THE ROYAL SOCIETY

Received 23 December 2003
Accepted 9 February 2004
Published online 28 April 2004

Bayesian analysis of experimental epidemics of foot-and-mouth disease

George Streftaris* **and Gavin J. Gibson**

Actuarial Mathematics and Statistics, School of Mathematical and Computer Sciences, Heriot-Watt University, Riccarton, Edinburgh EH14 4AS, UK

We investigate the transmission dynamics of a certain type of foot-and-mouth disease (FMD) virus under experimental conditions. Previous analyses of experimental data from FMD outbreaks in non-homogeneously mixing populations of sheep have suggested a decline in viraemic level through serial passage of the virus, but these do not take into account possible variation in the length of the chain of viral transmission for each animal, which is implicit in the non-observed transmission process. We consider a susceptible–exposed–infectious–removed non-Markovian compartmental model for partially observed epidemic processes, and we employ powerful methodology (Markov chain Monte Carlo) for statistical inference, to address epidemiological issues under a Bayesian framework that accounts for all available information and associated uncertainty in a coherent approach. The analysis allows us to investigate the posterior distribution of the hidden transmission history of the epidemic, and thus to determine the effect of the length of the infection chain on the recorded viraemic levels, based on the posterior distribution of a *p*-value. Parameter estimates of the epidemiological characteristics of the disease are also obtained. The results reveal a possible decline in viraemia in one of the two experimental outbreaks. Our model also suggests that individual infectivity is related to the level of viraemia.

Keywords: Bayesian inference; foot-and-mouth disease; Markov chain Monte Carlo; stochastic epidemic modelling; transmission chain

1. INTRODUCTION

The study of the infection dynamics of animal pathogens is not straightforward, especially in the case of asymptomatic diseases, and many issues regarding the transmission dynamics and characteristics of foot-and-mouth disease (FMD) remain open. One such issue was raised following the FMD outbreak in Greece during 1994. The relatively fast diminishing of the epidemic suggested that the viral strain responsible for the disease (type O) may not be able to maintain itself through time, as animals infected in the later stages of the epidemic may not be able to produce sufficient amounts of the infectious agent for new infections to occur (Mackay *et al.* 1995), owing to a suspected lower blood viral load under a hypothesis of viraemia being related to exposure dose.

To investigate the persistence of the disease through time, we analyse data on the transmission of the type O viral strain to sheep. In two conducted experiments, animals were randomly allocated to four groups (G1 to G4), and the first-group sheep were inoculated with a constant FMD viral dose. The virus was then passed to animals of the remaining groups, so that, throughout the experiment, each group spent 24 h mixing with a given group of 'donor' animals, followed by 24 h in the presence of a given group of 'recipient' animals. It has been reported that this specific FMD viral isolate may not be able to maintain itself through time, as mean peak viraemic levels are significantly different in sheep of different passage groups (Hughes *et al.* 2002*b*). This is based on the assumption that the 'donating' group is the only possible source of infection for all members of a given group, and implies that the length of the transmission chain, and therefore the relative temporal position in the epidemic process, is identical for all animals in a given group and correlates well with group membership.

This assumption may not be valid even under the specific experimental conditions, and in this paper we allow the source of infection of members of a given group to vary, so that the virus may be acquired from any animal contacting a susceptible at the moment of transmission, given the restrictions imposed by the non-homogeneous daily mixing patterns in the experimental design. The aim is to obtain a more realistic representation of a possible decline in the peak viraemic levels as a result of an increasing number of contacts in the transmission chain leading to an individual infection. Futhermore, to attribute the eradication of the disease to a diminishing viral load, an association between viraemia and infectiousness must be established. We investigate the presence of such a relationship through suitable modelling of the infection-challenge process. Relevant studies to assess whether the transmission process of the FMD virus to susceptible sheep is dose related (e.g. Hughes *et al.* 2002*a*) have not been conclusive in establishing a clear relationship between dose level and infectivity. Here, we also address the issue of the characterization of the disease through the quantification of appropriate epidemiological parameters such as the transmission coefficient and the duration of the latent (incubation) period, which, if the times of the events related to the epidemic outbreak are unobserved or censored, cannot be tackled with the use of standard statistical techniques.

To address these issues we fit a stochastic compartmental epidemic model (Bailey 1975; Gibson & Renshaw

* Author for correspondence (g.streftaris@ma.hw.ac.uk).

