



ELSEVIER

Available online at www.sciencedirect.com

SCIENCE @ DIRECT®

Preventive Veterinary Medicine 74 (2006) 180–193

PREVENTIVE
VETERINARY
MEDICINE

www.elsevier.com/locate/prevetmed

A simulation model for the study of the within-animal infection dynamics of *E. coli* O157

J.C. Wood^a, I.J. McKendrick^{a,*}, G. Gettinby^b

^a *Biomathematics & Statistics Scotland, King's Buildings, Edinburgh EH9 3JZ, United Kingdom*

^b *Department of Statistics and Modelling Science, University of Strathclyde, Glasgow G1 1XH, United Kingdom*

Received 11 August 2004; received in revised form 3 November 2005; accepted 23 November 2005

Abstract

Escherichia coli O157 can cause serious illness, even death, in humans. There is some consensus that the main reservoirs of this harmful bacterium are the rumens and intestines of cattle. Hence, a stochastic model of the bovine gut was developed to investigate the in vivo population dynamics of *E. coli* O157. Because bacterial numbers can reach minimal levels, a stochastic system was considered, with a birth–death process being used to represent bacterial growth and decay dynamics throughout the gastrointestinal tract. Reinfection through ingestion of bacteria present in the environment was allowed to occur and the required clustered distribution of inter-event times was implemented through the use of a random hazard doubly stochastic Poisson process. Due to the inclusion of multiple compartments, a feedback mechanism and an interest in the non-equilibrium dynamics of the process, it was not possible to obtain an analytical representation of the process and therefore, a simulation study was used to obtain results. The within-animal model can be used to explore the efficacy of control measures which act at an individual animal level.

© 2005 Elsevier B.V. All rights reserved.

Keywords: Mathematical model; Gastrointestinal tract; Poisson process; Birth–death process; *Escherichia coli* O157

DOI of related article: [10.1016/j.prevetmed.2005.11.010](https://doi.org/10.1016/j.prevetmed.2005.11.010).

* Corresponding author. Tel.: +44 131 650 4894; fax: +44 131 650 4901.

E-mail address: iajn@bioss.ac.uk (I.J. McKendrick).

0167-5877/\$ – see front matter © 2005 Elsevier B.V. All rights reserved.

doi:[10.1016/j.prevetmed.2005.11.011](https://doi.org/10.1016/j.prevetmed.2005.11.011)

1. Introduction

Human cases of *Escherichia coli* O157 infection are relatively rare. However the consequences can be severe since a significant number of these patients either die or develop long-term complications. One method which could help to decrease the incidence of such infections is to reduce the prevalence of the organism at the points at which people come into contact with it. Since many early outbreaks were associated with the consumption of hamburgers (Bell et al., 1994) and a large volume of research has revealed that direct or indirect contact with animal faeces is a major risk factor (Parry et al., 1998; Reilly et al., 2000), it is generally accepted that the main source of *E. coli* O157 is in the rumens and intestines of cattle (Chapman et al., 1992; Chapman and Ackroyd, 1997). While Wood et al. (2006b) have considered the infection dynamics at a herd level, a greater understanding of the within-animal infection dynamics could ensure that the most appropriate control strategies are implemented. Therefore this paper concentrates on the infection dynamics at an individual animal level. Experimental studies have been carried out to improve our knowledge of the infection process, but they are often restricted by financial and/or ethical considerations. Mathematical modelling can be used as an alternative approach to gain further insight into a system (Renshaw, 1991). Hence, this paper presents a simplified model of the passage of *E. coli* O157 through the bovine gastrointestinal tract. Despite the simplified nature of the model, it was not possible to derive analytical results, necessitating the use of simulation techniques. A program was written using the C++ programming language, which keeps track of the bacterial population sizes at key sites along the gut over time. The model was designed to enable the investigation of the in vivo population dynamics, with a flexible structure which can allow the efficacy of potential control strategies to be assessed.

2. Model development

The lack of knowledge of the underlying biology and the absence of similar models of bacterial populations in the literature meant that, in order to produce a within-animal model, it was necessary to make several assumptions about how the bacteria behave inside an animal.

Data obtained from in vivo experimental studies have shown that bacterial numbers can reach minimal levels in the gastrointestinal tract (Cray and Moon, 1995; Brown et al., 1997; Woodward et al., 1999), and therefore a stochastic model was assumed most appropriate. Ruminant digestion literature indicated that a compartmentalised model would best capture the variation in the bacterial populations throughout the gut (Matis and Hartley, 1971; France et al., 1985) and so experimental data were used to help determine the number of compartments needed to represent the bovine gut. As the data demonstrated notable differences in size between the populations located in the rumen and abomasum contents and in faeces collected from inside the animal, a stochastic three-compartment model was adopted to represent the within-animal *E. coli* O157 population dynamics. Fig. 1 shows a diagrammatic representation of this model which attempted to capture the nature of the system whilst minimising the number of compartments, where X_1 , X_2 , and X_3 represent the rumen, abomasum and faeces, respectively, of an adult cow.

