Epidemiol. Infect. (2011), **139**, 1029–1038. © Cambridge University Press 2010
doi:10.1017/S0950268810002165

The epidemiology of rubella in Mexico: seasonality, stochasticity and regional variation

C. J. E. METCALF[1,2]*, O. N. BJØRNSTAD[1,3], M. J. FERRARI[1], P. KLEPAC[1,2], N. BHARTI[1,2], H. LOPEZ-GATELL[4] AND B. T. GRENFELL[2,3]

[1] *Center for Infectious Disease Dynamics, The Pennsylvania State University, University Park, PA, USA*
[2] *Department of Ecology and Evolutionary Biology, Eno Hall, Princeton University, Princeton NJ, USA*
[3] *Fogarty International Center, National Institutes of Health, Bethesda, MD, USA*
[4] *Directorate General for Epidemiology, Ministry of Health, México*

(*Accepted 17 August 2010; first published online 15 September 2010*)

SUMMARY

The factors underlying the temporal dynamics of rubella outside of Europe and North America are not well known. Here we used 20 years of incidence reports from Mexico to identify variation in seasonal forcing and magnitude of transmission across the country and to explore determinants of inter-annual variability in epidemic magnitude in rubella. We found considerable regional variation in both magnitude of transmission and amplitude of seasonal variation in transmission. Several lines of evidence pointed to stochastic dynamics as an important driver of multi-annual cycles. Since average age of infection increased with the relative importance of stochastic dynamics, this conclusion has implications for the burden of congenital rubella syndrome. We discuss factors underlying regional variation, and implications of the importance of stochasticity for vaccination implementation.

Key words: Analysis of data, epidemiology, rubella, statistics.

INTRODUCTION

Rubella is primarily a mild childhood disease, but acquiring the infection during the first 16 weeks of pregnancy is associated with the risk of fetal death or birth of a child with congenital rubella syndrome (CRS). CRS is a condition associated with a range of problems, from hearing impairment to brain damage. Estimates of the burden of CRS are high in many developing countries [1] and may frequently go unreported [2]. Since vaccination drives up the average age of infection, increasing the proportion of women of child-bearing age at risk [3–7], many countries in the developing world are not vaccinating against rubella [8], particularly given recent work indicating how partial vaccination of populations can increase the population-wide risk of CRS by allowing build-up of susceptible individuals in older age groups [9]. Multi-annual outbreak cycles are another factor that may allow such build-up because the deepened post-epidemic troughs may lead to local extinction and periods of lowered force of infection [10, 11].

Across much of the developing world, recent efforts to reduce the burden of measles mortality [12] via measles vaccination campaigns has lead to an increase in national vaccination coverage capacity and second-dose availability. In this context, it is of interest to revisit the risks associated with introduction of the

* Author for correspondence: Dr C. J. E. Metcalf, Department of Ecology and Evolutionary Biology, Eno Hall, Princeton University, Princeton NJ 0854, USA.
(Email: cmetcalf@princeton.edu)

Table 1. *Different mechanisms proposed to explain multi-annual cycles in childhood infections*

Driver	Dynamic characteristics	Predictions	Ref.
School term-time forcing of transmission	Single basin of attraction, corresponding to cycles with the observed period. For rubella, different starting conditions lead to annual or 5-year cycles	Deterministic seasonally forced dynamics should correspond to the observed multi-annual cycles	[33]
Seasonal forcing interacting with transient dynamics	Where two spectral peaks are observed, perturbation analysis shows that the resonant peak is determined by seasonality and is generally annual, the non-resonant peak can be predicted from transient dynamics of the system	Asymptotically, seasonally forced dynamics should be annual (irrespective of birth rate). The relative magnitude of resonant and non-resonant peaks should relate to the degree of demographic stochasticity, for which population size is a proxy	[19]

rubella vaccine [13]. A major challenge to quantifying the potential CRS burden is that details of the drivers of rubella dynamics remain unclear. Where rubella dynamics have been described, multi-annual cycles of varying periodicity are generally observed [14] and competing explanatory mechanisms have been put forward (summarized in Table 1). Depending on the particular mechanism underlying multi-annual cycles, vaccination might either increase or decrease the period between major epidemics by shifting dynamics between multi-annual and annual regimens. Identifying both whether rubella dynamics are multi-annual, and if so, the mechanism underlying such multi-annual epidemics is therefore of programmatic interest. The mechanisms hypothesized hang on the treatment of seasonal patterns of transmission (Table 1), currently generally assumed to reflect term-time forcing, i.e. increases in transmission associated with aggregation of children in schools, and decreases during vacations [15]. However, whether this assumption is valid is unclear, since estimation of seasonal variation of transmission from time-series has been rare, and estimates from countries outside Europe and North America are non-existent. There is also as yet very little information on the critical community size (CCS) of rubella, i.e. the population size required for the infection to persist locally without stochastic extinction [16].

The ideal dataset to tackle these questions would be spatio-temporal time-series of incidence for rubella, coupled with information on the age structure of incidence. Data on rubella are relatively rare compared to those on measles because it is a milder disease which is somewhat harder to diagnose. The notification record from Mexico coordinated by the Dirección General de Epidemiología of the Mexican Ministry of Health (described in [17]) provides an excellent opportunity to explore rubella dynamics. Rubella has been a reportable disease in Mexico since 1985 and vaccination against rubella with the MMR vaccine was implemented in 1998 as part of a broader Pan-American Health Organisation (PAHO) effort, which successfully reduced incidence to very low levels [18].

