

## Statistical analysis of the dynamics of antibody loss to a disease-causing agent: plague in natural populations of great gerbils as an example

Siyun Park, Kung-Sik Chan, Hildegunn Viljugrein, Larissa Nekrassova, Bakhtiyor Suleimenov, Vladimir S Ageyev, Nikolay L Klassovskiy, Sergey B Pole and Nils Chr. Stenseth

*J. R. Soc. Interface* 2007 **4**, 57-64  
doi: 10.1098/rsif.2006.0160

### Supplementary data

["Data Supplement"](http://rsif.royalsocietypublishing.org/content/suppl/2009/03/25/4.12.57.DC1.html)

<http://rsif.royalsocietypublishing.org/content/suppl/2009/03/25/4.12.57.DC1.html>

### References

[This article cites 11 articles, 4 of which can be accessed free](http://rsif.royalsocietypublishing.org/content/4/12/57.full.html#ref-list-1)

<http://rsif.royalsocietypublishing.org/content/4/12/57.full.html#ref-list-1>

Article cited in:

<http://rsif.royalsocietypublishing.org/content/4/12/57.full.html#related-urls>

### Email alerting service

Receive free email alerts when new articles cite this article - sign up in the box at the top right-hand corner of the article or click [here](#)

To subscribe to *J. R. Soc. Interface* go to: <http://rsif.royalsocietypublishing.org/subscriptions>

# Statistical analysis of the dynamics of antibody loss to a disease-causing agent: plague in natural populations of great gerbils as an example

Siyun Park<sup>1</sup>, Kung-Sik Chan<sup>2,\*</sup>, Hildegunn Viljugrein<sup>3</sup>, Larissa Nekrassova<sup>4</sup>,  
Bakhtiyor Suleimenov<sup>4</sup>, Vladimir S. Ageyev<sup>4</sup>, Nikolay L. Klassovskiy<sup>4</sup>,  
Sergey B. Pole<sup>4</sup> and Nils Chr. Stenseth<sup>3</sup>

<sup>1</sup>Department of Statistics, Seoul National University, Sillim-dong, Gwanak-gu,  
Seoul 151-742, South Korea

<sup>2</sup>Department of Statistics and Actuarial Science, University of Iowa, Iowa City,  
IA 52242, USA

<sup>3</sup>Centre for Ecological and Evolutionary Synthesis (CEES), Department of Biology,  
University of Oslo, PO Box 1066 Blindern, N-0316 Oslo, Norway

<sup>4</sup>M. Aikimbaev's Kazakh Scientific Centre for Quarantine and Zoonotic Diseases,  
14 Kapalskaya Street, Almaty 480074, Republic of Kazakhstan

We propose a new stochastic framework for analysing the dynamics of the immunity response of wildlife hosts against a disease-causing agent. Our study is motivated by the need to analyse the monitoring time-series data covering the period from 1975 to 1995 on bacteriological and serological tests—samples from great gerbils being the main host of *Yersinia pestis* in Kazakhstan. Based on a four-state continuous-time Markov chain, we derive a generalized nonlinear mixed-effect model for analysing the serological test data. The immune response of a host involves the production of antibodies in response to an antigen. Our analysis shows that great gerbils recovered from a plague infection are more likely to keep their antibodies to plague and survive throughout the summer-to-winter season than throughout the winter-to-summer season. Provided the seasonal mortality rates are similar (which seems to be the case based on a mortality analysis with abundance data), our finding indicates that the immune function of the sampled great gerbils is seasonal.

**Keywords:** continuous-time Markov chain; generalized nonlinear mixed-effect model; specific immunity; serology; *Yersinia pestis*

## 1. INTRODUCTION

We consider the problem of studying the dynamics of the immune responses of wildlife hosts to infectious disease, based on monitoring time-series data. Our motivating case study concerns the modelling of the dynamics of antibodies to plague in the blood of great gerbils (*Rhombomys opimus*) in Kazakhstan. Plague in great gerbils is of general interest, not the least since the great gerbil is the main host in a large part of Central Asia where most of the larger epidemics might have arisen; see Achtman *et al.* (2004). There are many gaps in the knowledge about the dynamics and the natural history of plague in its natural reservoirs. Yet, as remarked by Begon *et al.* (2006) ‘The control of plague,

in particular the control of the risk of it spreading from its wildlife hosts into peridomestic animals and humans, depends on understanding the dynamics and natural history of plague in those wildlife hosts’.

Our new approach provides a framework for studying the dynamics of the immune response of the hosts in terms of past and recent prevalence rates of the disease, and the host population fluctuations. It allows us to assess the interesting question of whether or not the dynamical structure of the immune response is seasonal. The significance of understanding the seasonal structure of the epidemics is as follows. Plague cannot easily persist in a population if infected individuals all die of plague. The presence of specific immunity in the host species helps the host individuals to survive an infection. Moreover, during periods with a lower level of specific immunity (population average), infected individuals are likely to suffer more severe bacteraemias, thus infecting more fleas and hence facilitating the spread of plague

\*Author for correspondence (kchan@stat.uiowa.edu).

Electronic supplementary material is available at <http://dx.doi.org/10.1098/rsif.2006.0160> or via <http://www.journals.royalsoc.ac.uk>.

among gerbils and to alternative host species (increased risk for human plague cases). Consequently, the seasonal dynamics of increased and decreased specific immunity may facilitate the persistence of plague in the host population. Note that the new methodology is rather general and applicable to the study of the immune response to other infectious diseases.

This article is organized as follows. In §2, we provide some general background information about the plague system as it is found in Kazakhstan. In §3, we provide further details on the monitoring data, the protocol of the bacteriological and the serological tests. In §4, we derive a generalized nonlinear mixed-effect model for analysing the serological test data. Details of the derivation are collected in the electronic supplementary material. Section 5 presents the results of the data analysis and §6 presents the conclusion.

