

## Research



**Cite this article:** Fonnesebeck CJ, Shea K, Carran S, Cassio de Moraes J, Gregory C, Goodson JL, Ferrari MJ. 2018 Measles outbreak response decision-making under uncertainty: a retrospective analysis. *J. R. Soc. Interface* **15**: 20170575.  
<http://dx.doi.org/10.1098/rsif.2017.0575>

Received: 4 August 2017

Accepted: 26 February 2018

### Subject Category:

Life Sciences – Mathematics interface

### Subject Areas:

biometrics

### Keywords:

measles epidemic, Bayesian disease model, outbreak control, vaccine intervention, decision analysis

### Author for correspondence:

Christopher J. Fonnesebeck

e-mail: [fonnesbeck@gmail.com](mailto:fonnesbeck@gmail.com)

Electronic supplementary material is available online at <https://dx.doi.org/10.6084/m9.figshare.c.4024459>.

# Measles outbreak response decision-making under uncertainty: a retrospective analysis

Christopher J. Fonnesebeck<sup>1</sup>, Katriona Shea<sup>2,3</sup>, Spencer Carran<sup>3</sup>, Jose Cassio de Moraes<sup>4</sup>, Christopher Gregory<sup>5</sup>, James L. Goodson<sup>6</sup> and Matthew J. Ferrari<sup>2,3</sup>

<sup>1</sup>Department of Biostatistics, Vanderbilt University Medical Center, Eleventh Floor, Suite 11000, 2525 West End Avenue, Nashville, TN, USA

<sup>2</sup>Department of Biology and Intercollege Graduate Degree Program in Ecology, 208 Mueller Laboratory, The Pennsylvania State University, University Park, PA, USA

<sup>3</sup>Center for Infectious Disease Dynamics, Department of Biology, Eberly College of Science, The Pennsylvania State University, University Park, PA, USA

<sup>4</sup>Faculdade de Ciências Médicas da Santa Casa de São Paulo, São Paulo, Brazil

<sup>5</sup>Arboviral Diseases Branch, Division of Vector-Borne Diseases, National Center for Emerging and Zoonotic Infectious Diseases, US Centers for Disease Control and Prevention, Fort Collins, CO, USA

<sup>6</sup>Accelerated Disease Control and Vaccine Preventable Disease Surveillance Branch, Global Immunization Division, Center for Global Health, US Centers for Disease Control and Prevention, Atlanta, GA, USA

CJF, 0000-0002-7107-140X

Resurgent outbreaks of vaccine-preventable diseases that have previously been controlled or eliminated have been observed in many settings. Reactive vaccination campaigns may successfully control outbreaks but must necessarily be implemented in the face of considerable uncertainty. Real-time surveillance may provide critical information about at-risk population and optimal vaccination targets, but may itself be limited by the specificity of disease confirmation. We propose an integrated modelling approach that synthesizes historical demographic and vaccination data with real-time outbreak surveillance via a dynamic transmission model and an age-specific disease confirmation model. We apply this framework to data from the 1996–1997 measles outbreak in São Paulo, Brazil. To simulate the information available to decision-makers, we truncated the surveillance data to what would have been available at 1 or 2 months prior to the realized interventions. We use the model, fitted to real-time observations, to evaluate the likelihood that candidate age-targeted interventions could control the outbreak. Using only data available prior to the interventions, we estimate that a significant excess of susceptible adults would prevent child-targeted campaigns from controlling the outbreak and that failing to account for age-specific confirmation rates would underestimate the importance of adult-targeted vaccination.

## 1. Introduction

Reports of resurgent outbreaks of vaccine-preventable diseases following long periods of relative absence are increasingly common [1–4]. Several factors may contribute to the occurrence of such outbreaks. McLean & Anderson [5] predicted that such outbreaks should be expected because of the ‘honeymoon’ phenomenon following the introduction of vaccination, whereby post-vaccination cohorts no longer experience high rates of natural immunization to supplement population immunity following vaccination activities [6]. Further, population-level vaccination rates may decline over time as immigration from areas of low vaccination coverage lead to a build-up of susceptible individuals, or the reduction in individual infection risk [7] leads to apathy about vaccination within the population. Also, local stochastic extinction may result in temporary breakdown of local transmission, even though populations remain susceptible to subsequent outbreaks upon the reintroduction of infection [8].

Increasingly, in the event of a measles outbreak, outbreak response immunization (ORI) is recommended as an intervention. The goals of these ORIs are twofold: (i) to protect high risk groups (i.e. young children) and (ii) to attenuate the current outbreak [9,10]. To achieve the former goal, ORI campaigns routinely target children 6–59 months of age [9]. To achieve the latter goal, the campaign must reach some target level of immunization ( $P_c$ )—i.e. a percentage reduction of the susceptible population—such that effective reproductive number,  $R_e$ , will be below 1 and the outbreak will end. From the standard susceptible–infected–removed (SIR) model, this level of immunization is  $P_c = 1 - 1/R_e$  [11]. To identify this target coverage and appropriately plan a campaign, one must estimate both the value of  $R_e$  itself, which determines the necessary reduction of the susceptible population required to end the outbreak, and the age distribution of the susceptible population, which allows us to identify the critical age classes to be targeted in a campaign. For example, if  $R_e = 1.5$  one must then reduce the susceptible fraction by 33%; if 80% immunization of susceptibles in the target age groups can be achieved (considering both campaign coverage and vaccine efficacy), then at a minimum, the intervention should target the first  $a$  age groups whose summed proportion ( $\sum_{i=1}^a s_i$ )/ $S$  of the susceptible population  $S$  exceeds  $0.33/0.8 = 0.41$ .

Experience with past outbreaks can provide guidance about likely values of  $R_e$  [12] and the likely distribution of the susceptible population [13,14]. However, in the case of resurgent outbreaks, which follow periods of relatively low measles incidence, there may be insufficient data on which to base estimates of  $R_e$ . In these settings, it is the current outbreak itself which may provide the most relevant empirical information [12,15,16]. However, clinical confirmation of measles cases through case-based surveillance systems has relatively low specificity, particularly in settings of low prevalence [17–20] when other illnesses that result in fever and rash (e.g. rubella, Dengue) may be misdiagnosed as measles [18]. Given that the various aetiologies of febrile illness may disproportionately affect different age classes, reliance on a clinical definition alone may result in a biased assessment of the age classes at risk [12,17]. Further, misdiagnosis of cases may bias the assessment of the rate of increase of total cases and lead to a biased estimate of  $R_e$  and hence the vaccination coverage necessary to limit the outbreak. Though serological confirmation of measles cases is preferred to clinical confirmation, resources often limit the proportion of clinical cases that can be confirmed by serology in outbreak settings in time to be of use to decision-makers.

A variety of methods are available to estimate  $R_e$  from early surveillance data [12,21–23]. The use of mathematical modelling to estimate the age distribution of the susceptible population during an outbreak is rare; decisions about age targeting have been classically based on prior experience or early evaluation of cases confirmed earlier in the outbreak [24]. Here we present an epidemic model that uses serological confirmation on a subset of cases to estimate age-specific confirmation, and use these estimates to correct the observed number of reported cases before being used to estimate epidemic parameters.