Here we use district-level monthly data from Mexico from 1985 to 2007 and yearly data on age structure to explore the critical dynamical questions. We first estimate some basic parameters of the epidemiology of rubella in Mexico to compare with results from other countries [1] and explore regional patterns of extinction ('fade-out') to establish whether stochastic dynamics are likely to be important. We then use spectral analysis to explore the magnitude of multi-annuality in the dynamics, and explore whether this can be attributed to transient dynamics [19]. We fit a time-series Susceptible–Infected–Recovered (TSIR) model to estimate seasonal swings in transmission [20]. With this, we finally explore the determinants of patterns of seasonality, and whether these vary by district size or socioeconomic characteristics. We discuss the implications of our results for identifying populations most at risk for CRS and for the implementation of vaccination campaigns.

METHODS

The data

Rubella incidence reports were obtained from ref. [21]. Monthly reported cases between 1985 and 2007 were available for each of the 32 states of Mexico (Fig. 1), at the spatial scale of districts. For 1990–2007, district-level yearly incidence is available in age groups of 5–10 years (e.g. 1–4, 5–14, 15–24, etc.). The MMR

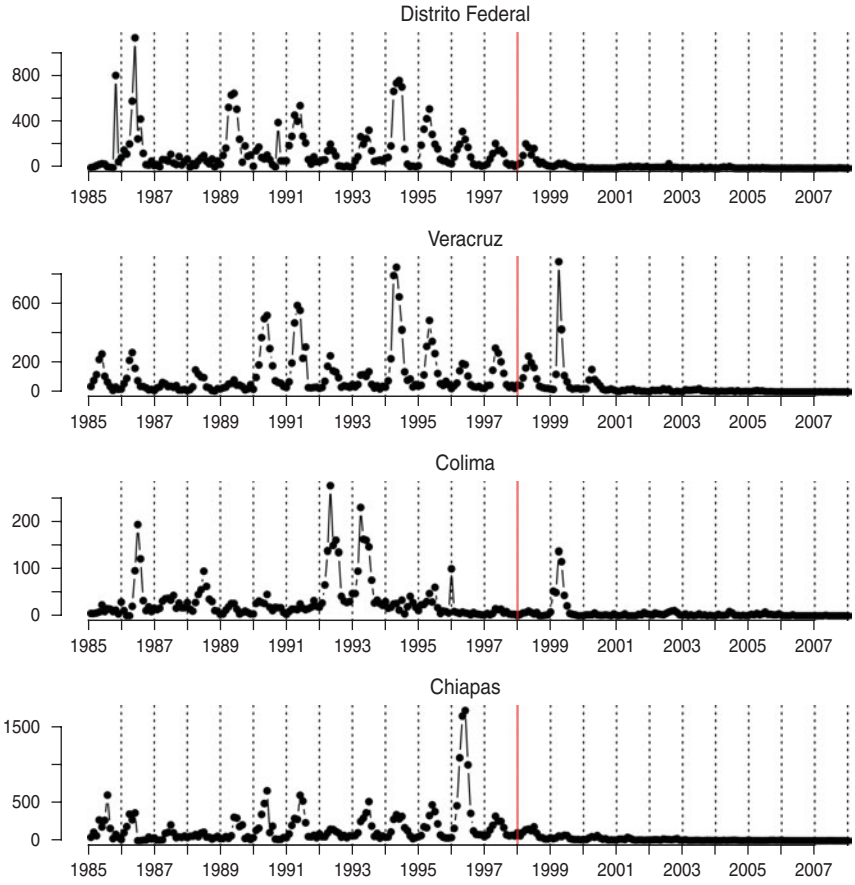

Fig. 1. Time-series of rubella from four districts in Mexico, representing a range of population size (see Table 2); the year at the start of vaccination is shown by a red vertical line.

vaccine was introduced in 1998 and resulted in relatively high coverage [18]. Birth numbers and population size for each district in each month were obtained from ref. [22]; as were socioeconomic indicators of each of the districts.

Average age of infection and R_0 from the age-structured data

Assuming negligible mortality in the disease-relevant age groups and a constant force of infection, the mean age of infection A is defined by $A = \int x\sigma(x)\mathrm{d}x$, where x is age and $\sigma(x)$ is the proportion susceptible at age x. From this, using 1 − [the cumulative proportion of case numbers over age] as a proxy for the proportion susceptible in each discrete age group, we calculated A for each district, and explored spatial patterns of average age of infection. This estimate of the average age of infection can also be used to obtain a crude estimate of the basic reproduction number, R_0, for every district. In a growing population, A is related to R_0 by $R_0 \sim G/A$ where G is the inverse of the *per capita* birth rate [23], here set to the inverse of the mean birth rate of each district. The value of R_0 estimated in this way may be biased by a range of factors, particularly differences in the force of infection over age [24]. Nevertheless, the relative magnitude of R_0 across districts should remain comparable, as long as district-specific differences in the force of infection over age are not too great. To verify this, we fitted a logistic regression to the cumulative proportion of cases over age across the entire dataset, with age fitted as a continuous covariate and district fitted as a factor. Significant variability attributable to district in the profile of infection over age would imply that transmission varies over age such that the average age is a poor indicator of the relative magnitude of R_0.