## 2. THE PLAGUE SYSTEM

Plague exists in nature as a disease of wild rodents caused by infection of the bacterium, *Yersinia pestis*. The infection is maintained in natural foci of the disease in wild rodent colonies through transmission between rodents by their flea ectoparasites. Rodents are the primary hosts of *Y. pestis*; however, other mammals, including humans, may be infected. Other species, especially rodent-consuming predators, may play an ecologically important role by transporting infected fleas from one area to another.

Plague is widespread in many countries in Africa, the Americas and Asia. In desert and semidesert areas of Kazakhstan (and Central Asia, in general), *R. opimus* (the great gerbil) and the fleas inhabiting their burrows (mainly of the genus *Xenopsylla*) are considered to be the main host and vectors, respectively, of plague (Pollizer 1966).

Great gerbils are social desert rodents living in family groups that occupy discrete, permanent burrow systems (colonies). The winter death rate of adults (i.e. the ones that have been breeding) is high (Korneyev & Karpov 1977), but single individuals may live as long as 3–4 years. Breeding takes place from April to September, with one to three litters and each ranging in size from 1 to 14 (but usually 4–7), depending on vegetation and whether there is a free space (i.e. non-occupied burrow colonies; Naumov & Lobachev 1975). Great gerbils exhibit extensive population fluctuations between years (coefficient of variation = 2.23). Years of high abundance are necessary for the large-scale spread of plague that produces epizootics (Davis *et al.* 2004).

Great gerbils are fairly resistant to plague, but there is an extensive individual variation (Rothschild 1978). Antibodies to plague do not provide full protection, but if infected with plague, the disease is likely to be mild. Resistance to plague may differ dramatically among individuals and populations, and the rate may change within populations depending on how recently they were exposed to plague. Resistant great gerbils rarely die of plague, but they are likely to become bacteremic (i.e. with bacteria present in the blood); therefore, they may serve as sources of infectious blood meals for feeding fleas. The rodents are infectious only in

the relatively short period of bacteremia of any intensity. The incubation period is 3–5 days.

## 3. THE MONITORING DATA FROM KAZAKHSTAN

The great gerbil populations constitute several natural plague foci (caused by the bacteria *Y. pestis*) in Kazakhstan where the disease may be transmitted to humans by vectors, mainly fleas. A long-term monitoring study of this natural plague system was undertaken from 1949 to 1995, for tracking the prevalence of plague in the great gerbil population; see Davis *et al.* (2004), Frigessi *et al.* (2005), Samia *et al.* (in press) and Stenseth *et al.* (2006). Particularly, in both spring and autumn of each year, samples of great gerbils (mean sample size = 678.6, coefficient of variation = 1.71, maximum = 4429) provided bacteriological and serological test data for plague symptoms. While bacteriological tests may detect the presence of plague bacteria and hence the plague disease in great gerbils at the time of sampling, serological tests may detect the presence of antibodies to plague bacteria. Consequently, the serological test data are indicative of past infections and may shed light on the dynamical structure of the specific immune system of the great gerbils against plague.

The survey area is located southeast of Lake Balkhash in southeastern Kazakhstan, being part of the PreBalkhash plague focus (figure 1). The PreBalkhash focus is separated into specific landscape-epizootological regions, a landscape-epizootological region being an area of a particular type of landscape, soil, vegetation, density of the gerbils and their fleas, as well as by the level of epizooticity. For the purpose of monitoring plague, the whole of Kazakhstan was divided into  $40 \times 40$  km squares, here referred to as large squares (see figure 1). Each large square comprises four  $20 \times 20$  km primary squares which in turn are divided into four sectors. Within a given sector, data are maximally recorded twice a year providing information on the results of bacteriological and serological tests (prevalence data) as well as great gerbil density.

The sampling was done during spring and autumn from 1949 to 1995. The great gerbils were mainly caught during May–June in spring and during September–October in autumn, with the bacteriological and serological tests administered to the caught gerbils. For simplicity, it is assumed that the sampling was carried out at time  $\Delta_t$  in the  $t$ th season which spans from  $b_t$  to  $L_t$ . For ease of mathematical analysis, we assume that  $\Delta_t = L_t$ , i.e. the seasons are operationally defined as the inter-sampling periods. Thus, data sampled in May–June (September–October) provide information for the autumn-to-spring (spring-to-autumn) seasons. Standardized serological tests have been carried out since 1975; hence, the analyses reported below are confined to the data collected from the period starting from 1975. The bacteriological test was positive when *Y. pestis* was isolated by planting rodent samples (blood, liver and spleen) on agar media. The bacteriological test represented a combined testing of blood, liver and spleen samples taken from each single rodent. Bacteriologically

