Epidemiol. Infect. (2002), **129**, 479–489. © 2002 Cambridge University Press
DOI: 10.1017/S0950268802007896 Printed in the United Kingdom

Kinetics of the IgG antibody response to pertussis toxin after infection with *B. pertussis*

P. F. M. TEUNIS[1]*, O. G. VAN DER HEIJDEN[2], H. E. DE MELKER[2], J. F. P. SCHELLEKENS[3], F. G. A. VERSTEEGH[4] AND M. E. E. KRETZSCHMAR[2]

[1] *Computerization and Methodological Consultancy Unit, National Institute for Public Health and the Environment, PO Box 1, 3720 BA Bilthoven, The Netherlands*
[2] *Centre for Infectious Diseases Epidemiology, National Institute for Public Health and the Environment, PO Box 1, 3720 BA Bilthoven, The Netherlands*
[3] *Laboratory of Infectious Diseases Screening, National Institute for Public Health and the Environment, PO Box 1, 3720 BA Bilthoven, The Netherlands*
[4] *Groene Hart Ziekenhuis, Department of Pediatrics, Gouda, The Netherlands*

(*Accepted 29 August 2002*)

SUMMARY

We aimed to provide a quantitative description of decay in pertussis antibody levels to aid in finding a serological estimate of the incidence of pertussis. The serum IgG response against pertussis toxin was studied in a group of clinically diagnosed patients. Individual records consisted of repeated serum IgG measurements at irregular intervals for up to 10 years post diagnosis. These data were analysed with a nonlinear regression model taking into account censoring at upper and lower threshold levels, measurement errors, and individual variation in the shape and magnitude of the immune response. There was considerable variation between individual responses, both in strength (amplitude) and duration (shape). The inverse model relating IgG levels to time from infection (diagnosis) can be applied to cross-sectional IgG data to generate distributions of times from infection, which may be used to calculate infection rates and their variation, in populations sampled for cross-sectional IgG data.

INTRODUCTION

Pertussis (whooping cough) is an infection of the respiratory tract caused by the highly contagious bacterium *Bordetella pertussis*. Despite high vaccination coverages, *B. pertussis* is still circulating in the Dutch population. Notification data show that most symptomatic cases occur among vaccinated children aged 4–9 years [1]. However, symptoms are most severe among infants who are too young to be vaccinated. In order to protect this group at risk it is important to determine which people are most likely to transmit infection to newborns. We aimed to identify those age groups in the population in which most of the circulation of *B. pertussis* takes place. The age profile of notified cases may not reflect the age distribution of infection with *B. pertussis* because the case/infection ratio is probably higher in younger age groups [2]. It was necessary to determine the infection rates in each age category, rather than notification rates.

In response to an infection, IgG titres typically show a rapid increase, followed by a steady, slow decline over a long period (several years) [1, 3–9]. Also IgG-PT induced after three or four doses of acellular or whole cell pertussis vaccine in the first year of life declines rapidly, mostly within one year, to low levels [10, 11]. Therefore, it is likely that in individuals in whom the last vaccination with pertussis vaccine has been administered more than one year ago, the finding of

* Author for correspondence.

moderate or high levels of IgG-PT indicates infection with *Bordetella pertussis*. Because the magnitude of the IgG-PT level is inversely related to time elapsed since infection, this suggests that in a patient with a given response, time from infection can be estimated from IgG levels.

If this approach to estimating times from infection were feasible, cross-sectional studies of IgG-PT-levels could be used to estimate the incidence of infection, independent of case notification. This requires quantitative characterization of the IgG-PT-response to pertussis, inclusive of its variation among individual patients.

We used data from an ongoing long term follow-up study in which blood samples were taken at irregular intervals from patients with clinically and serologically confirmed pertussis. At present, this study comprises 85 patients with follow-up times ranging from 6 months to 11 years. Numbers of samples were unequally distributed among patients. In an earlier report on the first 57 patients in this longitudinal study it was shown that despite large variation in responses the general pattern appeared to be a rapid increase in antibody levels followed by a slow decrease.

We tested several mathematical functions for their ability to fit these observed changes in IgG-PT levels with time. Here we present results of this analysis, taking into account censoring at upper and lower threshold levels, measurement errors, as well as individual variation in the shape and magnitude of the IgG-PT immune response.

The selected model describes a functional relationship between time since last infection and level of antibody titre. The model was applied to a cross-sectional population based study of IgG-pertussis antibody titres, permitting estimation of a distribution of infection dates for individuals in this population. Such model-based age-specific distributions of times since infection assist in identifying those age groups in which circulation of *B. pertussis* is most prevalent.

METHODS

Data used

During the period 1989–2000, a collection of follow-up serum samples was obtained from 85 patients clinically diagnosed with pertussis (paroxysmal cough lasting more than 2 weeks) in whom the clinical diagnosis had been confirmed by the finding of an IgG-PT level of 75 U/ml in the first or the second serum obtained in the symptomatic stage. The specificity of IgG-PT of 75 U/ml as an indicator of recent infection with *B. pertussis* has been estimated to be $>97{\cdot}5\%$ while the sensitivity was around 80% [1]. For participation in this study a minimum follow-up period of 3 months was required.

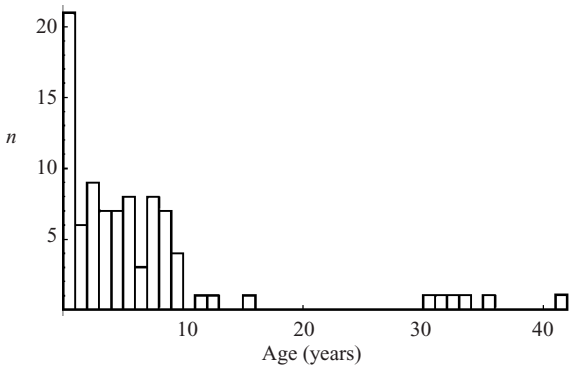

Fig. 1. Age distribution of patients.

