks in a Posteradication Setting. American Journal of Epidemiology .

[2] Nathanson N, Kew OM (2010) From Emergence to Eradication: The Epidemiology of Poliomyelitis Deconstructed. American Journal of Epidemiology 172: 1213–29.

[3] Hird TR, Grassly NC (2012) Systematic Review of Mucosal Immunity Induced by Oral and Inactivated Poliovirus Vaccines Against Virus Shedding Following Oral Poliovirus Challenge. PLoS Pathogens 8: e1002599.

[4] Hull HF, Ward Na, Hull BP, Milstien JB, De Quadros C (1994) Paralytic Poliomyelitis: Seasoned Strategies, Disappearing Disease. Lancet 343: 1331–1337.

[5] Jafari H, Deshpande JM, Sutter RW, Bahl S, Verma H, et al. (2014) Efficacy of Inactivated Poliovirus Vaccine in India. Science 345: 922–5.

[6] Melnick JL (1996) Current Status of Poliovirus Infections. Clinical Microbiology Reviews 9: 293–300.

[7] Combelas N, Holmblat B, Joffret ML, Colbère-Garapin F, Delpeyroux F (2011) Recombination Between Poliovirus and Coxsackie A Viruses of Species C: A Model of Viral Genetic Plasticity and Emergence. Viruses 3: 1460–84.

[8] Jegouic S, Joffret ML, Blanchard C, Riquet FB, Perret C, et al. (2009) Recombination Between Polioviruses and Co-circulating Coxsackie A viruses: Role in the Emergence of Pathogenic Vaccine-Derived Polioviruses. PLoS Pathogens 5: e1000412.

[9] World Health Organization (2013) Polio Eradication & Endgame Strategic Plan 2013-2018. Technical report.

[10] Martinez-Bakker M, King AA, Rohani P (2015) Unraveling the Transmission Ecology of Polio. PLoS Biology (in press).

[11] Chumakov M, Voroshilova M, Vasilieva K, Bakina M, Dobrova I, et al. (1959) Preliminary Report on Mass Oral Immunization of Population Against Poliomyelitis with Live Virus from A.B. Sabin's Attenuated Strains. In: Live Polivirus Vaccines: Papers Presented and Discussion Held at the First International Conference of Live Polio Vaccines. pp. 517–529.

[12] Chumakov M, Voroshilova M, Drozdov S, Dzagurov S, Lashkevich V, et al. (1960) On the Course of Mass Immunization of the Population in the Soviet Union with the Live Poliovirus Vaccine from Albert B. Sabin's Strains. Report No. 3 as of 1 June 1960. In: Second International Conference on Live Poliovirus Vaccines. pp. 413–428.

[13] Chumakov M, Voroshilova M, Drozdov S, Dzagurov S, Lashkevich V, et al. (1961) Some Results of the Work on Mass Immunization in the Soviet Union with Live Poliovirus Vaccine Prepared from Sabin Strains. Bulletin of the World Health Organization : 79–91.

[14] Langmuir A, Nathanson N, Hall WJ (1956) The Surveillance of Poliomyelitis in the United States in 1955. American Journal of Public Health 46: 75–88.

[15] Francis T, Napier J, Voight R, Hemphill F, Wenner H, et al. (1955) Evaluation of the 1954 Field Tial of Poliomyelitis Vaccine: Final Report. Technical report, Poliomyelitis Evaluation Center, University of Michigan, Ann Arbor, Michigan.

[16] Nathanson N, Hall W, Thrupp L, Forester H (1957) Surveillance of Poliomyelitis in the Unites States in 1956. Public Health Reports 72: 381–392.

[17] Schonberger LB, Kaplan J, Kim-farley R, Moore M (1984) Control of Paralytic Poliomyelitis in the United States. Reviews of Infectious Diseases 6: S424–S426.

[18] Horstmann DM (1991) The Sabin Live Poliovirus Vaccination Trials in the USSR, 1959. Yale Journal of Biology and Medicine 64: 499–512.

[19] Debre R, Thieffry S (1955) Symptomatology and Diagnosis of Poliomyelitis. In: Poliomyelitis.

[20] King AA, Nguyen D, Ionides E (2015) Statistical Inference for Partially Observed Markov Processes via the R Package pomp. Journal of Statistical Software (in press).

[21] Trevelyan B, Smallman-Raynor M, Cliff AD (2005) The Spatial Dynamics of Poliomyelitis in the United States: From Epidemic Emergence to Vaccine-Induced Retreat, 1910-1971. Annals of the Association of American Geographers 95: 269–293.

[22] Grassly NC, Jafari H, Bahl S, Sethi R, Deshpande JM, et al. (2012) Waning Intestinal Immunity after Vaccination with Oral Poliovirus Vaccines in India. The Journal of Infectious Diseases 205: 1554–61.

[23] PSU (1980) Poliomyelitis Surveillance Report 1977-1978 Summary. Technical report, US Centers for Disease Control and Prevention.

[24] Magpantay F, Riolo M, Domenech De Celles M, King A, Rohani P (2014) Epidemiological Consequences of Imperfect Vaccines for Immunizing Infections. SIAM Journal of Applied Mathematics 74: 1810–1830.

[25] Edens C, Dybdahl-Sissoko NC, Weldon WC, Oberste MS, Prausnitz MR (2015) Inactivated Polio Vaccination using a Microneedle Patch is Immunogenic in the Rhesus Macaque. Vaccine : 8–15.

CHAPTER VI

Conclusion

An open challenge in disease ecology is to reveal the mechanisms and implications of infectious disease seasonality [1, 2, 3, 4, 5]. My dissertation research was a starting point for addressing disease seasonality in what will likely be a career-long endeavor. I will now generally discuss disease seasonality and present ideas for extending my dissertation research. Seasonal drivers of disease can be separated into four interconnected categories: (1) environmental drivers, (2) host behavior, (3) host phenology, and (4) exogenous biotic drivers. I will briefly discuss each of these four categories and frame research approaches for studying seasonal cycles.

6.1 Environmental Drivers

Environmental drivers are abiotic conditions that influence transmission via their effects on hosts and/or parasites. Such drivers include quantitative features of the environment; classical examples are temperature and rainfall, which influence a variety of infectious diseases [2]. Environmental drivers can impact pathogen survival during transitions between hosts. Transitions can take place during short time windows (e.g., for droplet transmitted infections) or long time windows (e.g., for parasites with environmental life stages). In addition to their impact on pathogens, environmental drivers can also influence host susceptibility to infection. For example, environmental

conditions can moderate the host immune response and increase the susceptibility of cells to infection [6] or pose seasonal challenges (such as food limitations) that leave hosts vulnerable to infection or pathology [7]. Environmental drivers, along with host behavior, are the seasonal drivers that have received the most attention from the scientific community. Presumably, environmental drivers have been a main focus because they strongly covary with seasonal disease incidence. Associations between environmental conditions and disease incidence are causal in some cases, with influenza transmission being the most notable example [8]. However, covariation is expected between seasonal phenomena and we should also consider seasonal features of the biotic environment in addition to the abiotic.

6.2 Host Behavior

Transmission seasonality is sometimes due to seasonal host behavior, most commonly fluctuations in host contact rates throughout the year. School-term transmission of measles is an excellent example of seasonal contact rates [9]. Studies of human diseases have been enlightening in this area; however, I would argue that wild animal systems offer a richer arena for studying seasonal host behavior and disease transmission. In the case of measles, children contact other children year-around, and school-terms increase this rate; whereas, in non-human systems, risky contacts could be seasonally isolated in time.

Let us consider sexually transmitted diseases in seasonally breeding mammals. Seasonal transmission would not simply modulate a pre-existing non-zero contact rate, but there is a complete absence of sexual contacts (and thus transmission) outside of the breeding season. I propose that isolation in time is the most extreme form of seasonal transmission, and it is an important area for future research. I predict seasonally isolated transmission exists for STDs in wildlife, but may also occur for parasites with other transmission modes. Discrete windows of transmission are likely

to have evolutionary consequences for parasite life history. Cattadori et al. [10] pointed out that, when transmission is restricted to a short seasonal window, natural selection will favor parasites with "long-lived infective stages". I further speculate that transmission isolated in time will have dynamical consequences that make these disease systems unique from those with year-around transmission.

Sexual contacts are not the only seasonal "risky" behavior. Seasonal engagement in risky behavior may also occur in other disease contexts, including for infections transmitted during bouts of fighting. For example, the Tasmanian devil contact network varies between the mating and non-mating season and could influence transmission of devil face tumor disease, the transmission of which is facilitated by aggressive behavior [11, 12]. Mating and aggression can elevate disease risk, but it is important to acknowledge that some behaviors mitigate disease risk. Disease mitigation behaviors include grooming to remove ectoparasites [13] and self-medicating [14], but to my knowledge there are few studies of seasonality in risky behavior and disease mitigation behavior. In the very least, disease ecology would benefit from a comprehensive review of the ways in which animal behaviors affect disease transmission and how influential behaviors are structured in space (i.e., across populations), through time (i.e., seasonally), and how they vary with age. By bringing a disease ecology focus to analyzing observational studies of animal behavior and the growing amount of telemetry data collected from wild animals we could form hypotheses regarding how the ecology of infection and behavior overlap.

6.3 Host Phenology

Host behavior can also seasonally structure disease risk via the geographical localization of hosts. Hosts seasonally engage with different aspects of their environment. For hosts that migrate or hibernate, contact with risky environments can be seasonal [15]. Migration and hibernation are part of host phenology, which includes host life

history, annual cycles, and endogenous circannual rhythms. As discussed in Chapter III, in disease ecology, relevant host phenology includes, but is not limited to, birth seasonality, seasonal re-structuring of immunity, cycles of nutrition and body condition, hibernation, and migration. Unlike environmental drivers and host behavior, host phenology can affect diseases dynamics by means other than transmission. In SIR-type models of disease transmission, host phenology can generate seasonality in (1) susceptible recruitment, (2) host susceptibility to infection, (3) host infectiousness, (4) the recovery rate, (5) the mortality rate (both natural and disease-induced), and (6) symptomatology. An import extension to my dissertation research is to test whether is it possible to determine which of these six seasonal mechanisms is acting in specific infectious disease systems. The six mechanisms presumably have different dynamical effects and would therefore leave imprints in long-term incidence data. Extensive simulation studies are required to identify the dynamical effects of these seasonal mechanisms acting in isolation and in combination. Simulation studies will provide a foundation for determining the types of data required for distinguishing among seasonal mechanisms. This challenge becomes greater when considering chronic infection with parasites that have complex life histories and phenology of their own.

6.4 Exogenous Biotic Drivers

In addition to abiotic, behavioral, and phenological features of host-parasite systems, hosts and their parasites are embedded in ecological networks that have additional seasonal aspects. I refer to interactions within ecological networks as exogenous biotic drivers because they are exogenous to the host-parasite dyad. Exogenous biotic drivers include parasite-parasite interactions and host community ecology. I did not include exogenous biotic drivers in my dissertation research. In recent years, however, studies highlighting the importance of immune mediated parasite interactions

and host heterogeneity in multi-host systems have demonstrated the need to consider community ecology of hosts and their parasites [16, 17]. Interactions between hosts/parasites and members outside of the dyad undoubtedly display seasonality, as do nearly all aspects of ecology. Since seasonality is an inherent feature of ecological systems, and seasonal incidence is a feature of infectious diseases, I propose that progress can be made if we (the scientific community) embrace the assumption that "everything is seasonal". Working under this assumption would prevent us from futilely establishing correlative relationships among seasonal phenomena *ad nauseam*. Instead, we can focus our attention on building theory that will provide a deeper understanding of seasonal mechanisms, and we can learn how to identify imprints of seasonal drivers in disease data. My general conclusion is, since "everything is seasonal", then everything will covary (usually with some phase shifts). Therefore, seasonal covariance alone is not useful for establishing seasonal drivers. Long-term parallel data of potential seasonal drivers and disease incidence could be used to establish causality or suggest causal mechanisms if we look beyond seasonal covariance. I predict that the information contained in interannual variation and anomalous years holds the key to establishing causality.

6.5 Understanding Seasonality

Let us do a thought experiment to illustrate how anomalous years and interannual variation could be used establish the causal mechanism of disease seasonality. Consider a human disease with seasonal peak incidence in the summer, since incidence peaks in summer, it would have a strong positive relationship with temperature, photoperiod, and many other summer-related features of the environment and host population. To highlight how strongly non-causal seasonal factors correlate with disease incidence, we could quantify the positive relationship between disease incidence and the sales of bathing suits or the frequency of back-yard barbecues, both of which

presumably go up in summer. Now imagine we build transmission models for this infection and test the four potential seasonal drivers (i.e., temperature, photoperiod, swim suite sales, and barbecues) as modulators of host susceptibility to infection. If the disease displays either interannual variation in epidemic size or anomalous years with differences in epidemic timing, then only the data from the causal driver would aid us in predicting the variation observed across years, assuming the causal driver is responsible for some of the variation.

In this thought experiment, photoperiod might be ruled out as the causal driver because there is no interannual variation in photoperiod for any given latitude. Therefore, we need to consider a better metric of light conditions, such as sunlight hours calculated using data on photoperiod and cloud cover. After correcting the sunlight metric, we find that all four models capture the seasonal structure of epidemics, because they all covary with disease incidence. However, only the model with sunlight hours accurately predicts the between-year variation in epidemic size and the anomalous years. This suggests that sunlight is the primary (causal) seasonal driver, and the model indicates that sunlight acts by modulating host susceptibility. As is the case with any scientific endeavor, we would be left with a new problem: to identify the mechanism of action on a biological level. We would need to address whether sunlight is modulating host susceptibility by modifying host immunity, or if it is instead influencing the transmission rate by affecting the pathogen or some other feature of the host-pathogen system. The effects of seasonal drivers are multi-layered. Thus, in order to better understand seasonality we must work at multiple organizational levels of science. Geophysical factors, host population ecology, and within-host biology will need to be integrated in the practice of studying seasonality. In my own work, I have identified several new avenues for exploration that would require interdisciplinary research.

6.6 Expanding on Birth Seasonality

My work on birth seasonality had one major downfall, the lack of an empirical example demonstrating the effect of birth seasonality in real-world epidemics. In Chapter II, I studied the effect of birth seasonality on New York City measles epidemics. In hindsight, the data from New York were unlikely to prove insightful. Based on my simulation studies, the birth amplitude needed to be fairly high in a population (i.e., $> 20\%$) in order to have an observable effect on measles incidence. However, New York City had a birth amplitude well below 20% during the time period I studied. I did not observe an effect of birth seasonality on estimates of measles transmission in New York City, and I attributed this to the low amplitude of birth seasonality.

I chose to investigate measles for two reasons. First, measles was the focus of foundational studies in disease ecology [9, 18, 19, 20, 21, 22] making it familiar to many disease ecologists. Second, the data were readily available for New York City. However, I recognize that my study of birth seasonality and disease needs to be repeated using data from populations with higher amplitude birth seasonality. Furthermore, it is likely that birth seasonality plays a larger role in disease systems with an infant age of first infection, such as rotavirus or hand foot and mouth disease, but this remains to be tested.

The difficulty of studying the impact of birth seasonality on childhood diseases—which I will now refer to as early-life infections to include wildlife diseases—is that births do not directly enter the susceptible pool. My work on measles and polio required me to grapple with the presence of maternal immunity in infants. In mammals, maternal immunity is transferred from mothers to their offspring during pregnancy and lactation. Females transfer maternal antibodies and leukocytes (i.e., immune cells) that help fight early-life infections in their infants [23, 24]. The exciting next step in the study of birth seasonality is to determine how it interfaces with the evo-

lutionary ecology of maternal immunity.

Maternal immunity is short-lived, usually lasting weeks to several months [23]. In seasonally breeding mammals, maternal immunity is present in the population at the same time as the birth pulse. Since birth pulses replenish the susceptible pool, and can fuel outbreaks, I hypothesize that maternal immunity is an evolutionary response for mitigating the dynamical consequences of birth seasonality. Furthermore, there is evidence that the interaction between maternal immunity and infection is not a one-way effect, but a feedback. A study of leukocytes in breastmilk found that the abundance of leukocytes increased if either the mother or the infant was infected [24]. This suggests there are dynamical feedbacks between infection and maternal immunity. Transmission models that bridge the population-level and the individual-level could be used to study the interaction between, and evolution of, birth seasonality, maternal immunity, and disease seasonality. This work would require an interdisciplinary group of disease ecologists, immunologists, and clinicians. The study of maternal immunity and seasonal cycles of reproduction should not be studied in isolation; additional seasonal features of host life history, including rhythms in immunity and nutrition, would also require consideration.

6.7 Exploring Biological Rhythms

My work on biological rhythms and infection raised more questions than it answered. Seasonal cycles of gene expression in the human immune system, which I discussed in the introduction, are suggestive that the human immune system has endogenous circannual cycles. In order for rhythms to be endogenous, they must persist under constant environmental conditions. The test of this requires long-term experiments where subjects are kept in an environment absent of seasonality, which is one of the reasons laboratory studies of circannual rhythms are rare. In my view, though unconventional, if desired it would be possible to test if seasonal rhythms in

immunity are endogenous by studying crews working on the international space station. NASA is currently conducting its first 1-year mission, which would encompass a full circannual cycle. Astronauts are particularly promising for studies of immunity rhythms because (a) they live in a relatively constant environment, and (b) they are not exposed to pathogens that would stimulate an immune response and bias results. The study of biological rhythms in space is promising, NASA is currently conducting studies on circadian rhythms; therefore, the opportunities space flight present for chronobiology have been acknowledged [25].

Switching to the parasite perspective, F. Hawking extensively documented the circadian and seasonal cycles of parasites in the 1970s [26], but the field has progressed little since. Endoparasites pose a difficult problem for studying rhythms because they live within other organisms. It is unclear how parasite rhythms are entrained. Are parasite rhythms entrained to the rhythms of their host? If parasites rhythms entrain to host rhythms, then parasites can be placed in a "constant" environment by infecting hosts lacking functioning biological clocks. Fortunately, knockouts of circadian clock genes (e.g., *Per2* and *Bmal1*) exist in mouse models and are often used in circadian clock studies. Regardless of whether parasite rhythms or seasonal host immunity rhythms are endogenous, it is important to determine whether they have consequences for disease.

The next step for studying biological rhythms and infection will be to expand upon the pre-existing work on circadian rhythms. Circadian studies are data rich because they are short in period, and the circadian clock is well understood. Whereas, the mechanistic underpinnings of the circannual clock remain a mystery. There are several research groups studying circadian rhythms in malaria parasites. Over the next several years, the rodent malaria system will undoubtedly progress the study of rhythms and infection. As a complement to experimental studies, an open research area is to develop dynamic models of the multi-oscillator system that includes host

rhythms, parasite rhythms, and the environmental cycles to which these rhythms entrain. Models of adaptive dynamics could be used to investigate the evolutionary trajectories of interacting host-parasite rhythms; whereas, within host models of infection dynamics could be used to explore the effect of rhythms on parasite populations and pathology. Theory could also be developed to study circannual rhythms in hosts and parasites. However, since data are limited for circannual rhythms, it would be difficult to ground the theory with data.

6.8 Polio's Temporal Niche

Polio's temporal niche is the summer and autumn. Poliovirus transmission is limited during the wintertime low season but the mechanism responsible for this remains unknown. Together with previous studies of polio seasonality [27, 28, 29, 4], my results narrowed the list of potential mechanisms to an environmental or host factor that displayed latitudinal variation in its amplitude and phase. I propose temperature, relative humidity, sunlight hours, and seasonal immunity as viable seasonal drivers. Temperature, humidity, and sunlight hours (i.e., photoperiod taken together with cloud cover) vary seasonally and latitudinally. Temperature and humidity have been shown to have a strong relationship with polio cases, with the rise in temperature and humidity preceding the rise in cases by 2–3 months [29, 28]. My fitted transmission models from the US estimated the peak in poliovirus transmission to precede the incidence peak by $\sim$ 2 months. Since temperature and humidity have strong seasonal covariance with poliovirus transmission, they meet the minimum criteria for potential seasonal drivers. Temperature and humidity also have an effect on poliovirus survival in aerosol, with poliovirus surviving best in conditions reflective of northern hemisphere summers, conditions opposite those that favor influenza virus survival [30].

As shown in the introduction, human immunity is seasonally structured with

distinct summertime and wintertime gene expression profiles [31]. If the seasonal structure of immunity varies with latitude, it could account for polio's spatiotemporal transmission structure. Polio transmission models could be used to test whether temperature, humidity, sunlight, or immunity affect transmission. The models I used in Chapter IV could be extended to include seasonal drivers as covariates. Based on my one-step ahead-predictions, which moderately over- or under-predicted peak cases each year, the model-data fit could be improved for polio in the US. If the model contained the correct seasonal driver, it could better fit the data by predicting the size of the epidemic peak, which varies significantly from year to year. As previously discussed, I anticipate that the information contained in interannual variation in epidemic size and anomalous years holds the key to establishing seasonal causality.

It is important to identify polio's seasonal driver because knowledge of the driver would allow transmission models to be extended to contemporary studies of polio in endemic countries. By using historical data with recurrent epidemics, I had data sufficient to estimate the seasonal transmission rate. Today, due to widespread vaccination, polio cases are rare and contemporary data are depauperate of information needed to estimate transmission seasonality. If historical data were used to establish the relationship between environmental conditions and polio transmission, then environmental conditions from polio endemic countries could be used in models of current outbreaks. Polio's environmental driver is the biggest knowledge gap in polio's ecology. Non-polio enteroviruses, such as coxsackievirus and echovirus, also display elevated incidence in the summer [32]. I am therefore optimistic that uncovering the mechanisms of poliovirus seasonality will result in a better understanding of the enteroviruses in general. This sentiment was shared in 1949 by the epidemiologist H. Gear, who wrote: *"It must be admitted that the reasons for the seasonal incidence of poliomyelitis remain obscure. When they have been elucidated perhaps much of the epidemiology of this disease will be solved"* [33].

Bibliography

[1] Dowell SF (2012) Seasonality - Still Confusing. Epidemiology and Infection 140: 87–90.

[2] Altizer S, Dobson A, Hosseini P, Hudson P, Pascual M, et al. (2006) Seasonality and the Dynamics of Infectious Diseases. Ecology Letters 9: 467–484.

[3] Fisman DN (2007) Seasonality of Infectious Diseases. Annual Review of Public Health 28: 127–143.

[4] Dowell SF (2001) Seasonal Variation in Host Susceptibility and Cycles of Certain Infectious Diseases. Emerging Infectious Diseases 7: 369–374.

[5] Grassly NC, Fraser C (2006) Seasonal Infectious Disease Epidemiology. Proceedings of the Royal Society B: Biological Sciences 273: 2541–2550.

[6] Foxman EF, Storer JA, Fitzgerald ME, Wasik BR, Hou L, et al. (2015) Temperature-dependent Innate Defense Against the Common Cold Virus Limits Viral Replication at Warm Temperature in Mouse Airway Cells. Proceedings of the National Academy of Sciences of the United States of America 112: 827–832.

[7] Pedersen AB, Greives TJ (2008) The Interaction of Parasites and Resources Cause Crashes in a Wild Mouse Population. The Journal of Animal Ecology 77: 370–377.

[8] Lowen AC, Mubareka S, Steel J, Palese P (2007) Influenza Virus Transmission is Dependent on Relative Humidity and Temperature. PLoS Pathogens 3: 1470–1476.

[9] Soper HE (1929) The Interpretation of Periodicity in Disease Prevalence. Journal of the Royal Statistical Society 92: 34–73.

[10] Cattadori IM, Boag B, Bjørnstad ON, Cornell SJ, Hudson PJ (2005) Peak Shift and Epidemiology in a Seasonal Host-Nematode System. Proceedings of the Royal Society B: Biological Sciences 272: 1163–1169.

[11] Hamede RK, Bashford J, McCallum H, Jones M (2009) Contact Networks in a Wild Tasmanian Devil (Sarcophilus harrisii) Population: Using Social Network Analysis to Reveal Seasonal Variability in Social Behaviour and its Implications for Transmission of Devil Facial Tumour Disease. Ecology Letters 12: 1147–1157.

[12] Bostanci A (2005) A Devil of a Disease. Science 307: 2005.

[13] Mooring MS, Blumstein DT, Stoner CJ (2004) The Evolution of Parasite-defence Grooming in Ungulates. Biological Journal of the Linnean Society 81: 17–37.

[14] Choisy M, de Roode JC (2014) The Ecology and Evolution of Animal Medication: Genetically Fixed Response versus Phenotypic Plasticity. The American Naturalist 184 Suppl: S31–S46.

[15] Altizer S, Bartel R, Han Ba (2011) Animal Migration and Infectious Disease Risk. Science 331: 296–302.

[16] Ezenwa VO, Jolles AE (2011) From Host Immunity to Pathogen Invasion: The Effects of Helminth Coinfection on the Dynamics of Microparasites. Integrative and Comparative Biology 51: 540–551.

[17] Pedersen AB, Fenton A (2006) Emphasizing the Ecology in Parasite Community Ecology. Trends in Ecology and Evolution 22: 133–139.

[18] Bartlett MS (1957) Measles Periodicity and Community Size. Journal of the Royal Statistical Society Series A 120: 48–70.

[19] London WP, Yorke JA (1973) Recurrent Outbreaks of Measles, Chickenpox and Mumps: I. Seasonal Variation in Contact Rates. American Journal of Epidemiology 98: 453–468.

[20] Grenfell BT, Kleczkowski A, Ellner S, Bolker BM (1994) Measles as a Case Study in Nonlinear Forecasting and Chaos. Philosophical Transactions of the Royal Society of London A 348: 515–530.

[21] Earn DJ, Rohani P, Bolker BM, Grenfell BT (2000) A Simple Model for Complex Dynamical Transitions in Epidemics. Science 287: 667–70.

[22] Grenfell BT, Bjørnstad ON, Kappey J (2001) Travelling Waves and Spatial Hierarchies in Measles Epidemics. Nature 414: 716–23.

[23] Hasselquist D, Nilsson JA (2009) Maternal Transfer of Antibodies in Vertebrates: Trans-Generational Effects on Offspring Immunity. Philosophical Transactions of the Royal Society of London Series B, Biological Sciences 364: 51–60.

[24] Hassiotou F, Hepworth AR, Metzger P, Tat Lai C, Trengove N, et al. (2013) Maternal and Infant Infections Stimulate a Rapid Leukocyte Response in Breastmilk. Clinical & Translational Immunology 2: e3.

[25] National Aeronautics and Space Administration (2015). NASA: International Space Station: Circadian Rhythms (Circadian Rhythms) - 06.24.15. URL `http://www.nasa.gov/mission_pages/station/research/experiments/892.html`.

[26] Hawking F (1975) Circadian and Other Rhythms of Parasites. Advances in Parasitology 13: 123–182.

[27] Armstrong C (1950) Seasonal Distribution of Poliomyelitis. American Journal of Public Health 40: 1296–1304.

[28] Armstrong C (1951) Atmospheric Conditions and the Spread of Poliomyelitis. American Journal of Public Health 41: 1231–1237.

[29] Nathanson N, Martin J (1979) The Epidemiology of Poliomyelitis: Enigmas Surrounding its Appearance, Epidemicuty, and Disapperance. American Journal of Epidemiology 110: 672–692.

[30] Hemmes JH, Winkler KC, Kool SM (1962) Virus Survival as a Seasonal Factor in Influenza and Poliomyelitis. Antonie van Leeuwenhoek 28: 221–233.

[31] Dopico XC, Evangelou M, Ferreira RC, Guo H, Pekalski ML, et al. (2015) Widespread Seasonal Gene Expression Reveals Annual Differences in Human Immunity and Physiology. Nature Communications 6: 1–13.

[32] Melnick J, Dalldorf G, Enders J, Hammon W, Sabin A, et al. (1957) The Enteroviruses: Committee on the Enteroviruses, National Foundation for Infantile Paralysis. American Journal of Public Health 47: 1556–1566.

[33] Gear J (1949) The Virus of Poliomyelitis: Its Distribution and Methods of Spread. The Journal of the Royal Sanitary Institute 69: 149–153.

APPENDICES

APPENDIX A

Human Birth Seasonality: Latitudinal Gradient and Interplay with Childhood Disease Dynamics

A.1 Human Birth Seasonality: Materials & Methods

A.1.1 Birth Timing and Amplitude

In our analyses we followed the work of Rosenberg, who stated that adjusting for the differing number of days in each month had little effect on analyses of birth seasonality [1] . Thus, we did not make any adjustments of our time series to account for the different number of days in each month.

In the wavelet spectral analysis we tested for birth periodicity with periods ranging from 2 months to one-third the length of each data series. Since a significant 1 year period was observed, we constructed monthly phase angle time series for each data series using an 11-13 month period. The phase angle time series were subsequently used to determine the timing of the annual birth peak for each location. Peak-birth months were then averaged for each individual data series, with the U.S. states mapped to visualize the geographical variation in the timing of the annual birth

peak. When biannual peak-births occurred in the U.S., they were separated into two 6 month periods: summer (May-Oct) and winter (Nov-April) (Supp. Figs. A.1 & A.2).

The analysis of seasonal birth amplitude, or percent deviation from the annual mean, was done using seasonally decomposed time series. The stl function in the stats package in R was used to decompose the data into seasonal ($\mathcal{S}$), trend ($\mathcal{T}$), and noise ($\mathcal{N}$) components for each data series. The noise free time series were constructed as:

$$\mathcal{F} = \mathcal{S} + \mathcal{T} \tag{A.1}$$

The deviation from the mean during the birth peak was calcualted for each year, i, as:

$$x = \max(\mathcal{F}_i) - \text{mean}(\mathcal{T}_i) \tag{A.2}$$

The deviation from the mean during the birth trough was calculated for each year, i, as:

$$y = \min(\mathcal{F}_i) - \text{mean}(\mathcal{T}_i) \tag{A.3}$$

Thus, the one-half peak-trough difference is:

$$z = \frac{x - y}{2} \tag{A.4}$$

The seasonal amplitude, measured as a percent deviation from the mean, was calculated as:

$$\text{amplitude} = \frac{z}{\text{mean}(\mathcal{T}_i)} \tag{A.5}$$

A.1.2 Simulation study

For the simulation study we used a daily discrete-time SEIR model of measles adopted from Earn et al. 2000 [2]. The model has a daily time step and uses school-term forcing of seasonal transmission based on the school terms of England & Wales. The models assume transition probabilities follow a Poisson process. The difference equations are as follows:

$$S_{t+1} = \mu_t N_t + S_t e^{-(\beta_t I_t + \delta)} \tag{A.6}$$

$$E_{t+1} = S_t \left(1 - e^{-(\beta_t I_t + \delta)}\right) \frac{\beta_t I_t}{\beta_t I_t + \delta} + E_t e^{-(\phi+\delta)} \tag{A.7}$$

$$I_{t+1} = E_t \left(1 - e^{-(\phi+\delta)}\right) \frac{\phi}{\phi + \delta} + I_t e^{-(\gamma+\delta)} \tag{A.8}$$

$$R_{t+1} = I_t \left(1 - e^{-(\gamma+\delta)}\right) \frac{\gamma}{\gamma + \delta} + R_t e^{-\delta} \tag{A.9}$$

$$N_t = S_t + E_t + I_t + R_t \tag{A.10}$$

$$Incidence_t = E_t \left(1 - e^{-(\phi+\delta)}\right) \frac{\phi}{\phi + \delta} \tag{A.11}$$

$$\mu_t = \frac{\nu + A\sin(\omega t + \sigma)}{30} \tag{A.12}$$

$$\beta_t = \frac{B_0}{\frac{1}{365}\left((1 + b_1)273 + (1 - b_1)92\right)} \left(1 + b_1 Term_t\right) \tag{A.13}$$

$Term_t$ is based off the the school term schedule. When school is in session $Term_t = 1$ and when students are on holiday $Term_t = -1$. See Table A.1 for the school term schedule. The parameter values used for the simulation study can be found in Table A.2. All simulations were run for 100 - 150 yrs. to ensure that the trajectories were past the transient phase.