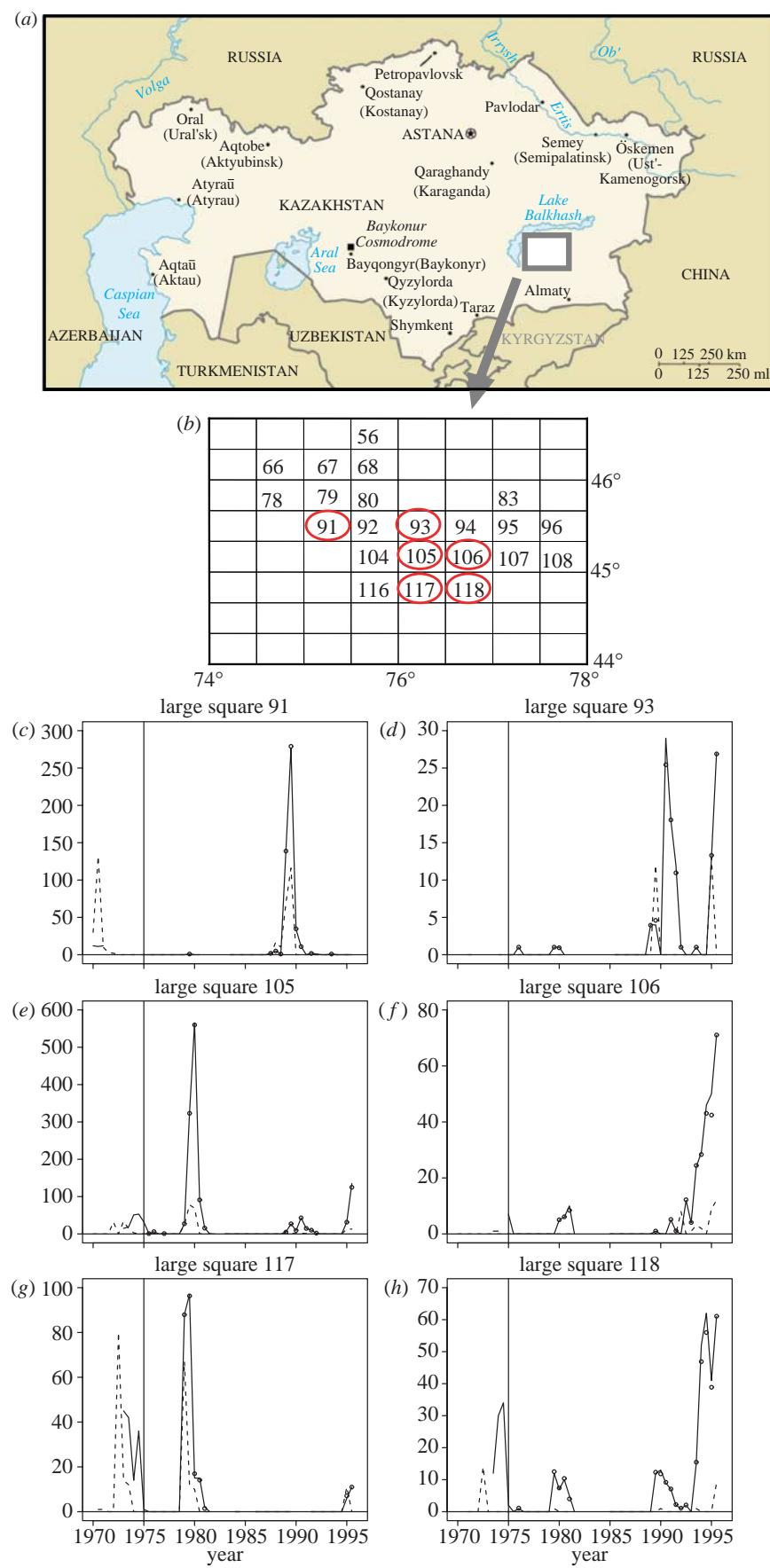


Figure 1. (a) PreBalkhash focus marked as a square on a map of Kazakhstan. (b) Large squares in the PreBalkhash focus from which we have data on the monitoring of plague prevalence. (c–h) Time-series plots of counts of bacteriological positives (dashed line) and serological positives (solid line), for six large squares. Analyses reported in the paper are based on data collected on or after 1975, i.e. for data to the right of the vertical line at 1975. Open circles represent fitted values from a fitted model reported in table 2.

positive results were found for rodents with local forms of infection in organs (liver or spleen) or with expressed bacteremia (i.e. live bacteria present in blood).

Gerbil density data were estimated in April (spring) and September (autumn) and were derived from the number of burrows per hectare in the primary square, the proportion of inhabited burrows, and sit-count observations from 10 of the inhabited burrows. The proportion of inhabited burrows and sit-count observations were measured at different locations with great gerbils different from those trapped for the bacteriological and serological tests. Consequently, the gerbil density estimates are stochastically independent of the prevalence data.

We analyse data from six large squares that have adequate data for the model estimation reported in §4. The time-series plots of counts of serological positives and those of the bacteriologically positives are depicted in figure 1. These plots display a variety of temporal patterns and diverse levels of serologically and bacteriologically positive counts. Since 1975, most large squares had at most two epizootics.

The more *Y. pestis* there are in the sample, the more likely there will be a positive test result. However, samples taken from infected individuals may not include live bacteria because the likelihood for bacteria to be included in the sample is dependent on the strength of the infection (the amount and the virulence of bacteria), the resistance to plague and the phase of the infectious process in the great gerbil caught. Hence, only a proportion of the infected gerbils will give rise to a positive bacteriological test result. Additionally, it is possible that sick animals (owing to reduced above-ground activity) are more difficult to trap than the more resistant individuals. Fortunately, the reduced sensitivity of the bacteriological test does not affect the statistical analysis presented in §4; see the electronic supplementary material.

Serological tests were done by analysing blood samples from the caught great gerbils for the presence of anti-F1 antibodies by a passive haemagglutination test and confirmed by haemagglutination inhibition with F1 antigen (MacIntyre *et al.* 2004). The F1 antigen is specific to *Y. pestis* (Perry & Fetherston 1997). Trapped individuals were killed and hence there were no repeated tests on the individuals. We note that seropositivity refers to whether an animal exhibited demonstrable levels of antibody (Begon *et al.* 2006). A positive serological test indicates recent or past exposure to plague. Negative serological tests are not evidence for the absence of plague because individual rodents may not react serologically or have very low diagnostic titres (corresponding to a low infectious dose, initial or final stages of antibody production and/or low reactivity of an individual; Suleimenov *et al.* 2001). Intensity of antibody production (and the length of their presence) depends on the immune reactivity of a rodent. The later the elimination of the plague bacteria, the clearer is the immune response of the rodent (Kanatov 1974). The immune response may also be boosted by repeated contacts with plague microbes.