In one of the participating patients a second symptomatic infection with *B. pertussis* occurred 7 years after the first, confirmed by positive pertussis PCR and a strong rise of IgG-PT. For analysis in this study we considered this record to be two patients: the first one connected to the first episode of pertussis and its follow-up until the last sampling before the second episode, and one connected to the second episode and the follow-up thereafter. Thus, in 85 patients, 86 episodes of pertussis and the course of IgG-PT thereafter were analysed. Recently, an additional three patients appeared to be re-infected [12]. These were not included in the present analysis.

The follow-up period ranged from 6 months to 11 years and the number of serial sera per patient ranged from 2 to 11. The age distribution of the patients is given in Figure 1. Most were children between 0·5 and 17 years of age (69/86). Eleven infants were less than 6 months old at the time they contracted pertussis. Also included were six adults with ages ranging from 30 to 41 years. Vaccination status as reported by the physician at the time of onset of pertussis was in 1 of 4 categories: negative, incomplete (1 or 2 vaccinations), complete (3 or 4 vaccinations) or unknown. Four infants whose vaccination had not been completed at the time of pertussis were vaccinated shortly thereafter. We therefore decided to use five categories of patients: infants (0–0·5 years) vaccinated after infection (4); infants not vaccinated after infection (7); vaccinated juveniles (0·5–20 years) (62); unvaccinated juveniles (7) and adults (6). These subgroups were analysed separately.

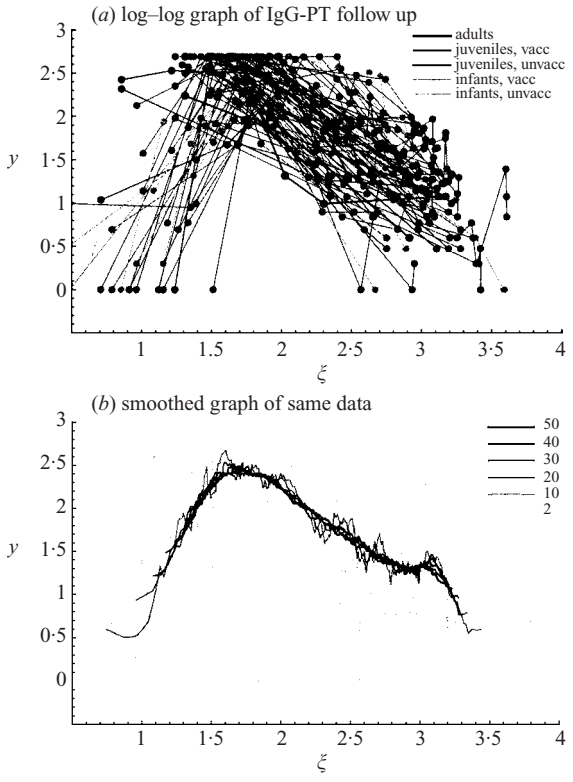

Fig. 2. Data set of log-transformed IgG-values against log time (in days) from all patients diagnosed at time 0 with symptomatic infection by *Bordetella pertussis*. Individual responses are connected, and log-transformed responses of vaccinated juvenile patients against log-time, after application of a smoothing kernel (moving average, bandwidth as in inset), showing linear increase and decrease.

Only changes in the highly specific anti-pertussis IgG-titres were analysed. Titres were determined by enzyme-linked immunosorbent assay (ELISA) [11, 13], and expressed in arbitrary units ('dutch units per ml'). In Figure 2 all measurements from all included patients are shown. Reproducibility of measurements was checked according to Standard Operating Procedure (SOP: coefficient of variation in log-transformed readings from three control sera (low, medium and high titre) less than 20%). For quantitative analysis, the scatter in these measurements therefore was considerable.

Below 5 U/ml the ELISA test is considered insensitive, therefore 5 U/ml was interpreted as 5 U/ml or less. Serial dilutions were used to an upper limit corresponding to a concentration of 500 U/ml: a value of 500 U/ml was therefore interpreted as 500 U/ml or higher.

Each patient record started with the time the first symptoms occurred. Strictly speaking, we did not estimate times from infection, but times from first symptoms (or even first diagnosis). Since incubation periods may show variation among individual patients, individual responses could have been slightly offset relative to each other.

Response model

During the decreasing phase, the change in IgG titre with time appeared to be less steep than an exponential decay function. In a log–log graph, an exponential relation is a convex function with increasing negative slope with log-time. Our data did not seem to indicate such behaviour. Instead, in a log–log graph, decay appeared to be more or less linear (Fig. 2). This called for a power function as a model, which on a log–log scale would be a straight line with arbitrary slope. However, observations were taken at random points in time, usually starting somewhere during the rising phase of the response. Fitting a straight line to the log–log transformed data, as reported by de Melker [14], only accounted for the decaying phase of the response, and required omission of these initial observations. Since there was no clearly defined criterion by which observations would be excluded, we chose to use a response model that included the initial rising phase.

Let $g(t;\boldsymbol{\theta})$ denote the IgG-response on a linear scale, and $f(\xi;\boldsymbol{\theta})$ its $\log_{10}$ as a function of log time $\xi = \log_{10}(t)$ with parameter vector $\boldsymbol{\theta}$

$$f(\xi;\boldsymbol{\theta}) = d + c\left[\xi + b\left(1 - \sqrt{1 + \frac{\xi^2}{a}}\right)\right] \tag{1}$$

is a skewed hyperbola with linear asymptotes with arbitrary slopes for $\xi \to -\infty$ and $\xi \to \infty$, respectively, and parameter vector $\boldsymbol{\theta} = (a, b, c, d)$. For $\xi \to -\infty$ and $\xi \to \infty$ this function approaches the asymptotes

$$\lim_{\xi \to \mp\infty} d + bc + c\xi\left(1 \pm \frac{b}{\sqrt{a}}\right) - f(\xi;\boldsymbol{\theta}) = 0. \tag{2}$$

The parameter d can be interpreted as the amplitude of the response, and the parameter a mainly determines the long-term decrease of the response.

