Estimation of measles vaccine efficacy and critical vaccination coverage in a highly vaccinated population

Michiel van Boven^{1,*}, Mirjam Kretzschmar^{1,2}, Jacco Wallinga^{1,2}, Philip D O'Neill³, Ole Wichmann⁴ and Susan Hahné¹

¹*Centre for Infectious Disease Control, National Institute for Public Health and the Environment, PO Box 1, 3720 Bilthoven, The Netherlands*

²*Julius Center for Health Research & Primary Care, University Medical Center Utrecht, PO Box 85500, 3508 Utrecht, The Netherlands*

³*School of Mathematical Sciences, University of Nottingham, Nottingham NG7 2RD, UK*

⁴*Department for Infectious Disease Epidemiology, Robert Koch-Institute, DGZ-Ring 1, 13086 Berlin, Germany*

Measles is a highly infectious disease that has been targeted for elimination from four WHO regions. Whether and under which conditions this goal is feasible is, however, uncertain since outbreaks have been documented in populations with high vaccination coverage (more than 90%). Here, we use the example of a large outbreak in a German public school to show how estimates of key epidemiological parameters such as the basic reproduction number (R_0), vaccine efficacy (VE_s) and critical vaccination coverage (p_c) can be obtained from partially observed outbreaks in highly vaccinated populations. Our analyses rely on Bayesian methods of inference based on the final size distribution of outbreak size, and use data which are easily collected. For the German public school the analyses indicate that the basic reproduction number of measles is higher than previously thought ($\hat{R}_0 = 30.8$, 95% credible interval: 23.6–40.4), that the vaccine is highly effective in preventing infection ($VE_s = 0.997$, 95% credible interval: 0.993–0.999), and that a vaccination coverage in excess of 95 per cent may be necessary to achieve herd immunity ($\hat{p}_c = 0.971$, 95% credible interval: 0.961–0.978). We discuss the implications for measles elimination from highly vaccinated populations.

Keywords: measles; vaccination; reproduction number; vaccine efficacy; Bayesian inference; MCMC

1. INTRODUCTION

Measles virus has been targeted for elimination by the World Health Organization (WHO) from four of the six WHO regions (the Americas, Europe, Eastern Mediterranean, Western Pacific). To date this goal arguably has only been achieved in the American region. A key reason for this is the fact that measles is among the most transmissible of infectious diseases (Anderson & May 1991). Therefore, successful elimination hinges on the availability and use of an effective vaccine, and a sustained very high vaccination coverage.

The question arises as to how high the vaccination coverage should be to ensure that no major outbreaks can occur. Early estimates of measles basic reproduction numbers based on population-based serological studies and data on the average age at infection are in the range 5–18 (Anderson & May 1991, p. 70). For a vaccine that provides complete protection against infection epidemiological theory states that the critical vaccination coverage in a homogeneously

mixing population is given by $p_c = 1 - 1/R_0$, where R_0 is the basic reproduction number and p_c is the critical vaccination coverage. Hence, early estimates suggest that the critical vaccination coverage should be at least 80–94%. Based on this, the WHO recommends that at least 95 per cent of children should receive two doses of measles virus containing vaccine (MVCV) with the first dose in the first year of life (WHO 2003).

Even though measles vaccines are known to be highly effective in preventing infection, it is also known that occasionally vaccinated persons may be infected (Yeung *et al.* 2005; Muscat *et al.* 2009). This is of importance since it could imply that in effect vaccination coverage may need to be higher than predicted above. Furthermore, recent analyses suggest that transmissibility of measles virus may be even higher than early estimates indicate (Wallinga *et al.* 2003, 2005). This finding is supported by the observation that outbreaks have been recorded in populations with high vaccination coverage (Sutcliffe & Rea 1996; Paunio *et al.* 1998; Lynn *et al.* 2004; Yeung *et al.* 2005; Ong *et al.* 2007; table 1).

*Author for correspondence (michiel.van.boven@rivm.nl).

Table 1. Overview of measles outbreaks in schools with a high prior vaccination coverage.

outbreak (place, year)	vaccination coverage ^a		infected			reference
	≥1 dose	2 doses	unvaccinated	vaccinated once	vaccinated twice	
Honkajoki, Finland, 1989	338/417 (81%)	149/417 (35%)	18/79	11/189	6/149	Paunio <i>et al.</i> (1998)
Toronto, Canada, 1995	1063/1135 (94%)	116/1135 (10%)	1/39	82/947	2/116	Sutcliffe & Rea (1996)
Alaska, USA, 1998	3653/3679 (99%)	1781/3679 (48%)	0/8	17/1872	1/1781	Lynn <i>et al.</i> (2004)
Pennsylvania, USA, 2003	655/663 ^b (99%)	629/663 ^b (95%)	2/3	0/26	6/627	Yeung <i>et al.</i> (2005)
Singapore, 2004	171/184 ^c (93%)		7/13		2/171 ^c	Ong <i>et al.</i> (2007)

^aVaccination status at the start of the outbreak.
^bDuring the outbreak 33 students were vaccinated.
^cNo information as to whether this concerns one or two doses.