Critical community size

Since population size is a key determinant of stochastic extinction for strongly immunizing infections,

Table 2. *Average age at infection, median transmission rate estimated from the TSIR analyses, β, and the corresponding estimated reporting rate, p_{obs}, and heterogeneity parameter, α (see text); with R_0 estimated from the age data for each district, and the district's rank by population size (the range of population sizes are shown in Fig. 4)*

	Average age, A (pre-1999)	β	α	p_{obs}	R_0	Population size rank
Aguascalientes	3·68	2·95	0·95	0·077	9·36	5
Baja California	6·28	6·39	0·97	0·05266	6·18	14
Baja California Sur	5·03	3·51	0·99	0·1821	7·33	1
Campeche	7·56	2·83	0·98	0·0594	4·63	4
Coahuila	4·96	6·46	0·96	0·0273	7·98	17
Colima	8·02	3·34	0·97	0·0382	4·85	3
Chiapas	7·54	4·09	0·98	0·0031	3·41	22
Chihuahua	5·14	6·88	0·96	0·0188	7·34	20
Distrito Federal	4·55	7·98	0·98	0·0247	9·04	31
Durango	3·83	5·22	0·98	0·0124	7·77	13
Guanajuato	3·91	7·82	0·96	0·0136	8·44	27
Guerrero	7·66	4·17	0·99	0·0087	3·36	23
Hidalgo	4·51	5·28	0·96	0·0104	6·67	16
Jalisco	5·23	4·87	0·97	0·0135	7·29	29
Mexico	4·26	6·48	0·96	0·0200	9·55	32
Michoacan	5·22	6·69	0·96	0·0040	5·63	26
Morelos	5·48	4·11	0·99	0·0205	7·43	9
Nayarit	5·93	6·34	0·95	0·0111	5·78	7
Nuevo Leon	5·17	3·92	0·99	0·0523	7·95	25
Oaxaca	6·84	3·67	0·98	0·0117	4·37	24
Puebla	5·78	9·65	0·95	0·0067	5·22	28
Queretaro	4·40	5·33	0·97	0·0293	7·08	8
Quintana Roo	9·30	2·53	0·97	0·0771	3·76	2
San Luis Potosi	5·34	6·7	0·96	0·0112	6·35	18
Sinaloa	6·41	5·25	0·96	0·0155	5·32	21
Sonora	4·52	10·37	0·97	0·0210	8·38	15
Tabasco	8·54	5·68	0·98	0·0061	3·47	11
Tamaulipas	9·25	4·50	0·98	0·0497	4·23	19
Tlaxcala	4·97	8·36	0·96	0·0169	6·58	6
Veracruz	10·18	3·83	0·98	0·0137	3·26	30
Yucatan	8·32	4·21	0·98	0·0394	4·58	12
Zacatecas	3·68	8·59	0·96	0·0141	9·49	10

population size is strongly negatively related with the number of fade-outs (or proportion of zeros) in the time-series of incidence [16, 20]. The point where this curve intercepts with zero provides an indication of the CCS, or population size below which the infection is subject to stochastic fade-outs. Since under-reporting could lead to apparent fade-outs where there are none [16], we define fade-outs as corresponding to a full month with zero reported cases. Since case reporting is at the district level, rather than the city level, and since cities are more likely to be the epidemiologically relevant unit, this estimate is likely to be an upper bound on the CCS.

Spectral analysis

To test whether predicted relationships between population size and periodic features of the time-series were upheld (Table 1) we calculated periodograms of the rubella time-series for each district using modified Daniell smoothers of width 2 [25]. We located the main peaks in spectral density and identified the multi-annual peak with a period closest to an integer multiple of a year as 'resonant' [19]. For six time-series, only a single peak could be identified. For the remaining 26 districts, we tested for correlations between population size and the ratio between the

resonant and the second largest peak ('non-resonant peak') [19]. If stochastically excited transients are a driver of multi-annuality this ratio should decrease with populations size (Table 1).

The TSIR model

Seasonal transmission rates can be estimated using TSIR methods [20] in a Bayesian state-space framework [26]. The generation time (serial interval) of rubella (approximately the latent plus infectious period) is $\sim$18 days [14], so we assumed that the time-scale of the epidemic process was approximately 2 weeks, resulting in two unobserved epidemic time-steps for each observed monthly report (repeating the analysis taking a monthly time-step does not alter the major conclusions). Denoting Y_m as the number of cases reported in month m, and I_m the unobserved total number of cases, the observation process is defined by $Y_m \sim \text{binomial}(I_m, p_{\text{obs}})$, where p_{obs} is the reporting rate. The unobserved epidemic follows the TSIR process model where the number of infected individuals at time $t+1$, I_{t+1} depends stochastically on I_t, and the number of susceptible individuals S_t, with expectation $\lambda_{t+1}=\beta_m S_t I_t^{\alpha}/N_t$, where β_m is the transmission rate in every month and the exponent α, usually a little less than 1, captures heterogeneities in mixing not directly modelled by the seasonality [20, 27] and the effects of discretization of the underlying continuous time process [28]. Note that inasmuch as the time-step used represents the true average time from infection to recovery, β_m estimated in this way corresponds to the seasonally varying basic reproductive number R_0 (since it indicates the number of new infections caused by a single infected individual in a wholly susceptible population), and thus the median value of β_m can be compared to estimates of R_0 obtained using age-based calculations. We model I_{t+1} as a negative binomial random variable with expectation λ_t and clumping parameter I_t [20]. Susceptible individuals are depleted by infections and replenished by births, B_t, and are modelled with the renewal equation $S_{t+1}=S_t+B_t-I_{t+1}$. Since vaccine coverage levels are thought to be around 80%, and vaccinated individuals avoid susceptibility, births are discounted by 0·8 following the start of vaccination (i.e. post-vaccination, $S_{t+1}=S_t+0{\cdot}2B_t-I_{t+1}$). We assumed that infection predominantly occurs early in life, so that depletion of susceptible individuals by mortality can be neglected.