Figure 3 illustrates the shape of this function, in a non-transformed (lin–lin) graph to show the very slow decline with time.

Model-fitting procedure

The response model [equation (1)] can be used to calculate expected values of the logarithm of the IgG titre.

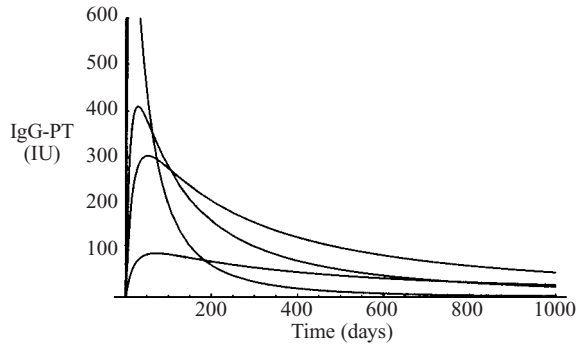

Fig. 3. Representative individual response curves generated by the model in equation (1), fitted to data from vaccinated juvenile patients. These curves illustrate characteristics of the responses: rapidly rising titres, followed by very slow decrease, and variation in both levels and steepness of the curves.

Assuming normally distributed measurement errors, with standard deviation σ, the log-likelihood function can be written as

$$L_1(\boldsymbol{\theta},\sigma) = \sum_i \log(\phi[X_i; f(\Xi_i;\boldsymbol{\theta}); \sigma]) + \sum_j \log\big(\Phi[Y_j; f\big(\Xi_j;\boldsymbol{\theta}\big); \sigma]\big) + \sum_k \log(1-\Phi[Z_k; f(\Xi_k;\boldsymbol{\theta}); \sigma]) \quad (3)$$

with $\phi[X_i; f(\Xi_i;\boldsymbol{\theta}), \sigma]$ a normal probability density function (expected value $f(\Xi_i;\boldsymbol{\theta})$, standard deviation σ) for the contribution of a measured log-titre X_i at log-time Ξ_i. For a log-titre Y_j equal to, or below the lower measurement threshold, the contribution is given by the cumulative normal distribution function $\Phi[Y_j; f\big(\Xi_j;\boldsymbol{\theta}\big), \sigma]$. For a measurement Z_k equal to or above the upper threshold, the appropriate contribution is $1-\Phi[Z_k; f(\Xi_k;\boldsymbol{\theta}), \sigma]$. In our case the lower threshold in the ELISA test was 5 U/ml, and the upper threshold was 500 U/ml. Parameter values of the response model (a, b, c, d) and the error of the detection method would be determined using this likelihood function.

Initially, our response model was fitted to the pooled data, neglecting individual variation. In this 1-level model measurement error was the only possible source of uncertainty. In order to assess variation in responses among individual subjects we would have fitted our model to the measurements of each individual response, thereby generating sets of parameters $(\boldsymbol{\theta}_i, \sigma_i)$ for every individual i in the population. But since individual responses often consisted of only a few measurements, we decided to allow only one or two parameters to vary among individual subjects. Leaving aside the measurement error parameter σ, fitting a model equation with four parameters to a population of n subjects thus generates $n+3$ parameter values (or $2n+2$, where two parameters vary among individual responses), when 3 (2) parameters are equal for all subjects, and n ($2n$) estimates for the (third and) fourth parameter to describe variation among subjects.

In our response model in equation (1), the parameters d describing the vertical offset on a log-scale (or the amplitude of the response), and a describing the shape of the response could be employed to model variation among individuals; the parameters b and c were shared by the whole population. The parameter σ (the measurement error) was assumed to be independent of the individual whose serum was being analysed.

Therefore, our final model included individual variation in both a and d, leaving us with a set $\{(a_1, \ldots, a_n), b, c, (d_1, \ldots, d_n), \sigma\}$ to be determined by optimization of the likelihood function

$$L_2[(a_1, \ldots, a_n), b, c, (d_1, \ldots, d_n), \sigma] = \sum_{i=1}^{n} L_1(a_i, b, c, d_i, \sigma). \quad (4)$$

In order to avoid cumbersome likelihood optimization procedures this two-level model was analysed by a Markov chain Monte Carlo (MCMC) method, using the Metropolis–Hastings algorithm [15]. Parameters were log-transformed, initial values were set at the values found for the one-level model (individual values for $\log(a_i)$ and $\log(d_i)$ all equal) and prior distributions for the log-transformed parameters were uniform probability distributions with an interval of $(-3, +3)$ log-units about the initial value (wider intervals were checked and did not produce different results). For each iteration in the Markov chain the likelihood value was tabulated. The model was then allowed to run for about 10 000 iterations and a trend test was used to confirm stationarity of the series of likelihood values (equal means in successive bisections of the series). The iteration with the highest likelihood value was then chosen from this set of likelihood values

$$\hat{L}_2[(\hat{a}_1, \ldots, \hat{a}_n), \hat{b}, \hat{c}, (\hat{d}_1, \ldots, \hat{d}_n), \hat{\sigma}] = \max_{j\in 1,\ldots,m} L_2[(a_{1,j}, \ldots, a_{n,j}), b_j, c_j, (d_{1,j}, \ldots, d_{n,j}), \sigma_j]. \quad (5)$$

This was taken as an approximation of the true maximum likelihood parameter set. When the resulting response functions were plotted against separate

Table 1. *Maximum log-likelihoods for the two-level model applied to separate patient categories*

	No. patients	One-level		Two-level	
		No. par*	$-2\log(L)$	No. par	$-2\log(L)$
Juv, vacc	62	5	536·4	127	434·0
Juv, unvacc	7	5	48·9	17	33·6
Adults	6	5	25·0	15	6·6
Infants, vacc	4	5	10·5	11	0·5
Infants, unvacc	7	5	44·5	17	4·6

* No. par = number of parameters.