The age distribution of the susceptible population is usually unknown at the start of a measles outbreak. The mean and variance of the age distribution of susceptible

individuals are expected to increase as the prevalence of infection declines [13,25] and during periods of measles absence [12]. As a result, historical surveillance data may be misleading with respect to the current distribution of susceptibles. Initial estimates of the susceptible population can be reconstructed from demographic rates, historical records of routine and supplemental vaccination coverage, and measles incidence [26]; however, uncertainty in these rates and historical incidence mean that *a priori* estimates of the susceptible population may be significantly biased. Further, lack of data on migration, heterogeneity of vaccination coverage, and clustering of susceptibles means that these estimates will be highly uncertain in the best of circumstances. The 1997 measles outbreak in São Paulo, Brazil, presented in detail below had many more adolescent and adult cases than was expected based on historical rates [27]. This pattern of unexpectedly wide age distributions of cases has been recently seen in outbreaks in Malawi [28] and Mongolia (<http://www.wpro.who.int/mongolia/media-centre/releases/20160505-measles-outbreak-faqs/en/>).

This work comprises a retrospective analysis of a resurgent measles outbreak in São Paulo, Brazil in 1996–1997. This outbreak followed several years of relatively low measles incidence as a result of both routine and supplemental measles vaccination. Brazil began its national immunization plan in 1973, with a single dose of measles vaccine. In 1992, Brazil introduced a nation-wide recommendation of a second routine dose of measles vaccine, while adopting a goal of measles elimination by the year 2000. Also beginning in 1992, Brazil began conducting national supplemental measles vaccination campaigns. Notably, São Paulo state conducted a restricted campaign in 1992, targeting children between the ages of 9 months and 10 years (compared to a 9-month-to-14-year target elsewhere in the country) and did not participate in a 1995 national campaign targeting children between the ages of 1 and 3 years. The outbreak itself resulted in over 30 000 confirmed cases in São Paulo State, with an unexpectedly high proportion of cases in adults; 60% were in individuals greater than 20 years of age. Several limited vaccination campaigns targeting children under 4 years of age, health workers, and some adults were implemented between June and August, prior to a widespread campaign targeting all children between 6 months and 4 years of age that was conducted in August of 1997. As a consequence of the broad age distribution of susceptibles, typical childhood-based ORIs may not have resulted in sufficient immunity to limit the outbreak. Following the outbreak, Brazil implemented a Supplementary Emergency Measles Action Plan in 1999 with significant increases in surveillance, case investigation and rapid response [29]. Between 2001 and 2013 all cases that occurred in Brazil were attributed to importations. During 2013–2014 there was an outbreak of measles in Ceara, Brazil [30]. This outbreak was confirmed to have ended after the last case was registered in July 2015 and in September 2016 the International Expert Committee for Documenting and Verifying Measles, Rubella, and Congenital Rubella Syndrome Elimination in the Americas declared measles eliminated in the whole of the Americas.

We created a novel statistical model that combines an *a priori* model of the susceptible population based on available immunization coverage information, a time-series model of the progression of the outbreak, and an age-specific model

of IgM serological confirmation of suspected cases to estimate  $R_e$  and the age distribution of the susceptible population. To illustrate how real-time surveillance could be applied to inform the design of ORI age targets, we generated estimates of  $R_e$  and the resulting ORI target recommendations at both 1 and 2 months prior to the vaccination campaign that was conducted on 15 August. We argue that a flexible approach to ORIs can better incorporate the information gained in the early stages of an outbreak to identify campaign age targets and that unforeseen biases in clinical diagnosis can be mitigated through the incorporation of serological confirmation and high quality surveillance data.

## 2. Methods

### 2.1. Data

Case-based records of individuals presenting with clinical measles symptoms (fever, rash, and at least one of the following symptoms: cough, conjunctivitis or coryza [17]) during the calendar year of 1996 were provided by the Ministério da Saúde (Ministry of Health) in Brazil. Fields for each record included the county in which the case presented, reporting date, age in months or years, and the results (positive, negative or inconclusive) of a serological test for measles specific IgM, if one was conducted. After limiting the data to cases presenting in urban São Paulo and discarding those with incomplete records, there were 10 810 cases presenting with clinical symptoms only and 23 699 cases with serological tests, of which 1067 were discarded due to lack of reagent or improper collection and thus treated as clinical cases.

The age distribution of the population in São Paulo was extrapolated from the decadal census. Historical rates of routine vaccination coverage were taken from <http://apps.who.int/gho/data/node.main.A826>.

### 2.2. Confirmation bias model

We specified a structured case confirmation submodel to retrospectively determine the age group-specific probabilities of laboratory confirmation (i.e. laboratory positive for measles specific IgM) for measles in São Paulo, conditional on clinical diagnosis. Individual laboratory confirmation events  $c_i$  were modelled as Bernoulli random variables, with the probability of confirmation being allowed to vary by age group:

$$c_i \sim \text{Bernoulli}(p_{a[i]}),$$

where  $a[i]$  denotes the age group for the individual indexed by  $i$ . There were 16 age groups, the first 15 of which were 5-year age intervals [0, 5), [5, 10), ..., [70, 75), with the last interval including all individuals 75 years and older. Since our choices of age group boundaries were arbitrary, we allowed the probabilities of adjacent groups to be correlated with one another. To this end, the transformed set of probabilities was modelled as a multivariate normal random variable:

$$\beta^{(\text{conf})} \sim N(\mu, \Sigma),$$

where the transformation is the logit ( $\text{logit}(p) = \log[p/(1-p)]$ ), which serves to convert probabilities (defined on the [0, 1] interval) to the real line:

$$\text{logit}(p_a) = \beta_a^{(\text{conf})},$$

where the corresponding covariance matrix was tridiagonal, incorporating a correlation term  $\rho$  that was assumed, for simplicity, constant among groups. This was to allow confirmation

rates to be correlated between neighbouring age groups, since the age boundaries were arbitrarily defined.

$$\Sigma = \begin{bmatrix} \sigma^2 & \sigma^2\rho & 0 & \dots & 0 & 0 \\ \sigma^2\rho & \sigma^2 & \sigma^2\rho & \dots & 0 & 0 \\ 0 & \sigma^2\rho & \sigma^2 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \sigma^2 & \sigma^2\rho \\ 0 & 0 & 0 & \dots & \sigma^2\rho & \sigma^2 \end{bmatrix}.$$

To estimate the true (latent) number of cases for each age group  $I_a$ , the estimated probabilities were used to correct the clinically reported cases  $n_a^{(\text{clinic})}$ , as modelled by a binomial distribution, and the total number of cases then calculated as the sum of this estimated value  $x_a^{(\text{clinic})}$  and the laboratory-confirmed cases  $x_a^{(\text{lab})}$ :

$$\begin{aligned} x_a^{(\text{clinic})} &\sim \text{Bin}(n_a^{(\text{clinic})}, p_a) \\ I_a &= x_a^{(\text{clinic})} + x_a^{(\text{lab})}. \end{aligned}$$

The set of  $I = \{I_a\}$  was, in turn, used to inform the estimate of the latent age distribution of the infected class. A natural probability distribution to model age classes is the multinomial distribution, which is parametrized by a set of probabilities, here corresponding to the expected proportion in each age group; the corresponding prior for this vector of probabilities is the Dirichlet distribution. These are specified by

$$\begin{aligned} f^{(\text{age})} &\sim \text{Dirichlet}(\mathbf{1}) \\ \{I_a\} &\sim \text{Multinomial}(I, f^{(\text{age})}), \end{aligned}$$

where  $f^{(\text{age})}$  is a vector of age-specific proportions and  $\mathbf{1}$  a vector of ones, which is used as a non-informative prior for the Dirichlet.