Holiday	*Model Days*	*Calendar Days*
Christmas	356 - 6	Dec 21 - Jan 6
Easter	100 - 115	Apr 10 - 25
Summer	200 - 251	Jul 19 - Sept 8
Autumn Half Term	300 - 307	Oct 27 - Nov 3

Table A.1: School term schedule. When students are on holiday $Term_t = -1$ otherwise 1.

Parameter	*Value*	*Parameter*	*Value*
$\mathcal{R}_0$	16 (basic reproductive no.)	b_1	0.25
D	1 - 365 (birth peak day)	μ_0	$\frac{1}{18250}$ day^{-1}
δ	$\frac{1}{18250}$ day^{-1} (50 yr life span)	β_0	$\frac{\mathcal{R}_0}{5}$ day^{-1}
ϕ	$\frac{1}{8}$ day^{-1} (8 day latent period)	S_0	0.06
γ	$\frac{1}{5}$ day^{-1} (5 day infectious period)	E_0	0.001
ν	$\frac{30}{18250}$ month^{-1} (balances δ)	I_0	0.001
A	0 - 0.0009208 month^{-1} (0 - 56% birth amp.)	R_0	0.938
ω	$\frac{2\pi}{365}$ $\frac{\text{radians}}{\text{day}}$	N_0	$S_0 + E_0 + I_0 + R_0$
σ	$\frac{\pi}{2} - \frac{2\pi}{365}D$	$Incidence_0$	0
B_0	$\frac{\mathcal{R}_0}{5}$		

Table A.2: Parameters used in simulation study, main text Figures 2.4A & 2.4B.

A.1.3 Inference study using simulated data

For the inference study we coupled our SEIR model (Eqs. A.6 - A.13) with a stochastic measurement model. The measurement model is as follows:

$$\text{cases}_t \sim \text{normal}(\rho Incidence_t, \rho\tau) \tag{A.14}$$

In order to test whether the seasonality in births influences parameter estimation, we simulated case data using three parameterizations of our model, each differing in the timing of the birth peak. The parameters used to generate the data are given in Table A.3. Simulations were run to year 50 to ensure the transient phase had passed.

Parameter	*Value*	*Parameter*	*Value*
$\mathcal{R}_0$	$\frac{6250}{365}$ (basic reproductive no.)	b_1	0.25
D	162, 295, or 351 (birth peak day)	μ_0	$\frac{1}{18250}$ day^{-1}
δ	$\frac{1}{18250}$ day^{-1} (50 yr life span)	β_0	$\frac{\mathcal{R}_0}{5}$ day^{-1}
ϕ	$\frac{1}{8}$ day^{-1} (8 day latent period)	S_0	0.06
γ	$\frac{1}{5}$ day^{-1} (5 day infectious period)	E_0	0.001
ν	$\frac{30}{18250}$ month^{-1} (balances δ)	I_0	0.001
A	0.000456 month^{-1} ($\sim$ 28% birth amp.)	R_0	0.938
ω	$\frac{2\pi}{365}$ $\frac{\text{radians}}{\text{day}}$	N_0	1
σ	$\frac{\pi}{2} - \frac{2\pi}{365}D$	$Incidence_0$	0
B_0	$\frac{\mathcal{R}_0}{5}$		

Table A.3: Parameters used to generate data for study, main text Figure 2.4C.

The three time series generated using the stochastic SEIR model were then fit to the SEIR model with a 0% birth amplitude, i.e. $A = 0$. The mean transmission rate, B_0, was the only free parameter. All other parameters, aside from A and B_0, were fixed at the values used to generate the data. Only the last 6 years of the time series were used for fitting, thus, the initial conditions were set to match those at the beginning of data used for inference.

For each time series B_0 was profiled and the likelihoods of the parameter sets with varying values of B_0 were calculated using a particle filtering in the R package pomp [3]. A particle filter (a.k.a. Sequential Monte Carlo) is a method of integrating state variables of a stochastic system and estimating the likelihood of the model for a fixed parameter set, given the data. However, since our model lacked process noise, we were able to obtain the exact likelihood for each parameter set.

A.1.4 Inference study using New York City measles data

For the New York inference study we utilized a Partially Observed Markov Process (POMP) model which are suited for dealing with epidemiological data where the state variables (susceptible, infected, recovered individuals) are not observed in the data, rather the infected individuals are partially observed through case reports [3]. For our process model we used a stochastic biweekly discrete-time SIR model. Similar to the model used for the simulation study, transitions were modeled using a Poisson process. The process model is as follows:

$$\lambda_t = \left(\beta_t \frac{I_t}{N_t} + \psi\right) \epsilon_t \tag{A.15}$$

$$\varrho_t = e^{-dt(\lambda_t + \delta)} \tag{A.16}$$

$$S_{t+1} = dtB_t + \varrho_t S_t \tag{A.17}$$

$$I_{t+1} = (1 - \varrho_t)\, S_t \frac{\lambda_t}{\lambda_t + \delta} \tag{A.18}$$

The transmission rate β_t was modeled using a periodic B-spline with 6 bases, a degree of 2, and a period of 1 year. The process noise ϵ_t was modeled as $\epsilon_t \sim \text{normal}(1, \beta_{sd})$. The covariates B_t, monthly number of individuals entering the susceptible class, and N_t, population size, were taken from data. All parameters were estimated using iterated particle filtering [3] (discussed later), with the exception of the death rate which was fixed at $\delta = \frac{1}{600}\ month^{-1}$, i.e. 50 yr. life span, and $dt = \frac{1}{2}$ fixing the time step to biweekly. In order to couple our model with measles case data we overlaid the process model with a stochastic measurement model. The measurement model is as follows:

$$cases_t \sim \text{normal}(\rho I_t, \tau I_t) \tag{A.19}$$

The case data, gathered from [2], consisted of monthly measles cases for New York City from January 1949 - December 1962. Although we did not have birth data for New York City we did have per capita monthly births for the state of New York. Thus, we assumed the per capita monthly birth rate for New York City was equal to the per capita monthly birth rate for New York state. The population size of New York City was taken from the decadal census and the population size was interpolated for non-census years. Taking the New York City population size together with the time series of per capita monthly births we constructed a time series of the number of monthly births, $\mathcal{B}_t$, in New York City (not to be confused with B_t in Eqn. A.17, which will be explained in the next section).

In order to test whether the seasonality in births influences model parameterization, we used four variants of our model (Eqn. A.15-A.18), each differing in the susceptible recruitment covariate, B_t. The first three models contain birth seasonality and account for the existence of maternal antibodies for 3-9 months. Whereas, in the fourth model we removed the birth seasonality. The first model variant lags the births by 3 months to account for a scenario where maternal antibodies confer protection from measles for the first 3 months of life:

$$B_t = \mathcal{B}_{t-3} \tag{A.20}$$

In the second model variant we lag births by 6 months.

$$B_t = \mathcal{B}_{t-6} \tag{A.21}$$

In the third model variant we lag births by 9 months.

$$B_t = \mathcal{B}_{t-9} \tag{A.22}$$

In the fourth model variant we removed the seasonality of births by making the monthly births constant within each year by setting them equal to the mean monthly births for the year:

$$B_t = \frac{\sum_{i=1}^{12} \mathcal{B}_{i,j}}{12}; j \in [1948 : 1962], \tag{A.23}$$

where i is the month and j is the year.

Each of the four model variants were independently fit to the data using Maximization by Iterated particle Filtering (MIF) using the R package POMP. MIF is a state-of-the-art simulation based method for parameter estimation that uses likelihood as the objective function. The basis of MIF is particle filtering (a.k.a. Sequential Monte Carlo), which is a method of integrating state variables of a stochastic system and estimating the likelihood of the model for a fixed parameter set, given the data. Unlike particle filtering, which uses fixed parameter values, MIF varies parameter values throughout the filtering process and selectively propagates particles (in the simplest sense, parameter sets) that have the highest likelihood. Thus, by initializing MIF throughout parameter space one can get a picture of the likelihood surface and identify the maximum likelihood parameter combinations within that space. For each of our models, MIF was initialized with 80000 parameter sets generated using a Sobol design, which pseudo-randomly samples parameters across parameter space in order to evenly sample the space. After this initial phase of MIF, parameter sets were passed through 15 successive stages of MIF, which included profiling. In total, for each model MIF was initialized at over 424000+ locations in parameter space to estimate the shape of the likelihood surface and identify the maximum likelihood parameter set(s). Table A.4 provides the maximum likelihood estimate (MLE) parameter set for each model.

A.2 Human Birth Seasonality: Results.

A.2.1 Detailed Results

Biannual birth pulses. During the pre-baby boom and baby boom eras, we found that some states had two birth pulses per year, i.e. they had a significant biannual period. All states significant for the biannual period were clustered together in the lower-midwest, deep south, and southeast (Figs. A.1 & A.2). In the baby boom era,

some of the states lost their significant biannual period and transitioned to having only a single seasonal birth pulse (Figs. A.1, & A.2). In the modern era, Arkansas remained the only state with a biannual period. The clustering of the states with a significant 6 month period in the southeastern U.S. may have been due to now defunct cultural factors (Figs. A.1 & A.2).

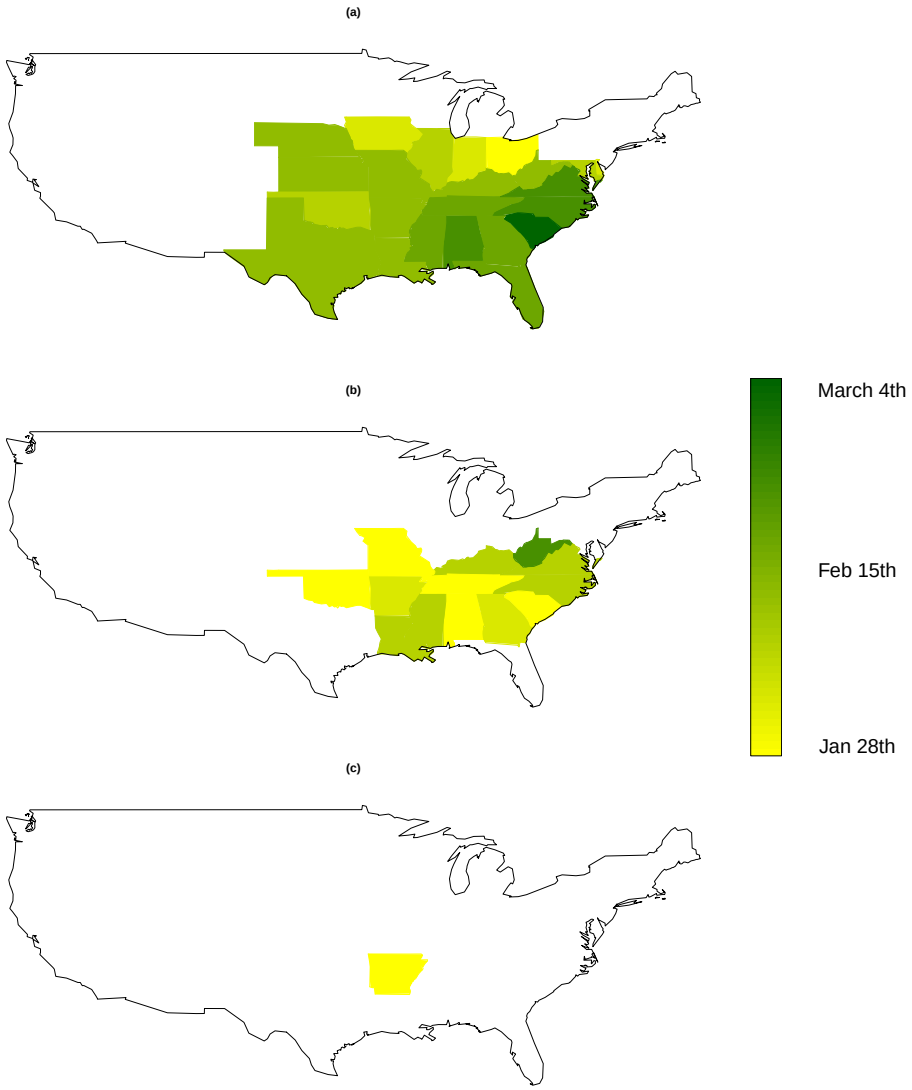


Figure A.1: Bi-annual winter (November-April) birth peak timing by period. (a) pre-baby boom (1931-1945), (b) baby boom (1946-1965), and (c) present period (1965-2008). States shown in white were not significant.

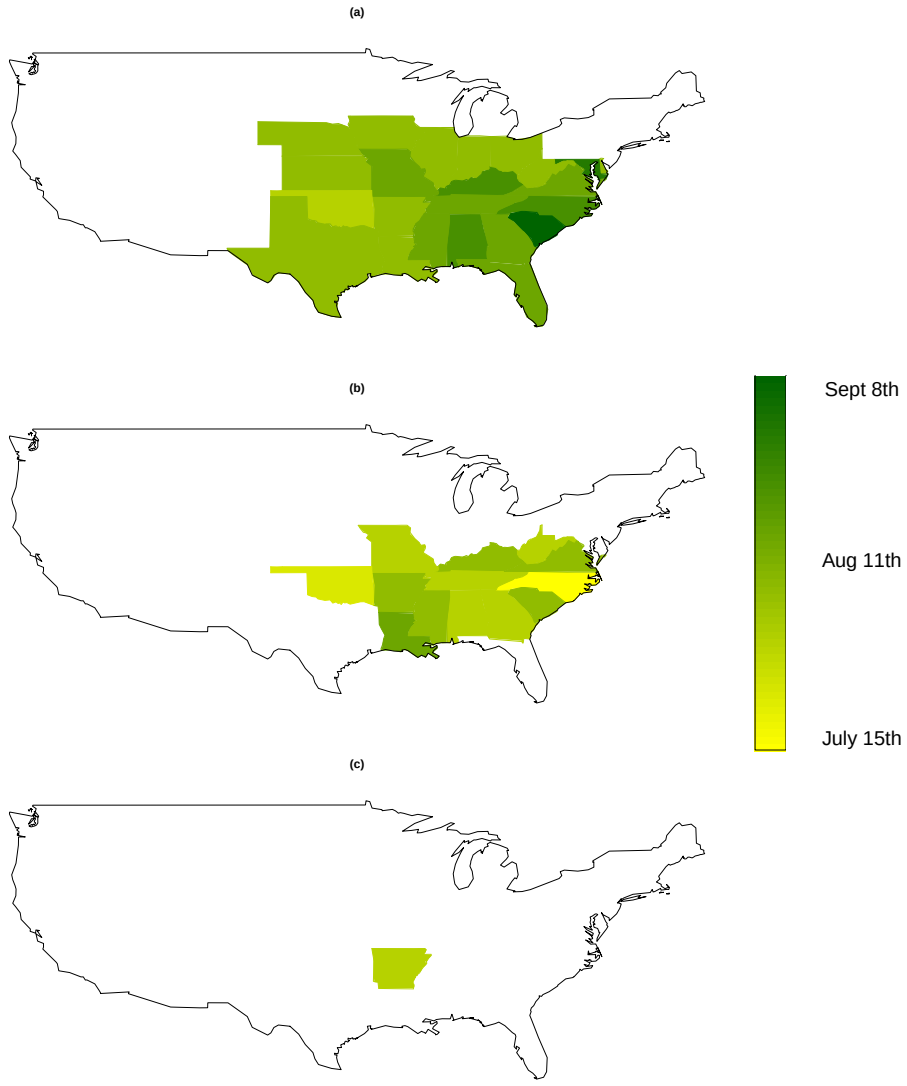


Figure A.2: Bi-annual summer (May-October) birth peak timing by period. (a) pre-baby boom (1931-1945), (b) baby boom (1946-1965), and (c) present period (1965-2008). States shown in white were not significant.

Birth rates. Raw birth rates in the pre-baby boom era ranged from 0.89/1000/month in Nevada (February, 1936) to 2.80/1000/month in New Mexico (May, 1932), with the mean and median both approximately 1.60/1000/month; while in the baby boom era Maryland had the lowest birth rate at 1.13/1000/month (April, 1950), New Mexico with the highest birth rates at 3.36/1000/month (October, 1946), and the mean and median both approximately 2.04/1000/month. In the present period Vermont

had the lowest birth rate at 0.67/1000/month (July, 2005), and Utah had the highest at 2.61/1000/month (July, 1977) with the mean and median falling in around 1.27/1000/month. Worldwide birth rates were not calculated, because we did not have population size data for our 200+ countries, rather raw birth values per month were used for wavelet spectral analysis. See Figure A.3 for maps of the mean birth rates in each state and each era.

The seasonal birth pulse. Examining the phase angle time series at a period ranging from 11-13 months, the U.S. data had twenty four states significant in the pre-baby boom era, whereas all were significant in the baby boom and present eras. Of the 210 worldwide data series analyzed, 132 (63%) were significant at an 11-13 month period. Many of those found insignificant were shorter time series (5-7 years) or countries with extremely low birth values (<100 individuals/month). Those states and countries found to be significant were then analyzed for the timing of the peak birth month using a wavelet spectral analysis.

U.S. timing of the seasonal birth peak. During the pre-baby boom era, of the states with a significant 1 year period, Oregon (June 12th) and Maine (June 10th) had the earliest peak birth timing —excluding New Mexico, which is an outlier because it has an early peak in all eras yet is located in the southern U.S. Florida (Nov 10th) had the latest peak birth timing in early November. The median peak birth timing was July 3rd, and the mean was July 26th. The range of peak birth timing was at a maximum in the pre-baby boom era with a length of 156 days, approximately 5 months. In this era there was a subtle pattern in the birth peak timing, with the northeast and northwest having earlier peaks than that of the deep south, the southeast and California.

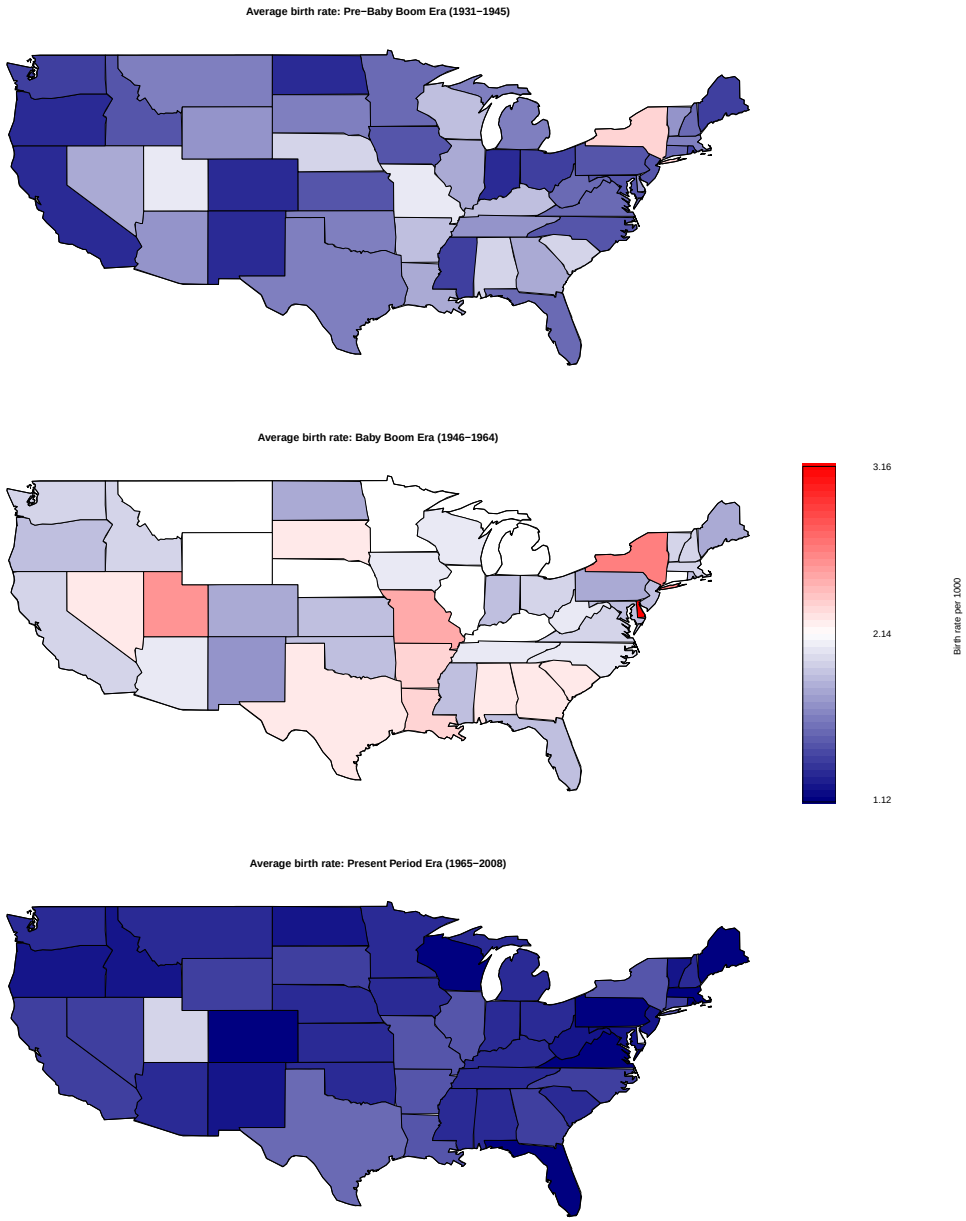


Figure A.3: Maps of mean birth rates for each state, in each era. Top to bottom: pre-baby boom, baby boom, and present era. No geographic pattern could be easily discerned.

In the baby boom era, when all states were significant at the 1 year period, there was a clear latitudinal gradient in birth amplitude and peak birth timing. Northern states saw an earlier peak birth timing; Utah beginning with a peak birth timing in

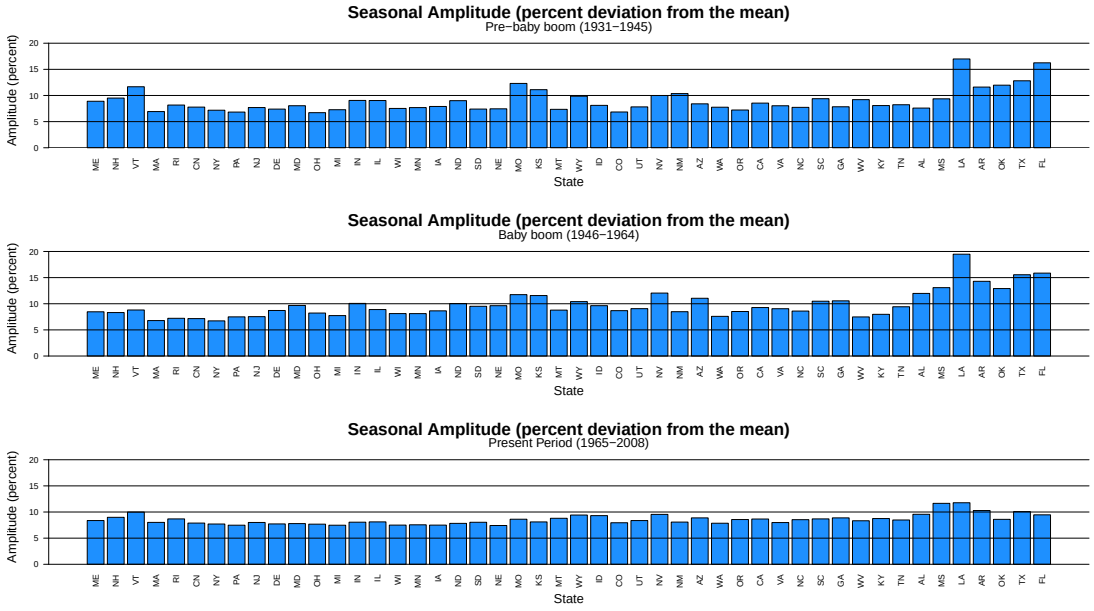


Figure A.4: Seasonal amplitude of births in U.S. states during each era. Note the high amplitude in the southern states (far right in all panels).

mid-July (July 13th), followed by Washington with a peak birth timing a few days later (July 18th) with no other peak birth timing occurring for at least another week after that. Again, Florida had the latest peak birth timing 3 months later in mid-October (Oct 21st). The mean (Sept 4th) and median (Sept 8th) both occurred in early September. In the baby boom era, other than Maine and those states already mentioned, the only states to have a peak prior to August 1st were also located in the northwestern U.S. (Idaho, Montana, North Dakota, and Oregon).

In the modern era, the spatial pattern of peak births was further elucidated with the mid-latitude states acting as a gradient for the northern and southern states. Utah again had the earliest peak birth timing in late June (June 26th), while Florida had the latest peak birth timing in early October (Oct 5th). The mean (Aug 11th) and median (Aug 7th) in this era both occurred in early August. Maine and Vermont also had peak birth timing prior to July 15th, which were the earliest peaks east of the Rocky Mountains. As with the baby boom era, many of the northwest states (Utah,

Idaho, Montana, Oregon, Washington, and Wyoming) had peaks prior to July 15th in the present era. Over all eras, Florida consistently had the latest peak ranging from early October in the present period, to early-November in the pre-baby boom era.

Worldwide timing of the seasonal birth peak. The worldwide birth peak timing followed a similar pattern as observed in the U.S., with countries at higher latitudes having an earlier birth peak than those closer to the equator. The earliest peak was in Italy during the period 1970-1985 (March 22nd) followed closely by Tajikstan in the period 1989-1994 (April 15th). The latest birth peak occurred in Saint Vincent and the Grenadines during the period 1992-2005 (November 17th). The mean worldwide peak (U.S. states not included) was mid-August (August 17th). The overall pattern is clear, with Europe (high latitude) having an earlier birth peak, and the Caribbean having a later peak. Both the Asian/Middle Eastern and Non-U.S. data are difficult to categorize as they span broad geographical ranges.

Seasonal birth amplitude. In the U.S., the largest seasonal amplitude in the pre-baby boom era was 20% from the mean (Louisiana, 1945), and the minimum was 4.2% from the mean in 1931 South Dakota (Fig. A.4 top panel). These values increased during the baby boom era with a maximum 21.3% variation from the mean (Louisiana, 1954), and a minimum of 5.5% (Connecticut, 1957) (Fig. A.4 middle). In the modern era, the maximum variation from the mean dropped to 17.4% (Louisiana, 1965) with the minimum staying relatively similar 5.4% (Delaware, 2004) (Fig. A.4 bottom). The mean percent deviation from the mean in all states during the pre-baby boom era was 9.0%, 9.8% during the baby boom, and 8.5% in the modern era.

As shown in the main text Figure 2.2, we observed a latitudinal gradient in the seasonal birth amplitude in each era in the U.S. However, the latitudinal gradient in

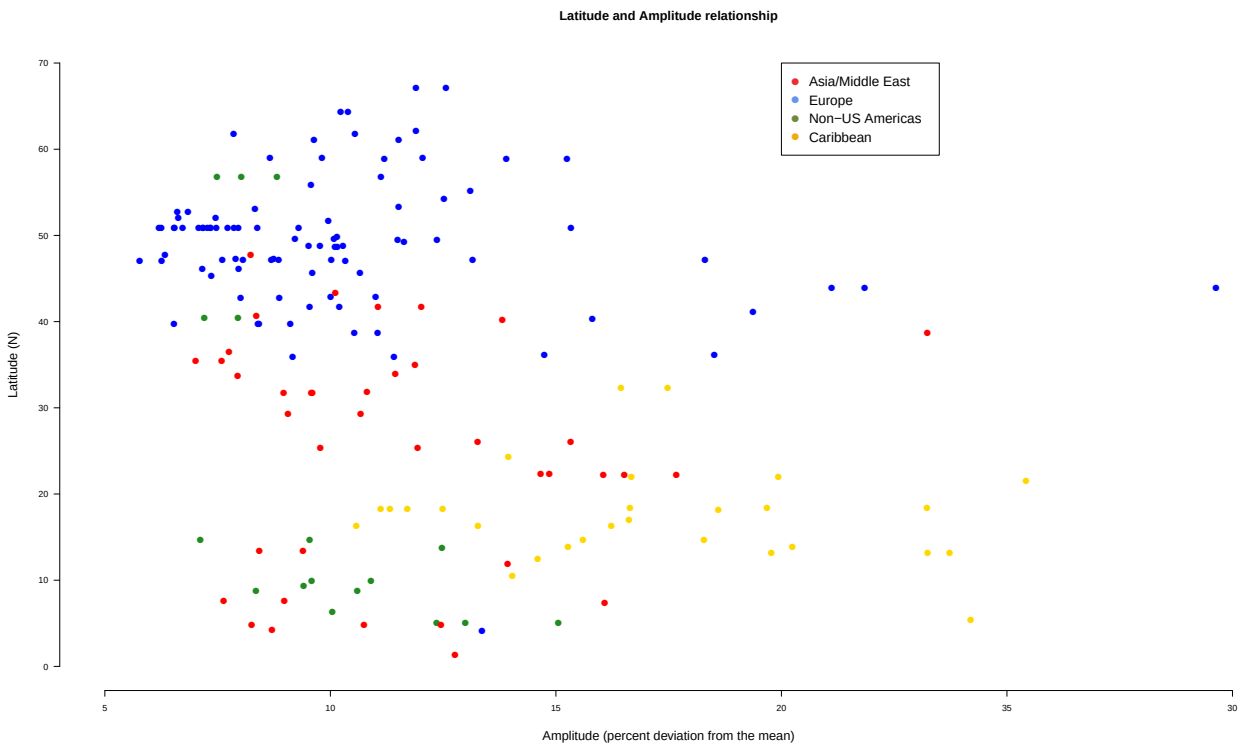


Figure A.5: Seasonal (intra-annual) amplitude of northern hemispheric data plotted vs latitude.

birth amplitude was not observed outside of the U.S. (Fig. A.5). European seasonal amplitudes tended to be low, with a mean of 10.3%. Non-U.S. Americas had an approx. 9.8% amplitude, Asian/Middle Eastern countries having approx. 12.6%, and Caribbean countries approx. 17.2% (Fig. A.5). However, due to the high variation in countries grouped into regions it is difficult to draw any conclusions from this.

