

# Resolving pertussis immunity and vaccine effectiveness using incidence time series

*Expert Rev. Vaccines* 11(11), 1319–1329 (2012)

Jennie S Lavine\*<sup>1,2</sup> and Pejman Rohani<sup>1,2,3</sup>

<sup>1</sup>Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, MI 48109, USA

<sup>2</sup>Fogarty International Center, National Institutes of Health, Bethesda, MD 20892, USA

<sup>3</sup>Center for the Study of Complex Systems, University of Michigan, Ann Arbor, MI 48109, USA

\*Author for correspondence: [jennie.lavine@gmail.com](mailto:jennie.lavine@gmail.com)

Resolving the long-term, population-level consequence of vaccine-induced immunity to pertussis is a key challenge for control strategies and vaccine development. Controlled vaccine efficacy studies provide invaluable information; however, they are limited in scope by their sample size and follow-up duration. Long-term time series of incidence data collected by public health institutions provide insight at a broader scale, especially when the data are spatially explicit and age stratified. By using modern ecological and statistical methodologies, which are reviewed in this paper, new insights into the duration of transmission-blocking immunity and the age-specific patterns of transmission can be gained. Recent advances in computing power and statistical software development will increasingly make these methods available to public health practitioners, vaccine developers and academics alike.

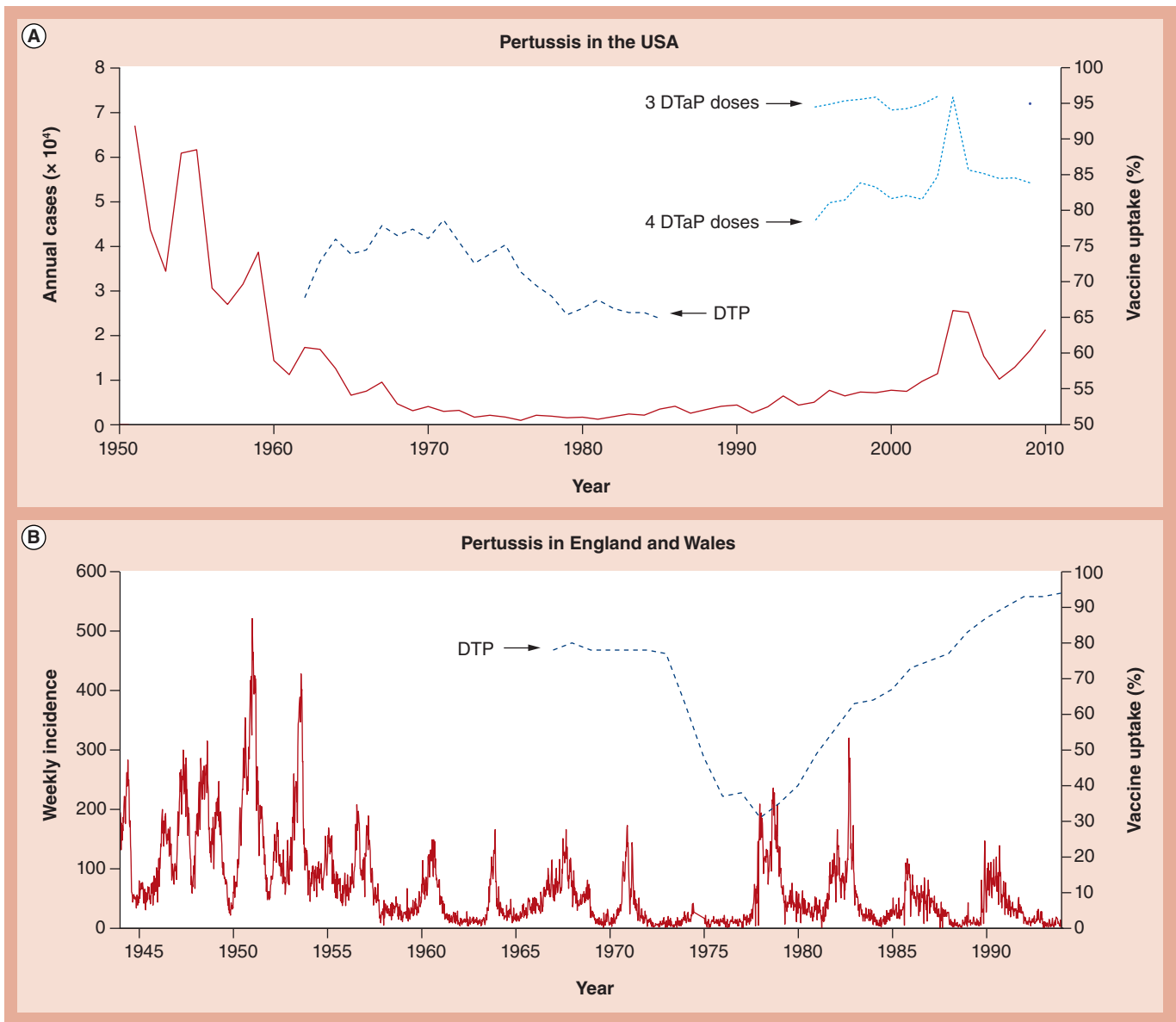
**KEYWORDS:** age-stratified incidence • disease ecology • dynamical signature • likelihood • mathematical model • reporting efficiency • statistical inference • time-series data • transmission • whooping cough

Pertussis, also known as whooping cough, is a respiratory infection that causes close to 200,000 deaths in children per year, despite the widespread use of a vaccine [1]. It is caused primarily by the Gram-negative coccobacillus *Bordetella pertussis*. Based on data from the pre-vaccine era, pertussis was thought of as a classic, immunizing childhood infection [2], with few reports of a second bout of whooping cough in adults. The mean age of infection in urban populations was approximately 4 years, and more than 90% of infections were observed in children younger than the age of 10 years [3]. In addition, the observed cyclic dynamics in pertussis reports (FIGURE 1) were consistent with infections that generate long lasting, sterilizing immunity. However, even before the vaccine was introduced, anecdotal evidence suggested that some adults were carrying and transmitting subclinical pertussis infections to young children, who then got a symptomatic primary infection [4].

Due to the high prevalence and resultant morbidity and mortality caused by pertussis in young children, a whole-cell vaccine was developed in the early 20th century. Large-scale vaccination campaigns were rolled out throughout much of the developed world in the 1940s and 1950s.

For the next few decades, incidence declined steadily, and it appeared that the vaccine was working as hoped for, protecting individuals from disease and the population from spread of infection. Beginning in the 1970s, some regions began to see a rise in cases [5]. In some locations, such as the UK and Sweden, this coincided with a major vaccine scare (FIGURE 1B) [6,7]. However, reported incidence increased despite sustained high coverage in many parts of the world including much of Europe, the USA, Australia and Taiwan [5,8–13].

