THE ROYAL SOCIETY

Feed-borne transmission and case clustering of BSE

T. J. Hagenaars*, N. M. Ferguson, C. A. Donnelly, A. C. Ghani and R. M. Anderson

The Wellcome Trust Centre for the Epidemiology of Infectious Disease, Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3FY, UK

An unresolved issue in the epidemiology of bovine spongiform encephalopathy (BSE) in the UK is what precisely determines the degree to which cases of disease in cattle are clustered within herds throughout the course of the epidemic. This paper presents an analysis of feed-borne transmission at the herd level and tests various models of case-clustering mechanisms, associated with heterogeneity in exposure to infectious feed, against observed epidemic pattern. We use an age-structured metapopulation framework in which the recycling of animal tissue between herds via feed producers is explicitly described. We explore two alternative assumptions for the scaling with herd size of the within-herd risk of exposure of an animal to infectious material. We find that whereas exposure heterogeneity caused by variation in feed and offal processing methods and by variation in per-animal feed uptake can explain the pattern of case clustering seen in the BSE epidemic, exposure heterogeneity due to the aggregation of infectivity within feed cannot.

Keywords: BSE; transmission dynamics; clustering; epidemiology; mathematical model

1. INTRODUCTION

Bovine spongiform encephalopathy (BSE) was first identified in Great Britain (GB) in November 1986 (Wells *et al.* 1987). By late 1999, over 175 000 cases of BSE had been confirmed in GB, 1787 in Northern Ireland, 411 in the Republic of Ireland, and 323 in Portugal. The BSE epidemic has had a lasting impact on the European agricultural industries. In March 1996, the UK Secretary of State for Health announced that the most probable cause of a new variant of Creutzfeldt–Jakob disease (vCJD) in humans was exposure to meat products infected with the aetiological agent of BSE (Calman 1996). Since this time, scientific evidence has accumulated, strengthening the link between BSE and vCJD (Hill *et al.* 1997; Collinge *et al.* 1996; Bruce *et al.* 1997).

The dominant route of BSE infection in cattle is thought to have been via consumption of feedstuff containing meat and bone meal (MBM) rendered from infected cattle offal, with maternal transmission having been shown to play a minor role (Wilesmith *et al.* 1997; Donnelly *et al.* 1997*a*,*d*,*e*; Curnow *et al.* 1997; Gore *et al.* 1997). Whilst the transmission dynamics of BSE at the population level have been extensively analysed in previous work (Anderson *et al.* 1996; Donnelly *et al.* 1997*b*,*c*; Ferguson *et al.* 1997, 1999; De Koeijer *et al.* 2000), this paper is the first to model feed-borne infection processes at the herd level. Our aim is to explore possible mechanisms that could have given rise to the significant case clustering within cattle holdings and herds seen in the BSE case data: the distribution of observed cases within cattle herds is highly overdispersed, with over 80% of cases occurring in 20% of herds. Such a distribution is characterized, for example, by a variance that is (much) larger than its mean. Also, data from GB and the Republic of Ireland (Griffin *et al.* 1997) indicate that cases tend to cluster in large herds; more precisely, that the per-capita incidence of BSE increases with herd size (see Donnelly *et al.* 1997*b*). For a discussion of how the conclusions from the previous population-level analyses might be affected by the occurrence of case clustering see Ferguson *et al.* (1999).

The summary statistics we use characterize case clustering in data (and model results); they are the variance-to-mean ratio $\mathcal{K} = \sigma^2/\mu$ of the distribution of observed cases within herds and the mean per-herd incidence, stratified by cohort and herd size. The degree of case clustering is measured by how much $\mathcal{K}$ exceeds unity, the value corresponding to a Poisson distribution (no clustering). In figure 1 we show the variance-to-mean ratios versus mean, observed for the birth cohorts 1985–1994. The different symbols represent different herd-size categories. We observe that the degree of clustering tends to increase with μ. Also, we see that the data for all herd-size categories, except for the smallest, lie on roughly one and the same curve. We note that the finite width of the herd-size categories used here affects the variance-to-mean ratio expected in the absence of clustering (i.e. when cases are Poisson-distributed for any given herd size). However, as illustrated in figure 1 (dotted and dashed lines), the difference with the $\mathcal{K} = 1$ (long dashed) line is very small and will therefore be ignored henceforth.

A variety of mechanisms may contribute to the observed clustering of cases. The most important is likely to be heterogeneity in the exposure to infectious feed, generated by variation in feed processing methods, aggregation of infectivity within MBM, spatial locality in the production and distribution of feed, together with holding-level variation in husbandry practices. Other factors such as varying age-specific survival probabilities and genetically determined variations in susceptibility may also contribute.

The first mechanism we consider is aggregation of infectivity in feed. This might arise at scales ranging from that of the bite (due to, for example, the presence or absence of

*Author for correspondence (thomas.hagenaars@ceid.ox.ac.uk).

Received 2 September 1999 *Accepted* 12 October 1999

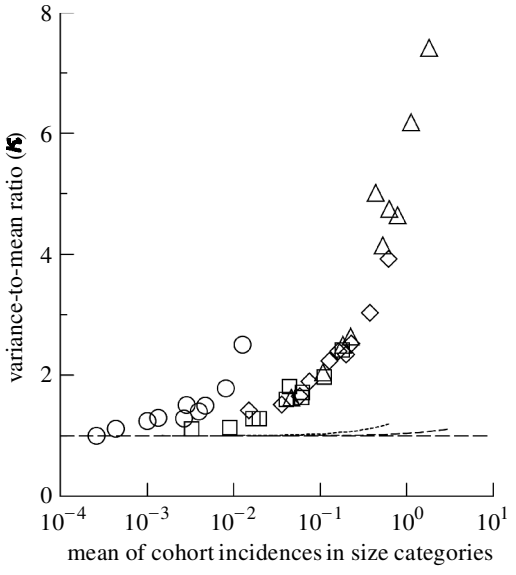

Figure 1. Variance-to-mean ratios (symbols) versus mean of cohort-herd incidences as observed in the epidemic in GB (see Donnelly *et al.* 1997*b*, fig. 12; here we show the results when ignoring data with missing size category). Circles, 1–29 adult animals; squares, 30–49 adult animals; diamonds, 50–99 adult animals; triangles, 100+ adult animals. Thin lines, expected mean–variance relationships for the 1–29 (dotted) and the 50–99 (dashed) adult animals' size categories under the assumption of Poisson-distributed cases. Long dashed line, $\mathcal{K} = 1$ (mean equals variance).

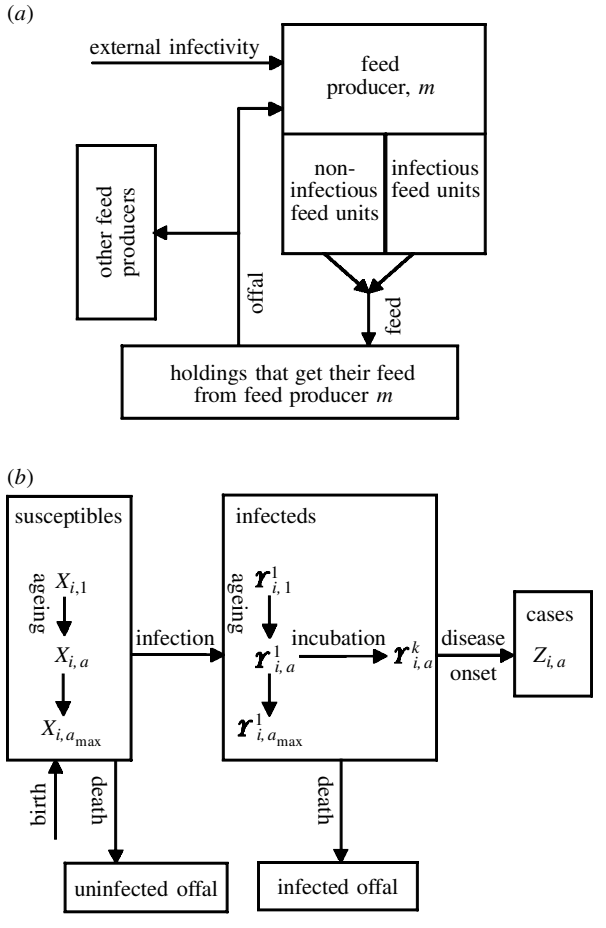

Figure 2. Schematic representation of the model structure considered in this paper. (*a*) The recycling of animal tissue, (*b*) animal population and demographic, infection, and incubation processes. Demographic (α_a, μ_a) and incubation (ν) parameters are kept fixed, whereas the parameters controlling recycling and infection processes ($\delta_m, \theta_m, \eta_m, \gamma, c, \zeta, b$) are being varied.