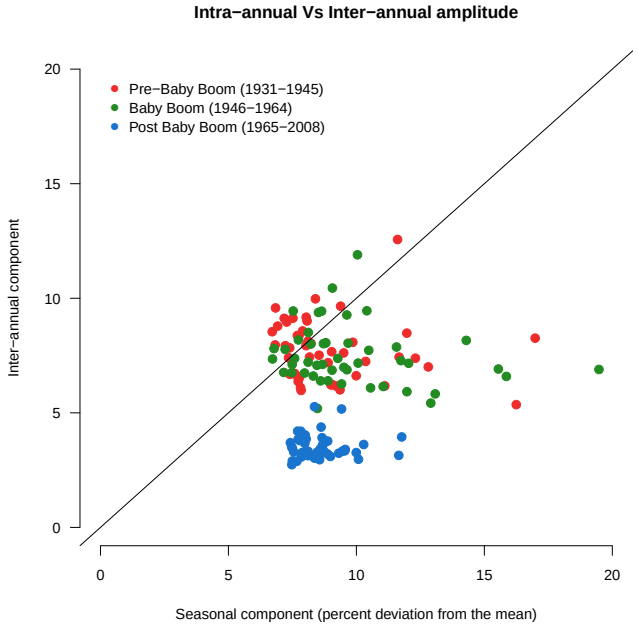


Figure A.6: The seasonal (intra-annual) amplitude vs the annual (inter-annual) amplitude. In both the pre-baby boom and baby boom most states had a stronger seasonal component, whereas in the present period all states have a stronger seasonal component.

Intra- vs. inter-annual variation in birth rates. We examined the seasonal (intra-annual) variation in births and compare that to the inter-annual variation (percent change from one year to the next) for each state in every era. We found that the seasonal (intra-annual) variation is generally larger (Fig. A.6), with the intra-annual variation in the modern era 2-3 times larger than the inter-annual value.

A.2.2 Simulation study

Fig. A.7 shows the effect of increasing birth amplitude on measles incidence with the birth peak occurring at various times of the year. The difference in incidence resulting from measles models with either constant births or seasonal births increases with birth amplitude. However, depending on the phase relationship between peak susceptible recruitment (i.e. the birth peak) and the peak in seasonal transmission, birth seasonality can have the effect of either enhancing or dampening the epidemic year incidence.

Fig. A.8 shows the effect of birth seasonality for various values of $\mathcal{R}_0$. Birth seasonality in this range of amplitude, $< 28\%$, has a pronounced effect on incidence when epidemics are biennial or triennial, as opposed to annual.

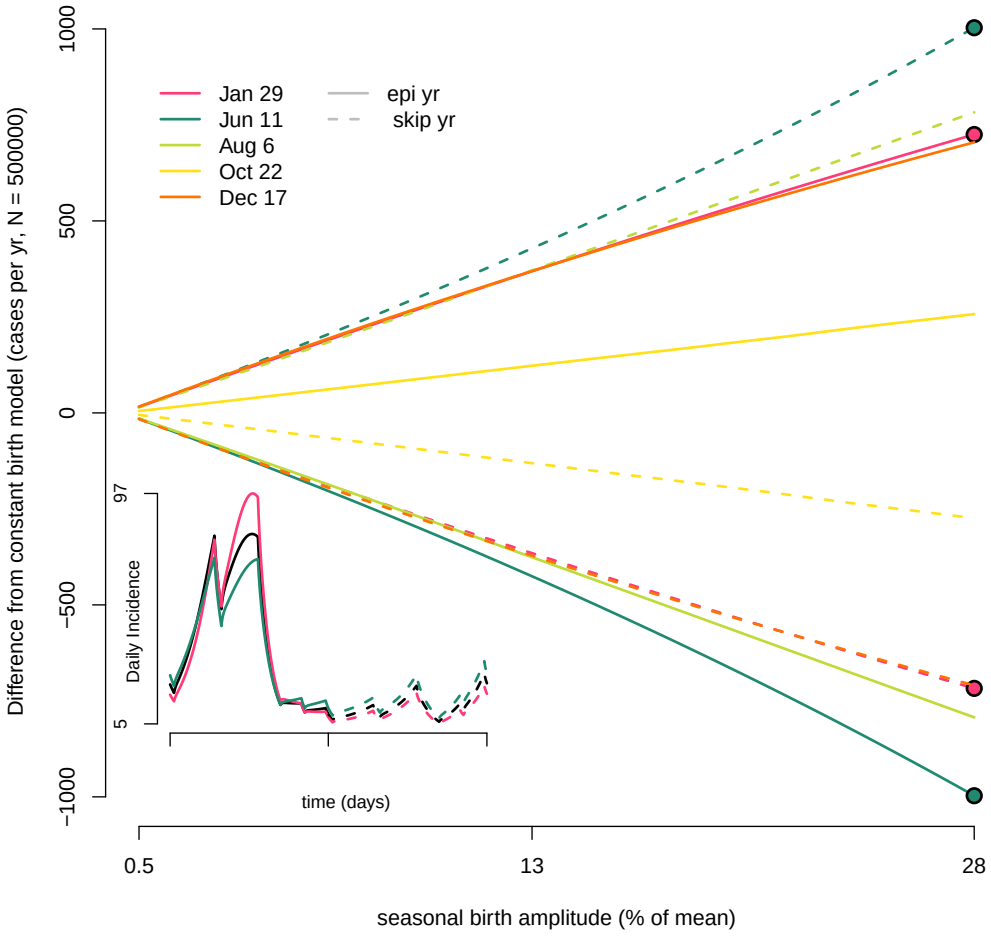


Figure A.7: Effect of increasing seasonal birth amplitude on measles incidence. The main graph show the change in epidemic and skip year incidence as a function of birth amplitude for 5 different phases of births seasonality. Phases were set such that the birth peak occurred in either Jan, Jun, Aug, Oct, or Dec. The turquoise and the fuschia points in the main graph correspond to the turquoise and fuchsia time series in the inset. Here $\mathcal{R}_0 = 17$ and the birth amplitude ranged from 0-28%, all parameters are those from Table A.2.

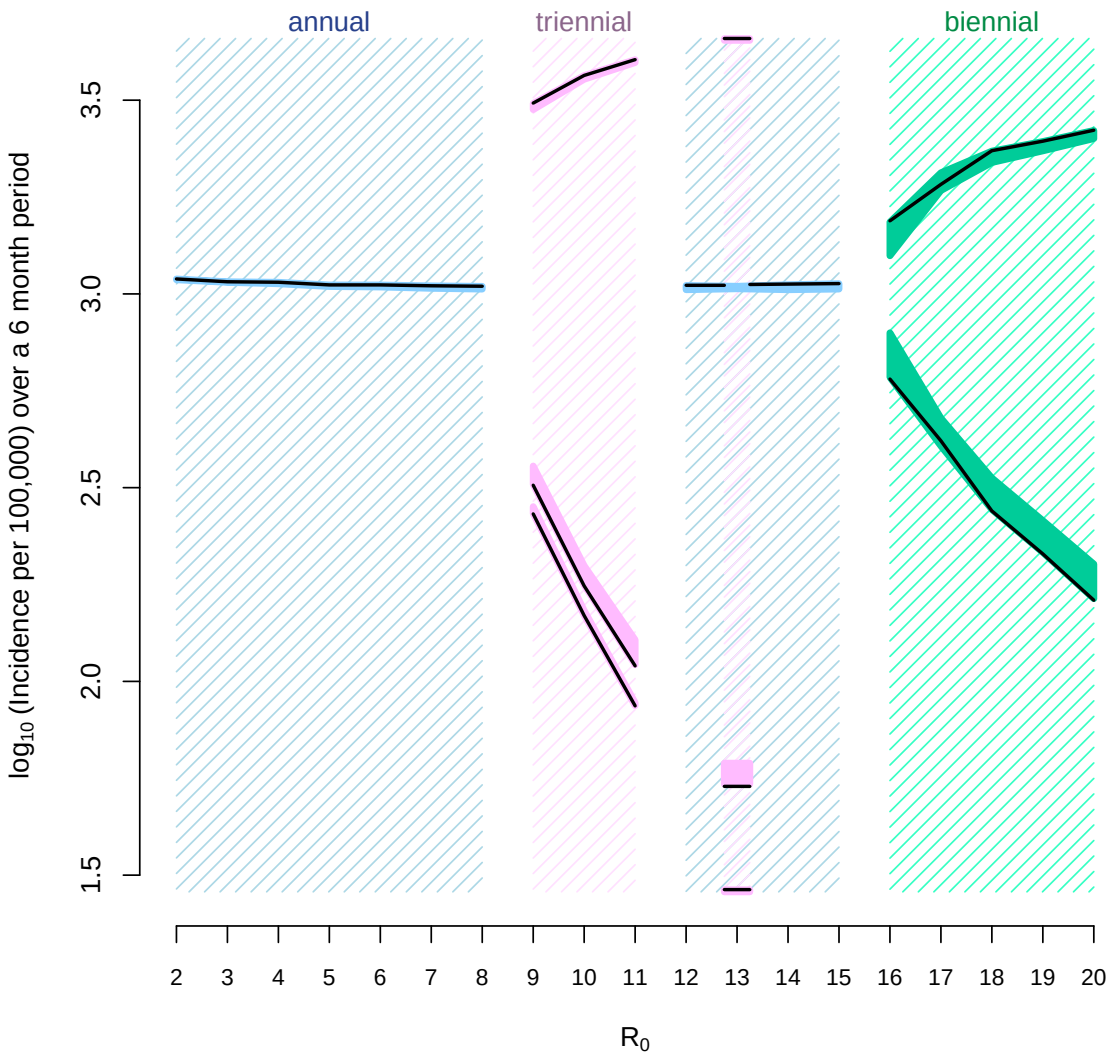


Figure A.8: The effect of birth seasonality on diseases with varying basic reproductive numbers ($\mathcal{R}_0$). This is a bifurcation diagram our SEIR model with varying $\mathcal{R}_0$ and varying seasonal birth amplitude. The black lines are the incidence for the model with no birth seasonality (i.e. birth amplitude = 0%). The solid shaded intervals indicate the regions containing the incidence of the model with seasonal births, where the birth peak is in early June and the amplitude ranges from 1.4 - 27.7%. Birth seasonality in this range of amplitude has a pronounced effect on incidence when epidemics are biennial or triennial, as opposed to annual. Here $\mathcal{R}_0 \in [2:20]$, and $S_0 = 1/\mathcal{R}_0$, otherwise all other parameters are those in Table A.2.

A.2.3 Inference study using New York City measles data

Maximum likelihood parameter estimates were obtained for each of our four measles models: seasonal births with a lag of 3 months from birth to susceptible to account for maternal antibodies, seasonal births with a lag of 6 months, seasonal births with a lag of 9 months, and constant births throughout the year.

Parameter	*Model Seas-3*	*Model Seas-6*	*Model Seas-9*	*Model No Seas*
LogLikelihood	-1080.99 ± 0.20	-1080.86 ± 0.12	-1081.14 ± 0.12	-1080.95 ± 0.14
$\mathcal{R}_0$	19.3	19.5	20.3	19.7
β_{coef_1}	55.4	55.6	58.9	56.5
β_{coef_2}	45.8	47.2	48.6	46.8
β_{coef_3}	45.6	45.5	48.5	47.8
β_{coef_4}	44.0	44.2	46.8	45.3
β_{coef_5}	17.5	18.5	18.7	18.3
β_{coef_6}	28.7	27.6	28.3	27.7
ψ	3.4×10^{-4}	3.8×10^{-4}	4.2×10^{-4}	4.1×10^{-4}
β_{sd}	0.121	0.118	0.121	0.124
dt	1/2	1/2	1/2	1/2
δ	1/600	1/600	1/600	1/600
ρ	0.238	0.236	0.239	0.238
τ	3.8×10^{-2}	3.8×10^{-2}	3.9×10^{-2}	3.8×10^{-2}
S_0	819214	947306	937014	1719583
I_0	16602	22758	17265	27629
R_0	6238289	6104041	6119826	5326893

Table A.4: Maximum likelihood parameter estimates for each model. Rates are given in units of $month^{-1}$. All parameters were estimated using MIF with the exception of δ and dt, which were fixed. Note, $\mathcal{R}_0$ is the basic reproductive number and R_0 is the initial number of recovered individuals.

Country	Years	No. Years	Latitude	Mean peak month	Significant	Group	Amplitude
Albania	1981-2007	27	41.17	6.15	Yes	Europe	19.37%
Algeria	1998-2002	5	33.10	8.80	No	Africa	10.11%
American Samoa	1984-1988	5	-14.30	5.20	Yes	Asia	14.77%
American Samoa	1996-2006	11	-14.30	4.45	Yes	Asia	14.81%
Antigua and Barbuda	1979-1986	8	17.05	10.50	Yes	Caribbean	16.62%
Armenia	1987-1999	13	40.29	7.31	No	Europe	15.81%
Aruba	2002-2007	6	12.52	9.17	Yes	Caribbean	14.60%
Australia	1973-2008	36	-32.35	6.56	No	Asia	6.06%
Austria	1973-2011	38	47.77	6.50	Yes	Europe	6.33%
Azerbaijan	1992-2004	13	40.18	2.31	No	Asia	13.81%
Bahamas	1972-1979	8	24.32	10.50	Yes	Caribbean	13.95%
Bahrain	1975-1985	11	26.03	6.09	No	Asia	15.33%
Bahrain	1986-2002	17	26.03	10.35	Yes	Asia	13.26%
Barbados	1969-1976	7	13.16	7.86	No	Caribbean	23.73%
Barbados	1982-1991	10	13.16	11.00	Yes	Caribbean	19.78%
Belarus	1987-1999	13	53.33	5.31	Yes	Europe	11.52%
Belgium	1971-1995	25	50.84	6.08	Yes	Europe	7.18%
Belgium	1998-2008	11	50.84	7.64	No	Europe	6.54%
Belgium-Bruxelles	1998-2008	11	50.84	8.09	No	Europe	7.35%
Belgium-Flamande	1998-2008	11	50.84	6.73	No	Europe	6.54%
Belgium-Anvers	1998-2008	11	50.84	6.64	No	Europe	7.33%
Belgium-Limbourg	1998-2008	11	50.84	6.45	No	Europe	7.96%
Belgium-Flandreorientale	1998-2008	11	50.84	7.18	No	Europe	6.25%
Belgium-Brabantflamand	1998-2008	11	50.84	6.91	Yes	Europe	7.47%
Belgium-Flandreoccidentale	1998-2008	11	50.84	5.64	No	Europe	7.17%
Belgium-Wallonne	1998-2008	11	50.84	8.55	Yes	Europe	6.73%
Belgium-Germanophone	1998-2008	11	50.84	5.18	No	Europe	15.34%
Belgium-Brabantwallon	1998-2008	11	50.84	6.91	Yes	Europe	7.86%
Belgium-Hainaut	1998-2008	11	50.84	9.09	Yes	Europe	6.20%
Belgium-Liege	1998-2008	11	50.84	8.27	No	Europe	7.08%
Belgium-Luxembourg	1998-2008	11	50.84	8.55	Yes	Europe	8.38%
Belgium-Namur	1998-2008	11	50.84	8.91	Yes	Europe	7.27%
Bermuda	1984-1991	8	32.30	9.38	Yes	Caribbean	17.48%
Bermuda	1995-2001	7	32.30	9.00	Yes	Caribbean	16.45%
BVI	1980-1986	7	18.43	10.43	Yes	Caribbean	23.23%
Brunei Darussalam	1972-1976	5	4.82	8.00	No	Asia	12.46%
Brunei Darussalam	1980-1992	13	4.82	9.31	No	Asia	10.75%
Brunei Darussalam	1996-2002	7	4.82	9.71	Yes	Asia	8.26%
Bulgaria	1973-1978	6	42.75	4.83	No	Europe	8.01%
Bulgaria	1980-1990	11	42.75	6.18	Yes	Europe	8.87%
Canada	1973-1990	18	56.76	6.06	Yes	Americas	7.49%
Canada	1992-1997	6	56.76	6.00	Yes	Americas	8.82%
Canada	1999-2008	10	56.76	6.80	Yes	Americas	8.03%
Cape Verde	1968-1975	8	15.11	3.63	No	Africa	12.94%
Cape Verde	1980-1985	6	15.11	11.50	No	Africa	17.40%
Caymen Islands	1986-1995	10	5.36	10.20	Yes	Caribbean	24.20%
Chile	1967-2008	42	-35.12	9.29	Yes	Americas	7.83%
China-Hong Kong	1973-1977	5	22.30	10.00	Yes	Asia	14.66%
China-Hong Kong	1979-2009	31	22.30	10.00	Yes	Asia	14.85%
China-Macao	1971-1975	5	22.17	9.60	Yes	Asia	16.52%
China-Macao	1984-1989	6	22.17	9.83	Yes	Asia	17.67%
China-Macao	1991-2010	20	22.17	9.70	Yes	Asia	16.06%
Cook Islands	1983-1988	6	-21.20	6.50	No	Asia	19.36%
Costa Rica	1987-1991	5	9.92	10.40	Yes	Americas	9.59%
Costa Rica	2003-2010	8	9.92	9.88	Yes	Americas	10.90%
Croatia	1988-2004	17	45.32	8.18	No	Europe	7.36%
Cuba	1976-1988	13	22.03	10.00	Yes	Caribbean	16.68%
Cuba	1990-2009	20	22.03	10.20	Yes	Caribbean	19.93%
Cyprus	1973-2009	37	34.97	8.49	Yes	Asia	11.88%
Czech Republic	1991-2010	14	49.82	5.36	Yes	Europe	10.15%
Denmark	1972-2005	34	55.85	6.03	Yes	Europe	9.57%
Egypt	1972-1982	11	28.80	7.00	Yes	Africa	30.08%
Egypt	1987-1999	13	28.80	10.85	No	Africa	35.30%
Egypt	2003-2009	7	28.80	9.29	No	Africa	12.60%
El Salvador	1973-2007	35	13.72	10.46	Yes	Americas	12.48%
Estonia	1989-1997	9	58.96	4.89	Yes	Europe	12.05%
Estonia	1999-2005	6	58.96	5.67	Yes	Europe	8.67%
Estonia	2007-2011	5	58.96	6.40	Yes	Europe	9.82%
Faeroe Islands	1972-1987	16	62.09	7.25	No	Europe	11.90%

Table A.5: Data used in national-level analyses of birth seasonality. Significance refers to the annual period. Mean birth peak timing and amplitude were estimated from the data.

Country	Years	No. Years	Latitude	Mean peak month	Significant	Group	Amplitude
Finland	1972-1988	17	61.76	4.94	Yes	Europe	10.55%
Finland	1994-2004	14	61.76	5.93	Yes	Europe	7.86%
France	1974-1989	16	47.14	5.75	Yes	Europe	10.02%
France	1991-1997	7	47.14	7.00	Yes	Europe	8.69%
France	1999-2004	6	47.14	7.83	Yes	Europe	8.06%
French Guiana	1977-1986	10	5.09	10.10	No	Americas	13.00%
French Guiana	1997-2003	7	5.09	10.71	No	Americas	12.35%
French Polynesia	1985-1992	8	-17.53	4.63	Yes	Asia	9.81%
Germany	1991-1997	7	50.86	7.14	Yes	Europe	7.72%
Germany	2004-2010	7	50.86	8.00	Yes	Europe	9.30%
Gibraltar	1973-1988	16	36.14	7.50	No	Europe	14.75%
Gibraltar	2002-2008	7	36.14	10.00	No	Europe	18.51%
Greece	1974-1985	12	38.69	6.00	Yes	Europe	10.53%
Greece	1990-2001	12	38.69	7.67	Yes	Europe	11.05%
Greenland	1972-1987	16	67.10	7.63	No	Europe	12.57%
Greenland	1992-2010	19	67.10	6.95	No	Europe	11.90%
Guadeloupe	1975-1980	6	16.27	9.33	No	Caribbean	10.58%
Guadeloupe	1982-1986	5	16.27	11.00	Yes	Caribbean	13.27%
Guadeloupe	1997-2003	7	16.27	10.57	Yes	Caribbean	16.23%
Guam	1973-1982	10	13.45	10.10	No	Asia	8.43%
Guam	1988-1992	5	13.45	9.80	No	Asia	9.40%
Guatemala	1972-1979	8	14.72	7.38	No	Americas	7.12%
Guatemala	1981-1999	19	14.72	10.79	No	Americas	9.54%
Guernsey	1973-1979	7	49.48	2.71	No	Europe	12.37%
Guernsey	1992-2000	9	49.48	8.22	No	Europe	11.49%
Guyana	1967-1971	5	6.35	9.60	Yes	Americas	10.04%
Hungary	1973-1992	20	47.29	6.40	Yes	Europe	8.74%
Hungary	1994-2004	11	47.29	7.73	No	Europe	7.90%
Iceland	1972-1980	9	64.37	5.67	Yes	Europe	10.39%
Iceland	1982-2004	22	64.37	6.77	Yes	Europe	10.23%
Iran	1999-2004	6	33.68	2.83	No	Asia	7.95%
Ireland	1972-2004	33	53.11	5.94	Yes	Europe	8.33%
Isle of Man	1973-1988	16	54.19	7.38	No	Europe	12.52%
Israel	1973-1981	9	31.78	9.33	Yes	Asia	9.58%
Israel	1983-1988	6	31.78	9.00	Yes	Asia	9.60%
Israel	1990-2009	20	31.78	9.60	Yes	Asia	8.97%
Italy	1970-1985	15	42.87	3.20	Yes	Europe	10.01%
Italy	1988-2009	22	42.87	8.14	Yes	Europe	11.01%
Jamaica	1999-2007	9	18.13	10.78	Yes	Caribbean	18.60%
Japan	1972-1992	21	35.41	7.48	Yes	Asia	7.59%
Japan	1994-2010	17	35.41	8.24	Yes	Asia	7.01%
Jersey	1973-1989	17	49.22	6.59	No	Europe	11.64%
Kazakhstan	1987-2008	22	43.35	6.27	No	Asia	10.11%
Korea Republic	1996-2009	24	36.47	4.83	No	Asia	7.75%
Kuwait	1975-1987	13	29.33	10.23	No	Asia	9.06%
Kuwait	1991-2008	18	29.33	9.44	No	Asia	10.68%
Kyrgyzstan	1985-2004	20	41.76	4.85	No	Asia	11.06%
Kyrgyzstan	2005-2009	5	41.76	7.40	Yes	Asia	12.02%
Latvia	1989-2005	17	56.83	5.12	Yes	Europe	11.12%
Lebanon	2003-2010	8	33.93	8.88	No	Asia	11.44%
Libya	1972-1981	10	29.96	1.00	Yes	Africa	12.01%
Libya	1989-1996	8	29.96	7.00	No	Africa	10.04%
Liechtenstein	1978-1987	10	47.15	6.80	No	Europe	13.16%
Liechtenstein	2000-2005	6	47.15	5.83	No	Europe	18.31%
Lithuania	1987-2011	25	55.22	5.72	Yes	Europe	13.11%
Luxembourg	1973-1989	17	49.64	6.47	No	Europe	9.22%
Luxembourg	1998-2010	13	49.64	6.23	Yes	Europe	10.08%
Malaysia	1994-2008	15	4.19	7.67	No	Asia	8.71%
Maldives	1978-2009	32	4.17	7.31	No	Europe	13.36%
Moldova Republic	1987-1992	6	47.17	5.67	No	Europe	7.60%
Moldova Republic	1998-2010	13	47.17	8.62	No	Europe	8.86%
Malta	1973-1988	16	35.90	9.06	Yes	Europe	11.41%
Malta	1992-2004	13	35.90	9.85	No	Europe	9.17%
Martinique	1975-1992	18	14.67	10.94	Yes	Caribbean	15.60%
Martinique	1998-2003	6	14.67	11.00	Yes	Caribbean	18.29%
Mauritius	1994-2010	17	-20.16	5.76	Yes	Africa	11.34%
Mongolia	1994-2003	10	47.77	5.10	No	Asia	8.24%
Netherlands	1973-1988	16	52.07	6.06	Yes	Europe	7.46%
Netherlands	1990-2010	21	52.07	7.62	Yes	Europe	6.63%
New Caledonia	1970-1977	8	-21.50	5.63	Yes	Asia	11.36%
New Caledonia	1982-2007	26	-21.50	5.54	Yes	Asia	12.12%
New Zealand	1972-2009	39	-41.44	7.46	Yes	Asia	23.52%
Niue	1982-1987	6	-19.06	7.83	No	Asia	45.89%
Norway	1976-1987	12	61.13	5.00	Yes	Europe	11.51%
Norway	1995-2004	10	61.13	5.50	Yes	Europe	9.64%
Occ. Palestinian Territory	1997-2007	11	31.88	10.64	Yes	Asia	10.81%

Country	Years	No. Years	Latitude	Mean birth month	Significant	Group	Amplitude
Palau	1997-2003	7	7.35	7.29	No	Asia	16.08%
Panama	1973-1999	27	8.75	10.78	Yes	Americas	8.35%
Panama	2005-2009	5	8.75	10.00	Yes	Americas	10.60%
Phillipines	1997-2007	11	11.87	10.00	Yes	Asia	13.93%
Poland	1978-2005	28	51.71	5.43	Yes	Europe	9.95%
Portugal	1973-1993	21	39.75	6.67	No	Europe	8.42%
Portugal	1999-2009	11	39.75	8.64	Yes	Europe	9.11%
Puerto Rico	1967-1985	19	18.26	10.05	Yes	Caribbean	11.33%
Puerto Rico	1987-1992	6	18.26	10.00	Yes	Caribbean	12.49%
Puerto Rico	1996-2000	5	18.26	10.00	Yes	Caribbean	11.71%
Puerto Rico	2002-2008	7	18.26	10.00	Yes	Caribbean	11.12%
Qatar	1985-1990	6	25.30	9.50	No	Asia	11.93%
Qatar	1999-2009	11	25.30	10.18	No	Asia	9.78%
Reunion	1977-1986	10	-21.11	6.40	Yes	Africa	7.53%
Reunion	1998-2003	6	-21.11	4.33	No	Africa	6.77%
Romania	1986-1992	7	45.69	6.29	Yes	Europe	9.60%
Romania	1994-2010	16	45.69	6.25	Yes	Europe	10.66%
Saint Helena ex dep	1981-1986	6	-15.93	6.83	No	Africa	31.44%
Saint Lucia	1976-1986	11	13.90	10.82	Yes	Caribbean	15.27%
Saint Lucia	1994-2002	9	13.90	11.00	Yes	Caribbean	20.24%
Saint Vincent and the Grenadines	1992-2005	14	13.20	11.07	Yes	Caribbean	23.24%
San Marino	1973-1978	6	43.94	7.33	No	Europe	21.85%
San Marino	1984-1989	6	43.94	9.50	No	Europe	29.63%
San Marino	2000-2004	5	43.94	7.60	No	Europe	21.12%
Sao Tome and Principe	1967-1971	5	0.32	4.20	No	Africa	13.27%
Seychelles	1973-1993	21	-4.63	6.24	Yes	Africa	14.81%
Seychelles	1995-2010	16	-4.63	5.94	Yes	Africa	18.63%
Singapore	1973-2009	36	1.37	9.36	Yes	Asia	12.77%
Slovakia	1988-1995	8	48.66	5.75	Yes	Europe	10.10%
Slovakia	1998-2002	5	48.66	6.20	Yes	Europe	10.16%
Slovenia	1988-1996	9	46.17	6.67	No	Europe	7.17%
Slovenia	1998-2005	8	46.17	7.38	Yes	Europe	7.97%
Spain	1970-1985	16	39.72	6.38	Yes	Europe	8.40%
Spain	1991-2005	15	39.72	7.87	No	Europe	6.53%
Sri Lanka	1973-1986	14	7.57	9.14	No	Asia	7.64%
Sri Lanka	2005-2010	6	7.57	10.00	Yes	Asia	8.98%
Suriname	1989-2009	21	5.06	11.05	Yes	Americas	15.05%
Sweden	1973-1990	18	58.91	4.56	Yes	Europe	15.25%
Sweden	1992-2002	11	58.91	4.91	Yes	Europe	13.90%
Sweden	2004-2010	7	58.91	5.71	Yes	Europe	11.19%
Switzerland	1973-1982	10	47.03	5.00	Yes	Europe	10.33%
Switzerland	1984-1990	7	47.03	5.86	Yes	Europe	6.26%
Switzerland	1998-2002	5	47.03	6.80	No	Europe	5.78%
Tajikstan	1989-1994	6	38.70	4.00	Yes	Asia	23.24%
Macedonia	1989-1993	5	41.74	8.20	No	Europe	9.54%
Macedonia	1999-2010	12	41.74	7.83	Yes	Europe	10.20%
Tonga	1993-2000	8	-19.70	4.25	Yes	Asia	16.00%
Trinidad and Tobago	1972-1995	24	10.55	10.83	Yes	Caribbean	14.04%
Tunisia	1971-1976	6	35.42	2.83	Yes	Africa	16.04%
Tunisia	1978-1982	5	35.42	4.00	Yes	Africa	13.83%
Tunisia	1985-1995	11	35.42	5.18	No	Africa	13.40%
Tunisia	2001-2007	7	35.42	7.00	Yes	Africa	16.75%
Turks and Caicos Islands	1997-2005	9	21.51	9.56	No	Caribbean	25.43%
Ukraine	1980-1986	7	48.81	6.00	Yes	Europe	9.52%
Ukraine	1989-1996	8	48.81	5.75	Yes	Europe	9.77%
Ukraine	2003-2010	8	48.81	8.25	Yes	Europe	10.28%
UK + NI	1982-1988	7	52.75	6.71	Yes	Europe	6.85%
UK + NI	1990-2006	15	52.75	7.47	Yes	Europe	6.61%
USA	1969-1975	7	40.42	8.86	Yes	Americas	7.21%
USA	1978-2006	29	40.42	8.07	Yes	Americas	7.95%
USVI	1969-1973	5	18.33	10.80	Yes	Caribbean	19.68%
USVI	1980-1997	18	18.33	10.61	Yes	Caribbean	16.64%
Uruguay	1980-1988	9	-33.00	8.33	No	Americas	5.55%
Uzbekistan	1993-1997	5	40.68	5.80	No	Asia	8.36%
Venezuela	1972-2001	30	9.39	9.40	Yes	Americas	9.41%
Wallis and Futuna Islands	1973-1978	6	-13.30	4.00	No	Asia	10.72%

Bibliography

[1] Rosenberg, H. M., 1966 Seasonal Variation of Births United States, 1933-63. *National Center for Health Statistics* **21**, 1–59.

[2] Earn, D. J., Rohani, P., Bolker, B. M. & Grenfell, B. T., 2000 A Simple Model for Complex Dynamical Transitions in Epidemics. *Science* **287**, 667–670. ISSN 0036-8075.