Some researchers question whether the increase may in part be a result of the change from whole-cell to acellular vaccines. However, Argentina has recently experienced an increase in cases similar to that observed in other countries, despite continued, exclusive use of the whole-cell vaccine [14]. The resurgence led many to wonder whether the vaccine was working as well as had been initially assessed. In particular, researchers and public health workers began to question how long vaccine-induced immunity lasts, whether immunity prevents transmission or only disease, and whether there has been significant vaccine-driven evolution of the pathogen population. Alternatively, some suggest that increased awareness of adult pertussis, coupled



**Figure 1. Pertussis incidence in the USA and England and Wales. (A)** Annual incidence from the USA from 1951 until 2010 (solid line). Pediatric immunization is thought to have been implemented in the late 1940s, though uptake information is not available. Historical DTP vaccine estimates from 1962 to 1985 obtained from [61] are plotted using a dark, dashed line (axis on right) and DtaP estimates for three and four doses have been obtained from the National Immunization Survey and plotted in light, dashed line. **(B)** Weekly incidence of pertussis in England and Wales (solid line), accompanied by estimated DTP uptake (dark, dashed line). The national immunization program was initiated in 1957. Note the drop in vaccine uptake starting in 1974.

with improvements in diagnosis and reporting underlie increased cases, rather than a rise in transmission [15].

Assessing the diagnostic process is a key challenge in understanding pertussis transmission and immunity, as many infections are subclinical, and therefore never reported [1]. Reporting fidelity is related to disease severity, and it is thought that disease severity decreases with age and history of exposure to pertussis [16,17]. In addition, secondary infections may be less transmissible than primary ones. Explicit transmission experiments are impossible for obvious ethical reasons, rendering these quantities hard to identify; however, they are critical for

designing successful vaccination strategies. This becomes especially important, given the new appreciation for the import of age-specific contacts in pertussis transmission [18]. It seems that infants, who are at the highest risk of severe morbidity, and even mortality, are most commonly getting infected by older family members, many of whom have few or no symptoms [19–21]. The aforementioned questions regarding changes in vaccine-induced immunity and reporting efficiency, due to their broad scope and inherently dynamical nature, lend themselves to large-scale ecological analyses of incidence data that supplement traditional vaccine efficacy studies.

Long-term public health records of infectious disease incidence hold dynamical information on susceptibility and transmission, and, therefore, help us assess the long-term impacts of vaccination efforts. Vaccine trials provide a controlled system by which vaccine efficacy can be measured; however, they are necessarily small in scope, following limited subsets of a population for only a few years. These studies generate high-resolution data on individual effects, but are less probable to capture emergent phenomena, such as herd immunity. Incidence data, on the other hand, are often collected at a broader scale than traditional vaccine trials, tracking an entire population over many decades. Time series of disease incidence provide data at a scale commensurate with the breadth and duration of vaccine-induced protection, and are therefore more likely to provide information concerning the impact of these mechanisms at the population scale. Time-series data additionally hold information regarding the dynamics of transmission that are often lost in typical vaccine trials. Therefore, the authors argue that dynamical systems methods developed for analyzing ecological systems are an invaluable tool for assessing the long-term, population-wide impacts of vaccination.

In this perspective piece, the authors identify metrics that hold information about immunity in pertussis incidence time series and serosurveys, and discuss what they have learned from them, propose the use of modern methodologies for efficiently extracting information from dynamic data, and set goals for future pertussis data collection and analysis. The authors emphasize the importance of combining different data types and statistical methodologies to gain insight into the invisible processes, such as subclinical transmission, which drive pertussis epidemiology and have prevented widespread immunization from achieving the level of control they initially anticipated.

## Statistical signatures in surveillance data

### Disease prevalence

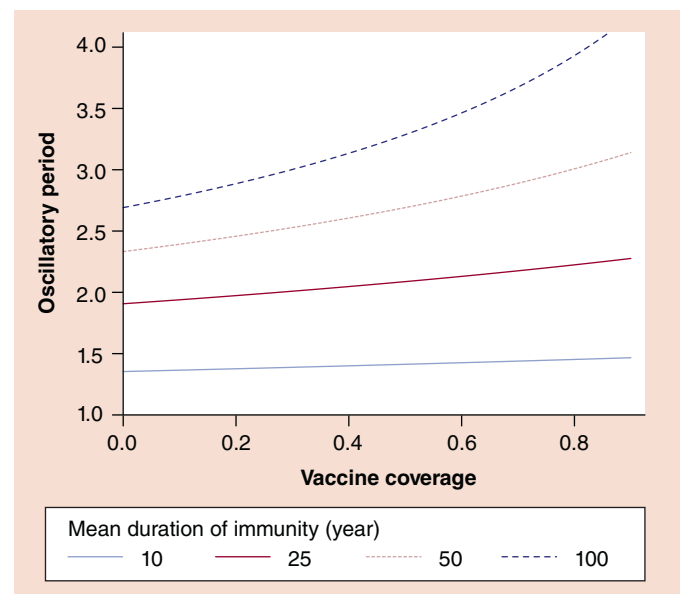
An obvious signature in long-term infectious disease surveillance data of a vaccine that produces some effect is a decrease in disease incidence following its introduction. This occurred globally with pertussis; every vaccinated region that has been studied experiences lower incidence on average now than in the prevaccine era (see FIGURE 1 for examples from the USA, and England and Wales). However, inference on vaccine efficacy based on average trends over time suffers from a variety of drawbacks. First, the mean incidence is highly sensitive to changes in the reporting probability. Second, it will be affected by vaccines that reduce either disease severity or transmission, making it an unreliable measure of the utility of a vaccine to generate herd immunity.

Long-term data sets of disease incidence over the span of years or decades hold far more information than just the mean trend, and these other aspects can help us assess the effectiveness and duration of vaccine-induced transmission-blocking immunity. By understanding the various vaccine effects, there is potential for more predictive power regarding the possibility of local or global eradication and vaccine-driven pathogen evolution.

### Periodicity

A key characteristic of time series of infectious diseases is the interepidemic period or the time between cyclic outbreaks. Directly transmitted, highly infectious, immunizing pathogens tend to produce recurrent multiannual epidemics (FIGURE 1) due to the inherently nonlinear process of transmission. An infection sparks an outbreak and depletes the population of susceptible hosts, which causes the chain of transmission to slow down until newborns replenish the susceptible population, generating the conditions necessary for a subsequent outbreak. Therefore, the speed of susceptible replenishment is key in determining the interepidemic period [22]. For a lifetime immunizing infection, this is determined by the birth rate and effective vaccine coverage. In general, slow susceptible recruitment, either due to low birth rates or high effective vaccine coverage, leads to a longer interepidemic period (FIGURE 2). Importantly, this result only holds when vaccination blocks transmission. If vaccination only prevents disease, the size of the observed outbreaks will decrease with higher vaccine coverage, but the interepidemic period will remain the same. Relatedly, the periodicity is insensitive to changes in reporting over time so long as they are not directly tied to incidence; this is an important feature for a robust metric given the notoriously low, age-specific and variable reporting rates for pertussis [23].

