

# Seasonality and comparative dynamics of six childhood infections in pre-vaccination Copenhagen

C. Jessica E. Metcalf<sup>1,2,\*</sup>, Ottar N. Bjørnstad<sup>1,3</sup>, Bryan T. Grenfell<sup>1,2,3</sup>  
and Viggo Andreasen<sup>3,4</sup>

<sup>1</sup>*Centre for Infectious Disease Dynamics, The Pennsylvania State University, 208 Mueller Laboratory, University Park, PA 16802, USA*

<sup>2</sup>*Department of Ecology and Evolutionary Biology, Woodrow Wilson School, Eno Hall, Princeton University, Princeton, NJ 08544-2016, USA*

<sup>3</sup>*Fogarty International Center, National Institutes of Health, Bethesda, MD 20892, USA*

<sup>4</sup>*Department of Science, Roskilde University, 4000 Roskilde, Denmark*

Seasonal variation in infection transmission is a key determinant of epidemic dynamics of acute infections. For measles, the best-understood strongly immunizing directly transmitted childhood infection, the perception is that term-time forcing is the main driver of seasonality in developed countries. The degree to which this holds true across other acute immunizing childhood infections is not clear. Here, we identify seasonal transmission patterns using a unique long-term dataset with weekly incidence of six infections including measles. Data on age–incidence allow us to quantify the mean age of infection. Results indicate correspondence between dips in transmission and school holidays for some infections, but there are puzzling discrepancies, despite close correspondence between average age of infection and age of schooling. Theoretical predictions of the relationship between amplitude of seasonality and basic reproductive rate of infections that should result from term-time forcing are also not upheld. We conclude that where yearly trajectories of susceptible numbers are perturbed, e.g. via waning of immunity, seasonality is unlikely to be entirely driven by term-time forcing. For the three bacterial infections, pertussis, scarlet fever and diphtheria, there is additionally a strong increase in transmission during the late summer before the end of school vacations.

**Keywords:** measles; susceptible reconstruction; time-series susceptible–infected–recovered

## 1. INTRODUCTION

Many infectious diseases show seasonal patterns of incidence (Altizer *et al.* 2006; Grassly & Fraser 2006). Mechanisms underlying seasonality are perhaps best understood for vector or aquatically transmitted infections (Altizer *et al.* 2006). For directly transmitted respiratory infections, the cause of seasonality may be either environmental or social. Environmental forcing may operate through enhanced persistence or dispersal of infectious particles by certain conditions of temperature or humidity (Lowen *et al.* 2007), or environmental conditions may affect the condition of the host in ways that favour the pathogen (Adams & Hewison 2007). Social forcing operates via enhanced transmission through aggregation of individuals at certain times of year (Schenzle 1984; Fine & Clarkson 1986; Bjørnstad *et al.* 2002; Ferrari *et al.* 2008). Identifying the drivers of seasonal variation in disease transmission rates is important, since seasonal variation is a strong force in shaping disease dynamics (Earn *et al.* 2000). For strongly

immunizing childhood infections in developed countries, a large body of work on measles has pointed to the effects of aggregation of school children on disease transmission rates as a key driver of seasonality and dynamics (London & Yorke 1973; Schenzle 1984; Bjørnstad *et al.* 2002). Recently, Keeling & Rohani (2008) showed that since strongly immunizing diseases lead to a specific age profile of cases, if seasonality is predominantly due to term-time forcing, the expected magnitude of seasonality of an infection and resulting dynamics and periodicity can be predicted from the proportion of cases expected to occur during schooling age. However, the empirical importance of term-time forcing for directly transmitted childhood infections other than measles has not been fully quantified.

Here we explore these issues by comparing across multiple childhood infections in a longterm dataset from a single population from the city of Copenhagen, spanning from 30 to 60 years starting at the turn of the twentieth century. This unique dataset is of particular value for exploring the importance of term-time forcing, as vaccination was initiated only more recently, and all time series are therefore pre-vaccination. Furthermore, the age distribution of infectious individuals is available at an annual scale allowing us to estimate the average age at infection for each disease, and the proportion of infected individuals who are school aged. This gives us a more direct opportunity to understand the importance of term-time forcing in generating the observed

\* Author and address for correspondence: Department of Ecology and Evolutionary Biology, Woodrow Wilson School, Eno Hall, Princeton University, Princeton, NJ 08544-2016, USA (cjm53@psu.edu).

Electronic supplementary material is available at <http://dx.doi.org/10.1098/rspb.2009.1058> or via <http://rspb.royalsocietypublishing.org>.