[3] King, A. A., Ionides, E. L., Bretó, C. M., Ellner, S., Kendall, B., Wearing, H., Ferrari, M. J., Lavine, M. & Reuman, D. C., 2010 *POMP: Statistical Inference for Partially Observed Markov Processes (R package).*

APPENDIX B

Unraveling the Transmission Ecology of Polio

B.1 Polio Transmission Ecology: Materials & Methods

B.1.1 Data

Our data consisted of monthly cases reported in the weekly US Public Health Reports and the CDC Morbidity and Mortality Weekly Report from January 1931–December 1954 for each US state and the District of Columbia. The data were first published in [1], which includes discussion of data quality during this period. Data are included in CSV files as Supporting Information. Estimates of paralytic polio reporting during this time period are 60–80%, while the reporting of non-paralytic cases varied city-to-city. Monthly state–level time series of live births from 1931–1954 were downloaded from the Vital Statistics of the United States [2]. Annual state–level population size data were collected from the Population Distribution Branch of the US Census Bureau [3].

B.1.2 Data Analysis

An epidemic fadeout was defined as three consecutive months without reported cases. The portion of fadeout months was calculated for each state and each year as the number of months that were part of a fadeout, divided by the total months in the era. The relationship between the portion of fadeout months and state population size (taken as the mean population size for the era) was fit using a negative exponential to the fadeout data. A negative exponential was fit to each era independently using Nonlinear Least Squares, nls, in R. The following is the equation used to relate fadeouts to population size:

$$\text{portion of fadeout months} = be^{-aN}, \tag{B.1}$$

where N is $\log_{10}$(population size). The pairwise epidemic synchrony between states was measured using the spatial nonparametric covariance function in the R package ncf [4]. The distance between states was measured as the distance between the population center of each state. Population centers were gathered from the US Census Bureau from the 2000 US Census. Data from https://www.census.gov. Population centers were used from the year 2000 because, at the time of this project, mid-20th century estimates of population centers were not available. However, the US Census has now made historical population centers available.

In order to test for periodicity in epidemics, a wavelet spectral analysis was conducted for each state using the R package biwavelet [5]. Significance of the 1 year period was tested using the lag-1 autocorrelation, with state-specific lag-1 autocorrelation calculated for each time series. Since the 1 year period was significant for each state, phase angle time series were constructed for the 1 year period in order to measure the seasonal timing of epidemics. For each year, the states were ranked 1 – 49, with 1 being the state for which the epidemic peaked earliest in the year, and 49

being the state with the latest epidemic peak. The states were ranked according to their phase angle in the month of August. The phase angle is zero at the epidemic peak, negative before the peak, and positive after the peak. Thus, the state with the largest positive phase angle was ranked 1. The epidemic timing was regressed with latitude and mapped in order to test for spatial patterns.

B.1.3 Model

To mechanistically model polio epidemiology we utilized a Partially Observed Markov Process (POMP) model which is suited for dealing with epidemiological data where the state variables (susceptible, infected, recovered individuals) are not observed in the data; rather the infected individuals are partially observed through case reports. For our process model we used a seasonally-forced stochastic monthly discrete-time SIR model where transitions followed a Poisson process. The model has a single recovered class that accounts for life-long immunity. The model contains 6 classes (S_i^B) of infants susceptible to infection but protected from clinical illness by maternally antibodies. The 6 susceptible infant classes contain 0–1 month olds, 1–2 month olds, etc., up to 6 month olds. The model had a single infected class for infants (I^B). The older age class contains individuals > 6 months of age, and these individuals have their own susceptible (S^O) and infected class (I^O). The measurement model translated infections to clinical reported cases. We assume that infections in individuals under 6 months of age are asymptomatic, and only individuals over 6 month of age can be symptomatic and reported as a clinical case. Refer to the main text for model schematic, which was made with the R program diagrammeR (http://rich-iannone.github.io/DiagrammeR/) [6].

B.1.3.1 Process Model

The force of infection, also referred to as the risk of infection, was

$$\lambda_t = \left(\beta_t \frac{I_t^O + I_t^B}{N_t} + \psi \right) \epsilon_t, \tag{B.2}$$

with the following condition imposed,

$$\lambda_t = \begin{cases} \lambda_t, & \text{if } \lambda_t \geq 0 \\ 0, & \text{otherwise} \end{cases}. \tag{B.3}$$

The first term of the force of infection represents transmission that occurs locally by individuals infected in the state at time t. Whereas, the second term, ψ, encompasses WPV that arises in the population in a way that is divorced from the local infection dynamics. These infections can include immigration of infected individuals from other geographic locations, environmental sources of WPV, and individuals shedding with an infectious period longer than a month. The transmission parameter β_t was parameterized using a B-spline, giving it the flexibility to have either a constant transmission rate or a seasonal transmission rate,

$$\beta_t = \exp \sum_{i=1}^{6} \beta_i \xi_{i_t}, \tag{B.4}$$

where each ξ_{i_t} is a periodic B-spline basis with a 1 year period. The periodic.bspline.basis() function in the pomp package was used to construct the B-spline bases. The process noise, ϵ_t, was gamma distributed.

$$\epsilon_t \sim \Gamma \left(\frac{1}{\Theta}, \Theta \right) \tag{B.5}$$

where the scale parameter, Θ, of the gamma distribution accounts for both environmental and demographic stochasticity.

$$\Theta = \left(\frac{\beta_{sd_1}}{\sqrt[2]{\beta_t \frac{I_t^O + I_t^B}{N_t}} + \beta_{sd_2}} \right)^2 \tag{B.6}$$

The first term of Θ represents demographic stochasticity that is modulated by the force of infection, λ_t. Whereas, the second term represents environmental stochasticity. Formulating the process noise in this way gives us some useful properties,

$$\mathbb{E}\left[\epsilon_t\right] = 1 \tag{B.7}$$

$$\mathrm{Var}\left[\epsilon_t\right] = \Theta. \tag{B.8}$$

Thus, the variance can capture purely environmental stochasticity, under the following parameterization:

$$\mathrm{Var}\left[\epsilon_t\right] = \beta_{sd_2}^2, \text{ if } \beta_{sd_1} = 0 \tag{B.9}$$

and, it can represent purely demographic stochasticity under alternate parameterization:

$$\mathrm{Var}\left[\epsilon_t\right] = \frac{\beta_{sd_1}^2}{\beta_t \frac{I_t^O + I_t^B}{N_t} + \psi}, \text{ if } \beta_{sd_2} = 0 \tag{B.10}$$

Due to the computational intensity of this project, we used a discrete time model with a 1 month time step to speed-up simulation. We implemented the model as a poisson process. There is one probability governing the movement of individuals out of their susceptible class: p_{S_t}, which is the probability that a susceptible individual remains susceptible,

$$p_{S_t} = e^{-(\delta + \lambda_t)}. \tag{B.11}$$

One minus the probability of remaining susceptible is the probability of either being infected or dying. The parameter δ is the natural death rate. The superscripts ^{B} and ^{O} will indicate infants (i.e., babies) and older individuals, respectively. The equations for the first infant class is:

$$S^B_{1_{t+1}} = B_{t+1}, \tag{B.12}$$

where B_{t+1} are the births in month $t+1$. Similarly, the other five infant classes are tracked using the following equations:

$$S^B_{j_{t+1}} = S^B_{j-1_t} p_{S_t};\ \text{for } j \in 2:6. \tag{B.13}$$

The equation above describes that infants in the $j-1$ susceptible infant class at time t move into infant class j at time $t+1$ if they are not infected or die of natural death. The equation for the susceptible non-infant age group is:

$$S^O_{t+1} = \left(S^B_{6_t} + S^O_t\right) p_{S_t}. \tag{B.14}$$

The first term of the S^O equation represents the movement of infants from the oldest infant class S^B_6 to the susceptible non-infant class. Whereas, the second term represents susceptible non-infants from time t remaining susceptible at time $t+1$. The infected infant class is tracked using the following equation:

$$I^B_{t+1} = \sum_{j=1}^{6} S^B_{j_t}(1 - p_{S_t})\frac{\lambda_t}{\delta + \lambda_t}. \tag{B.15}$$

Infants from each of the six susceptible infant classes have a probability of being infected or dying, $1 - p_{S_t}$. The probability that an individual is infected, rather than death is $\frac{\lambda_t}{\delta+\lambda_t}$. The equation for infected non-infants is similar in structure,

$$I^O_{t+1} = S^O_t(1 - p_{S_t})\frac{\lambda_t}{\delta + \lambda_t}. \tag{B.16}$$

Note, infected individuals are infected for exactly 1 month, which is the typical duration of shedding [7]. There was one rounding condition imposed on infections in the process model. If I_t^B or I_t^O were values between 0 and 1, they were rounded to 0 or 1. This was in order to prevent fractions of infected individuals.

B.1.3.2 Measurement Model

In order to model the stochastic process of infected individuals becoming symptomatic and subsequently being reported as a case, we drew cases from a normal distribution.

$$\text{cases}_t \sim \text{normal}(\rho_t I_t^O, \tau I_t^O), \tag{B.17}$$

with the following condition imposed,

$$\text{cases}_t = \begin{cases} \text{round}(\text{cases}_t), & \text{if cases}_t \geq 0 \\ 0, & \text{otherwise} \end{cases} \tag{B.18}$$

For calculating likelihood we used a binned-normal probability density.

If $\text{cases}_t > 0$,

$$\text{Likelihood}_t = \text{pnormal}(\text{cases}_t{+}0.5, \rho_t I_t^O, \tau I_t^O) - \text{pnormal}(\text{cases}_t{-}0.5, \rho_t I_t^O, \tau I_t^O) + 1e^{-18}. \tag{B.19}$$

If $\text{cases}_t = 0$,

$$\text{Likelihood}_t = \text{pnormal}(\text{cases}_t + 0.5, \rho_t I_t^O, \tau I_t^O) + 1e^{-18}. \tag{B.20}$$

The report rate, ρ_t, was estimated for the Pre-Baby Boom era and the Baby Boom era, with the change in report rate occurring in January 1946, the start of the Baby

Boom. The parameter τ captures the variation in the process of observing infections via reported cases. The scaler $1e^{-18}$ was used for practical purposes to put a lower bound on the likelihood in order to ensure finite log-likelihood values. Initially we tried using a continuous normal measurement model, however, we discovered that a continuous normal model over-inflated the likelihood of parameter sets that produced zero cases in trough months, making it difficult to identify parameter regimes that could capture both the epidemic cases and the trough cases. Thus, we used the binned-normal distribution.

B.1.4 Statistical Inference

We fit our SIR model to the data from each state independently using Maximization by Iterated particle Filtering (MIF) in the R package pomp [8, 9]. pomp tutorials and vignettes may be found at http://pomp.r-forge.r-project.org/. MIF is a simulation-based likelihood method for parameter estimation. The basis of MIF is particle filtering, which integrates state variables of a stochastic system and estimates the likelihood for fixed parameters. Instead of fixing parameters, MIF varies them throughout the filtering process and selectively propagates particles (parameter sets) that have the highest likelihood. By initializing MIF throughout parameter space we estimated the shape of the likelihood surface for each US state and identified the Maximum Likelihood parameter Estimates (MLEs). MIF was initialized from 1 million parameter sets for a global search, followed by additional phases of increasingly localized searches, which included profiling. In total, for each US state, MIF was initialized from tens-of-thousands of locations in parameter space to estimate the shape of the likelihood surface and identify the MLEs. We characterized seasonal transmission using a B-spline with six bases. Although the transmission rate fluctuated seasonally, there was no inter–annual variation in transmission in our model. In order to account for improved reporting brought about by the Foundation for Infantile Paralysis, we

estimated two report rates, one for the Pre-Baby Boom Era and another for the Baby Boom Era. Inference was done using the data from May 1932 to January 1953, with the exception of South Dakota and Texas, whose inference started in May 1933 and 1934, respectively.

For each state inference was initiated in May of the year following the first full year of available data on polio cases and births. This was done in order to construct the infant initial conditions directly from the birth data. The month of May is the tail-end of the polio off-season. Thus, the 0-6 month old infants in May were born between November and April, when polio transmission was low and they were unlikely to be infected. Based on this, we assumed that the initial number of susceptible infants in the 6th infant class was the number of individuals born in November of the previous year, individuals born in December were in the 5th infant class, January births were in the the 4th infant class, February births in the 3rd infant class, March births in the 2nd infant class, and April births in the the 1st infant class. We set the initial number of infected infants to zero. Constructing the infant initial conditions prevented us from having to estimate an addition seven parameters.

B.1.4.1 Phase I: Global Search

To initialize a global search of parameter space, 1 million parameter combinations, with initial conditions reflecting a low incidence month, were generated to cover the range of parameter values given in Table B.1. For each parameter set, ten replicate particle filters were run, each with 2000 particles, to estimate the likelihood of the data from the state of Illinois. Illinois was chosen because of the high incidence of polio in Illinois and because Illinois had an increase in incidence following the World War II baby boom, which was a characteristic feature of most US states. The reporting transition was fixed to January 1946, the start of the Baby Boom era and about the time when non-paralytic polio cases started being reported. Each parameter set was

placed in combination with an increase in report rate of 0% and 1%. The 1 million likelihoods were used to provide a rough estimate of the global shape of the likelihood surface and identify regions of parameter space with high likelihood. Of the 1 million parameter sets, the 50000 with the highest likelihoods were used to initialize Phase II. It is important to note that the particle filter gives an unbiased estimate of the likelihood, but some extra algebra needs to be done to get the estimates of the log-likelihood from replicate particle filters. We calculated the unbiased estimate of the log-likelihood using the following equations,

$$\log \vec{\mathcal{L}}(\Theta) = \vec{\mathbf{x}}. \tag{B.21}$$

Given the vector of log-likelihoods, $\vec{\mathbf{x}}$, produced from n replicate particle filters under the parameter set Θ (not to be confused with Θ in earlier equations), we find the mean:

$$y = \text{mean}(\vec{\mathbf{x}}) \tag{B.22}$$

then we use the mean to estimate of the log-likelihood:

$$\log \mathcal{L} = y + \log(\text{mean}(e^{\vec{\mathbf{x}}-y})) \tag{B.23}$$

and we can also calculate the standard error of the estimate of the log-likelihood:

$$\log \mathcal{L}_{\text{se}} = \frac{\text{sd}(e^{\vec{\mathbf{x}}-y})}{\sqrt{n}} \frac{1}{e^{\log \mathcal{L}-y}} \tag{B.24}$$

where sd is the standard deviation.

B.1.4.2 Phase II: Initial Condition Improvement

Each of the 50000 parameter sets obtained from Phase I was taken in pairwise combination with the data from each of the 49 continental US states. For each

parameter	interpretation	lower bound	upper bound	search scale
ρ	report rate	0.0001074816	0.9999401	logit
$\Delta\rho$	increase in report rate	0	0.01	logit
τ	measurement stochasticity	1.024683e-06	2390.506	log
ψ	external source of infection	1.843172e-07	114.648	log
δ	natural death rate	0.001388889	NA	NA
β_{sd_1}	demographic stochasticity	1.835635e-07	21.59023	log
β_{sd_2}	environmental stochasticity	1.217956e-07	15.25643	log
β_i	B-spline basis coefficients	9.152956e-12	770910.8	log

Table B.1: Parameter Bounds used in Global Search. Values in the table are on the natural scale, rates are given in months. ρ is the report rate for the Pre-Baby Boom era, and $\Delta\rho$ is the increase in report rate during the Baby Boom. The natural death rate δ was fixed. The search scale indicates whether parameters were transformed into the log or logit scale for searching parameter space. The transformation ensured the breadth of the search. Note, that parameter searches were not restricted to these bounds, rather these bounds were used to initialize searches for parameter space.

parameter set and US state combination, maximization by iterated particle filtering (MIF) was used to improve the estimates of the initial conditions using the first 12 months of data. Each initial condition had a random walk standard deviation of 0.02. Each MIF had 45 iterations, 2000 particles, hyperbolic cooling with a variance factor of 2, and a cooling fraction of 0.5. This resulted in 50000 parameter sets for each state, each with unique initial conditions. For each parameter set, ten replicate particle filters were run, each with 2000 particles, to estimate the likelihood of the full time series (i.e., the fitted region of the data) from the state. All searches of parameter space were done on the log or logit scale. The log scale was used to enforce parameter values > 0 and the logit scale was used to bound parameters, such as the report rate, between 0 and 1. See Equations below for expit and logit functions.

$$\text{logit}(p) = \log\left(\frac{p}{1-p}\right) \tag{B.25}$$

$$\text{expit}(x) = \frac{1}{1+e^{-x}} \tag{B.26}$$

B.1.4.3 Phase III–V: Local Searches of Parameter Space

In Phase III, for each US state, the top 25000 likelihood parameter sets from Phase II were carried forward into Phase III. Therefore, in Phase III, MIF was initiated from a total of 1225000 parameter sets. Each initial condition and all other estimated parameter were allowed to vary in MIF with a random walk standard deviation of 0.02, with the exception of $\Delta\rho$ which was given a random walk standard deviation of 0.03. It is important to note that the random walk standard deviations for ρ and $\Delta\rho$ were time-varying. ρ was estimated using the data up through December 1945, and $\Delta\rho$ was estimated using the data beginning in January 1946. Thus, the random walk standard deviation for ρ was zero from Jan 1946 onward. Likewise, the random walk standard deviation for $\Delta\rho$ was zero up until January 1946. In order to define time-varying random walk standard deviations, the MIF2 method was used in the developer version of pomp, version 0.44-1. Refer to Table B.1 for a list of estimated parameters. Each MIF run had 45 iterations, 2000 particles, hyperbolic cooling with a variance factor of 2, and a cooling fraction of 0.5. Here the reporting transition remained fixed at January 1946, although $\Delta\rho$ was allowed to vary. Once again, for each parameter set, ten replicate particle filters were run, each with 2000 particles, to estimate the likelihood of the data from the state.

In Phase IV, the 1225000 parameter sets from Phase III were culled to remove all parameter combinations that had a likelihood 20 log-likelihood units below the maximum for that state. Using the culled parameter sets, MIF was run with 45 iterations, 2000 particles, hyperbolic cooling with a variance factor of 2, and a cooling fraction of 0.5. Random walk standard deviations were the same as Phase III. To estimate the likelihood, ten replicate particle filters were run, each with 2000 particles.

In Phase V, 49000 parameter sets were chosen using the 1000 top likelihood parameter sets for each US state from Phase IV. In Phase V, MIF was run with 45 iterations for each parameter set using the same MIF settings as Phase IV, with the exception of the $\Delta\rho$ random walk standard deviation, which was 0.02.

B.1.4.4 Phase VI & VII: Profiling Report Rate and Immigration

In Phase VI, we found the range of ρ and ψ in the top parameter sets for each state. For profiling, ρ and ψ values were generated to span the range. The ranges were [0.001, 0.1] and [6e-6, 2.5e-3], for ρ and ψ, respectively. There were 37 values of ρ, $(\rho_1, \rho_2, ..., \rho_{37})$, with 0.001 interval sampling in the range of [0.001,0.03] and 0.01 interval sampling in the range of [0.04,0.1]. There were 50 values of ψ, $(\psi_1, \psi_2, ..., \psi_{50})$ evenly sampled across the range. In order to make the profiles two dimensional, we took every pairwise combination of ρ and ψ, which resulted in 1850 profile points, each being a unique combination of ρ and ψ. Next, we needed to identify the other parameters to join with our 1850 combinations of ρ and ψ. At each of the 1850 profile points we used four parameter set variants for each US state. The first variant was the ρ_i and ψ_j combination substituted into the maximum likelihood parameter set from Phase V (US state-specific maximum). The second variant was the ρ_i and ψ_j combination substituted into the highest likelihood parameter set for which $\rho_{i-1} \leq \rho \leq \rho_{i+1}$. The third was the ρ_i and ψ_j combination substituted into the highest likelihood parameter set for which $\psi_{j-1} \leq \psi \leq \psi_{j+1}$. The fourth was the ρ_i and ψ_j combination substituted into the parameter set with the minimum Euclidean distance from ρ_i and ψ_j.

With four parameter set variants at each of the 1850 profile values, there were a total of 7400 parameter sets used to initialize the 2-D profile for each state. ρ and ψ were fixed at the given profile values, and MIF was used to maximize the likelihood along

the other parameter and initial condition dimensions. Starting at each of the 7400 parameter set variants, two replicates of MIF were run. Each estimated parameter and initial condition had a random walk standard deviation of 0.02. Once again, the random walk standard deviations for ρ and $\Delta\rho$ were time-varying. There were 45 iterations of MIF per replicate, with 2000 particles, hyperbolic cooling with a variance factor of 2, and a cooling fraction of 0.5. The reporting transition remained fixed at January 1946. This resulted in 14800 parameter sets for each state, and 725200 parameter sets across all US states. For each state, the likelihood of the 14800 parameter sets was estimated by running 10 replicate particle filters, each with 2000 particles.

In Phase VII, we found the highest likelihood parameter set at each state's 1850 profile points. The ρ and ψ values remained fixed at each profile point and two replicate MIF runs were preformed. Now each estimated parameter and initial condition had a random walk standard deviation of 0.01.There were 60 iterations of MIF per replicate, with 4000 particles, hyperbolic cooling with a variance factor of 2, and a cooling fraction of 0.5. This resulted in 3700 parameter sets for each state, and 181300 parameter sets across all US states. For each state, the likelihood of the 3700 parameter sets was estimated by running 10 replicate particle filters, each with 4000 particles.

B.1.4.5 Phase VIII: Profiling Report Rate and Immigration with Constant Reporting Assumption

In order to test whether the increased report rate during the Baby Boom Era, represented by $\Delta\rho$, is necessary to account for the increased incidence during the Baby Boom Era, we repeated the profiling in Phase VII, with $\Delta\rho = 0$. For each US state we then used AIC to determine which model best fit the data, the model with

$\Delta\rho = 0$ or the model with $\Delta\rho > 0$. For each state, the MLE was taken from the model with the lower AIC.

B.2 Polio Transmission Ecology: Results

B.2.1 Patterns in the Data

The negative exponential functions representing the relationship between fadeouts and population size for the Pre-Baby Boom and Baby Boom are given as,

$$\text{portion of fadeout months}_{\text{ Pre-Baby Boom}} = 9195e^{-1.8N} \mid N \text{ is } \log_{10}(\text{population size}), \tag{B.27}$$

$$\text{portion of fadeout months}_{\text{ Baby Boom}} = 10230e^{-2N}. \tag{B.28}$$

B.2.2 Inference Results

Model Validation. In order to test the fit of the model for each state we first used the MLE for each state to calculate the expected number of infected non-infants for each month of the fitted data, May 1932–January 1953 for all states except South Dakota and Texas whose inference started in May 1933 and 1934, respectively. The expected number of infections was obtained by running a particle filter with 2000 particles and taking the prediction mean of the particles at each time point. The prediction mean is the expected state value at time t, given the data and the state values up to time $t-1$.

$$\mathbb{E}(\mathbf{X}_t \mid \mathbf{X}_{t-1}, \text{cases}_{t-1}) \tag{B.29}$$

In order to get the expected number of cases, we took the expected number of non-infant infections and multiplied it by the state- and time-specific report rate.

$$\mathbb{E}(\text{cases}_t) = \rho_t \mathbb{E}(I_t^O \mid \mathbf{X}_{t-1}, \text{cases}_{t-1}) \tag{B.30}$$

For model validation, the r-squared was quantified for the model as the reduction in the sum-of-squared deviations obtained when using the expected cases from the state-specific MLE versus a state-specific null model of the mean number of cases.

$$R^2 = 1 - \frac{\sum_t \left[\mathbb{E}(\text{cases}_t) - \text{cases}_t\right]^2}{\sum_t \left[\text{mean}(\text{cases}) - \text{cases}_t\right]^2} \tag{B.31}$$

The generalized r-squared was quantified for the model as the reduction in the sum-of-squared deviations across all US states obtained when using the expected cases from the state-specific MLE versus a state-specific null model of the mean number of cases.

$$R^2_{\text{generalized}} = 1 - \frac{\sum_{i=1}^{49} \sum_t \left[\mathbb{E}(\text{cases}_t) - \text{cases}_t\right]^2}{\sum_{i=1}^{49} \sum_t \left[\text{mean}(\text{cases}) - \text{cases}_t\right]^2}. \tag{B.32}$$

The out-of-fit predictions were done using the same procedure as was used for the fitted data region. The out-of-fit predictions were done using the data from the last two epidemic years, January 1953–December 1954.

Predicated Latitudinal Gradient. In order to determine whether or not the model reproduces the latitudinal gradient in the timing of polio epidemics, 10 stochastic simulations were run using the MLE for each state. For these 10 simulations, we used the inference time period inclusive to all states, May 1934–January 1953. The simulations were taken together to test for a gradient in the simulated data. A wavelet analysis was run on the simulated time series and the phase angles associated with the 1 year period were used to rank the states 1-49, 1 being the earliest epidemic peak and 49 being the latest epidemic peak. We found that the model reproduces the latitudinal gradient in the timing of polio epidemics (main text Fig 4.3) with epidemics peaking earlier in the southern US and later in the northern US (observed gradient shown in Fig B.1). In order to determine whether the latitudinal gradient in the timing of polio epidemics is driven by the latitudinal gradient in birth peak tim-

ing, and/or seasonal birth amplitude, we simulated the model with birth seasonality removed. We did this by decomposing the birth time series using the stl function in R and using the trend in births as a model covariate, rather than using the actual births. We found that the model without birth seasonality (Fig B.4) also displays a latitudinal gradient (Fig B.5). The latitudinal gradient in the timing of polio epidemics is largely driven by the latitudinal variation in the timing of peak polio transmission. Fig B.7 shows that the peak transmission rate for polio occurs between May and July in the US. The peak transmission rate in southern states typically occurs in May and the peak transmission rate in northern states occurs in July, a couple months earlier than their respective peak in incidence. In order to test whether the predicted timing of peak polio incidence matches the observed timing, we compared the mean rank calculated from the data and the mean rank based off simulation. See main text Fig 4.3D, showing the correlation between the observed and predicted mean rank. You will also see in Fig B.5C–D that the simulated slope and R^2 values of latitude vs. mean rank reflect a latitudinal gradient.

While geographical variation in birth seasonality was insufficient to explain the latitudinal gradient seen in epidemic timing, birth seasonality played a role in shaping seasonal incidence of polio. Epidemiological theory indicates that birth seasonality can have important dynamical consequences for childhood diseases [10, 11, 12]. To quantify the influence of birth seasonality (Fig B.4) on infant infections, we compared simulations of the fitted models to simulations for which seasonal fluctuations in births were removed. In the presence of birth seasonality, infant infection incidence was generally higher during the epidemic peak (Fig B.6); however, since infant infections were assumed to be asymptomatic, this did not affect the incidence of disease directly, and no indirect effect was observed. We attribute the negligible effect of birth seasonality on disease incidence to the low amplitude of birth seasonality, which

was $\sim 10\%$ in the US at this time[11].

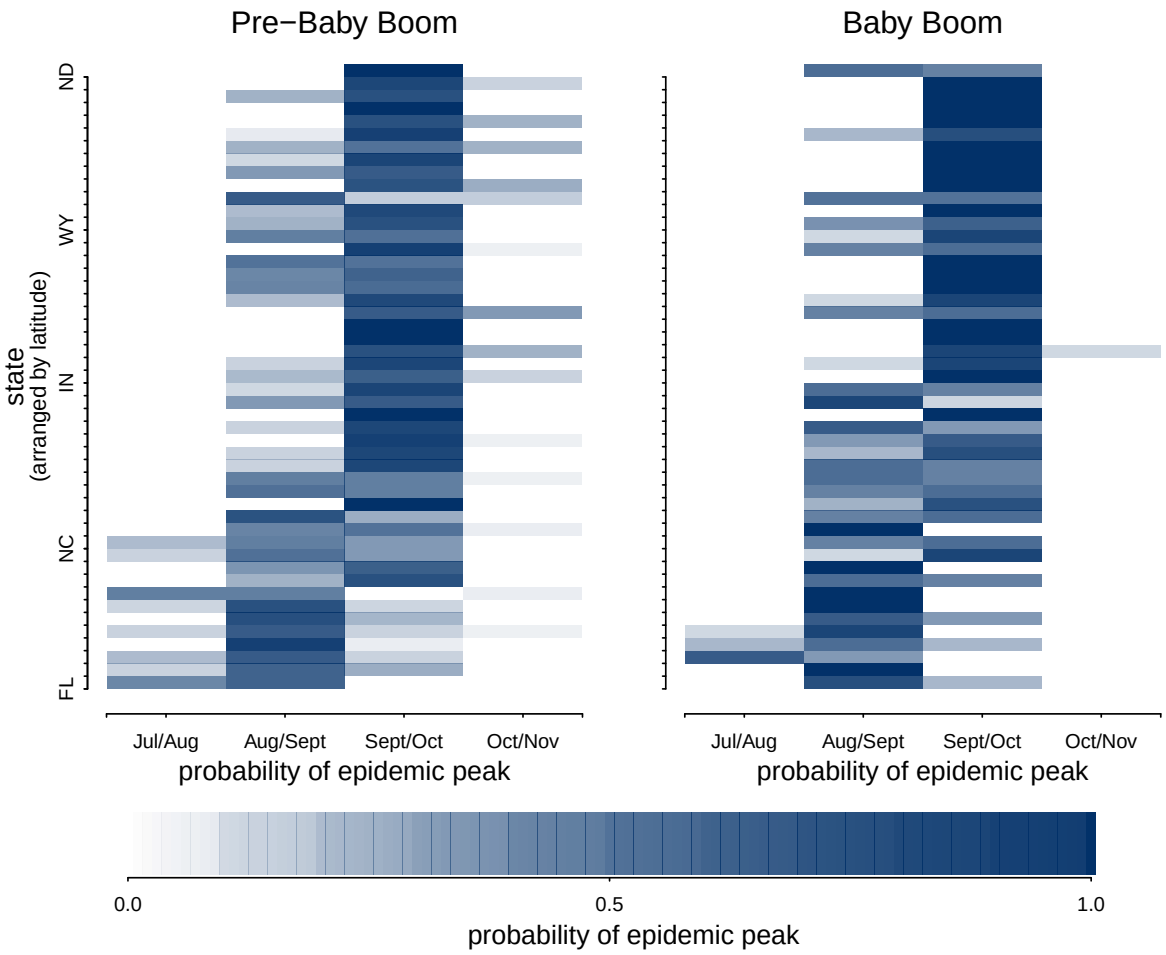


Figure B.1: Probability densities of the timing of the epidemic peak for each state during the pre-baby boom and baby boom eras. For each state, the timing of the epidemic peak was determined for each year with 20+ polio cases. The probability was measured as the portion of years (in each era) for which the peak occurred between July/Aug, Aug/Sept, Sept/Oct, and Oct/Nov. Peak timing was measured using the 1 yr wavelet band phase angle.