Many studies have fruitfully used this underlying theory to assess the effectiveness of pertussis vaccines in preventing transmission. The first study to do so [24], published in 1982, used aggregated



**Figure 2. Impact of vaccine coverage and loss of immunity on a predicted interepidemic period.** The dark, dashed line shows that when immunity is long lasting, the interepidemic period is sensitive to vaccine coverage, ranging from 2.7 years in the absence of vaccination to 4.2 years with 90% coverage. For scenarios in which immunity is lost more rapidly, the interepidemic period is shorter and less sensitive to vaccination. All curves were calculated according to [22] with the following parameter values: *per capita* annual birthrate ( $\mu$ ) of 0.02; mean infectious period ( $1/\gamma$ ) of 21 days; transmission rate ( $\beta$ ) of 200 per year, corresponding to a basic reproductive number ( $R_0$ ) of 11.5.

data from all of England and Wales, and cited little change in the range of observed interepidemic periods of pertussis after the introduction of vaccination as evidence that the vaccine only protected against disease, not transmission. Since then, various studies have demonstrated a positive association between the interepidemic period and vaccine coverage [25,26]. A more recent and detailed study on pertussis in England and Wales [27] that used data with higher spatial resolution showed evidence for a systematic increase in the interepidemic period following the inception of vaccination. Taken together, these studies lend strong credence to the theory that historical vaccines indeed provided some protection against transmissible disease.

Although the evidence for induction of transmission-blocking immunity is clear, details regarding its extent and duration remain obscure. A few studies have attempted to estimate these quantities from data on periodicity as well [26,28]; however, the large number of parameters that are not well identified makes rigorous inference difficult when using periodicity as the sole metric. In general, these studies have found that a variety of biological mechanisms can generate observed periodicity profiles. The variables shown to affect the interepidemic period other than birth rate and vaccine coverage include the durations of infection- and vaccine-induced immunity, the relative infectiousness of cases in previously immune hosts, the primary efficacy of the vaccine, the possibility of immune boosting without infection upon re-exposure and the noise in the system [28,29]. Recently developed methods for inference on dynamical systems allow for better identification of some of these mechanisms, which will be discussed later in this article.

### Extinction profile

The frequency and duration of local pathogen extinction events also hold information regarding immunity in a population. Although pertussis is nowhere near eradicated, weeks or months may pass with no case notifications in sufficiently small populations, suggesting that it is locally absent (termed a 'fade-out'), and then reintroduced from an outside source. As shown in FIGURE 3, even in the prevaccine era, some populations were below the critical community size, that is, they were too small to support sustained pertussis transmission [30]. If vaccination is transmission blocking, its introduction is predicted to increase the critical community size; a larger population will be necessary to prevent local extinction. This result holds even if immunity is not lifelong; however, a shorter duration of immunity decreases the critical community size [28] since the susceptible pool is replenished not only by births, but also by loss of immunity.

There are a few ways to empirically estimate the critical community size, but all relate to the frequency or duration of fade-outs [31]. This type of analysis has been used less frequently than the periodicity analysis for pertussis, perhaps, in part, because it is necessary to have fine spatial and temporal resolution, such as cases aggregated weekly from locations with a range of population sizes, and perhaps in part because it depends on immigration rates, which are often not well documented. Data from cities in England

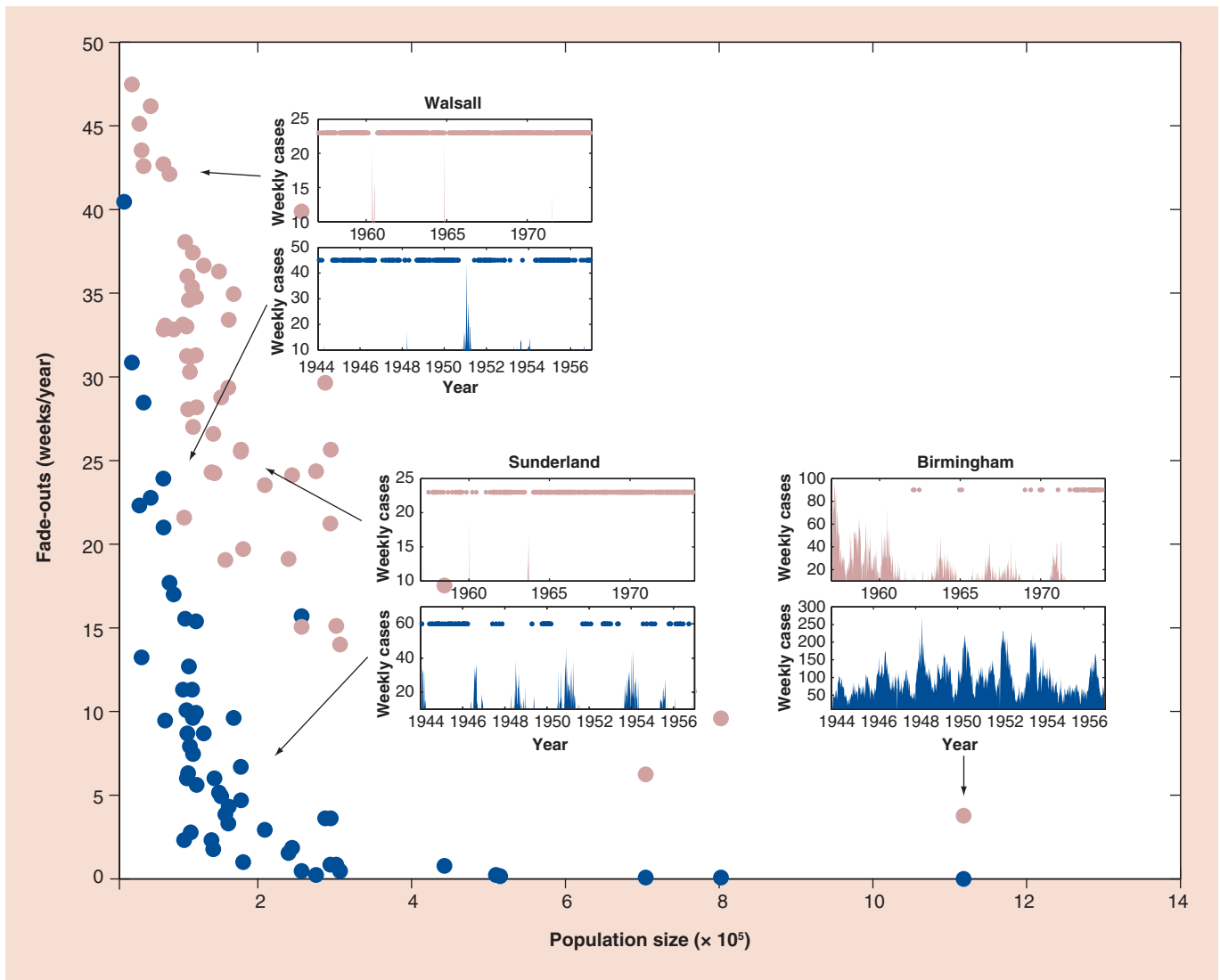
and Wales demonstrated a consistent increase in the number of weeks with zero cases and the duration of each individual fade-out from the pre- to post-vaccine eras, consistent with the predictions of a transmission-blocking vaccine (FIGURE 3) [27]. A similar study using data from small communities in Senegal also showed an increase in the duration of fade-outs with the introduction of vaccination [32].

