The impact of past vaccination coverage and immunity on pertussis resurgence

Matthieu Domenech de Cellès,1,2,* Felicia M. G. Magpantay,1,3 Aaron A. King,1,4,5† Pejman Rohani6,7,8†

The resurgence of pertussis over the past decades has resulted in incidence levels not witnessed in the United States since the 1950s. The underlying causes have been the subject of much speculation, with particular attention paid to the shortcomings of the latest generation of vaccines. We formulated transmission models comprising competing hypotheses regarding vaccine failure and challenged them to explain 16 years of highly resolved incidence data from Massachusetts, United States. Our results suggest that the resurgence of pertussis is a predictable consequence of incomplete historical coverage with an imperfect vaccine that confers slowly waning immunity. We found evidence that the vaccine itself is effective at reducing overall transmission, yet that routine vaccination alone would be insufficient for elimination of the disease. Our results indicated that the core transmission group is schoolchildren. Therefore, efforts aimed at curtailing transmission in the population at large, and especially in vulnerable infants, are more likely to succeed if targeted at schoolchildren, rather than adults.

INTRODUCTION

Pertussis, the highly transmissible respiratory infection caused primarily by the bacterium Bordetella pertussis (1), has unexpectedly reemerged in several countries with histories of sustained high vaccine coverage (2, 3). Responsible for roughly 195,000 infant mortalities annually, mostly in the developing world (4, 5), the recent spate of infant deaths in industrialized countries, such as the United States and the United Kingdom (6, 7), has emphasized the growing threat. In the United States, routine vaccination beginning in the 1940s led to a 100-fold reduction in pertussis incidence, and the disease appeared to be on the road to elimination (8). However, since the mid-1970s, the disease has made a surprising comeback (9), steadily increasing in incidence to 15.1 cases per 100,000 in 2012 (10). The latest U.S. Centers for Disease Control and Prevention estimates indicate that 20,762 individuals contracted the disease in 2015, including 2709 cases and three deaths in infants under 1 year of age (11), the population most at risk of severe complications (12). A variety of hypothetical explanations for the resurgence have been advanced, but its causes remain the source of much contention (2, 13, 14), reflecting long-acknowledged but poorly understood complexities of pertussis transmission and immunity (15). Increasingly, attention has focused on vaccine immunity as a driver of disease severity, transmission, and pathogen evolution (16–22). We sought to elucidate the nature of vaccine failure and quantify vaccine protection by harnessing the information present in epidemiological time series. Specifically, we formulated mechanistic models expressing the full range of hypothetical vaccine failure modes and fitted these to 16 years’ worth of age-stratified incidence data provided by the Massachusetts Department of Public Health (MDPH) (Fig. 1A and table S1). We examined three, not mutually exclusive, potential modes of failure (23–25): (i) Primary vaccine failure is the failure of the vaccine to “take” in some recipients; that is, some fraction of those vaccinated receive full protection, whereas the remainder receive none. (ii) Failure in duration is the waning of vaccine-induced protection with time (26). (iii) Failure in degree of protection, also known as vaccine “leakiness,” occurs when vaccine-induced protection is imperfect (22, 27), potentially due to antigenic evolution in the pathogen. We adopted the epidemiological structure proposed in (28) (fig. S1), which explicitly allows infections in naïve and vaccinated hosts to differ both in transmissibility and in disease severity (and hence observability). We additionally incorporated age-specific immunization and demographic data (figs. S2 and S3), age-specific contact network information (fig. S4, [29]), and age-specific reporting efficiencies [table S2, (30)] into the model.

RESULTS

Incidence data recapitulate reemergence of pertussis in Massachusetts

A timeline of pertussis surveillance in Massachusetts is presented in table S1. Notably, serological testing became available in 1987 for individuals ≥11 years of age, leading to an immediate and substantial increase in the number of reported cases in individuals 11 to 19 and ≥20 years of age (31). By contrast, the introduction of polymerase chain reaction (PCR) testing in January 2005 did not appear to have an immediate impact, with no noticeable increase in the number of reports in 2005 (fig. S5). In 2006, a reduced-dose booster of acellular pertussis vaccine combined with tetanus and diphtheria toxoids (Tdap) was recommended for adolescents 11 to 18 years old (32). We therefore restricted our analysis to data during 1990–2005 (2005 included, n = 16 years of monthly data), a period of stable surveillance before the introduction of Tdap.