We used a Bayesian state-space model to obtain parameter estimates. Flat priors were set on the number of susceptible individuals at every time-step, with the upper and lower limits set to constrain the time-series to reflect a 20% range around the known proportion of susceptible individuals in 1990 [29]. Uninformative priors were set for the β_m parameters (a normal distribution left-truncated at zero with a mean of 6 and a standard deviation of 10); and flat priors were set on the α parameter, constraining it to be between 0·9 and 1. We set an informed prior on the observation probability p_{obs}, centred at an initial estimate obtained from susceptible reconstruction [20, 27] for each district, following a beta distribution to allow for direct sampling.

We initiated the chain with constant monthly transmission set to 6, and $\alpha=0{\cdot}95$. At each iteration, we chose at random to update candidates of either one of the β_m parameters or the α parameter, or a subset of the unobserved process, I_t, also chosen at random. If the latter was chosen, we used the renewal equation to obtain the corresponding numbers of susceptible individuals through time, and checked that this violated none of the conditions or priors. If the candidate proved suitable, we then calculated the acceptance ratio for the candidate parameters (calculated by the product of the likelihood of the observation model, the likelihood of the process model and the prior density functions of the candidates, divided by the same three elements as calculated for the current parameter values). We accepted or rejected the candidates accordingly, and then used a Gibbs step to update observation probabilities. After a burn-in of 10 000 iterations we ran the chain for 100 000 iterations, sampling every 100th step to avoid autocorrelation.

RESULTS

Age of infection

The mean age of infection estimated for the pre-vaccination era (before 1999, see above) ranges between 3 (Zacatecas) and 11 (Veracruz) years, with highest average ages of infection concentrated in the south and coastal areas of the country (Fig. 2); age-based R_0 estimates are correspondingly low in these locations (Table 2). A logistic regression relating the cumulative proportion of infected individuals over age showed no significant effect of location ($\chi^2_{31}=2{\cdot}92$, $P>0{\cdot}5$, Fig. 3), suggesting that the relative magnitudes of R_0 estimated in this way are comparable across locations (although episodic dynamics will provide a further slight bias, see below).

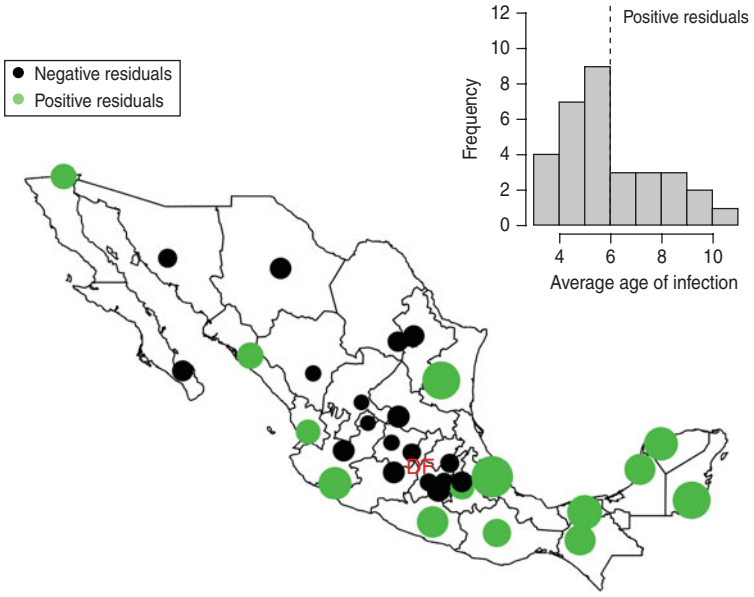

Fig. 2. The inset shows a histogram of average age at infection taken from years before 1999; the dotted vertical line indicates the countrywide average. The map shows the magnitude (point size) and sign of deviations from the countrywide average log age of infection (point colour) taken from this inset histogram. Points are placed in district capitals. Districts centred around the district containing the capital city ('Federal District' indicated as 'DF') have a lower than average age of infection; coastal and southern districts have a higher average age of infection.