### Age structure

A growing body of theory and data predicts how changes in vaccine coverage and disease transmission should affect age-stratified disease incidence [18,33–35]. The prediction from dynamical models of immunizing infectious diseases is that a transmission-blocking vaccine will decrease pathogen circulation, which reduces the force of infection (the rate at which susceptible individuals become infected) and leads to an increase in the proportion of primary infections occurring in older individuals [3,22]. The most naive model assumes homogeneous mixing among all individuals in a population. This implies, for example, people are equally likely to interact with all age groups, and there is no assortative mixing related to vaccination status. Although the impacts of these simplifications are important, the basic model provides a baseline against which observed data can be observed. In particular, it predicts that the ages of cases will be distributed exponentially in both the pre- and post-vaccine eras, but the mean will shift up in the presence of transmission-blocking vaccination. As with the previously discussed metrics, because this prediction is based on the dynamics of transmission, the shift in mean age will only occur if the vaccine blocks transmission.

One of the most widely noted changes in pertussis epidemiology in the past few decades has been the shift to cases in older individuals. In particular, since the beginning of the resurgence, there has been a dramatic increase in cases in teenagers [36,37], and epidemics have broken out in middle- and high-schools with highly vaccinated student populations [38,39]. As described earlier, the theory predicts that the proportion of cases in teenagers should increase simply due to a vaccine-induced reduction in transmission. However, the naive theory does not predict an increase in the number of cases in teenagers.

Based on the increase in teenage cases, many researchers postulated that immunity, at least to disease, is being lost increasingly rapidly due to either new pathogen strains that elicit shorter-lasting immunity [14,40,41] or reduced natural immune boosting arising from a reduction in pathogen circulation [35,42]. However, an alternative explanation has been proposed: the increase may result from an age-varying force of infection even if vaccine-induced transmission-blocking immunity is long lasting. A recent study by Rohani *et al.* showed that the documented high contact rates among teenagers, especially with each other [43], can explain the recent increase in teenage cases in Sweden, where pertussis vaccination was reintroduced in 1996 after a 17-year hiatus [18,44]. Studies have demonstrated that the pertussis vaccine prevents against transmission for at least a short time [45,46], and that its effectiveness at preventing disease wanes with time [39,46]. The extent to which disease in successfully vaccinated individuals



**Figure 3. Pertussis extinction profiles in England and Wales.** For the 60 largest cities, the authors plot the number of weeks per year during which no pertussis cases were reported against the respective population size. The prevaccine (1944–1957) and vaccine (1957–1974) era are plotted in dark and light filled circles, respectively. To illustrate the concept, the incidence reports for three different cities, Birmingham, Sunderland and Walsall, for both prevaccine (dark) and vaccine (light) eras have also been presented. The dots in these insets indicate that no pertussis cases were reported during that week.

contributes to transmission and maintenance of population-level immunity remains an open question [47].

### Expert commentary: methods for dynamical inference & hypothesis comparison

#### Comparing models & data

Thus far, we have described how a variety of metrics calculated from surveillance data – interepidemic period, extinction profile and age structure – can be compared with dynamical model predictions to make inference regarding various components of vaccine efficacy. Dynamic models of epidemics expressed as systems of differential equations date back to the early 20th century [48]. In the past, they were largely informed by epidemiological intuition regarding key processes and parameter estimates from outside sources, such as household transmission studies on the infectious

period. The early models primarily described large-scale, qualitative patterns in data. The methods described in the preceding section allow for quantitative contact between models and data. Although these methods do not allow for definitive rejection or acceptance of a hypothesis, like other statistical methods, they do allow us to assess the relative merits of different hypotheses regarding the underlying processes driving epidemic patterns.

Each metric has its strengths and weaknesses because it is sensitive to different changes in model assumptions and data collection. Looking at each individually, we are confronted with statistical trade-offs, for example, between the duration of immunity and the probability of immune boosting [28], that prevent us from choosing between two competing hypotheses. Modern statistical methodologies that take advantage of rapid computing power allow the data to speak for themselves; they



comprehensively search potentially high-dimensional parameter spaces to optimize a given model in the face of data, and provide a rigorous, quantitative framework for choosing between optimized models.

### **Probe matching**

We may consider each metric described earlier as a summary statistic of a set of spatially explicit, age-stratified time series of incidence data. In all of the aforementioned examples, the full data set was aggregated in some way and reduced to a particular statistic, which was then compared with a theoretical prediction. For example, in [28], all age groups were lumped together to form one time series for each location. These were then summarized by two metrics separately: the dominant period observed in each location and the proportion of weeks in which zero cases were observed. A number of parameter values (i.e., durations of immunity and rapidity of boosting) were chosen along with a corresponding transmission rate to ensure the mean age of infection was consistent with observations, and these values were then used to simulate from a dynamic model, governed by stochastic differential equations. Summary statistics of the simulated time series were compared with summary statistics of the data to identify which parameter value was most consistent with the data. This same basic idea has recently been developed further to allow for two improvements to the methodology and is known as ‘probe matching’ [49]. First, the method allows for simultaneous comparison of arbitrarily many summary statistics, and accounts for the covariation between them. Second, it is implemented in a likelihood-based framework, the synthetic likelihood, such that models with different numbers of parameters can be compared using standard likelihood theory.

### **Likelihood methods**

All of the methods discussed so far require summarizing time-series data and model outputs with individual statistics or ‘probes’. In addition, a variety of methods exist that allow for comparison of model predictions with each data point in a time series, rather than choosing certain metrics and discarding the rest of the information contained in the data. These methods, termed ‘sequential Monte Carlo’ or ‘plug-and-play’, utilize the statistical quantity of the likelihood, and are ideal for comparing biologically motivated hypotheses because they require that only one can encapsulate a hypothesis in a model in which each time step depends only on the preceding one (a so-called Markov chain) [50,51].