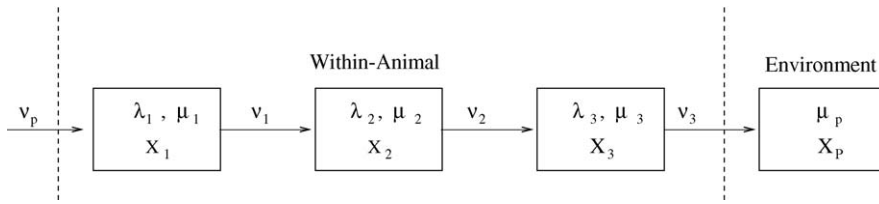


Fig. 1. Compartmental representation of within-cow bacterial dynamics, with X_1 = rumen, X_2 = abomasum, X_3 = faeces and X_p = environmental pool.

A simple stochastic linear birth–death process was used to represent the bacterial growth and decay that occurs within each site, with each compartment possessing its own site-specific set of birth and death parameters, denoted by λ and μ , respectively (Feller, 1939).

The within-animal model also allowed migration at rate v_i , $i = 1, 2, 3, p$, of individuals in one direction only, i.e. to the right, through the gastrointestinal tract, connecting the processes operating in neighbouring compartments. The migration of bacteria caused the hazard function for births and deaths in each compartment to be piecewise constant. Therefore, the distribution of times between migration events were conditionally exponential, with the memoryless property being retained.

The model, as illustrated in Fig. 1, was linked to an environmental pool compartment. Longitudinal studies suggest that infection with *E. coli* O157 is a fairly transient event, yet somehow the infection has the ability to persist within herds for at least several months. It is likely that the existence of an environmental pool of bacteria is an important factor in the maintenance of infection, since animals can inadvertently ingest bacteria whilst grazing and hence reinfest themselves (Harmon et al., 1999). As a result, it was felt necessary to include this feedback mechanism in the model. Hence, the bacteria shed from the third and final within-animal compartment entered an environmental pool consisting of all *E. coli* O157 living outwith the animal. The bacteria inhabiting this pool were subjected to a stochastic death process at a constant rate, and hence exhibited exponential decay.

Initially, to simplify matters mathematically, immigration from the environmental pool back into the first compartment of the animal was assumed to occur at a rate v_p , with these events being conditionally exponentially-distributed. However, despite developing a simple model, it was not possible to derive closed-form solutions since the presence of multiple compartments and the feedback loop resulted in an analytically intractable model given that, due to the transiency of typical infection, the primary interest is in the non-equilibrium states of the system. Therefore, it was necessary to resort to simulation techniques. Given the need for a simulation-based approach, the opportunity was taken to improve the biological realism of certain aspects of the model.

2.1. Clustered ingestion process

One such aspect was the use of the conditionally exponential distribution to determine the length of time between ingestion events, i.e. modelling the rate at which infection

occurs through the ingestion of environmental bacteria. The assumption of a memoryless process is a staple of simple mathematical models. However, given the environmental distribution of faecal pats and the faecal-avoidance behaviour of cattle, a more clustered distribution of inter-event times, the rate of which also depended on the total amount of environmental bacteria, seemed more realistic.

The clustered distribution was implemented through the use of a random hazard doubly-stochastic Poisson process, in which a two-state Markov chain, reflecting the infection status of the animal location, replaced the fixed intensity parameter ν_p (Cox and Isham, 1980; Taylor and Karlin, 1998). Hence, in this application, as an animal moved between areas infected with *E. coli* 0157 and areas that remained uninfected, the process Λ alternated between being ‘INFECTED’ and ‘CLEAN’. When the underlying intensity was ‘INFECTED’, ingestion events occurred according to a Poisson process of intensity λ , and when the underlying intensity was ‘CLEAN’, no events occurred.

The two-state Markov chain was defined as follows:

$$\begin{aligned} P_1 &= Pr(\Lambda(t + \Delta t) = 1 | \Lambda(t) = 0) = \alpha \Delta t \\ P_0 &= Pr(\Lambda(t + \Delta t) = 0 | \Lambda(t) = 1) = \beta \Delta t \end{aligned} \quad (1)$$

where $\Lambda = 0$ represented an animal foraging in a clean, uninfected area, whilst $\Lambda = 1$ represented an animal present in an infected area. Hence, $\alpha \Delta t$ is the transition probability of an animal moving into an infected area and $\beta \Delta t$ is the probability of an animal moving out of an infected area. The rate of change of the probability of being in an infected area can be written as:

$$\frac{dP_1(t)}{dt} = \alpha P_0(t) - \beta P_1(t),$$

which can be solved to give:

$$P_1 = \frac{\alpha}{\alpha + \beta} \quad (\equiv P_C).$$

The variable P_C , which is the proportion of the field that is contaminated with infected faeces, was introduced as a means of incorporating dependency on the size of the environmental pool of bacteria into the transition probabilities (1). Firstly, it was assumed that the probability of an animal being in state $\Lambda = 1$, i.e. an infected area, was equivalent to P_C . Although cattle do have a tendency to avoid pasture contaminated with faeces during grazing, they may be attracted to the rich, mature forage that grows up from around faecal deposits (Phillips, 1993). Alternatively, as clean pasture becomes depleted, cattle will increasingly be obliged to graze near infected forage. The assumption seems a reasonable first approximation given that it is most valid when the faecal (and hence the bacterial) load is large and thus ingestion events are more likely.