To obtain further insight into the critical vaccination coverage required to prevent outbreaks, we analyse data from a large outbreak of measles that occurred in a highly vaccinated population of students at a German public school (Wichmann *et al.* 2007, 2009). From this outbreak information is available on the vaccination and infection status of most students. The data have been analysed earlier to estimate vaccine efficacy by the cohort method (Wichmann *et al.* 2007, 2009). Here we reanalyse the outbreak using methods that are based on explicit susceptible-exposed-infectious-recovered (SEIR) type epidemic models. In these models the parameters to be estimated have a clear-cut biological interpretation (Becker *et al.* 2003; Becker & Britton 2004; Demiris & O'Neill 2005). In comparison with the classical approach that is based on estimation of the relative risk of infection in the vaccinated versus unvaccinated subpopulations this has the advantage that it is possible to meaningfully average over populations, and to extrapolate the findings to other situations (larger or smaller populations, populations with different contact patterns).

We estimate two key epidemiological parameters, the basic reproduction number and vaccine efficacy for susceptibility, using a framework that is based on the distribution of the outbreak final size, i.e. the number of persons that have become infected at the end of the outbreak. Since the vaccination status of a sizeable fraction of the population is unknown, we use Bayesian methods of analysis in which the unknown vaccination statuses are added as latent variables. Together, estimates of the basic reproduction number and vaccine efficacy for susceptibility allow us to estimate the critical vaccination coverage. An advantage of our analyses over previous studies on the critical vaccination coverage of measles is that by focusing on an outbreak in a population that is already highly vaccinated our analyses need not extrapolate far beyond the study population (i.e. from a study population with low vaccination coverage to populations with high vaccination coverage).

2. MATERIAL AND METHODS

2.1. Outbreak description

In 2006, a large measles outbreak occurred in Germany in the densely populated state of North Rhine-Westphalia. Over a 48-week period a total of 1749 measles cases were notified (Wichmann *et al.* 2009). A retrospective cohort study was conducted at a public school in Duisburg, where the first cluster of cases was observed, to estimate vaccination coverage and vaccine effectiveness (Wichmann *et al.* 2007). A total of 1250 students between 10 and 18 years of age attended the school. Questionnaires were returned by 1098 students, and vaccination records were abstracted from 859 students, of whom 820, 605 and 39 had received at least one dose, two doses and no dose of MVCV. A total of 53 measles cases were identified using the standard clinical case definition (WHO 2003). After excluding students with an earlier history of measles ($n = 62$) and students who were vaccinated during the outbreak ($n = 22$), data from 1014 students were used for further analysis. The attack rate among unvaccinated students was 52.8 per cent (19 of 36), 1 per cent among students with one MVCV dose (two of 199), 0.4 per cent among students with two MVCV doses (two of 561) and 13.8 per cent among students with unknown vaccination status (30 of 218). Vaccine efficacy was calculated using the cohort method (Orenstein *et al.* 1985), and based only on students with documented vaccination status ($n = 859$). Vaccine efficacy was estimated at 0.981 (95% CI: 0.92–1) in students with one and 0.994 (95% CI: 0.97–1) in students with two MVCV doses (Wichmann *et al.* 2007).

During the school outbreak, only a public awareness campaign was conducted to inform parents about the outbreak. No vaccination campaign was launched, due to limited personnel at the district public health office and to the hope that transmission would stop during the Easter holidays (Wichmann *et al.* 2009). The above study was launched a week before the Easter holidays when the outbreak at this particular school was already over, although it continued in the city another 18 weeks.

Based on the above, persons were classified as being either unvaccinated (36), vaccinated at least once (782), previously infected with measles virus (62), or having unknown vaccination status (370). Among the 152 students who did not return the questionnaire and did not participate in the study, two additional measles cases were identified through the routine national surveillance system (Wichmann *et al.* 2007). Thus, the total number of infections at the school was 19 (unvaccinated), four (vaccinated), zero (previously infected) and 32 (unknown vaccination status). Table 2 gives an overview of the data.