a fragment of nervous system tissue) to the bag or batch (with batches potentially being very large due to the scale of the feed production process). We therefore model feed in terms of discrete units which can be arbitrarily sized to explore the effect of aggregation scale on clustering, formulating a stochastic susceptible-infected (SI) model framework that explicitly describes the recycling of animal material. Our analyses use epidemiological parameter estimates obtained from back-calculation analyses of the BSE case data in GB (Anderson *et al.* 1996; Ferguson *et al.* 1997) and explore a wide range of transmission and mixing parameter combinations which produce results consistent with observed 1985 and 1986 cohort incidences in GB. Apart from infectivity aggregation, we also consider herd-level heterogeneity of per-animal feed uptake, and variation in feed production processes as mechanisms that can generate case clustering.

We characterize the effects of different parameter choices on the transmission dynamics of the epidemic in terms of the basic reproduction number and the 'generation time' of BSE. The basic reproduction number is defined as the average number of secondary infections generated by a primary infection in an entirely susceptible population. Similarly, the generation time is the average interval between the infection of a host and the transmission of that infection to a second host in a fully susceptible population.

2. MODEL FORMULATION

We study feed-borne transmission of BSE using a mathematical model that explicitly takes into account the recycling of animal tissue in the form of MBM. We may distinguish two epidemiological scales relevant to this process. The between-herd variation is represented by the structuring of the cattle population into herds that are linked through use of the same feed producers or suppliers (figure 2*a*). This metapopulation framework provides a simplified description of the actual feed production and distribution process. Within-herd dynamics are governed by demographic, infection and incubation processes (figure 2*b*).

(a) *Between-herd processes*

In our model we explicitly incorporate one or more feed production plants (labelled with index m), from which individual herds (labelled with index i) acquire animal feed. We characterize production plants by two parameters, η_m, reflecting the degree to which the infectiousness of contaminated material is reduced during rendering, and θ_m, reflecting the degree to which one animal's tissue is dispersed over several units of feed. Since we do not explicitly model the rendering process, these two parameters characterize the average MBM supplied to feed production plant m.

Animal feed is modelled as discrete units, each of which may be infectious. The average number of such units consumed per animal per week, c, describes the size

of units being used, and hence the scale on which infectivity can aggregate in feed. At each time-point, a number of feed units are added to the store of a producer equal to the number of units consumed by its customers at that moment. The proportion of feed units infected p_{I} is related to the relative infectiousness of offal $\chi_m(t)$ (expressed in terms of slaughtered infectives in the next section) processed at that time, in the following way:

$$p_{\mathrm{I}} = 1 - (1 - \chi_m(t))^{\theta_m}, \qquad (1)$$

for $0 \leqslant \chi_m(t) \leqslant 1$, where $\chi_m(t) = 0$ if there is no infectivity and unity if all MBM is infectious. θ_m characterizes the degree of mixing of tissue during rendering and feed production. For small θ_m, infectivity is concentrated in relatively few feed units, whilst as θ_m becomes large any level of infectivity will infect all feed units. The rationale for choosing the form of equation (1) for p_{I} is that each unit is thought to contain traces of a certain number (θ_m) of carcasses that each have a probability χ_m of containing the infective agent.

Feed distribution is modelled in a way that allows for some randomness. We assume that deliveries of batches of feed to each herd occur after time intervals that vary somewhat around a herd-independent mean interval (chosen to equal six weeks in the model calculations presented here). A batch equals the amount of feed required for feeding the herd during a mean time interval. In order to achieve variations in the time interval between deliveries, we use two threshold values for the filling of the herd's feed store. At each time-point, if this filling has fallen below the lower threshold (here chosen to be 50% of the feed required by the herd for six weeks), a new delivery will take place; if it is above the lower threshold, but below the upper threshold (here chosen to be 60% of the six-week feed amount) there is a certain non-zero probability (here chosen to be $(0.6 - \max(\pi, 0.5))/0.1$ where π is the proportion of feed remaining in the feed stock) that the herd receives a new delivery of feed. The number of infectious units in each delivery is a binomial random variable with probability p_{I} (equation (1)). The order in which herds receive feed deliveries at any one time-step is random. Each feed producer begins with a stock of feed large enough (here chosen to be 14 weeks' feed for its customers) to cope with chance fluctuations in the demand. For the parameter values used, the mean storage time of feed units is approximately three months. The qualitative character of the predictions of the model do not change when using different (but meaningful) values of the feed delivery parameters (e.g. a doubling of the mean storage time used leaves our results essentially unchanged).

Between-herd interaction is modelled through groups of herds acquiring feed from the same feed producer. For simplicity we assume that each herd i only buys feed from feed producer $m(i)$.

(b) *Within-herd processes*

Within each herd, animals are stratified by age a (discretized in quarter years) and incubation stage k. Let $X_{i,a}$, $Y^k_{i,a}$ and $Z_{i,a}$ denote the number of susceptible animals, infected animals in incubation stage k, and clinical cases, respectively, at age a in herd i.

We model a set of herds comprising four size categories (0–29, 30–49, 50–99 and >100 adult cattle). The herds are distributed over these categories according to the ratios 15:3:4:2, closely approximating the herd-size distribution in GB. Herds within one category start off being identical replicas. As the stochastic simulation progresses, demographic processes result in variation in herd sizes within each category. Initially, all animals in a herd are susceptible with an age distribution taken from the mean age distribution for demographic equilibrium in the absence of BSE-related mortality.

(i) *Population dynamics*

Deterministically, infection, incubation and demographic processes are governed by the following equations:

$$\left.\begin{aligned}
\frac{\mathrm{d}X_{i,1}}{\mathrm{d}t} &= \sum_{a'} \alpha_{a'} N_{i,a'} - \lambda_{i,1} X_{i,1} - \mu_1 X_{i,1} \\
\frac{\mathrm{d}X_{i,a}}{\mathrm{d}t} &= -\lambda_{i,a} X_{i,a} - \mu_a X_{i,a}, \qquad (1 < a \leqslant a_{\max}) \\
\frac{\mathrm{d}Y^1_{i,a}}{\mathrm{d}t} &= \lambda_{i,a} X_{i,a} - \mu_a Y^1_{i,a} \qquad (1 < a \leqslant a_{\max}) \\
\frac{\mathrm{d}Y^k_{i,a}}{\mathrm{d}t} &= -\mu_a Y^k_{i,a} + \nu_{k-1} Y^{k-1}_{i,a} - \nu_k Y^k_{i,a} \\
&\qquad (1 < a \leqslant a_{\max},\ 1 < k < k_{\max}) \\
\frac{\mathrm{d}Z_{i,a}}{\mathrm{d}t} &= \nu_{k_{\max}} Y^{k_{\max}}_{i,a}
\end{aligned}\right\}, \qquad (2)$$

where $N_{i,a} = (X_{i,a} + \sum_k Y^k_{i,a})$, μ_a is the age-dependent slaughter rate, α_a is the age-dependent calving rate, ν_k is the transition rate from incubation stage k to $k+1$, and λ_i denotes the force of infection in herd i.