2.2. Pool size-dependent transition probabilities

A function for P_C , in terms of the environmental pool size, had to be obtained in order to derive pool size-dependent transition probabilities for the two-state Markov chain. The

spatial distribution of animal excreta in the field has been found to be well-described by the negative binomial distribution since it allows for both the overlapping of cowpats and the tendency for an animal to defecate at a particular point to vary from one area of a pasture to another: e.g. increased amounts of defecation may occur in areas where animals gather to rest and ruminate (Petersen et al., 1956; Richards and Wolton, 1976; Morton and Baird, 1990). Hence, an appropriate function for the proportion of pasture surface covered by at least one faecal pat, P_C , was found using the negative binomial function in Petersen et al. (1956):

$$P_C = 1 - \frac{1}{[(D + K)/K]^K}, \quad (2)$$

where D represented the mean excreta density and the spatial distribution of excreta was summarised by an aggregation parameter, K .

The notional maximum area of infected faeces in the environment in square centimetres, F , was calculated by assuming that each bacterium passing out of an animal inside a faecal pat was surrounded by a very small, arbitrarily-chosen, quantity of faeces, δ . The area that this quantity δ covers on the field was calculated using measurements for a typical cowpat (Haynes and Williams, 1993), and added to the running total for F , which was divided by the area of the pasture to obtain the density, D .

Fig. 2 illustrates a practical problem that can arise when using this method to calculate the notional maximum infected area, F , within the simulation model. It is possible that another bacterial migration event could take place within the time-interval required for the faecal layer, δ , surrounding the first migrant bacteria to pass out of the animal. In such an instance, as the bacteria are very close together, their surrounding faecal layers would overlap. In order to account for such clustering, the overlapping quantity was not double-counted and instead the contribution to F was reduced proportionately.

The death of bacteria in the environment was modelled by reducing the area of infected faeces by a proportionate amount since, although the faeces might not have actually disappeared, the quantity of faeces surrounding the dead bacteria was no longer considered a threat to cattle in terms of reinfection.

The second parameter, K , which provided a measure of the excretal patch aggregation, varied as a function of the stocking density. For instance, at higher stock densities cattle have less opportunity to group together when, for example, resting, which results in a more uniform distribution of faecal pats over the field and thus a higher value of K . Given that the

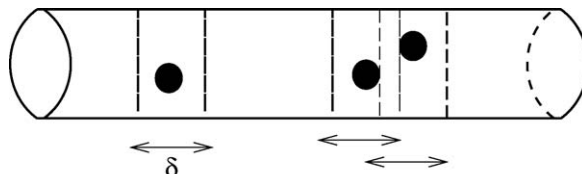


Fig. 2. Diagram of bacteria passing through the gut, where the tube represents the gut and solid circle enclosed by dotted lines represents a bacterium surrounded by a small quantity of faeces, δ .

stocking rate on a typical dairy farm is roughly three cows per hectare (Haynes and Williams, 1993) and, for simplicity, assuming cattle graze on the same field over the pasturing season, a value of 8 was chosen for K , based on measurements obtained in Petersen et al. (1956) and Richards and Wolton (1976).

Since it was assumed earlier that $P_C \equiv P_1$, expressions can be derived for the transition probabilities, as follows:

$$\alpha = b \left(\left[\frac{(D + K)}{K} \right]^K - 1 \right), \quad \beta = b,$$

where b is a temporal scaling constant.

2.3. Simulating the ingestion process

When the process was in state 1, i.e the ‘INFECTED’ state, it was straightforward to obtain the time at which the process switched to state 0 by generating a random variable from a stationary Poisson process with parameter β (Law and Kelton, 1991; Renshaw, 1991). However, once in state 0, calculation of the time to the next return to state 1 required the generation of a random variable from a non-stationary Poisson process, since the transition probability, α , involved the environmental pool size, X_p , which changed over time. A thinning algorithm with an adjustable maximum value was used to improve the simulation efficiency (Lewis and Shedler, 1979).

When the process was in the ‘INFECTED’ state a Poisson process was used to determine the length of time between ingestion events. Fig. 3 shows a graphical illustration of the entire ingestion process over the course of one realization. It can be seen that on

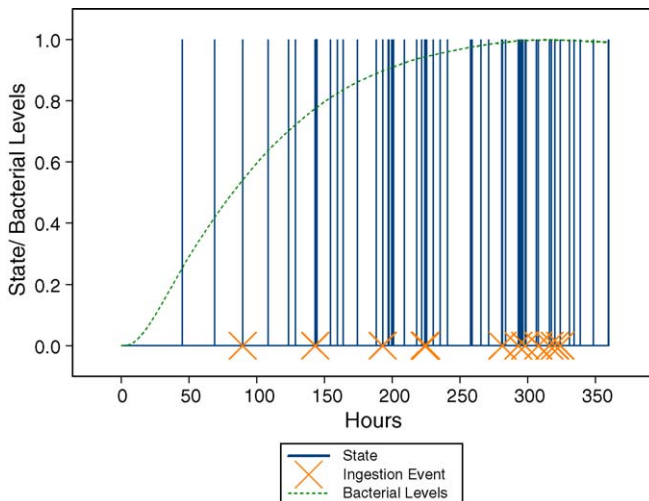


Fig. 3. Illustration of the ingestion process over the course of one realization. Solid lines represent the two-state Markov chain, which switched from state 0 to state 1, and crosses represent ingestion events. The dotted line gives an indication of the quantity of bacteria in the environment and is scaled from 0 to 1.

average, the Markov chain spent more time in state 1 when higher bacterial levels were present, and consequently more ingestion events occurred.