1998, 2001) to the data. We follow a Bayesian approach to incorporate the uncertainty associated with the unobserved times of exposure to the virus and the exact durations of the viraemic periods, and also to accommodate available knowledge and past experience regarding the disease characteristics. Our analysis is based on the posterior distribution of the model parameters, obtained with the use of Markov chain Monte Carlo (MCMC) methodology, which is being increasingly employed in epidemic modelling (e.g. Gibson 1997; O'Neill & Roberts 1999; Britton & O'Neill 2002). Höhle *et al.* (2003) also use MCMC to analyse data from disease transmission experiments in a Bayesian setting. Here, in addition to estimating the model parameters, we employ the posterior distribution of the unobserved exposure times to determine the lengths of the infection chains of infected animals, and therefore investigate their effect on detected viraemia by using analysis of variance (ANOVA) to test the hypothesis of no change in viraemic level as the chain length increases. This allows the derivation of a posterior distribution of p-values (cf. Meng 1994).

2. DATA

Data were collected during experiments carried out at the Institute of Animal Health, Pirbright, UK, in collaboration with the Centre for Tropical Veterinary Medicine, University of Edinburgh. Two experiments were performed under identical conditions, and each involved 32 susceptible sheep randomly divided into four groups of eight animals. Viraemic diagnosis was based on daily blood samples. Full details of the design and conditions of the experiments and the collected data can be found in Hughes *et al.* (2002*b*). The data used in this paper consist of individual records of the days of onset and cessation of viraemia and the peak viraemic level. Four sheep in experiment 1 and nine in experiment 2 did not exhibit detectable viraemia, but were treated as susceptible in our model, and their contribution to the transmission dynamics of the disease was taken into account in the analysis. When differences in viraemic levels among animals were tested, only naturally infected sheep (G2–G4) were included in the analysis.

3. MODEL AND METHODOLOGY

We represent the spread of the epidemic through a susceptible–exposed–infectious–removed model (e.g. Bailey 1975; Becker 1989), where each animal can be classified as susceptible to the disease (S), exposed to the infecting agent (E), infectious (I) or removed (recovered, R). The sojourn time in state E, i.e. the latent period, plays a significant role in the control of FMD, but for purposes of inference is often considered to be of fixed length in epidemic modelling (e.g. Becker 1983; O'Neill & Becker 2001). Also, the exponential distribution has been widely used to describe sojourn times in various compartments in the study of epidemics, partly owing to its mathematically convenient Markov property, as for example in the general susceptible–infectious–removed epidemic model (Bailey 1975). In this paper, following the work in Streftaris & Gibson (2004), we employ the two-parameter Weibull(ν, λ) distribution with probability density function $f(x) = \nu\lambda x^{\nu-1}\exp(-\lambda x^{\nu})$, with x, ν, $\lambda > 0$, whose properties offer a more flexible framework for realistic epidemic modelling. Non-Markovian models have been used in the literature to describe sojourn times in the epidemic modelling of diseases such as acquired immune deficiency syndrome (AIDS) (e.g. De Angelis *et al.* 1998) and in transmissible spongiform encephalopathy related epidemics (e.g. Ferguson *et al.* 1997; Ghani *et al.* 1998). The gamma distribution has also been considered in a similar context (O'Neill & Becker 2001).

(a) *Likelihood*

We use n to denote the number of viraemic animals in each population. The observation period of the epidemic is represented in our model by the time interval $[0, T]$, defining its start as the inoculation time and its end as the time of the last recorded event (last recovery). We assume that no G1-sheep infection was caused by natural transmission of the virus, and therefore G1 animals are not regarded as recipients in the model. The design of the experiments mimics a non-homogeneous population mixing pattern, according to which the groups mix in pairs. Thus, (G1, G2) and (G3, G4) mix on odd days of the experiment, and (G2, G3) mix on even days while G1 and G4 are kept separately. Transitions from compartment S to compartment E in the infinitesimal time increment $[t, t + \mathrm{d}t)$ occur according to a probability of the form

$$\Pr\{S_k(t + \mathrm{d}t) = S_k(t) - 1\} = \beta S_k(t) \sum_{l=1}^{n} \{v_l^{\alpha} i_l(k, t)\}\,\mathrm{d}t, \qquad (3.1)$$