## 2.3. Disease dynamics model

### 2.3.1. Initial susceptible population

We used the population age structure in 1996 and the history of vaccination (through routine immunization and campaigns) prior to 1997 to estimate the initial number of susceptible individuals in each age class at the beginning of the 1997 outbreak. We modelled the number of susceptible individuals in each annual age class at the beginning of 1997  $S_{a,0}$  as a finite mixture of local susceptibles  $S_{a,0}^{(L)}$  and excess susceptibles  $S_{a,0}^{(E)}$ , where excess susceptibles are defined as those that cannot be predicted from local demographic processes. We model local susceptibles  $S_a^{(L)}$  as a binomial draw from the São Paulo population:

$$S_a^{(L)} \sim \text{Binomial}(N_a, p_s),$$

where  $N_a$  is the population size in age class  $a$  and  $p_s$  the realized susceptibility. Uncertainty in this probability was specified with a beta distribution having expected value  $\mu_s$ , which was calculated based on historical immunization activity in the city:

$$\begin{aligned} \mu_s &= \pi^{(\text{vacc})} \pi^{(\text{SIA})} \pi^{(\text{nat})} \\ &= (1 - 0.85V_a) \prod_{j=1}^2 [1 - \psi_j I_{a,j}] \left[ \exp \sum_{y=0}^{a-1} \theta_{a-y} \right], \end{aligned}$$

where  $\pi^{(\text{vacc})}$ ,  $\pi^{(\text{SIA})}$  and  $\pi^{(\text{nat})}$  are proportions of residual susceptibility following routine vaccination, supplemental immunization activities (SIA) and natural immunity,  $V_a$  is the routine vaccination coverage experienced by age class  $a$  when they were eligible for routine vaccination between 9 and 12 months of age. We assume that efficacy of the routine dose is 85% [31],<sup>1</sup> while SIAs conducted in 1987 and 1992 were conservatively assumed to result in 80% reduction of susceptibles in their targeted age classes [32]. Further,  $\psi_j$  is the coverage of the  $j$ th SIA and  $I_{a,j}$  is an indicator function that is 1 if age class  $a$  was eligible for the  $j$ th SIA and 0 otherwise, and  $\theta_y$  the force of

infection in year  $i$  (i.e. the contribution of natural immunity). The force of infection is assumed to be  $1/A$  in the absence of vaccination, where  $A = L/R_0$  [11] is the mean age of infection in the absence of vaccination,  $L$  is the life expectancy and  $R_0$  is the basic reproduction number. The force of infection  $\theta_i$  in the presence of vaccination is then approximated as  $(1 - V_a)/A$  [11]. The basic reproduction number  $R_0$  is estimated below. We present an analysis of sensitivity to assumptions about historical vaccination coverage in the electronic supplementary material.

We modelled excess susceptibles using a data augmentation approach. A latent number of susceptibles, constrained to a value uniformly distributed between zero and 1 million, were added to the model. These excess susceptibles were assigned an age distribution, modelled as Gaussian with unknown mean  $\mu_{\text{age}}$  and standard deviation  $\sigma_{\text{age}}$  and added to the resident susceptibles for a total susceptible estimate in the São Paulo population:  $S_a = S_a^{(L)} + S_a^{(E)}$ . Because the number and distribution of excess susceptibles were modelled as latent variables, the estimates of their values were determined by the data, rather than being specified using prior information. This provided a means of accounting for infections that may not have been provided by a local population of susceptibles, but potentially by a pool of non-resident susceptibles migrating from areas of low immunity. We also fit an alternative version of the model assuming no excess susceptibles (i.e. the distribution of susceptibles is forced to be consistent with local demographics), to assess the effect of including excess susceptibles explicitly in our model.

We modelled the number of measles cases in each age class  $a$  and time step  $t$  as a Poisson random variable:

$$I_a \sim \text{Poisson}\left(S_{a,t-1} \frac{I_{t-1} B_a}{N}\right),$$

where  $S_{a,t-1}$  is the number of susceptibles in age class  $a$  at time  $t - 1$ ,  $I_{t-1}$  a row vector of the number of infected individuals in each age class at time  $t - 1$ , and  $B_a$  is the  $a$ th column of the who acquires infection from whom (WAIFW) matrix. We model the WAIFW matrix as an assortative matrix  $B$

$$B = \begin{bmatrix} \beta & \beta\delta & \dots & \beta\delta^{k-2} & \beta\delta^{k-1} \\ \beta\delta & \beta & \dots & \beta\delta^{k-3} & \beta\delta^{k-2} \\ \beta\delta^2 & \beta\delta & \dots & \beta\delta^{k-4} & \beta\delta^{k-3} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \beta\delta^{k-1} & \beta\delta^{k-2} & \dots & \beta\delta & \beta \end{bmatrix}$$

that assumes that interaction among age groups declines exponentially with difference in age. The basic reproduction number,  $R_0$ , in the model of the initial population size is taken as the dominant eigenvalue of the matrix  $B$ .

Our assumed contact matrix structure is simple, reflecting the fact that there is little information in the available data to inform a more complex parametrization. Nevertheless, the decay parameter allows for some flexibility in transmission dynamics; if there is little difference in the contact rate among age groups, the estimated  $\delta$  will be close to one, while lower (higher) contact among disparate groups will result in  $\delta$  estimates less (greater) than one. In previous work, Mossong *et al.* [33] used population survey data to estimate contact rates within eight European countries. Unfortunately, estimating a similar empirically driven contact matrix for the São Paulo population was outside the scope of this work. However, to examine the potential for more realistic patterns of mixing to explain the dynamics of the 1997 outbreak, we conducted a sensitivity analysis using a contact matrix for Brazil, calculated by [34]. For this, we used the following formulation:

$$B = \beta M^\alpha,$$

where  $M$  is the contact matrix of [34] and  $\alpha$  is a constant, estimated as a random variable in the model.  $\alpha$  was assigned an Exponential(0.1) prior, giving ample prior weight to any

plausible values it might take. Values of  $\alpha$  near one would imply that contact structure like that represented in  $M$  is relevant for predicting the observed data, while values closer to zero would eliminate age-specific transmission values, since all the elements would converge to one.