The actual dose of *E. coli* O157 ingested at each ingestion event was based on the average number of bacteria likely to be found within a mouthful of food. Given the aversion of animals to their own faeces, it was unlikely that the ingesta resulting from one bite would solely consist of faeces. Instead, the proportion of faeces inadvertently swallowed was much more likely to be skewed towards 0. Thus, the current average number of bacteria within a bite of size 12 cm² (Hutchings and Gordon, 2001) was calculated and this maximum value was then multiplied by a proportion sampled from the positively-skewed, but arbitrarily chosen, *Beta* (1, 10) distribution.

Finally, a mechanism was incorporated to ensure that it was not possible for ingestion events to occur when the amount of environmental bacteria was very low; although the likelihood of simulating such an event was very small, it was felt necessary to eliminate any possibility of it occurring since observations of cattle reveal that they will only graze in the vicinity of spoiled forage once there is a build-up of faeces over the whole pasture surface (Hutchings et al., 1999).

3. Parameterisation

After the initial arrival into the environment, it was assumed that bacteria were only subject to a death process, which eliminated the need for birth and migration parameters relating to the environmental pool. Survival rates of *E. coli* O157 in cattle faeces, manure, slurry and soil are well documented (Wang et al., 1996; Maule, 2000; Himathingham et al., 1999; Ogden et al., 2001). A moderate mean survival length of just under 10 weeks was selected from the reported range, which corresponded to a death rate in the environmental pool of 0.0006 per hour (h⁻¹).

However, estimation of the within-animal birth, death and migration parameters was less straightforward. Only published in vivo bacterial counts obtained from experimental studies were available for parameterisation purposes. The summarised results in Cray and Moon (1995) proved to be most informative for this exercise, although only eight data points could be extracted (Table 1) and therefore the more usual data-intensive methods of parameterisation could not be utilised. Instead, an approach was developed, intermediate between the use of a fully stochastic method and a purely deterministic method. This involved obtaining the first and second moments for the population size in each compartment, assuming that reinfection could not occur in the system, using the method outlined in Grenfell et al. (1995). The absence of immigration ensured that the solutions of

Table 1
Recovery of *E. coli* O157 from adult cattle inoculated with the organism 2–4 days previously: adapted from Cray and Moon (1995)

Compartment	Observed values (cfu/g)
Rumen	{80, 30, ≤ 20}
Faeces	{150, 512820, 13000000}
Abomasum	{320, ≤ 20}

the resulting equations were tractable and seemed a sensible simplification to the process for parameterisation purposes, given that the observed data were obtained from experimental studies where animals were inoculated with a dose of *E. coli* O157 and kept in isolation in order to investigate the effects of a single pulse of infection. The moments were evaluated at a time corresponding to the time elapsed between the reported inoculation of the animals with *E. coli* O157 and the collection of the sample from which the observed count was calculated in the experimental study. It was assumed that the counts were normally distributed, which seems reasonable since the observed similarity of the mean and variance for each compartment (not shown) is indicative of a Poisson distribution, which can be approximated by the normal distribution for large mean values. These counts were compared against rescaled moments for different sets of parameters, with the best combination of parameters being determined by maximisation of the log-likelihood. Rescaling of the moments was necessary to ensure that these quantities were directly comparable to the observed counts (measured in colony-forming units (cfu) per gram), as the original moments derived from the model related to the number of bacterial cfu per compartment.

As there were more parameters requiring estimation than observations, it was appropriate to fix the birth rates since there was information available on the growth rates of *E. coli* O157 under different pH conditions in vitro (Rasmussen et al., 1993; Duncan et al., 1999, 2000). Therefore, by assuming that the birth rate in this model represented the rate of growth under optimal conditions, while the death rate incorporated the deviations from optimal conditions that exist in vivo, the birth rates were estimated from the published literature, leaving only the death and migration parameters requiring estimation.

However, during the estimation of these parameters it was notable that the observed counts obtained from the faeces were so overdispersed that it proved impossible to find a set of parameters that generated a distribution from which it was plausible that all three counts could originate. This effect could be caused by clustering of bacteria in ingesta, but in a similar fashion to the Grenfell et al. (1995) model of parasite aggregation which failed to predict enough variability to match that observed in practice, the extremely large variation in the observed counts suggested that the model was missing a significant source of heterogeneity. It seems likely that this extra variation is due to heterogeneity between animals (for instance, genetic susceptibility or other density-independent variations in establishment of infection between animals), which is consistent with observations that a small proportion of animals appear to shed the majority of organisms (Omisakin et al., 2003).

Due to insufficient data, the only plausible solution to the problem of parameterising the third compartment was to add a large amount of excess variability to the variance obtained from the second moment equation, since it was known that an additional source of heterogeneity must exist even if the source was uncertain. As long as the additional variance component exceeded the contribution resulting from the relevant moment equation, a distribution could be found for which the three observed counts were plausible realizations. However, increasing the variance term by the required amount resulted in the parameterisation of the third compartment becoming essentially a deterministic process since the variance was so large that only the mean contributed useful information to the likelihood.