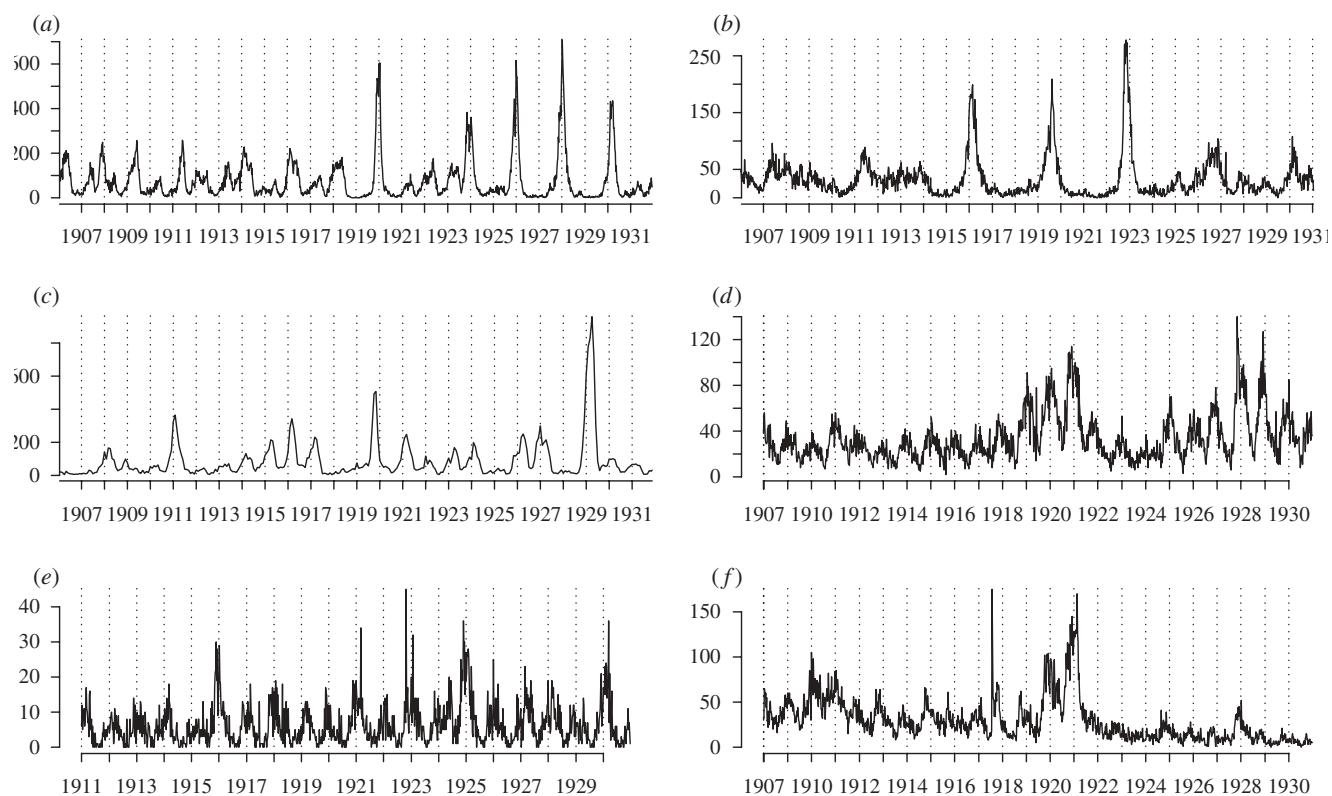


Figure 1. Detail of numbers of cases reported between 1907 (1911 for varicella) and 1930; reports are weekly for all infections except mumps, where reports are monthly. See electronic supplementary material for full time series. (a) Measles; (b) pertussis; (c) mumps; (d) diphtheria; (e) varicella; (f) scarlet fever.

seasonality as quantified using a time-series susceptible–infected–recovered (TSIR) model.

The dataset includes notification time series for three viral diseases, measles, mumps, and varicella (chicken pox), and three bacterial diseases, pertussis, diphtheria and scarlet fever, the latter two caused by infection with group A *Streptococcus* bacteria. All six infections are transmitted through direct contact or large droplet spread, and the incidence of infection is highest in young children. However, there is considerable variation in a number of life-history features across this guild of infections, including the generation time (serial interval) of the infection (approximately the latent plus infectious period; see Kenah *et al.* (2008) for more exact relations), which is around 10–14 days for measles (Bjørnstad *et al.* 2002; Grenfell *et al.* 2002), 16–26 days for mumps, 18–23 days for varicella, approximately 27 days for pertussis, 10–14 days for scarlet fever and around 26 days for diphtheria (Anderson & May 1991). There is also variation in transmission rates: measles and whooping cough are very highly transmissible and prevalent in young children; mumps, scarlet fever and diphtheria are less so and tend to be seen in children and teenagers (Anderson & May 1991). The degree of immunity conferred by the infection also varies—measles is strongly immunizing, but diphtheria (Galazka 2000) and pertussis (Rohani *et al.* 1999) show waning of immunity. A final life-history difference is that pertussis, diphtheria (Frost *et al.* 1936) and scarlet fever (Pichichero & Casey 2003; Feeney *et al.* 2005) are all known to feature asymptomatic individuals. Beyond direct life-history differences, all six infections may exhibit variable dynamics in space and time. For example, the combination of a short generation

time, high transmission rate and strong immunization usually binds measles dynamics strongly to annual or biennial attractors (Earn *et al.* 2000). By contrast, the longer generation time of pertussis and important stochastic transients lead to longer period multi-annual cycles (Rohani *et al.* 1999). Generally 3–4 year cycles are observed (Broutin *et al.* 2005). The dynamics of the remaining infections are less well characterized in the literature. We observe more or less irregular annual outbreaks in the historic Copenhagen data for these other infections (figures S2 and S3, electronic supplementary material).

In this paper, we take a comparative approach to understanding seasonal patterns of transmission and look for evidence of deviation from term-time forcing among these childhood infections. We apply a TSIR model (Bjørnstad *et al.* 2002) to estimate month-specific transmission rates, and we apply logistic regressions of cumulative age incidence curves to estimate mean age of infection and proportion of cases in school-age individuals. We then compare predictions of rankings of the relative magnitude of seasonality based on a model of term-time forcing with the observed patterns. Below we introduce the data, then outline the TSIR model and the approach used to estimate mean age of infection. Our results show intriguing deviations of seasonal transmission from term-time forcing, particularly for the bacterial infections.

## 2. MATERIAL AND METHODS

### (a) Data

We have two sources of data. First, we have counts of the numbers of reported infected individuals at a weekly scale for varying ranges of years for the different infections (figure 1