Animals age deterministically, moving from $X_{i,a}$ to $X_{i,a+1}$ every three months. Similarly, to obtain a two-year time delay at the start of the incubation period (i.e. no cases of BSE over the first two years post infection, as described in Ferguson *et al.* (1997)), the first eight quarter years of the incubation period are updated deterministically along with the ageing process. Other processes are implemented stochastically, using a discrete time approximation, i.e. the changes in $X_{i,a}$, $Y^k_{i,a}$ and $Z_{i,a}$ occurring at each (small) time-step are drawn from Poisson distributions. The rates governing these processes (birth, slaughter, infection, post-two-year incubation) are given by the relevant terms on the right hand side of equations (2).

We use multiple (stochastic) incubation stages to allow the post-two-year incubation period distribution to be non-exponentially distributed. The post-two-year transition probabilities are chosen to be independent of k, giving a gamma distribution. By taking $\nu_k = 0.0356$ per week the resulting overall incubation period distribution is similar to that estimated in the best-fitting model in back-calculation analyses of the GB case database (Ferguson *et al.* 1997). Animals leaving the last incubation stage reach clinical onset of BSE (equations (2)).

We assume that animals older than two years calve annually. Age-dependent slaughter rates are calculated from the survival function estimated in Anderson *et al.* (1996) and Donnelly *et al.* (1997*b*).

(ii) *Infectiousness of offal*

For the incubation-stage dependence of the infectiousness of offal we assume the the following mathematical form

$$\Omega_k = \zeta + (1-\zeta)\exp(-b\mathcal{T}(k)), \tag{3}$$

where ζ and b are parameters and $\mathcal{T}(k)$ is the mean time for animals to go from the kth incubation stage to the last stage. This form comprises scenarios ranging from an infectiousness independent of incubation stage to an infectiousness that is very small initially and increases dramatically towards the end of incubation.

The relative infectiousness of all offal at time t for feed producer m is then given by

$$\chi_m(t) = \sum_k \Omega_k p_m^k(t) + \delta_m(t), \tag{4}$$

where p_m^k is the proportion of animals slaughtered at time t that are in incubation stage k, and $\delta_m(t)$ is an external infectivity source, representing the infectivity of offal unrelated to feed-borne BSE infection (e.g. a scrapie-like or sporadic BSE-like agent). Such an external source is required to initiate the simulated epidemics.

We consider two types of external infectivity source. The first is a single event at the start of the simulation (corresponding biologically to the recycling into feed of a sporadic case of BSE infection in cattle), modelled by a $\delta_m(t)$ that is non-zero only during the first time-step. The second is a low-level source that remains present through time (corresponding biologically to a scrapie-like agent). Here we have chosen a linear onset over a period of three years (in contrast to the abrupt onset of the first type of source), after which the source remains constant.

In the situation where all slaughtered infectious animals have maximal infectiousness and where the relative contribution of δ_m to $\chi_m(t)$ is marginal, we may think of the tissue mixing parameter θ_m in equation (1) as the average number of different animals from which the MBM in one unit of feed originates.

(iii) *The within-herd force of infection*

The dynamics in different herds are coupled through the force of infection parameter $\lambda_{i,a}$, given by

$$\lambda_{i,a}(t) = g(a)\eta_{m(i)}(t)e_i(t), \tag{5}$$

where $g(a)$ is the estimated age-dependent exposure–susceptibility distribution from the best-fitting back-calculation model in Ferguson *et al.* (1997) (normalized to reach a maximum value of unity), $m(i)$ denotes the index of the feed producer that supplies herd i, $\eta_{m(i)}(t)$ is the relative infectiousness of feed units originating from feed producer $m(i)$ (representing the degree to which the production process reduces the infectivity of the raw material), and $e_i(t)$ is the proportion of animals in herd i that are exposed to infectious feed at time t.

When n feed units are delivered to herd i, from which $\tilde{n}$ are infectious, these are converted into $n/(d_i c)$ weekly portions added to the farmer's store out of which $\tilde{n}/(d_i c)$ are infectious. As explained above, the parameter c, the average number of feed units consumed per animal per week, controls the scale on which infectivity can aggregate in feed. The parameter d_i is the relative feed demand (its average over all herds equals unity), i.e. it describes a herd-level variation in per-animal feed uptake. The size-category means of d_i are taken equal to unity. The proportion of cows exposed in a week, $e_i(t)$, is a function of $f_i(t) = \tilde{m}_i(t)/N_i(t)$, the ratio of the number of infectious portions consumed, $\tilde{m}_i(t)$, and the total number of portions (equal to the number of cows) consumed in that week in herd i.

The mathematical relationship assumed between $e_i(t)$ and $f_i(t)$ determines the manner in which feed units are distributed amongst animals. We consider two different models for this relationship, one labelled mass action (MA) and one pseudo mass action (PMA). For mass-action feed consumption, the cows that are unexposed at time t are those that have not eaten any infectious portions. Hence,

$$e_i(t) = 1 - (1 - f_i(t))^{\gamma_{\mathrm{MA}}}, \tag{6}$$

where γ_{MA} (the 'feed consumption parameter') measures the average number of feed units per week from which an individual animal will eat (allowing that animals share feed units). For small values of $f_i(t)$, equation (6) may be approximated by

$$e_i(t) \approx \gamma_{\mathrm{MA}} f_i(t).$$

For the PMA feed-consumption model, the proportion exposed, $e_i(t)$, is taken to be

$$e_i(t) = 1 - (1 - \gamma_{\mathrm{PMA}})^{cN_i(t)f_i(t)}. \tag{7}$$

In this model, the feed is distributed in such a way amongst the animals that, for every infectious unit consumed, a proportion γ_{PMA} of the herd is exposed to the risk of infection. For small values of γ_{PMA} and $cN_i(t)f_i(t)$, equation (7) may be approximated by

$$e_i(t) \approx -\log(1 - \gamma_{\mathrm{PMA}})cN_i(t)f_i(t). \tag{8}$$

The nomenclature MA/PMA derives from the analogy with scaling assumptions for the contact rate in models for direct horizontal transmission (see De Jong *et al.* 1995). For (true) MA the infection hazard is taken to be proportional to the density or proportion of infectious units (individuals), whereas for PMA it is taken to be proportional to the number of such units (individuals), cf. equations (7) and (8).

Apart from their different clustering behaviour (discussed in § 4(b)), the MA and PMA models differ in the way the per capita incidence depends on holding size. This is because for the PMA model, given a sufficiently low infection prevalence, the infection risk experienced by an animal is proportional to the number of infectious feed units from which the animal has eaten (equation (8)), and it is assumed that feed is sufficiently well mixed within a holding for the probability of consuming a part of any one feed portion to be independent of the size of the holding. Because the total consumption of feed is proportional to holding size, this gives rise to a per capita infection risk which also scales with holding size. In contrast, for the MA model the per capita risk is always independent of holding size. Thus only the PMA model can reproduce the observed linear dependence of per capita incidence on holding size without assuming, for instance, that larger holdings were more likely to employ