Table 2. *Maximum log-likelihoods for joint categories, and differences* $\Delta[-2\log(L)]$ *with separately fitted model. The last column shows the significance of the difference* (*likelihood ratio test, as explained in the text, model fitting procedure*) *at the 0·95 level. Therefore, responses from all categories must be considered different, given the proposed model*

	(log)likelihood		(log)likelihood ratio		
	D.F.	$-2\log(L)$	Δ D.F.	$\Delta[-2\log(L)]$	Significance
Juv, vacc + unvacc	141	486·0	3	18·4	+
Juv, vacc + unvacc, adults	153	524·0	6	49·7	++
Infants, vacc + unvacc	25	17·8	3	12·7	+
All	175	588·2	12	113·1	++

individual observed responses, this seemed to be a reasonable assumption. The likelihood values also appeared to fall well below the values found for the one-level model and those of alternative two-level models with either a or d varying among individuals. Compared to a 'likelihood supremum' obtained by using the observed tires instead of the model function (using the measurement error from the fitted model) a significant improvement in goodness of fit was found (likelihood ratio tested against χ^2 deviate). This can be explained by the apparent lack of fit early in the responses (Fig. 5*a*). It should be noted that when individual responses were fitted, each response curve was based mostly on only a few (sometimes as few as three) observations. Therefore, the χ^2 approximation of the deviance function should be treated with caution.

For a sequence of log-times 5, 50 and 95% percentiles were determined for these maximum likelihood fits. The median curve was used to describe the time–titre relation, while the 5 and 95% percentiles indicated the magnitude of the variation among members of this population.

In addition to the maximum likelihood estimates, the Monte Carlo method provided information on parameter uncertainty, as the Markov chains for all parameters could be regarded as a sample from their joint posterior distribution.

Patients were categorized according to age (adults older than 20 years, infants younger than 6 months, and juveniles older than 6 months and younger than 20 years) and vaccination status (unvaccinated and vaccinated). To test whether the two categories could be merged, given a certain regression model, a likelihood ratio test was used as follows: first, calculate maximum likelihood values for each separate data category (say L_a and L_b). Then, pool the data and maximize the likelihood for the merged data (yielding a log-likelihood L_{ab}). The difference $-2(L_{ab}-L_a-L_b)$ would now be tested against a χ^2 variate with degrees of freedom equal to the number of parameters of the regression model [16]. Various combinations of categories were tested using this method.

RESULTS

Table 1 summarizes the results of application of the two-level model to the various patient categories (age, vaccination status). In Table 2 results of the likelihood ratio test for merging various combinations of categories are given. None of the response categories

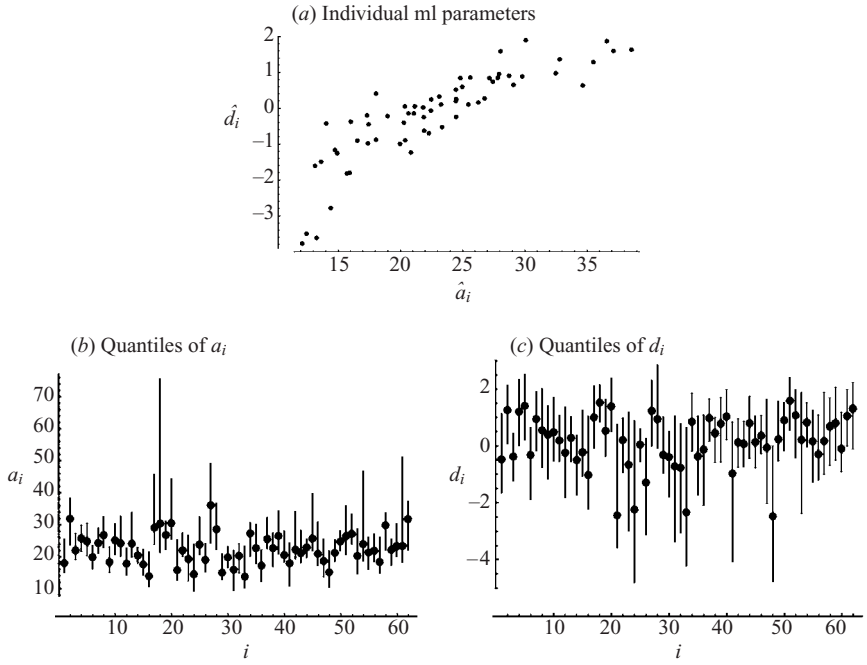

Fig. 4. (*a*) Maximum likelihood values of individual shape parameters ($\hat{a}_i$) and amplitude parameters ($\hat{d}_i$) for the two-level model fitted to vaccinated juvenile patients. (*b*) Variation in shape (parameter a_i) among patients (rank number i, random order) expressed by plotting MCMC-based median values and 0·05 and 0·95 quantiles. (*c*) Variation in amplitude (parameter d_i) among patients expressed by plotting MCMC-based median values and 0·05 and 0·95 quantiles.

could be merged; differences were significant for all the tested combinations. Given the applied model, the IgG-PT response to infection with *B. pertussis* was seen to change with age and with vaccination status.

Figure 5 shows two-level model fits to individual responses, illustrating large variations in shape and amplitude of responses in individual patients.