with β denoting the rate of infection per possible susceptible–infectious contact weighed by the associated infectivity, and $S_k(t)$ giving the number of susceptible animals in group k at time t. We consider the peak viraemic level of each infectious sheep as a potential factor affecting the infective challenge exerted on each susceptible animal. The possible influence is modelled as the sum of a power function of the individual viraemic levels, allowing the power level, denoted by α, to be estimated as a model parameter. This allows for the special case of a mass-action infective-challenge function when $\alpha = 0$. For comparisons with mass-action models and also to facilitate the MCMC algorithm presented in § 3b, we scale the original viraemic measurements to a mean level of one unit, and use v_l, $l = 1, \ldots, n$, to denote the normalized measurements. The function $i_l(k, t)$ is an indicator factor such that for $l = 1, \ldots, n$, $i_l(k, t) = 1$ if at time t animal l is infectious and mixing only with group k, or zero otherwise. We assume independent Weibull distributions for the durations of the latent and infectious periods. It is possible that sheep that acquired the disease by viral inoculation will have a shorter incubation period than naturally infected sheep (Sellers *et al.* 1977), and therefore we consider a Weibull(γ_1, δ_1) distribution for G1 sheep and a Weibull(γ_2, δ_2) distribution for G2–G4 animals. A Weibull(ν, λ) distribution is used to represent the length of the infectious period, and we assume that infectious and viraemic periods coincide in time. The likelihood of the complete data (assuming perfect observation of the epidemic) can be written as a function of the model parameters as

$$L(\alpha, \beta, \gamma_1, \delta_1, \gamma_2, \delta_2, \nu, \lambda; \boldsymbol{e}, \boldsymbol{s}, \boldsymbol{r}) = \prod_{j \in \mathcal{E}} \left[\beta \sum_{l=1}^{n} \{v_l^{\alpha} i_l(G_j, e_j)\} \right] \times \exp\left\{ - \int_0^T \beta C(t) \mathrm{d}t \right\} \times \prod_{j \in I_1} f_1(s_j - e_j; \gamma_1, \delta_1) \times \prod_{j \in I_{2,3,4}} f_2(s_j - e_j; \gamma_2, \delta_2) \times \prod_{j \in R} f_3(r_j - s_j; \nu, \lambda), \quad (3.2)$$

where e_j, s_j and r_j denote, respectively, the times of exposure, the start of the infectious period and recovery of animal j, and $\boldsymbol{e}$, $\boldsymbol{s}$ and $\boldsymbol{r}$ are the corresponding vectors; G_j is the group to which animal j belongs; $f_1(\cdot)$ and $f_2(\cdot)$ denote the Weibull densities for the latent G1 and G2–G4 periods, and $f_3(\cdot)$ is the Weibull density of the infectious period. Also, $\mathcal{E}$, I_1, $I_{2,3,4}$ and R denote the sets of exposed (G2–G4), infectious (G1, G2–G4) and recovered animals, respectively, at the end of the experiment, while $\beta C(t)$ represents the total infective force on the susceptible population at time t, given the mixing pattern and the infectious state of the population at that time. $C(t)$ is defined as

$$C(t) = \sum_{k=1}^{3} \left[\{S_k(t) + S_{k+1}(t)\} \times \sum_{l=1}^{n} \{v_l^{\alpha} i_l(k, k+1, t)\} \right] + S_4(t) \times \sum_{l=1}^{n} \{v_l^{\alpha} i_l(4, t)\}, \quad (3.3)$$

and uses information on all susceptible sheep, including those that remained uninfected through the course of the epidemic. Here $i_l(k, m, t)$ is defined as in equation (3.1) with the generalization that $i_l(k, m, t) = 1$ if at time t animal l is infectious and mixing with both groups k and m, or zero otherwise, $l = 1, \ldots, n$.

Perfect knowledge of all event times would allow direct use of the likelihood equation (3.2) to obtain estimates for the parameters of interest, for example by the method of maximum likelihood. However, the available information is only partial, as the exposure times for naturally infected animals, e_l where $l \in \mathcal{E}$, are not known. Furthermore, the recorded times of infectiousness onset (s_l) and recovery (r_l) result from sampling carried out every 24 h, and are therefore not exact. Standard statistical methodology would fail to account for the missing information and accommodate the associated uncertainty. In §§ 3b and 3c we present a coherent framework for tackling these issues.