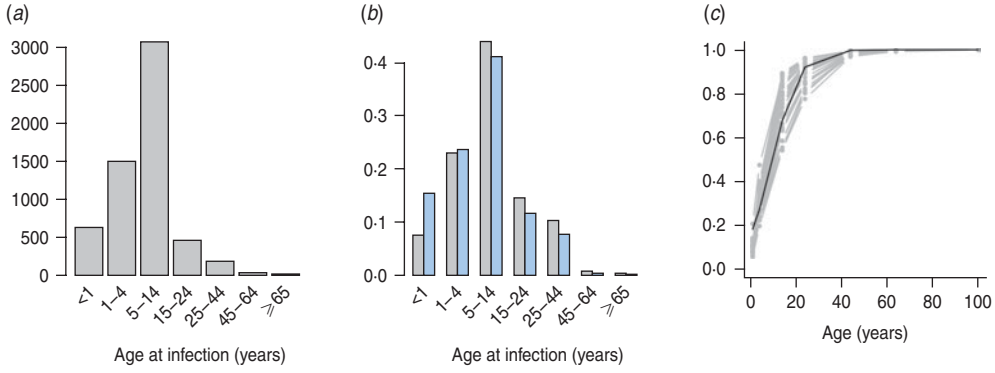

Fig. 3. The pattern of infected individuals across age showing (*a*) numbers reported from 1985 to 1999 across the whole of Mexico; (*b*) the proportion across age from 1985 to 1999 in the smallest (Baja California Sur) and largest (Federal District) districts, similar profiles are obtained in all other districts; and (*c*) the cumulative proportion over age (grey lines) and a logistic regression fitted to this data [black line, $\text{logit}(y) \sim -1{\cdot}67(0{\cdot}33)+0{\cdot}17(0{\cdot}02)\times \text{age}$, $\chi^2_1=7{\cdot}72$, D.F. $=223$, $P<0{\cdot}01$; parentheses indicate standard errors]. Incorporating district as a factor did not significantly improve this model ($\chi^2_{31}=2{\cdot}92$, $P>0{\cdot}5$), suggesting little regional difference in the pattern of force of infection over age.

Average age of infection was not significantly related to population size ($F_{1,30}=0{\cdot}20$, $P>0{\cdot}5$), or districts' socioeconomic indicators ($F_{1,30}=3{\cdot}51$, $P>0{\cdot}05$). Estimates of average age of infection are consistent with global patterns of results on seronegativity, which generally indicate that more than half of the population has been infected by age 13 years [1]; estimates are also in line with the incidence-based estimate of an average age of infection of approximately 9 years from Peru [11].

Critical community size

The relationship between proportion of zero incidence and log population size in the pre-vaccination era is very noisy, and more triangular than linear. However, the relationship is significantly negative ($F_{1,30}=8{\cdot}61$, $P<0{\cdot}01$), and provides an indicator of the order of magnitude of the CCS as being larger than a million (Fig. 4). This is comparable to estimates from Peru [30].

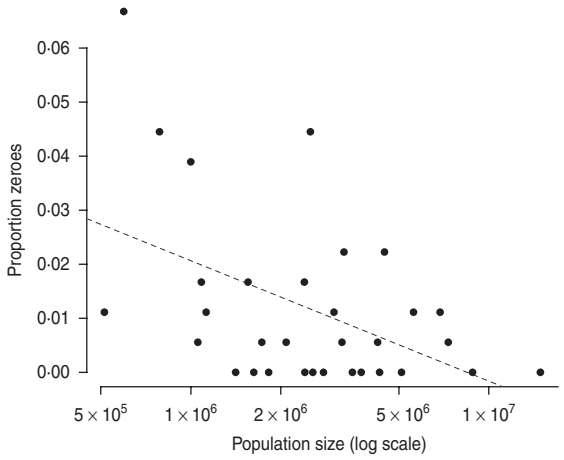

Fig. 4. The relationship between log population size and proportion of zeros in the time-series for years up to 1998 when vaccination was started (Fig. 1). The dashed line is $y = 0{\cdot}154 - 0{\cdot}009x$, $P < 0{\cdot}01$, $r^2 = 0{\cdot}22$. The line intersects zero at a population size of 8 500 000; no occurrence of zero fade-outs occurred in districts smaller than one million.

Spectral analysis

Across districts, the mean period of the non-resonant peak was 3·79 years with a standard deviation of 1·25 years. The resonant peak was 1 year for all districts, and the magnitude of the non-resonant peak was greater than the resonant peak in only three districts (Baja California Sur, Quintana Roo, Tabasco). Two of these are the smallest populations in the data. The relationship between population size and the ratio of the magnitude of the resonant and non-resonant peaks, denoted k, was significant (Fig. 5), supporting the role of stochasticity in driving inter-annual variability (Table 1) via transient dynamics [19]. Log average age of infection was also significantly and positively related to the ratio of the magnitude of the resonant and non-resonant peaks ($F_{1,25} = 10{\cdot}41$, $P < 0{\cdot}05$, $r^2 = 0{\cdot}22$, with $y = 4{\cdot}74 + 2{\cdot}00k$), i.e. districts with more strength in the non-resonant peak had a higher average age at infection. This suggests that stochasticity may increase the mean age of infection (and may slightly bias estimates of R_0 obtained above). Including the mean transmission rate β estimated via the TSIR (see below) did not significantly improve this regression model ($F_{1,24} = 0{\cdot}06$, $P > 0{\cdot}4$).

Seasonality and transmission magnitude

The overall magnitude of seasonal transmission estimated via the TSIR model (which can be equated with R_0, see above) concurred with the broad estimate of R_0 obtained via average age of infection (Table 2), suggesting relatively low transmission in this infection relative to the most contagious childhood infections such as whooping cough or measles [14]. Seasonal variation in transmission rates estimated by the Bayesian state-space TSIR showed a signal of term-time forcing, with low transmission in July; this pattern did not vary much across socioeconomic indicators (Fig. 6). The coefficient of variation in seasonality of transmission ranged from 0·16 (Oaxaca) to 8·57 (Durango). There were no clear geographical differences in the pattern of seasonality (not shown) and no correlation between magnitude of seasonality and socioeconomic indicators ($n = 32$, $\rho = -0{\cdot}20$, $P > 0{\cdot}1$). The coefficient of variation of seasonality and the average age of infection were significantly negatively correlated ($n = 32$, $\rho = -0{\cdot}37$, $P < 0{\cdot}05$), so that locations with higher amplitude transmission also had a lower average age of infection.