The likelihood of the data given fixed values of the governing parameters can be estimated by comparison with simulations. By exploring a wide range of parameter values and combinations, we can find the optimal parameters for a given model and data set. For example, in FIGURE 4B, we compute the maximum likelihood (y-axis) of models with different  $R_0$ s (x-axis) in the face of a time series from the prevaccine era in Copenhagen (Denmark) by fixing that parameter and optimizing the others. Even when all the other parameters in the model (duration of immunity, rapidity

of boosting and others) are allowed to vary to best fit the data, likelihood theory suggests that an  $R_0$  larger than 20 or less than 12 (range of x-values above dashed gray line) provides a significantly worse fit to the data.

Between the development of rapid optimization algorithms for maximizing the likelihood and the explosion in computing power, it is now feasible to search high-dimensional parameter spaces and identify even small differences in the explanatory power of hypotheses encapsulated in mathematical models against large and detailed data sets. The ability to optimize the likelihood grounds model comparison in well-developed statistical theory.

A pure likelihood-based iterated filtering algorithm [51,52] was recently used to test hypotheses regarding the duration of immunity and the importance of natural immune boosting based on a long time series from the prevaccine era in Copenhagen (in review, FIGURE 4). All parameters were estimated simultaneously, thereby allowing the data to speak for themselves. Despite many degrees of freedom, key quantities were tightly constrained. For example, the basic reproductive number,  $R_0$ , was estimated to be 16, with 95% CI: 12–19 (FIGURE 4B). Interestingly, values outside that range produced simulations that were not visibly worse (compare FIGURE 4A with FIGURES 4C and 4D); however, the statistical methods were able to harvest information from patterns not obvious via a visual comparison. A particularly useful result of simultaneously estimating all the parameters in this fashion was that the reporting efficiency, or observation probability, was well identified at approximately 15% (maximum likelihood estimate: 15.2%, FIGURE 5, dark, solid line), with a 95% CI only spanning 14.7–16.1% (FIGURE 5, light circles and arrow). This is an important point to emphasize because the value of analyzing incidence reports is often questioned on the grounds of incomplete notification. Unlike the metrics described in the previous section, this likelihood maximization method not only takes reporting efficiency into account, but can quantify it based on the dynamical information contained within the time-series data.

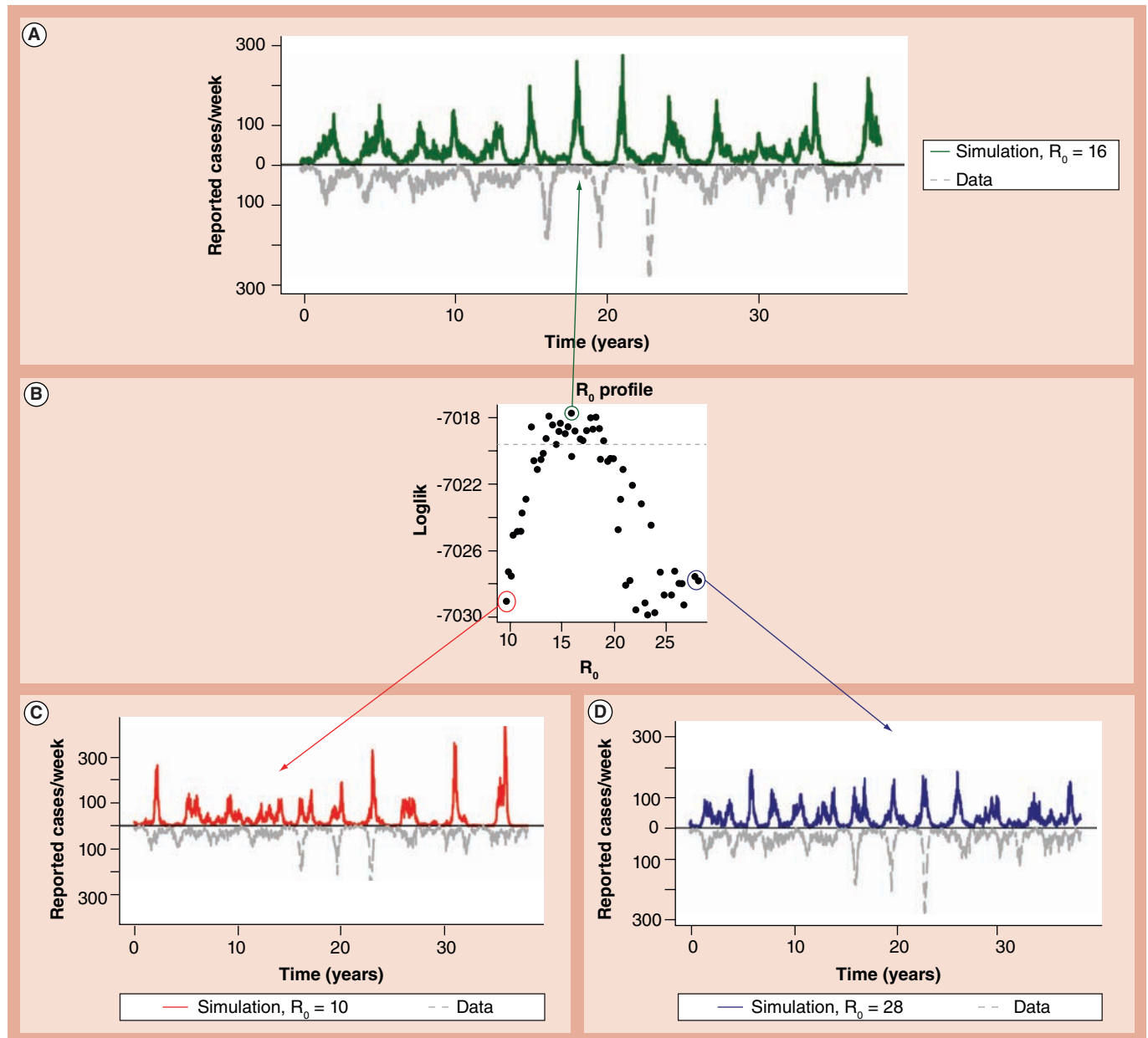
### **Five-year view**

In the same way that running a  $t$  test was once a laborious statistical procedure before the advent of the calculator, but is now simple and commonplace, so these more sophisticated statistical methods are increasingly accessible to public health workers, biologists and vaccine manufacturers alike with the advent of well-built statistical packages and cloud computing. Free and open-source implementations of these methods, such as the recently developed package *pomp* in the statistical environment R [52,53], make these tools available without a cost barrier. In conjunction with the burst in cloud computing options, very little overhead is necessary to implement them. The largest barrier to general use of these methods by researchers may be the skills required to convert biological hypotheses to mathematical models and write computer codes describing them. Recent efforts to create a web-based graphical user interface for these methods take us a step closer to making these methods generally accessible and

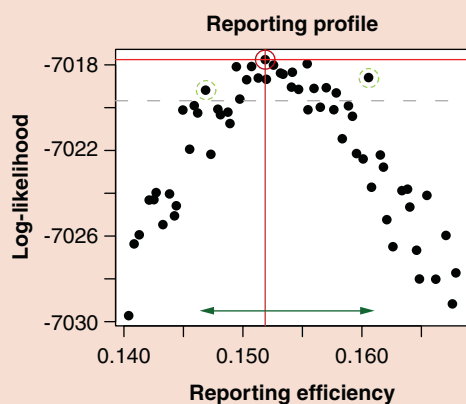
usable. The price tag and need for a specialized skill set are therefore diminishing impediments to using these methods to gain a better understanding of the emergent properties of immunity engendered by pertussis immunization efforts.