(b) *Bayesian inference*

We follow a Bayesian approach, under which the unobserved events in the transmission process of the disease are represented as nuisance parameters. The joint posterior density of these parameters and the model parameters (given the observed values) is investigated and inferences on model parameters are extracted from the respective marginal densities. We first specify appropriate prior distributions for the model parameters α, β, γ_1, δ_1, γ_2, δ_2, ν and λ. For the parameters of the Weibull distributions of latent and infectious periods we assume gamma priors with parameters determined such that the resulting distributions loosely reflect existing knowledge about the epidemic characteristics of FMD in sheep under conditions similar to those in the experiments. More specifically, based on information from previous studies, which suggest that the incubation period can be less than 1 day for inoculated sheep and less than 4 days for naturally infected sheep and also that the viraemic period usually lasts *ca*. 3 days (Sellers *et al.* 1977; Kitching & Mackay 1994; Cox *et al.* 1999), we determine the parameters of the gamma prior distributions in such a way that their moments provide Weibull distributions that are rather weakly concentrated around these values. The transmission parameter β is assigned a non-informative gamma distribution, while for the power parameter α we assume a vague exponential prior distribution. All prior distributions are assumed to be independent of each other. The prior distribution of the parameters is then updated using the information provided by the data through the likelihood function, to yield the posterior distribution. From Bayes' theorem the joint posterior density of the parameters is given as

$$p(\alpha, \beta, \gamma_1, \delta_1, \gamma_2, \delta_2, \nu, \lambda | \boldsymbol{e}, \boldsymbol{s}, \boldsymbol{r}) \propto L(\alpha, \beta, \gamma_1, \delta_1, \gamma_2, \delta_2, \nu, \lambda; \boldsymbol{e}, \boldsymbol{s}, \boldsymbol{r}) \times \pi(\alpha, \beta, \gamma_1, \delta_1, \gamma_2, \delta_2, \nu, \lambda), \quad (3.4)$$

where $L(\cdot)$ is the likelihood function in equation (3.2) and $\pi(\cdot)$ denotes the joint prior distribution of the parameters (where the dependence on hyperparameters has been suppressed). The posterior density is given in an analytically intractable form, and therefore inference will rely on computationally intensive estimation methods. We use MCMC methodology (e.g. Gelfand & Smith 1990; Tierney 1994), which offers tools for stochastic integration in problems of increased complexity and dimension. A single-component Metropolis–Hastings algorithm is used, in a manner similar to that described in Streftaris & Gibson (2004). Details of the algorithm implementation are given in electronic Appendix A.

(c) *Posterior investigation of the hidden infection process*

To investigate the effect of the length of the infection chain on the detected level of viraemia we consider posterior properties of the network of infectious contacts. These can be accessed via the posterior distribution of the unobserved times of exposure to the disease, conditional on the recorded events. We determine the length of the infection chain of each viraemic sheep by first reconstructing the whole infection transmission history of the epidemic. The lack of detailed observations of the infection process implies that this will be subject to uncertainty within a stochastic framework. The MCMC algorithm allows us to obtain stochastic realizations of the unobserved infection trees conditional on the unknown exposure times, by selecting an infecting animal for each exposure according to a probability distribution weighed by the infectiousness of available infectious individuals. Each sampled tree specifies a partition of the animals according to the length of the infection chain, and we assess the possible effect on the exhibited viraemia by using ANOVA to test a null hypothesis of no differences in viraemic levels with increasing length of the transmission chain. Each sample of individual infection chain is used for the evaluation of the appropriate test statistic based on data from the posterior distribution of exposure times, giving an associated p-value (cf. Meng 1994) for the null

Table 1. Posterior mean, standard deviation, median and equal-tailed 95% credible interval for disease-progression characteristics in the two experiments.