Table 2

Combinations of parameters that resulted in the maximum log-likelihood values

Parameter set	μ_{x_1}	μ_{x_2}	μ_{x_3}	ν_{x_1}	ν_{x_2}	ν_{x_3}	log-likelihood
A	0.48	0.001	0.76	0.01	0.14	0.05	-117.872
B	0.48	0.001	0.75	0.01	0.14	0.06	-117.872
C	0.48	0.001	0.74	0.01	0.14	0.07	-117.872
D	0.48	0.001	0.73	0.01	0.14	0.08	-117.872
E	0.48	0.001	0.72	0.01	0.14	0.09	-117.872
F	0.48	0.001	0.71	0.01	0.14	0.10	-117.872

4. Results

Table 2 details the sets of parameters that resulted in the maximum observed value of the log-likelihood. There is no criterion to select between these sets of parameters as their respective values for the log-likelihoods were identical, measured to three decimal places. Hence, it was decided to use a set of parameters towards the middle of the range of parameters detailed in Table 2, since this set can be considered representative. Hence, the parameters in set D were chosen and substituted into the simulation model. Further investigation confirmed that the simulation results were not sensitive to the choice of parameter set from these combination options.

The simulation model was initialised assuming an initial dose of 10^4 cfu of infection, which is typical of the magnitudes of infection found in the pastoral environment (Shere et al., 1998; Himathingkham et al., 1999), and Fig. 4 shows the resulting bacterial population sizes in each compartment plotted over time, averaged over 100 runs. The first graph (Fig. 4(a)) plots the mean population sizes for just the first 2 weeks following infection. As the bacteria passed through the animal, the population in each compartment in turn reached a peak. Importantly, the population size in the third (faeces) compartment exceeded that in the abomasum compartment for the majority of time, which makes

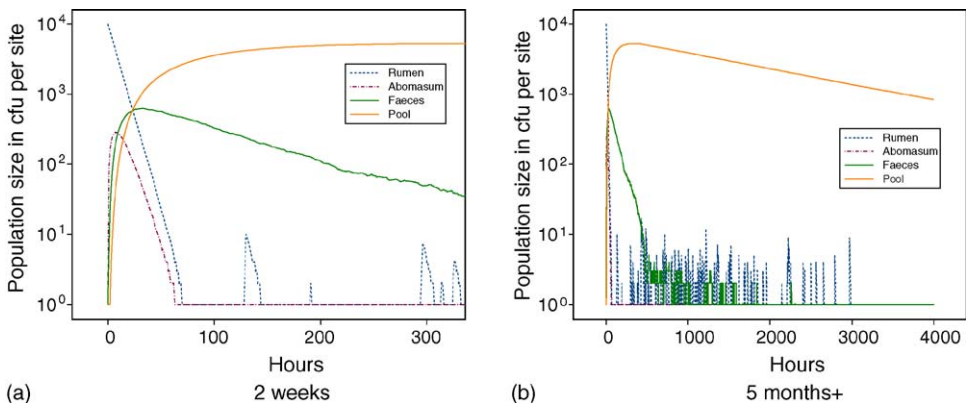


Fig. 4. Bacterial population sizes over time after an initial dose of 10^4 cfu, averaged over 100 runs, for the rumen, abomasums, faeces and environmental pool compartments: (a) 2 weeks and (b) more than 5 months.

Table 3
Summary statistics of ingestion events occurring in individual realizations

Number of runs	Minimum	Mean	Median	Maximum	S.D.
100	0	1.7	1	9	2.07

biological sense (given the acidic conditions in the latter which should present less favourable conditions) and matches experimental observations (Cray and Moon, 1995). The amount of bacteria being shed by the animal into the environmental pool resulted in ingestion events occurring despite the moderately low initial dose used. The ingestion events can be identified by the spikes in the rumen compartment, the first one occurring approximately 130 h after the initial infection. Note that ingestion events occurred randomly in time, unlike the initial inoculation, and hence each of these peaks reflected a much larger ingestion occurring in only one realization, averaged over all runs.

Fig. 4(b) shows the same plot over a longer timescale. From this graph it can be seen that on average, over 100 realizations, the animal shed continually for up to a maximum of 46 days following initial infection, as this was the first point at which the mean population size in the faeces compartment reached 0. This agrees with Rahn et al. (1997) who found the duration of shedding by naturally-infected animals to be less than 2 months (see also Besser et al. (1997)) and is just outside the range of shedding observed in Sanderson et al. (1999), who found that orally inoculated calves shed for a mean period of 30 days, with a range of 20–43 days. However, these animals were kept in isolation facilities and hence enjoyed a considerably lower risk of reinfection. In addition, the plot shows that once the animal was initially cleared of infection, there was a high risk of intermittent shedding for another month and a half (90 days post-initial infection) solely as a consequence of reinfection via ingestion events.

Examination of the individual realizations that contributed to the average population sizes revealed that in 44% of the runs an ingestion event did not occur. Furthermore, it can be seen from Table 3, which provides summary statistics of the number of ingestion events from these realizations, that if ingestion events did take place within a run, there were only likely to be one or two such events. Hence, despite the flurry of events suggested by the mean results (Fig. 4(b)), an ingestion event was actually a rare event for any particular animal over the course of the 5-month period considered. Nevertheless, the longer-term results show that the effect of just a few ingestion events can be considerable in maintaining the infection within an individual, justifying the inclusion of the improved ingestion process discussed in Section 2.1, while individual ingestion events will not be rare in a group of animals housed together, each of which may independently ingest bacteria.