2.2. Statistical analysis

Throughout, our methods of inference make use of the outbreak final size, i.e. the number of individuals that are ultimately infected. Methods for the analysis of final size distributions of SEIR type epidemic models are now well developed (Ball & O'Neill 1999; Demiris & O'Neill 2005). Here, we consider a three-type SEIR (susceptible-exposed-infected and infectious-recovered and immune) model in which the three types represent unvaccinated, vaccinated (either once or twice) and previously infected individuals. We assume that previously infected individuals are unavailable for infection, although they do take part in the contact process. We further assume that infectious contacts are only made at the level of the school, and not at other organizational levels such as class, household or community.

In the model individuals make contacts at a rate that is inversely proportional to the total school size N . Specifically, an infected person (unvaccinated or vaccinated) makes infectious contacts with each unvaccinated individual according to a Poisson process of rate λ/N , and with each vaccinated individual according to a Poisson process of reduced rate $(1 - \text{VE}_S)\lambda/N$, with all such Poisson processes being mutually independent. Hence, vaccinated and unvaccinated individuals are assumed to be equally infectious (since the data do not allow one to distinguish between variable infectiousness and the infection rate parameter, see §4), and the vaccine is assumed to reduce the probability of infection per infectious contact by a factor $1 - \text{VE}_S$ ($0 \leq \text{VE}_S \leq 1$), where VE_S represents vaccine efficacy for susceptibility (Basta *et al.* 2008).

Final size data contain no explicit temporal information and so in our analysis time can be arbitrarily rescaled (Ball & O'Neill 1999). Hence we may, without loss of generality, take the mean infectious period to be 1, and equate the contact rate parameter λ with the basic reproduction number R_0 in a fully susceptible unvaccinated population (Ball & O'Neill 1999). The choice of infectious period distribution is discussed later. We may therefore focus on a relatively simple model containing two parameters, namely the basic reproduction number R_0 and vaccine efficacy for susceptibility VE_S . In this context the reproduction number in a large population with vaccination coverage p , R_p , takes a simple form, and it is given by

$$R_p = R_0(1 - p\text{VE}_S) \quad (2.1)$$

Table 2. Data of the measles outbreak (after Wichmann *et al.* 2007).

	infected	total
unvaccinated	19	36
vaccinated ^a	4	782
previously infected	0	62
vaccination status not known	32	370

^a199 were vaccinated once. Of these, two were infected.

(Diekmann & Heesterbeek 2000). The critical vaccination coverage p_c above which no major outbreaks can occur is readily found by putting $R_p = 1$ and solving equation (2.1) for $p = p_c$, yielding $p_c = \text{VE}_S^{-1}(1 - R_0^{-1})$. Note that equation (2.1) applies exactly only in an infinitely large population. In a population of finite size one needs to substitute the parameters of the model into the final size equations to calculate the expected number of secondary infections (van Boven *et al.* 2007). We have carried out such calculations to ensure that for our study population, equation (2.1) provides an accurate approximation of the expected number of secondary cases during the first few infection generations in a highly vaccinated population (results not shown).

If complete infection and vaccination data are available then a natural method of analysis would be to maximize the likelihood with respect to the parameters R_0 and VE_S . We used this method to check the results of the Bayesian analyses described below by making assumptions on the persons with missing vaccination status. Since our population is large (1250 persons), the standard recursive method for calculating the final outcome distribution suffers from severe numerical instabilities which arise due to the presence of extremely small probabilities (Demiris & O'Neill 2006). It is therefore necessary to use very high-precision arithmetic. We calculated the values of the likelihood function on a $R_0 - \text{VE}_S$ grid of at least 300 grid points, where each point was calculated with 3000 digit precision. Subsequently, the likelihood values were transformed into a smooth likelihood surface by interpolation between the grid-points. To ascertain that our procedures yielded correct final size probabilities we checked (i) that total probability mass of the calculated final size distributions was 1, and (ii) that reducing operational precision from 3000 digits to 2500 digits yielded identical results. CIs and confidence regions were calculated using the profile likelihood (Pawitan 2001).

To take into account that the vaccination status of some persons is missing, we have analysed the data using the random directed graph ('digraph') method (Demiris & O'Neill 2005). In this approach, one does not write down the complete likelihood of the data, but instead focuses on the likelihood of a directed graph of potential infections. In such graphs directed edges (in the sequel also called infectious contacts) between pairs of individuals denote contacts which are sufficient for transmission if one person is infective and the other susceptible. Each such edge from person i to j , say, essentially ensures that j will become infected if i ever does. Note that an individual may have more than

one incoming edge, so the graph does not explicitly keep track of who actually infects whom. Since the data tell us who was actually infected, it is only necessary to construct the graph on the set of infected individuals: the likelihood that the remaining susceptibles are never infected is straightforward to calculate. Any realization of the digraph which disagrees with the data (e.g. the graph gives fewer than the required number of ultimately infected individuals) has zero likelihood.