	transmission coefficient, β	mean latent period G1	s.d. latent period G1	mean latent period G2–G4	s.d. latent period G2–G4	α
experiment 1						
mean	0.024	0.938	0.437	1.599	0.672	1.553
s.d.	0.011	0.265	0.270	0.440	0.364	0.762
median	0.023	0.915	0.370	1.507	0.581	1.477
95% interval	(0.005, 0.048)	(0.613, 1.360)	(0.185, 1.027)	(0.943, 2.695)	(0.255, 1.583)	(0.208, 3.259)
experiment 2						
mean	0.020	0.937	0.433	1.937	1.356	0.808
s.d.	0.008	0.231	0.258	0.732	1.012	0.493
median	0.020	0.910	0.378	1.882	1.137	0.732
95% interval	(0.004, 0.036)	(0.576, 1.469)	(0.183, 0.965)	(0.836, 3.403)	(0.313, 3.772)	(0.083, 1.998)

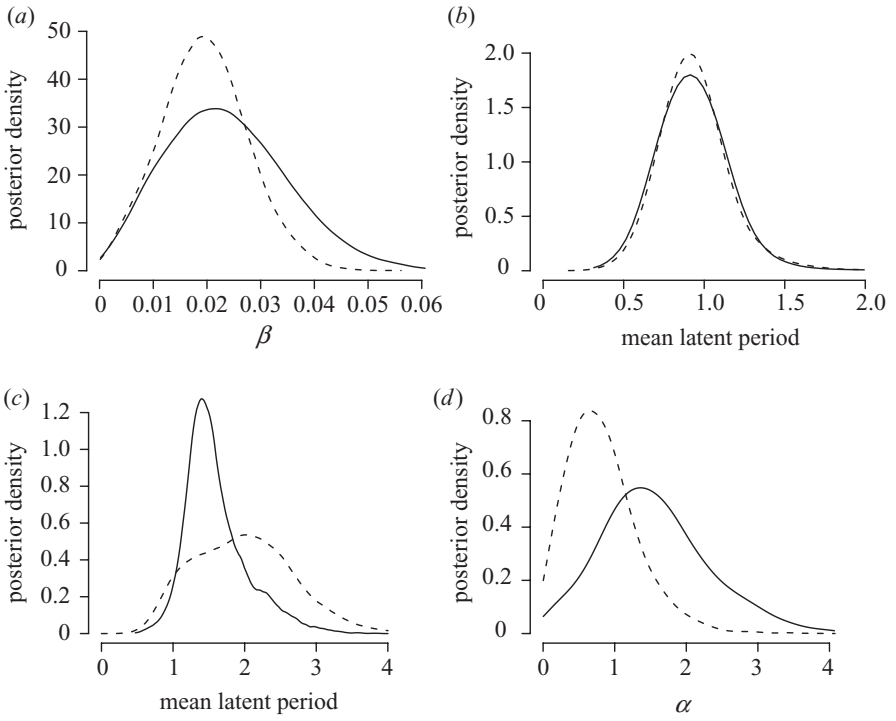

Figure 1. Posterior densities of the characteristics of the transmission of FMD virus in sheep under experimental conditions: (*a*) β; (*b*) mean latent period for G1; (*c*) mean latent period for G2–G4; and (*d*) α. The solid and dashed lines correspond to experiment 1 and experiment 2, respectively.

hypothesis. The whole posterior distribution of these *p*-values can then be obtained based on data from the MCMC output. The suggested methodology has been validated with the use of simulated data. Several epidemic outbreaks have been generated under the assumed model and experimental design, with epidemiological determinants corresponding to various scenarios. The analysis identified an effect of length of infection chain on viraemia, when viraemic levels were systematically generated to follow a decreasing trend (similar to that estimated from the experimental epidemics). Conversely, the posterior distribution of the *p*-values did not suggest any such trend when the viraemic levels were simulated independently of infection generation. Details of the simulated epidemics and relevant results are given in electronic Appendix A.

4. RESULTS

We first present estimates of the parameters quantifying the spread of FMD in the two studied experiments. Characteristics of interest are the transmission (or contact) parameter β, the durations of the latent (incubation) and infectious periods of the disease and the parameter α used to assess a possible relationship between blood viral load and the infectiousness of individual sheep. Estimates of the model parameters are given in table 1, with the corresponding marginal posterior densities shown in figure 1. The transmission parameter β represents the rate of viral transmission in relation to the number of possible infectious contacts and the viraemic level of infective animals. As the viraemia measurements were standardized to a mean level

