Human birth seasonality: latitudinal gradient and interplay with childhood disease dynamics

Micaela Martinez-Bakker1,2,†, Kevin M. Bakker1,2,†, Aaron A. King1,2,3,4 and Pejman Rohani1,2,4

1Department of Ecology and Evolutionary Biology, 2Center for the Study of Complex Systems, and 3Department of Mathematics, University of Michigan, Ann Arbor, MI 48109, USA 4Fogarty International Center, National Institutes of Health, Bethesda, MD 20892, USA

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 USA, and reveal pronounced seasonality in birth rates, with geographical 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 autumn peaks, a pattern we also observed throughout the Northern Hemisphere. Additionally, the amplitude of United States birth seasonality was more than twofold 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 susceptible-exposed-infected-recovered 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 owing 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.

1. Introduction

The ubiquity of seasonal variation in the incidence of infectious diseases has driven much epidemiological research focused on understanding the responsible underlying mechanisms [1–5]. Surprisingly, there remains much uncertainty regarding the drivers of seasonal incidence for numerous infections including polio, pertussis, scarlet fever, diphtheria, rotavirus, among others [5–8]. Early work on diphtheria and measles implicated elevated contact rates among children in school as the driver of pulsed transmission [1,9], 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 possibility that has been appreciated in studies of wildlife diseases [11,20–23]. For instance, studying cowpox virus in voles, Begon et al. [24] found that susceptible recruitment is seasonal, and higher breeding-season birth rates delayed epidemic peaks. However, despite evidence demonstrating the importance of host demography in recurrent epidemics [25–28], 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 [32–36]. Early studies of vital statistics in various United States 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 identified that northern and southern states have differences in their seasonal birth amplitude [32–35,37,38] and birth/conception minima [32,38–40]. Studies of births in Africa and Asia have been sparse, but seasonal peak–trough differences in conception ranging from 11% to 64% have been documented in Africa and 8–58% in Asia [36]. To date, there has been no long-term, large-scale, spatio-temporal analysis of births in either the USA or worldwide.

We have compiled the most extensive spatio-temporal dataset 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 demographical dataset 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 United States along with over 200 additional time series from countries spanning the Northern Hemisphere. We have analysed these data in combination with a transmission model and statistical inference tools to examine the dynamic implications of birth seasonality on childhood infection.

2. Data

Monthly state-level time series of live births from 1931 to 2008 were downloaded from United States Vital Statistics [41] and digitized. Annual state-level population size data were collected from the United States 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 United States 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 state in the USA in each era and for each country in the global dataset. The significance of each period was tested by comparing the power of each period against a noise background, using a 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 per cent of the annual mean (electronic supplementary material, equations S1–S5). Interannual variation was examined by analysing the per cent change in mean birth rates from one year to the next.

3. Measles models

We used a discrete-time susceptible-exposed-infected-recovered (SEIR) model of measles adapted from Earn et al. [25], with school-term forcing based on the England and Wales school year (electronic supplementary material, table S1; [11]). We incorporated seasonality in births using a sine function with varying amplitude and phase. The equations describing the model (electronic supplementary material, equations S6–S13) and parameter values (electronic supplementary material, table S2) are provided in the electronic supplementary material.

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 (electronic supplementary material, equations S6–S14), three time series were generated assuming the following parametrizations: a birth peak day of either 162, 295 or 351 and a 28% birth amplitude (see the electronic supplementary material, 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 the basic reproductive ratio, $R_0$. The transmission rate was profiled, and the likelihood was calculated using a particle filter (electronic supplementary material, Materials and methods) [45].

In order to test for bias in parameter estimation using real-world data, we used 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 three, six or nine months (see the electronic supplementary material for methods). We used both maximum-likelihood via iterated filtering [45] and the time-series susceptible-infected-recovered (TSIR) methodology of Finkenstadt & Grenfell [46,47] to quantify the impact of seasonal births on parameter estimates (figure 5).

4. The seasonal timing of births

Figure 1 and the electronic supplementary material, figure S3 provide an overview of birth rates in the USA. Most states had significant seasonal (1 year) birth pulses in the Pre-Baby Boom Era, whereas all states showed significant birth seasonality in the Baby Boom and Modern Eras. Of the 210
time series analysed outside the USA, 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 USA and throughout the Northern Hemisphere (figures 2 and 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 five 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 USA 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 (figure 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 (greater than 50° 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 Tajikistan, 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 occurring in Saint Vincent and the Grenadines during the period 1992–2005.

In addition to the annual birth peaks, in the USA, a significant biannual (six month) birth pulse was detected in 24 states. In the Pre-Baby Boom Era, all states with biannual periodicity were clustered in the lower midwest, deep south and southeast (figure 2 and electronic supplementary material, figures S1 and S2). In the Baby Boom Era, only 13 states continued to exhibit a biannual period. Arkansas is the only state where this biannual birth pulse persists in the Modern Era.