In order to investigate the behaviour of the system using the same set of parameters at a higher initial dose of infection, the simulation study was repeated using a dose of 10^7 cfu. The results obtained from just one realization are plotted in Fig. 5 since the computer run time to produce average results over a large number of runs proved to be prohibitive due to the large population sizes involved.

It can be seen that while the infection was highly transient in the first and second within-animal compartments (rumen and abomasum), the infection persisted within the faeces for much longer periods of time. In fact, in the absence of any ingestion events and assuming

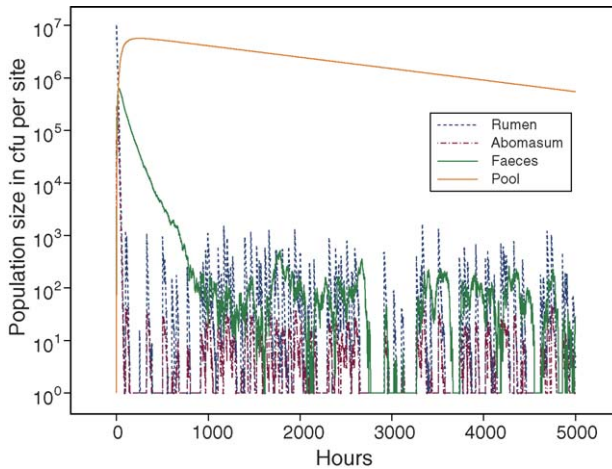


Fig. 5. Bacterial population sizes over time with an initial dose of 10^7 cfu, plotted for 5000 h, for the rumen, abomasums, faeces and environmental pool compartments.

that the decay within this compartment would continue at the initial rate, it could be predicted that the faeces would remain infected, and hence capable of continuing to contaminate the environment, for at least 2 months. However, the many ingestion events that occurred in this realization ensured that the infection persisted within-animal for a much longer length of time. While the results obtained from just one realization should not be assumed to typify a process, further investigation did reveal that other realizations generated similar results.

A closer examination of this realization reveals that over the 6 months plotted in Fig. 5, the model predicted that an animal would come into the vicinity of infected faeces on 3164 occasions (corresponding to $\Lambda = 1$ in the two-state Markov chain defining the ingestion process in Section 2.1). An ingestion event occurred on only 107 of these occasions. However, it is clear that increasing the initial dose from 10^4 to 10^7 cfu resulted in a massive increase in the risk of ingestion events.

5. Discussion

A simulation model of the bovine gut was developed to investigate the passage of *E. coli* O157, based on existing knowledge. The model provides an alternative to carrying out further large-scale experimental studies, which may be restricted by cost and ethical considerations. In addition, experimentally-infected animals might not provide an accurate representation of the in vivo infection dynamics in naturally-infected animals subject to reinfection. The simulation model was designed to also incorporate the assessment of control strategies which act at an individual animal level. It is likely that this model could be linked to others, perhaps forming part of a quantitative risk assessment exercise.

However, the model was constrained by the limits of our current knowledge and although it can be used to provide predictions about the system, its limitations should be

acknowledged. For instance, the model failed to incorporate seasonality effects and the daily variability in shedding levels that has been observed (Chapman and Ackroyd, 1997; Grau et al., 1969). While it would be relatively straightforward to model these effects, for example, by altering parameter values, this would not help to shed light on the underlying mechanisms of the processes. Therefore, since the causal mechanisms for their occurrence are not yet fully understood, these aspects of the epidemiology were not incorporated.

Nevertheless, it is felt that the model does capture as much of the biological essence of the process as is possible given current understanding, and that this knowledge was sufficient to generate sensible and valid results. This point can be demonstrated when an initial dose of a similar magnitude to that likely to be found in the environment of farms with persistent infection is assumed, since the nature and duration of shedding output from the simulation model matches observations. However, for higher initial doses of infection which exceed those levels, the results are more extreme. Housed animals might be expected to be exposed to a higher environmental dose of infection than animals at pasture, and the prevalence of shedding in these animals has indeed been observed to be statistically significantly higher (Synge et al., 2000). Also, the distribution of within-herd shedding prevalences has been observed to be highly over-dispersed, with a small percentage of groups of animals exhibiting very high prevalences (Gunn et al., 2004). Hence, the results seen in Fig. 5 may not be untypical of certain farms with a high environmental load.

It is inevitable that as understanding of the biological processes improves, certain aspects of the model will require modification. However, we anticipate that such alterations will be relatively straightforward to carry out due to the design of the model. For instance, it was found that if colon contents data were used instead of faecal data to parameterise the third within-animal compartment, the duration of shedding following infection and the amount of bacteria being shed into the environment both dropped to unrealistically low levels. This difference in behaviour exhibited by the two simulation studies is consistent with the existence of some region in the lower gastrointestinal tract which *E. coli* O157 is either able to exploit to achieve a very high growth rate, or is able to colonise. Recent experimental studies indicate that such a region does exist at the terminal rectum (Naylor et al., 2003). As understanding of this biological process improves, it will be relatively straightforward to modify the current model. For instance, colonization could be incorporated through the addition of a notional extra compartment, into which a fraction of the bacteria could migrate and be subjected to an appropriate growth rate, with a low migration rate back to the main compartment. Such a model would, however, require the acceptance of the existence of a single ‘colonization’ compartment, and would potentially be difficult to parameterise.