The approaches suggested here glean information from multiple, large and diverse data sets simultaneously. For example, it is now possible to take a broad range of location- and age-specified time series from the pre- and post-vaccine eras, and compare the ability of various hypotheses to explain their patterns, even

in the face of covarying parameters and under-reporting. A key goal for the future is therefore to incorporate diverse data sets into unified analyses to better understand the global drivers of pertussis epidemiology. Two actions are necessary for this to become a reality: collaboration among public health departments and researchers to bring together the necessary data on pertussis incidence, demographic characteristics and vaccine policies over time, and a digital platform for storing and sharing these data, with all proper privacy protections, metadata and credit given to



**Figure 4. Model simulations and likelihood profile over  $R_0$ .** (A) Sample model simulation from the maximum likelihood parameters (dark line) along with pertussis incidence data from Copenhagen (Denmark) from 1900 to 1937 (light line) [62]. (B) The likelihood profile for the basic reproductive number,  $R_0$ . The value of  $R_0$  is fixed at points along the x-axis, the likelihood maximized using iterated filtering and the corresponding estimated maximized likelihood shown on the y-axis. The point circled in green represents the maximum likelihood estimate. Points below the dashed line represent values that are outside the approximate 95% CI for  $R_0$ . (C & D) Simulations from models with low and high  $R_0$  values, respectively. Because these are simulations from stochastic models, other simulations with identical parameters can look quite different.



**Figure 5. Likelihood profile for the reporting efficiency obtained from analyses of the Copenhagen pertussis data described in Figure 4.** Log-likelihood values that fall below the dashed line are outside the approximate 95% CI of the maximum likelihood estimate. The dark, solid circled point has the highest log-likelihood, and therefore corresponds to the maximum likelihood estimate for the reporting efficiency, 0.152. The two light, dashed circled points correspond to the most extreme reporting probabilities that have likelihoods within the confidence bounds, and the arrow at the bottom shows the CI for reporting efficiency, 0.147–0.161.

the people and institutions who collected them. Organizations such as the pan-European pertussis research group, Eupert, offer a model for the type of international collaboration necessary, and others such as the Wellcome Trust and the journal *Epidemiology*, are supporting the move toward making public health data a public resource [54,101].

### Conclusion & the path forward

Despite the early development and widespread deployment of a pertussis vaccine, whooping cough still plagues even highly vaccinated human populations. The causes for the lack of control remain in question, in part, because key aspects of immunity and vaccine efficacy are poorly understood and challenging to study. We cannot directly measure how long vaccine-induced protection lasts. The closest we come to a direct observation is via serological data; however, the lack of antibodies against pertussis antigens does not imply lack of protection [55]. The transmissibility of a case resulting from secondary exposure is even more obscured from direct observation than reinfection itself, which may at least result in a serological signature of boosted antibody titers or mild symptoms that can be observed. Fortunately, these invisible events leave their footprints in large-scale dynamical, time-series data. Modern statistical methods for dynamical systems can x-ray data, extracting information about transmission and reporting rates contained in dynamical patterns that are invisible to the naked eye.

In addition to helping us quantify unobservable aspects of vaccine efficacy, dynamical models pinpoint the myriad factors that affect pertussis transmission in complex and sometimes counterintuitive ways. Recent models have demonstrated the epidemiological outcomes of pertussis vaccination are sensitive

to age-specific contact mixing patterns [18], the ease with which immunity may be boosted upon re-exposure [35], and the impact of the population-level profile of immunity on pathogen evolution [56]. A variety of other factors are likely to impact pertussis epidemiology as well. For example, congeneric species, such as *Bordetella parapertussis* and *Bordetella holmesii* have recently been implicated in many outbreaks of whooping cough-like disease [57–59]. In addition, different vaccines and vaccination schedules elicit different types and levels of immunity, and the large-scale effects of this variation require further exploration. We require further theoretical studies and models to assess which of these factors have the potential to be key drivers of pertussis dynamics. Thus far, these more complex models are largely informed by intuition, expert knowledge and external parameter estimates, as the simpler ones were 30 years ago, and provide primarily qualitative predictions and comparisons to data. We are now at the cusp of being able to test complex and increasingly realistic models against vast and diverse data sets. Techniques for statistical inference not only yield important quantitative information, they also provide a theoretically sound platform for hypothesis testing.

Each year, new datasets are collected and published on biological and sociological phenomena that affect disease transmission. High-resolution time-series data from various geographic locations may provide clues as to the key drivers of seasonality in pertussis dynamics. Sociological studies on the contact patterns among age groups in different locales and cultures provide a grounding for testing hypotheses regarding the age-specific routes of pertussis transmission, and, in particular, may offer strategies to provide indirect protection to infants [43,60]. At a different scale, the spatial and temporal patterns of *B. pertussis* population genetics are being uncovered; the growing bank of sequence data provides information on patterns of transmission, as well as the evolutionary processes that impact pertussis epidemiology. In the future, by considering models and methods that concurrently incorporate various types of data and encompass a broad range of hypotheses, we will gain new insights into the effects of vaccines on pertussis transmission, and thereby develop strategies to better control it.

### Acknowledgements

The authors thank A King for comments.

### Financial & competing interests disclosure

JS Lavine and P Rohani were supported by the Research and Policy for Infectious Disease Dynamics program of the Science and Technology Directorate, Department of Homeland Security, and the Fogarty International Center, NIH. P Rohani was also supported by the Vaccine Modeling Initiative of the Bill and Melinda Gates Foundation and by a grant from the NIH (1R01AI101155). The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.



## Key issues

- Pertussis, or whooping cough, remains a significant cause of morbidity and infant mortality despite high vaccine coverage.
- The causes of the poor control remain obscure, and questions abound regarding the duration and type of protection provided by currently used vaccines.
- Classical vaccine efficacy studies are invaluable, but limited in their ability to assess emergent phenomena, such as herd immunity.
- Ecological methods developed for dynamical studies of populations over time can be applied to time series of disease reports to gain new insights into long-term, population-level processes.
- Particularly informative features of time-series data for estimating the duration of transmission-blocking immunity include periodicity, frequency of fade-outs and age-stratified incidence.
- Rigorous model comparisons via simulation-based calculations of the likelihood are possible due to recent developments in statistical methodologies and new algorithms for optimization across many unknown quantities simultaneously.
- The aforementioned methods provide strong evidence for transmission-blocking immunity provided by vaccination and a relatively narrow estimate of the reporting efficiency based on dynamical feedbacks reflected in incidence data.