## 2.4. Model fitting

Our model was fitted using data truncated at two different dates during the outbreak, to simulate the information state at June 15 and July 15, 1997. These two dates, which include 3698 and 11 982 cases, respectively, represent two contrasting levels of available monitoring data on which to potentially base intervention decisions. Hence, excluding any potential lags in reporting, the model was fitted only to the information that would have been available at those time points (this includes the estimation of age-specific confirmation bias). Hence, four different management scenarios were modelled:

- June 15 information state, adjusted for age-specific confirmation bias
- June 15 information state, unadjusted for confirmation bias
- July 15 information state, adjusted for age-specific confirmation bias
- July 15 information state, unadjusted for confirmation bias

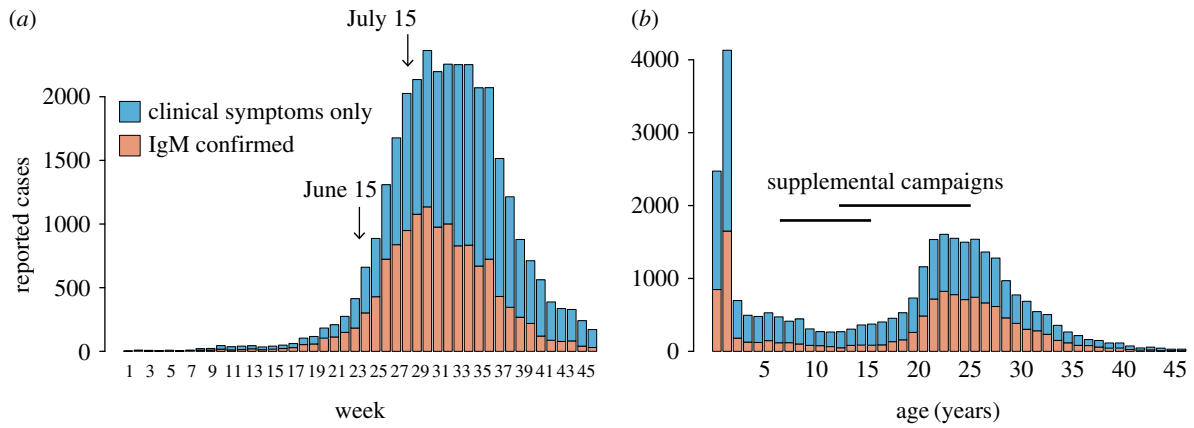
All models were fitted using PyMC 2.3 [35], a software package for the Python programming language that fits Bayesian statistical models using Markov chain Monte Carlo [36] sampling. Each model was sampled for 50 000 iterations using a Metropolis–Hastings sampling algorithm, with the first 40 000 samples discarded conservatively as a warmup period, and the remaining sample was assessed for lack of convergence using the Geweke diagnostic [37]. Hence, all inference was based on the final 10 000 samples from each model run.

## 2.5. Evaluating outbreak response campaigns

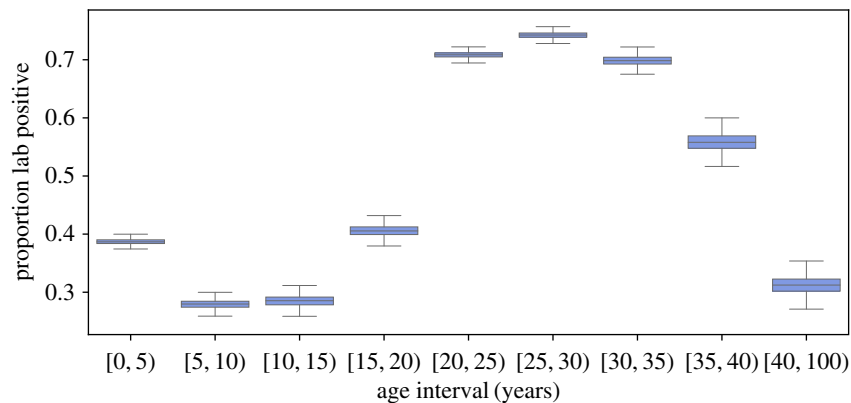
The classic result states that a fraction  $p_c > 1 - 1/R_e$  of susceptibles must be immunized (i.e. vaccinated and seroconverted) in order to reduce  $R_e < 1$  [11]. Young children are conventionally targeted in outbreak response vaccination campaigns because they are, on average, more likely to be susceptible and to experience severe complications because of measles infection. In populations with a long history of measles control, the age distribution of susceptibles is frequently wider [25] and wider age campaigns are often considered to reach larger proportion of susceptibles. As  $R_e$ , and thus the  $p_c$ , increases, then one might target a wider age range to increase the proportion of susceptibles immunized for a given coverage. Here, both the estimate of  $R_e$  and the age distribution of susceptibles are conditional on both date of the estimate (June or July) and the model used (all clinical cases or age-corrected). For each model, we calculated the necessary vaccination threshold assuming a campaign that achieves 90% coverage of the target population and 95% efficacy. We then evaluated whether there would have been empirical support on 15 June or 15 July that outbreak response campaigns targeting individuals from 6 months to 5 years, 6 months to 15 years, and 6 months to 30 years, or a mixed strategy targeting children 6 months to 5 years and adults 20 to 30 years would have met this necessary target.

## 3. Results

The 1997 outbreak in São Paulo began in December 1996 and spread throughout São Paulo State, peaking in August 1997 (figure 1a). The outbreak resulted in 25 393 total suspected cases, of which 13 516 were confirmed by serology (IgM).



**Figure 1.** (a) Weekly time series of suspected (blue) and laboratory confirmed (orange) measles cases in municipal São Paulo during 1997. (b) Number of suspected (blue) and laboratory confirmed (orange) measles cases by age in municipal São Paulo during 1997.



**Figure 2.** The estimated proportion of suspected measles cases in each 5 year age group that were IgM positive. Box plots indicate the central 50% of the posterior density and the posterior mean. Whiskers indicate the 2.5th and the 97.5th percentiles. (Online version in colour.)

The age distribution of cases was strongly bimodal, with 22% of suspected and confirmed cases below 2 years of age and a secondary mode at between 23 and 26 years of age (figure 1*b*).

The probability of a clinical case being confirmed by positive IgM serology was strongly age dependent (figure 2). Clinical cases less than 20 years were significantly more likely to be IgM negative (posterior mean between 27% and 42%, figure 2); individuals between age 20 and 40 years were significantly more likely to be IgM positive (greater than 50%). The consequence of this bias is that the observed age distribution of clinical cases over-represents the number and proportion of childhood cases.