In conclusion, the model described in this paper formed a reasonable starting point in the use of mathematical modelling as a tool to elucidate the in vivo infection dynamics of *E. coli* O157 upon which further development can be based as understanding of the processes involved increases. This model will allow the efficacy of potential control strategies to be assessed without facing the cost and ethical concerns raised by experimental studies. A subsequent paper will use the model to explore such control measures (Wood et al., 2006a).

Acknowledgements

JCW gratefully acknowledges the support of a BBSRC studentship. IJM acknowledges support from the Scottish Executive Environment and Rural Affairs Department, project BSS/028/99 and the Wellcome Trust International Partnership Research Award for Veterinary Epidemiology.

References

- Bell, B., Goldoft, M., Griffin, P., Davis, M., Gordon, D., Tarr, P., et al., 1994. A multistate outbreak of *E. coli* O157:H7-associated bloody diarrhea and hemolytic uremic syndrome from hamburgers: the Washington experience. *J. Am. Med. Assoc.* 272, 1349–1353.
- Besser, T., Hancock, D., Pritchett, L., McRae, E., Rice, D., Tarr, P., 1997. Duration of detection of fecal excretion of *Escherichia coli* O157:H7 in cattle. *J. Infect. Dis.* 175, 726–729.
- Brown, C., Harmon, B., Zhao, T., Doyle, M., 1997. Experimental *Escherichia coli* O157:H7 carriage in calves. *Appl. Environ. Microbiol.* 63, 27–32.
- Chapman, P., Ackroyd, H., 1997. Farmed deer as a potential source of verocytotoxin-producing *Escherichia coli* O157. *Vet. Rec.* 141, 315–316.
- Chapman, P., Siddons, C., Wright, D., Norman, P., Fox, J., Crick, E., 1992. Cattle as a source of verotoxigenic *Escherichia coli* O157. *Vet. Rec.* 132, 323–324.
- Cox, D., Isham, V., 1980. Point Processes: No. 12 in Monographs on Applied Probability and Statistics. Chapman and Hall, London.
- Cray, W., Moon Jr., H., 1995. Experimental infection of calves and adult cattle with *Escherichia coli* O157:H7. *Appl. Environ. Microbiol.* 61, 1586–1590.
- Duncan, S., Booth, I., Flint, H., Stewart, C., 2000. The potential for the control of *E. coli* O157 in farm animals. *Symp. Ser. Soc. Appl. Microbiol.* 88, 157S–165S.
- Duncan, S., Scott, K., Flint, H., Stewart, C., 1999. Commensal–pathogen interactions involving *E. coli* O157 and the prospects for control. In: *E. coli* O157 in Farm Animals. CABI Publishing, Oxon (Chapter 6).
- Feller, W., 1939. Die grundlagen der volterraschen theorie des kampfes ums dasein in wahrscheinlichkeitstheoretischer behandlung. *Acta Biotheor.* 5, 11–40.
- France, J., Thornley, J., Dhanoa, M., Siddons, R., 1985. On the mathematics of digesta flow kinetics. *J. Theor. Biol.* 113, 743–758.
- Grau, F., Brownlie, L., Smith, M., 1969. Effects of food intake on numbers of *Salmonella* and *Escherichia coli* in rumen and faeces of sheep. *J. Appl. Bacteriol.* 32, 112–117.
- Grenfell, B., Wilson, K., Isham, V., Boyd, H., Dietz, K., 1995. Modelling patterns of parasite aggregation in natural populations: trichostrongylid nematode–ruminant interactions as a case study. *Parasitology* 111, S135–S151.
- Gunn, G., McKendrick, I., Synge, B., 2004. Personal communication.
- Harmon, B., Doyle, M., Brown, C., Zhao, T., Tkalcic, S., Mueller, E., Parks, A., Jacobsen, K., 1999. Faecal shedding and rumen proliferation of *E. coli* O157:H7 in calves: an experimental model. In: *E. coli* O157 in Farm Animals. CABI Publishing, Oxon (Chapter 5).
- Haynes, R., Williams, P., 1993. Nutrient cycling and soil fertility in the grazed pasture ecosystem. *Adv. Agron.* 49, 119–199.
- Himathingkham, A., Bahari, S., Riemann, H., Cliver, D., 1999. Survival of *E. coli* O157:H7 and *Salmonella typhimurium* in cow manure and cow manure slurry. *FEMS Microbiol. Lett.* 178, 251–257.
- Hutchings, M., Kyriazakis, I., Gordon, I., Jackson, F., 1999. Trade-offs between nutrient intake and faecal avoidance in herbivore foraging decisions: the effect of animal parasitic status, level of feeding motivation and sward nitrogen content. *J. Anim. Ecol.* 68, 310–323.
- Hutchings, N., Gordon, I., 2001. A dynamic model of herbivore–plant interactions on grasslands. *Ecol. Model.* 136, 209–222.
- Law, A., Kelton, W., 1991. Simulation Modeling and Analysis, second ed. McGraw-Hill Inc., London.