The age-specific incidence records from the active surveillance program in Massachusetts during that period display the pattern typical of pertussis resurgence (Fig. 1 and figs. S5 to S7 (31, 33)). In particular, epidemics have increased in both size and frequency between 1990 and 2005, with a fourfold rise in overall incidence (a relative increase of 9.7%...
per year; table S3). Adolescents 10 to 20 years old accounted for more than 50% of these cases (Fig. 1, Band C, and table S3). The trend in adults was even more pronounced: Cases in those 20 years of age and older increased more than 16% per year and more than 10-fold over the 15-year interval. By contrast, trends were less obvious in children under the age of 10, and no increase was observed in infants under 1 year of age, despite the persistently high incidence rate in that age group (average, 57 cases per 100,000 per year). These trends continue the pattern of resurgence that began in the mid-1970s (9), and no effects of the switch from whole-cell to acellular vaccines in 1996 are immediately evident. The resulting age-specific incidence profile presents two peaks of comparable magnitude in infants and in adolescents (Fig. 1B). Conclusions regarding age-specific disease burden, however, must take into account the differential sensitivities of ascertainment methods used in different age groups. Specifically, the serological enzyme-linked immunosorbent assay used to identify infections in adolescents and adults is considerably more sensitive, and less specific, than the culture-based ascertainment methods used in children under the age of 11 (31).

**Resurgence of pertussis is linked to incomplete historical coverage and slowly waning vaccine immunity**

Using these data and likelihood-based inference methods ([34–37], Model formulation in Materials and Methods, and tables S4 and S5), we weighed the evidence for four alternative hypotheses of vaccine immunity after initial vaccine take: (i) Vaccine protection is perfect in both duration and degree (“no loss model”), (ii) vaccine-derived immunity is perfect in degree but transient (“waning model”), (iii) protection is permanent but imperfect in degree (“leaky model”), and (iv) vaccine immunity is imperfect in both degree and duration (“waning + leaky model”). In each of these, we allowed some primary vaccine failure (that is, failure in take) to be estimated along with the other parameters. We assumed that infection-derived immunity is perfect. The waning model received substantially higher support [as quantified by the Akaike Information Criterion (AIC)] than other models (ΔAIC > 140; Table 1, see also tables S6 and S7 and fig. S8). The best-fitting model predicts a primary vaccine failure probability of 4% [95% confidence interval (CI), 1 to 8%]. Under this model, vaccine protection wanes slowly on average; however, there is substantial variability among individuals. Specifically, there is a 10% risk (CI, 3 to 19%) of protection waning to zero within 10 years of completing routine vaccination and a 55% chance that protection remains lifelong. The results further suggest that post-vaccine infections (defined as infections in individuals in whom the vaccine took but whose immunity subsequently waned) are as transmissible as, but less visible than, naïve infections [relative transmissibility, 0.99 (CI, 0.40 to 1.00); relative observability, 0.39 (CI, 0.19 to 1.00)]. This finding is consistent with evidence from animal challenge studies.
(21, 22), although wide confidence intervals preclude definitive conclusions on this question. Also consistent with previous evidence (31), we estimated high detection rates of pertussis in adolescents and in adults, with 24% (CI, 10 to 66%) of post-vaccination infections reported. Although the waning + leaky model allows for a mixture of all three vaccine failure modes, its leakiness parameter is estimated at 0 (CI, 0 to 3%), and its waning rate is identical to that of the waning model. Thus, the additional complexity of this model is not supported by the data (ΔAIC = −2; Table 1).

To quantify vaccine effectiveness in reducing transmission, we computed \( R_0 \), the vaccine impact (28, 38), a population-wide summary measure accounting for all modes of vaccine failure. The estimated vaccine impact was 0.85 (CI, 0.75 to 0.93) for the waning model and similarly high for the leaky model [0.90 (CI, 0.81 to 0.95); leakiness, 0.06 (CI, 0.02 to 0.14); primary vaccine failure rate, 0.06 (CI, 0.02 to 0.14)] and the no-loss model [0.85 (CI, 0.70 to 0.95)]. The vaccine impact modifies the theoretical vaccination effort required for eradication according to \( \frac{1}{\psi} \left( 1 - \frac{1}{R_0} \right) \), where \( R_0 \) is the basic reproductive ratio (39). One of our key results, therefore, is that despite the effectiveness of the vaccine, eradication via routine immunization alone is not possible given the relatively large estimated value of \( R_0 \) [10.1 (CI, 6.5 to 17.2)].

### Table 1. Model comparison.

The maximum likelihood estimates (95% CI) are presented for the stochastic variant of the models, estimated using the MLE algorithm. The best AIC value is indicated in boldface. SE, standard error.

<table>
<thead>
<tr>
<th>Quantity</th>
<th>No-loss model</th>
<th>Leaky model</th>
<th>Waning model</th>
<th>Waning+Leaky model</th>
</tr>
</thead>
<tbody>
<tr>
<td>log ( \ell )</td>
<td>−3726.9 (SE: 0.4)</td>
<td>−3664.9 (SE: 0.9)</td>
<td>−3594.5 (SE: 0.6)</td>
<td>−3598.5* (SE: 1.8)</td>
</tr>
<tr>
<td>AIC</td>
<td>7474</td>
<td>7356</td>
<td>7215</td>
<td>7217*</td>
</tr>
<tr>
<td>ΔAIC</td>
<td>259</td>
<td>141</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>( R_0 )</td>
<td>2.4 (1.8, 2.7)</td>
<td>1.6 (1.3, 2.2)</td>
<td>1.8 (1.5, 2.0)</td>
<td>1.8 (1.6, 2.0)</td>
</tr>
<tr>
<td>Vaccine impact</td>
<td>0.85 (0.70, 0.95)</td>
<td>0.90 (0.81, 0.95)</td>
<td>0.85 (0.75, 0.93)</td>
<td>0.83 (0.70, 0.92)</td>
</tr>
</tbody>
</table>

*Because the two models are nested, the likelihood of the full model should be higher or equal to that of the waning model. The small difference indicates that the MLE of the leakiness for the full model is 0. Consequently, the AIC was calculated with the likelihood of the waning model.