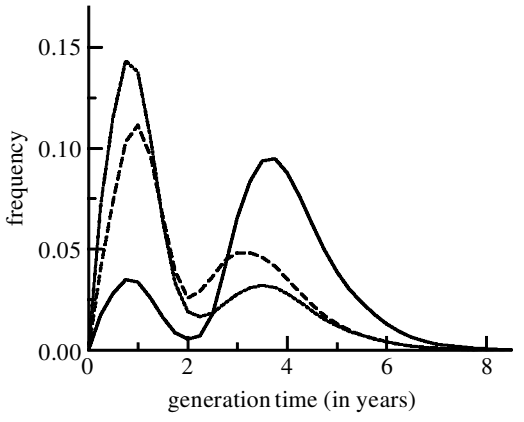

Figure 3. Distribution of the generation time between a primary and a secondary infection, for three different combinations of the offal infectiousness parameters ζ and b. The first peak of this distribution (the magnitude of which is controlled by the infectiousness level at around one year after infection) is due to the high slaughtering rates around two years of age, combined with the strong peak in susceptibility at one year of age. The second peak represents the contribution of animals that are slaughtered when adult and close to onset, and its bell shape reflects the shape of the incubation-time distribution. Solid line, $\zeta = 0.014$, $b = 0.70$, $T_g = 14.1$; dashed line, $\zeta = 0.022$, $b = 0.14$, $T_g = 8.7$; dotted line, $\zeta = 0.187$, $b = 2.42$, $T_g = 7.3$.

more intensive farming methods, and thus use more MBM feed. We note however that, as discussed by Donnelly & Ferguson (1999), the observed scaling of per capita incidence with holding size can also be explained by assuming that only a subset of holdings of any particular size were ever 'at risk' of BSE infection—in which case MA transmission may have occurred within the 'at risk' subset.

3. BASIC REPRODUCTION NUMBER AND GENERATION TIME DISTRIBUTION

The basic reproduction number R_0 is defined here as the expected number of secondary infections arising from a single primary infection in a totally susceptible population. R_0 measures the capability of the infectious agent to generate an expanding epidemic (Anderson & May 1991). In particular, in deterministic models, infection can only persist endemically if $R_0 > 1$.

For feed-borne transmission of BSE, it can be shown (Diekmann *et al.* 1990; Heesterbeek & Dietz 1996; Ferguson *et al.* 1999) that R_0 is given by

$$R_0 = n(0) \int S(a)g(a) \int A(\tau,a)\mathrm{d}\tau\mathrm{d}a, \tag{9}$$

where $n(0)$ is the number of newborn susceptibles in a unit of time, $S(a)$ is the survival probability to age a, $g(a)$ is the age-dependent exposure–susceptibility distribution, and $A(\tau,a)$ denotes the expected infectivity due to an individual that was infected τ units of time ago when its age was a, towards a susceptible with unit susceptibility.

For the deterministic versions of the models considered here, in the case of a single feed producer $A(\tau,a)$ is given by the product of the probability that the animal is slaughtered at age $a+\tau$ given that it has already survived to age a $(-(\frac{\mathrm{d}S}{\mathrm{d}a})(a+\tau)/S(a))$, the expected proportion of cows that will be exposed to the infectious feed produced from this animal's tissue (equal to $\theta\gamma_{\mathrm{MA}}/n(0)$ for MA), the relative feed infectiousness η, and the expected infectiousness of the individual's tissue:

$$\left.\begin{aligned} A_{\mathrm{MA}}(a'-a,a) &= -\frac{\theta\eta\gamma_{\mathrm{MA}}}{n(0)S(a)}\frac{\mathrm{d}S}{\mathrm{d}a}(a')\sum_k \Omega_k P_k(a'a), \\ A_{\mathrm{PMA}}(a'-a,a) &= \frac{c\sum_i \mathcal{N}_i^2(0)}{\sum_i \mathcal{N}_i(0)}\log(1-\gamma_{\mathrm{PMA}})\frac{\theta\eta}{n(0)S(a)} \\ &\quad \times \frac{\mathrm{d}S}{\mathrm{d}a}(a')\sum_k \Omega_k P_k(a',a) \end{aligned}\right\}.$$

Here $P_k(a',a)$ is the probability that an animal infected at age a is in incubation stage k at age a' in the absence of mortality; for post-two-year incubation stages it reads

$$P_k(a',a) = \frac{\left[\nu(a'-a-k_0)\right]^{k-k_0}}{(k-k_0)!}\mathrm{e}^{-\nu(a'-a-k_0)} \qquad (k_0 + a < a' \leqslant a_{\max}, k > k_0), \tag{10}$$

where k_0 denotes the incubation stage just before an infection age of two years. The first factor in the right-hand side of equation (10) measures the number of feed units per week from which an animal eats, averaged over all animals. As a result, R_0 for the PMA model depends on (the second moment of) the distribution of herd size.

The initial growth rate of an epidemic is determined by the magnitude of R_0 and by the generation time distribution, defined to be the distribution of the times between a primary and a secondary infection in a wholly susceptible population. The mean of this distribution is denoted $\mathcal{T}_g$, and is given by

$$\left.\begin{aligned} \mathcal{T}_g &= \frac{\int\int C(\tau,a)\tau\mathrm{d}\tau\mathrm{d}a}{\int\int C(\tau,a)\mathrm{d}\tau\mathrm{d}a} \\ C(\tau,a) &= g(a)\frac{\mathrm{d}S}{\mathrm{d}a}(a+\tau)\sum_k \Omega_k P_k(a+\tau,a) \end{aligned}\right\}. \tag{11}$$

Figure 3 illustrates the bimodal nature of the generation time distribution of BSE, and its dependence on the infectiousness parameters ζ and b. In general, the mean generation time is correlated with the rate at which tissue infectivity increases over the duration of the incubation period. If animals are equally infectious throughout the incubation period, the mean generation time is a little over one year (the first peak in the distributions shown in figure 3 contributes most to the average), as most animals are infected at around one year of age, and are slaughtered at around two years of age. However, if animals are only infectious in the late stages of the incubation, the mean generation time is closer to the mean incubation period of five years (the second peak in the generation time distribution contributes most to the average).

Some insight into the absolute infectiousness of infected animal tissue is gained by considering the average number of secondary infections generated (in an entirely susceptible population) by one maximally infectious animal which is slaughtered before disease onset, I_0. For the MA model this is given by

$$I_0 = -\theta\eta\gamma_{\mathrm{MA}} \int\int g(a)\frac{\mathrm{d}S}{\mathrm{d}a}(a+\tau)\mathrm{d}\tau\mathrm{d}a. \tag{12}$$

Table 1. *Parameter value ranges used for sampling parameter combinations.*

symbol	quantity	range sampled from
δ	external infectivity	$1.0 \times 10^{-7} < \delta < 5.0 \times 10^{-5}$ (continuous source)
		$1.0 \times 10^{-4} < \delta < 5.0 \times 10^{-1}$ (single event)
θ	tissue mixing	$1.0 < \theta < 100.0$
η	feed producer infectiousness	$0.001 < \eta < 1.0$
γ	feed consumption	$1.0 < \gamma_{\mathrm{MA}} < 100$
		$0.0 < \gamma_{\mathrm{PMA}} < 1.0$
c	unit consumption rate	$0.002 < c < 10.0$ (units cow^{-1} week^{-1})
ζ	infectiousness base-line level	$0.001 < \zeta < 1.0$
b	infectiousness exponent	$0.0 < b < 2.5$ (1/(quarter-year))

4. MODEL RESULTS

(a) *Sensitivity analysis*

In this section we explore the parameter ranges for external infectivity, tissue mixing, feed producer infectiousness, feed consumption, tissue infectiousness and infectivity aggregation, which give model results that are consistent with the observed total incidence in GB arising from the 1985 and 1986 birth cohorts. Later cohorts were not used because the incidence in these cohorts was affected by the MBM ban introduced in July 1988, whilst the case data from earlier cohorts suffer from bias due to under-reporting. We are thus more concerned with exploring epidemic patterns generated by the recycling of animal tissue that are broadly consistent with these case data, rather than fitting the model to the entire epidemic—which would necessitate allowing infectiousness–consumption parameters to vary with time. Back-calculation models (Anderson *et al.* 1996; Ferguson *et al.* 1997) are more easily suited to detailed fitting. In this analysis, we therefore do not consider temporal variations in any parameters.