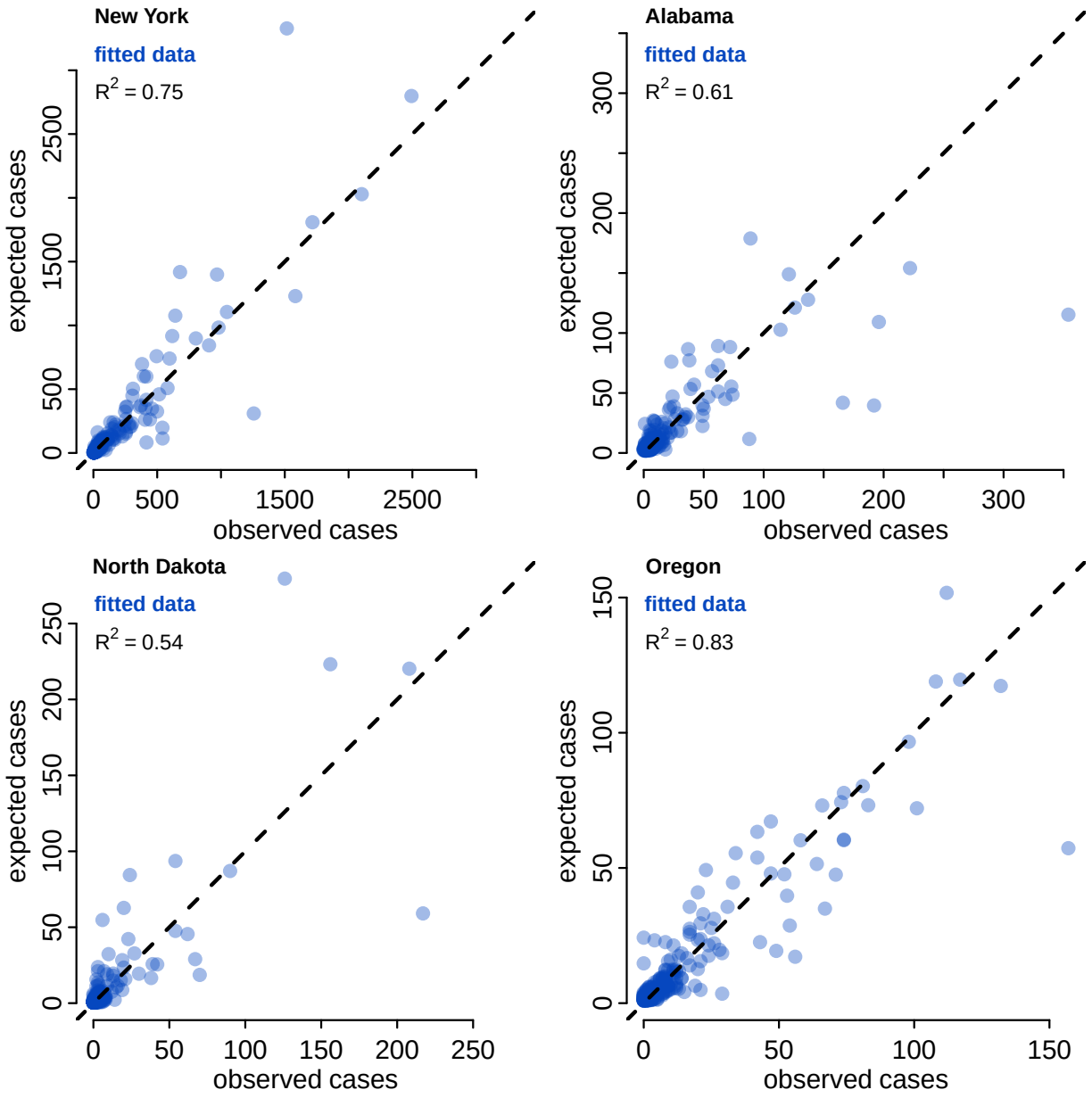


Figure B.2: One-step-ahead predictions based on the MLE model for each of four states. Observed and expected cases are shown on the natural scale, in contrast to Fig 4.3B in the main text, which is on the $\log_{10}$ scale.

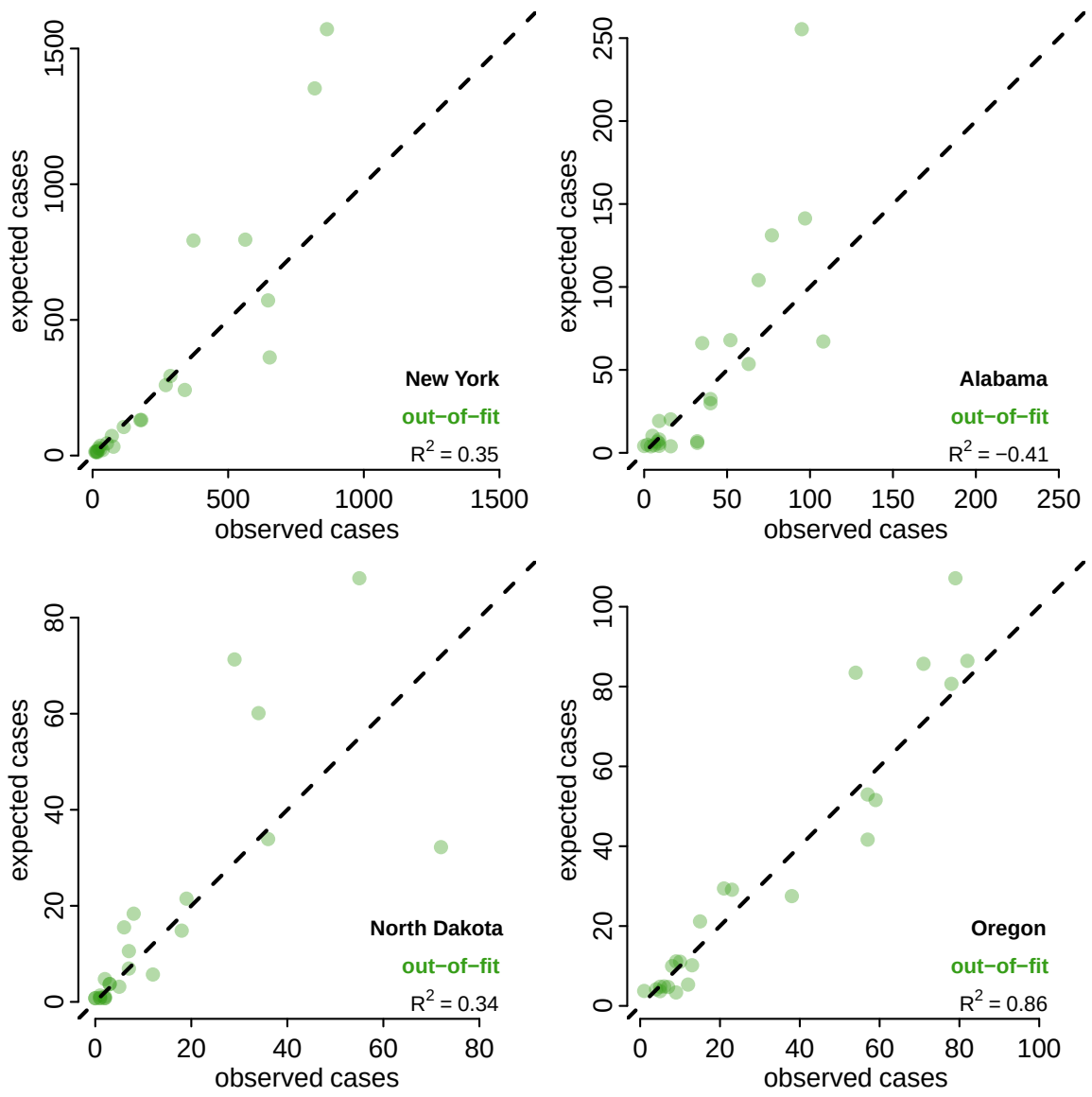


Figure B.3: Out-of-fit predictions based on the MLE model for each of four states. Observed and expected cases are shown on the natural scale, in contrast to Fig 4.3C in the main text, which is on the $\log_{10}$ scale. A negative R^2 value indicates that the null model had a lower sum of squared deviations than the fitted model for the out-of-fit predictions, which was the case for 3 of the 49 states (Alabama, Connecticut, and Delaware). Alabama was shown as an example of one of the "worst-fit" states.

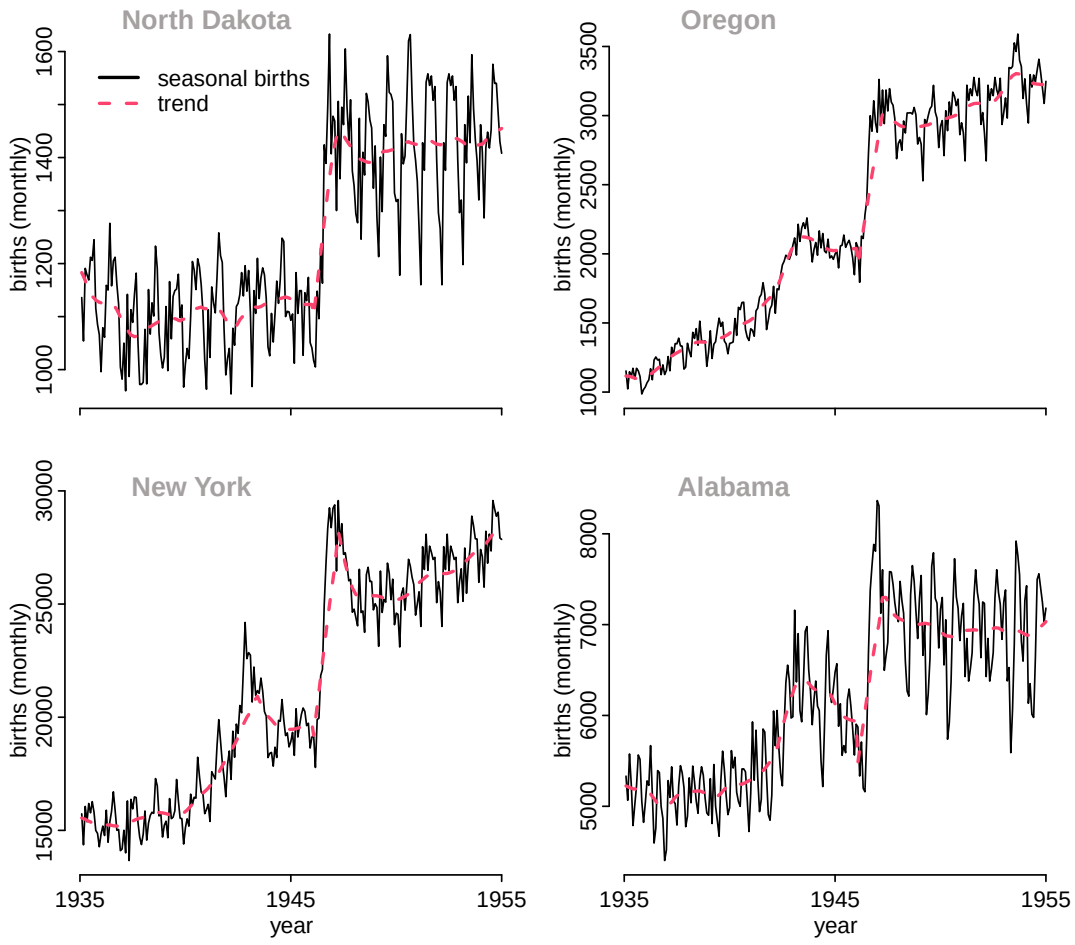


Figure B.4: Monthly births (black time series) are seasonal. Monthly births were used as covariates for the fitted models. Th trend in births is shown by the dashed line (fuchsia). Four states from different geographic regions are shown.

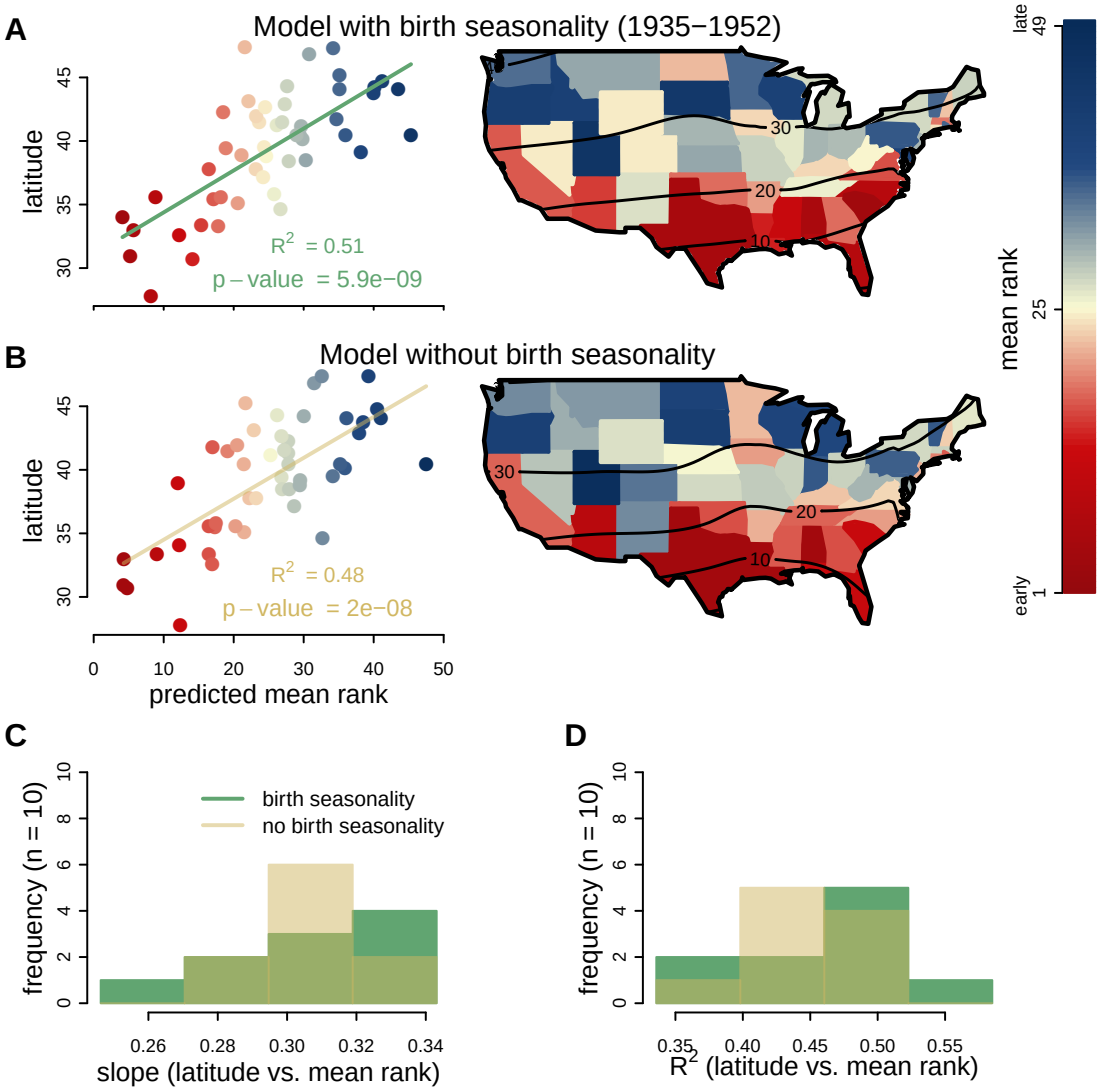


Figure B.5: Latitudinal gradient predicted from model simulations using maximum likelihood parameter estimates for each state. (A) Latitudinal gradient in simulations of the fitted models with seasonal births. (B) Latitudinal gradient in simulations with birth seasonality removed (i.e., trend in births was used). (C) The distribution of the latitudinal gradient slopes for 10 simulations for each US state for the model with birth seasonality (i.e. using the raw birth data) and the model with birth seasonality removed (i.e. using the trend in births). (D) The R^2 for the latitudinal gradient for the 10 simulations with and without birth seasonality.

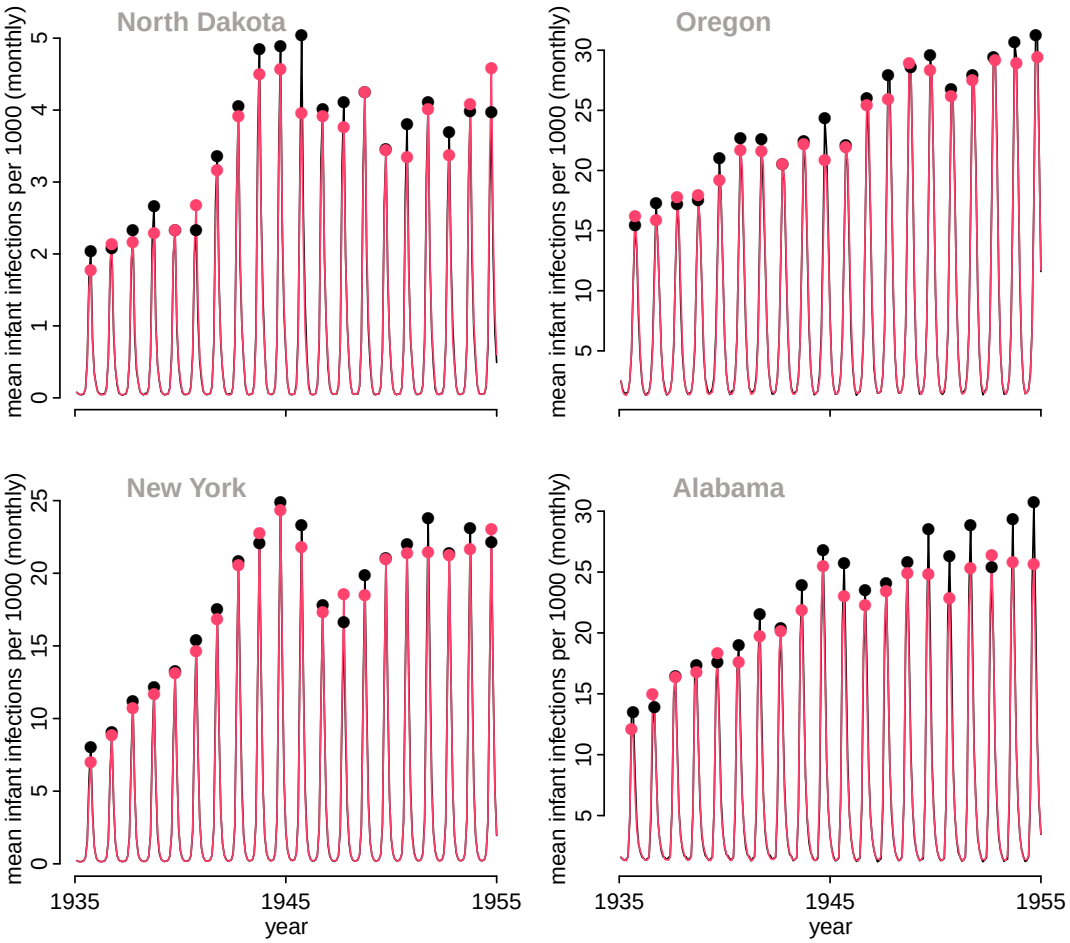


Figure B.6: Mean monthly incidence of infant-infections per 1000 infants, with and without birth seasonality. Model with birth seasonality is shown in black, model without birth seasonality in fuchsia. 500 simulation were run for each state and model combination. Points indicate annual peak incidence. The model with birth seasonality generally displayed higher peak infant-infections. Four states from different geographic regions are shown.

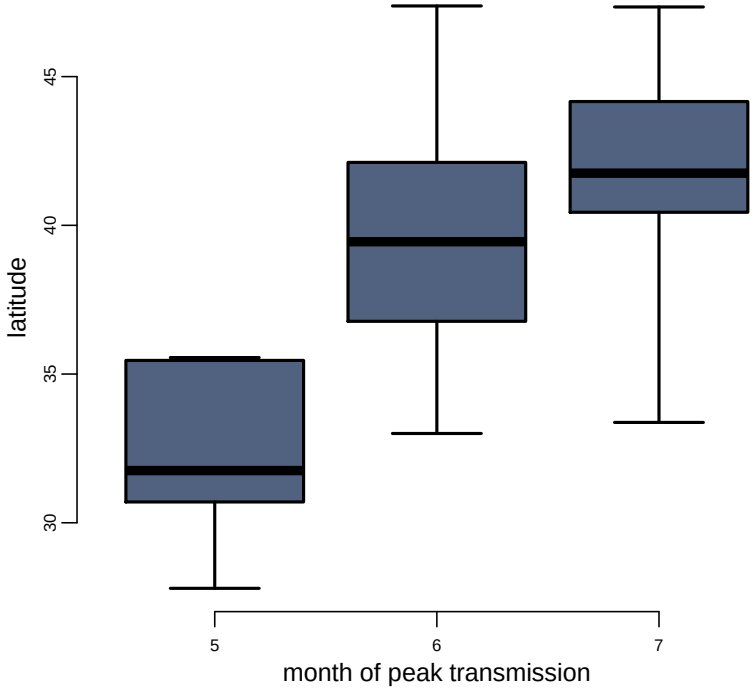


Figure B.7: Latitudinal variation in MLEs of peak transmission timing.

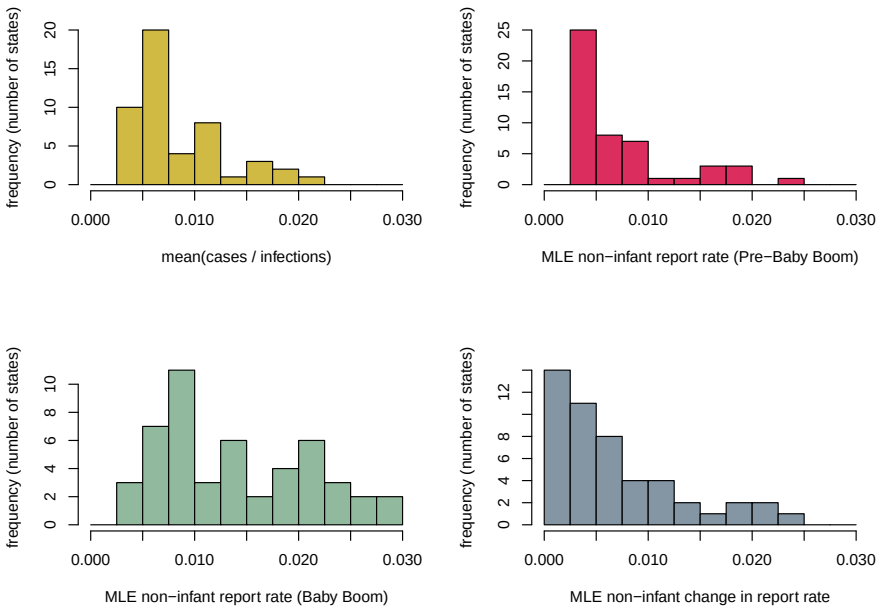


Figure B.8: MLE report rates. The report rate is a composite parameter that encompasses the probability of an infected individual becoming symptomatic, and the subsequent probability that symptomatic infections are reported. (top-left) The distribution of mean report rates for infants and non-infants across states. Non-infant report rates during the pre-baby boom era (top-right) were lower than that of the baby boom era (bottom-left), due to the increase shown in (bottom-right).

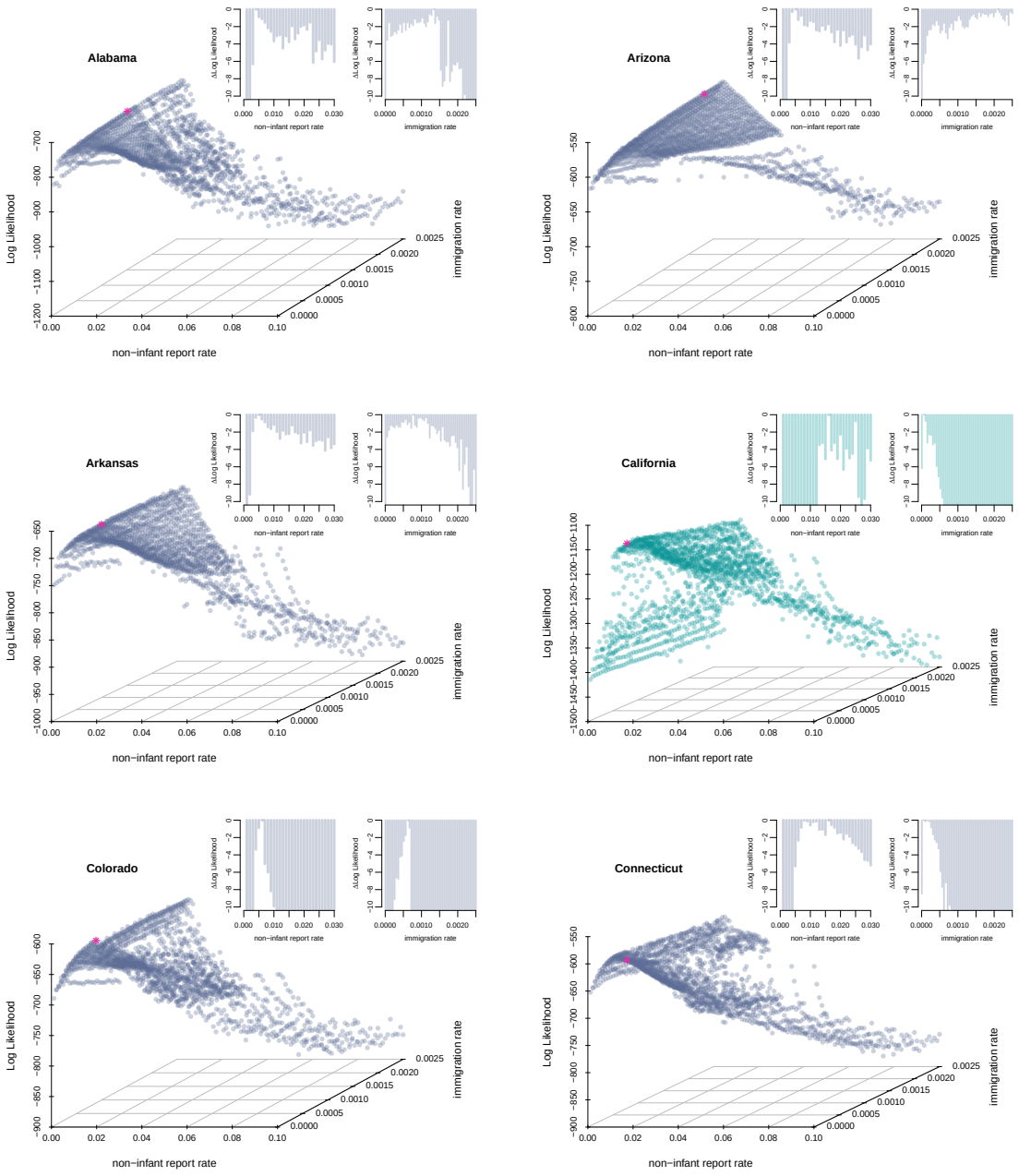


Figure B.9: Likelihood profiles for the pre-baby boom non-infant report rate (ρ_t) and the immigration rate (ψ). MLE indicated by pink asterisk. Profile color indicates whether the report rate was increased during the baby boom (purple profiles) or if the report rate was constant through time (green profiles).

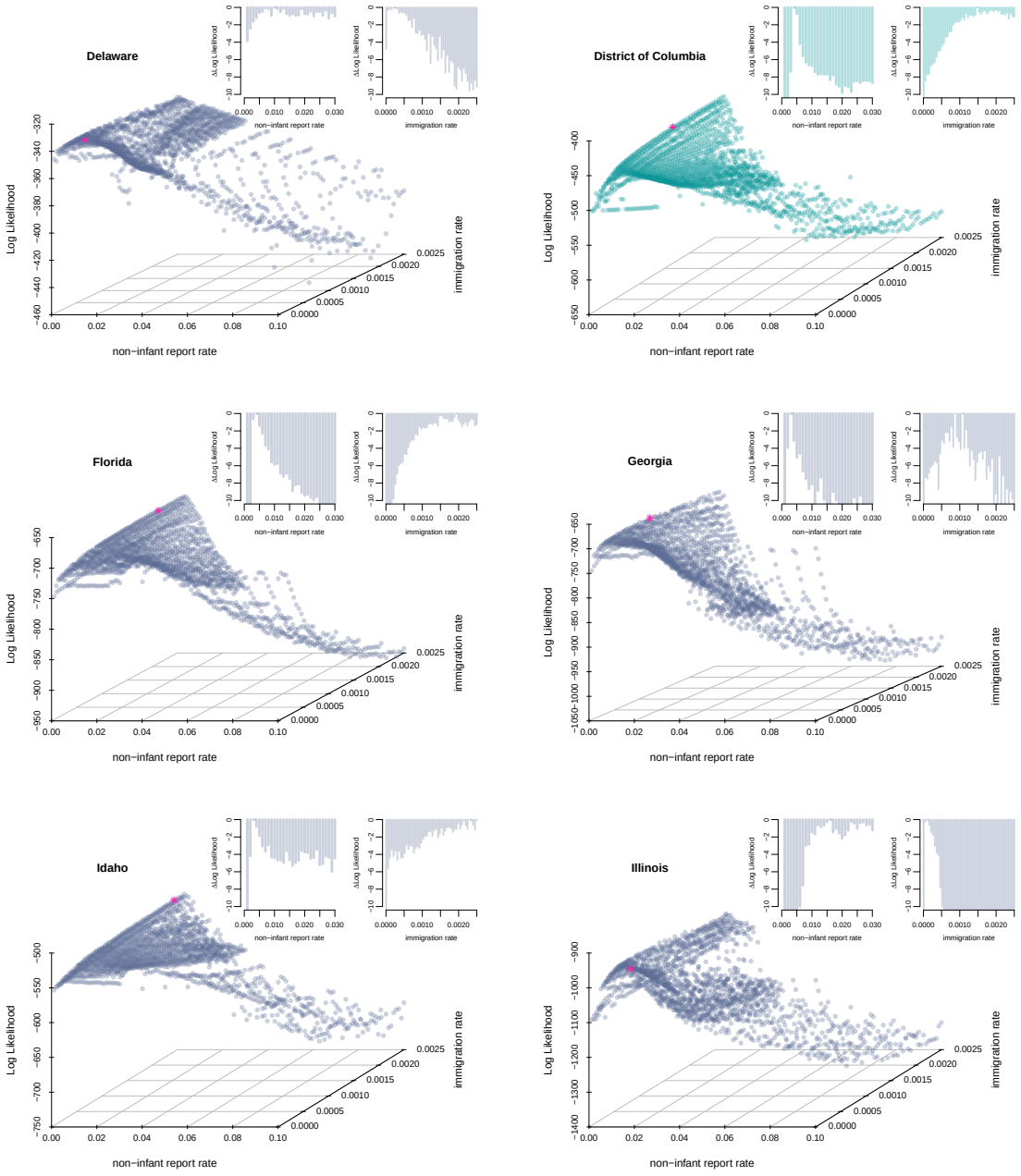


Figure B.10: Likelihood profiles for the pre-baby boom non-infant report rate (ρ_t) and the immigration rate (ψ). MLE indicated by pink asterisk. Profile color indicates the whether the report rate was increased during the baby boom (purple) or if the report rate was constant through time (green).