To quantify model-data agreement, we computed \( \Delta \psi \) and \( \Delta R_0 \) for each model. The no-loss model led to a substantial underestimate of \( \Delta R_0 \) compared to the other models (Fig. 2A). We also computed the AIC for each model (Table 1). The waning model was preferred to the leaky model (ΔAIC = 24), with estimates comparable—though more uncertain—to those of the base model (table S9). Hence, we found little evidence for a marked epidemiological effect of the switch to DTaP in Massachusetts in these data, although our results do indicate a moderately reduced efficacy of the DTaP vaccine compared to the wP vaccine (tables S9 and S10).

### Model simulations replicate key aspects of pertussis epidemiology in Massachusetts

To assess the adequacy of the best-fitting model’s explanation of the data, we compared the data to model simulations (Fig. 2A). Despite high variability in simulated dynamics, the model successfully captured key features of the data. In particular, incidence trends in adolescents and in adults were reproduced in simulations; in the younger age groups, where the trends are more obscure, the data were consistent with model simulations. To quantify model-data agreement, we computed a generalized \( R^2 \) for 1-month-ahead forecasts [see Model assessment in Supplementary Materials and Methods, (41)]. The value of 0.35 indicated a modest degree of forecasting skill, with evident underestimation of epidemic peaks in adolescents and adults in years 2000, 2003, and 2004 (Fig. 2B). To set this in context, we carried out a simulation study to assess the intrinsic predictability of pertussis. This revealed that infections like pertussis with generation times exceeding 3 weeks are inherently relatively unpredictable: \( R^2 \) is typically less than 0.45, even in the ideal circumstance that the true model is known (Fig. 2C and figs. S10 and S11).

Our fitted model can exhibit similar prediction skill for longer forecast horizons (figs. S12 and S13). For example, 6-month-ahead forecasts initiated annually in August generate an \( R^2 \) of 0.4 (Fig. 2D). Finally, we examined the model predictions for the seven decades preceding the data interval (1990–2005), a period spanning the late pre-vaccine and early vaccine eras. These hindcasts were consistent with historical incidence data (9, 42), showing an average incidence of 1500 cases per 100,000 with 2- to 4-year cycles in the pre-vaccine era (43), followed by a 100-fold reduction in cases upon introduction of routine vaccination, followed in turn by a resurgence beginning in the 1970s (Fig. 3E). In sum, despite its simplicity, the model both replicates key historical aspects of pertussis epidemiology and displays nontrivial prediction skill.
The increase in adolescent and adult pertussis cases is due to incomplete historical vaccination with slowly waning vaccine-conferred immunity

In epidemiological systems, generally, one cannot directly observe important variables such as population immunity profile or the age distribution of real infections (as opposed to reported cases). However, we can interrogate the fitted model, inquiring as to the likely history of these hidden variables by examining their expected values conditioned on the data (Fig. 3, A to D). These reveal a marked shift in the age-specific immunological profile over this period (Fig. 3, C and D). In particular, as time goes on, we see older individuals become increasingly unlikely to have been infected and, consequently, more likely to be susceptible to infection. According to the model, the introduction of vaccination in the 1940s led to an overall reduction in transmission, reducing the risk of natural infection during childhood not only to those vaccinated but also to the population generally. Those who escaped vaccination as children (or for whom the vaccine did not take) increasingly achieved adulthood having avoided natural infection as well. Concurrently, older cohorts, with their long-lived immunity derived from natural infections experienced during the pre-vaccine period, were gradually dying out. The resulting rise in the number of susceptible adults sets the stage for the pertussis resurgence, especially among adults (44). Thus, the fitted model explains the current pertussis resurgence as a legacy of incomplete vaccination with effective, but imperfect, vaccines against a background of slow demographic turnover, that is, as an end-of-honeymoon effect (20, 45).

To identify the core transmission groups, we performed numerical experiments in which a single pulse of age-targeted vaccination was applied, and we measured the subsequent total incidence of infections in infants 0 to 4 months of age—the population most vulnerable to severe
disease and death (26). We show that booster vaccination of 25% of the adult population (20 to 40 or >40 years of age) leads to a modest decrease in the incidence of pertussis in infants (Fig. 4). By contrast, a similar effort focused on children ages 5 to 10 or 10 to 20 years is predicted to have much higher impact, leading to a drop in infant cases on the order of 25% (Fig. 4, see also fig. S14).