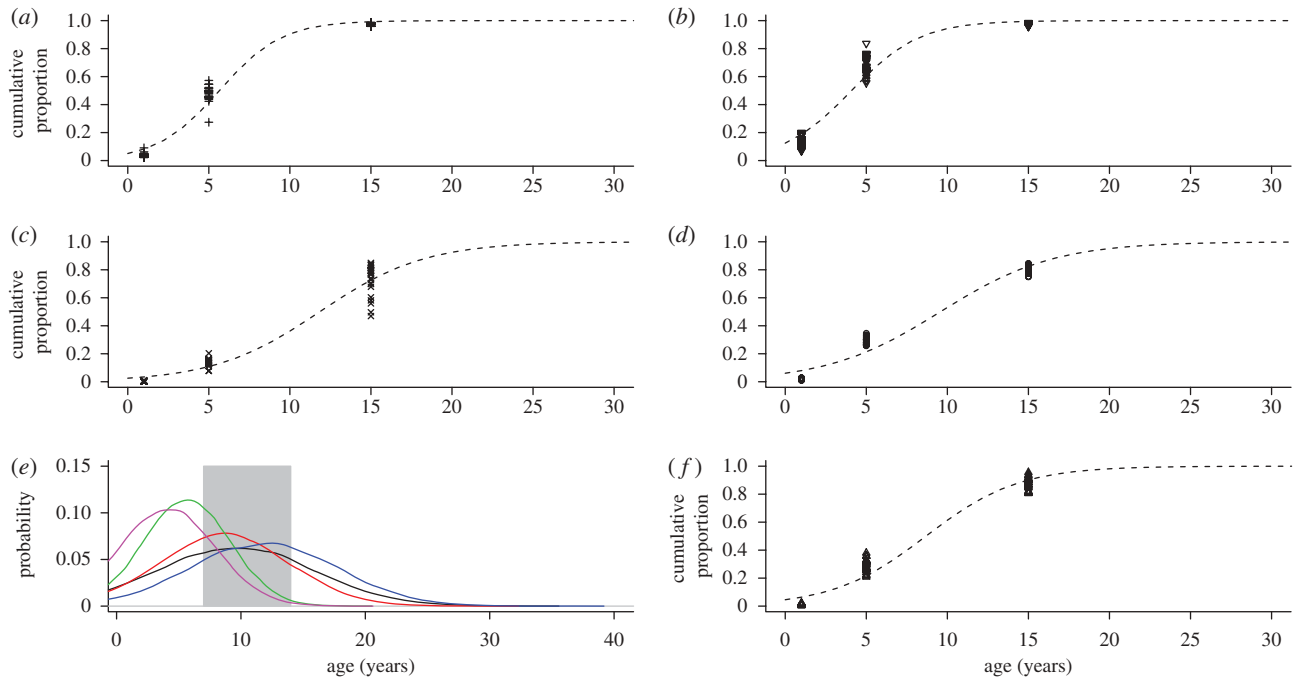


Figure 2. The cumulative proportion of reported cases at each age obtained by summing the data over age and dividing by the total sum to obtain proportions for (a) measles; (b) pertussis; (c) mumps; (d) diphtheria; (f) scarlet fever. Different points for the same age correspond to different years. Fitted logistic regressions are shown as a dashed line. No data are available for varicella. (e) Inferred probability density functions for different infections (black line, diphtheria; brown line, scarlet fever; green line, measles; blue line, mumps; pink line, pertussis). The grey area indicates ages between 7 and 15 years and approximately encapsulate the years of schooling. The area under each curve found between these lines is provided in table 1.

and figures S2 and S3 in the electronic supplementary material) from a system of weekly reporting of cases of epidemic infections by primary-care physicians in Copenhagen in place since 1900 (Stadslægen 1900–1970; see electronic supplementary material for further details). Second, we have the age category of the reported infected individuals at a yearly scale, in age ranges of less than 1, 1–5, 6–15, 16–65 and greater than 65 (figure 2), obtained from Anonymous (1900–1969) taken for years between 1907 and 1931. See *Andreasen et al.* (2008) and the electronic supplementary material for more details on data compilation.

### (b) The TSIR model

For sufficiently immunizing infections, if reporting rates are stable through time and all individuals eventually succumb to infection, numbers of infected individuals track births (*Finkenstadt & Grenfell 2000*) according to the difference equation

$$S_{t+1} = S_t + B_t - \frac{I_t^{(r)}}{\rho}, \quad (2.1)$$

where the time step is the generation time,  $S_t$  is the number of susceptible individuals at time  $t$ ,  $B_t$  is the number of births,  $I_t^{(r)}$  is the number of infected individuals reported and  $\rho$  reflects under-reporting. If we ignore observational uncertainty, the actual number of infected individuals at time  $t$ ,  $I_t$ , is  $I_t = I_t^{(r)}/\rho$ . Rearranging equation (2.1) results in a relationship from which the rate of reporting,  $\rho$ , and the pattern of change in the number of susceptible individuals can be inferred. First, from equation (2.1),

$$S_t = \bar{S}N_t + D_0 + \sum_{k=0}^{t-1} B_k - \sum_{k=0}^{t-1} \frac{I_k^{(r)}}{\rho}, \quad (2.2)$$

where  $\bar{S}$  is the average proportion of all individuals  $N_t$  that are susceptible and  $D_0$  is the unknown deviation around the average number of susceptible individuals at the beginning of the time series. To both estimate  $\rho$  and reconstruct a time series  $D_t$ , which defines how numbers of susceptible individuals vary around the mean number of susceptible individuals,  $S_t = \bar{S}N_t + D_t$ , we write (*Bjørnstad et al. 2002*; *Finkenstadt et al. 2002*)

$$\sum_{k=0}^{t-1} B_k = \left(\frac{1}{\rho}\right) \sum_{k=0}^{t-1} I_k^{(r)} + D_t - D_0. \quad (2.3)$$

From this,  $D_t$  can be estimated as the residuals from the (potentially locally varying) regression of the cumulative number of births on the cumulative number of cases and  $\rho$  is the inverse of the slope of this regression (*Finkenstadt et al. 2002*). Temporal variation in the reporting rate  $\rho$  during the time period considered can be modelled by allowing local variation in the regression using a smoothing spline (*Bjørnstad et al. 2002*). Independent estimation of  $\bar{S}$  is not possible, as  $\bar{S}$  and  $D_0$  are both subsumed into the intercept of the regression. For all infections, new births will only become available for infection after a delay set by the length of maternal immunity. Accordingly, maternal immunity period was used to offset recorded births relative to infection dynamics when using equation (2.3) to estimate  $D_t$ .