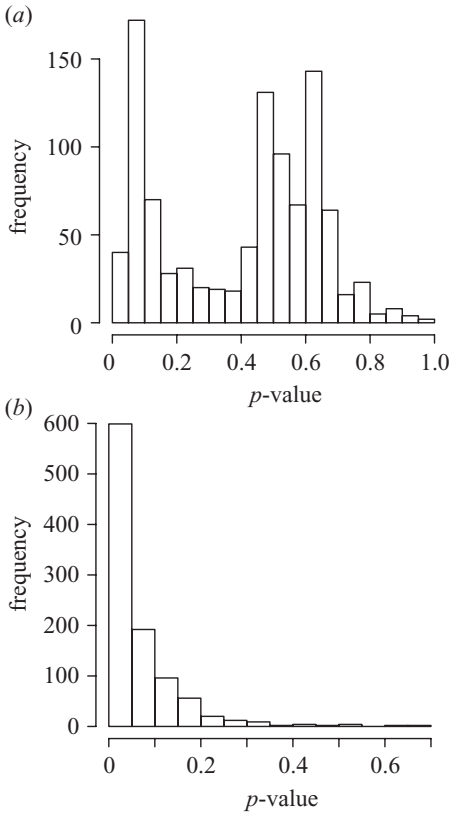

Figure 2. Histograms of posterior distributions of p-values associated with the hypothesis of no differences among average peak viraemic levels of sheep classified according to the length of the infection chain. (*a*) Experiment 1 and (*b*) experiment 2. Summary statistics for (*a*) experiment 1: first quartile = 0.12, median = 0.48, mean = 0.40, third quartile = 0.62; and (*b*) experiment 2: first quartile = 0.01, median = 0.04, mean = 0.06, third quartile = 0.08.

of one unit in the likelihood computation, our β estimates are comparable to the rate of infection per possible susceptible–infectious contact under a mass-action infective-challenge function of the form $\beta S(t)I(t)$. For sheep infected by inoculation, both experiments reveal an average latent period that concurs with empirical observations reported in the literature: that the incubation period for sheep inoculated with the FMD virus can be less than 24 h (Sellers *et al.* 1977). The posterior densities of all model parameters are consistent with the assumption of the same underlying epidemic process in the two experimentals. Estimation of the mean latent period of sheep in G2–G4 in experiment 1 is more precise than in experiment 2, as shown by the dispersion of its posterior distribution in figure 1*c*, which could partly be attributed to the greater estimation uncertainty present in the second experiment as a result of fewer G2–G4 animals being detected as viraemic and thus entering the process of transition between different compartments. This is not the case in the estimation of the latent period of G1 sheep, for which uncertainty is also reduced as a result of the exposure times being known.

(a) *Relationship between viraemia and infectiousness*

The posterior distribution of the α parameter is shown in figure 1*d*. As an almost flat non-informative exp(0.001) distribution was used to express our *a priori* knowledge about α, the location of the posterior distribution indicates that the information in the data supports non-zero values of the parameter. Experimentation with simulated epidemics verified that, when no relationship between viraemia and infectivity is assumed in the data-generation process ($\alpha = 0$), the posterior probability density of α is significantly shifted towards the origin, and its location can be clearly distinguished from the case where values of α close to 1 are used in the simulation. In addition, we employed a discrete prior model for α, with the parameter taking values over a non-negative grid, to assess the ability of the purely continuous exponential model to identify a possible concentration of point mass at $\alpha = 0$. Histograms of the posterior distribution of α (shown in electronic Appendix A) revealed the same characteristics with both models. Therefore our analysis shows that individual peak viraemic levels affect the infective challenge exerted on susceptible animals in both experiments, suggesting a possible association between blood viral load and the infectivity of sheep infected with FMD.

(b) *Decline of viraemia through the infection chain*

Work by Hughes *et al.* (2002*b*) on the same data showed that viraemia declines along passage groups. However, with the course of the infection process being largely unobserved, a more powerful analysis is required to suggest the same effect over the infection chain. Assuming an association between viraemia and infectivity, a decrease in viraemic levels can be interpreted as the disease not being able to maintain itself through a chain of viral transmissions. By determining the length of the infection chain of each viraemic sheep in the way described in § 3c we obtain a more realistic representation of the relative temporal position of each sheep in the infection process. We then perform ANOVA tests and obtain p-values for the null hypothesis of no differences in the average peak viraemic level among sheep classified according to the length of their infection chain. The posterior distributions of these p-values, based on histograms obtained from 1000 samples for each experiment, are shown in figure 2. Our analysis reveals possible differences in the viraemic levels of animals with different lengths of infection chain in experiment 2, as the histogram in figure 2*b* favours small p-values. Graphical inspection of the viraemic levels of sheep grouped according to infection-chain length verified that small p-values corresponded to differences caused by decreasing patterns of viraemia. However, the shape of the distribution for experiment 1 implies that, on this occasion, the available information on the epidemic gives rise to transmission patterns that are consistent with both changing and unaffected viraemic levels, as reflected by the range of small and large p-values in the histogram.