DISCUSSION

Our results point to considerable regional variability both in patterns of seasonality and average age of infection of rubella across Mexico, with for example low average age of infection around Mexico City (in agreement with a previous serosurvey in this district [31]), and higher average age of infection in the southern and coastal regions (Fig. 2). This variation in the average age of infection could result from either regional variability in R_0, or regional variability in the importance of episodic dynamics, as Ferrari *et al.* [10] have demonstrated that such dynamics can inflate the upper tail of the age-incidence curve. Various lines of evidence support the latter explanation (see below). Our results also indicate some general patterns for the epidemiology of rubella in Mexico such as a signal of term-time forcing [15] with lower transmission during school holidays found across the country (Fig. 6), as has also been indicated by a recent analysis of rubella in Peru [11, 30].

Our results provide general support for stochasticity as a driver of rubella dynamics. Our analysis suggests a CCS for rubella larger than 1 000 000, considerably larger than the estimate for measles in the UK, estimated to be between 300 000 and 500 000 [16]. This figure is likely to be an overestimate since the geographic unit used is districts rather than cities, but the large CCS is not unexpected given the comparatively weaker transmission rate of rubella relative to measles, and is consistent with one other estimate

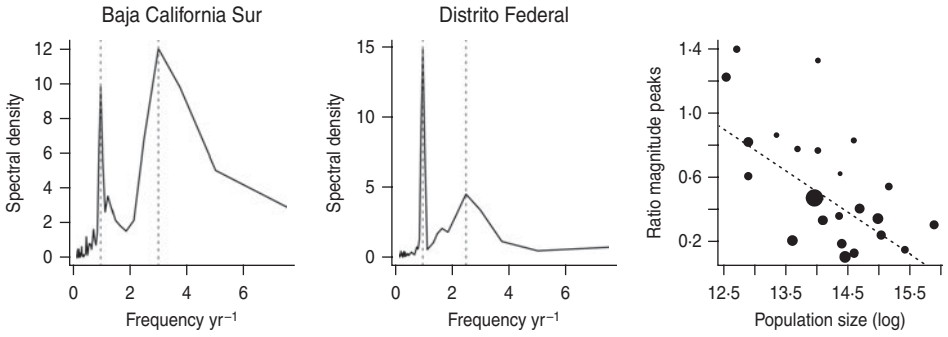

Fig. 5. Spectral density of log reported rubella incidence +1 for two districts in Mexico, with the resonant and non-resonant peaks shown as vertical dashed lines; and the ratio of the non-resonant peak over the resonant peak across population size for districts where the non-resonant peak was identifiable (26/32 districts). Point size scales inversely with the precision of the second peak (large points correspond to sharp peaks, small points correspond to broad peaks). The equation of the line is $y = 3{\cdot}49 - 0{\cdot}21x$, $P < 0{\cdot}01$, $r^2 = 0{\cdot}26$, and corresponds to a regression weighted by the precision of the second peak (the unweighted regression is also significant).