Given  $\rho$  and  $D_t$ , estimators for other epidemiological parameters including the seasonal transmission factor ( $\beta_s$ ) can be constructed. The epidemic process is described by

$$E[I_{t+1}] = \frac{\beta_s I_t^\alpha S_t}{N_t}, \quad (2.4)$$

where  $E[\ ]$  represents the expectation for the number of infected individuals one infection generation time in the

future, the exponent  $\alpha$ , usually a little less than 1, captures heterogeneities in mixing not directly modelled by the seasonality (Finkenstadt & Grenfell 2000; Bjørnstad *et al.* 2002) and the effects of discretization of the underlying continuous time process (Glass *et al.* 2003). By log-transforming equation (2.4), and using the relationships defined above, we obtain

$$\log(E[(I_{t+1})]) = \log(\beta_s) + \alpha \log(I_t) + \log(\bar{S} N_t + D_t) - \log(N_t). \tag{2.5}$$

Given  $I_t$ ,  $N_t$  and  $D_t$ , we use regression to estimate  $\beta_s$  and  $\alpha$  and marginal profile likelihoods to estimate  $\bar{S}$  (Bjørnstad *et al.* 2002). In our initial analysis of seasonality,  $\alpha$  was constrained to  $\alpha = 1$  to facilitate comparison of patterns of seasonality between the different diseases. For simplicity, the seasonal transmission rate was estimated as the linear regression without an intercept of log incidence in generation  $t + 1$  on month-specific  $\beta_s$  as a factor and log-incidence, log-susceptibles and minus log population size as offsets. More exact estimation of equation (2.4) would use a generalized linear model with Gaussian error and a log-link but results are robust to error distribution and link used, so we retained the linear regression. In as far as the time step used represents the true average time from infection to recovery, the estimates of  $\beta_s$  correspond to the seasonally varying basic reproductive number  $R_0$  (the expected number of new infections caused by a single infected individual in an entirely susceptible population) and  $\beta$  represents the time-averaged  $R_0$ . Note that the transmission rate  $\beta_s$  estimated in this way is a proxy for perhaps many complicated mixing patterns (in particular, heterogeneity across age) (Bjørnstad *et al.* 2002).

For each infection, data were aggregated into the time step chosen as reflective of the generation time of the infection, taken as either two weeks or four weeks based on the value closest to the reported sum of infectious plus latent periods (table 1). Altering the generation time used for different infections did not alter the qualitative conclusions regarding seasonality (see electronic supplementary material figures S8 and S9); we also verified that estimation was not affected by reasonable levels of waning of immunity using simulated data. Finally, although the main aim is to analyse seasonal patterns, we also tested the ability of the TSIR parameters to recapture long-term dynamics of the different infections (see electronic supplementary material).

**(c) Age at infection**

We estimated the mean and variance in the age of infection through aggregating the age-specific data across years and fitting a logistic regression to the resulting cumulative proportions  $y$  at each age  $x$  (figure 2), assuming  $y \sim \text{Binomial}(\sum I, c(x))$ , where  $c(x) = \exp(b_0 + b_s x) / (1 + \exp(b_0 + b_s x))$ . This gives a mean age of infection of  $-b_0/b_s$  and a variance of  $\pi^2 / (3b_s^2)$ .

The thus estimated mean age of infection provides a complementary approach to our TSIR approach to exploring the relative magnitude of  $R_0$  across infections. This is useful because the TSIR-based estimate can be confounded with estimates of the proportion of susceptible individuals (Bjørnstad *et al.* 2002). Assuming negligible mortality trajectories in the disease-relevant age classes and a constant force of infection, in a growing population, the mean age of infection  $A$  is related to  $R_0$  by  $R_0 \simeq \mu^{-1}/A$ ,

Table 1. Infection characteristics used for and estimated by the TSIR model and estimated by analysis of age data. (The approximate generation time of each infection sets the discrete interval (in weeks); the maternal immunity period sets the offset of recorded births relative to infection dynamics (also in weeks),  $\rho$  is the reporting rate estimated via susceptible reconstruction (standard error (s.e.) in brackets),  $\bar{\beta}$  provides the median value of the seasonal transmission rate  $\beta$  (per generation time) obtained via the TSIR (figure 3) (maximum s.e. in brackets) and  $\bar{S}$  is the estimated mean proportion of susceptible individuals (s.e. estimated via Fisher Information. Note that profile likelihoods are very right-skewed for some infections so that this quadratic approximation can be relatively poor). We also provide the intercept and slope of the fitted logistic regression (s.e. in brackets), the corresponding average age at infection,  $A$  (in years), and estimate of  $R_0$ ; as well as the estimated proportion of cases occurring in children at the start of school age (older than 6 but less than 7  $p_{6-7}$ ) and children during schooling years ( $p_{7-15}$ ). Age information is not available for varicella.)

	gen. time	mat. imm	$\rho$	$\bar{\beta}$	$\bar{S}$	intercept	slope	$A$	$R_0$	$p_{6-7}$	$p_{7-15}$
measles	2	16	0.40 (0.05)	8.44 (0.03)	0.12 (0.03)	-2.93 (0.76)	0.52 (0.15)	5.65	8.25	0.109	0.401
mumps	4	12	0.12 (0.04)	5.36 (0.06)	0.28 (0.06)	-3.65 (0.79)	0.31 (0.07)	11.84	3.94	0.044	0.575
varicella	2	12	0.02 (0.005)	16.59 (0.10)	0.05 (0.05)						
pertussis	4	12	0.17 (0.11)	10.06 (0.03)	0.09 (0.03)	-1.95 (0.51)	0.48 (0.12)	4.09	11.40	0.085	0.261
diphtheria	4	12	0.16 (0.04)	2.32 (0.05)	0.26 (1.01)	-2.74 (0.58)	0.29 (0.06)	9.57	4.87	0.056	0.530
scarlet fever	2	52	0.14 (0.04)	1.91 (0.05)	0.54 (1.2)	-3.04 (0.62)	0.35 (0.07)	8.68	5.37	0.070	0.585