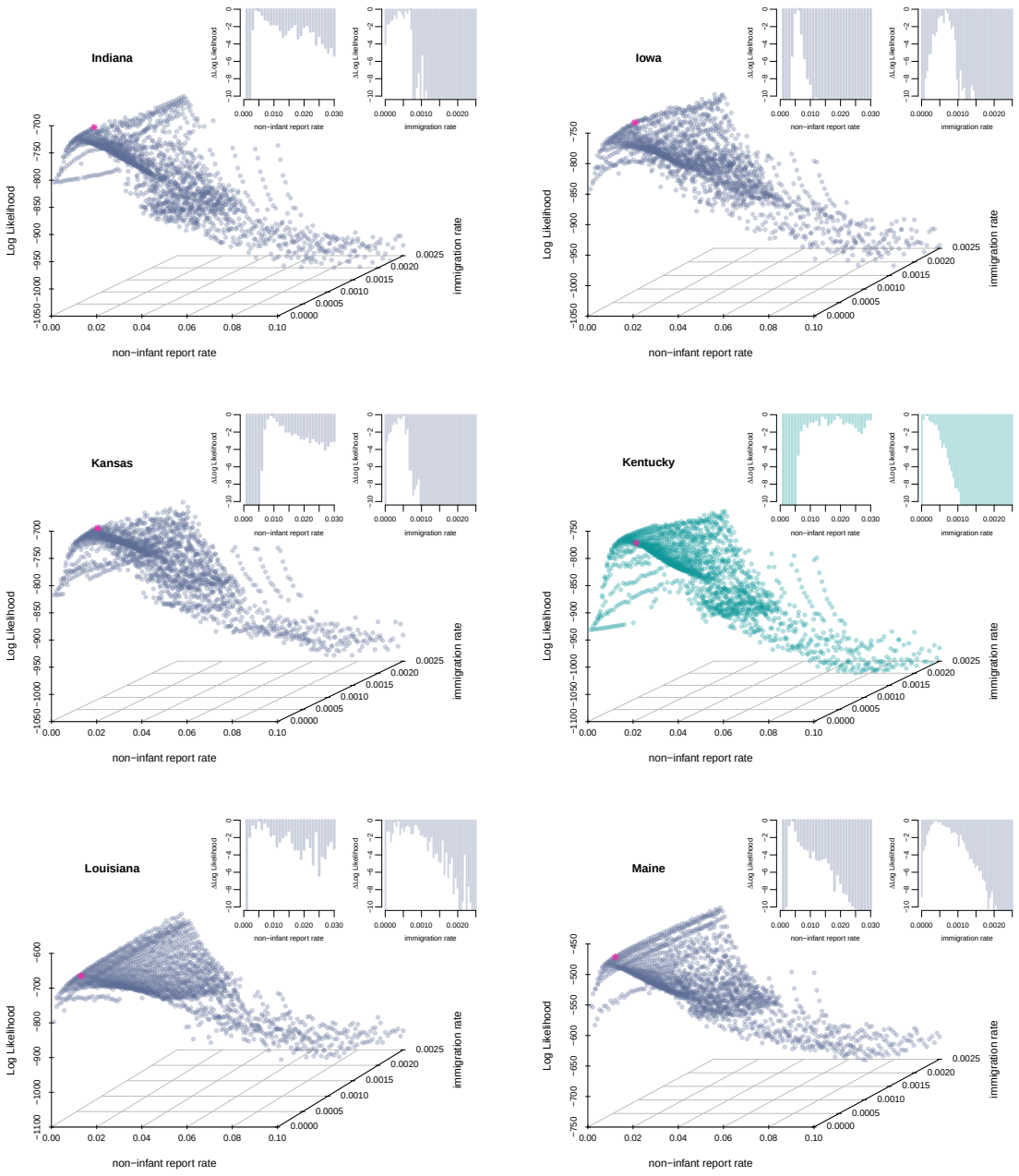


Figure B.11: Likelihood profiles for the pre-baby boom non-infant report rate (ρ_t) and the immigration rate (ψ). MLE indicated by pink asterisk. Profile color indicates the whether the report rate was increased during the baby boom (purple) or if the report rate was constant through time (green).

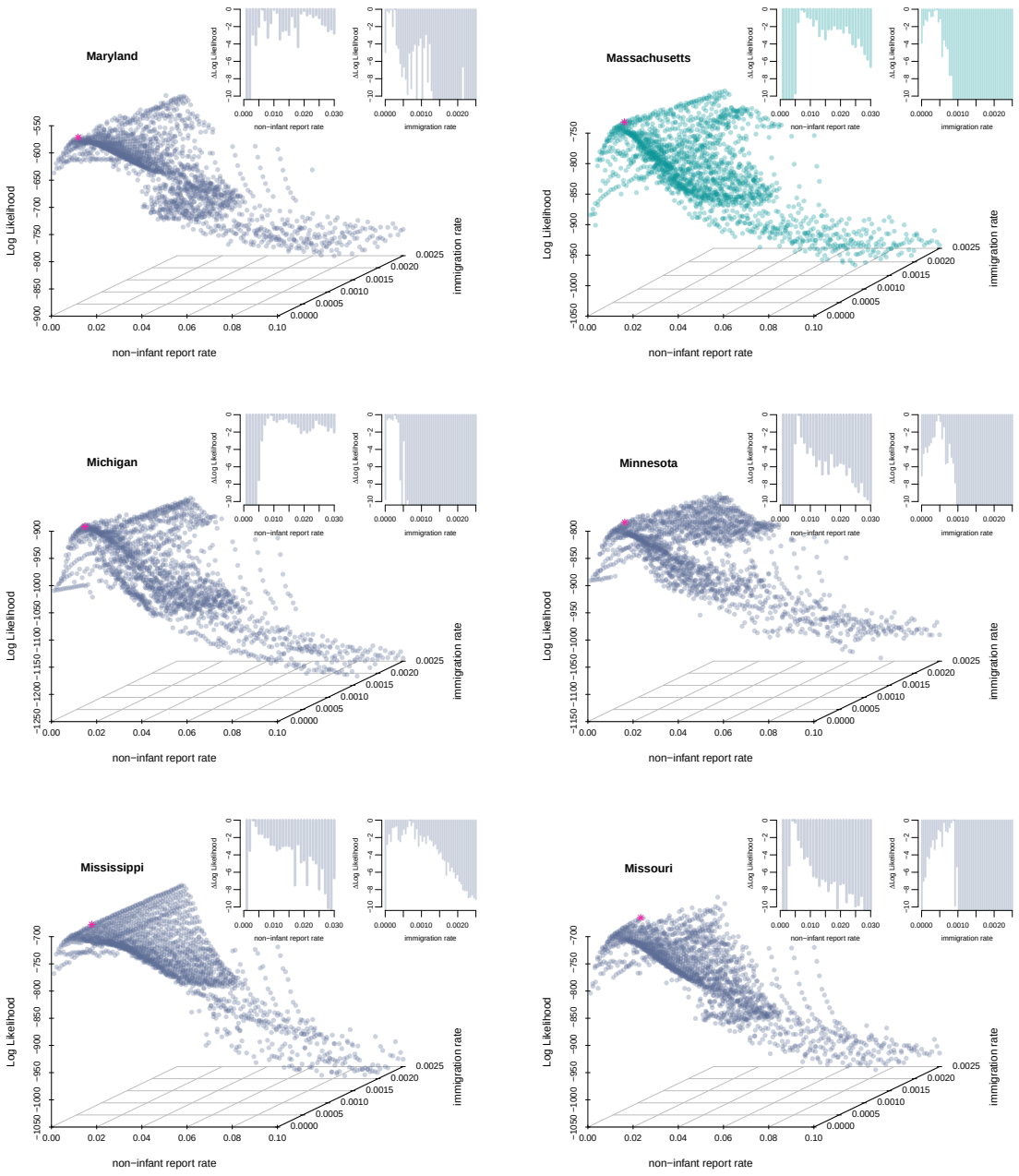


Figure B.12: Likelihood profiles for the pre-baby boom non-infant report rate (ρ_t) and the immigration rate (ψ). MLE indicated by pink asterisk. Profile color indicates the whether the report rate was increased during the baby boom (purple) or if the report rate was constant through time (green).

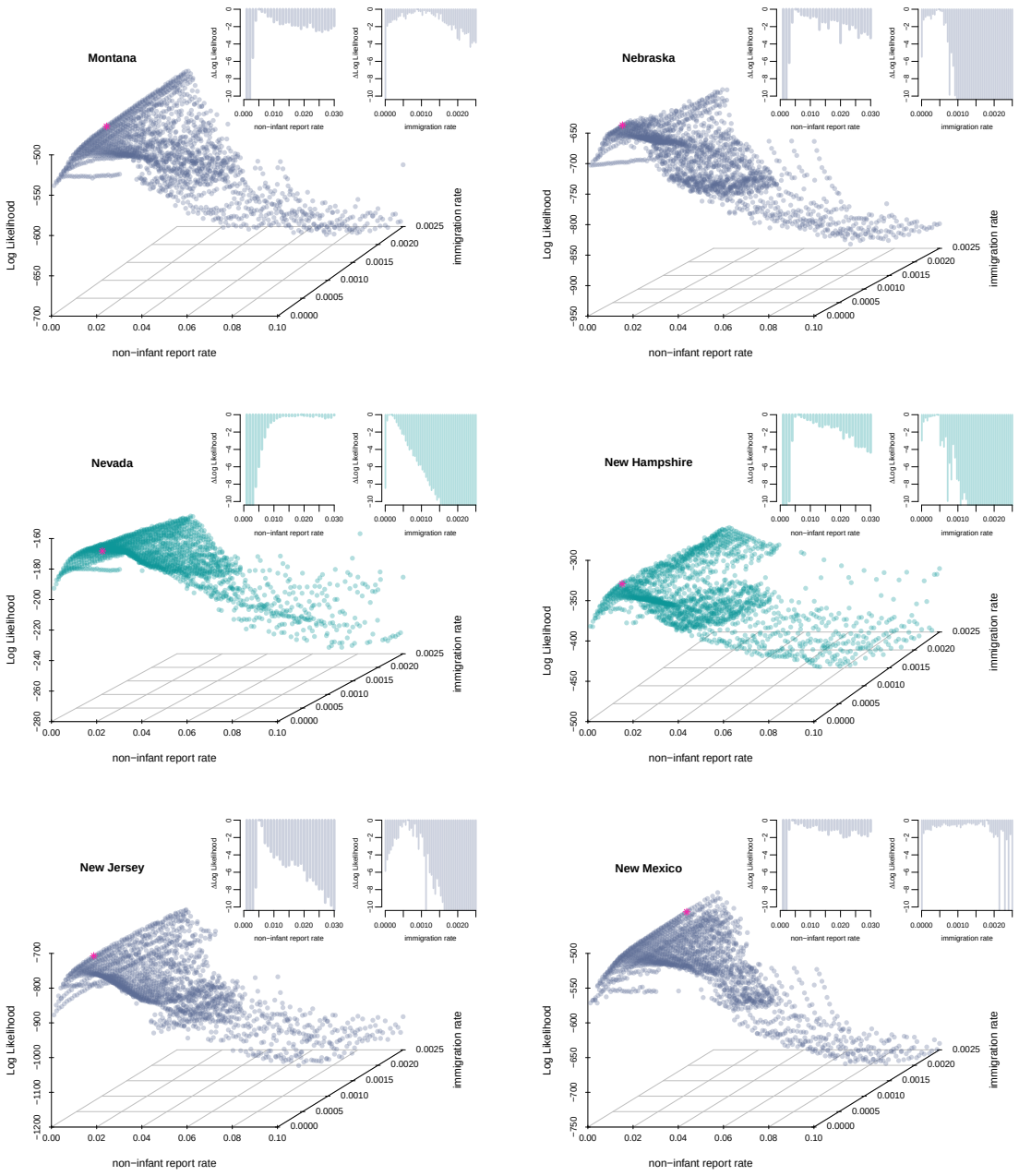


Figure B.13: Likelihood profiles for the pre-baby boom non-infant report rate (ρ_t) and the immigration rate (ψ). MLE indicated by pink asterisk. Profile color indicates the whether the report rate was increased during the baby boom (purple) or if the report rate was constant through time (green).

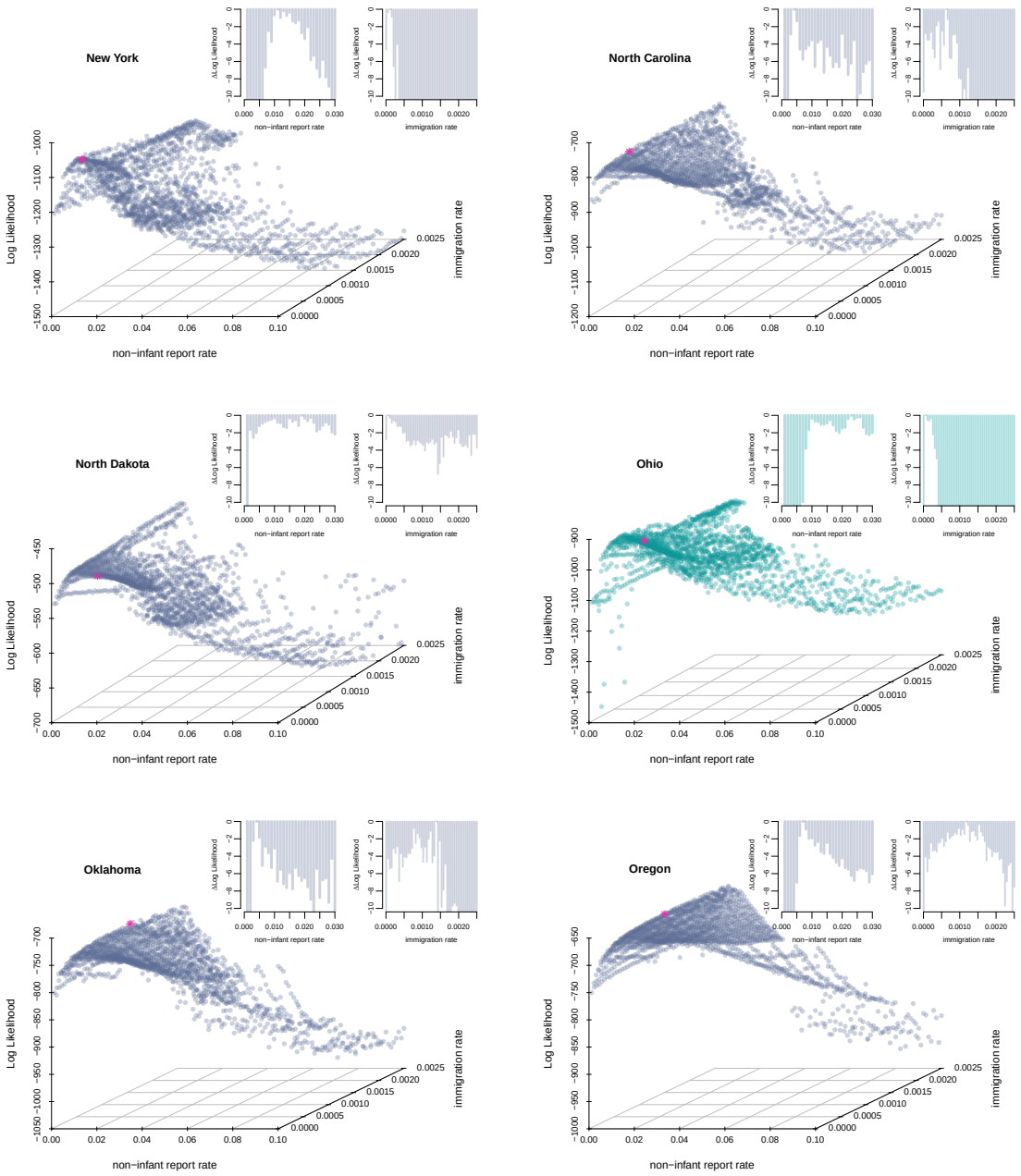


Figure B.14: Likelihood profiles for the pre-baby boom non-infant report rate (ρ_t) and the immigration rate (ψ). MLE indicated by pink asterisk. Profile color indicates the whether the report rate was increased during the baby boom (purple) or if the report rate was constant through time (green).

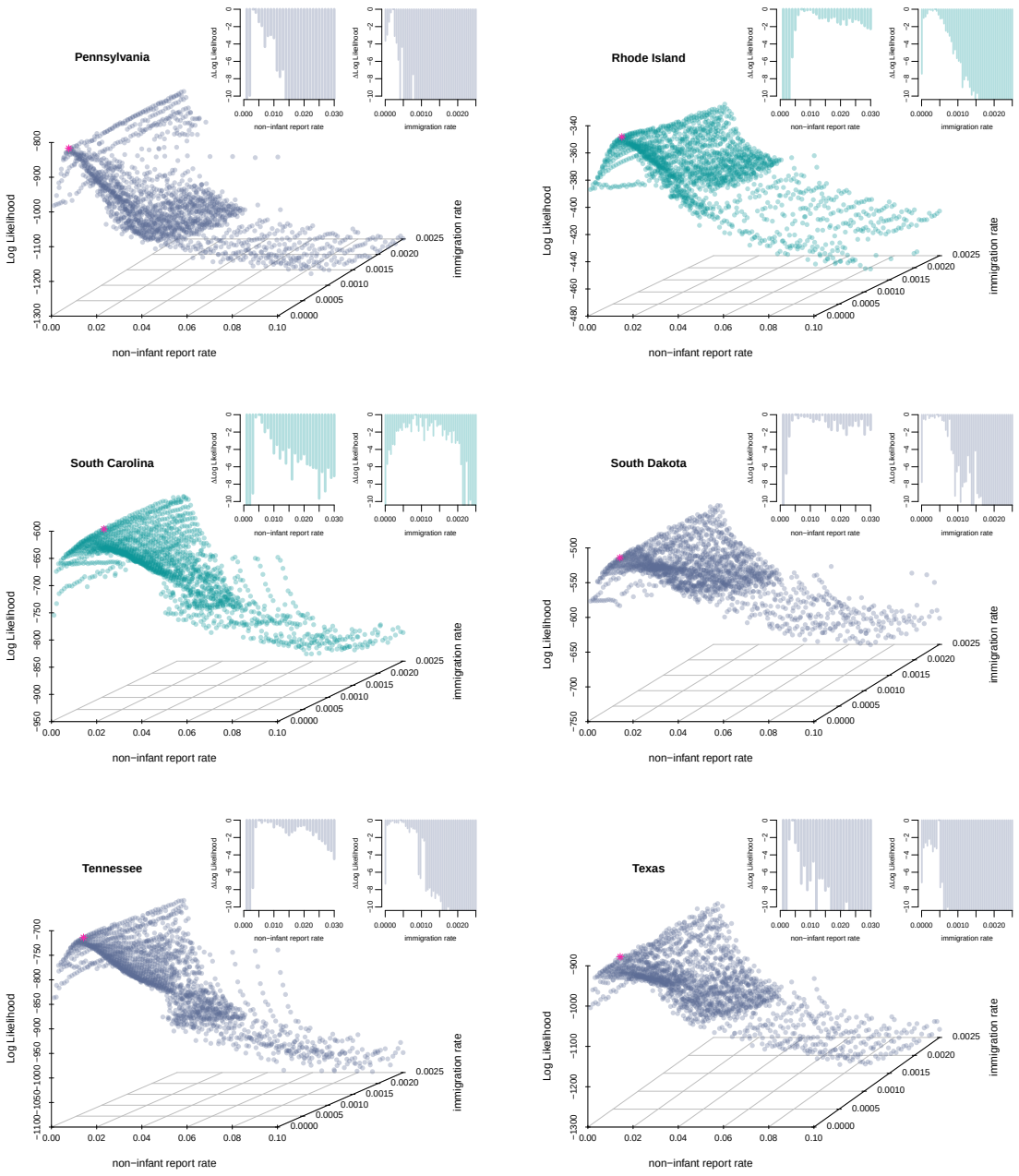


Figure B.15: Likelihood profiles for the pre-baby boom non-infant report rate (ρ_t) and the immigration rate (ψ). MLE indicated by pink asterisk. Profile color indicates the whether the report rate was increased during the baby boom (purple) or if the report rate was constant through time (green).

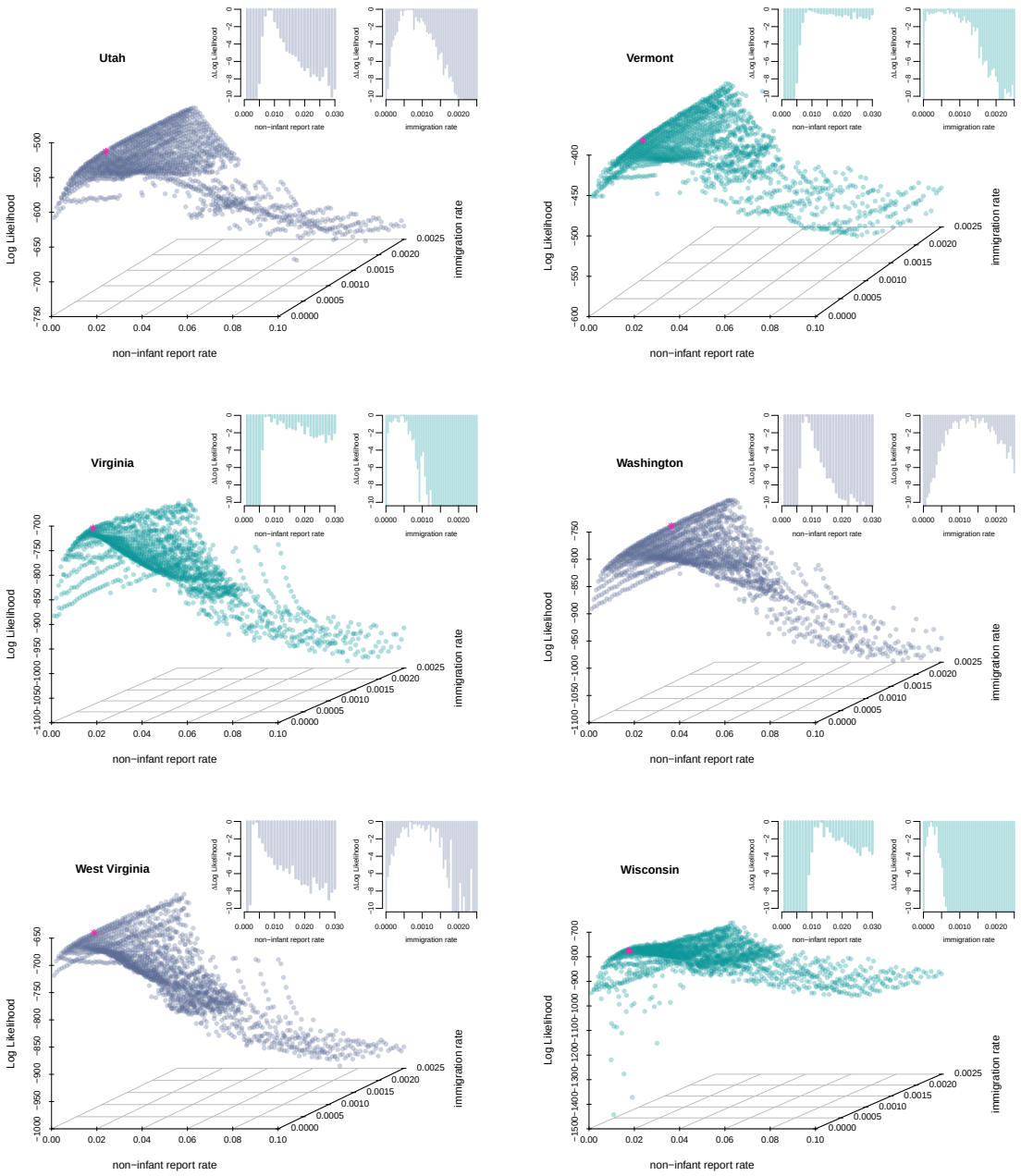


Figure B.16: Likelihood profiles for the pre-baby boom non-infant report rate (ρ_t) and the immigration rate (ψ). MLE indicated by pink asterisk. Profile color indicates the whether the report rate was increased during the baby boom (purple) or if the report rate was constant through time (green).

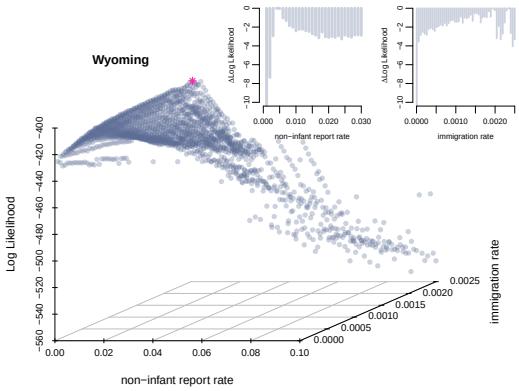


Figure B.17: Likelihood profiles for the pre-baby boom non-infant report rate (ρ_t) and the immigration rate (ψ). MLE indicated by pink asterisk. Profile color indicates the whether the report rate was increased during the baby boom (purple) or if the report rate was constant through time (green).

Bibliography

[1] Trevelyan, B, Smallman-Raynor, M, & Cliff, A. D. (2005) The Spatial Dynamics of Poliomyelitis in the United States: From Epidemic Emergence to Vaccine-Induced Retreat, 1910-1971. *Annals of the Association of American Geographers* **95**, 269–293.

[2] United States Centers for Disease Control and Prevention. (2010) United States Vital Statistics (http://www.cdc.gov/nchs/products/vsus.htm). Accessed: March 1, 2010.

[3] United States Census Bureau. (2010) United States Population Estimates: Historical Estimates (http://www.census.gov/popest/data/historical/index.html). Accessed: March 1, 2010.

[4] Bjørnstad, O. N. (2013) Spatial Nonparametric Covariance Function (R package ncf).

[5] Gouhier, T. (2014) *biwavelet: Conduct univariate and bivariate wavelet analyses.* (Version 0.17.4).

[6] Iannone, R. (2015) DiagrammeR: Create Diagrams and Flowcharts Using R.

[7] Alexander, J. P, Gary, H. E, & Pallansch, M. A. (1997) Duration of Poliovirus Excretion and its Implications for Acute Flaccid Paralysis Surveillance: A Review of the Literature. *The Journal of Infectious Diseases* **175, Suppl**, S176–S182.

[8] King, A. A, Ionides, E. L, Breto, C. M, Ellner, S, Kendall, B, Wearing, H, Ferrari, M. J, Lavine, M, & Reuman, D. C. (2010) pomp: Statistical Inference for Partially Observed Markov Processes (R package).

[9] King, A. A, Nguyen, D, & Ionides, E. (in press) Statistical Inference for Partially Observed Markov Processes via the R Package pomp. *Journal of Statistical Software*.

[10] Keeling, M & Rohani, P. (2008) *Modeling Infectious Diseases in Humans and Animals*. (Princeton University Press, Princeton).

[11] Martinez-Bakker, M, Bakker, K, King, A. A, & Rohani, P. (2014) Human Birth Seasonality: Latitudinal Gradient and Interplay with Childhood Disease Dynamics. *Proceedings of the Royal Society B: Biological Sciences* **281**.

[12] Dorélien, A, Ballesteros, S, & Grenfell, B. T. (2013) Impact of Birth Seasonality on Dynamics of Acute Immunizing Infections in Sub-Saharan Africa. *PLoS ONE* **8**, e75806.

APPENDIX C

Both Salk and Sabin Vaccines Effectively Reduce Polio Transmission in Epidemic Settings

C.1 Salk and Sabin Vaccines: Data from the US

Polio case data from the US were separated among 3 data sets, (1) the US Morbidity and Mortality Weekly Report (MMWR) monthly polio case data, (2) the US National Office of Vital Statistics (NOVS) annual polio case data, and (3) the Centers for Disease Control and Prevention Poliomyelitis Surveillance Unit (PSU) annual polio case data. There were 2 demographic data sets, (1) monthly births from the US National Office of Vital Statistics, and (2) annual population sizes from the US Census Bureau. There were 6 sets of data regarding the vaccines: the Inactive polio vaccine (IPV), also known as the Salk Vaccine, and the Oral polio vaccine (OPV), also known as the Sabin Vaccine. The 6 vaccine data sets were (1) annual age-structured IPV coverage in the US, by dose of IPV, (2) quarterly or annual shipments of IPV in the US, (3) annual age-structured OPV coverage in the US, by dose of OPV, (4) annual shipments of OPV in the US, (5) annual-age structured coverage of dual vaccination with IPV and OPV, and (6) annual age-structured percent of the population unvac-

cinated or under-vaccinated with both IPV and OPV. The data sets are discussed in more detail in the following sections.

C.1.1 MMWR Polio Data.

The MMWR data consisted of monthly polio cases reported in the weekly US Public Health Reports and the CDC Morbidity and Mortality Weekly Report from 1931–1968 for each US state and the District of Columbia. The data were first published in [1], which includes discussion of data quality during this period. The data for the contiguous US states and the District of Columbia were aggregated to construct a single time series of "national" data, which is used for the entirety of this project. For details on the state-level data please refer to our previous study [2]. The MMWR case data were aggregated on the national-level because our vaccine data were national-level data.

C.1.2 NOVS Polio Data.

The NOVS data contained the number of reported polio cases in the US each year from 1951–1965. These annual cases were broken down by symptomatology (i.e., paralytic, non-paralytic, and unspecified). These data were digitized from the Communicable Disease Center Poliomyelitis Surveillance Report Number 268, September 21, 1962, Supplement: The Association of Cases of Poliomyelitis with the use of Type III Oral Poliomyelitis Vaccines: A Technical Report by Luther L. Terry, Surgeon General, Public Health Service, US Department of Health, Education, and Welfare, September 20 1962. The data from 1951–1954 and 1962–1965 came from Public Health Reports volume 82, Number 5, May 1967 titled "Surveillance of Poliomyelitis in the United States, 1962–1965" by Morris, Witte, Gardner, Miller, and Henderson. The data in the aforementioned reports were compiled by the Poliomyelitis Surveillance Unit, Communicable Disease Center, Public Health Service, US Department of

Health, Education, and Welfare. The NOVS polio case data and the MMWR polio case data are not completely independent. It is the National Office of Vital Statistics that provided the polio case data reported in the MMWR. The NOVS data are an annual breakdown of the MMWR data by symptomatology.

C.1.3 PSU Polio Data.