Full details of the digraph method are given in Demiris & O'Neill (2005). Here we have extended their approach to allow for missing vaccination status of part of the individuals in the population. The vaccination statuses of these individuals are added in the analyses as latent variables.

Details of the Bayesian analyses are as follows. We use a Markov chain Monte Carlo method to explore the joint posterior density of the model parameters and the digraph. Chains are initialized with a digraph that is compatible with the data, in which each infected individual has a directed edge to precisely one other infected individual so that the initial digraph is a tree, and in which the initially infected individual is assumed to be unvaccinated. The infection rate parameter and vaccine efficacy are assigned independent uninformative uniform prior distributions, and are updated using Gaussian proposal distributions with variances of 0.5 and 0.001, respectively ($0 < R_0 < 100$, $0 \leq \text{VE}_S < 1$). In the analyses presented here we have assumed that the infectious period is of fixed duration, and that time is measured in units of the infectious period. Hence, there is no need to update the infectious periods, and the infection rate parameter can directly be equated with the basic reproduction number. We have also performed analyses with exponentially distributed infectious periods, and obtained similar estimates for the basic reproduction number and vaccine efficacy (results not shown). The digraph is updated by proposing to add and delete edges at random from the current digraph (Demiris & O'Neill 2005). Updating of the individuals with unknown vaccination status is performed separately for infected and uninfected individuals. For uninfected individuals (338 in the German dataset) the proposed number of unvaccinated individuals is increased by 1 with probability 0.5 each cycle. Otherwise it is decreased by 1. Further, for the infected individuals the proposed vaccination status of a randomly selected individual is flipped. Updating is performed in blocks, in the order (i) vaccination status of uninfected individuals, (ii) vaccination status of infected individuals, (iii) transmission rate parameter (the basic reproduction number), (iv) adding an edge to the digraph, and (v) deleting an edge from the digraph.

After running a number of exploratory analyses output was generated for a single chain of length 55 000. The first 5000 cycles were discarded as burn-in, and thinning was applied by taking output from each tenth cycle. Inspection of convergence of the chain was performed visually and by comparison of summary statistics of cycles 5000–10 000 with cycles 50 000–55 000 (distribution of the basic reproduction number and vaccine efficacy, distribution of the number of edges of the digraph, distribution of vaccination statuses).

To assess the distribution of outbreak sizes for the estimated parameters, and to evaluate whether the outbreak in the German school is typical for the estimated parameters we performed stochastic simulations using the Sellke construction (Ball & O'Neill 1999). Each simulation run was seeded with a random draw from the posterior distribution of the Markov chain (i.e. basic reproduction number, vaccine efficacy, number of vaccinated individuals). In the simulations the expected infection tolerance level of vaccinated individuals is $1/(1 - \text{VE}_S)$ times that of unvaccinated individuals. In each simulation a single infected individual was introduced into an otherwise susceptible population. All analyses were carried out using MATHEMATICA 7.0.

3. RESULTS

3.1. Bayesian estimation of the critical vaccination coverage

Within 2500 cycles the Markov chain has reached the equilibrium distribution. Figure 1 shows the initial (cycle 0) and final digraph (cycle 55 000). The infected persons with unknown vaccination status have almost invariably become unvaccinated, while the uninfected persons with unknown vaccination status have mostly become vaccinated. In all, the number of persons who are vaccinated ranges from 67 to 75, with relative posterior frequencies of 0.04 (67 vaccinated persons), 0.71 (68 vaccinated persons), 0.18 (69 vaccinated persons), 0.05 (70 vaccinated persons), 0.01 (71 vaccinated persons) and less than 0.002 if the number of vaccinated persons is 72 or more.

The posterior distribution of the number of infectious contacts (edges) differs markedly between vaccinated and unvaccinated persons. In fact, infected vaccinated persons have almost always only one incoming infectious contact (in-edge), while infected unvaccinated persons have 1.9 incoming infectious contacts on average. Based on these figures, the average number of out-edges of unvaccinated persons, which gives an indication of the reproduction number at the start of an outbreak in a vaccinated population, is almost 2.