Deterministic versions of the model are used to examine predictions based on different parameter combinations. Whilst the epidemic produced by a deterministic model is not typically identical to the mean of all possible epidemic realizations of the equivalent stochastic model, in the absence of strong infectivity aggregation in feed the differences during the phase of epidemic growth that are matched to the 1985 and 1986 cohorts are small enough to justify the approximation, given the dramatic reduction in the computational burden required to average over many realizations for each parameter combination. In the case of strong infectivity aggregation (small c) the differences can be big; however, we shall see below that such scenarios can not produce the observed pattern of clustering irrespective of how well the initial phase of growth matches the overall incidence in the 1985 and 1986 cohorts. Table 1 gives the bounds within which the parameter values are sampled (using a Latin hypercube technique (Stein 1987; McKay *et al.* 1979)). A parameter combination is accepted when the incidence in two consecutive cohorts (scaled for population size) is consistent with the observed 1985 and 1986 cohort incidences on the basis of a likelihood ratio test (with d.f. = 2).

Figure 4*a*,*b* illustrates the combinations of R_0 and mean generation time (T_g) consistent with the case data, for MA with continuous (*a*) and with single event (*b*) external infection source (both for uniform feed demand d_i and a single feed producer). Corresponding results for PMA (not shown) are very similar. The (T_g, R_0) combinations consistent with the case data are those that give epidemic growth rates which match that seen between the 1985 and 1986 cohorts. In figure 4*a* four scenarios are highlighted (symbols) for which the cohort incidences are shown in figure 4*c*, which illustrates how different scenarios that have very similar right-censored incidences (grey bars) diverge in later birth years.

The broadening of the range of accepted R_0-values with increasing T_g is due to oscillations in infection incidence in the early stages of the epidemic corresponding to identifiable generations of infection. These identifiable generations arise from the onset of the external infectivity. As the epidemic progresses, the incubation period and survival distributions cause these once discrete generations of infection to increasingly overlap, thus damping the oscillations. However, such oscillations add sufficient variability to the epidemic growth rate for the 1985–1986 cohorts to broaden the range of accepted R_0-values. This effect increases with the mean generation time, because the oscillations become slower, allowing both the 1985 and 1986 cohorts to be on the rising phase (where the real rate of epidemic growth is greater than the mean rate determined by R_0 and T_g) or flattening phase (where the real growth rate is lower than the mean) of a single oscillation. The weaker broadening of the accepted R_0 range for the continuous external source model is due to the smoother onset of primary infections.

Figure 5 shows the relationships between R_0, T_g and the average number of secondary infections caused by one maximally infectious bovine slaughtered before disease onset, I_0, for the accepted parameter combinations from figure 4*a* (mass action, continuous external source). We observe that I_0 takes a wide range of values, increasing rapidly in magnitude with increasing R_0 and T_g. For the smallest accepted R_0 and T_g, I_0 is of the same order of magnitude as R_0, but for the largest R_0 and T_g it is much bigger than R_0. This is because in the latter regime, infected animals are only substantially infectious close to onset, so that the next generation of infections arises entirely from the small fraction of the previous generation that is slaughtered just before onset (see also Ferguson *et al.* 1999).

In figure 6 we explore how the onset of epidemics varies with external infectiousness δ and inverse infectivity aggregation scale c. We show the median duration from the onset of the external infectivity to the occurrence of the first case as well as that to the 100th case. This duration as well as its variation (as measured by the 10th and 90th percentile) increase with decreasing δ and c.

(b) *Case clustering*

Exploration of the case clustering generated by the complex transmission dynamics of BSE requires stochastic simulations of the epidemic model. These simulations use a representative subset of parameter sets, as discussed in

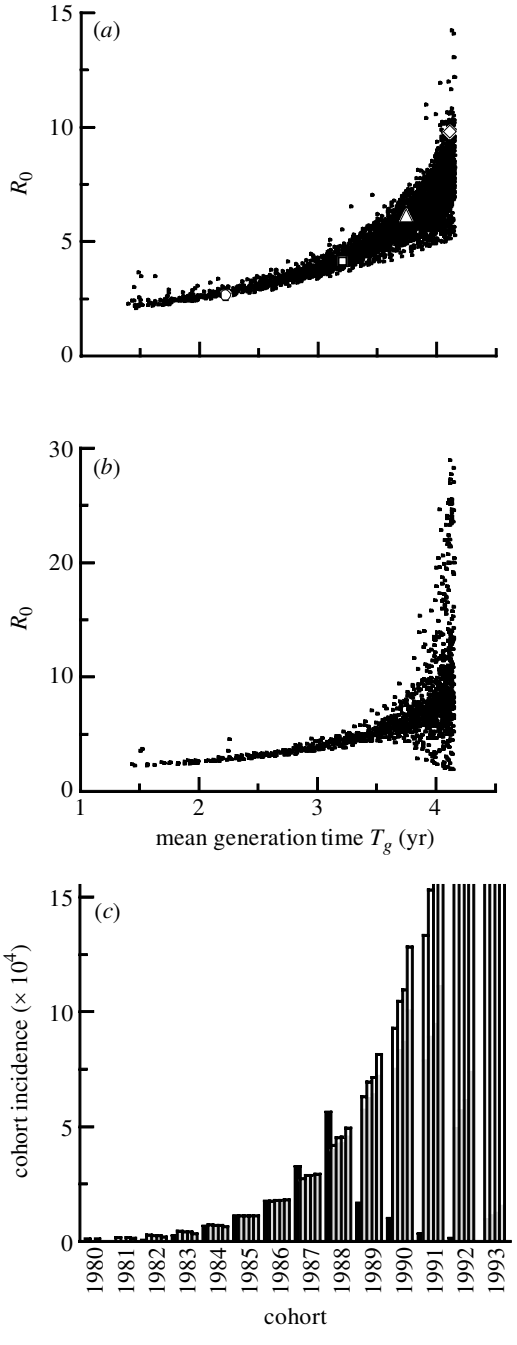

Figure 4. 6000 accepted parameter combinations, shown as dots labelled by their (T_g,R_0)-values. (*a*) Mass action with continuous external source (for open symbols see (*c*)); (*b*) mass action with single event. (*c*) Cohort case incidences for four accepted parameter sets from (*a*) compared with the corresponding incidences from the case data. Histogram bars from left to right: case data (black), parameters sets indicated in (*a*) with circle, square, triangle and diamond. Grey bars are cohort incidences subject to the same right censoring as the case data in black (i.e. cases up to 1 January 1998). The data and model results start to deviate very strongly after the 1988 cohort due to our choice not to model the effect of the 1988 MBM ban.

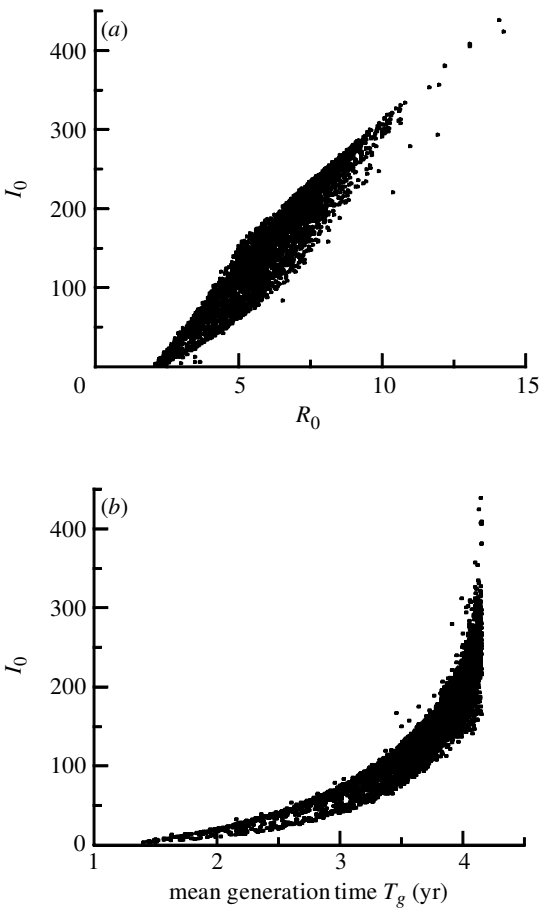

Figure 5. Accepted parameter combinations for mass action with continuous external source (see figure 4*a*), shown labelled by their (*a*) (R_0,I_0)-values, (*b*) (T_g,I_0)-values. Here I_0 is the average number of animals infected by one maximally infectious bovine in an entirely susceptible population.

the previous section, which generate results consistent with the case data. Epidemics are simulated in systems of up to 3600 holdings (distributed between four size categories in proportion to their respective frequencies calculated from agricultural census data for GB). This number of holdings represents a representative sample of up to 3% of the national herd in GB.