#### 4. A GENERALIZED NONLINEAR MIXED-EFFECT MODEL

In this section, we aim to derive a generalized nonlinear mixed-effect model for the time-series of counts of positive serological cases in terms of past prevalence rates and rodent density estimates. Let  $B_t$  denote the number of bacteriological positives and  $S_t$  the number of serological positives out of  $G_t$  great gerbils trapped at time  $\Delta_t$  in the  $t$ th season. Let  $N_t$  denote the true number of great gerbils in the  $t$ th season and  $D_t$  the corresponding gerbil density estimate. Below, the notation  $B(n, p)$  denotes the binomial distribution with  $n$  trials and the probability of success being  $p$ . Assume that  $B_t$  has the binomial distribution  $B(G_t, p_t)$  while  $S_t$  is distributed as  $B(G_t, q_t)$ . Thus,  $p_t$  denotes the true (instantaneous) prevalence rate of plague at  $\Delta_t$  and  $q_t$  the corresponding probability of a positive serological test. Since a serologically positive case must result from the recovery of an earlier infection,  $q_t$  bears some relationship with  $p_t$  and its lags. Under some simplifying assumptions that are biologically plausible for the plague monitoring data, it is shown in the electronic supplementary material that

$$q_t = C + \sum_{j=0}^M \frac{D_{t-j}}{D_t} \tau_{t-j} p_{t-j} \prod_{k=1}^j \theta_{t-k+1} + \eta_t, \quad (4.1)$$

where a product over an empty index set is defined to be 1;  $\tau_t$  is proportional to the instantaneous recovery rate (the proportionality constant is, however, generally region specific); the maximum value of  $M$  is set to be 7 because a great gerbil rarely lives longer than 4 years (Naumov & Lobachev 1975);  $\theta_t$  is the probability of a recovered great gerbil keeping immunity (i.e. plague antibodies) and surviving throughout the  $t$ th season;  $C$  is a term for accommodating possible migration of great gerbils between neighbouring squares; and  $\eta_t$  are independent  $N(0, \sigma^2)$  errors.

Equation (4.1) can be derived heuristically as follows.  $N_t q_t$  equals the number of serologically positive great gerbils, which is the sum of the number of great gerbils most recently infected in season  $t-j$  and continue to keep the immunity until the sampling time in the  $t$ th season, where the summation is over  $j=0, 1, 2, \dots, M$ . The parameter  $M$  must be smaller than the typical life expectancy of a great gerbil, and hence set as 7 (seasons). The probability that a great gerbil acquired the most recent infection during the  $(t-j)$ th season, survived and kept immunity until the  $t$ th season can be shown to equal  $\tau_{t-j} p_{t-j} \prod_{k=1}^j \theta_{t-k+1}$ , thanks to a Markov assumption and other simplifying assumptions of the underlying process. Specifically, we show in the electronic supplementary material that the probability of a great gerbil acquired an infection during the  $(t-j)$ th season and had immunity at the end of that season equals  $\tau_{t-j} p_{t-j}$ , and given the immunity of the great gerbil to plague at the end of the  $(t-j)$ th season, the probability that it continued to have immunity until the  $t$ th season is the product  $\prod_{k=1}^j \theta_{t-k+1}$ . Thus,  $N_t q_t$  equals  $\sum_{j=0}^M N_{t-j} \tau_{t-j} p_{t-j} \prod_{k=1}^j \theta_{t-k+1}$ , and hence  $q_t = \sum_{j=0}^M \frac{N_{t-j}}{N_t} \tau_{t-j} p_{t-j} \prod_{k=1}^j \theta_{t-k+1}$ .

Since  $N_t$ 's are unknown, they will be replaced by great gerbil density estimates. Specifically, let  $D_t$  be the estimated great gerbil density. Then, a simple model

Table 1. AIC<sub>c</sub> of the model defined by equation (4.1) for four cases and with two dummy variables accounting for two outliers; see text. Case 1, constant  $\tau$  and seasonal  $\theta$ ; case 2, seasonal  $\tau$  and constant  $\theta$ ; case 3, both  $\tau$  and  $\theta$  are constants; and case 4, both  $\tau$  and  $\theta$  are seasonal. Minimum AIC<sub>c</sub> is bold-faced, while entries for models whose Hessian matrices are singular or non-negative definite are given in italics.

model	M							
	7	6	5	4	3	2	1	0
case 1	34745.0	22597.0	1855.1	<b>729.2</b>	729.4	730.1	731.6	736.7
case 2	34747.0	22599.0	1858.1	732.2	732.2	732.4	732.9	736.7
case 3	34761.0	22614.0	1874.9	749.2	749.4	749.8	749.9	755.0
case 4	34748.0	22600.0	1857.6	731.6	732.1	732.7	734.5	<i>739.6</i>

relating the density estimates to  $N_t$  is:  $d_t = K + n_t + \epsilon_t$ , where  $K$  is a constant;  $d_t = \log(D_t)$ ;  $n_t = \log(N_t)$ ; and  $\epsilon_t$  are iid  $N(0, \sigma^2)$  random variables. Consequently,  $D_{t-j}/D_t$  equals

$$\exp(n_{t-j} - n_t) \exp(\epsilon_{t-j} - \epsilon_t) = \frac{N_{t-j}}{N_t} \exp(\epsilon_{t-j} - \epsilon_t).$$

This implies that

$$q_t = \sum_{j=0}^M \frac{D_{t-j}}{D_t} \tau_{t-j} p_{t-j} \prod_{k=1}^j \theta_{t-k+1} \exp(\epsilon_t - \epsilon_{t-j}).$$

The rather complex stochastic structure can be simplified by (i) noting that  $\exp(\epsilon_t - \epsilon_{t-j})$  can be approximated by  $1 + (\epsilon_t - \epsilon_{t-j})$ , and (ii) representing the combined random terms by a single term, i.e. we specify equation (4.1). (Strictly speaking, the error terms in equation (4.1) are heteroscedastic, but they are assumed to be homoscedastic for simplicity.) Equation (4.1) now follows by adding an intercept that accounts for between-square migration of the great gerbil.