Figure 2 shows samples from the posterior distribution for the basic reproduction number and vaccine efficacy for susceptibility. The median of the basic reproduction number is 30.7 (95% credible interval: 23.6–40.4) and the median of vaccine efficacy is 0.997 (95% credible interval: 0.993–0.999). Hence, measles virus is highly transmissible, and the vaccine is highly effective in preventing infection per contact that would have caused transmission to an unvaccinated person.

The posterior distribution of the critical vaccination coverage is given in figure 3. The median of the critical vaccination coverage is 0.971 (95% credible interval: 0.961–0.978), indicating that a vaccination coverage in excess of 95 per cent would be required for elimination.

3.2. Maximum likelihood estimation of the critical vaccination coverage

If complete information were available on the vaccination and infection status of the individuals the

Figure 1. Initial ((a), cycle 0) and final ((b), cycle 55 000) directed infection graph (digraph). Arrows or edges denote infectious contacts and nodes represent individuals that are infected. Red and blue nodes correspond to individuals that are known to be unvaccinated (red) or vaccinated (blue), while lighter-shaded nodes represent individuals with unknown vaccination status. The colour of these nodes corresponds with the vaccination status of the individuals in the Markov chain.

outbreak could be analysed by maximization of the likelihood of the final size distribution. To obtain a more intuitive understanding of the results of the Bayesian analyses we reanalysed the data from the German outbreak, assuming that the vaccination statuses of all individuals are known. Motivated by the above results we assume that all infected persons with unknown vaccination status are unvaccinated, while all uninfected persons with unknown vaccination status are vaccinated. This yields a final size of 51/68 (infected/total) unvaccinated and 4/1120 (infected/total) vaccinated persons.

Figure 4 shows maximum likelihood estimates of the basic reproduction number and vaccine efficacy for susceptibility for an infectious period of fixed duration, with associated 95 per cent confidence area. The figure also shows the relation between the basic reproduction number and vaccine efficacy for the lower and upper bound of the 95 per cent confidence limits of the critical vaccination coverage. Maximum likelihood estimates of

Figure 2. Joint posterior distribution of the basic reproduction number (R_0) and vaccine efficacy (VE_s). The posterior median is indicated by a red dot.

the basic reproduction number and vaccine efficacy for susceptibility are 32.1 (95% CI: 23.7–42.7) and 0.9975 (95% CI: 0.9938–0.9992), respectively. For these parameter estimates the most likely value of the critical vaccination coverage is 0.971 (95% CI: 0.961–0.979).

The assumptions of figure 4 are based on the results of the Bayesian analyses. These results indicate that it is unlikely that any of the uninfected persons with unknown vaccination status are actually unvaccinated. The intuitive reason is that if vaccine efficacy is high it is much more probable that an uninfected person with unknown vaccination status is vaccinated rather than unvaccinated. To investigate how the results are affected if some of the uninfected persons with unknown vaccination status are unvaccinated, we have carried out additional analyses in which the vaccination coverage among the 338 uninfected persons with unknown vaccination status corresponds to the vaccination coverage among the persons with known vaccination status. Hence, we assume that the vaccination coverage among the uninfected persons with unknown vaccination status is $36/(36 + 782) = 0.956$, and that $(1 - 0.956) \times 338 \approx 15$ uninfected persons with missing vaccination status are vaccinated.

Figure 5 shows the results of the analyses with more unvaccinated uninfected persons. The maximum likelihood estimate of the basic reproduction number is 22.1 (95% CI: 16.4–29.0), and the maximum likelihood estimate of vaccine efficacy for susceptibility is 0.9963 (95% CI: 0.9908–0.9989). For these parameter estimates the most likely value of the critical vaccination coverage is 0.958 (95% CI: 0.943–0.969).

3.3. Simulated outbreaks

To evaluate whether the observed German outbreak is compatible with the estimated parameters and vaccination statuses, we performed 10 000 simulations of

Figure 3. Relative posterior frequency distribution of the critical vaccination coverage. The median of the critical vaccination coverage is 0.971 (95% credible interval: 0.961–0.978).

outbreaks in a population of 1250 individuals. For each simulation a set of parameters (basic reproduction number, vaccine efficacy, number of vaccinated individuals) was drawn from the posterior distribution.

The results show a bimodal pattern typical for infectious disease outbreaks (figure 6). Despite the fact that the reproduction number exceeds the threshold value of 1, there is a substantial probability that an outbreak will remain small and does not affect more than, say, 20 persons. If, however, the epidemic does take off, then the number of infected persons usually is in the range 35–75 persons, and rarely is it that more than 85 are infected. Interestingly, even though vaccine efficacy is high it is not uncommon to find infected vaccinated persons. For instance, if the number of infected persons is in the range 35–55 the average number of infections in vaccinated persons ranges from 1 to 4, while if the number infected is in the range 56–75 the average number of infected vaccinated persons ranges from 4 to 13. In the unlikely case that more than 75 persons are infected the number of infected vaccinated persons may be more than 15.