For n_h herds of approximately the same size, case clustering within herds can be tested for each cohort using a dispersion statistic:

$$\hat{r} \equiv \frac{s^2}{\bar{x}}(n_h - 1),$$

where $\bar{x}$ and s^2 are the sample mean and variance of the case incidence, respectively. The statistic $\hat{r}$ is approximately χ^2-distributed (with $n_h - 1$ d.f.) under the null hypothesis that cases are randomly distributed in herds.

Significant clustering was observed for highly aggregated infectivity (small c; i.e. large feed units), as shown in figure 7, because there is greater between-holding variability in exposure when infectivity enters holdings in large units. Given highly aggregated infectivity, the degree of clustering is secondarily sensitive to the relative infectiousness η, small values of η giving rise to less strong clustering.

The clustering pattern generated by (strong) infectivity aggregation in feed is compared with the observed data in figure 8*a*,*b*. The variance-to-mean ratios (estimated by $s^2/\bar{x}$) obtained for MA (figure 8*b*) and PMA (figure 8*a*) models with highly aggregated infectivity in feed, are shown as full lines. We first note that

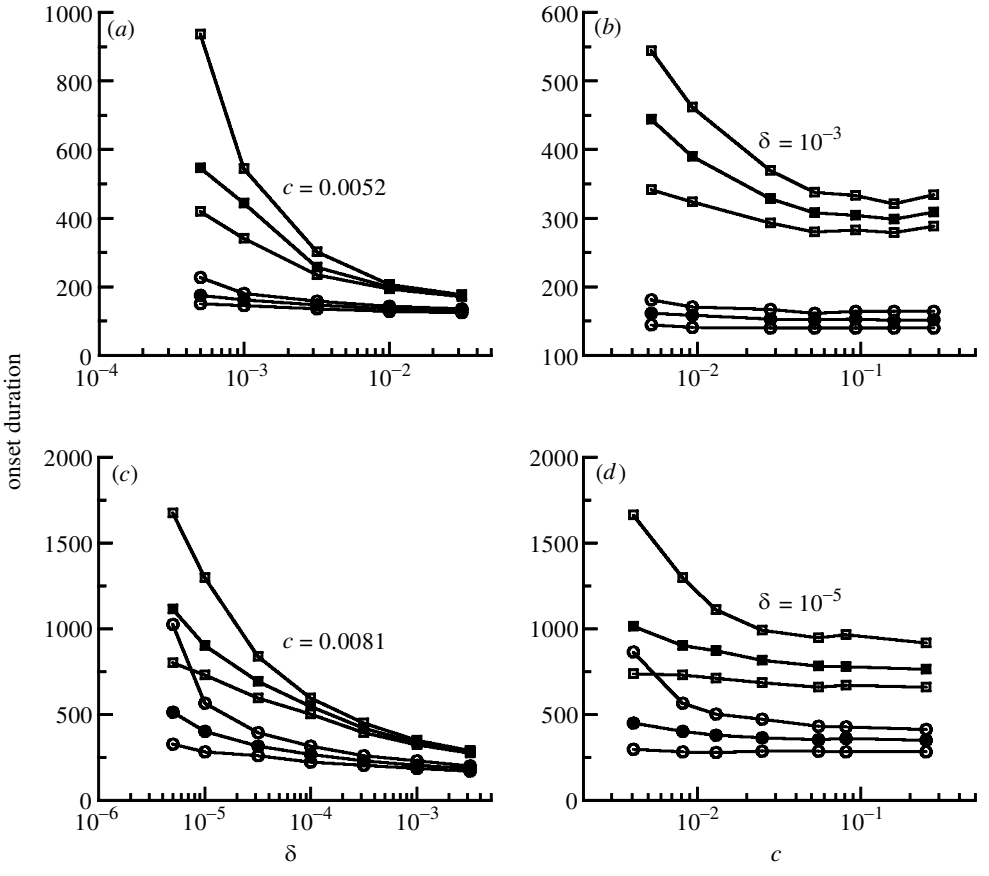

Figure 6. Mean and variance of the duration from the onset of external infectivity to the occurrence of the first (circles) and 100th (squares) case, as a function of the external infectivity δ ((*a*) and (*c*)) and the inverse infectivity aggregation scale c ((*b*) and (*d*)), calculated for a mass-action model from 200 stochastic realizations for each data point, in a system of a single feed producer and 240 herds with uniform feed demand d_i. Filled symbols, medians; open symbols, tenth and 90th percentiles. (*a*,*b*) Single event external infectivity source, parameter values: $\theta = 1.68$, $\eta = 0.52$, $\gamma_{MA} = 3.34$, $\zeta = 0.0072$, $b = 0.68$ ($R_0 = 6.24$, $T_g = 15.0$). (*c*, *d*) Continuous external infectivity source, parameter values: $\theta = 5.55$, $\eta = 0.20$, $\gamma_{MA} = 4.05$, $\zeta = 0.0017$, $b = 1.04$ ($R_0 = 7.5$, $T_g = 16.3$).

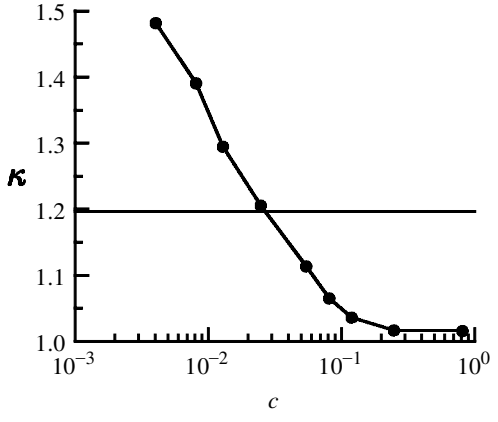

Figure 7. Mass action model results (for a continuous infectivity source) illustrating how the degree of clustering depends on c. Shown is $\mathcal{K}$ for a mean incidence of close to unity per herd in the smallest size category (1–29 animals) averaged over 200 realizations. Parameter values as in figure 6*d*. Above the dashed line $p < 0.05$.

the downward trend of $\mathcal{K}$ for high mean incidences is due to infection saturation (i.e. in all herds the number of infections approaches the full cohort size). Furthermore, it should be noted that the fall in $\mathcal{K}$ to unity for low incidence is a model artifact caused by the finite number of holdings modelled. If M holdings are simulated in a size category, then the smallest detectable incidence in any one realization is one case in all the holdings simulated, giving a lower per-holding incidence detectability threshold of $1/M$. In general, for a mean per holding incidence C in M holdings, there can be at most CM cases in one herd, and it is easy to demonstrate that this value also represents the absolute maximum detectable level of clustering, $\mathcal{K}$ for this incidence level. Thus, close to the detectability threshold, the degree of clustering is increasingly underestimated as mean incidence declines. Indeed, once this bias is allowed for, figure 8*a*,*b* is in fact consistent with a simple 'Poisson' model considered in Donnelly & Ferguson (1999), in which $\mathcal{K}$ is virtually constant for incidences C below those for which infection saturation occurs. This effect is in essence due to the fact that the number of infectious feed units encountered by a certain cohort in a herd of size $\mathcal{N}$ is Poisson distributed, and the average number $m_{\mathcal{N}}$ of cases arising from the consumption of a single infectious unit is essentially constant for these incidences. Because the case distribution will have a mean $m_{\mathcal{N}}$ times as large as the mean number of infectious units encountered and a variance $(m_{\mathcal{N}})^2$ times as large as the variance in the number of those units, thus $\mathcal{K}$ will equal $m_{\mathcal{N}}$, independent of incidence.