For parsimony of modelling, it is furthermore assumed that  $\tau_t$  and/or  $\theta_t$  are seasonal, i.e.

$$\tau_t = \begin{cases} \tau_S & \text{for spring (winter-to-summer season)} \\ \tau_F & \text{for autumn (summer-to-winter season)} \end{cases},$$

with  $\theta_t$  similarly specified. Later on, we shall consider the further simplification that  $\theta_t$  and/or  $\tau_t$  are constants.

The probability model introduced above provides a useful framework for analysing the bacteriological and serological test data. Ideally, a joint analysis may be carried out with both the bacteriological and the serological test data. However, such a joint analysis is not pursued here because there were just a few epizootics in each large square making a fair number of  $p_t$  equal to zero. The lag structure of  $p_t$  entering in  $q_t$  defined in equation (4.1) further complicates the joint analysis. Owing to these difficulties, we first estimate  $p_t$  based on the bacteriological test data. Since  $B_t$  has the binomial distributions  $B(G_t, p_t)$ , the maximum-likelihood estimates of  $p_t$  equal  $\hat{p}_t = B_t/G_t$ . We then treated  $p_t$  as if they were the true  $p_t$  and proceeded to analyse the serological test data, for those seasons since autumn of 1975 and with non-zero serological positives. (Because all epizootic events were comfortably bracketed within the periods when positive serological tests were recorded, for all series since 1975 autumn (figure 1), selection bias owing to subsetting the data is negligible.) In other words, we model  $S_t$  given  $G_t$  as  $B(G_t, q_t)$ , where  $q_t$  is given

by equation (4.1) with  $p_t$  there replaced by  $\hat{p}_t$ . This is a generalized nonlinear mixed-effect model, and the models reported in the next section were fitted using proc nlmixed of SAS v. 9.1.3. The estimation was done with the parameters  $\tau$ s parameterized to be positive and  $\theta$ s between 0 and 1; furthermore, we truncate the right side of equation (4.1) if it falls below 0 or above 0.99. For recent surveys of generalized nonlinear mixed-effect models, see, for instance, Pinheiro & Bates (1995) and Davidian & Giltinan (2003).

We now briefly comment on the theoretical framework used in the electronic supplementary material. The derivation of the nonlinear mixed-effect model defined by equation (4.1) assumes that the time course of the immune response of a (random) great gerbil follows a four-state continuous-time Markov chain (Cox & Miller 1968; Bhattacharya & Waymire 1990), the four states being susceptible, infected, recovered and death. Unlike the classical approaches to infectious disease modelling (e.g. the SIR model; see Grenfell & Dobson 1995; Dickmann & Heesterbeek 2000), our approach does not model the interactions between the infected, the recovered and the susceptibles. This omission is, however, deliberate and it simplifies the analysis, which is appropriate as our goal is not to model the epidemiological time course, but to understand the dynamics of the immune response, conditional on the prevalence rates. Moreover, the analysis is conditional on the observed population fluctuations, which bypasses the need for explicitly modelling the birth and death of the great gerbils. Indeed, it does not seem straightforward to extend the SIR model to account for extensive host population fluctuations, as will be needed for the plague analysis below. In addition, the SIR approach requires a joint analysis of both the bacteriological and the serological test data, which may be infeasible owing to the small number of epizootics. We note that maximum-likelihood estimation of the underlying continuous-time Markov chain model (Kalbfleisch & Lawless 1985; Chan & Munoz-Hernandez 2003) cannot be done with the plague-monitoring data because the data were sampled only twice per year and owing to complications due to the birth-death process of the great gerbils.

## 5. STATISTICAL ANALYSIS AND DISCUSSION

We considered four cases, namely, case 1: constant  $\tau$  and seasonal  $\theta$ ; case 2: seasonal  $\tau$  and constant  $\theta$ ; case 3:

Table 2. Parameter estimates for the best model based on table 1.

parameter	estimate	s.e.	p-value
$\theta_F$	0.491	0.196	0.015
$\theta_S$	0.103	0.029	0.001
$\tau$	9.351	4.040	0.023
$\sigma^2$	0.006	0.001	<0.0001
$\theta_F - \theta_S$	0.388	0.187	0.041
$C_{91}$	0.000	<0.001	0.999
$C_{93}$	0.023	0.026	0.379
$C_{105}$	0.039	0.020	0.058
$C_{106}$	0.074	0.023	0.002
$C_{117}$	0.033	0.041	0.426
$C_{118}$	0.095	0.020	<0.0001
$D_1$	0.572	0.114	<0.0001
$D_2$	0.349	0.109	0.002