4. DISCUSSION

We have obtained estimates of the reproduction number and vaccine efficacy of measles in a highly vaccinated public school. Our results confirm that measles virus is among the most transmissible of infectious diseases (Anderson & May 1991, p. 70), and suggest that it may even be more transmissible than hitherto thought. This indicates that even though estimated vaccine efficacy is high (0.997), this may not be enough to prevent major outbreaks unless the vaccination coverage is nearly 100 per cent. In fact, our results indicate that the vaccination coverage should be at least 0.971 (95% credible interval: 0.961–0.978). This high value

Figure 4. Maximum likelihood estimate of the basic reproduction number (R_0) and vaccine efficacy (VE_S), assuming that 51 out of 68 unvaccinated individuals and 4 out of 1120 vaccinated individuals are infected (red dot). The grey area represents the 95% confidence area, and blue lines correspond to the lower and upper bound of the 95% CI of the critical vaccination coverage. The estimate of the critical vaccination coverage based on the maximum likelihood estimates of the basic reproduction number and vaccine efficacy for susceptibility is 0.971.

is consistent with the experience of elimination in the Americas, where a high vaccination coverage augmented by catch-up campaigns was required for elimination, and it may help explain that outbreaks continue to occur in highly vaccinated populations in other parts of the World (Sutcliffe & Rea 1996; Paunio *et al.* 1998; Lynn *et al.* 2004; Yeung *et al.* 2005; Ong *et al.* 2007).

In our Bayesian analyses the unknown vaccination statuses are included as latent variables in a consistent framework. These analyses indicate that the 32 infected persons with unknown vaccination status are almost always unvaccinated (range: 31–32; median: 32), while the 338 uninfected persons with unknown vaccination status are mostly vaccinated (range: 332–338; median 338; table 2). These findings may seem surprising but can be understood if one realizes that if vaccine efficacy for susceptibility is high, it is far more probable that infected persons are unvaccinated than vaccinated, while the reverse is true for uninfected persons. Nevertheless, to investigate how the results are affected if some of the uninfected persons with unknown vaccination status are unvaccinated we have carried out analyses assuming that vaccination coverage in the uninfected persons with missing vaccination status corresponds to the observed vaccination coverage ($782/(782 + 36) \approx 0.96$; figure 5). The analyses indicate that the basic reproduction number, vaccine efficacy, and critical vaccination coverage all decrease with increasing number of uninfected unvaccinated persons. Still, the estimated critical vaccination coverage remains higher than the vaccination coverage typically achieved in developed countries (0.958; 95% CI: 0.943–0.969).

Figure 5. Maximum likelihood estimate of the basic reproduction number (R_0) and vaccine efficacy (VE_S), assuming that 51 out of 83 unvaccinated individuals and 4 out of 1105 vaccinated individuals are infected (red dot). The grey area and blue lines represent the 95% confidence area and bounds of the 95% CI of the critical vaccination coverage, respectively. The estimate of the critical vaccination coverage is 0.958.

More generally, in our main analyses (figures 2–4) imputation of the missing information is based on the likelihood contributions of the infectious contacts. One could take an alternative approach, based on the assumption that the probability that a person with missing vaccination status is vaccinated is determined by the observed vaccination coverage. For the specific case of the measles outbreak in Germany, however, it is difficult to reconcile the number of infections in the group with unknown vaccination status with the observed number of infections in the vaccinated and unvaccinated groups. To be specific, arguing along these lines one would expect $(36/(36 + 782)) \times 370 \approx 16$ out of 370 persons with unknown vaccination status to be unvaccinated. It is difficult to see how this figure can be matched with the observed number of infections in this group (32) and the infection attack rates in the unvaccinated and vaccinated groups (19/36 and 4/782).

Our definition of vaccine efficacy is based on an explicit epidemiological model, and has a clear-cut biological interpretation (reduction of the probability of infection per contact), making it possible to meaningfully average over populations and to extrapolate beyond the study population. This contrasts with traditional estimates of vaccine efficacy that are based on a comparison of attack rates in vaccinated and unvaccinated individuals, or by simply using the vaccination status of the infected individuals together with the population vaccination coverage (the screening method) (Orenstein *et al.* 1985). Vaccine efficacies estimated by these methods are lacking a clear biological interpretation, making it difficult to interpret the results and to extrapolate to different situations

Figure 6. Frequency distribution of the number of infected individuals in 10 000 simulated outbreaks. In each simulation a set of parameters is drawn from the posterior distribution (basic reproduction number, vaccine efficacy, number of vaccinated individuals). The red parts of the bars refer to the proportion of infected persons that have been vaccinated.