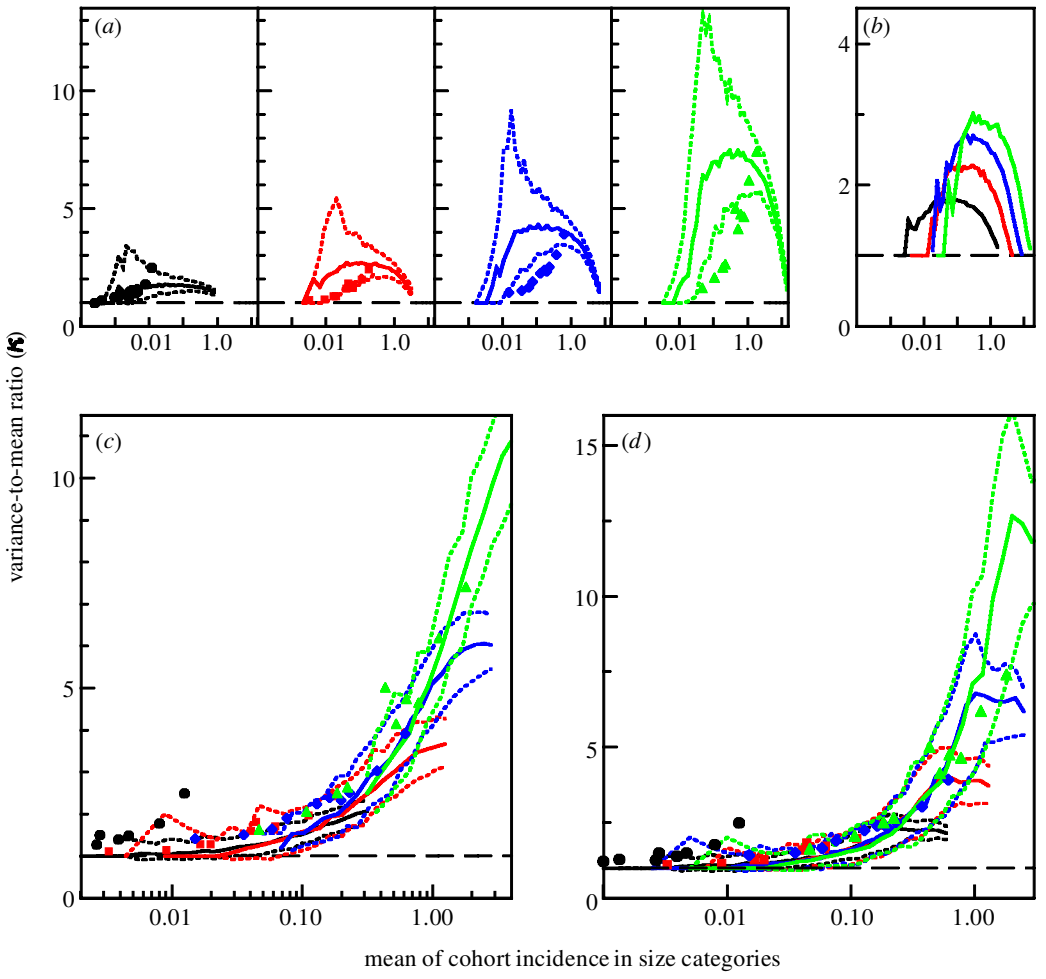

Figure 8. Within size category and cohort variance-to-mean ratios $\mathcal{K}$ as a function of the mean. Shown are median values of $\mathcal{K}$ (full lines) together with tenth and 90th percentiles (dotted lines) obtained from 300 stochastic realizations (cohorts with mean incidences falling within logarithmic spacings of 0.08 units were lumped together). Data for GB are shown as filled symbols for comparison. Colour coding: black, 1–29 adult animals; red, 30–49 adult animals; blue, 50–99 adult animals; green, 100+ adult animals. In (*a*) and (*b*) we show results for parameter combinations with significant clustering due to infectivity aggregation only (uniform d_i, and a single feed producer). In (*c*) we consider clustering due only to herd-level variation in per-animal feed uptake, and in (*d*) clustering due only to variation in infectiousness of feed between different producers. (*a*) Pseudo-mass-action feed consumption, $\theta = 2.60$, $\eta = 0.86$, $\gamma_{\mathrm{PMA}} = 0.83$, $c = 0.0046$, $\zeta = 0.0091$, $b = 1.03$ ($R_0 = 5.7$, $T_{\mathrm{g}} = 14.8$); the system size is 3600 herds. (*b*) Mass action, $\theta = 5.55$, $\eta = 0.20$, $\gamma_{\mathrm{MA}} = 4.05$, $c = 0.0081$, $\zeta = 0.0017$, $b = 1.04$; the system size is 1800 herds. (*c*) The d_i were sampled from a Weibull distribution (CDF $(x) = (\lambda x)^{\kappa}$) with $\lambda = 1$ and $\kappa = 1/7$, truncated at $x = 20$. Pseudo mass action, $\theta = 1.16$, $\eta = 0.89$, $\gamma_{\mathrm{PMA}} = 0.023$, $c = 1.0$, $\zeta = 0.018$, $b = 2.36$; 1800 herds and a single feed producer. (*d*) Ten feed producers. For $m = 1, \ldots, 7$, η_m is sampled uniformly from the range $[0,0.43\eta]$, for $m = 8$ from the range $[0.4\eta,1.6\eta]$, for $m = 9$ from the range $[1.6\eta,3.4\eta]$, and for $m = 10$ from the range $[3.4\eta,6.6\eta]$, and renormalized such that the average over the ten producers equals η. Mass action, $\theta = 3.14$, $\eta = 0.13$, $\gamma_{\mathrm{MA}} = 3.26$, $c = 0.72$, $\zeta = 0.0106$, $b = 1.00$; 2400 herds.

For PMA, $m_N \propto N$, which explains the scaling with holding size of the maximum $\mathcal{K}$ observed for PMA in figure 8*a*. A similar but less pronounced scaling occurring for MA is due to the scaling with N of the number of animals that get old enough to show clinical signs. This limits both the maximum number of observed cases and the degree of case clustering. In fact, this limiting effect is also partly responsible for the scaling in the PMA model.

On the basis of the above results, we tentatively conclude that the mechanism of infectivity aggregation cannot reproduce the type of clustering seen in the data. For the higher incidences observed, it leads (both for MA and PMA) to a scaling of $\mathcal{K}$ not present in the data, and for small incidences it would (in the absence of a detectability threshold) give rise to a variance-to-mean ratio independent of the mean incidence, which is also at odds with observation.

In the remainder of this paper, we consider two more clustering-generating mechanisms. The first is to assume that the demand for MBM-containing feed varied between holdings, with the distribution of per-holding feed consumption being highly skewed. In studying the effects of this heterogeneity, we concentrate on the PMA formulation. It is easy to see that, within the MA model, variation in feed demand only cannot reproduce the clustering pattern observed in the BSE case data. This is because the infection hazard is defined in terms of the proportion of feed units that are infectious; hence, herds with a finite but different per-animal feed uptake will for

a given cohort experience on average the same infection hazard. As a result the case distributions conditional on at least one case will be similar to a Poisson distribution, unlike those observed (Donnelly *et al.* 1997*b*).