## References

Papers of special note have been highlighted as:

• of interest

•• of considerable interest

- Crowcroft NS, Stein C, Duclos P, Birmingham M. How best to estimate the global burden of pertussis? *Lancet Infect. Dis.* 3(7), 413–418 (2003).
- Gordon JE, Hood RI. Whooping cough and its epidemiological anomalies. *Am. J. Med. Sci.* 222(3), 333–361 (1951).
- Anderson RM, May RM. *Infectious Diseases of Humans: Dynamics and Control*. Oxford University Press, Oxford, UK (1991).
- Luttinger P. The epidemiology of pertussis. *Am. J. Dis. Child* 12, 290–315 (1916).
- Rohani P, Drake JM. The decline and resurgence of pertussis in the US. *Epidemics* 3, 183–188 (2011).
- Gangarosa EJ, Galazka AM, Wolfe CR *et al.* Impact of anti-vaccine movements on pertussis control: the untold story. *Lancet* 351(9099), 356–361 (1998).
- Baker JP. The pertussis vaccine controversy in Great Britain, 1974–1986. *Vaccine* 21, 4003–4010 (2003).
- de Melker HE, Schellekens JF, Neppelenbroek SE, Mooi FR, Rümke HC, Conyn-van Spaendonck MA. Reemergence of pertussis in the highly vaccinated population of The Netherlands: observations on surveillance data. *Emerg. Infect. Dis.* 6, 348–357 (2000).
- Skowronski DM, De Serres G, MacDonald D *et al.* The changing age and seasonal profile of pertussis in Canada. *J. Infect. Dis.* 185, 1448–1453 (2002).
- Pebody RG, Gay NJ, Giammanco A *et al.* The seroepidemiology of *Bordetella pertussis* infection in Western Europe. *Epidemiol. Infect.* 133(1), 159–171 (2005).
- Dudman SG, Trøseid M, Jonassen TO, Steinbakk M. Kikhoste – et økende problem i Norge. *Tidsskr Nor Lægeforen* 3, 305–308 (2006).
- Celentano LP, Massari M, Paramatti D, Salmasso S, Tozzi AE. Resurgence of Pertussis in Europe. *J. Pediatr. Infect. Dis.* 24, 761–765 (2005).
- Lin YC, Yao SM, Yan JJ *et al.* Epidemiological shift in the prevalence of pertussis in Taiwan: implications for pertussis vaccination. *J. Med. Microbiol.* 56(Pt 4), 533–537 (2007).
- Hozbor D, Mooi F, Flores D *et al.* Pertussis epidemiology in Argentina: trends over 2004–2007. *J. Infect.* 59(4), 225–231 (2009).
- Provides evidence for a resurgence in the face of continued whole-cell vaccination.
- Cherry JD. The science and fiction of the ‘resurgence’ of pertussis. *Pediatrics* 112, 405–406 (2003).
- Storsaeter J, Hallander HO, Gustafsson L, Olin P. Low levels of antipertussis antibodies plus lack of history of pertussis correlate with susceptibility after household exposure to *Bordetella pertussis*. *Vaccine* 21, 3542–3549 (2003).
- Begins to tease apart possible serological correlates of protection, and suggests that recent previous exposure may provide it.
- Préziosi MP, Halloran ME. Effects of pertussis vaccination on disease: vaccine efficacy in reducing clinical severity. *Clin. Infect. Dis.* 37, 772–779 (2003).
- Rohani P, Zhong X, King AA. Contact network structure explains the changing epidemiology of pertussis. *Science* 330(6006), 982–985 (2010).
- Highlights the importance of age-specific mixing patterns and transmission rates in pertussis disease dynamics.
- Bisgard KM, Pascual FB, Ehresmann KR *et al.* Infant pertussis: who was the source? *Pediatr. Infect. Dis. J.* 23(11), 985–989 (2004).
- De Greeff SC, Mooi FR, Westerhof A *et al.* Pertussis disease burden in the household: how to protect young infants. *Clin. Infect. Dis.* 50, 1339–1345 (2010).
- Lavine J, Broutin H, Harvill ET, Bjørnstad ON. Imperfect vaccine-induced immunity and whooping cough transmission to infants. *Vaccine* 29, 11–16 (2010).
- Keeling MJ, Rohani P. *Modelling Infectious Diseases: in Humans and Animals*. Princeton University Press, NJ, USA (2008).
- Sutter RW, Cochi SL. Pertussis hospitalizations and mortality in the United States, 1985–1988. Evaluation of the completeness of national reporting. *JAMA* 267(3), 386–391 (1992).
- Fine PE, Clarkson JA. The recurrence of whooping cough: possible implications for assessment of vaccine efficacy. *Lancet* 1(8273), 666–669 (1982).
- Broutin H, Viboud C, Grenfell BT, Miller MA, Rohani P. Impact of vaccination and birth rate on the epidemiology of pertussis: a comparative study in 64 countries. *Proc. Biol. Sci.* 277(1698), 3239–3245 (2010).
- Provides a nice example of the use of interepidemic period across many data