Although statistically significant, the differences due to vaccination status within the two youngest age categories, infants and juveniles, were not large (Figs 5*a*, *b*, and 5*c*, *d*, respectively). A small group of adult patients (Fig. 5*e*) appeared to have a response that differed from the younger patients. The most prominent difference was the lack of early measurements: these adult patients apparently visited their physician later than the younger patients. This may have been associated with milder symptoms in adults (for instance causing a parent to infect a child who then developed severe enough symptoms for the infection to be detected in both the parent and the child). The decay of the response in these adults did not differ greatly from that of the vaccinated juveniles.

Table 3 and Figure 4 show estimated parameter values for the largest category, vaccinated juvenile patients (comprising 62 patients, 72% of all patients). Maximum likelihood parameter values (with median; 5%, 95% percentiles) are shown in Table 3. For parameters a_i (variation in shape of the response) and d_i (variation in amplitude of the response), fitted parameter values for all patients in this category are shown in Figure 4. The uncertainty in these parameters is illustrated in Figure 4*b*, *c*, which show median values with (MCMC-based) 95% intervals.

Measurement error

In the one-level model, all variation is interpreted as measurement error. Therefore, the value of the parameter σ is larger than in the two-level model, where part of the variation is attributed to heterogeneity in patient responses. Despite the considerable individual variation, the estimated measurement error is very large. In the vaccinated juvenile patients, the logarithm of this error factor decreased from 0·62 in the one-level model to 0·46 in the two-level model. This would mean that a 95% confidence interval for the IgG-levels would extend over approximately a factor five up or down, for the two-level model.

Time since infection as a function of titre

In order to describe the time elapsed since infection (numbers of days since the appearance of symptoms) as a function of titre, we determined the inverse function of (1).

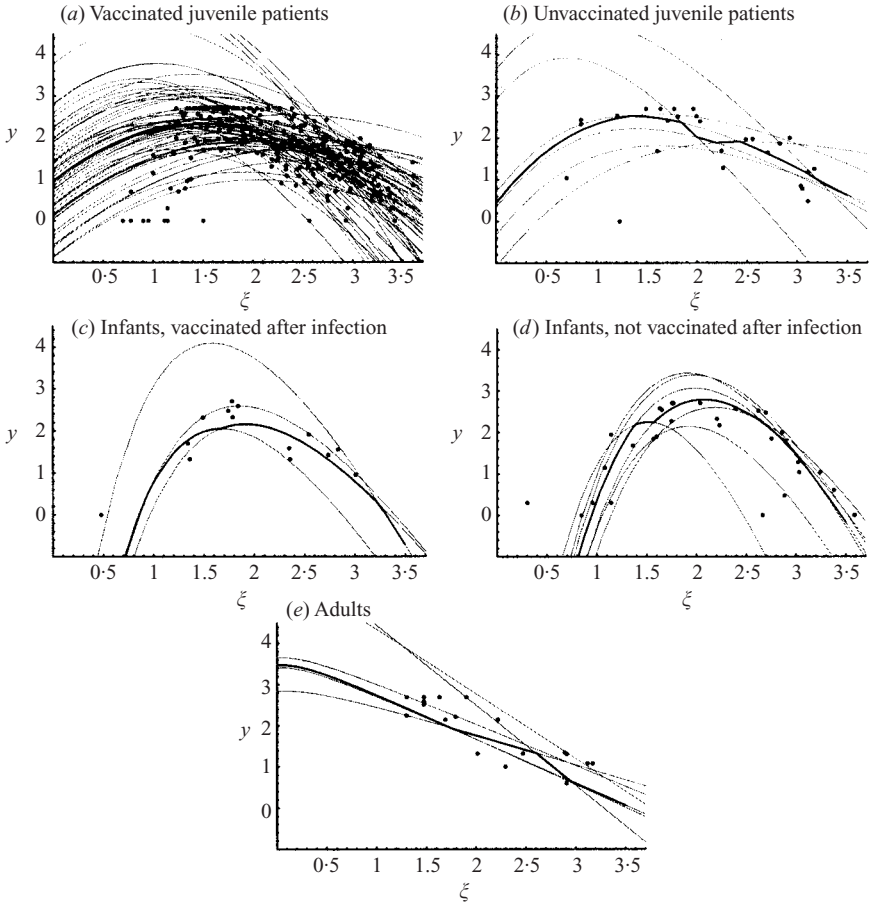

Fig. 5. (*a*) Log–log graph of individual responses of the two-level model, fitted to data from vaccinated juvenile patients. Also shown data and (heavy line) median response. (*b*) Log–log graph of individual responses of the two-level model, fitted to data from unvaccinated juvenile patients. Also shown (*c*) infants, vaccinated after infection, (*d*) infants, not vaccinated after infection, and (*e*) adults.

Table 3. *Parameter estimates for the one-level* (*1-l*) *and the two-level* (*2-l*) *model, applied to responses from patients in all categories*

	$\hat{a}$		$\hat{b}$		$\hat{c}$		$\hat{d}$		$\hat{\sigma}$	
	1-l	2-l	1-l	2-l	1-l	2-l	1-l	2-l	1-l	2-l
Juv, vacc	22·8	Fig. 4	14·0	15·7	2·68	1·90	1·00	Fig. 4	0·65	0·56
Juv, unvacc	2·15	—	2·00	2·66	6·85	3·09	2·86	—	0·61	0·46
Adults	0·12	—	1·87	4·23	0·25	0·11	−2·81	—	0·45	0·32
Infants, vacc	1·51	—	1·46	1·58	22·3	19·5	13·5	—	0·34	0·22
Infants, unvacc	2·59	—	2·08	2·29	17·2	18·9	11·4	—	0·54	0·31

Given the short duration of the rising part of the response, we considered only the decreasing part of the response. Each response curve had a maximum titre, dependent on the parameter values for an individual patient. Any titres higher than this maximum therefore had no corresponding time since infection. On a log scale ($y = \log_{10}$(IgG titre)), the inverse relation was

$$h(y;a,b,c,d) = \frac{b\sqrt{a}\sqrt{ac^2+(d-y)(2bc+d-y)}-a(bc+d-y)}{c(b^2-a)},$$

with $y < d+bc-c\sqrt{b^2-a}$.