For PMA, between-holding variation in feed demand can generate clustering which closely resembles the observed pattern in the BSE epidemic. In figure 8*c* we show results for the case when the d_i are picked from a Weibull distribution with $\lambda = 1/6$, truncated at $d = 20$ to avoid unrealistically high per-animal feed consumption. Also the mean-variance relationship conditional on having at least one case (not shown) is closely similar to that observed in the case data.

The last form of exposure heterogeneity we consider is variability in the relative infectiousness (i.e. the degree to which infectiousness is removed from offal) of feed from different producers. We find that a highly skewed between-producer infectiousness distribution can generate clustering similar to the case data within both the MA and the PMA formulations. Figure 8*d* shows results for the MA assumption based on ten feed producers, where the distribution of the relative infectiousness of the ten feed products is aggregated. The observed $\mathcal{K}$ versus mean profile closely resembles that of the BSE case data.

5. CONCLUSIONS

We have constructed a model framework for studying the feed-borne transmission of BSE in cattle. We have modelled infection processes at the herd level in order to be able to explore mechanisms that generate case clustering within herds. Our analyses give insight into the possible mechanism that could underlie the observed strong clustering of BSE cases in GB. We have demonstrated that patterns of case clustering within herds similar to those observed can be generated both by an over-dispersed distribution of feed usage and by between-producer variability in the extent to which the infectivity in feed was reduced during feed production. It is likely that both these mechanisms contributed to the clustering seen in the BSE epidemic in GB. The clustering patterns produced by exposure heterogeneity due to infectivity aggregation in feed were found not to be consistent with the observed pattern. We presented two model variants, based on MA and PMA assumptions for the within-herd distribution of infection risk over individual animals. The PMA model exhibits a scaling of the per-animal incidence with herd size, as was observed in the case data for GB and in those for the Republic of Ireland.

T.J.H., C.A.D., A.C.G. and R.M.A. thank the Wellcome Trust and Ministry of Fisheries and Food (MAFF). N.M.F. thanks The Royal Society and MAFF for research grant support. We are grateful to John Wilesmith and other members of the BSE Research Team of the Epidemiology Department, Central Veterinary Laboratory, for providing information and support in the analysis of the main epidemiological database.

REFERENCES

Anderson, R. M. & May, R. M. 1991 *Infectious diseases of humans: dynamics and control.* Oxford University Press.

Anderson, R. M. (and 14 others) 1996 Transmission dynamics and epidemiology of BSE in British cattle. *Nature* **382**, 779–788.

Bruce, M. E., Will, R. G., Ironside, J. W., McConnell, I., Drummond, D., Suttie, A., McCardle, L., Chree, A., Hope, J., Birkett, C., Cousens, S., Fraser, H. & Bostock, C. J. 1997 Transmissions to mice indicate that 'new variant' CJD is caused by the BSE agent. *Nature* **389**, 498–501.

Calman, K. 1996 A new variant of CJD. Department of Health, 20th March, CEM/CMO/g6/1.

Collinge, J., Sidle, K. C. L., Meads, J., Ironside, J. & Hill, A. F. 1996 Molecular analysis of prion strain variation and the aetiology of 'new variant' CJD. *Nature* **383**, 685–690.

Curnow, R. N., Hodge, A. & Wilesmith, J. W. 1997 Analysis of the bovine spongiform encephalopathy maternal cohort study: the discordant case-control pairs. *Appl. Statist.* **46**, 345–349.

De Jong, M. C. M., Diekmann, O. & Heesterbeek, H. 1995 How does transmission of infection depend on population size? In *Epidemic models: their structure and relation to data* (ed. D. Mollison), pp. 84–96. Cambridge University Press.

De Koeijer, A., Schreuder, B., Heesterbeek, H., Oberthur, R., Wilesmith, J. & De Jong, M. C. M. 2000 BSE risk assessment by calculating the basic reproduction ratio for infection among cattle. (Submitted.)

Diekmann, O., Heesterbeek, J. A. P. & Metz, J. A. J. 1990 On the definition and the computation of the basic reproduction ratio R_0 in models for infectious diseases in heterogeneous populations. *J. Math. Biol.* **28**, 365–382.

Donnelly, C. A. & Ferguson, N. M. 1999 *Statistical aspects of BSE and vCJD; models for epidemics.* Monographs on Statistics and Applied Probability, no. 84. London: Chapman & Hall/CRC.

Donnelly, C. A., Ferguson, N. M., Ghani, A. C., Wilesmith, J. W. & Anderson, R. M. 1997*a* Analysis of dam–calf pairs of BSE cases: confirmations of a maternal risk enhancement. *Proc. R. Soc. Lond.* B **264**, 1647–1656.

Donnelly, C. A., Ferguson, N. M., Ghani, A. C., Woolhouse, M. E. J., Watt, C. J. & Anderson, R. M. 1997*b* The epidemiology of BSE in GB cattle herds. I. Epidemiological processes, demography of cattle and approaches to control by culling. *Phil. Trans. R. Soc. Lond.* B **352**, 781–801.

Donnelly, C. A., Ghani, A., Ferguson, N. M. & Anderson, R. M. 1997*c* Recent trends in the BSE epidemic. *Nature* **389**, 903.

Donnelly, C. A., Gore, S. M., Curnow, R. N. & Wilesmith, J. W. 1997*d* The bovine spongiform encephalopathy maternal cohort study: its purpose and findings. *Appl. Statist.* **46**, 299–304.

Donnelly, C. A., Ghani, A. C., Ferguson, N. M., Wilesmith, J. W. & Anderson, R. M. 1997*e* Analysis of the bovine spongiform encephalopathy maternal cohort study: evidence for direct maternal transmission. *Appl. Statist.* **46**, 321–344.

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.

Ferguson, N. M., Donnelly, C. A., Woolhouse, M. E. J. & Anderson, R. M. 1999 Estimation of the basic reproduction number of BSE: the intensity of transmission in British cattle. *Proc. R. Soc. Lond.* B **266**, 23–32.

Gore, S. M., Gilks, W. R. & Wilesmith, J. W. 1997 Bovine spongiform encephalopathy maternal cohort study—exploratory analysis. *Appl. Statist.* **46**, 305–320.

Griffin, J. M., Collins, J. D., Nolan, J. P. & Weavers, E. D. 1997 Bovine spongiform encephalopathy in the Republic of Ireland: epidemiological observations 1989–1996. *Irish Vet. J.* **50**, 593–600.

Heesterbeek, J. A. P. & Dietz, K. 1996 The concept of R_0 in epidemic theory. *Statist. Neerlandica* **50**, 89–110.

Hill, A. F., Desbruslais, M., Joiner, S., Sidle, K. C. L., Gowland, I., Collinge, J., Doey, L. J. & Lantos, P. 1997

The same prion strain causes vCJD and BSE. *Nature* **389**, 448–450.

McKay, M. D., Beckman, R. J. & Conover, W. J. 1979 A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics* **21**, 239–245.

Stein, M. 1987 Large sample properties of simulations using Latin hypercube sampling. *Technometrics* **29**, 239–245.

Wells, G. A. H., Scott, A. C., Johnson, C. T., Gunning, R. F., Hancock, R. D., Jeffrey, M., Dawson, M. & Bradley, R. 1987 A novel progressive spongiform encephalopathy in cattle. *Vet. Rec.* **121**, 419–420.

Wilesmith, J. W., Wells, G. A. H., Ryan, J. B. M., Gavier-Widen, D. & Simmons, M. M. 1997 A cohort study to examine maternally-associated risk factors for bovine spongiform encephalopathy. *Vet. Rec.* **141**, 239–243.

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