**DISCUSSION**

To place our results within the context of recent pertussis epidemiology, we note that Klein et al. (46) reported a 42% annual increase in the odds of acquiring pertussis after the fifth booster dose, which has been interpreted as evidence for rapid loss of DTaP immunity, in apparent contradiction of our results. To investigate this, we simulated our best-fitting model hindcasts before 1990. Five stochastic realizations of the total incidence (naïve plus post-vaccine infections) are presented. The dotted line at year 1940 indicates the start date of mass vaccination assumed in the model. In the absence of immunization data before 1970, we assume that the vaccine coverage had ramped up between 1940 and 1955 (see Model formulation in Materials and Methods). For visual clarity, the y axis does not start at 0. (F) Comparison of model predictions with empirical studies that quantified DTaP vaccine failure by estimating relative changes (over age) in the odds of acquiring pertussis. We simulated the waning model (with identical wP- and DTaP-derived immunity) during 2006–2015 and used log-linear regression to calculate the yearly relative change (over age) in the odds of acquiring pertussis in children aged 5 to 10 years, that is, 0 to 5 years after receipt of the fifth vaccine dose (see model predictions in the DTaP era in Materials and Methods). The distribution is based on 10^4 simulations, accounting for parametric uncertainty by sampling from the bootstrap distribution. Also presented are estimates from three empirical studies in the United States [Klein et al. (46), Misegades et al. (48), and Tartof et al. (49)] and from a meta-analysis [McGirr and Fisman (47)].
Fig. 4. Predicted impact of single-booster vaccination. Simulations of the waning model were run until the end of 2005, at which point a 25% fraction of susceptible individuals in a target age group (5 to 10, 10 to 20, 20 to 40, or ≥40 years old) was moved to the vaccinated class. The model was run for the subsequent 10 years (2006–2015), and the age-specific total annual infections (that is, naïve and post-vaccine infections, calculated before applying the observation model) were compared to a control scenario without booster vaccination. Each boxplot is based on 10^4 stochastic simulations, accounting for parametric uncertainty of the waning model by sampling parameters from the bootstrap distribution. For each intervention, the number indicates the relative difference between the median simulated incidence and that of the control scenario. The figure shows the predicted impact in unvaccinated infants aged 0 to 4 months, the age group most at risk of severe disease (see fig. S14 for the corresponding figure in every age group). For visual clarity, the y axis does not start at 0.

Our model from 2006 to 2015 and estimated the annual change in the odds of acquiring pertussis. As shown in Fig. 3F (see also model predictions in the DTaP era in Materials and Methods), our model, with its assumption of a slow-waning, high-impact DTaP vaccine, predicts an odds increase of 32% per year, comparable to the Klein et al. study and others (46–49). Thus, not only are our results, perhaps surprisingly, quite consistent with these case-control and cohort studies, but they also show that time series data aggregated at the population scale can be more informative about the quantities of interest than the data from these smaller-scale studies. Our finding calls into question the standard but naïve interpretation of odds ratio/age slopes as quantification of the speed of loss of vaccine-derived immunity: It demonstrates that the slopes observed in case-control and cohort studies are consistent with immunity as long-lived as we estimated from the time series. Further, our fitted model represents an alternative explanation for empirical observations: Increased infection risk in schoolchildren need not be due to rapid decreases in vaccine protection and can arise from increased contact rates following school entry.

Our model makes other testable predictions with consequences for policy. Despite low reported incidence in children aged 5 to 10, the model predicts that true incidence of post-vaccine infections in that age group is comparable to that in adolescents (Fig. 3B). By contrast, the model suggests a much lower rate of infection in adults. Our numerical experiments also indicate that the bulk of transmission occurs in schoolchildren, instead of adults (Fig. 4). Although these results are tempered by the absence of household structure in our model, they are concordant with results from a 2015 U.S. household transmission study, suggesting that the most important immediate source of infection in infants is siblings (50). These findings may help explain the documented failure of postpartum vaccination of immediate family contacts at reducing infant pertussis (51, 52). The lack of a strong transmission link between infants and older age groups is also evident in the asynchrony of pertussis seasonality among these groups (age-specific seasonality in Supplementary Materials and Methods and figs. S6 and S7). Specifically, patterns of seasonality are suggestive of core transmission groups among infants and their young siblings (1 to 5 years old) and among teenagers; incidence in adults does not show a seasonal peak, instead displaying a modest increase in incidence throughout late-summer/early-fall months, perhaps as a result of transmission from the younger age groups. More generally, our results indicate a key role for children and adolescents and at most a minor role for adults in pertussis transmission, perhaps due to the differences in the frequency of contacts at different ages and the assortative structure of the contact network. Although it is beyond the scope of this study, we suggest that the design of optimal vaccine schedule should take these results into account.