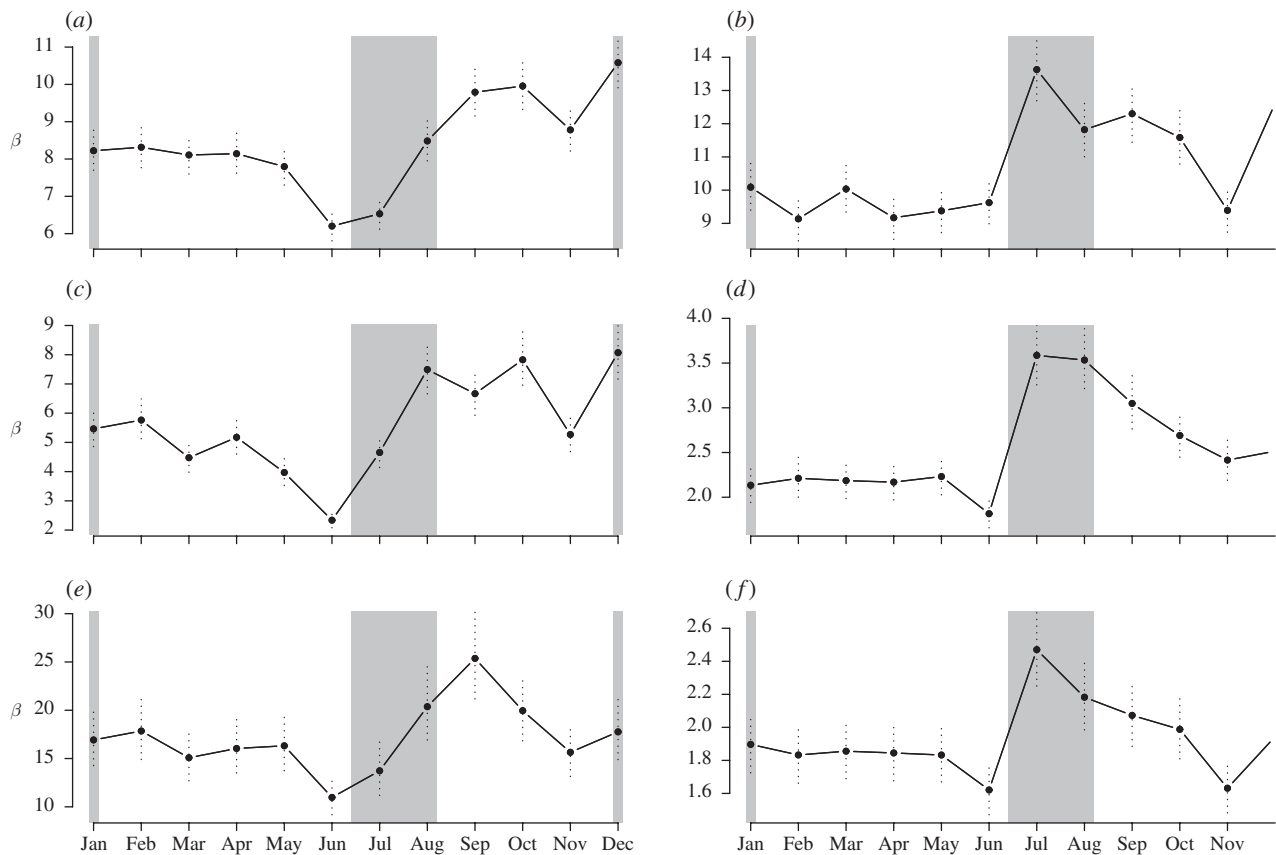


Figure 3. The seasonal transmission coefficient,  $\beta$  with confidence intervals ( $\pm 1.96$  s.e.) shown as dotted vertical lines; other parameters as in table 1. The grey rectangles show months during which the major school holidays occur. (a) Measles; (b) pertussis; (c) mumps; (d) diphtheria; (e) varicella; (f) scarlet fever.

where  $\mu$  is the *per capita* birth rate (Anderson & May 1991). We set  $\mu = 0.021$ , i.e.  $\mu^{-1} = 47.6$ , based on the mean from historical data on Copenhagen birth rates. This method of estimation of  $R_0$  may be biased by a range of potential complications and in particular, heterogeneity of transmission across age (Farrington *et al.* 2001), but will provide some indication of relative magnitude of  $R_0$  across different infections.

#### (d) Theoretical predictions of seasonality

If seasonality is driven by term-time forcing, the amplitude of seasonality is likely to be proportional to the ratio of mixing of children of the same age within school, relative to random mixing outside of schools. The basic susceptible–infected–recovered model predicts that this ratio will depend on  $R_0$  and the rate of population turnover  $\mu$  (reflecting birth and death rates), as these two parameters determine the equilibrium proportions of susceptible and infected individuals across age (Anderson & May 1991; Keeling & Rohani 2008). If we assume negligible mortality in the disease-relevant age classes and negligible change in the age structure of infected individuals over time, and transmission is homogenous across age and space, the proportion of susceptible individuals,  $X$ , changes over age according to

$$\frac{dX}{da} = -\frac{X(a)}{A}, \quad (2.6)$$

where  $A$  is the average age of infection (Anderson & May 1991). Solving this yields  $X(a) = \exp(-\mu(R_0 - 1)a)$ . The

fraction of infected individuals of age  $a$ ,  $Y(a)$ , is then

$$Y(a) = \frac{\mu(R_0 - 1)}{\gamma} \exp^{-\mu(R_0 - 1)a}, \quad (2.7)$$

assuming  $dY(a)/da \sim 0$ , and with  $Y(a) = 1/\gamma dX/da$ , where  $\gamma$  is the rate of recovery from infection. Using these results, we can then estimate the ratio of mixing of children of the same age within school relative to random mixing (Keeling *et al.* 2001; Keeling & Rohani 2008), according to

$$\begin{aligned} \phi &= \frac{\int_{A_S}^{A_L} X(a)Y(a)da}{\int X(a)da \int Y(a)da} \\ &= \mu R_0 (\exp(-2\mu R_0 A_S) - \exp(-2\mu R_0 A_L)), \end{aligned} \quad (2.8)$$

where  $A_S$  is the age of starting school (7 years in Copenhagen; see electronic supplementary material) and  $A_L$  is the age of leaving school (here taken as 15 years; see electronic supplementary material; using 14 does not change the results) and the population turnover rate  $\mu$  can be estimated from the yearly *per capita* birth rate, which ranges between 0.014 and 0.028 over the time course of the study in Copenhagen. If seasonality is chiefly driven by term-time forcing, the magnitude of seasonality across infections (as measured by their variance in  $\beta_s$  across the year) should be positively correlated with the  $\phi$  quantity.