The PSU data were digitized from the US Centers for Disease Control and Prevention Poliomyelitis Surveillance Unit Reports, these reports were bound as periodicals and are housed at the Centers for Disease Control and Prevention Library in the Clifton Landmark Collection, with the call number WC 555 C397p. Each table from the PSU reports from 1955 onward was evaluated for data. The earliest reports from which we digitized data were from 1959. The latest reports from which we digitized data were from the late 1970s. The PSU polio case data contained an annual breakdown of polio cases by age, symptomatology, and vaccine status. The Poliomyelitis Surveillance Unit was tasked with gathering the details on every polio case in the US and ensuring vaccine safety following the Cutter incident. Therefore, the Poliomyelitis Surveillance Unit collected their own case data, independent of the National Office of Vital Statistics, and the number of cases reported differed among the PSU and NOVS data.

C.1.4 Demography Data.

Monthly state–level time series of live births from 1931–1954 were downloaded from the Vital Statistics of the United States [3]. Annual state–level population size data were collected from the Population Distribution Branch of the U.S. Census Bureau [4]. The state-level population size and birth data for the contiguous US were aggregated to form national-level demography time series.

C.1.5 Vaccine Data.

Vaccination data were also gathered from the Poliomyelitis Surveillance Unit Reports housed at the CDC Library, Clifton Landmark Collection. The main text shows the raw vaccine coverage data, which are age-stratified vaccine coverage for the age classes: under 1 year old, 1–4 years old, 5–9, 10–14 years old. Data for these age classes were used to construct the vaccine coverage for the 1–14 year old age group as a whole (Figure S1). The under 1 year old age class was excluded due to the lack of data for this age class from 1962 onward, and because vaccination efforts were concentrated on the school-aged children. Age-structured population size and vaccine coverage data for the 1–14 year old age group constructed from the population size and coverage data for the 1–4, 5–9, and 10–14 year old age classes. The annual vaccine coverage data for the 1–14 year old age group was used for statistical inference. Interpolations are shown for visualization. We lacked IPV coverage data from 1955-1958; however, the cumulative distribution of IPV suggests a linear increase in vaccine coverage from this time period. Therefore, we fit a single IPV per capita vaccine uptake rate for the US from April 1955–December 1959.

C.2 Data from the USSR

In order to track the success of the massive polio immunization campaign, the USSR's Poliomyelitis Research Laboratory tracked vaccine coverage and clinical cases. Polio case data and OPV data from the USSR were digitized from *Some Results on the Work on Mass Immunization in the Soviet Union with Live Poliovirus Vaccine prepared from Sabin Strains* in **Bulletin of the World Health Organization**, 1961, 25, 79-91. Demography data for the USSR was obtained from the United Nations Demography database. We downloaded crude birth rates (per 1000 people) for the RSFSR, Ukraine, Estonia, and Lithuania. The crude birth rate indicates the

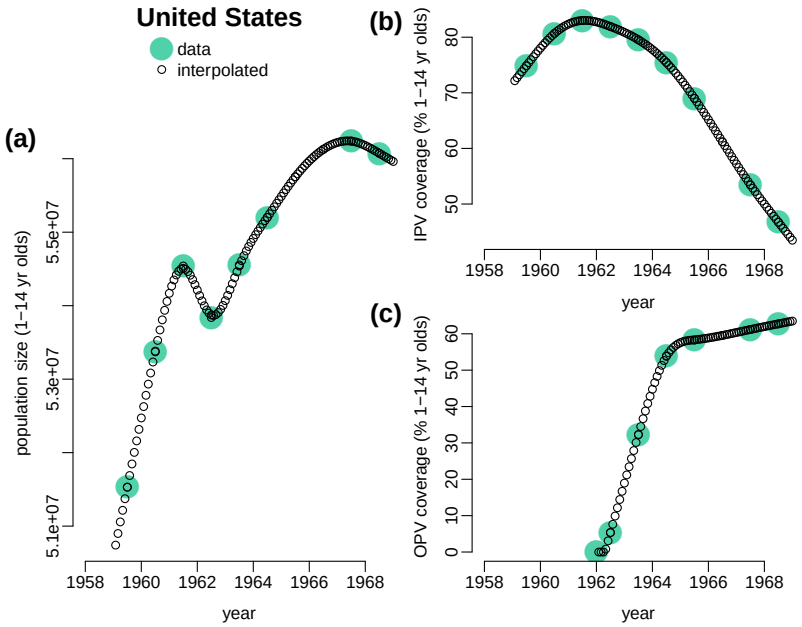


Figure C.1: Age-structured population size and vaccine coverage data for the 1–14 year old age group constructed from the population size and coverage data for the 1–4, 5–9, and 10–14 year old age classes. The annual vaccine coverage data for the 1–14 year old age group was used for statistical inference. Interpolations are shown for visualization.

number of live births occurring during the year, per 1000 population estimated at midyear. Data available online at https://data.un.org. Population data were also downloaded for each region in the USSR. Population data are available online at https://data.un.org.

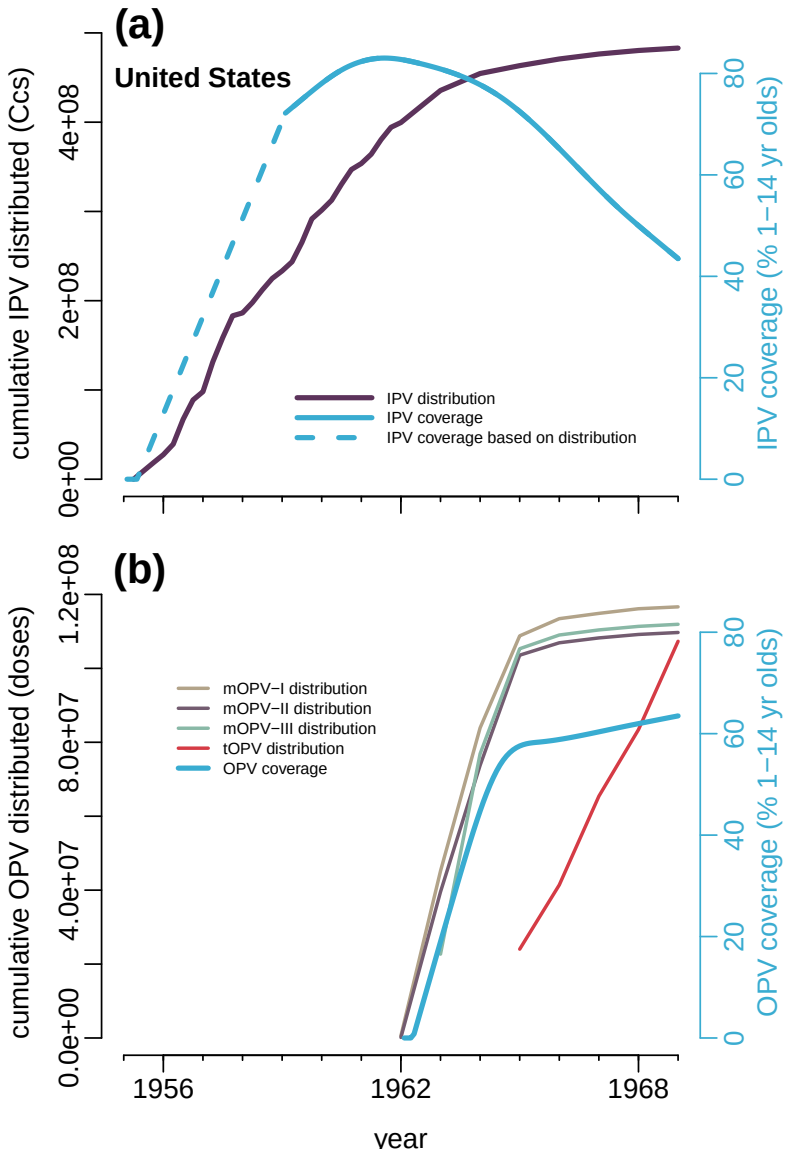


Figure C.2: IPV coverage and distribution in the US. We lacked IPV coverage data from 1955-1958; however, the cumulative distribution of IPV suggests a linear increase in vaccine coverage from this time period. Therefore, we fit a single IPV per capita vaccine uptake rate for the US from April 1955–December 1959.

C.3 Modeling Vaccine Uptake Rate

In order to capture the vaccine coverage in the population of each country/region, we estimated the per capita monthly vaccine uptake rates using the vaccine coverage data. In the pre-vaccine era, the vaccine uptake rate was set to zero and it was estimated beginning with the introduction of IPV/OPV in the US and USSR.

C.3.1 US

In the US, the vaccine uptake rates were tracked as state variable in our process model. The IPV uptake rate was

$$\text{IPV uptake rate}_t = (IPV^1_{uptake} IPV^1_{covar}) + (IPV^2_{uptake} IPV^2_{covar}) - (IPV^3_{uptake} IPV^3_{covar} t), \tag{C.1}$$

where the first term represents the IPV uptake rate from 1955–1959, the second term represents the IPV uptake rate from 1960–1962, and the last term represents the gradual decline of the IPV uptake rate during the OPV era, 1963 onward. The covariates, IPV^i_{cover}, are time-varying and take on the value of either 0 or 1, allowing us to shut off or turn on each component of the uptake rate. Similarly, for the OPV uptake rate in the US, we used the equation:

$$\text{OPV uptake rate}_t = (OPV^1_{uptake} OPV^1_{covar}) - (OPV^2_{uptake} OPV^2_{covar} t). \tag{C.2}$$

However, since individuals with 3+ doses of IPV were also vaccinated with OPV, we had an IPV to OPV uptake rate of:

$$\text{IPV to OPV uptake rate}_t = (\text{IPV to OPV}_{uptake} OPV^1_{covar}) - (OPV^2_{uptake} OPV^2_{covar} t). \tag{C.3}$$

Once again, the covariates OPV^i_{covar} allow us to shut off or turn on each term in the equation. OPV^1_{covar} was turned on at the onset of licensing of all monovalent OPVs, and OPV^2_{covar} allowed us to account for the decline in the OPV uptake rate toward the end of the OPV era. We constrained the uptake rates such that they had a value ≥ 0.

C.3.2 USSR

Modeling vaccine uptake in the USSR was much more simple than the US, because only OPV was in use. The uptake rate for OPV in the USSR was either zero (before OPV was launched) or it was a fixed value estimated form the data. For each region in the USSR, OPV uptake was modeled as:

$$\text{OPV uptake rate}_t = OPV_{uptake} OPV_{covar}, \tag{C.4}$$

where, OPV_{uptake} was estimated from the data and OPV_{covar} was zero before the use of OPV in the USSR, and was 1 after the launch of OPV.

C.4 Transmission Models

To mechanistically model polio epidemiology we utilized Partially Observed Markov Process (POMP) models which are suited for dealing with epidemiological data where the state variables (susceptible, infected, recovered individuals) are not observed in the data; rather the symptomatic infected individuals are partially observed through case reports. For our process models we used seasonally-forced stochastic monthly

discrete-time SIR models where transitions followed a Poisson process. The models had a single recovered class that accounts for life-long immunity. The models contained 6 classes (S_i^B) of infants susceptible to infection but protected from clinical illness by maternally antibodies. The 6 susceptible infant classes contained 0–1 month olds, 1–2 month olds, etc. up to 6 month olds. The models had a single infected class for unvaccinated or under-vaccinated infants (I^B). The older age class contained individuals > 6 months of age, and these individuals had their own unvaccinated/under-vaccinated susceptible (S^O) and infected class (I^O). Infected individuals > 6 months of age, that were no longer protected from disease by maternal antibodies, had a probability of having paralytic polio or non-paralytic polio. Each process model had a measurement model that translated symptomatic (i.e., paralytic and non-paralytic) infections to clinical reported cases. Based on clinical data, we assumed that infections in individuals under 6 months of age are asymptomatic, and only individuals over 6 month of age can be symptomatic and reported as a case.

C.5 Epidemiological Process Model

A similar version of this model was originally validated in our previous study to understand the state-level transmission dynamics of polio [2]. The difference between this model and our previous model is now we explicitly accounted for symptoms in the process model (i.e., paralytic and non-paralytic polio), and we now included IPV and OPV vaccination in the population. In our model, the force of infection, also referred to as the risk of infection, was:

$$\lambda_t = \left(\beta_t \left(\frac{I_{O_t} + I_{B_t} + \zeta_1 I_{IPV_t} + \zeta_2 I_{OPV_t}}{N_t} \right)^{\alpha} + \psi \right) \epsilon_t \tag{C.5}$$

The first term of the force of infection represents transmission that occurs locally by individuals infected in the country/regions at time t. Whereas, the second term,

ψ, encompasses WPV that arises in the population in a way that is divorced from the local infection dynamics. These infections can include immigration of infected individuals from other geographic locations, environmental sources of WPV, and individuals shedding with an infectious period longer than a month. The parameters ζ_1 and ζ_2 are the relative infectiousness of individuals vaccinated with IPV or OPV, respectively. Therefore, if $\zeta = 1$ the vaccine does not reduce infectiousness. The parameter α was used in the model of the US to account for deviations from homogenous mixing due to the board geographic scale. A value of $\alpha \leq 1$ saturates the force of infection as the number of infected individuals in the population increases. For the USSR, we fixed $\alpha = 1$. The transmission parameter β_t was parameterized using a B-spline, giving it the flexibility to have either a constant transmission rate or a seasonal transmission rate.

$$\beta_t = exp \sum_{i=1}^{6} \beta_i \xi_{i_t}, \tag{C.6}$$

where each ξ_{i_t} is a periodic B-spline basis with a 1 year period. Although the transmission rate was able to fluctuate seasonally, there was no inter–annual variation in transmission in our model.The process noise, ϵ_t, was gamma distributed.

$$\epsilon_t \sim \Gamma\left(\frac{1}{\Theta}, \Theta\right), \tag{C.7}$$

where the scale parameter, Θ, of the gamma distribution accounts for both environmental and demographic stochasticity.

$$\Theta = \left(\frac{\beta_{sd_1}}{\sqrt[2]{\beta_t \left(\frac{I_{O_t} + I_{B_t} + \zeta_1 I_{IPV_t} + \zeta_2 I_{OPV_t}}{N_t}\right)^{\alpha} + \psi}} + \beta_{sd_2} \right)^2 \tag{C.8}$$

The first term of Θ represents demographic stochasticity that is modulated by the force of infection, λ_t. Whereas, the second term represents environmental stochasticity. Formulating the process noise in this way gives us some useful properties:

$$\mathbb{E}\left[\epsilon_t\right] = 1, \tag{C.9}$$

$$\mathrm{Var}\left[\epsilon_t\right] = \Theta. \tag{C.10}$$

Thus, the variance can capture purely environmental stochasticity, under the following parameterization:

$$\mathrm{Var}\left[\epsilon_t\right] = \beta_{sd_2}^2, \text{ if } \beta_{sd_1} = 0, \tag{C.11}$$

and, it can represent purely demographic stochasticity under alternate parameterization:

$$\mathrm{Var}\left[\epsilon_t\right] = \frac{\beta_{sd_1}^2}{\beta_t \left(\frac{I_{O_t} + I_{B_t} + \zeta_1 I_{IPV_t} + \zeta_2 I_{OPV_t}}{N_t}\right)^{\alpha} + \psi}, \text{ if } \beta_{sd_2} = 0. \tag{C.12}$$

Due to the computational intensity of this project, we used a discrete time model with a 1 month time step to speed-up simulation. We implemented the model as a poisson process. There is one probability governing the movement of individuals out of susceptible infant classes, p_{B_t}, which is the probability that a susceptible infant remains susceptible,

$$p_{B_t} = e^{-(\delta + \lambda_t)}. \tag{C.13}$$

The parameter δ is the natural death rate. One minus the probability of remaining susceptible is the probability of either being infected or dying. The equations for the first infant class is:

$$S_{B_{1_{t+1}}} = B_{t+1}, \tag{C.14}$$

where B_{t+1} are the births in month $t+1$. Similarly, the other five infant classes are tracked using the following equations:

$$S_{B_{j_{t+1}}} = S_{B_{j-1_t}} p_{B_t};\ \text{for } j \in 2:6 \tag{C.15}$$

The equation for $S_{B_{j_{t+1}}}$ states that infants in the $j-1$ susceptible infant class at time t move into infant class j at time $t+1$ if they are not infected or die of natural death. Similarly, there is one probability governing the movement of individuals out of the susceptible non-infant class, p_{O_t}, which is the probability that a susceptible remains susceptible. One minus the probability of remaining susceptible is the probability of either being infected, vaccinated, or dying.

$$p_{O_t} = e^{-(\delta+\lambda_t+\text{IPV uptake rate}_t+\text{OPV uptake rate}_t)}, \tag{C.16}$$

where "IPV uptake rate$_t$" and "OPV uptake rate$_t$" are monthly per capita vaccine uptake rates that account for vaccine roll-out. See earlier section for calculation of vaccine uptake rates. The equation for the susceptible non-infant age group is:

$$S_{O_{t+1}} = S_{B_{6_t}} p_{B_t} + p_{O_t} S_{O_t} \tag{C.17}$$

The first term of the S_O equation represents the movement of infants from the oldest infant class S_{B_6} to the susceptible non-infant class. Whereas, the second term represents susceptible non-infants from time t remaining susceptible for time $t+1$. For the vaccinated classes, the probability that an individual remains in the IPV class, p_{IPV_t}, is dependent on the force of infection and the OPV uptake rate of IPV-vaccinated individuals, i.e., the "IPV to OPV uptake rate$_t$". Thus,

$$p_{IPV_t} = e^{-(\delta+\zeta_3\lambda_t+\text{IPV to OPV uptake rate}_t)}. \tag{C.18}$$

Here, ζ_3 is the susceptibility to infection of IPV-vaccinated individuals, relative to unvaccinated individuals. The equation for the IPV class is:

$$IPV_{t+1} = S_{O_t}(1 - p_{O_t})\frac{\text{IPV uptake rate}_t}{\delta + \lambda_t + \text{IPV uptake rate}_t + \text{OPV uptake rate}_t} + p_{IPV_t}IPV_t. \tag{C.19}$$

Similarly, the probability that an individual remains in the OPV class is dependent on the force of infection,

$$p_{OPV_t} = e^{-(\delta+\zeta_4\lambda_t)} \tag{C.20}$$

The equation for the OPV class is,

$$\begin{aligned} OPV_{t+1} = S_{O_t}(1 - p_{O_t})\frac{\text{OPV uptake rate}_t}{\delta + \lambda_t + \text{IPV uptake rate}_t + \text{OPV uptake rate}_t} \\ +IPV_t(1 - p_{IPV_t})\frac{\text{IPV to OPV uptake rate}_t}{\delta + \zeta_3\lambda_t + \text{IPV to OPV uptake rate}_t} + p_{OPV_t}OPV_t; \end{aligned} \tag{C.21}$$

As for infections, the infected infant class is tracked using the following equation:

$$I_{B_{t+1}} = \sum_{j=1}^{6} S_{B_{j_t}}(1 - p_{B_t})\frac{\lambda_t}{\delta + \lambda_t}. \tag{C.22}$$

Infants from each of the six susceptible infant classes have a probability of being infected or dying, $1 - p_{B_t}$. The probability that an infant is infected, rather than death is $\frac{\lambda_t}{\delta+\lambda_t}$. The equation for infected non-infants is similar in structure,

$$I_{O_{t+1}} = S_{O_t}(1 - p_{O_t})\frac{\lambda_t}{\delta + \lambda_t + \text{IPV uptake rate}_t + \text{OPV uptake rate}_t}. \tag{C.23}$$

Vaccinated individuals are also able to be infected, according to the reports of paralytic and non-paralytic polio in IPV and OPV vaccinated individuals in the PSU data. The infected IPV class was modeled as:

$$I_{IPV_{t+1}} = IPV_t(1 - p_{IPV_t})\frac{\zeta_3\lambda_t}{\delta + \zeta_3\lambda_t + \text{IPV to OPV uptake rate}_t}. \tag{C.24}$$

Similarly, the infected OPV class is:

$$I_{OPV_{t+1}} = OPV_t(1 - p_{OPV_t})\frac{\zeta_4\lambda_t}{\delta + \zeta_4\lambda_t}. \tag{C.25}$$

Note, infected individuals are infected for exactly 1 month, which is the typical duration of shedding [5]. The infections were then broken down based on symptomatology. The majority of polio infections are asymptomatic, but some will manifest as paralytic or non-paralytic polio. We modeled paralytic polio for the unvaccinated/undervaccinated individuals as,

$$\text{Paralytic}_{U_t} = p_p I_{O_t}; \tag{C.26}$$

and for non-paralytic polio as,

$$\text{NParalytic}_{U_t} = p_{np} I_{O_t}, \tag{C.27}$$

where p_p and p_{np} are the probability of paralytic and non-paralytic polio, respectively, in an unvaccinated infected individual. Paralytic and non-paralytic infections were also tracked for IPV individuals. Paralytic IPV infections are:

$$\text{Paralytic}_{IPV_t} = \zeta_5 p_p I_{IPV_t}, \tag{C.28}$$

while, non-paralytic IPV was,

$$\text{NParalytic}_{IPV_t} = \zeta_6 p_{np} I_{IPV_t}, \tag{C.29}$$

where ζ_5 and ζ_6 are the reduction in the paralytic and non-paralytic probability provided by IPV. Since the PSU reports no cases of non-paralytic polio in OPV

individuals, we only tracked paralytic polio in OPV individuals,

$$\text{Paralytic}_{OPV_t} = \zeta_7 p_p I_{OPV_t}. \tag{C.30}$$

There was one rounding condition imposed in the model of infections. If the number of individuals in I_{B_t}, I_{O_t}, I_{IPV_t}, I_{OPV_t}, paralytic, or non-paralytic infection classes were values between 0 and 1, they were rounded to 0 or 1. This was in order to prevent fractions of infected individuals. Since the MMWR data reports both paralytic and non-paralytic infections, we also had a symptomatic class,

$$\text{Symptomatic}_t = \text{Paralytic}_{U_t} + \text{NParalytic}_{U_t} + \text{Paralytic}_{IPV_t} + \text{NParalytic}_{IPV_t} + \text{Paralytic}_{OPV_t}. \tag{C.31}$$

The NOVS data had the breakdown of cases by paralytic and non-paralytic symptoms, so we also tracked those classes for NOVS,

$$\text{Paralytic}_t = \text{Paralytic}_{U_t} + \text{Paralytic}_{IPV_t} + \text{Paralytic}_{OPV_t}, \tag{C.32}$$

and,

$$\text{NParalytic}_t = \text{NParalytic}_{U_t} + \text{NParalytic}_{IPV_t}. \tag{C.33}$$

For data sets where cases were reported annually (i.e., NOVS and PSU), in the model the cases were accumulated from January–December and reported in December.

C.6 Vaccine Tracker Model

Since the majority of polio infections are asymptomatic, when the vaccines were rolled-out, naturally-immune individuals (i.e., individuals in the recovered class) were undoubtedly vaccinated and counted in the vaccine coverage data. Therefore, in order to estimate the vaccine uptake rates from the vaccine coverage data, we needed to

track vaccine coverage in the population, not just the susceptible individuals. To do so we constructed what we refer to as vaccine tracker classes, these classes tracked the demography of each country/region and are similar to the epidemiological process model states, with the exception that they do not include the infection process. For the US, the vaccine tracker classes tracked the individuals >6 months old and <15 years old; this is because our vaccine data for the US were based on the 1–14 year age group. In contrast, in the USSR, the vaccine tracker classes tracked all individuals >6 months old, because vaccine coverage data in the USSR were for the entire population. In the vaccine tracker model, the probability that an individual remained in the unvaccinated class, p_U, was:

$$p_{U_t} = e^{-(\Lambda+\delta+\text{IPV uptake rate}_t+\text{OPV uptake rate}_t)}; \tag{C.34}$$

where, the new parameter Λ is the rate of aging out of the vaccine tracker class. In the US the rate of aging out was 14 yrs^{-1}, because we were tracking the vaccine coverage in the population under 15. The unvaccinated/under-vaccinated class was tracked as:

$$U_{t+1} = S_{B6_t} p_{B_t} + U_t p_{U_t}. \tag{C.35}$$

The first term represents 6 month olds entering the unvaccinated class after their material antibodies have waned, and the second term represents individuals from time t remaining in the class at time $t+1$. For individuals vaccinated with IPV, the probability of them remaining in the vaccine tracker IPV class was,

$$p_{Salk_t} = e^{-(\Lambda+\delta+\text{IPV to OPV uptake rate}_t)}; \tag{C.36}$$

and the probability of individuals remaining in the OPV or the dual-vaccine IPV & OPV vaccine tracker class was,

$$p_{Sabin_t} = e^{-(\Lambda+\delta)}. \tag{C.37}$$

The IPV vaccine tracker class was,

$$\text{Salk}_{t+1} = U_t(1-p_{U_t})\frac{\text{IPV uptake rate}_t}{\Lambda+\delta+\text{IPV uptake rate}_t+\text{OPV uptake rate}_t} + p_{Salk_t}\text{Salk}_t. \tag{C.38}$$

The OPV vaccine tracker class was,

$$\text{Sabin}_{t+1} = U_t(1-p_{U_t})\frac{\text{OPV uptake rate}_t}{\Lambda+\delta+\text{IPV uptake rate}_t+\text{OPV uptake rate}_t} + p_{Sabin_t}\text{Sabin}_t. \tag{C.39}$$

The final vaccine tracker class contained individuals with dual vaccine coverage with both IPV & OPV,

$$\text{IPVOPV}_{t+1} = \text{Salk}_t(1-p_{Salk_t})\frac{\text{IPV to OPV uptake rate}_t}{\Lambda+\delta+\text{IPV to OPV uptake rate}_t} + p_{Sabin_t}\text{IPVOPV}_t; \tag{C.40}$$

For measuring vaccine coverage, we were interested in the fraction of individuals vaccinated with IPV, OPV, or dual vaccinated. Therefore, we tracked the fraction of individuals in the vaccine tracker classes. The fraction of individuals in the IPV class was,

$$\text{IPV}_{frac_t} = \frac{\text{Salk}_t+\text{IPVOPV}_t}{U_t+\text{Salk}_t+\text{Sabin}_t+\text{IPVOPV}_t}. \tag{C.41}$$

The OPV fraction was,

$$\text{OPV}_{frac_t} = \frac{\text{Sabin}_t+\text{IPVOPV}_t}{U_t+\text{Salk}_t+\text{Sabin}_t+\text{IPVOPV}_t}; \tag{C.42}$$

the dual vaccine fraction was,

$$\text{IPVOPV}_{frac_t} = \frac{\text{IPVOPV}_t}{U_t + \text{Salk}_t + \text{Sabin}_t + \text{IPVOPV}_t}; \tag{C.43}$$

and the unvaccinated/under-vaccinated fraction was,

$$\text{U}_{frac_t} = \frac{U_t}{U_t + \text{Salk}_t + \text{Sabin}_t + \text{IPVOPV}_t}. \tag{C.44}$$

C.6.1 Observation Models

In order to model the stochastic process of symptomatic infected individuals being reported as a case, we drew cases from a normal distribution. In the US, Lithuania, and Estonia symptomatic cases were reported monthly. In the RSFSR and Ukraine paralytic cases were reported monthly. In addition to the monthly cases reported in the US in the MMWR, non-paralytic and paralytic cases were reported annually in NOVS and PSU. To represent observations in each data set, the reporting of the appropriate state variable X_t^i was drawn from a normal distribution with a report rate ρ_t^i and dispersion parameter τ^i,

$$\text{cases}_t \sim \text{normal}(\rho_t^i X_t^i, \tau^i X_t^i), \tag{C.45}$$

with the following condition imposed,

$$\text{cases}_t = \begin{cases} \text{round}(\text{cases}_t), & \text{if cases}_t \geq 0 \\ 0, & \text{otherwise} \end{cases} \tag{C.46}$$

For calculating the likelihood for each case data set, we used a binned-normal probability density and used the lower tail of the cumulative density function to discretize the density.

If $\text{cases}_t > 0$,

$$\text{Likelihood}_t^i = \text{pnormal}(\text{cases}_t + 0.5, \rho_t^i X_t^i, \tau^i X_t^i) - \text{pnormal}(\text{cases}_t - 0.5, \rho_t^i X_t^i, \tau^i X_t^i) \tag{C.47}$$

If $\text{cases}_t = 0$,

$$\text{Likelihood}_t^i = \text{pnormal}(\text{cases}_t + 0.5, \rho_t^i X_t^i, \tau^i X_t^i). \tag{C.48}$$

For all data sets, with the exception of the US MMWR data, the report rate, ρ_t^i, was constant. For the MMWR data, there was a report rate for the pre-baby boom era in the US (May 1934–Dec 1945), when mostly paralytic polio was reported; and there was a higher report rate in the baby boom era (1946 onward), with the increase in report rate occurring in January 1946. The elevated report rate accounted for the reporting of non-paralytic polio cases that began sometime in 1945. The parameter τ^i captures the variation in the process of observing infections via reported cases. In addition to the observation models for disease, we also had observation models for vaccine coverage.

To represent vaccine coverage in each country/region, the reporting of the appropriate vaccine tracker state variable V_t^i was drawn from a normal distribution with a dispersion parameter τ^i,

$$\text{coverage}_t \sim \text{normal}(100V_t^i, \tau^i V_t^i). \tag{C.49}$$

The vaccine tracker state variables, V_t^i, are fractions, but we were fitting to vaccine coverage as the percent of the population; therefore we multiplied by 100. We also

imposed the following condition,

$$\text{coverage}_t = \begin{cases} \text{round}(\text{coverage}_t), & \text{if } 0 \leq \text{coverage}_t \leq 100 \\ 0, & \text{if coverage}_t < 0 \\ 100, & \text{if coverage}_t > 100 \end{cases} \tag{C.50}$$

For calculating the likelihood for each vaccine coverage data set, we used a binned-normal probability density.

If $0 < \text{coverage}_t < 100$, we used the lower tail of the cumulative density function to discretize the density,

$$\text{Likelihood}_t^i = \text{pnormal}(\text{coverage}_t + 0.5, \rho_t^i V_t^i, \tau^i V_t^i) - \text{pnormal}(\text{coverage}_t - 0.5, \rho_t^i V_t^i, \tau^i V_t^i) \tag{C.51}$$

If $\text{coverage}_t = 0$, we once again used the lower tail of the cumulative density function to discretize the density,

$$\text{Likelihood}_t^i = \text{pnormal}(\text{coverage}_t + 0.5, \rho_t^i V_t^i, \tau^i V_t^i). \tag{C.52}$$

In contrast, if $\text{coverage}_t = 100$, we used the upper tail of the cumulative density function to discretize the density,

$$\text{Likelihood}_t^i = \text{pnormal}(\text{coverage}_t - 0.5, \rho_t^i V_t^i, \tau^i V_t^i). \tag{C.53}$$

Because we were fitting the data to multiple data sets, the likelihood was taken as the product of the likelihood of each individual data set. Therefore the joint likelihood

was,

$$\text{Likelihood}_t = \left(\prod_{i=1}^{n} \text{Likelihood}_t^i \right) + 1e^{-18}, \tag{C.54}$$

where n is the appropriate number of data sets for that country/region. The scaler $1e^{-18}$ was used for practical purposes to put a lower bound on the likelihood in order to prevent arithmetic underflow. In the US, we fit the vaccine uptake rates to the four time series of vaccine coverage (IPV, OPV, dual vaccinated, and unvaccinated/under-vaccinated). In the US we also fit to the two NOVS time series and the MMWR time series. The state variables used to report to NOVS were the paralytic and non-paralytic infections. Whereas, the symptomatic infections were reported in the US MMWR, and in Lithuania, and Estonia. The paralytic infections were reported in the RSFSR and Ukraine.