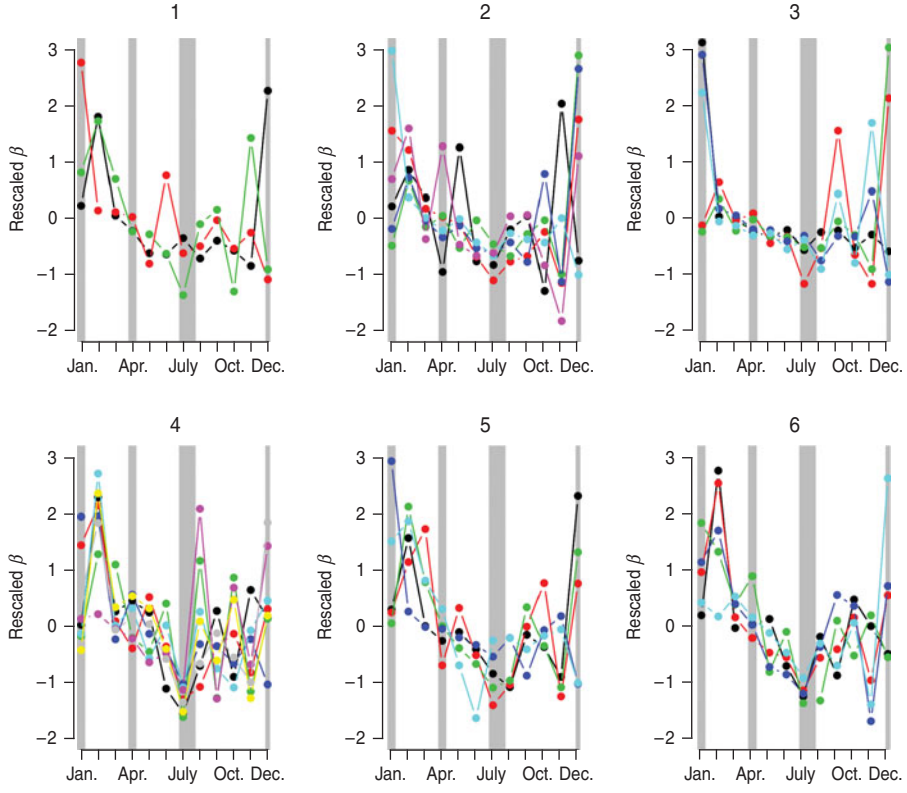

Fig. 6. Pattern of seasonal transmission rates across districts, organized according to the socioeconomic status rankings (see Methods section): '1' indicates low, and '6' indicates high. Transmission rates have been scaled to have the same mean and variance to facilitate comparison of seasonality. Colors are (1) Chiapas = black, Guerrero = red, Oaxaca = green; (2) Campeche = black, Hidalgo = red, Puebla = green, San Luis Potosi = blue, Tabasco = turquoise, Veracruz = pink; (3) Durango = black, Guanajuato = red, Michoacan = green, Tlaxcala = blue, Zacatecas = pink; (4) Colima = black, Mexico = red, Morelos = green, Nayarit = blue, Queretaro = turquoise, Quintana Roo = pink, Sinaloa = yellow, Yucatan = grey; (5) Baja California = black, Baja California Sur = red, Chihuahua = green, Sonora = blue, Tamaulipas = turquoise; (6) Aguascalientes = black, Coahuila = red, Jalisco = green, and Nuevo Leon = blue, Distrito Federal = turquoise. Approximate school holidays are shown in grey (2 weeks at Christmas, 2 weeks at Easter, and July to mid-August).

for rubella in Peru [30]. This suggests that local extinction is likely to be frequent, inducing more episodic dynamics and allowing build-up of susceptible individuals in later age groups [10], implying that connectivity between population centres is a key direction of research in establishing the burden of CRS [11].

In contrast to the multi-annual dynamics observed for rubella in the pre-vaccination era in a range of developed country contexts [14, 19], dynamics of rubella in Mexico were predominantly annual (i.e. in Figure 5, the ratio is very rarely >1, indicating that the annual resonant peak dominates). This could reflect the scale of our observation, i.e. district-level observations could be averaging highly erratic local dynamics, as in the case of measles in Niger, where country-level dynamics appear annual, but dynamics at the local scale are near-chaotic [26]. However, a discernible non-resonant peak was detectable for all but six districts, and the relative magnitude of the district-level non-resonant and resonant peaks was significantly related to the district average age of infection. This relationship with average age of infection is not predicted if local dynamics deviate from district-level observations, but is expected if local dynamics within a district are synchronized and episodic [11].

Since dynamic features do significantly correlate with the average age of infection (even though the annual features dominate), the mechanism underlying episodic dynamics remains a question of considerable interest in the context of rubella. Our results support a role for stochasticity [19]; and in fact the low magnitude of the non-resonant peak relative to the resonant peak would be expected in the context of this mechanism given the relatively high population densities across districts in Mexico. This mechanism is also supported by results including (i) the relationship between population size and relative magnitude of the resonant and non-resonant peaks (Fig. 5); and (ii) estimates of R_0 which are sufficiently low (~ 6) as to be unlikely to drive complex multi-annual cycles alone [32], unless birth rates are much higher than those observed in Mexico. This importance of stochasticity in determining the dynamics of rubella is of public health significance because it provides a second mechanism by which moderate vaccination cover can enhance CRS risk; vaccination will break local chains of transmission and thus increase stochasticity in dynamics, potentially exacerbating the predicted increase in mean age relative to less stochastic settings [11].

The generality of the importance of stochasticity in rubella dynamics is an interesting question for future research. Serology data from Gabon have suggested R_0 values as high as 16 [14], and similar magnitudes have been reported for Ethiopia and China [9]. Such robust transmission would considerably reduce the importance of stochasticity. Distinguishing whether deterministic high transmission or transient dynamics and fade-outs are the norm for rubella is important in considering implementation of vaccination strategies. For example, in Mexico spatially explicit deterministic models that ignore demographic stochasticity and local heterogeneities (Fig. 2) will not accurately represent disease dynamics by failing to capture frequent local extinctions, or inter-annual variability, and attempts to model optimal vaccination strategies may consequently go awry. Overall, the relatively low estimates of R_0 obtained here (Table 2), and a short generation time indicate that interrupting rubella transmission may be relatively straightforward, and this is supported by the success of the Mexican vaccination campaign (Fig. 1). However, in rubella, unless vaccination is maintained at high levels, local extinction may be followed by increases in CRS incidence resulting from the build-up of susceptible individuals in older age groups [11].

ACKNOWLEDGEMENTS

We thank the staff of Dirección General de Epidemiología and the Mexican National Surveillance System for their efforts in conducting surveillance activities and providing the rubella data used in this study. This work was funded by the Bill and Melinda Gates Foundation and grant NIH R01 GM083983-01 (C.J.E.M., B.T.G., O.N.B, M.J.F.), and the RAPIDD program of the Science & Technology Directorate, Department of Homeland Security and the Fogarty International Center, National Institutes of Health (B.T.G., O.N.B, M.J.F.).

DECLARATION OF INTEREST

None.

REFERENCES

1. **Cutts FT, Vynnycky E.** Modelling the incidence of congenital rubella syndrome in developing countries. *International Journal of Epidemiology* 1999; **28**: 1176–1184.
2. **Lawn JE, *et al.*** Unseen blindness, unheard deafness, and unrecorded death and disability: congenital rubella in Kumasi, Ghana. *American Journal of Public Health* 2000; **90**: 1555–1561.
3. **Knox EG.** Strategy for rubella vaccination. *International Journal of Epidemiology* 1980; **9**: 13–23.