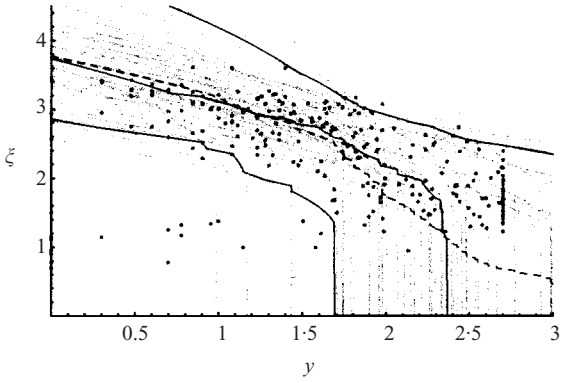

Fig. 6. Expected log-time since infection (ξ) as a function of anti-pertussis log-IgG titre (y). Inverse function of the two-level regression model, using maximum likelihood values for all parameters. Also shown (heavy lines) median and 95% range, and (hatched) arithmetic mean response. Each light grey curve represents an individual response, with its own maximum level. Titres higher than this individual maximum cannot be reached and the inverse response jumps to zero.

This inverse function described (log-)time since last infection as a function of (log-)titre. Figure 6 shows that this inverse function was also subject to considerable individual variation. Given a certain measured IgG-level we now estimated the (variation in) time elapsed since infection.

When doing this, we took into account the fact that not every individual response reached the same maximum IgG titre. Therefore, with IgG titres rising above a certain level, the inverse function did not exist for an increasing number of patients. This is shown in Figure 7*a*, where the fraction of the individual responses (of vaccinated juvenile patients) that reached a given log-titre Y, is shown as a function of Y.

Figure 7*b* shows histograms of the times from infection for a range of IgG-titres, constructed from the inverse responses in Figure 6. Returned estimates were based on decreasing numbers of responses: only those that reached the level of interest were used. Some technical details are given in the Appendix.

DISCUSSION

The serum IgG-response in patients diagnosed with pertussis can be described by a simple parametric model, but variation among subjects is a prominent phenomenon that must be included in the analysis. IgG titres rise steeply after infection to reach a maximum within a few weeks, and then decline slowly over several years. Although this is probably true in any infected subject in a population, incidental measurement of a single IgG titre gives little information about the time since infection. In particular, maximum IgG titres appear to show considerable heterogeneity; a modestly enhanced IgG titre of 20 IU/ml may correspond to very recent infection, or may equally correspond to infection several months earlier, depending on the particular response of the patient. Current practice of characterizing population immune responses with their geometric mean titres neglects the importance of heterogeneity [8, 9, 17].

If the group of patients in the study used here is representative of the general population, at least with respect to their IgG response to pertussis toxin, our results may be employed to estimate times from infection from cross-sectional samples of IgG-PT titres. Any given (measured) IgG-PT titre may thus be assumed to correspond with a time from infection (first diagnosis). Our results also indicate the existence of considerable variation among individual responses. With increasing IgG-PT levels, there also seems to be an increasing fraction of subjects who do not reach

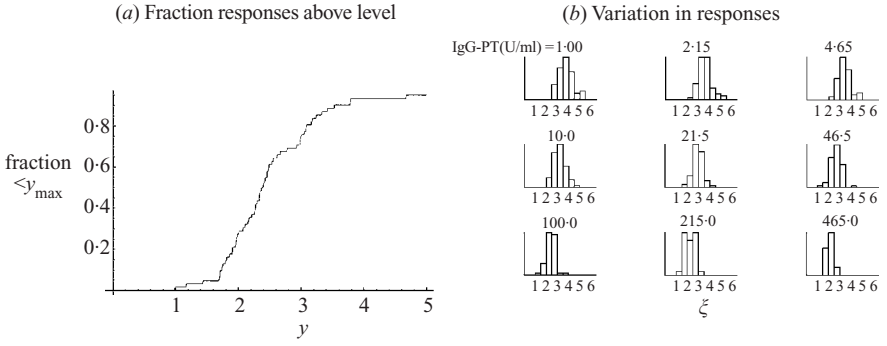

Fig. 7. (*a*) Fraction of the patients (vaccinated juveniles) who failed to reach a level y (log-titre). Even at low titres a considerable proportion of the patients do not reach that level and cannot contribute to the estimated time since infection. (*b*) Variation in log-time since infection (ξ) for various levels of anti-pertussis IgG titre.

that level at any time during their post-infection response. High IgG-PT levels therefore indicate recent infection, but only in a small part of the population.

Data set

To apply the results of this analysis to the estimation of incidences of pertussis in the general population, the observed responses should be representative of the responses that would be found in a random sample of this general population. The age distribution in our test population differed substantially from that of the population at large, with juveniles aged less than 20 years strongly overrepresented. Older subjects were hardly present. The heterogeneity among all patient categories in this study complicates the use of these results for back-calculation of infection from a cross-sectional sample of a completely different population. Estimation of times from infection only involved the decaying phase of the response, where differences may have been least marked. This would easily be tested in a practical application, which we intend to report shortly.

Newborns may be important for illness burden, and they should possibly be treated separately when back-calculating times from infection. Unfortunately, we had responses from only a few of these patients, but since infection cannot have occurred before birth, time from infection was limited in these infants.