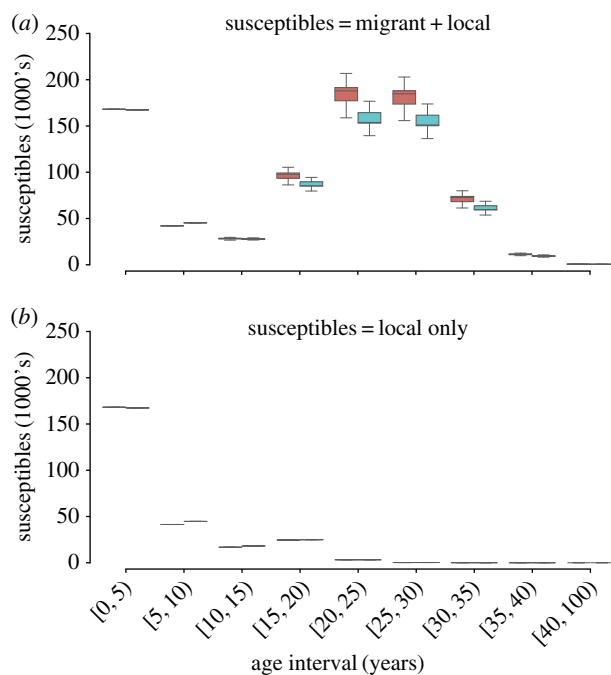
If we assume that the entire pool of susceptibles is from the local population (no excess susceptibles), the best fit model for 15 June data yields an estimate of  $R_0 = 57$  (95% credible interval: 54–60). This estimate is not consistent with previously published estimates of  $R_0$  for measles [38,39]; thus we interpret this as a result of poor model fit and do not consider this model parametrization further (see electronic supplementary material for details).

We fit the full epidemic model, including resident and excess susceptibles, using data through either 15 June or 15 July under the assumption either that all clinical cases were true measles cases or that true measles cases were a sub-set of the reported clinical cases (i.e. corrected for confirmation bias), where the confirmation rate was determined by the age-specific confirmation model. The age distribution

estimate of total susceptibles (both local and excess) was similar at both observation points (figure 3*a,b*). The model fit using the age confirmation model consistently estimated larger numbers of susceptibles in the [0, 5) and [20, 25) year age classes and fewer susceptibles in the [15,20) year age classes (figure 3*a,b*). Estimates of  $R_0$  from the 15 July data were 12.4 (95% CI 11.5–13.5) for the age confirmation model and 12.75 (95% CI 12–13.5) for model fitted to data using only clinical confirmation. The corresponding estimate using data truncated to 15 June was 11.75 (95% CI 10–14) and 11.5 (95% CI 10–12.5) for the clinical only and age confirmation models respectively.

The effective reproduction number  $R_e$ , which describes the rate at which the epidemic spread in 1997 in a partially immune population, was estimated to be clearly greater than 1 for both 15 June and 15 July using the age confirmation model (figure 4). Notably, while the posterior mean estimate for  $R_e$  on 15 June was 1.125, the 95% credible interval did not overlap the estimate that results from including the data through 15 July. Ignoring the age confirmation model and treating all clinical cases as measles cases yielded a 15 June estimate of  $R_e = 1.02$  with a 95% credible interval (0.95–1.08) that includes 1. Using all data through 15 July, the estimate of  $R_e$  was comparable, irrespective of whether clinical cases were corrected using the age confirmation model or not.

These differences in  $R_e$ , with posterior means ranging from 1.02 to 1.25 depending on time (June or July) and

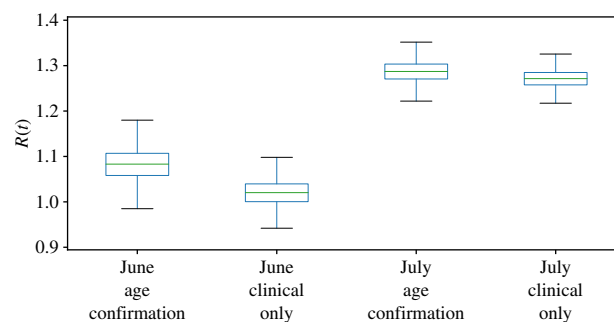


**Figure 3.** (a) The estimated total number of susceptible individuals in each age class based on data observed on 15 July using the age confirmation model (red) and the clinical only model (blue). (b) The estimated number of resident susceptibles—total susceptibles minus excess susceptibles—in each age class based on data observed on 15 July using the age confirmation model (orange) and the clinical only model (blue).

observation model (clinical or age-corrected), though small in absolute value, reflect larger practical differences in the vaccination response required to stop the outbreak. If we use these estimates to calculate minimum vaccination thresholds (assuming campaigns that reach 90% of the target population and achieve 95% effectiveness), this establishes vaccination targets ranging from 2 to 20% of susceptibles; conservatively using the 95th percentile of the posterior distribution of  $R_e$  estimates, this shifts the target range to 8–24% (figure 5). From this, we define an ORI campaign as *sufficient* if it is expected to reduce the susceptible population by at least the threshold amount. Based on the July estimates, neither confirmation model predicts a 5-and-under strategy to be sufficient for stopping the outbreak (figure 5). Using the June estimates under the age confirmation model, there is the suggestion that targeting children under 5 years would be sufficient, while the June estimate fitted to the clinical cases only predicts all strategies would be comfortably above the threshold. In all cases, ORI strategies that target individuals up to 15 or 30 years of age, as well as a mixed strategy that targets children under 5 and adults between 20 and 30 years of age, are predicted to be sufficient to stop the outbreak.

## 4. Discussion

The 1997 measles outbreak in São Paulo was unexpected, having followed several years of high routine vaccination coverage, SIAs, and relatively low incidence. Further, the age distribution of cases, with a secondary mode among adults, had not been observed in previous outbreaks. Thus, while historical precedent often serves as a guide for outbreak response, in this case, such precedent would have

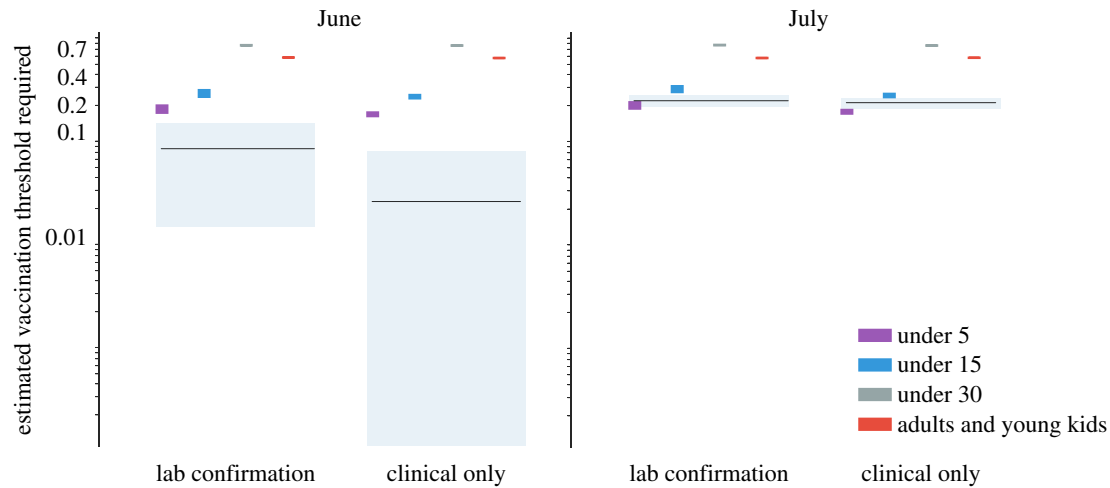


**Figure 4.** Estimated effective reproduction number,  $R_e$ , estimated by the age confirmation model and the clinical only model, using data available on 15 June and 15 July. (Online version in colour.)

greatly underestimated outbreak risk and the vaccination targets necessary to control the outbreak. Here we have presented a novel approach for integrating real-time outbreak surveillance into the evaluation of an evolving outbreak in order to evaluate candidate response strategies. In doing so, we have developed a model for interpreting clinical measles surveillance that acknowledges that the correlation between clinical measles symptoms and laboratory confirmation of positive measles IgM serology is age-specific. Further, we have shown that relying on clinical confirmation alone can significantly bias inference about transmission rate ( $R_e$ ) and the minimal vaccination targets required to stop an outbreak.