both  $\tau$  and  $\theta$  are constants and case 4: both  $\tau$  and  $\theta$  are seasonal. All the analyses reported in this section were based on joint statistical modelling of the data from large squares 91, 93, 105, 106, 117 and 118 (see figure 1), starting from 1975; other large squares were not included because they did not have adequate positive serological data. Given the relatively few epizootics, we assume a common  $\theta$  and  $\tau$  model for all the six large squares, but the constant  $C$  is allowed to be square specific. We restrict  $M$  to be 7 or less, as a great gerbil rarely lives beyond 4 years. Exploratory data analysis suggested that there are two outliers in LSQ 93, in the spring of 1991 and the autumn of 1995. These outliers were accounted for by including two dummy variables for the outliers in the model, with their coefficients labelled as  $D_1$  and  $D_2$ . Table 1 reports the  $AIC_c$  of the fitted models for the four cases, with various values of  $M$  in the model defined by equation (4.1). The minimum  $AIC_c$  value is boldfaced, while entries for models whose Hessian matrices are singular or non-negative definite are given in italics. Based on  $AIC_c$ , a seasonal  $\theta$  and constant  $\tau$  model with  $M=4$  is selected as the best model with a common immune response structure, although  $M=3$  or 2 is very competitive. In other words, the best model chosen includes a four-seasonal-lag disease structure where the antibody loss rate differs between the winter-to-summer season and the summer-to-winter season, while the recovery rate is constant between the two seasons. The coefficient estimates and some statistics related to the estimates of the best model are summarized in table 2. The estimates of the main parameters,  $\theta_F$ ,  $\theta_S$ ,  $\tau$  and  $\sigma^2$ , are statistically significant. Note that the model fit is very similar between  $M=3$  (unreported results) and  $M=4$ .

Figure 1 indicates possible spatial dependence in the data. Thus, it is important to check the need to incorporate spatial correlation beyond that induced by the covariates in the model. We computed the contemporaneous correlations (unreported) between residuals from each pair of the six large squares and found that none are significant at 5% level, suggesting that there is no remaining spatial correlation structure.

In figure 1, the open circles represent the fitted values (i.e. the predicted value of each case given the other data

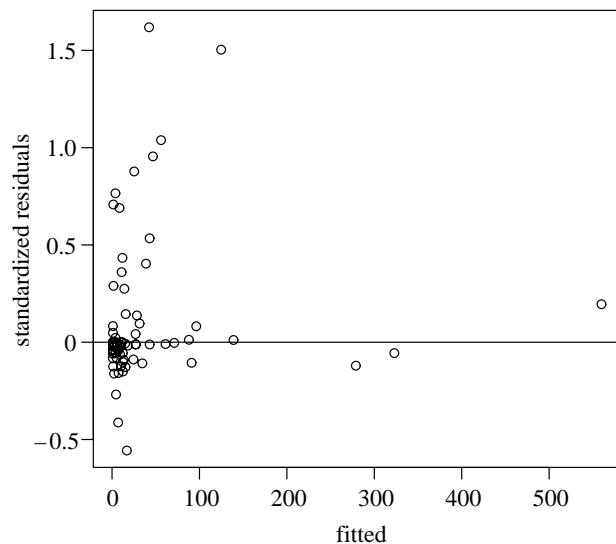


Figure 2. Standardized residual plot of the model reported in table 2.

Table 3. Assessing the robustness of the model reported in table 2 to three potentially influential data cases.

parameter	estimate	s.e.	p-value
$\theta_F$	0.543	0.238	0.025
$\theta_S$	0.089	0.027	0.002
$\tau$	9.548	4.458	0.036
$\sigma^2$	0.005	0.001	<0.0001
$\theta_F - \theta_S$	0.454	0.228	0.050
$C_{91}$	0.006	0.034	0.853
$C_{93}$	0.023	0.024	0.341
$C_{105}$	0.029	0.022	0.176
$C_{106}$	0.070	0.022	0.002
$C_{117}$	0.041	0.037	0.267
$C_{118}$	0.093	0.019	<0.0001
$D_1$	0.576	0.111	<0.0001
$D_2$	0.350	0.105	0.001
$D_3$	-0.151	0.088	0.089
$D_4$	-0.063	0.079	0.425
$D_5$	0.157	0.078	0.048

cases) based on the best-fitted model; these fitted values appear to closely track the counts of serological positives. Figure 2 plots the standardized residuals against the fitted values. The plot suggests that there are no more outliers beyond the two outliers already accounted for in the model. However, three cases (autumn, 1985 in LSQ 91; autumn, 1979 and spring, 1980 in LSQ 105) have fitted values more than 200, so these cases are potentially influential observations. We assessed these potential influential observations by including three extra dummy variables with their coefficients labelled as  $D_3$ ,  $D_4$  and  $D_5$  in the model. We then re-calculated the  $AIC_c$  for various models listed in table 1, but we obtained similar results with the minimum  $AIC_c$  attained for  $M=4$  and the case of constant  $\tau$  and seasonal  $\theta$ . Moreover, the estimates are essentially unchanged (see table 3). Indeed,  $D_3$  and  $D_4$  are not significant and  $D_5$  is marginally significant. A joint test of  $D_3=D_4=D_5=0$  has  $p$ -value equal to 0.064, suggesting that these three cases may not be influential.

Hence, we conclude that the fitted model reported in table 2 is robust to the three influential cases.

The estimate of the parameter  $\theta$  is significantly higher in autumn than in spring, suggesting that antibodies were boosted in recovered great gerbils owing to repeated contacts with plague microbes during the summer-to-winter season. The higher the level of antibodies, the longer it may take before the antibodies disappear from the blood. During winter, on the other hand, plague activity is reduced. Fleas are less active and gerbils move less around, reducing the potential for plague transmission between colonies. When there is no repeated plague contact, the level of antibodies in blood will steadily decrease. It thus makes sense that antibodies are lost faster during winter than during summer and autumn.