There are three caveats of our analysis that are worth noting. First, more precise estimates of vaccine traits may be possible with longer time series. We chose to restrict our attention to the period ending in 2005 to bypass the need to accommodate (i) the introduction of anacellular vaccine booster (Tdap) for teenagers in 2006 and (ii) the switch to PCR as the method of infection ascertainment among adolescents and adults, with concomitant increased sensitivity and diminished specificity (53). Better characterization of differential vaccine traits would need to account for the additional uncertainties associated with these complications. Second, in contrast to some recent modeling studies (16–18), we found little evidence for differences between whole-cell and acellular vaccines. Resolution of this incongruence, however, is thwarted by differences in pertussis vaccination history among study populations, as well as important differences in model structure and statistical estimation choices. Third, we used two quantifications of the age-specific patterns of contact, neither of which is ideal. A better fit was obtained by using self-reported contact data from the POLYMOD study in Great Britain (29), rather than estimates from household survey data from Massachusetts (40). Ideally, the contact matrix would be derived from more direct measures of behavior in the focal population.

With serological correlates of protection as yet unidentified (54), the nature and frequency of pertussis vaccine failure and its role in the resurgence have remained uncertain (14). Here, we have characterized these traits via a rigorous and computationally ambitious high-dimensional
statistical inference approach exploiting age-stratified incidence data. We found that both vaccines induce immunity that, on average, wanes slowly over time, with no evidence for the switch to acellular vaccines as the driver of pertussis resurgence in Massachusetts irrespective of our assumptions regarding the differential efficacy of the D'TaP vaccine. Our results suggest that the train of events leading to the resurgence of pertussis was set in motion well before the shift to the D'TaP vaccine. However, we note that there is substantial heterogeneity among vaccine recipients in terms of the durability of the protection they receive. Crucially, we find that the vaccine is effective at reducing pathogen circulation but not so effective that eradication of this highly contagious bacterium should be possible without targeted booster campaigns (55). In the design of such campaigns, we anticipate that models, such as those presented here, when rigorously confronted with data, will prove to be valuable tools.

MATERIALS AND METHODS

Study design
This study aimed to test a number of hypotheses on the nature and the degree of protection conferred by the whole-cell and acellular pertussis vaccines. We considered whether pertussis vaccines failed to confer immunity in some recipients; whether vaccine-induced immunity waned with time; and whether vaccines may have induced some, but imperfect, protection against the disease. These different hypotheses were tested with dynamic transmission models based on immunization and demographic data from Massachusetts and age-specific daily contact rates from a study in Great Britain. We confronted these models with age-stratified data on pertussis incidence from Massachusetts during 1990–2005.

Immunization data
Immunization levels of children entering kindergarten were available from the MDPH, from school year 1975/1976 for children having received ≥4 doses and from school year 1995/1996 for children having received 5 doses [based on 60,000 to 90,000 annual records, (56)]. As shown in fig. S2, the vaccine coverage was approximately constant during this period. Therefore, we assumed constant vaccine coverages $v_1 = 0.97$ and $v_2 = 0.93$, where $v_1$ represents the vaccine coverage for the primary course and $v_2$ represents the conditional probability of having received a fifth dose given that four doses have been received.

In the absence of vaccination data before 1970, we made pragmatic assumptions based on available evidence. Although mass production of the WP vaccine began in 1950 in Massachusetts (31), it was already distributed across the United States from 1940 (57). This is consistent with historical incidence data from Massachusetts, which show that pertussis began decreasing in the 1940s and steeply declined after 1950 (42). We therefore assumed that vaccination had started in 1940 and that the vaccine coverage ramped up from 0 in 1940 to $v_1$ in 1955. We also assumed that the preschool booster dose began to be administered in 1967, based on the earliest record we could find in the literature (58).

Demographic data
Age-stratified mid-year population estimates in Massachusetts were available from the U.S. Census Bureau for 1990–2005 (59). Data on annual number of births during 1990–2005 were available from (60). These demographic data, plotted in fig. S3, were interpolated using smoothing splines (with 10 degrees of freedom) to calculate the time-varying annual number of births, $B(t)$; the age-stratified population sizes, $N_j(t)$; and the first derivative of the age-stratified population sizes, $\frac{dN_j(t)}{dt}$. These quantities were used to calculate age-specific migration rates, so that the simulated population sizes approximately equaled the actual values (see model equations in Supplementary Materials and Methods).

Contact network data
The model incorporated empirical age-specific contact rates from the POLYMOD study in Great Britain (29), corrected for reciprocity as detailed in the supplementary materials of (20). Let $C_{ij}$ be the average number of daily contacts (both physical and conversational) reported by a participant of age group $i$ with members of age group $j$ [here, individuals are categorized by 5-year age groups from age 0 to age 75, so that $1 \leq (i, j) \leq 15$]. Denoting $N_i$ the number of individuals in age group $i$ in Massachusetts, the average total number of contacts between age groups $i$ and $j$ is $E_{ij} = N_i C_{ij}$. Because of the necessary symmetry in the total number of contacts between age groups, the matrix $E = (E_{ij})$ was made symmetric: $E \leftarrow \frac{1}{2}(E + E^T)$. The individual average number of daily contacts between age groups $i$ and $j$, corrected for reciprocity, was then given by $\forall (i, j)$. $C_{ij} = \frac{E_{ij}}{N_i}$. The corrected matrix $C = (C_{ij})$ was used in all simulations and is plotted in fig. S4.