In the case of São Paulo in 1997, estimates of  $R_e$  were similar as was the assessment of which campaign age targets were sufficient to reduce  $R_e$  below 1, regardless of whether confirmation bias was corrected, when fitted to data on 15 July. Thus, while our proposed model accounting for age-specific bias in serologic confirmation explicitly estimates the uncertainty in clinical diagnosis, it results in little practical difference in the interpretation of risk  $R_e$  or candidate interventions on 15 July. However, using only the data available on 15 June, estimates based only on clinical confirmation data would have grossly underestimated risk and overestimated the benefit of a vaccination campaign targeting children below 5 years of age.

Although outbreak risk can be evaluated *a priori*, outbreaks themselves are often the first indication of the build-up of susceptibles or gaps in immunity. In 1997, the age distribution of cases in São Paulo indicated a dangerous gap in immunity among individuals between 15 and 35 years of age. The SIA conducted in 1987, targeting children below 14 years of age, would be expected to have immunized individuals below 23 years of age in 1997; those older than 23 years would have been born prior to a national immunization system in Brazil and would be expected to have experienced natural infection during their childhood. We estimated that excess susceptibles between 15 and 35 years of age may have accounted for 63% of all susceptibles during the 1997 outbreak. We term these as *excess* susceptibles because they are excess relative to the expected age distribution of susceptibles based on historical rates of natural infection, routine vaccination and SIAs. We are unable to positively identify the source of these excess susceptibles; they may have been the result of over-estimating the coverage of previous vaccination programmes, or migrants from low coverage or low transmission risk areas that were unlikely to be exposed to vaccination or natural infection. While the former explanation



**Figure 5.** A comparison of the estimated immunization threshold that would be achieved by four vaccination strategies (each targeting different age classes) to the minimum vaccination threshold necessary to reduce the effective reproduction number below one. Estimates of the necessary threshold and 95% credible interval are shown by black lines and grey regions, respectively, based on data available on 15 June and 15 July for the age-confirmation and clinical only models. Coloured rectangles give the 95% credible intervals for the estimated immunization achieved by each of the four vaccination strategies. If the estimated immunization achieved by a campaign is higher than the minimum vaccination threshold required, then the campaign is assumed to be sufficient to stop the outbreak.

is possible, insufficient vaccination coverage would be expected to result in more circulating infection, which would still likely result in exposure to natural infection, and thus immunity, by adulthood. The latter explanation requires that individuals were recent immigrants to São Paulo and had not been exposed to either vaccination or natural infection as children in the region that they emigrated from. Measles rarely persists endemically in small populations below some critical community size [40,41]; thus it is possible that recent migrants from small villages might have not been exposed to natural infection. A case-control trial conducted after the 1997 outbreak found that recent immigration to São Paulo was a significant risk factor for measles infection during the outbreak [27]. Further, immigration rates into São Paulo in 1991 were highest among individuals between 15 and 30 years of age [42], which is consistent with the age distribution of the excess susceptibles estimated by our models. While this does not confirm that immigration or gaps in prior immunization were the source of the adult susceptibles during the 1997 outbreak, this analysis does suggest that these adult susceptibles may have played a significant role in the outbreak; absent the excess susceptibles,  $R_e$  at the start of the outbreak would have been comfortably less than 1. Other recent measles outbreaks have exhibited this same age profile, with an unexpectedly large number of adult cases (e.g. Malawi [28], Mongolia (<http://www.wpro.who.int/mongolia/mediacentre/releases/20160505-measles-outbreak-faqs/en/>), China [43]). Thus, strategies for monitoring and targeting immunity gaps in adults (e.g. migrant susceptibles, in this case) may be useful in preventing future outbreaks. Moreover, outbreak response strategies should consider adult-targeted vaccination when surveillance indicates a large number of adult susceptibles.

Though our models account for the age distribution of susceptibles, we make very simplistic assumptions about age-specific transmission; namely that within age-class transmission is the same for all ages, and between age-class transmission decays exponentially with difference in ages. Important recent work has shown that age-specific mixing

rates are likely to vary considerably and may be culturally specific [33]. It is possible that higher contact (and thus, transmission) rates among adults mean that adult susceptibles disproportionately contributed to this outbreak. Fitting an age-specific transmission matrix with differential age-specific rates would have greatly inflated the model complexity and is beyond the scope of the present work, but may be worthwhile in cases where age-specific contact network structure information is available [44]. In our sensitivity analysis based on the contact matrix for Brazil estimated by [34], the value of the exponent  $\alpha$  for scaling the matrix was estimated to be 0.01 (95% BCI = [0.00, 0.04]). This resulted in the virtual elimination of contact structure in favour of a constant transmission parameter. Moreover, the estimate of  $R_0$  under this parametrization was 52 (95% BCI = [50, 55]), which suggests a poor fit relative to what would be expected of a measles outbreak. Thus, we conclude that there is no support for a structured generational matrix such as that presented by [34] to have generated the observed patterns in the age distribution of cases in this outbreak. However, we do not claim that structured mixing was not happening in São Paulo, only that it is not necessary to explain the age distribution that was observed, nor is it preferred to a simpler model.

There are several potential explanations for the age-specific serological confirmation bias for cases with clinical symptoms. There are many aetiologies that may generate fever and rash symptoms in young children [17–20], which would increase the rate of false positives in children based on clinical symptoms alone. Further, young children may be more likely to be brought to clinic shortly after the onset of symptoms, when IgM titres may not yet have reached detectable levels. Regardless of the cause of this bias, the result is that assessing the overall age distribution of cases may tend to underestimate the role of adults if based on clinical confirmation alone. Further, as the age distribution of cases presenting with symptoms may change over the course of an outbreak (in São Paulo, there were relatively more young children near the start and end of the outbreak, and more adults in the middle), clinical cases may reflect

time-varying confirmation and bias estimates of transmission rates. If true, by correcting for age-specific confirmation bias, we also necessarily correct for temporal variation in false positives. We see the effect of this in the higher estimated  $R_e$  using the confirmation bias corrected model with the 15 June data (figure 4), relative to that of the clinical confirmation model; though the model predicts fewer true measles cases (i.e. false positives are removed), the exponential increase from the start of the outbreak to 15 June is steeper and this estimate is consistent with the higher  $R_e$  estimate from the longer time series available on 15 July.