(Becker *et al.* 2003; Becker & Britton 2004; Halloran *et al.* 2010).

We have based our analyses on a single outbreak of measles in a single school, and one could argue that this limits the scope of our results. However, in recent years numerous measles outbreaks have been described in highly vaccinated populations (table 1), and we therefore believe that the critical vaccination coverage for measles virus may have been systematically underestimated. This could be due to the fact that estimates of measles reproduction numbers are often still based on serological studies from the pre-vaccination era. More recent analyses of outbreaks and long-term time series in the highly vaccinated populations of the Netherlands and the UK suggest that the fraction of the population that should be immune to prevent major outbreaks is approximately 96 per cent (Wallinga *et al.* 2003, 2005), and that the basic reproduction number is in the range 21–57 (Finkenstadt & Grenfell 2000; He *et al.* 2010; table 2).

Of course, we do not claim that measles basic reproduction numbers and vaccine efficacies are invariant, and do not differ between populations with differing contact intensities (e.g. households versus schools versus public space), vaccination composition and coverage, vital statistics (birth rates, life expectancy) and other demographic characteristics (e.g. age distribution, gender, socioeconomics). Still, we believe that it is of importance to analyse measles outbreaks of measles in high-contact settings such as households and schools because these may act as a multiplier of infection and driver of epidemics in heterogeneously vaccinated populations.

While we were able to obtain estimates of the efficacy of the vaccine in preventing infection in vaccinated persons, the design of the study makes it difficult to obtain

estimates of the efficacy of the vaccine reducing the infectiousness of infected vaccinated persons (Basta *et al.* 2008). The reason is that with final size information from a single population it is difficult to disentangle the effect of the vaccine in reducing the probability of infection from the effect it may have in reducing infectiousness of infected vaccinated persons. To this end, other schools with different vaccination coverages and different attack rates in the vaccinated and unvaccinated subgroups should have been included. However, since the infection attack rates in vaccinated persons are very low in the German outbreak (Wichmann *et al.* 2009) the precision with which one can estimate the efficacy of the vaccine in reducing the infectiousness of infected vaccinated persons is likely to be poor.

Our finding that infected vaccinated individuals are to be expected despite the fact that vaccine efficacy is very high shows that the effectiveness of a vaccine in a population cannot be judged solely by its (estimated) efficacy, but also depends on the pathogen's transmissibility. If vaccines are judged by their ability to induce herd immunity, then the requirements for what constitutes an adequate vaccine are much more stringent for a highly transmissible infectious disease such as measles than for other, less transmissible, infectious diseases. For instance, while for most infectious diseases a vaccine that is able to prevent transmission in more than 99.5 per cent of infectious contacts will certainly be able to prevent major outbreaks with current vaccination coverages, this is arguably not always the case for measles.

We would like to acknowledge the constructive comments of three anonymous reviewers which helped improve the manuscript.