Recorded responses started at the first date of symptoms, as reported by the patients (or their parents). In very young children, this may have been earlier than in older patients, because of the severity of symptoms. On the other hand, in newborns symptoms may be non-specific, making diagnosis difficult or delaying diagnosis. An alterative model employing a variable shift in onset (at the expense of yet another parameter per patient) only produced minor offsets, indicating that the data offered little support for such differences. Nevertheless, the difference in response shapes between infants and adults (Fig. 5*c–e*) is striking and seems to support an effect of late onset of symptoms.

The patients included in this study all presented symptoms of respiratory illness. Asymptomatic or mild infection is probably more frequent in adults, and could be associated with a different IgG response: a smaller amplitude, or different time course. No information was available to test this hypothesis. Most of the juvenile patients also suffered from chronic respiratory symptoms (asthma); it remains unclear whether this condition increases the susceptibility to infection or could lead to a different immune response to pertussis. Are younger subjects more susceptible to infection, or is there a higher probability that infection is symptomatic [18, 19]? Recently, an animal model with infection and clinical symptoms similar to those in humans has been described [20]. It is conceivable that such a model could be utilized to study details of the differences in susceptibility between adults and infants.

Regression procedure

Heterogeneity in immune responses has been studied for various pathogens. In order to assess immunity, Gay used mixture models to describe the variation in age stratified IgG levels against parvovirus B19 [21]. Hierarchical Bayesian models have been used to describe the decline in immunity and its variation in hepatitis B [15] and *Haemophilus influenzae* [22, 23]. These models assumed linear decline of IgG titres (on a log scale). The immune response against hepatitis A vaccine has been studied by van Herck and colleagues [17, 24]. They studied the decline in geometric mean antibody titre of their subjects with time, but also reported on individual responses, presenting evidence of considerable variation among subjects [17].

Since our data included not only the decline but in many cases also the rising slope of the immune response, we needed a model that accommodated both parts of the response, necessarily a non-linear model. Our data also contained strong evidence of censoring; in accounting for this, the fitted responses were steeper than they would have been without correction for censoring.

Part of the variation in the measurements appeared to be explained by individual variation, in amplitude and/or in shape (descending slope) of the response, rather than by some measurement error. In addition to this, other parameters, like incubation period, could also have contributed to the heterogeneity of the responses. We investigated this by incorporating individual variation in shift along the time axis (not shown here). However, this did not result in significant improvement in goodness of fit (judged by log-likelihood); estimated delay times were also very small and appeared to vary little among individual patients.

Biological interpretation of the hyperbola model

The extremely slow decline in IgG antibody titres with time following infection, probably associated with

long-term protection, precludes use of a first order model of antibody decay. Such a model would lead to a convex curve, with an increasing downward slope on a log time–log titre scale. For our data, such a slope is much too steep, with poor fit to the measured responses. This may not be always a problem. Tiru et al. [25] reported successful use of a simple exponential model to describe the immune response to diphtheria toxin. The model we used had asymptotically linear decline on a log–log scale at long times from infection.

Although at first sight perhaps biologically unattractive, such a hyperbolic response function may result from an intrinsically first order system, when there is heterogeneity. Suppose IgG production and removal are distributed among several locations, all with different properties (different time constants and amplitudes). Sufficient heterogeneity among these sites could then result in a varying contribution of any given subsystem with time, with fast systems providing the bulk of the IgG in the initial phase of the response, and 'recruiting' a smaller fraction of slower systems in the tail of the response.

Our analysis indicates that the IgG-PT response to an infection with *Bordetella pertussis* shows a typical pattern, with a rapid transient increase over a few days to weeks, followed by slow decay extending over several years. This response can be described with a mathematical model; there appears to be considerable variation among responses from individual patients. The remarkably large variation in responses cannot be neglected when these are employed to estimate the time since infection from a given IgG-titre in a randomly chosen subject.

APPENDIX

Calculation of times since infection

From the above analysis of the longitudinal study we have a function describing the IgG-PT level against time from infection (first diagnosis)

$$y = f(x; \boldsymbol{\Theta})$$

in which the parameter vector $\boldsymbol{\theta}$ is stochastic, describing the variation in responses among patients. Suppose we are able to invert this function, i.e. given $\boldsymbol{\Theta}$ and y (IgG-PT level) we can find a corresponding time since infection

$$x = f^{-1}(y; \boldsymbol{\Theta}) = h(y; \boldsymbol{\Theta}).$$

We are only interested in the decreasing part of this function, leaving a monotonically descending function of titre with time from infection.

If times to infection ξ are distributed as $g(\xi; \boldsymbol{\lambda})$ the response of patient with response function parameter vector $\boldsymbol{\Theta}$ would lead to a titre density

$$\gamma(y; \boldsymbol{\Theta}, \boldsymbol{\lambda}) = \begin{cases} 0 & \text{if } y > y_{\max}(\boldsymbol{\Theta}) \\ -g(h(y; \boldsymbol{\Theta}); \boldsymbol{\lambda}) h'(y; \boldsymbol{\Theta}) & \text{otherwise} \end{cases}$$

$y_{\max}(\boldsymbol{\Theta})$ is the maximum IgG-PT level reached by this patient (having parameter vector $\boldsymbol{\Theta}$). This is different for each individual patient.

If, further, any patient's response is equally likely, the probability density of titres is

$$\gamma(y; \boldsymbol{\lambda}) = \int \gamma(y; \boldsymbol{\Theta}, \boldsymbol{\lambda}) \, \mathrm{d}\boldsymbol{\Theta}.$$

Integration is done over all possible parameter combinations, and may be based on the posterior Markov chain.