While we have demonstrated the potential for integrating real-time monitoring information to inform and improve outbreak response, the ability to provide model-based predictions is just one component for an operational system for adaptive control of outbreaks. For example, our approach does not take into account the investment of time and resources needed to prepare for ORI, such as vaccine procurement, microplanning, training or social mobilization. A complete decision support system would fully integrate costs and constraints into the planning process so that management objectives can be adequately met. However, because we implemented our model using an extensible, open-source Bayesian modelling framework (PyMC), one can readily integrate it into a larger decision support system, as appropriate.

This work highlights the value of an integrated approach to model-based inference and prediction in supporting decision-making for the control of measles outbreaks. We used a Bayesian hierarchical model structure to integrate three critical sources of uncertainty that are common in emerging and re-emerging outbreaks: the structure and parameter values of the underlying disease dynamics model, partial observability of epidemic process due to imperfect reporting or diagnostics, and the current level of risk due to uncertainty in historical levels of vaccine coverage and efficacy. The full model is comprised of component sub-models, which are linked explicitly via conditional probability statements; this results in parameter estimates and evaluation of candidate vaccination targets that fully incorporate the uncertainty in the dependent variables, ultimately yielding predictions that more fully account for the lack of complete information at hand for decision-making. For example, the likely presence

of a pool of excess susceptibles in the population was not evident based on historical immunization records, but was made apparent only when this information was integrated with outbreak monitoring data, via a realistic model of measles dynamics. From this, the model predicted that only campaigns targeting adults, as well as children, were likely to be effective in stopping the outbreak, and this was evident from data available as early as July 15. In practice, while targeted efforts to vaccinate school children and adults were made in 1997, before and after the August campaign targeting children between 6 months and 4 years, this analysis suggests, retrospectively, that there may have been evidence to support broader vaccination efforts targeting older age groups. Despite the apparent complexity of the integrated model, it was straightforward to implement and fit using readily available, open-source software, which allowed us to iteratively evaluate key assumptions regarding model structure and prior information in terms of their influence on model estimates and predictions. We believe this approach holds promise as a structured decision-making tool for a variety of disease outbreak scenarios, by efficiently incorporating all available information and incorporating monitoring information to provide robust decision support under uncertainty.

**Data accessibility.** The full model used to obtain estimates for this manuscript is available as a Jupyter notebook: <https://goo.gl/EowM1u>. The notebook includes all data preparation steps, the full PyMC model and diagnostic outputs.

**Authors' contributions.** C.J.F., K.S., S.C., J.C.d.M., C.G., J.L.G. and M.J.F. designed the research; C.F. and M.J.F. performed analyses and wrote the paper. All authors gave final approval for publication.

**Competing interests.** We declare we have no competing interests.

**Funding.** C.J.F., K.S. and M.J.F. are supported by a grant from the Ecology and Evolution of Infectious Disease Program of the National Science Foundation/National Institutes of Health (Award no. 1 R01 GM105247- 01) ([www.nsf.gov](http://www.nsf.gov)). M.J.F. is funded through a grant from the Bill and Melinda Gates Foundation.

## Endnote

<sup>1</sup>We assume efficacy is a function of age, hence 85% at 9 months (largely due to interference with maternal immunity), with higher efficacy for older age groups.

## References

- Hersh BS, Fine PEM, Kent WK, Cochi SL, Kahn LH, Zell ER, Hays PL, Wood CL. 1991 Mumps outbreak in a highly vaccinated population. *J. Pediatr.* **119**, 187–193. (doi:10.1016/s0022-3476(05)80726-7)
- Cherry JD. 2012 Epidemic pertussis in 2012 the resurgence of a vaccine-preventable disease. *N. Engl. J. Med.* **367**, 785–787. (doi:10.1056/nejmp1209051)
- Celentano LP, Massari M, Paramatti D, Salmaso D, Tozzi AE. 2005 Resurgence of pertussis in Europe. *Pediatr. Infect. Dis. J.* **24**, 761–765. (doi:10.1097/01.inf.0000177282.53500.77)
- Shibeshi ME *et al.* 2014 Measles resurgence in southern Africa: challenges to measles elimination. *Vaccine* **32**, 1798–1807. (doi:10.1016/j.vaccine.2014.01.089)
- McLean AR, Anderson RM. 1988 Measles in developing countries part I. Epidemiological parameters and patterns. *Epidemiol. Infect.* **100**, 111. (doi:10.1017/s0950268800065614)
- Jansen VAA. 2003 Measles outbreaks in a population with declining vaccine uptake. *Science* **301**, 804. (doi:10.1126/science.1086726)
- Omer SB, Salmon DA, Orenstein WA, deHart MP, Halsey N. 2009 Vaccine refusal mandatory immunization, and the risks of vaccine-preventable diseases. *N. Engl. J. Med.* **360**, 1981–1988. (doi:10.1056/nejmsa0806477)
- Ferrari MJ, Grais RF, Bharti N, Conlan AJK, Bjørnstad ON, Wolfson LJ, Guerin PJ, Djibo A, Grenfell BT. 2008 The dynamics of measles in sub-Saharan Africa. *Nature* **451**, 679–684. (doi:10.1038/nature06509)
- Cairns KL, Perry RT, Ryman TK, Nandy RK, Grais RF. 2011 Should outbreak response immunization be recommended for measles outbreaks in middle- and low-income countries? An update. *J. Infect. Dis.* **204**, S35–S46. (doi:10.1093/infdis/jir072)
- Grais RF, Strebel P, Mala P, Watson J, Nandy R, Gayer M. 2011 Measles vaccination in humanitarian emergencies: a review of recent practice. *Confl. Health* **5**, 21. (doi:10.1186/1752-1505-5-21)
- Anderson RM, May RM. 1981 The population dynamics of microparasites and their invertebrate hosts. *Phil. Trans. R. Soc. B* **291**, 451–524. (doi:10.1098/rstb.1981.0005)