REFERENCES

- Anderson, R. M. & May, R. M. 1991 *Infectious diseases of humans. Dynamics and control*. Oxford, UK: Oxford University Press.
- Ball, F. & O'Neill, P. D. 1999 The distribution of general final state random variables for stochastic epidemic models. *J. Appl. Probab.* **36**, 473–491. (doi:10.1239/jap/1032374466)
- Basta, N. E., Halloran, M. E., Matrajt, L. & Longini, I. M. 2008 Estimating influenza vaccine efficacy from challenge and community-based study data. *Am. J. Epidemiol.* **168**, 1343–1352. (doi:10.1093/aje/kwn259)
- Becker, N. G. & Britton, T. 2004 Estimating vaccine efficacy from small outbreaks. *Biometrika* **91**, 363–382. (doi:10.1093/biomet/91.2.363)
- Becker, N. G., Britton, T. & O'Neill, P. D. 2003 Estimating vaccine effects on transmission of infection from household outbreak data. *Biometrics* **59**, 467–475. (doi:10.1111/1541-0420.00056)
- Demiris, N. & O'Neill, P. D. 2005 Bayesian inference for stochastic multitype epidemics in structured populations via random graphs. *J. R. Statist. Soc. B* **67**, 731–745. (doi:10.1111/j.1467-9868.2005.00524.x)
- Demiris, N. & O'Neill, P. D. 2006 Computation of final outcome probabilities for the generalized stochastic epidemics. *Statist. Comput.* **16**, 309–317. (doi:10.1007/s11222-006-8320-4)
- Diekmann, O. & Heesterbeek, J. A. P. 2000 *Mathematical epidemiology of infectious diseases. model building, analysis and interpretation*. New York, NY: Wiley.
- Finkenstadt, B. F. & Grenfell, B. T. 2000 Time series modelling of childhood diseases: a dynamical systems approach. *J. R. Statist. Soc.* **49**, 187–205.
- Halloran, M. E., Longini, I. M. & Struchiner, C. J. 2010 *Design and analysis of vaccine studies*. New York, NY: Springer.
- He, D., Ionides, E. L. & King, A. A. 2010 Plug-and-play inference for disease dynamics: measles in large and small populations as a case study. *J. R. Soc. Interface* **7**, 271–283. (doi:10.1098/rsif.2009.0151)
- Lynn, T. V., Beller, M., Funk, E. A., Middaugh, J. P., Ritter, D., Rota, P. A., Bellini, W. J. & Torok, T. J. 2004 Incremental effectiveness of 2 doses of measles-containing vaccine compared with 1 dose among high school students during an outbreak. *J. Infect. Dis.* **189**, S86–S90. (doi:10.1086/377699)
- Muscat, M., Bang, H., Wohlfart, J., Glismann, S. & Mølbak, K. 2009 Measles in Europe: an epidemiological assessment. *Lancet* **373**, 383–389. (doi:10.1016/S0140-6736(08)61849-8)
- Ong, G., Rasidah, N., Wan, S. & Cutter, J. 2007 Outbreak of measles in primary school students with high first dose MMR vaccination coverage. *Singapore Med. J.* **48**, 656–661.
- Orenstein, W. A., Bernier, R. H. & Dondero, T. J. 1985 Field evaluation of vaccine efficacy. *Bull. World Health Organ.* **63**, 1055–1068.
- Paunio, M., Peltola, H., Valle, M., Davidkin, I., Virtanen, M. & Heinonen, O. P. 1998 Explosive school-based measles outbreak: intense exposure may have resulted in high risk, even among revaccinees. *Am. J. Epidemiol.* **148**, 1103–1110.
- Pawitan, Y. 2001 *In all likelihood: statistical modelling and inference using likelihood*. Oxford, UK: Oxford University Press.
- Sutcliffe, P. A. & Rea, E. 1996 Outbreak of measles in a highly vaccinated secondary school population. *Can. Med. Assoc. J.* **155**, 1407–1413.
- van Boven, M., Koopmans, M., Du Ry van Beest Holle, M., Meijer, A., Klinkenberg, D., Donnelly, C. A. & Heesterbeek, J. A. P. 2007 Detecting emerging transmissibility of avian influenza virus in human households. *PLoS Comput. Biol.* **3**, e145. (doi:10.1371/journal.pcbi.0030145)
- Wallinga, J., Teunis, P. & Kretzschmar, M. 2003 Reconstruction of measles dynamics in vaccinated populations. *Vaccine* **21**, 2643–2650. (doi:10.1016/S0264-410X(03)00051-3)
- Wallinga, J., Heijne, J. C. & Kretzschmar, M. 2005 A measles epidemic threshold in a highly vaccinated population. *PLoS Med.* **2**, e316. (doi:10.1371/journal.pmed.0020316)
- WHO 2003 Strategic plan for measles and congenital rubella infection in the European Region of WHO. www.euro.who.int/document/e81567.pdf.
- Wichmann, O., Hellenbrand, W., Sagebiel, D., Santibanez, S., Ahlemeyer, G., Vogt, G., Siedler, A. & Van Treeck, V. 2007 Large measles outbreak at a German public school, 2006. *Pediatr. Infect. Dis. J.* **26**, 782–786. (doi:10.1097/INF.0b013e318060aca1)
- Wichmann, O., Siedler, A., Hellenbrand, W., Sagebiel, D., Santibanez, S., Mankertz, A., Vogt, G., van Treeck, U. & Krause, G. 2009 Further efforts needed to achieve measles elimination in Germany—results of an outbreak investigation. *Bull. World Health Organ.* **87**, 108–115. (doi:10.2471/BLT.07.050187)
- Yeung, L. F., Lurie, P., Dayan, G., Eduardo, E., Britz, P. H., Redd, S. B., Papania, M. J. & Seward, J. F. 2005 A limited measles outbreak in a highly vaccinated US boarding school. *Pediatrics* **116**, 1287–1291. (doi:10.1542/peds.2004-2718)