A seasonal difference in mortality rate may, however, contribute to the seasonal variation in  $\theta$ . Using seasonal age-structure data of great gerbils (i.e. counts of adults and juveniles caught in spring and autumn) from our study area during 1978–1989 and the corresponding rodent density estimates, we calculated average seasonal survival rates through a linear mixed-effects model fit (S. Park, K.-S. Chan & H. Viljugren 2006, unpublished results). The overall autumn-to-spring mean survival rate (0.54, s.e. = 0.10) was almost equal to the spring-to-autumn mean survival rate of adults (0.51, s.e. = 0.12). The spring-to-autumn survival rate of juveniles was confounded with the net birth rate (the estimated spring-to-autumn net birth rate of juveniles: 2.2, s.e. = 0.47). However, juveniles are not likely to have a higher summer survival rate than adults. Hence, we may conclude that the great gerbils were unlikely to have a lower mean survival rate in the autumn-to-spring season than in the spring-to-autumn season during 1978–1989. Thus, we may interpret the higher  $\theta$  in autumn as indicative of a higher rate of loss of antibodies in the winter-to-summer season than in the summer-to-winter season.

Additionally, rather than reflecting a seasonal immunity response within great gerbil individuals, the seasonal  $\theta$  might reflect average seasonal reproductive and survival patterns within the great gerbil population. From autumn to spring there is a generation shift in the great gerbil population. Many of the adults from the previous season will not survive the winter. There is also evidence that mortality is higher for seropositive great gerbils than for those testing seronegative (Begon *et al.* 2006). Furthermore, in late spring/early summer, there is an influx of juveniles (susceptibles), which have not yet been in contact with the plague microbe. In the PreBalkhash area, reproduction occurs mainly in April–June. As winters progress and nutritional or climatic stressors act on the great gerbils, we may expect decline in the body mass and the level of immunocompetence. In addition, immunity is generally lower during the breeding season than the non-breeding season (Nelson *et al.* 2002), suggesting that adult gerbils may lose resistance (innate and acquired) and become susceptible to plague at higher rates in spring than in autumn. Subsequently, this may facilitate transmission of plague to juveniles in early summer. The high proportion of susceptible juveniles

during early summer may contribute to the higher  $\theta$  in autumn. The maximum level of antibodies is produced with some delay (approximately a month) after the infection, so that juveniles infected in early summer may not be detected as seropositive before the autumn sampling. A high proportion of susceptible individuals is likely to facilitate plague transmission and increase the chance for new and repeated plague infections in the population.

The apparent immigration effects are insignificant for large squares 91, 93, 105 and 117, but they are significant for large squares 106 and 118. Indeed, migration of great gerbils from other large squares into the latter two squares may account for the surge in the serologically positive cases towards the end of the sampling period for these three squares; see figure 1. Moreover, while the bacteriological time-series always led the serological series in large squares 91, 93, 105 and 117, this was not the case for large squares 106 and 118. Instead, the serological series sometimes led the bacteriological series; immigration is a possible cause of this phenomenon. However, long-distance migration among great gerbils in the PreBalkhash area is not thought to be common (S. Pole & N. Klassovskiy 2006, personal communication). The low sensitivity of the bacteriological test (i.e. the low probability of obtaining samples with live microbes from plague-infected individuals) is a more reasonable explanation. In addition, only a small fraction of the study area was covered in any sampling epoch, so some cases of plague might have gone undetected.

## 6. CONCLUSION

According to our analysis, great gerbils having recovered from a plague infection were more likely to keep their antibodies to plague and survive throughout the summer-to-winter season than throughout the winter-to-summer season. Provided that the seasonal mortality rates were similar, this suggests that the immune response of the sampled great gerbils was seasonal. As remarked earlier in the introduction section, the seasonal dynamics of increased and decreased specific immunity may facilitate the persistence of plague in the host population. Thus, understanding the seasonal structure of the epidemics is greatly needed. In this context, our paper provides some important insights.

Earlier studies of viral, bacterial and parasitic infections suggest that seasonality in the immune response may be mediated by changes in food availability, temperature, photoperiod and social behavioural changes (see Klein *et al.* 2002; Nelson *et al.* 2002; Rogovin *et al.* 2003). The present study suggests that decreased plague activity during winter, a generation shift in the population from autumn to spring and a possibly lower survival rate during winter than during summer are other factors likely to contribute to the lower probability of keeping plague antibodies throughout the winter-to-autumn season than throughout the summer-to-winter season. Indeed, the importance of the heightened plague activities in the summer-to-winter season is consistent with the finding by Stenseth *et al.* (2006) that the prevalence during both spring and

autumn is best predicted by, respectively, the great gerbil population size during the autumn 1.5 or 2 years earlier. The higher probability of keeping antibodies throughout the summer-to-winter season was furthermore likely to be a result of antibodies being boosted in great gerbils owing to repeated contacts with plague microbes. A possible high proportion of susceptible individuals in spring and early summer was likely to increase plague activity, so that the chances for new and secondary plague infections in the population were increased.

A major contribution of this paper consists of the development of an epidemiologically based time-series analysis of the dynamical structure in antibody loss using aggregate time-series disease data. An interesting statistical problem is how to extend our approach to include covariates, so that we may identify the important factors that may explain the observed seasonal structure. Such an extension will profoundly help us improve our understanding of the epidemiological dynamics of wildlife-borne diseases like plague.

This work was partially supported by the European Union as projects ISTC K-159 and STEPICA (INCO-COPERNICUS, ICA 2-CT2000-10048), as well as the authors' institutes. The work of S.P. was supported by the Post-doctoral Fellowship Program of Korea Science and Engineering Foundation (KOSEF). H.V. was supported by the Research Council of Norway (RCN). We thank M. Begon, H. Leirs and S. Davis for comments on the manuscript and D. Ehrlich for valuable work as an interpreter and translator. We are thankful to two anonymous referees whose helpful comments on an earlier draft led to this streamlined version.