5. DISCUSSION

Owing to the asymptomatic nature of FMD, control measures have to rely on estimation of the transmission dynamics, especially in cases of insufficient early diagnosis (e.g. caused by large outbreaks, time restrictions or the non-availability of suitable tests). In this paper, we have developed estimation methodology that can be applied when complex modelling of the infection process is used to represent an epidemic outbreak. Such explicit

modelling is required to relax assumptions that are either unrealistic or difficult to substantiate. The Bayesian analysis of the presented model tackles the complex interactions between rate of infection, quantity of virus replicated, durations of latent and infectious periods, and infectivity of the host in a coherent way, so that the conclusions are based on the entire available information on the epidemic outbreak. The power of our approach to distinguish between different infection dynamics of various epidemic processes was validated through simulations of epidemic data from appropriate outbreak scenarios.

Our estimates of the mean latent period for in-contact infections (G2–G4) suggest a shorter incubation time for naturally infected sheep than that reported in the literature (between 3 and 8 days; e.g. Kitching & Mackay 1994). Possible explanations include the dependence of the incubation period on the infectious dose received by susceptible sheep and also the specific experimental conditions leading to a highly intensive process. The assumption that infectious and viraemic periods coincide in time is common in studies of the course of FMD infection (e.g. Gibson & Donaldson 1986; Cox *et al.* 1999) and is employed here for reasons of model parsimony and to avoid problems of parameter identifiability. The use of a more complicated model to account for a possible time delay between the onset (or end) of infectiousness and viraemia would be problematic for inference purposes under the restricted information in the available data.

We have modelled the infective challenge exerted by each infectious sheep as a function of its blood viral load (standardized to a mean level of one), assuming that the recorded peak viraemic levels represent accurately the overall viral load. Therefore, the non-zero estimates of the power parameter α should be regarded as an indication of a departure from the hypothesis of no relationship between viraemia and infectiousness, rather than as an effort to quantify such an association.

Our results regarding a suspected decline in levels of blood viral load in sheep infected by the FMD virus involve ANOVA techniques employing a test statistic evaluated using realizations of the unobserved exposure times based on their posterior distribution. The resulting distributions of p-values can be viewed as an indication of how extreme the 'observed' random representations of the epidemic tree would be under the null hypothesis of no change in viraemia over the transmission chain. We also note here that the data (viraemic levels) involved in the ANOVA computations were transformed to the logarithmic scale to satisfy the assumptions required for the tests. Our analysis relies on the stochastic reconstruction of the unobserved infection process, which is based on identifying individual infectious in-contact animals as the possible infecting source. However, because inhalation of infectious droplets is considered to be a common route of FMD infection in sheep, a more complicated representation of the transmission process provides scope for possible extensions to our model.

We thank two referees for their constructive suggestions that improved the paper. We are also grateful to Mark Woolhouse for many valuable discussions and comments, and to Gareth Hughes for providing the data. This work was funded by the 213/MAF12261 BBSRC grant.

REFERENCES

Bailey, N. T. J. 1975 *The mathematical theory of infectious diseases and its applications*, 2nd edn. London: Griffin.

Becker, N. G. 1983 Analysis of data from a single epidemic. *Aust. J. Statist.* **25**, 191–197.

Becker, N. G. 1989 *Analysis of infectious disease data*. London: Chapman & Hall.

Britton, T. & O'Neill, P. D. 2002 Statistical inference for stochastic epidemics in populations with random social structure. *Scand. J. Statist.* **29**, 375–390.

Cox, S. J., Barnett, P. V., Dani, P. & Salt, J. S. 1999 Emergency vaccination of sheep against foot-and-mouth disease: protection against disease and reduction in contact transmission. *Vaccine* **17**, 1858–1868.