If we now have a cross-sectional sample of IgG-PT titres:

$$\{Y_1, \ldots, Y_M\}$$

the likelihood

$$\ell(\boldsymbol{\lambda}) = \prod_{i=1}^{M} \gamma(Y_i; \boldsymbol{\lambda})$$

allows estimation of the parameter vector $\boldsymbol{\lambda}$ of the distribution of times to infection $g(\xi; \boldsymbol{\lambda})$.

REFERENCES

1. de Melker HE, Versteegh FGA, Conyn van Spaendonck MAE, et al. Specificity and sensitivity of high levels of immunoglobulin G antibodies against pertussis toxin in a single serum sample for diagnosis of infection with *Bordetella pertussis*. J Clin Microbiol 2000; **38**: 800–6.
2. Guris 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 1999; **28**: 1230–7.
3. Tomoda T, Ogura H, Kurashige T. Immune responses to *Bordetella pertussis* infection and vaccination. J Infect Dis 1991; **163**: 559–63.
4. Tomoda T, Ogura H, Kurashige T. The longevity of the immune response to filamentous hemagglutinin and pertussis toxin in patients with pertussis in a semiclosed community. J Infect Dis 1992; **166**: 908–10.
5. Isacson J, Trollfors B, Hedvall G, Taranger J, Zackrisson G. Response and decline of serum IgG antibodies to

pertussis toxin, filamentous hemagglutinin and pertactin in children with pertussis. Scand J Infect Dis 1995; **27**: 273–7.
6. Giammanco A, Taormina S, Genovese M, Mangiaracina G, Giammanco G, Chiarini A. Serological responses to infection with *B. pertussis*. Dev Biol Standard 1997; **89**: 213–20.
7. Di Tommaso A, Bartalini M, Peppoloni S, Podda A, Rappuoli R, De Magistris MT. Acellular pertussis vaccines containing genetically detoxified pertussis toxin induce long-lasting humoral and cellular responses in adults. Vaccine 1997; **15**: 1218–24.
8. Giuliano M, Mastrantonio P, Giammanco A, et al. Antibody kinetics and long-term sero-prevalence in the Italian clinical trial of acellular pertussis vaccines. Dev Biol Standard 1997; **89**: 275–8.
9. Giuliano M, Mastrantonio P, Giammanco A, Piscitelli A, Salmaso S, Wassilak SGF. Antibody responses and persistence in the two years after immunization with two acellular vaccines and one whole-cell vaccine against pertussis. J Pediatr 1998; **132**: 983–8.
10. Bell F, Martin A, Blondeau C, Thronton C, Chaplais J, Finn A. Combined diphtheria, tetanus, pertussis and *Haemophilus influenzae* type b vaccines for primary immunisation. Arch Dis Child 1996; **75**: 298–303.
11. de Melker HE, Schellekens JFP, Neppelenbroek SE, Mooi FR, Rümke HC, Conyn van Spaendonck MAE. Reemergence of pertussis in the highly vaccinated population of the Netherlands: observations on surveillance data. Emerg Infect Dis 2000; **6**: 348–57.
12. Versteegh FGA, Schellekens JFP, Nagelkerke AF, Roord JJ. Laboratory-confirmed reinfections with *Bordetella pertussis*. Acta Paediatr 2002; **91**: 95–7.
13. Nagel J, de Graaf S, Schijf Evers D. Improved serodiagnosis of whooping cough caused by *Bordetella pertussis* by determination of IgG anti-LPF antibody levels. Dev Biol Standard 1985; **61**: 325–30.
14. de Melker HE. Seroepidemiology of diphtheria, tetanus, poliomyelitis and pertussis. PhD thesis, Wageningen University, 1999.
15. Gilks WR, Richardson S, Spiegelhalter DJ, editors. Markov Chain Monte Carlo in practice. London: Chapman and Hall, 1996.
16. Teunis PFM, Evers EC, Slob W. Analysis of variable fractions resulting from microbial counts. Quant Microbiol 1999; **1**: 63–88.
17. van Herck K, Beutels P, van Damme P, et al. Mathematical models for assessment of long-term persistence of antibodies after vaccination with two inactivated hepatitis A vaccines. J Med Virol 2000; **60**: 1–7.
18. Huang LM, Lee CY, Lin TY, Chen JM, Lee PI, Hsu CY. Responses to primary and a booster dose of acellular, component, and whole-cell pertussis vaccines initiated at 2 months of age. Vaccine 1996; **14**: 916–22.
19. Van Savage J, Decker MD, Edwards KM, Sell SH, Karzon DT. Natural history of pertussis antibody in the infant and effect on vaccine response [see comments]. J Infect Dis 1990; **161**: 487–92. J Infect Dis 1990; **161**: 473–9.
20. Hall E, Parton R, Wardlaw AC. Time-course of infection and responses in a coughing rat model of pertussis. J Med Microbiol 1999; **48**: 95–8.
21. Gay NJ. Analysis of serological surveys using mixture models: application to a survey of parvovirus B19. Stat Med 1996; **15**: 1567–73.
22. Auranen K, Eichner M, Kayhty H, Takala AK, Arjas E. A hierarchical Bayesian model to predict the duration of immunity to *Haemophilus influenzae* type b. Biometrics 1999; **55**: 1306–13.
23. Auranen K. Back-calculating the age-specific incidence of recurrent subclinical *Haemophilus influenzae* type b infection. Stat Med 2000; **19**: 281–96.
24. van Herck K, van Damme P. Inactivated hepatitis a vaccine-induced antibodies: follow-up and estimates of long-term persistence. J Med Virol 2001; **63**: 1–7.
25. Tiru M, Hallander HO, Gustafsson L, Storsaeter J, Olin P. Diphtheria antitoxin response to DTP vaccines used in Swedish pertussis vaccine trials, persistence and projection for timing of booster. Vaccine 2000; **181**: 2295–306.