## REFERENCES

Achtman, M. *et al.* 2004 Microevolution and history of the plague bacillus, *Yersinia pestis*. *Proc. Natl Acad. Sci. USA* **101**, 17 837–17 842. (doi:10.1073/pnas.0408026101)

Begon, M., Klassovsky, N., Ageyev, V., Suleimenov, B. & Bennett, M. 2006 Epizootiologic parameters for plague in Kazakhstan. *Emerg. Infect. Dis.* **12**, 268–273.

Bhattacharya, R. N. & Waymire, E. C. 1990 *Stochastic processes with applications*. New York, NY: Wiley.

Chan, K. S. & Munoz-Hernandez, B. 2003 A generalized linear model for repeated ordered categorical response data. *Stat. Sin.* **13**, 207–226.

Cox, D. & Miller, H. 1968 *The theory of stochastic processes*. London, UK: Chapman and Hall.

Davidian, M. & Giltinan, D. M. 2003 Nonlinear models for repeated measurement data: an overview and update. *J. Agr. Biol. Environ. Stat.* **8**, 387–419. (doi:10.1198/1085711032697)

Davis, S. *et al.* 2004 Predictive thresholds for plague in Kazakhstan. *Science* **304**, 736–738. (doi:10.1126/science.1095854)

Dickmann, O. & Heesterbeek, J. A. P. 2000 *Mathematical epidemiology of infectious disease: model building, analysis and interpretation*. Chichester, UK: Wiley.

Frigessi, A., Marshall, C., Holden, M., Viljugrein, H., Stenseth, N. Chr., Holden, L., Ageyev, V. & Klassovskiy, N. L. 2005 Bayesian population dynamics of interacting species: great gerbils and fleas in Kazakhstan. *Biometrics* **61**, 231–239. (doi:10.1111/j.0006-341X.2005.030536.x)

Grenfell, B. T. & Dobson, A. P. 1995 *Ecology of infectious diseases in natural populations*. New York, NY: Cambridge University Press.

Kalbfleisch, J. & Lawless, J. 1985 The analysis of panel data under a Markov assumption. *J. Am. Stat. Assoc.* **80**, 863–871. (doi:10.2307/2288545)

Kanatov, Yu. V. 1974 Systems of serological tests using sensitized erythrocytes for detecting antigens and antibodies. Thesis of doctoral dissertation. Saratov. 36 [In Russian].

Klein, S. L., Bird, B. H., Nelson, R. J. & Glass, G. E. 2002 Environmental and physiological factors associated with seoul virus infection among urban populations of Norway rats. *J. Mammal.* **83**, 478–488. (doi:10.1644/1545-1542(2002)083<0478:EAPFAW>2.0.CO;2)

Korneyev, G. A. & Karpov, A. A. 1977 Dynamics of the sex and age structure of the large gerbil population. In *Ecology and medical value of the gerbils of the USSR fauna. Materials of All-Union meeting*. M, pp. 144–146.

MacIntyre, S., Knight, S. D. & Fooks, L. J. 2004 Structure assembly and applications of the polymeric F1 antigen of *Yersinia pestis*. In *Yersinia molecular and cellular biology* (ed. E. Carniel & B. J. Hinnebusch), **110**, 363–407. Norfolk, UK: Horizon Bioscience.

Naumov, N. P. & Lobachev, V. S. 1975 Ecology of desert rodents of the U.S.S.R. (Jerboas and Gerbils). Great Gerbil. In *Rodents in desert environments* (ed. I. Prakash & P. K. Ghosh), pp. 549–598. The Hague, The Netherlands: Dr. W. Junk b.v. Publishers.

Nelson, R. J., Demas, G. E., Klein, S. L. & Kriegsfeld, L. J. 2002 *Seasonal patterns of stress, immune function, and disease*. New York, NY: Cambridge University Press.

Perry, R. D. & Fetherston, J. D. 1997 *Yersinia pestis*—etiological agent of plague. *Clin. Microbiol. Rev.* **10**, 35–66.

Pinheiro, J. C. & Bates, D. M. 1995 *Mixed effects models in S and S-Plus*. New York, NY: Springer.

Pollizer, R. 1966 *Plague and plague control in the Soviet Union*. New York, NY: Fordham University, Bronx.

Rogovin, K., Randall, J. A., Kolosova, I. & Moshkin, M. 2003 Social correlates of stress in adult males of the great gerbil, *Rhombomys opimus*, in years of high and low population densities. *Horm. Behav.* **43**, 132–139. (doi:10.1016/S0018-506X(02)00028-4)

Rothschild, V. E. 1978 *Spatial structure of plague natural focus and methods of its study*, p. 192. Moscow, Russia: Publishing House of Moscow University. [In Russian.]

Samia, N. I., Chan, K. S. & Stenseth, N. C. In press. A generalized threshold mixed model for analyzing non-normal nonlinear time series: plague in Kazakhstan as an illustration. *Biometrika*.

Stenseth, N. C. *et al.* 2006 Plague dynamics are driven by climate variation. *Proc. Natl Acad. Sci. USA* **103**, 13 110–13 115. (doi:10.1073/pnas.0602447103)

Suleimenov, B. M., Isin, Zh. M., Atshabar, B. B., Klassovskiy, N. L., Kogay, O. V., Kopbayev, E. Sh. & Novikov, G. S. 2001 Immunophysiological structure of acute plague epizootics. *Quarantine and Zoonotic Infections in Kazakhstan*. Almaty, **4**, 261–266. [In Russian.]