- Lewis, P., Shedler, G., 1979. Simulation of nonhomogenous poisson processes by thinning. *Naval Res. Logistics Q.* 26, 403–413.
- Matis, J., Hartley, H., 1971. Stochastic compartmental analysis: model and least squares estimation from time series data. *Biometrics* 27, 77–102.
- Maule, A., 2000. Survival of verocytotoxigenic *E. coli* O157 in soil, water and on surfaces. *J. Appl. Microbiol. Symp. Suppl.* 88, 71S–78S.
- Morton, J., Baird, D., 1990. Spatial distribution of dung patches under sheep grazing. *N. Z. J. Agric. Res.* 33, 285–294.
- Naylor, S., Low, C., Besser, T., Mahajan, A., Gunn, G., Pearce, M., McKendrick, I., Smith, D., Gally, D., 2003. Lymphoid follicle-dense mucosa at the terminal rectum is the principal site of colonization of enterohemorrhagic *E. coli* O157:H7 in the bovine host. *Infect. Immun.* 71, 1505–1512.
- Ogden, I., Fenlon, D., Vinten, A., Lewis, D., 2001. The fate of *E. coli* O157 in soil and its potential to contaminate drinking water. *Int. J. Food Microbiol.* 66, 111–117.
- Omisakin, F., MacRae, M., Ogden, I., Strachan, N., 2003. Concentration and prevalence of *E. coli* O157 in cattle feces at slaughter. *Appl. Environ. Microbiol.* 69 (5), 2444–2447.
- Parry, S., Salmon, R., Willshaw, G., Cheasty, T., 1998. Risk factors for and prevention of sporadic infections with verocytotoxin (shigatoxin) producing *E. coli* O157. *Lancet* 351, 1019–1022.
- Petersen, R., Lucas, H., Woodhouse, W., 1956. The distribution of excreta by freely grazing cattle and its effect on pasture fertility. Part I: excretal distribution. *Agron. J.* 48, 440–444.
- Phillips, C., 1993. *Cattle Behaviour*. Farming Press Books.
- Rahn, K., Renwick, S., Johnson, R., Wilson, J., Clarke, R., Alves, D., McEwen, S., Lior, H., Spika, J., 1997. Persistence of *Escherichia coli* O157:H7 in dairy cattle and the dairy farm environment. *Epidemiol. Infect.* 119, 251–259.
- Rasmussen, M., Cray, J.W., Casey, T., Whipp, S., 1993. Rumen contents as a reservoir of enterohemorrhagic *Escherichia coli*. *FEMS Microbiol. Lett.* 114, 79–84.
- Reilly, W., Locking, M., O'Brien, S., Wright, E., Campbell, D., Coia, J., Browning, L., Ramsay, C., 2000. Risk factors for sporadic cases of *E. coli* O157 infection: the importance of contact with animal excreta. In: *Zoonotic Infections in Livestock and the Risk to Public Health VTEC O157*, Edinburgh. Scottish Centre for Environmental Health, Glasgow, Scotland, pp. 41–47.
- Renshaw, E., 1991. *Modelling Biological Populations in Space and Time*. Cambridge Studies in Mathematical Biology. Cambridge University Press, Cambridge.
- Richards, I., Wolton, K., 1976. The spatial distribution of excreta under intensive cattle grazing. *J. Br. Grassl. Soc.* 31, 89–92.
- Sanderson, M., Besser, T., Gay, J., Gay, C., Hancock, D., 1999. Fecal *E. coli* O157:H7 shedding patterns of orally inoculated calves. *Vet. Microbiol.* 69, 199–205.
- Shere, J., Bartlett, K., Kaspar, C., 1998. Longitudinal study of *E. coli* O157:H7 dissemination on four dairy farms in Wisconsin. *Appl. Environ. Microbiol.* 64, 1390–1399.
- Synge, B., Gunn, G., Ternent, H., Hopkins, G., Thompson-Carter, F., Foster, G., McKendrick, I., 2000. Preliminary results from epidemiological studies in cattle in Scotland. In: *Zoonotic Infections in Livestock and the Risk to Public Health VTEC O157*, Edinburgh. Scottish Centre for Environmental Health, Glasgow, Scotland, pp. 10–17.
- Taylor, H., Karlin, S., 1998. *An Introduction to Stochastic Modeling*. Academic Press, New York.
- Wang, G., Zhao, T., Doyle, M., 1996. Fate of enterohemorrhagic *E. coli* O157:H7 in bovine feces. *Appl. Environ. Microbiol.* 62, 2567–2570.
- Woodward, M., Gavier-Widen, D., McLaren, I., Wray, C., Sozmen, M., Pearson, G., 1999. Infection of gnotobiotic calves with *Escherichia coli* O157:H7 strain A84. *Vet. Rec.* 144, 466–470.
- Wood, J., McKendrick, I., Gettinby, G., 2006a. Assessing the efficacy of within-animal control strategies against *E. coli* O157: a simulation study. *Prev. Vet. Med.* 74, 194–211.
- Wood, J., McKendrick, I., Gettinby, G., 2006b. A simulation model to assess herd-level intervention strategies against *E. coli* O157. *Epidemiol. Infect.*, submitted for publication.