4. **Anderson RM, Grenfell BT.** Quantitative investigations of different vaccination policies for the control of congenital rubella syndrome (CRS) in the United Kingdom. *Journal of Hygiene* (*Cambridge*) 1986; **96**: 305–333.
5. **Anderson RM, May RM.** Vaccination against rubella and measles: qualitative investigations of different policies. *Journal of Hygiene of Cambridge* 1983; **90**: 259–325.
6. **Panagiotopoulos T, Antoniadou I, Valassi-Adam E.** Increase in congenital rubella occurrence after immunisation in Greece: retrospective survey and systematic review. *British Medical Journal* 1999; **319**: 1462–1467.
7. **Massad E, *et al.*** A model-based design of a vaccination strategy against rubella in a non-immunized community of Sao Paulo State, Brazil. *Epidemiology and Infection* 1994; **112**: 579–594.
8. **Robertson SE, *et al.*** Control of rubella and congenital rubella syndrome (CRS) in developing countries, part 2: vaccination against rubella. *Bulletin of the World Health Organization* 1997; **75**: 69–80.
9. **Vynnycky E, Gay NJ, Cutts FT.** The predicted impact of private sector MMR vaccination on the burden of Congenital Rubella Syndrome. *Vaccine* 2003; **21**: 2708–2719.
10. **Ferrari MJ, *et al.*** Episodic outbreaks bias estimates of age specific force of infection: a corrected method using measles in Niamey, Niger as an example. *Epidemiology and Infection* 2010; **138**: 108–116.
11. **Metcalf CJE, *et al.*** Rubella meta-population dynamics and importance of spatial coupling to the risk of congenital rubella syndrome in Peru. *Journal of the Royal Society Interface* (in press).
12. **Wolfson LJ, *et al.*** Has the 2005 measles mortality reduction goal been achieved? A natural history modelling study. *Lancet* 2007; **369**: 191–200.
13. **Plotkin SA.** Rubella eradication. *Vaccine* 2001; **19**: 3311–3319.
14. **Anderson RM, May RM.** *Infectious Diseases of Humans*. Oxford: Oxford University Press, 1991.
15. **Fine PEM, Clarkson JA.** Seasonal influences on pertussis. *International Journal of Epidemiology* 1986; **15**: 237–247.
16. **Bartlett MS.** Measles periodicity and community size. *Journal of the Royal Statistical Society, Series A* (*General*) 1957; **121**: 48–70.
17. **Vargas MH.** Ecological association between scarlet fever and asthma. *Respiratory Medicine* 2006; **100**: 363–366.
18. **Santos JI, *et al.*** Measles in Mexico, 1941–2001: interruption of endemic transmission and lessons learned. *Journal of Infectious Diseases* 2004; **189**: S243–S250.
19. **Bauch CT, Earn DJD.** Transients and attractors in epidemics. *Proceedings of the Royal Society of London, Series B* 2003; **270**: 1573–1578.
20. **Bjørnstad ON, Finkenstadt B, Grenfell BT.** Endemic and epidemic dynamics of measles: Estimating epidemiological scaling with a time series SIR model. *Ecological Monographs* 2002; **72**: 169–184.
21. **Dirección General de Epidemiología.** Yearly Morbidity Report. México, 2008.
22. **National Institute of Statistics and Geography.** Statistical Information > Source/Project > Administrative records > Vital statistics > Birth Statistics > Interactive data query, 2008.
23. **McLean AR, Anderson RM.** Measles in developing countries. Part I: Epidemiological parameters and patterns. *Epidemiology and Infection* 1988; **100**: 111–133.
24. **Farrington CP, Kanaan MN, Gay NJ.** Estimation of the basic reproduction number for infectious diseases from age-stratified serological survey data. *Applied Statistics* 2001; **50**: 251–292.
25. **Venables WN, Ripley BD.** *Modern Applied Statistics with S*. New York: Springer, 2003.
26. **Ferrari MJ, *et al.*** The dynamics of measles in sub-Saharan Africa. *Nature* 2008; **451**: 679–684.
27. **Finkenstadt B, Grenfell BT.** Time series modelling of childhood diseases: a dynamical systems approach. *Journal of the Royal Statistical Society, Series C* 2000; **49**: 187–205.
28. **Glass K, Xia Y, Grenfell BT.** Interpreting time-series analyses for continuous-time biological models-measles as a case study. *Journal of Theoretical Biology* 2003; **223**: 19–25.
29. **Trujillo GG, Munoz O, Tapia-Conyer R.** Seroepidemiology of rubella in Mexican women. *National Survey of Public Health Probabilities in Mexico* 1990; **32**: 623–631.
30. **Rios-Doria D, *et al.*** Spatial and temporal dynamics of rubella in Peru, 1997–2006: geographic patterns, age at infection and estimation of transmissibility. In: Chowell EA, ed. *Mathematical and Statistical Estimation Approaches in Epidemiology*. Springer: Netherlands, 2009.
31. **Golubjatnikov R, Elsea WR, Leppla L.** Measles and rubella hemagglutination-inhibition antibody patterns in Mexican and Paraguayan children. *American Journal of Tropical Medicine and Hygiene* 1971; **20**: 958–963.
32. **Earn DJD, *et al.*** A simple model for complex dynamical transitions in epidemics. *Nature* 2000; **287**: 667–670.
33. **Keeling MJ, Rohani P, Grenfell BT.** Seasonally forced disease dynamics explored as switching between attractors. *Physica D* 2001; **148**: 317–335.