## 3. RESULTS

### (a) Dynamics

Two-yearly plots obtained from the weekly incidence data for all infections are shown in figure S3 in the electronic

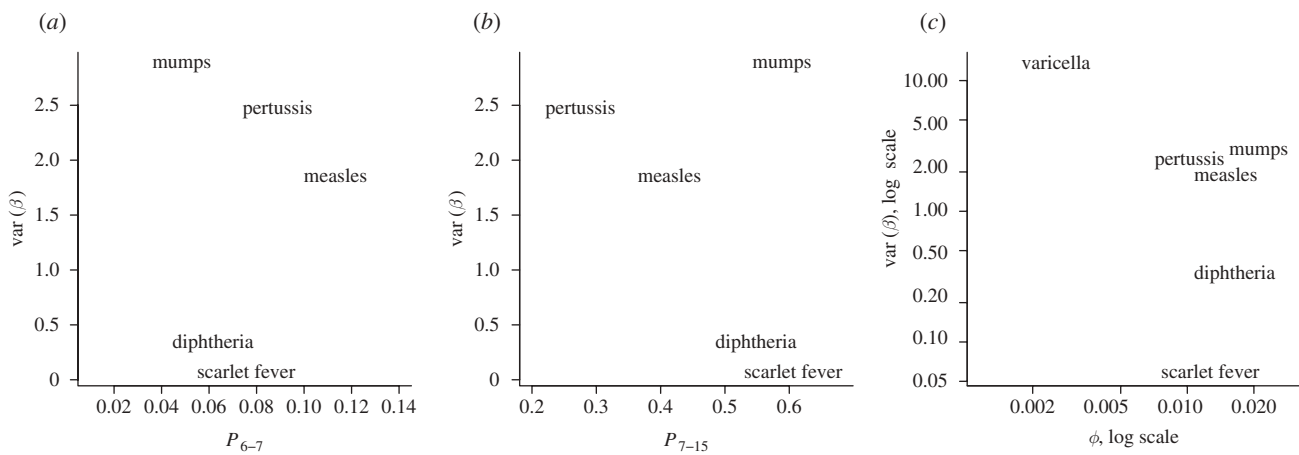


Figure 4. Relationship between the strength of seasonality, measured as variance of  $\beta$ , and (a) the proportion of cases occurring in individuals starting school ( $p_{6-7}$ ), (b) the proportion of cases in school age individuals ( $p_{7-15}$ ) and (c) a measure of the expected relative magnitude of seasonality if seasonality is entirely driven by term-time forcing (i.e. if term-time forcing dominates seasonality in transmission, this relationship should be positive). Note that the scale of the y-axis is different in the third figure, since by using  $\beta$  as a measure of  $R_0$ , we can include varicella in calculations of  $\phi$ . Using  $R_0$  based on the average age of infection (table 1) does not alter the basic patterns.

supplementary material. The generally biennial nature of the measles dynamics and more-or-less erratic annual patterns of the other infections are clear (see also full time series in figure S1, electronic supplementary material). All three viral infections show low incidence during periods of school closure. Their behaviour at the peak of incidence is somewhat different: measles and varicella peak fairly consistently between November and January (figure S2 in the electronic supplementary material). The peak of epidemics for mumps tends to be later and is quite irregular. The bacterial infections peak in incidence between October and December, and all three start increasing earlier during the summer than the viral infections.

### (b) The TSIR model

For all infections, the TSIR model provides a good fit to the short-term behaviour of the data (electronic supplementary material figure S5). In the viral infections, transmission rates are generally estimated to be low during July (figure 3a,c,e), although mumps shows a slight increase in July relative to June. The month of July corresponds to school summer vacations in Copenhagen (see electronic supplementary material). Transmission increases in August, the month where the school term starts. The bacterial infections, diphtheria, scarlet fever and pertussis, exhibit a surprising increase in transmission rates in July, i.e. much less signature of the school year (figure 3b,d,f). These results are not sensitive to the assumed generation time (figures S8 and S9 in the electronic supplementary material) or period of maternal immunity or exact alignment of school holidays and weeks.

To ensure that the timing of increases and decreases in the transmission rate observed in, say, measles did not result from a folding in of major and minor epidemics (Earn *et al.* 2000), we also fitted biennial  $\beta$  schedules to all infections. There was no evidence that the difference between diphtheria and scarlet fever, on the one hand, and infections with more biennial-like cycles, on the

other, was owing to the effects of alternating major and minor epidemics (not shown).

A range of sensitivity tests (see electronic supplementary material) support the extension of the TSIR to infections that differ from measles (via generation time and the presence of waning of immunity). Although this was not the main aim of the paper, simulations of the TSIR also capture long-term dynamics for measles (electronic supplementary material figure S7), providing further corroboration of the TSIR model (as also shown in Grenfell *et al.* (2002)). These results are promising, although more work remains for the other infections.

### (c) Age patterns of infection

Fitted age–incidence curves are shown in figure 2. Deduced parameters are given in table 1. The resulting inferred distributions of the age of infection are shown in figure 2e. Measles and pertussis have relatively narrow distributions concentrated at younger ages than diphtheria, scarlet fever and mumps. Given the breadth of the age distribution of cases for diphtheria, scarlet fever and mumps, the proportion of cases of school age is higher than that for measles and pertussis. However, the proportion of cases among individuals at ages corresponding to the start of schooling, when individuals are most likely to be susceptible, is higher in measles and pertussis (table 1). The order of  $R_0$  across infections inferred from the age–incidence curves is similar to that obtained on the basis of the  $\beta_s$  estimated from the TSIR model (pertussis > measles > scarlet fever and diphtheria), except that based on the TSIR, mumps has a higher transmission rate  $\beta$  than expected from its average age at infection (table 1). Age-specific data are not available for varicella.