12. Durrheim DN, Crowcroft NS, Strebel PM. 2014 Measles the epidemiology of elimination. *Vaccine* **32**, 6880–6883. (doi:10.1016/j.vaccine.2014.10.061)
13. Goodson JL, Masresha BG, Wannemuehler K, Uzicanin A, Cochi S. 2011 Changing epidemiology of measles in Africa. *J. Infect. Dis.* **204**, S205–S214. (doi:10.1093/infdis/jir129)
14. Takahashi K, Kanda H, Kim JY. 2014 Reasons for non-vaccination among patients who acquired measles: lessons from the local measles epidemics in Japan. *West Indian Med. J.* **63**, 647–649. (doi:10.7727/wimj.2013.310)
15. Merl D, Johnson LR, Gramacy RB, Mangel M. 2009 A statistical framework for the adaptive management of epidemiological interventions. *PLoS ONE* **4**, e5807. (doi:10.1371/journal.pone.0005807)
16. Shea K, Tildesley MJ, Runge MC, Fonnesebeck CJ, Ferrari MJ. 2014 Adaptive management and the value of information: learning via intervention in epidemiology. *PLoS Biol.* **12**, e1001970. (doi:10.1371/journal.pbio.1001970)
17. Hutchins SS *et al.* 2004 Evaluation of the measles clinical case definition. *J. Infect. Dis.* **189**, S153–S159. (doi:10.1086/379652)
18. Ho HJ, Low C, Ang LW, Cutter JL, Tay J, Chan KP, Ooi PL, Thoon KC, Goh KT. 2014 Progress towards measles elimination in Singapore. *Vaccine* **32**, 6927–6933. (doi:10.1016/j.vaccine.2014.10.046)
19. Guy RJ, Andrews RM, Kelly HA, Leydon JA, Riddell MA, Lambert SB, Catton MG. 2004 Mumps and rubella: a year of enhanced surveillance and laboratory testing. *Epidemiol. Infect.* **132**, 391–398. (doi:10.1017/S0950268804001955)
20. Dietz V, Rota J, Izurieta H, Carrasco P, Bellini W. 2004 The laboratory confirmation of suspected measles cases in settings of low measles transmission: conclusions from the experience in the Americas. *Bull. World Health Organ.* **82**, 852–857.
21. Chiew M, Gidding HF, Dey A, Wood J, Martin N, Davis S, McIntyre P. 2013 Estimating the measles effective reproduction number in Australia from routine notification data. *Bull. World Health Organ.* **92**, 171–177. (doi:10.2471/blt.13.125724)
22. Ferrari MJ, Bjørnstad ON, Dobson AP. 2005 Estimation and inference of  $R_0$  of an infectious pathogen by a removal method. *Math. Biosci.* **198**, 14–26. (doi:10.1016/j.mbs.2005.08.002)
23. Lipsitch M. 2003 Transmission dynamics and control of severe acute respiratory syndrome. *Science* **300**, 1966–1970. (doi:10.1126/science.1086616)
24. Minetti A, Hurtado N, Grais RF, Ferrari M. 2013 Reaching hard-to-reach individuals: nonselective versus targeted outbreak response vaccination for measles. *Am. J. Epidemiol.* **179**, 245–251. (doi:10.1093/aje/kwt236)
25. Ferrari MJ, Grenfell BT, Strebel PM. 2013 Think globally act locally: the role of local demographics and vaccination coverage in the dynamic response of measles infection to control. *Phil. Trans. R. Soc. B* **368**, 20120141. (doi:10.1098/rstb.2012.0141)
26. Takahashi S, Metcalf CJE, Ferrari MJ, Moss WJ, Truelove SA, Tatem AJ, Grenfell BT, Lessler J. 2015 Reduced vaccination and the risk of measles and other childhood infections post-Ebola. *Science* **347**, 1240–1242. (doi:10.1126/science.aaa3438)
27. Camargo MCC, de Moraes KC, Souza VAUF, Matos MR, Pannuti CS. 2000 Predictors related to the occurrence of a measles epidemic in the city of São Paulo in 1997. *Revista Panamericana de Salud Pública* **7**, 359–365. (doi:10.1590/S1020-4989200000600001)
28. Minetti A *et al.* 2013 Lessons and challenges for measles control from unexpected large outbreak, Malawi. *Emerg. Infect. Dis.* **19**, 202–209. (doi:10.3201/eid1902.120301)
29. Rebecca Prevots D. 2003 Interruption of measles transmission in Brazil 2000–2001. *J. Infect. Dis.* **187**, S111–S120. (doi:10.1086/368030)
30. Leite RD, Barreto JLTMS, Sousa AQ. 2015 Measles reemergence in Ceará Northeast Brazil 15 years after elimination. *Emerg. Infect. Dis.* **21**, 1681–1683. (doi:10.3201/eid2109.150391)
31. Uzicanin A, Zimmerman L. 2011 Field effectiveness of live attenuated measles-containing vaccines: a review of published literature. *J. Infect. Dis.* **204**, S133–S149. (doi:10.1093/infdis/jir102)
32. Pannuti CS, Moraes JC, Souza VA, Camargo MC, Hidalgo NT. 1991 Measles antibody prevalence after mass immunization in São Paulo, Brazil. *Bull. World Health Organ.* **69**, 557–560.
33. Mossong J *et al.* 2008 Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS Med.* **5**, e74. (doi:10.1371/journal.pmed.0050074)
34. Prem K, Cook AR, Jit M. 2017 Projecting social contact matrices in 152 countries using contact surveys and demographic data. *PLoS Comput. Biol.* **13**, e1005697. (doi:10.1371/journal.pcbi.1005697)
35. Patil A, Huard D, Fonnesebeck CJ. 2010 PyMC: Bayesian stochastic modelling in Python. *J. Statist. Softw.* **35**, 1–81. (doi:10.18637/jss.v035.i04)
36. Geyer C. 2011 Introduction to Markov chain Monte Carlo. In *Handbook of Markov chain Monte Carlo* (eds S Brooks, A Gelman, G Jones, X-L Meng), pp. 3–48. London, UK: Chapman & Hall/CRC.
37. Geweke J, Berger JO, Dawid AP. 1992 Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In *Bayesian statistics 4* (eds JM Bernardo, JO Berger, AP Dawid, AFM Smith), pp. 169–193. Oxford, UK: Oxford University Press.
38. Edmunds WJ, Gay NJ, Kretzschmar M, Pebody RG, Wachmann H. 2000 The pre-vaccination epidemiology of measles mumps and rubella in Europe: implications for modelling studies. *Epidemiol. Infect.* **125**, 635–650. (doi:10.1017/S095026880004672)
39. Mossong J, Muller CP. 2000 Estimation of the basic reproduction number of measles during an outbreak in a partially vaccinated population. *Epidemiol. Infect.* **124**, 273–278. (doi:10.1017/S0950268899003672)
40. Conlan AJK, Grenfell BT. 2007 Seasonality and the persistence and invasion of measles. *Proc. R. Soc. B* **274**, 1133–1141. (doi:10.1098/rspb.2006.0030)
41. Keeling MJ, Grenfell BT. 1997 Disease extinction and community size: modeling the persistence of measles. *Science* **275**, 65–67. (doi:10.1126/science.275.5296.65)
42. de Moraes JC, Camargo MCC, de Mello MLR, Hersh BS, Glasser JW. 2016 The 1997 Measles outbreak in Metropolitan São Paulo Brazil: strategic implications of increasing urbanization. In *Mathematical and statistical modeling for emerging and re-emerging infectious diseases* (eds G Chowell, J Hyman), pp. 269–289. Berlin, Germany: Springer. (doi:10.1007/978-3-319-40413-4\_16)
43. Zheng X *et al.* 2015 Investigation of a measles outbreak in China to identify gaps in vaccination coverage routes of transmission, and interventions. *PLoS ONE* **10**, e0133983. (doi:10.1371/journal.pone.0133983)
44. Bhattacharyya S, Ferrari MJ. 2017 Age-specific mixing generates transient outbreak risk following critical-level vaccination. *Epidemiol. Infect.* **145**, 12–22. (doi:10.1017/S0950268816002016)