C.7 Pre-Vaccine Era Statistical Inference

We fit our SIR models to the data using maximization by iterated particle filtering (mif) in the R package pomp [6]. mif is a simulation-based likelihood method for parameter estimation. The basis of mif is particle filtering, also known as Sequential Monte Carlo, which integrates state variables of a stochastic system and estimates the likelihood for a fixed parameter set. Instead of fixing parameters, mif varies the parameters throughout the filtering process using a random walk and selectively propagates particles (parameter sets) that have high likelihood. By initializing mif throughout parameter space, we identified regions of high likelihood and identified the maximum likelihood parameter estimates. We estimated parameters independently for the US and the four regions of the USSR. The inference conducted in this study was an extension to model fitting that was done in a previous study of WPV transmission in the US [2]. In our baseline study, in order to get a handle on WPV epidemiological parameters in the pre-vaccine era, mif was initialized from 1 million parameter sets

for a global search, followed by additional phases of increasingly localized searches, which included profiling.

To estimate WPV epidemiological parameters for the US, inference was done using the data from May 1934 to December 1954, the pre-vaccine era. Inference was initiated in May of the year following the first full year of available data on polio cases and births. This was done in order to construct the infant initial conditions directly from the birth data. The month of May is the tail-end of the polio off-season. Thus, the 0–6 month old infants in May were born between November and April, when polio transmission was low and they were unlikely to be infected. Based on this, we assumed that the initial number of susceptible infants in the 6th infant class was the number of individuals born in November of the previous year, individuals born in December were in the 5th infant class, January births were in the the 4th infant class, February births in the 3rd infant class, March births in the 2nd infant class, and April births in the the 1st infant class. We set the initial number of infected infants to zero. Constructing the infant initial conditions prevented us from having to estimate an addition seven parameters. We used a similar procedure for the USSR. We initialized inference in January of the first year of available data. Once again, we assumed all infants were susceptible. Unlike the US where we had monthly birth data, for the USSR we had birth data on the annual basis. Therefore, we had to assume births were distributed evenly throughout the year in the USSR. We know that births are in fact seasonal, as shown in the US birth data in the main text. However, without monthly birth data for the USSR we had no way to constructing the birth seasonality in the USSR.

C.7.1 WPV Global Search

In our original study of WPV transmission [2], to initialize a global search of parameter space, 1 million parameter combinations, with initial conditions reflecting

a low incidence month, were generated to cover a broad range of parameter values. For each parameter set, ten replicate particle filters were run, each with 2000 particles, to estimate the likelihood of the data from the state of Illinois. Although the state of Illinois was used to initially search the global space for WPV epidemiological parameters, the US national-level data and the USSR regional-level data were used to find the country/region specific MLEs. Illinois was chosen in our previous study because of the high incidence of polio in Illinois and because Illinois had an increase in incidence following the World War II baby boom, which was a characteristic feature of many US states. The 1 million likelihoods were used to provide a rough estimate of the global shape of the WPV likelihood surface and identify regions of parameter space with high likelihood. Of the 1 million parameter sets, the 50000 with the highest likelihoods were used to initialize local searches of parameter space using mif. It is important to note that the particle filter gives an unbiased estimate of the likelihood, but some extra algebra needs to be done to get the estimates of the log-likelihood from replicate particle filters. We calculated the unbiased estimate of the log-likelihood using the following equations,

$$\log\vec{\mathcal{L}}(\Theta) = \vec{\mathbf{x}}. \tag{C.55}$$

Given the vector of log-likelihoods, $\vec{\mathbf{x}}$, produced from n replicate particle filters using the parameter set Θ, we find the mean:

$$y = \text{mean}(\vec{\mathbf{x}}), \tag{C.56}$$

then we use the mean to estimate of the log-likelihood:

$$\log\mathcal{L} = y + \log(\text{mean}(e^{\vec{\mathbf{x}}-y})); \tag{C.57}$$

and we can also calculate the standard error of the estimate of the log-likelihood:

$$\log\mathcal{L}_{\text{se}} = \frac{\text{sd}(e^{\vec{\mathbf{x}}-y})}{\sqrt{n}} \frac{1}{e^{\log\mathcal{L}-y}}, \tag{C.58}$$

where sd is the standard deviation.

C.7.2 Local Searches of Parameter Space

Each of the 50000 "best" likelihood parameter sets obtained from the Global search were used to initiate local searches of WPV transmission parameter space. mif was used for local searches of parameter space. In general, for each run of mif, we used a random walk standard deviation of 0.02, 45+ iterations of mif, 2000+ particles, hyperbolic cooling with a variance factor of 2, and a cooling fraction of 0.5. Monte Carlo replicate mif runs were also used in local searches. All searches of parameter space were done on the log or logit scale. The log scale was used to enforce parameter values >0 and the logit scale was used to bound parameters, such as the report rate between 0 and 1. See Equations below for expit and logit functions.

$$\text{logit}(p) = \log\left(\frac{p}{1-p}\right) \tag{C.59}$$

$$\text{expit}(x) = \frac{1}{1+e^{-x}} \tag{C.60}$$

It is important to note that the random walk standard deviations for some parameters were time-varying, such as for the pre-baby boom report rate for the US MMWR data. This is because, if the parameter was only relevant to a particular time period (e.g., the pre-baby boom), then that parameter was only allowed to vary via the random walk during the time points for which the parameter was relevant. This was done in order to align the inference procedure with the data containing information to guide the estimation of each parameter. In order to define time-varying random walk

standard deviations, the mif2 method was used in the developer version of the package pomp. As parameters passed through our multiphase inference pipeline, the number of parameter sets expanded and contracted. This is because the use of Monte Carol mif replicates expanded the parameter sets, but we would also cull parameter sets to remove all parameter combinations that had a likelihood >> 20 log-likelihood units below the current maximum.

C.7.3 US Profiling Report Rate and Immigration

In the US, we identified that the MMWR report rate of infections (i.e., not broken down by symptoms, I_{O_t}), in the pre-baby boom era, ρ^{mmwr}, to be important; along with the immigration into the force of infection, ψ. Therefore, we decided to profile ρ^{mmwr} and ψ. The profile values spanned the ranges [0.001, 0.1] and [6e-6, 2.5e-3], for ρ^{mmwr} and ψ, respectively. There were 37 values of ρ^{mmwr}, $(\rho_1, \rho_2, ..., \rho_{37})$, with 0.001 interval sampling in the range of [0.001,0.03] and 0.01 interval sampling in the range of [0.04, 0.1]. There were 50 values of ψ, $(\psi_1, \psi_2, ..., \psi_{50})$ evenly sampled across the range. In order to make the profiles two dimensional, we took every pairwise combination, which resulted in 1850 profile points, each being a unique combination. Monte Carlo replicate mif runs were used to maximize the likelihood along the other parameter and initial condition dimensions. We found the highest likelihood parameter set at each of the 1850 profile points and profiling was repeated to refine the parameter estimates. Fig C.3 shows the resulting likelihood profile for the MMWR report rate during the pre-baby boom and the the immigration term. The MLE from this profile was used as the MLE for the WPV epidemiological parameters in the US. Therefore, these MLEs remained fixed from here out in order to do inference on symptomatology and vaccine efficacy in the US.

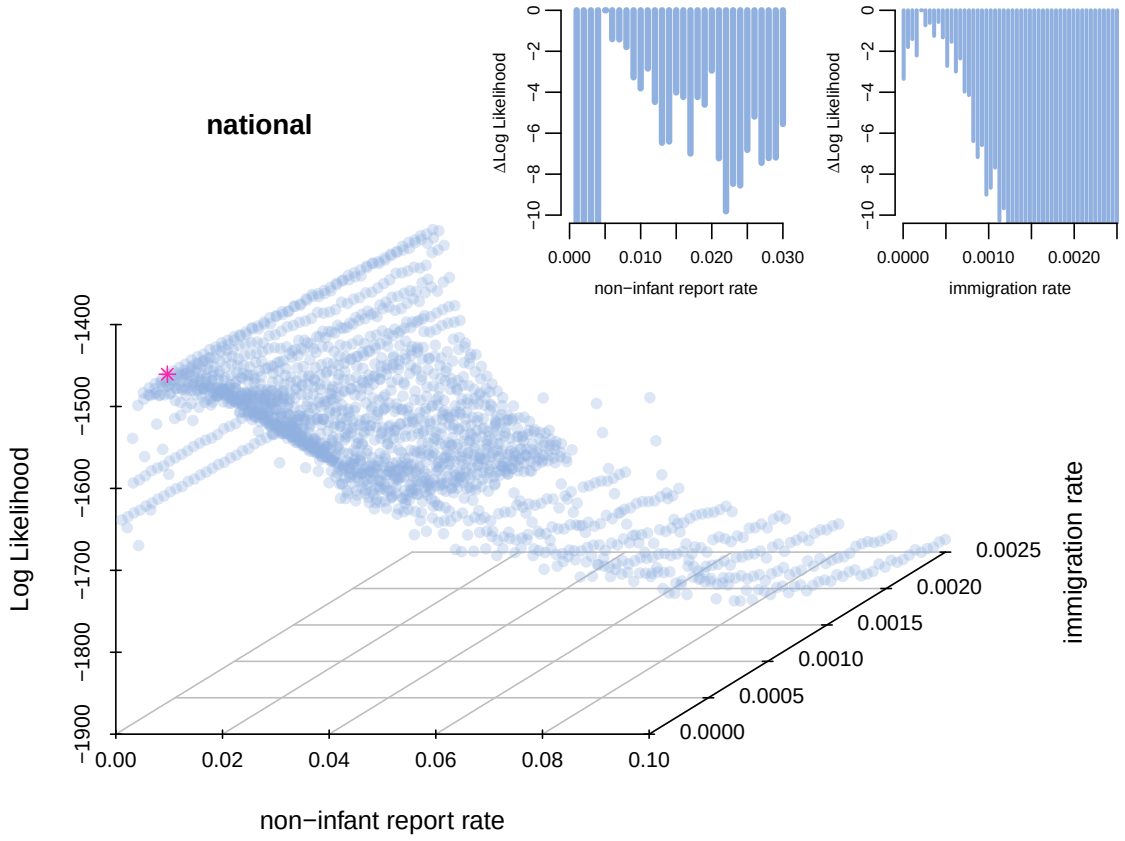


Figure C.3: **US national MMWR data likelihood profile.** Profile contains the pre-baby boom non-infant report rate (ρ^{mmwr}) and the immigration rate (ψ). MLE indicated by pink asterisk.

C.7.4 US Profiling Paralytic and Non-paralytic Probabilities.

Using the monthly MMWR data from May 1934–December 1954 and the NOVS data from 1951–1954, we conducted 2-dimensional profiling of the probability of paralytic symptoms, p_p, and the probability of non-paralytic symptoms, p_{np}, in infected unvaccinated individuals. The profile values ranged from 0.001–0.04 for the paralytic probability and 0.008–0.15 for the non-paralytic probability. For pairwise profile combinations, we used the constraint that the the combined probability of paralytic or non-paralytic polio upon infection was $\geq \rho^{mmwr}$. This is because ρ^{mmwr} was the combined probability of symptoms and reporting. Therefore, the probability of either a paralytic or non-paralytic infection (i.e., the probability of symptoms) must be greater than or equal to the joint probability of symptoms and subsequent reporting.

Since we fixed all epidemiological parameters based on the profile of ρ^{mmwr} and ψ, we estimated 7 parameters while profiling p_p and p_{np}. The new parameters estimated were the report rate of symptomatic infections (now infections are being distinguished by symptoms) in the MMWR data (during the pre-baby boom and baby boom eras), the associated dispersion parameter, the report rate of paralytic polio in NOVS, the reporting of non-paralytic polio in NOVS, and the associated dispersion parameters for NOVS reports. The NOVS parameters had non-zero random walks from January 1951 onward, and zero otherwise. mif2 was initiated from each of the 22451 profile points for the paralytic probability versus the non-paralytic probability. We used Monte Carlo replicate runs for each profile point. Each mif2 run had 60 iterations, 8000 particles, a variance factor of 2, hyperbolic cooling, and a cooling fraction of 0.5. We took the highest likelihood parameter set from each profile point and initiated the profiles multiple times.

The profiles of the paralytic and non-paralytic probability were fairly flat across a large region of parameter space. We discovered that this was due to a tradeoff between the probability of symptoms and the probability of reporting to MMWR and

NOVS. The combined probability of symptoms and reporting was well identified (see Fig C.4). This tradeoff results in a "ridge" in the likelihood surface (see Fig C.5) along 6 dimensions of parameter space (i.e., paralytic probability, non-paralytic probability, pre-baby boom reporting to the MMWR, the increase in MMWR reporting in the baby boom, reporting of paralytic cases to NOVS, and reporting of non-paralytic polio to NOVS). In order to deal with this tradeoff, we assumed that the report rate of paralytic polio was 60–80% as taken from the literature [1], and we then profiled intensely within the region of parameter space that fit this assumption. We specifically profiled along the likelihood ridges because we knew they were areas of high likelihood. We then drew the MLE from this final profile. The profile shown in Fig C.6 is the combination of the top parameter sets from all of the profiles done on the paralytic and non-paralytic probability. The densely profiled area in the forward left corner is the profile region for which the report rate for paralytic polio was 60–80%.

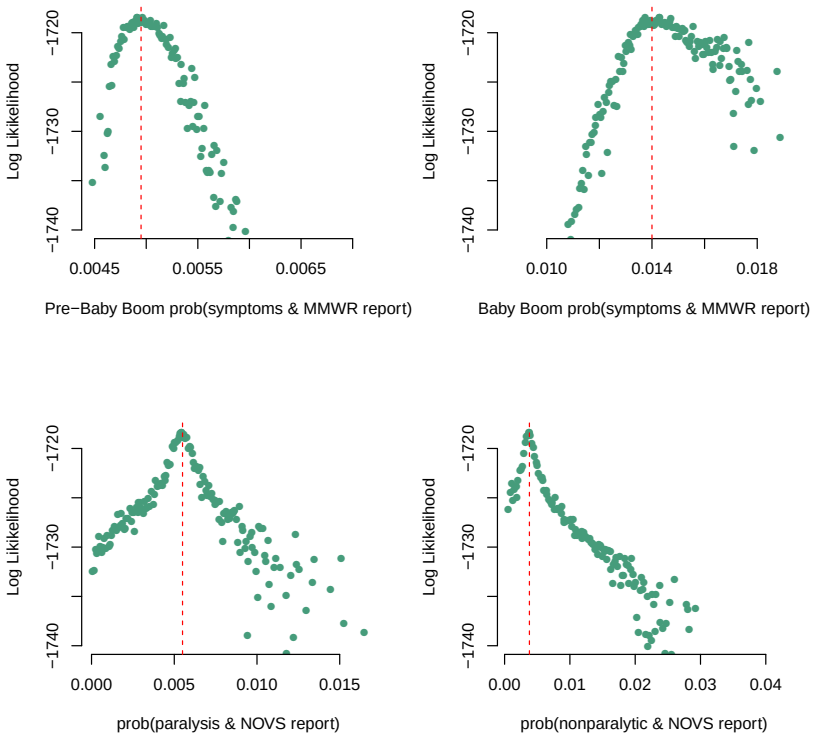


Figure C.4: **Symptoms and Subsequent Reporting.** In the US pre-vaccine era, the combined probability of symptoms and subsequent reporting to MMWR and NOVS were well identified, but the individual probabilities of symptoms versus reporting traded off with one another. Shown here are quasi-profiles.

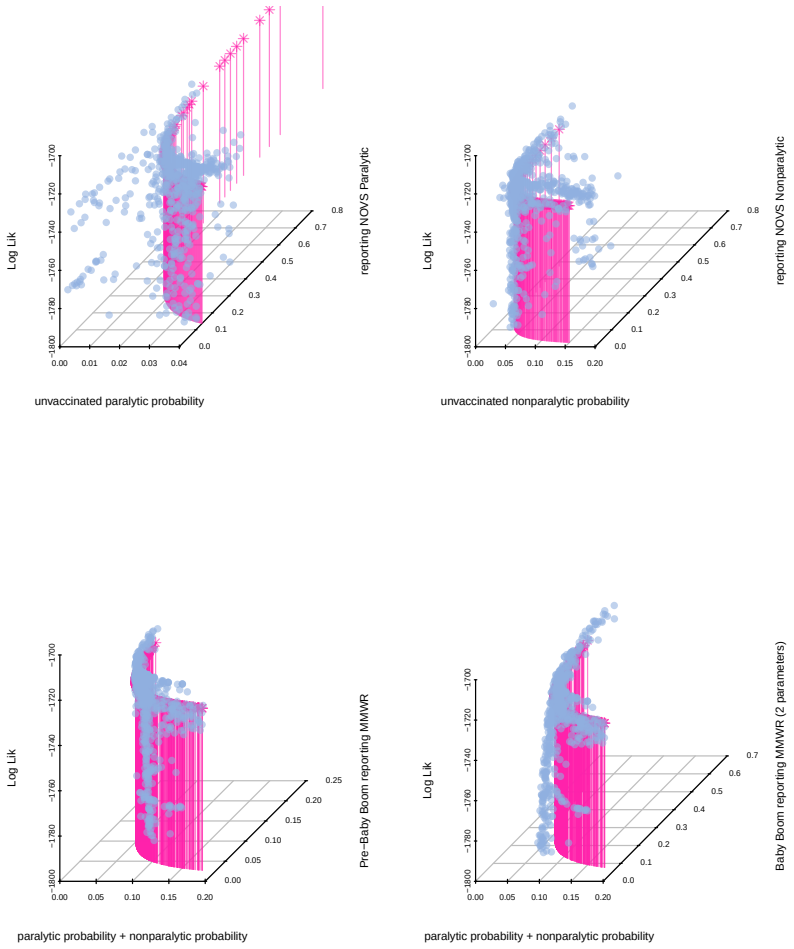


Figure C.5: In the US pre-vaccine era, the combined probability of symptoms and reporting to MMWR and NOVS were well identified. However, there are 6 parameters that combined to results in the 4 well-identified symptomatic and reporting probabilities in the previous figure. The tradeoffs between these parameters result in a "ridge" along 6 dimensions of the likelihood surface. Components of this ridge are shown here. The parameters sets from the quasi-profiles in the previous figure are plotted and fall along the ridges.

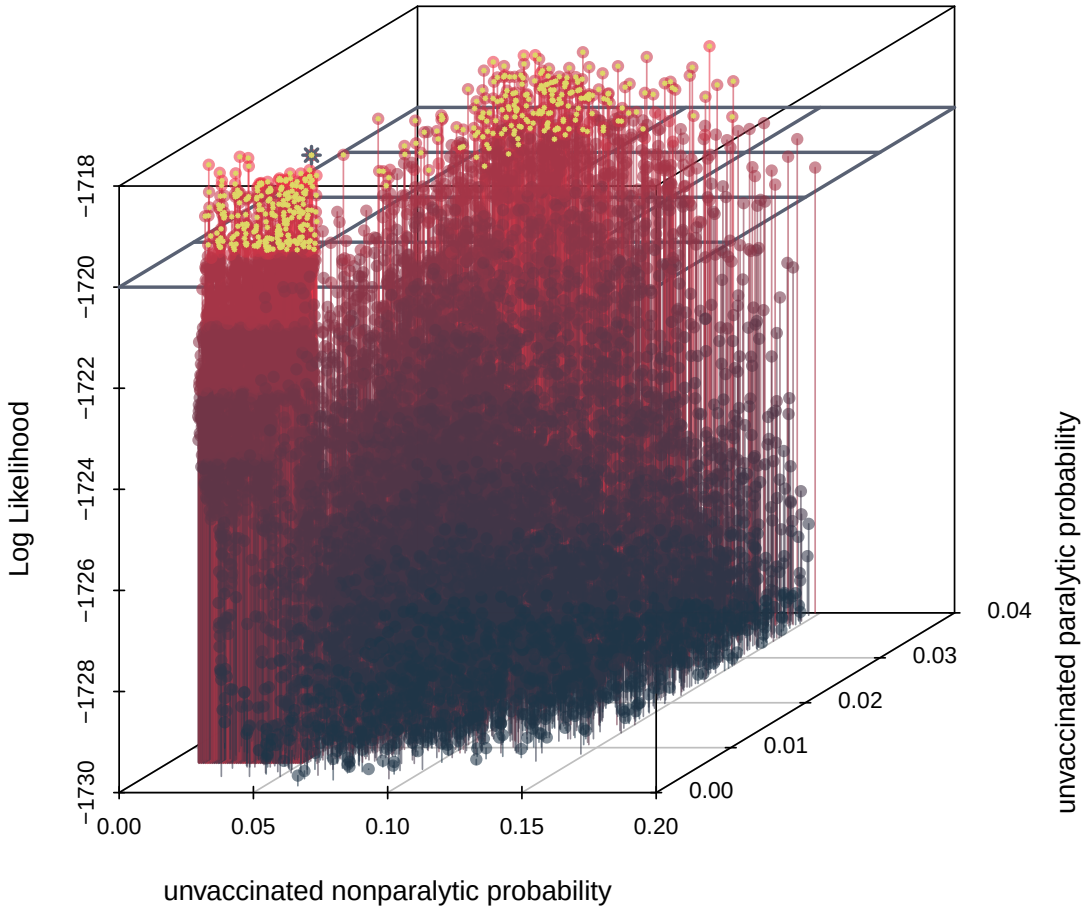


Figure C.6: Multi-phase profiles of the unvaccinated paralytic probability and non-paralytic probability. The blue asterisk indicates the MLE, the yellow dots indicate parameter sets within 2 log-likelihood units of the MLE, demonstrating that a broad range of parameter combinations could explain these dynamics due to the tradeoffs between the probability of symptoms and the probability of reporting. Therefore, we did high resolution profiling in the region with the reporting of paralytic polio to MMWR between 60–80% and the MLE was drawn from there. The high resolution profiling can be seen in the forward left corner of the profile.

C.7.5 USSR Pre-Vaccine Era Statistical Inference

Using MLEs from each US state provided in [2], we initialized mif2 for the pre-vaccine era in each of the USSR regions. The data from Estonia and Lithuania from January 1955–December 1958 were used to estimate the epidemiological parameters. The data from the RSFSR and Ukraine did not begin until January 1957. Due to the similarity in the epidemiological dynamics in the US and the USSR, and the limited size of the USSR data sets, we were quickly able to identify parameter sets that captured the USSR pre-vaccine dynamics. We did not do any pre-vaccine era profiling in the USSR or convergence diagnostics. However, no tradeoffs in the pre-vaccine parameters were observed when fitting this model to US states. Therefore, we assumed our pre-vaccine MLEs for the USSR are in the neighborhood of the true MLEs. We decided to wait to do pre-vaccine profiling in USSR until we have extended our pre-vaccine data for the USSR.

C.8 Vaccine Efficacy Statistical Inference

C.8.1 US Vaccine Phase 1: Sobol Design

Importantly, all epidemiological parameters (i.e., transmission parameters, MMWR and NOVS report rates, symptomatic probabilities, noise terms, etc.) were fixed based on the pre-vaccine era MLEs. In order to get a broad overview of the shape of the likelihood surface under various parameter combinations of vaccine efficacy, a Sobol design with > 4 million parameter sets was generated. The Sobol design included both IPV and OPV efficacy parameters. The likelihood of the full vaccine era model was estimated for each parameter set. The full likelihood was a joint likelihood based on the MMWR, NOVS, PSU, IPV coverage, and OPV coverage data. The parameter sets within 100 log likelihood units of the top parameter set, were carried forward into Phase 2 (there were 2914 parameter sets carried forward). The parameters used in

the Sobol design were: the IPV and OPV parameters measuring efficacy and effects on pathology, ζ_{1-7}, as well as the two reporting (i.e., paralytic and non-paralytic) and two dispersion parameters for PSU. We used rough estimates of the vaccine uptake rates for calculating the likelihoods in the Sobol design.

C.8.2 US Vaccine Phase 2: MIF

We were unable to reliably run mif using the full vaccine era likelihood model because the joint likelihood calculation for all 12 time series resulted in numerics below that computationally possible. Therefore we ran what we will refer to as a mif cascade, which involved using the same process model but fitting non-overlapping parameters to various data set combinations sequentially. We (1) improved the fit of the vaccine uptake rates using the vaccine coverage data, (2) fit the IPV efficacy parameters using the MMWR and NOVS data, and (3) fit the report rate and the reporting standard deviation of the paralytic cases to PSU using the paralytic cases in the PSU data. We continued to hit the limits of floating point number representation when incorporating the PSU data. When hitting the floating point number representation limit, the likelihood returned was the tolerance level, 1e-18, in C.54. Without a way to calculate the PSU likelihoods reliably, we decided not to particle filter using the PSU data. Therefore, we did all subsequent fitting using the MMWR and NOVS data and we focused our attention on the estimation of IPV efficacy in the US. By exploring the results from our Sobol design and running mif to estimate IPV efficacy parameters, we identified a tradeoff between estimates of IPV's reduction in susceptibility to infection, ζ_3, and IPV's reduction in infectiousness, ζ_1. Fig C.7 shows high likelihood parameter sets resulting from the Sobol design and the tradeoff between susceptibility and infectiousness. This tradeoff persisted after running mif using the MMWR and NOVS data, as shown in the main text.

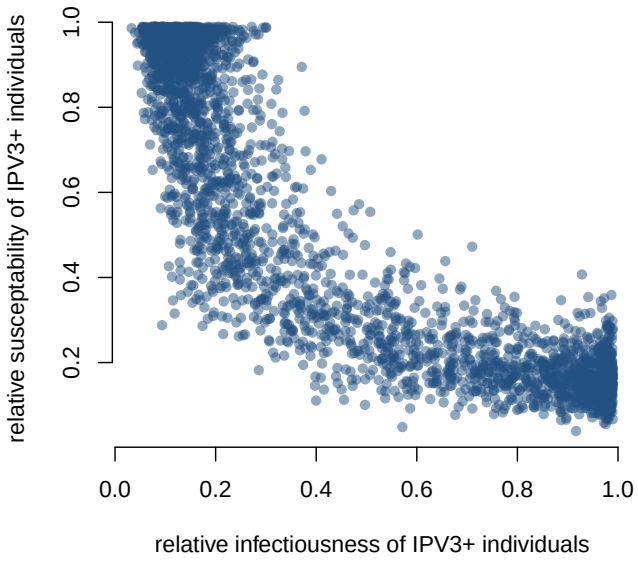


Figure C.7: **Tradeoff observed between IPV efficacy parameters.** The relative infectiousness of IPV-vaccinated individuals, relative to unvaccinated individuals, and the relative susceptibility to infection of IPV individuals, as compared to fully susceptible individuals.

C.8.3 US Vaccine Phase 3: Profiling IPV Efficacy

Due to the tradeoff between IPV efficacy parameters, we decided to fix the susceptibility to infection in IPV individuals. We assumed that IPV individuals are equally susceptible to infection as unvaccinated individuals, an assumption supported by clinical data [7]. Under this assumption, we profiled the reduction in infectiousness in IPV individuals, ζ_1. The profile results are shown in the main text. We profiled ζ_1 in the range [0.1,0.7] by units of 0.01. For all parameters with the exception of ζ_1, we used the highest likelihood parameter combination from Phase 2 at each profile point, as well as the highest likelihood parameter sets from Phase 2 that fell within each profile slice, to initialize the profiling. Each profile parameter set was run with 25 Monte Carlo replicates, each with 60 mif2 iterations, and 2000 particles.

C.8.4 USSR Vaccine Efficacy Statistical Inference

We fit parameters relating OPV efficacy and the reduction in the paralytic probability to the data from the RSFSR. We did 2-dimensional profiling of the reduction in infectiousness, ζ_2, and the reduction in the paralytic probability, ζ_7. Both ζ_2 and ζ_7 were profiled in the range of [0.001,0.999] in a 50x50 profile. All epidemiological and reporting parameters were fixed at the MLEs estimated using the pre-vaccine data. The parameters maximized during profiling were the OPV uptake rate and the dispersion parameter for reporting the OPV coverage. We also did 2-dimensional profiling of the reduction in susceptibility, ζ_4, and the reduction in the paralytic probability, ζ_7. ζ_4 was profiled in the range of [0.001,0.5] and ζ_7 was profiled in the range of [0.001,0.999] in a 50x50 profile. All epidemiological and reporting parameters were fixed at the MLEs estimated using the pre-vaccine data. Once again, the parameters maximized during profiling were the OPV uptake rate and the dispersion parameter for reporting the OPV coverage. We used the MLE OPV uptake rate from the 2-dimensional profiles and constructed a 3-dimensional 50x50x50 profile of reduction in infectiousness, reduction in susceptibility, and reduction in paralytic probability. Particle filters were run to calculate the likelihood at each of the 125000 profile points and the MLE was taken from the 3-dimensional profile. For independent validation of the OPV parameters, we then took the MLEs of ζ_2, ζ_4, and ζ_7 from the RSFSR and used them to predict the vaccine era cases in Ukraine, Lithuania, and Estonia. Before doing independent validation, we fit the OPV uptake rate and the OPV coverage dispersion parameter for Ukraine, Lithuania, and Estonia.

Bibliography

[1] Trevelyan B, Smallman-Raynor M, Cliff AD (2005) The Spatial Dynamics of Poliomyelitis in the United States: From Epidemic Emergence to Vaccine-Induced Retreat, 1910-1971. Annals of the Association of American Geographers 95: 269–293.

[2] Martinez-Bakker M, King AA, Rohani P (2015) Unraveling the Transmission Ecology of Polio. PLoS Biology (in press).

[3] United States Centers for Disease Control and Prevention CDC U.S. Births. URL `http://www.cdc.gov/nchs/products/vsus.htm`.

[4] United States Census Bureau Census Bureau U.S. Population Estimates.

[5] Alexander JP, Gary HE, Pallansch MA (1997) Duration of Poliovirus Excretion and its Implications for Acute Flaccid Paralysis Surveillance: A Review of the Literature. The Journal of Infectious Diseases 175, Suppl: S176–S182.

[6] King AA, Nguyen D, Ionides E (2015) Statistical Inference for Partially Observed Markov Processes via the R Package pomp. Journal of Statistical Software (in press).

[7] Hird TR, Grassly NC (2012) Systematic Review of Mucosal Immunity Induced by Oral and Inactivated Poliovirus Vaccines Against Virus Shedding Following Oral Poliovirus Challenge. PLoS Pathogens 8: e1002599.