Incidence data
Monthly pertussis incidence data were available from the MDPH. The data were stratified by age (1-year breakdown for individuals under 20, 5-year breakdown for individuals over 20). For simplicity, we aggregated the data into seven epidemiologically relevant age groups: infants 0 to 1 (that is, [0, 1]) year old; preschool children 1 to 5 years old; school-aged children 5 to 10, 10 to 15, and 15 to 20 years old; and adults 20 to 40 and ≥40 years old. According to the MDPH (http://tinyurl.com/p98mmhvw), the case definition for pertussis in non-outbreak settings is as follows: laboratory confirmation by culture in a patient with any cough illness; a cough illness lasting ≥2 weeks, with laboratory confirmation by serology in a person not vaccinated with a pertussis-containing vaccine in the three previous years; a cough illness lasting ≥2 weeks with one or more of the following: paroxysms of coughing, inspiratory whoop, or post-tussive vomiting, without other apparent cause, in an individual who has a positive PCR test; a cough illness lasting ≥2 weeks with one or more of the following: paroxysms of coughing, inspiratory whoop, or post-tussive vomiting, without other apparent cause, without appropriately timed negative laboratory test, in an individual who is epidemiologically linked to a laboratory-confirmed case.

Model formulation
We implemented an age-stratified, compartmental model of pertussis transmission, building on previously described models (41, 61, 62). The model is an extension of the classic Susceptible Exposed Infected Recovered (SEIR) model that allows for post-vaccination infections in previously vaccinated or infected individuals. The population of susceptibles is divided into those naïve to exposure, $S^{(1)}$, and those whose immune system has been previously primed by vaccination or natural infection, $S^{(2)}$. Exposed and infected individuals are similarly divided into those who experience a naïve infection $[E^{(1)}$ and $I^{(1)}]$ or a post-vaccination infection $[E^{(2)}$ and $I^{(2)}]$. Upon recovery from either type of infection, individuals move to the recovered class, $R$. To account for possible differences between infection- and vaccine-derived immunity, vaccinated individuals are explicitly modeled (V). As previously proposed
(23, 24, 27, 28, 38), we considered two possible modes of failure of infection- or vaccine-derived immunity:

1. Waning (failure in duration): Immunized individuals lose their immunity and become susceptible \([S^{(2)}]\) at a rate \(\alpha_t\) for infection-derived immunity and \(\alpha_V\) for vaccine-derived immunity.

2. Leakiness (failure in degree): Immunized individuals (\(R\) or \(V\)) remain susceptible to a post-vaccination infection \([E^{(2)}]\), but at a lower degree than susceptible individuals. The degree of susceptibility is \(\epsilon_I\) for infection-derived immunity and \(\epsilon_V\) for vaccine-derived immunity.

For vaccine-derived immunity, we additionally considered a third mode of failure, for which, with probability \(\epsilon_{pV}\), vaccinated individuals immediately fail to mount an immune response and move to the \(S^{(1)}\) class [failure in take or primary vaccine failure \((23, 24, 38)\)]. A schematic of the model structure is presented in fig. S1. Individuals are categorized by 5-year age groups from age 5 to age 75; the 0- to 5-year age group is further divided into 1 to 5 years and infants aged 0 to 4 months and 4 to 12 months. The 0- to 4-month age group is included to represent the fact that infants are fully vulnerable to infection before receiving the second dose of DTP at 4 months (41). Overall, the model consists of 17 age groups, labeled \(i = 1, \ldots, 17\). Aging occurs continuously, at rates \(\delta_i = \frac{1}{\Delta a_i}\), where \(\Delta a_i\) is the age span in age group \(i\). To model the effect of the primary vaccination course, a fraction \(V_1\) of susceptible individuals \([S^{(1)}]\) is moved to the vaccinated class on aging from 0 to 4 months to 4 to 12 months. Similarly, the effect of the preschool booster dose is modeled by moving a fraction \(V_2\) of susceptible individuals aging from 1 to 5 to 5 to 10 to the vaccinated class. Because the pediatric booster dose (at age 15 to 18 months) is administered shortly after the primary course, we ignored the effect of this dose.