5. 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 per cent of the mean annual birth rate (electronic supplementary material, Materials and methods). As with the seasonal timing, in the USA, the amplitude of birth seasonality displays a latitudinal gradient. Figure 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 ($p < 4.6 \times 10^{-5}$). However, the amount of variation in birth amplitude explained by latitude decreased through time, perhaps owing to the decline in birth amplitude throughout southern states during the Modern Era (figure 2 and electronic supplementary material, figure S4). We did not observe a latitudinal gradient in birth amplitude outside of the USA (electronic supplementary material, figure S5).

As expected, the mean combined amplitude across all states was found to be comparable with 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, owing to the geographical variation in birth peak timing, state-level births are out of phase. Thus, aggregated United States birth data have a deceptively low amplitude that is not reflective of individual states. Birth amplitudes more than 15% were observed in many southern states throughout the time series (electronic supplementary material, figure S4).

Prior work has demonstrated that the levels of interannual variation in births observed in the USA 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 interannual variation. Interannual variation was measured for each state as the per cent change in mean birth rate from one year to the next. We found that in almost every instance, seasonal variation exceeded interannual variation, with seasonal variation in the Modern Era being two to three times larger than the variation from year to year (electronic supplementary material, figure S6).

6. The effects of birth seasonality on epidemic dynamics

We investigated the impact of birth seasonality on epidemics of childhood disease by using models of measles transmission. As shown in figure 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, then the epidemic is amplified owing to the availability of susceptibles. By contrast, if the peak in susceptible recruitment occurs at the end of the school year, when children are entering summer break, then the epidemic is dampened (figure 4a).

Independent of the timing of the birth peak, the effect of birth seasonality on measles epidemics depends on the birth amplitude (figure 4b and electronic supplementary material, figure 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 (more than 40%) can alter incidence to such an extent that they can drive dynamical transitions (figure 4b).

Statistical inference on simulated data led to small biases in the estimate of $R_0$ for measles (figure 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 overestimating $R_0$ in order to capture the elevated epidemic year incidence. By contrast, when the birth peak was set to either early June (day 162) or late October (day 295), times at which susceptible recruitment dampens epidemic year incidence and elevates skip-year incidence, we underestimated $R_0$. However, the bias in $R_0$ was small, 0.4–1.3%.

We found that models with seasonal births effectively capture measles dynamics in New York City (figure 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 (figure 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 to 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

Figure 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 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 colour-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 approximately 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 reference [49]. (c) Bias in $R_0$ estimates owing to the exclusion of birth seasonality in an SEIR model. Time series were generated using an SEIR model with 28% birth amplitude and 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 (MLE) of $R_0$ is either over- or underestimated when birth seasonality is excluded from the model. 95% confidence intervals for MLE are indicated on profiles. (Online version in colour.)
whelmed by uncertainty in parameter estimates (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 over-

within the standard error of the MLE are represented by horizontal lines. (c) Legend applies to all of figure 5. (d) With a three month lag (blue), seasonal births with a six month lag (green), seasonal births with a nine month lag (yellow) and births with no seasonality

Figure 5. 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 three month lag (blue), seasonal births with a six month lag (green), seasonal births with a nine month lag (yellow) and births with no seasonality (maroon). Legend applies to all of figure 5. (b) The shape of the likelihood surface with respect to \( R_0 \). The MLE \( R_0 \)s are indicated by points, and the values within the standard error of the MLE 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 over-

whelmed by uncertainty in parameter estimates (b,c). No difference in transmission parameters was observed using the TSIR method. (Online version in colour.)

New York City can be captured by the model without birth seasonality (figure 5a).

7. 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 USA for the entirety of our data, and was reflective of a much broader geographical 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 USA and 6–35% in other Northern Hemispheric countries. Along with the latitudinal gradient in peak birth timing, in the USA, we also observed a latitudinal gradient in birth amplitude. States in the southern United States 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 the USA, suggesting this may either be a localized phenomenon or strongly correlated with social, economic and/or cultural factors in the USA.

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 biannual peaks across the lower midwest, deep south and southeastern United States in the Pre-Baby Boom Era. This biannual pulse was lost over time, with only Arkansas exhibiting biannual periodicity in the Modern Era. Biannual fluctuations in births have been documented in previous studies [38], but our data suggest that biannual birth pulses in the USA 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 [38,50–56]. 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 analyses focused solely on the USA 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 neighbours 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 [11,25]. 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 [7,11,58], 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 (electronic supplementary material, 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 probably 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 southeast Asia. Clearly, these infant immunization campaigns will be inefficient if they occur prior to the birth pulse. Thus, it is our hope that future studies aimed at mitigating childhood diseases will use 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 and the Rohani/King laboratories for their helpful comments on this work.

Funding statement. M.M.B. is supported by the NSF Graduate Research Fellowship Programme. P.R. and A.A.K. are supported by the Research and Policy in Infectious Disease Dynamics programme 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).

References

29. Forrest J, Miller-Rushing AJ. 2010 Toward a