De Angelis, D., Gilks, W. R. & Day, N. E. 1998 Bayesian projection of the acquired immune deficiency syndrome epidemic (with discussion). *Appl. Statist.* **47**, 449–498.

Ferguson, N. M., Donnelly, C. A., Woolhouse, M. E. J. & Anderson, R. M. 1997 The epidemiology of BSE in GB cattle herds. II. Model construction and analysis of transmission dynamics. *Phil. Trans. R. Soc. Lond.* B **352**, 803–838. (DOI 10.1098/rstb.1997.0063.)

Gelfand, A. E. & Smith, A. F. M. 1990 Sampling-based approaches to calculating marginal densities. *J. Am. Statist. Assoc.* **85**, 398–409.

Ghani, A. C., Ferguson, N. M., Donnelly, C. A., Hagenaars, T. J. & Anderson, R. M. 1998 Epidemiological determinants of the pattern and magnitude of the vCJD epidemic in Great Britain. *Proc. R. Soc. Lond.* B **265**, 2443–2452. (DOI 10.1098/rspb.1998.0596.)

Gibson, C. F. & Donaldson, A. I. 1986 Exposure of sheep to natural aerosols of foot-and-mouth disease virus. *Res. Vet. Sci.* **41**, 45–49.

Gibson, G. J. 1997 Markov chain Monte Carlo methods for fitting spatiotemporal stochastic models in plant epidemiology. *Appl. Statist.* **46**, 215–233.

Gibson, G. J. & Renshaw, E. 1998 Estimating parameters in stochastic compartmental models using Markov chain Monte Carlo methods. *IMA J. Math. Appl. Med. Biol.* **15**, 19–40.

Gibson, G. J. & Renshaw, E. 2001 Likelihood estimation for stochastic compartmental models using Markov chain methods. *Statist. Comput.* **11**, 347–358.

Höhle, M., Jorgensen, E. & O'Neill, P. D. 2003 Inference in disease transmission experiments using stochastic epidemic models. Research Report 03-13, Division of Statistics, School of Mathematical Sciences, University of Nottingham, UK.

Hughes, G. H., Kitching, R. P. & Woolhouse, M. E. J. 2002*a* Dose-dependent responses of sheep inoculated intranasally with type O foot-and-mouth disease virus. *J. Comp. Pathol.* **127**, 22–29.

Hughes, G. H., Mioulet, V., Haydon, D. T., Kitching, R. P., Donaldson, A. I. & Woolhouse, M. E. J. 2002*b* Serial passage of foot-and-mouth disease virus in sheep reveals declining levels of viraemia over time. *J. Gen. Virol.* **83**, 1907–1914.

Kitching, R. P. & Mackay, D. K. 1994 Foot and mouth disease. *State Vet. J.* **4**, 7–10.

Mackay, D. K. J., Newman, B. & Sachpatzidis, A. 1995 Epidemiological analysis of the serological survey for antibody to FMD virus, Greece 1994. Report of the Institute for Animal Health, Pirbright, UK, and the Animal Health Office, Xanthi, Greece.

Meng, X. L. 1994 Posterior predictive p-values. *Ann. Statist.* **22**, 1142–1160.

O'Neill, P. D. & Becker, N. G. 2001 Inference for an epidemic when susceptibility varies. *Biostatistics* **2**, 99–108.

O'Neill, P. D. & Roberts, G. O. 1999 Bayesian inference for partially observed stochastic epidemics. *J. R. Statist. Soc.* A **162**, 121–129.

Sellers, R., Herniman, K. & Gumm, I. 1977 The airbourne dispersal of foot and mouth disease virus from vaccinated and recovered pigs, cattle and sheep after exposure to infection. *Res. Vet. Sci.* **23**, 70–75.

Streftaris, G. & Gibson, G. J. 2004 Bayesian inference for stochastic epidemics in closed populations. *Statist. Model.* **4**, 63–75.

Tierney, L. 1994 Markov chains for exploring posterior distributions (with discussion). *Ann. Statist.* **22**, 1701–1762.

As this paper exceeds the maximum length normally permitted, the authors have agreed to contribute to production costs.

Visit www.journals.royalsoc.ac.uk and navigate to this article through *Proceedings*: Biological Sciences to see the accompanying electronic appendix.