Overall, there is no evidence of a relationship between amplitude of seasonality (as measured by variance in  $\beta$  through the year) and proportion of cases occurring in ages at the beginning of schooling (although measles and pertussis rank highly for both relative to the other infections; figure 4a). Infections with the lowest proportion of infected individuals of school age actually

have among the highest amplitudes of seasonality (figure 4*b*), and infections with the highest theoretically expected degree of term-time seasonality ( $\phi$ ) have among the lowest estimated values of magnitude of seasonality (figure 4*c*). To calculate  $\phi$ , we used  $\mu = 0.021$ , which is the average yearly *per capita* birth rate in Copenhagen over the period considered. Patterns obtained for the year with lowest (0.014) or highest (0.028) recorded yearly *per capita* birth rate are similar, which is also the case if  $R_0$  estimates obtained from the average age of infection are used (table 1) rather than  $\bar{\beta}$ .

#### 4. DISCUSSION

Strong seasonality was documented in all six directly transmitted childhood infections from Copenhagen. However, theoretical predictions of the magnitude of seasonality if term-time forcing is the dominant seasonal driver (Keeling & Rohani 2008) were not met (figure 4*c*). Although it is possible that the quantity used ( $\phi$ ) is simply a poor measure of term-time forcing, other evidence also supports the conclusion that seasonal forcing may be a result of many factors, rather than just patterns of mixing at schools. In particular, although seasonal transmission rates of the three viral infections (measles, mumps and varicella) show a clear signal of school vacations (i.e. low transmission during July, which corresponds to the summer holidays), for the bacterial infections, the seasonal transmission pattern deviates considerably from term-time forcing expectations, with an increase in transmission in July, despite the fact that schools are closed during this month (figure 3). Overall, this weakens the prospect of predicting the magnitude of seasonality and thus periodicity of dynamics in different settings. It is likely that information on population mixing beyond  $R_0$ , or the average age at infection (e.g. used in Keeling & Rohani (2008)) will be required.

The deviations from the well-defined term-time forcing seen in pre-vaccination measles in the UK is puzzling (Bjørnstad *et al.* 2002). A possible explanation is that school entry age is slightly higher in Copenhagen than in pre-vaccination England and Wales. In Copenhagen, susceptible depletion during term time might act more rapidly as a result of higher infection rates among pre-schoolers, resulting in the low rates of transmission in June before school closure (all diseases except pertussis; figure 3). This is an intriguing direction for more detailed age-specific modelling.

Among the viral infections, the term-time forcing pattern in July is weakest for mumps, despite the fact that for this infection, the proportion of infected individuals of school age is one of the highest among all six infections considered (table 1). There are long irregular intervals between major epidemics in mumps (figure S2 in the electronic supplementary material), perhaps caused by an interaction between the longer generation time of this infection and stochasticity, akin to that previously described for pertussis dynamics (Rohani *et al.* 2002). These irregularities may determine the weakness of the term-time forcing pattern: numbers of susceptible individuals will fluctuate across years in line with long-term dynamics, and this may swamp within-year variation. Although varicella has a similar generation time, it has a much higher transmission rate (table 1) which may

more tightly tie it to the annual attractor, leading to a profile closer to term-time forcing (figure 3).

Among the bacterial infections, pertussis appears to exhibit the most complex long-term dynamics (figure S3 in the electronic supplementary material) which may be part of the explanation for the deviation from term-time forcing. By contrast, diphtheria and scarlet fever have strongly annual epidemics, so numbers of susceptible individuals should show similar trajectories within each year, all other things being equal. Complexities such as waning of immunity (diphtheria and pertussis), asymptomatic individuals (diphtheria, scarlet fever, pertussis) and stochastic transients may contribute to irregular fluctuations in numbers of susceptible individuals through the year, but they do not shed any light on the apparent increase in transmission rates during the summer holidays (figure 3). The increase in transmission in July relative to June may be a signature of some important environmental driver in the dynamics of these bacterial respiratory infections and is an interesting direction for further work.

To conclude, our results indicate that term-time forcing is a contributing component of seasonal dynamics of many childhood infections (figure 3). Transmission is low during the months of summer vacation for the viral infections, and also low in November for all infections, attributable to susceptible exhaustion after the start of the school year (Bjørnstad *et al.* 2002). However, even for infections that are contracted at the age of schooling, significant deviations from term-time forcing are apparent. The cause of this may be some combination of (i) susceptible depletion owing to the later age of schooling in Copenhagen, (ii) long-term fluctuations in epidemic numbers and (iii) waning immunity and the presence of asymptomatic individuals. That the signal of term-time forcing (figure 3) is clearest in measles, and that this is also found across other populations, may be attributed to the strongly resonant period of measles dynamics with a natural decay time of 2 years, which will tightly determine the number of susceptible individuals in each month of the year. Infections with more complex dynamics and natural history features such as waning of immunity are likely to show less clear term-time forcing, as numbers of susceptible individuals will experience more erratic trajectories within years, thus potentially revealing the effects of environmental seasonal drivers. These complexities will affect our capacity to predict dynamics and particularly dominant epidemic periods across diseases from limited information (Keeling & Rohani 2008).

This work was funded by the Danish Medical Research Council (V.A.; grant 271-07-0555), the Bill and Melinda Gates Foundation (C.J.E.M., B.G., O.N.B.), the RAPIDD program of the Science & Technology Directorate (B.G.) and the National Institute of Health (B.G., O.N.B.; grant NIH/GM R01-GM083983-01).