### Estimated parameters and estimation procedure

To determine the mode of vaccine-derived immunity, we used likelihood-based inference to evaluate the support of three models:

1. No loss of vaccine-derived immunity (no-loss model). After an initial failure in take, vaccine-derived immunity is hypothesized to be perfect, so that no post-vaccine infections are possible. For this model, the estimated parameters were \(\epsilon_{pA}\) (fraction of primary vaccine failures), \(q_{10}, q_{20}, q_{30}\) (susceptibility factors), \(\alpha_{12}^C\) (seasonality coefficients in children aged 5 to 10), \(\alpha_{12}^T\) (seasonality coefficients in adolescents aged 10 to 20), \(\rho_1(10+)\) (reporting probability of naïve infections in \(\geq 10\) years old), and \(\tau\) (reporting overdispersion).

2. Waning vaccine-derived immunity (waning model). After an initial failure in take, vaccine-derived immunity is hypothesized to wane at rate \(\alpha_V\). The estimated parameters are those of the no-loss model, plus \(\alpha_V, \theta\) (transmissibility of post-vaccine infections relative to that of naïve infections), and \(\eta\) (reporting probability of post-vaccine infections relative to that of naïve infections).

3. Leakly vaccine-derived immunity (leaky model). After an initial failure in take, vaccine-derived immunity is hypothesized to be leaky, with degree of leakiness \(\epsilon_V\). The estimated parameters are those of the no-loss model, plus \(\epsilon_V, \theta,\) and \(\eta\).

For each model, the parameters were estimated in two steps:

1. Trajectory matching. The deterministic variant of the model was fitted to the data using maximum likelihood estimation via trajectory matching. In this case, the observation model is the only source of variability in simulated observations, and the likelihood can be calculated exactly. The likelihood was maximized using the subplex algorithm, implemented in the R package nloptr (63). The search was initiated over \(10^6\) starting points generated using Latin hypercube sampling over broad parameter ranges (table S5). To ensure convergence to the maximum likelihood estimate (MLE), the optimization was repeated on the 500 best parameter sets. A parametric bootstrap was then used to assess uncertainty in parameter estimates. For each model, 500 synthetic time series of simulated data were generated at the MLE. For each of these 500 synthetic data sets, parameters were reestimated as described above, resulting in a bootstrap distribution of parameter estimates. We then calculated 95% confidence intervals for the estimated parameters and for the derived parameters \(R_0\) (basic reproduction number), \(R_p\) (vaccine reproduction number), \(q\) (vaccine impact), and \(\rho_1(10+) = \eta\rho_1(10+)\) (reporting probability of post-vaccine infections in \(\geq 10\) years old) from the bootstrap distribution.

2. Maximum iterated filtering (MIF). The stochastic variant of the model was fitted using the MIF algorithm (35), implemented in the R pomp package (36). The following algorithmic parameters were used: 2000 particles, 50 MIF iterations, and a random walk intensity of \(10^{-6}\) during the first MIF iteration and \(10^{-2}\) for the next 49 iterations. Because the model was simulated for a long period before the first data point (January 1990), a time-varying random walk was used for the parameters, with no perturbation until the first data point. For each MIF run, the log-likelihood was computed as the log of the average likelihood of 20 replicate particle filters, each with \(5 \times 10^4\) particles; the SE of the log-likelihood estimate was computed from these replicates using a jackknife implemented in the function logmeanexp in the pomp package. Because each MIF run required about 24 hours of computation, we sought to find good starting parameter values to initiate the algorithm. To do this, we calculated for each model the 95% multivariate confidence interval around the MLEs from trajectory matching. The search was then initiated over 100 starting points generated using Latin hypercube sampling over these ranges. As for trajectory matching, the estimations were repeated from the best parameter sets to ensure convergence to the MLE and a parametric bootstrap was used to generate a bootstrap distribution of size 100.

### Sensitivity analyses

We conducted three sensitivity analyses. First, we fitted an alternative model in which infection- and wP-derived immunities were assumed to be identical (with protected individuals in the \(R\) compartment) but possibly different from DTaP-derived immunity (with protected individuals in the \(V\) compartment). For this analysis, we assumed that infection/wP-derived immunity was waning with an average duration of protection of 75 years (61). During the vaccine transition period, we also assumed, for simplicity, that the first dose of vaccine received determined the nature of subsequent immunity (3). Thus, all infants vaccinated before October 1996 were assumed to be protected by wP, whereas those vaccinated after were assumed to be protected by DTaP. We then repeated the estimations to determine which mechanism of loss of immunity (here specific to DTaP) was best supported by the data (table S9).

Second, we extended the base model (with identical wP- and DTaP-derived immunity) to make the primary vaccine failure time-dependent. Hence, we estimated two parameters for wP and DTaP. The results are presented in table S10 and indicated a higher primary vaccine failure of DTaP, although the absolute difference was modest (0.06 versus 0.03) and the improvement in model fit (compared with the base model) was small (\(\Delta \log L = 2.8\)).