- sets to provide strong evidence for the impact of vaccination on pertussis transmission.
- 26 Korobeinikov A, Maini PK, Walker WJ. Estimation of effective vaccination rate: pertussis in New Zealand as a case study. *J. Theor. Biol.* 224(2), 269–275 (2003).
- 27 Rohani P, Earn DJ, Grenfell BT. Impact of immunisation on pertussis transmission in England and Wales. *Lancet* 355(9200), 285–286 (2000).
- 28 Wearing HJ, Rohani P. Estimating the duration of pertussis immunity using epidemiological signatures. *PLoS Pathog.* 5, e1000647 (2009).
- **Exemplifies many of the methods described for using dynamical signatures to compare hypotheses regarding pertussis immunity and transmission.**
- 29 Rohani P, Keeling MJ, Grenfell BT. The interplay between determinism and stochasticity in childhood diseases. *Am. Nat.* 159, 469–481 (2002).
- 30 Bartlett MS. The critical community size for measles in the United States. *J. Roy. Stat. Soc. Series A (General)* 123, 37–44 (1960).
- 31 Conlan AJ, Rohani P, Lloyd AL, Keeling M, Grenfell BT. Resolving the impact of waiting time distributions on the persistence of measles. *J. R. Soc. Interface* 7, 623–640 (2010).
- 32 Broutin H, Simondon F, Guégan JF. Whooping cough metapopulation dynamics in tropical conditions: disease persistence and impact of vaccination. *Proc. Biol. Sci.* 271 (Suppl. 5), S302–S305 (2004).
- 33 Grenfell BT, Anderson RM. Pertussis in England and Wales: an investigation of transmission dynamics and control by mass vaccination. *Proc. R. Soc. Lond., B, Biol. Sci.* 236(1284), 213–252 (1989).
- 34 Wallinga J, Teunis P, Kretzschmar M. Using data on social contacts to estimate age-specific transmission parameters for respiratory-spread infectious agents. *Am. J. Epidemiol.* 164, 936–944 (2006).
- 35 Lavine JS, King AA, Bjørnstad ON. Natural immune boosting in pertussis dynamics and the potential for long-term vaccine failure. *Proc. Natl Acad. Sci. USA* 108(17), 7259–7264 (2011).
- **Shows that decreased circulation in the presence of rapid immune boosting could cause changes in the disease dynamics that lead to increased teenage incidence.**
- 36 Yih WK, Lett SM, des Vignes FN, Garrison KM, Sipe PL, Marchant CD. The increasing incidence of pertussis in Massachusetts adolescents and adults, 1989–1998. *J. Infect. Dis.* 182(5), 1409–1416 (2000).
- 37 Güris D, Strebel PM, Bardenheier B *et al.* Changing epidemiology of pertussis in the United States: increasing reported incidence among adolescents and adults, 1990–1996. *Clin. Infect. Dis.* 28(6), 1230–1237 (1999).
- 38 Brennan M, Strebel P, George H *et al.* Evidence for transmission of pertussis in schools, Massachusetts, 1996: epidemiologic data supported by pulsed-field gel electrophoresis studies. *J. Infect. Dis.* 181(1), 210–215 (2000).
- 39 Berger F, Njamkepo E, Minaberry S *et al.* Investigation on a pertussis outbreak in a military school: risk factors and approach to vaccine efficacy. *Vaccine* 28(32), 5147–5152 (2010).
- 40 van Boven M, de Melker HE, Schellekens JF, Kretzschmar M. Waning immunity and sub-clinical infection in an epidemic model: implications for pertussis in The Netherlands. *Math. Biosci.* 164, 161–182 (2000).
- 41 Mooi FR. *Bordetella pertussis* and vaccination: the persistence of a genetically monomorphic pathogen. *Infect. Genet. Evol.* 10(1), 36–49 (2010).
- **Reviews the existing evidence for vaccine-driven evolution and verbally lays out a mechanism by which the vaccine could drive the observed changes in pertussis epidemiology.**
- 42 Aguas R, Gonçalves G, Gomes MG. Pertussis: increasing disease as a consequence of reducing transmission. *Lancet Infect. Dis.* 6(2), 112–117 (2006).
- 43 Mossong J, Hens N, Jit M *et al.* Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS Med.* 5, e74 (2008).
- 44 Carlsson RM, Trollfors B. Control of pertussis—lessons learnt from a 10-year surveillance programme in Sweden. *Vaccine* 27(42), 5709–5718 (2009).
- 45 Préziosi MP, Halloran ME. Effects of pertussis vaccination on transmission: vaccine efficacy for infectiousness. *Vaccine* 21, 1853–1861 (2003).
- 46 Lavine JS, Bjørnstad ON, de Blasio BF, Storsaeter J. Short-lived immunity against pertussis, age-specific routes of transmission, and the utility of a teenage booster vaccine. *Vaccine* 30(3), 544–551 (2012).
- 47 Schellekens J, von König CH, Gardner P. Pertussis sources of infection and routes of transmission in the vaccination era. *Pediatr. Infect. Dis. J.* 24, S19–S24 (2005).
- 48 Kermack WO, McKendrick AG. Contributions to the mathematical theory of epidemics—I. 1927. *Bull. Math. Biol.* 53(1–2), 33–55 (1991).
- 49 Wood SN. Statistical inference for noisy nonlinear ecological dynamic systems. *Nature* 466, 1–6 (2010).
- 50 Doucet A, De Freitas N, Gordon N. Sequential Monte Carlo methods in practice. *Statistics for Engineering and Information Science*. Springer, NY, USA (2001).
- 51 Ionides EL, Bhadra A, Atchade Y, King AA. Iterated Filtering. *Ann. Stat.* 39, 1776–1802 (2011).
- 52 King AA, Ionides EL, Bretó CM *et al.* *Pomp: statistical inference for partially observed Markov processes*. (R package) (2010).
- 53 R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria (2010).
- 54 Hernán MA, Wilcox AJ. Epidemiology, data sharing, and the challenge of scientific replication. *Epidemiology* 20, 167–168 (2009).
- **Concise and persuasive plea for data sharing in epidemiology as the next best thing to repeated experiments.**
- 55 Hewlett EL, Halperin SA. Serological correlates of immunity to *Bordetella pertussis*. *Vaccine* 16(20), 1899–1900 (1998).
- 56 van Boven M, Mooi FR, Schellekens JF, de Melker HE, Kretzschmar M. Pathogen adaptation under imperfect vaccination: implications for pertussis. *Proc. Biol. Sci.* 272, 1617–1624 (2005).
- 57 Bokhari H, Said F, Syed MA *et al.* Whooping cough in Pakistan: *Bordetella pertussis* vs *Bordetella parapertussis* in 2005–2009. *Scand. J. Infect. Dis.* 43, 818–820 (2011).
- 58 Miranda C, Porte L, García P. *Bordetella holmesii* in nasopharyngeal samples from Chilean patients with suspected *Bordetella pertussis* infection. *J. Clin. Microbiol.* 50, 1505 (2012).
- 59 Njamkepo E, Bonacorsi S, Debryne M, Gibaud SA, Guillot S, Guiso N.

- Significant finding of *Bordetella holmesii* DNA in nasopharyngeal samples from French patients with suspected pertussis. *J. Clin. Microbiol.* 49, 4347–4348 (2011).
- 60 Eames KT, Tilston NL, Brooks-Pollock E, Edmunds WJ. Measured dynamic social contact patterns explain the spread of H1N1v influenza. *PLoS Comput. Biol.* 8, e1002425 (2012).
- 61 Simpson DM, Ezzati-Rice TM, Zell ER. Forty years and four surveys: how does our measuring measure up? *Am. J. Prev. Med.* 20(Suppl. 4), 6–14 (2001).
- 62 Metcalf CJ, Bjørnstad ON, Grenfell BT, Andreasen V. Seasonality and comparative dynamics of six childhood infections in pre-vaccination Copenhagen. *Proc. Biol. Sci.* 276(1676), 4111–4118 (2009).

**Website**

- 101 Wellcome Trust. Sharing research data to improve public health: full joint statement by funders of health research electronic. [www.wellcome.ac.uk/About-us/Policy/Spotlight-issues/Data-sharing/Public-health-and-epidemiology/WTDV030690.htm](http://www.wellcome.ac.uk/About-us/Policy/Spotlight-issues/Data-sharing/Public-health-and-epidemiology/WTDV030690.htm)