#### REFERENCES

- Adams, J. S. & Hewison, M. 2007 Unexpected actions of vitamin D: new perspectives on the regulation of innate and adaptive immunity. *Nat. Clin. Pract.* **4**, 80–90. (doi:10.1038/ncpendmet0716)
- Altizer, S., Dobson, A., Hosseini, P., Hudson, P., Pascual, M. & Rohani, P. 2006 Seasonality and the dynamics of infectious diseases. *Ecol. Lett.* **9**, 497–484. (doi:10.1111/j.1461-0248.2005.00879.x)

- Anderson, R. M. & May, R. M. 1991 *Infectious diseases of humans*. Oxford, UK: Oxford University Press.
- Andreasen, V., Viboud, C. & Simonsen, L. 2008 Epidemiologic characterization of the 1918 influenza pandemic summer wave in Copenhagen: implications for pandemic control strategies. *J. Infect. Dis.* **197**, 270–278. (doi:10.1086/524065)
- Anonymous 1900–1969 Medicinal rapport for Kongeriget Danmark (Medical report for the kingdom of Denmark). Medicinal Rapport, Sundhedsstyrelsen, Copenhagen, Denmark.
- Bjørnstad, O. N., Finkenstadt, B. & Grenfell, B. T. 2002 Endemic and epidemic dynamics of measles: estimating epidemiological scaling with a time series SIR model. *Ecol. Monogr.* **72**, 169–184.
- Broutin, H., Guegan, J.-F., Elguero, E., Simondon, F. & Cazelles, B. 2005 Large scale comparative analysis of pertussis population dynamics: periodicity, synchrony, and impact of vaccination. *Am. J. Epidemiol.* **161**, 1159–1167. (doi:10.1093/aje/kwi141)
- Earn, D. J. D., Rohani, P., Bolker, B. M. & Grenfell, B. T. 2000 A simple model for complex dynamical transitions in epidemics. *Science* **287**, 667–670. (doi:10.1126/science.287.5453.667)
- Farrington, C. P., Kanaan, M. N. & Gay, N. J. 2001 Estimation of the basic reproduction number for infectious diseases from age-stratified serological survey data. *Appl. Stat.* **50**, 251–292. (doi:10.1111/1467-9876.00233)
- Feeney, K. T., Dowse, G. K., Keil, A. D., Mackaay, C. & McLellan, D. 2005 Epidemiological features and control of an outbreak of scarlet fever in a Perth primary school. *Commun. Dis. Intell.* **29**, 71–75.
- Ferrari, M. J., Grais, R. F., Bharti, N., Conlan, A. J. K., Bjørnstad, O. N., Wolfson, L. J., Guerin, P. J., Djibo, A. & Grenfell, B. T. 2008 The dynamics of measles in sub-Saharan Africa. *Nature* **451**, 679–684. (doi:10.1038/nature06509)
- Fine, P. E. M. & Clarkson, J. A. 1986 Seasonal influences on pertussis. *Int. J. Epidemiol.* **15**, 237–247 (doi:10.1093/ije/15.2.237)
- Finkenstadt, B. & Grenfell, B. T. 2000 Time series modelling of childhood diseases: a dynamical systems approach. *J. R. Statist. Soc. Ser. C* **49**, 187–205.
- Finkenstadt, B., Bjørnstad, O. N. & Grenfell, B. T. 2002 A stochastic model for extinction and recurrence of epidemics: estimation and inference for measles outbreaks. *Biostatistics* **3**, 493–510. (doi:10.1093/biostatistics/3.4.493)
- Frost, W. H., Frobisher, M., Van Volkenburgh, V. A. & Levin, M. L. 1936 Diphtheria in Baltimore; a comparative study of morbidity, carrier prevalence and antitoxin immunity in 1921–1924 and 1933–1936. *Am. J. Hyg.* **24**, 568–586.
- Galazka, A. 2000 The changing epidemiology of diphtheria in the vaccine era. *J. Infect. Dis.* **181**, S2–S9. (doi:10.1086/315533)
- Glass, K., Xia, Y. & Grenfell, B. T. 2003 Interpreting time-series analyses for continuous-time biological models—measles as a case study. *J. Theor. Biol.* **223**, 19–25. (doi:10.1016/S0022-5193(03)00031-6)
- Grassly, N. & Fraser, C. 2006 Seasonal infectious disease epidemiology. *Proc. R. Soc. B* **273**, 2541–2550. (doi:10.1098/rspb.2006.3604)
- Grenfell, B. T., Bjørnstad, O. N. & Finkenstadt, B. 2002 Endemic and epidemic dynamics of measles: scaling predictability, noise and determinism with the time series SIR model. *Ecol. Monogr.* **72**, 185–202.
- Keeling, M. J. & Rohani, P. 2008 *Modeling infectious diseases in humans and animals*. Princeton, NJ: Princeton University Press.
- Keeling, M., Rohani, P. & Grenfell, B. T. 2001 Seasonally forced disease dynamics explored as switching between attractors. *Physica D* **148**, 317–335. (doi:10.1016/S0167-2789(00)00187-1)
- Kenah, E., Lipstich, M. & Robins, J. M. 2008 Generation interval contraction and epidemic data analysis. *Math. Biosci.* **213**, 71–79 (doi:10.1016/j.mbs.2008.02.007)
- London, W. P. & Yorke, J. A. 1973 Recurrent outbreaks of measles, chicken-pox and mumps. I. The role of seasonality. *Am. J. Epidemiol.* **98**, 453–468.
- Lowen, A. C., Mubareka, S., Steel, J. & Palese, P. 2007 Influenza virus transmission is dependent on relative humidity or temperature. *PLoS Pathog.* **10**, e151. (doi:10.1371/journal.ppat.0030151)
- Pichichero, M. & Casey, J. 2003 Defining and dealing with carriers of group A streptococci. *Contemp. Pediatr.* **1**, 46.
- Rohani, P., Earn, D. J. D. & Grenfell, B. T. 1999 Opposite patterns of synchrony in sympatric disease metapopulations. *Science* **286**, 968–971. (doi:10.1126/science.286.5441.968)
- Rohani, P., Keeling, M. & Grenfell, B. T. 2002 The interplay between noise and determinism in childhood diseases. *Am. Nat.* **159**, 469–481. (doi:10.1086/339467)
- Schenzle, D. 1984 Estimation of the basic reproduction number for infectious diseases from age-stratified serological survey data. *IMA J. Math. Appl. Med. Biol.* **1**, 161–191.
- Stadslægen 1900–1970 Ugentlig oversigt over fødsler, sygdomme og dødsfald i København. (Weekly summary of births, diseases, and deaths in Copenhagen), Københavns Kommune, Copenhagen, Denmark.