Third, in the absence of empirical contact data in the United States, the results presented in the main text were obtained using the POLYMOD contact matrix in Great Britain. To assess the robustness of our results to this critical assumption, we calculated a contact matrix in Massachusetts using the method described in (40). Briefly, the method uses highly
detailed census and demographic data to build a matrix of “effective” contacts $M = (M_{ij})$. This matrix defines contacts among age groups up to a scaling constant $N_{ma}$ usually absorbed in the transmission rate ($\gamma_0$). In our simulations, we fixed this constant to the total contact rate in the POLYMOD matrix. As shown in fig. S9, the matrix presented fewer intergenerational contacts between children and adults but more contacts between adults than the POLYMOD matrix. Repeating the estimations using that matrix, we found our main results to be robust (table S8).

The waning model was preferred to the leaky model ($\Delta AIC = 153$), with estimates comparable to those obtained with the POLYMOD contact matrix [waning rate, 0.007 (0.001, 0.020) per year; primary vaccine failure, 0.03 (0.01, 0.05); vaccine impact, 0.90 (0.80, 0.96)]. We note, however, that the waning model was less consistent with the data using that matrix ($\Delta \log L = -43.3$, cf. tables S6 and S7).

**Model predictions in the DTaP era**

We sought to compare our model predictions with recent empirical studies in the United States that quantified DTaP vaccinal failure by estimating relative changes (over age) in the odds of acquiring pertussis (46–48). To this end, we extended our base model (with identical wP- and DTaP-derived immunity) to have a higher age resolution (76 age groups overall, 0 to 0.33, 0.33 to 1, 1 to 2, 2 to 3, ..., 74 to 75 years old). For simplicity, we here assumed no demographic changes over time, so the population size remained approximately constant (total population size, $N = 5 \times 10^6$; birth rate, $b = \frac{1}{75}$ per year). To incorporate parametric uncertainty, we ran stochastic simulations at each parameter set from the waning model parametric bootstrap distribution (table S7). For each simulation, we then calculated the overall incidence of post-vaccine infections (including primary vaccine failures) during 2006–2015 (that is, a 10-year period after the last time point used for the fits):

$$H_a = \frac{\sum_{y=2006}^{2015} \epsilon_a H_{a,y}^{(1)} + H_{a,y}^{(2)}}{\sum_{y=2006}^{2015} N_{a,y}}$$

where $\epsilon_a$ is the proportion of primary vaccine failures, $H_{a,y}^{(1)}$ [or $H_{a,y}^{(2)}$] is the yearly number of naïve (or post-vaccine) infections at age $a$ during year $y$, and $N_{a,y}$ is the population size of age $a$ during year $y$. In keeping with previous studies (46–48), we considered children aged 5 to 10 (five age groups: 5 to 6, 6 to 7, 7 to 8, 8 to 9, and 9 to 10 years old), that is, 0 to 4 years after receipt of the fifth vaccine dose. This choice is also convenient because these age groups experience the same force of infection in our model, so that, all else being equal, differences of incidence reflect differences of susceptibility caused by waning of vaccinal immunity. We calculated the odds of acquiring pertussis at age $a$, $O_a = \frac{H_a}{1 - H_a}$, and fitted a log-linear regression model:

$$\log O_a = \alpha + \beta(a - \text{min}a)$$

Therefore, the quantity $\beta$ represents the average relative change in the odds of acquiring pertussis after every year of age, which can be compared directly to estimates from empirical studies. The results of these simulations are presented in Fig. 3F.

**REFERENCES AND NOTES**


**SUPPLEMENTARY MATERIALS**

www.sciencetranslationalmedicine.org/cgi/content/full/10/434/eaaj1748/DC1

Materials and Methods

Fig. S1. Pertussis transmission model schematic.


Author contributions: A.A.K. and P.R. conceived the study. M.D.d.C. and F.M.G.M. developed the model. M.D.d.C. implemented the model. M.D.d.C., F.M.G.M., A.A.K., and P.R. analyzed the results. M.D.d.C., F.M.G.M., A.A.K., and P.R. wrote the manuscript.

Competing interests: The authors declare that they have no competing interests.

Data and materials availability: R codes are available from https://github.com/kingaa/massachusetts-honeymoon. The data used in this paper (that is, monthly age-specific case counts during 1990–2005) are available from the MDPH upon request to the Office of Integrated Surveillance and Informatics Services help desk (isishelp@state.ma.us).

The impact of past vaccination coverage and immunity on pertussis resurgence
Matthieu Domenech de Cellès, Felicia M. G. Magpantay, Aaron A. King and Pejman Rohani

Sci Transl Med 10, eaaj1748,
DOI: 10.1126/scitranslmed.aaj1748

The problem of pertussis
The recent rise of pertussis in developed countries has generated controversy as to its cause. Domenech de Cellès et al. modeled pertussis transmission using incidence data from Massachusetts, United States. They found little evidence that the switch to the acellular vaccine contributed to the Massachusetts outbreaks. Instead, waning vaccine-conferred immunity, as opposed to vaccine failure to mount a full or even partial immune response, best explained the local rise in pertussis cases along with a historical gap in vaccination coverage. Simulations suggested that administering existing boosters to children may be an effective strategy to halt pertussis transmission.