

Published in final edited form as:

J Biol Dyn. 2012 ; 6(2): 1052–1066. doi:10.1080/17513758.2012.722232.

The strain-specific dynamics of *Escherichia coli* O157:H7 faecal shedding in cattle post inoculation

R. Gautam^{a,*}, M. Kulow^b, D. Döpfer^b, C. Kaspar^c, T. Gonzales^b, K.M. Pertzborn^b, R.J. Carroll^d, W. Grant^e, and R. Ivanek^a

M. Kulow: mkulow@svm.vetmed.wisc.edu; D. Döpfer: dopferd@vetmed.wisc.edu; C. Kaspar: cwkaspar@wisc.edu; T. Gonzales: tixgirl3@hotmail.com; K.M. Pertzborn: kpertz@iastate.edu; R.J. Carroll: carroll@stat.tamu.edu; W. Grant: wegrant@tamu.edu; R. Ivanek: rivanek@cvm.tamu.edu

^aDepartment of Veterinary Integrative Biosciences, College of Veterinary Medicine and Biomedical Sciences, Texas A&M University, College Station, TX 77843, USA

^bSchool of Veterinary Medicine, University of Wisconsin, Madison, WI 53706, USA

^cDepartment of Bacteriology, University of Wisconsin, Madison, WI 53706, USA

^dDepartment of Statistics, Texas A&M University, College Station, TX 77843, USA

^eDepartment of Wildlife and Fisheries Sciences, College of Agriculture and Life Sciences, Texas A&M University, College Station, TX 77843, USA

Abstract

This study reports analysis of faecal shedding dynamics in cattle for three *Escherichia coli* O157:H7 (ECO157) strains (S1, S2 and S3) of different genotype and ecological history, using experimental inoculation data. The three strains were compared for their shedding frequency and level of ECO157 in faeces. A multistate Markov chain model was used to compare shedding patterns of S1 and S2. Strains S1 and S2 were detected seven to eight times more often and at 10^4 larger levels than strain S3. Strains S1 and S2 had similar frequencies and levels of shedding. However, the total time spent in the shedding state during colonization was on average four times longer for S1 (15 days) compared to S2 (4 days). These results indicate that an ECO157 strain effect on the frequency, level, pattern and the duration of faecal shedding may need to be considered in control of ECO157 in the cattle reservoir.

Keywords

Escherichia coli O157:H7; multistate Markov chain model; shedding; colonization

1. Introduction

Escherichia coli O157:H7 (ECO157) is a member of the enterohaemorrhagic *E. coli* and is an important cause of haemorrhagic colitis in humans worldwide [39,40]. Faecal shedding of ECO157 by the colonized (we are using the term ‘colonize’ instead of infection, because ECO157 does not cause clinical symptoms in cattle) cattle is a major public health concern because it presents a risk for human exposure and illness through direct and indirect contacts. Human illness from direct contact is possible during farm visits [6] or consumption

© 2012 R. Gautam

*Corresponding author. rgautam@cvm.tamu.edu.

This paper is based on an invited talk given at the 3rd International Conference on Math Modeling & Analysis, San Antonio, USA, October 2011.

of contaminated milk and meat products [9]. Indirect contact-related illness may occur due to consumption of fruits and vegetables that are contaminated from application of cattle manure as fertilizers or irrigation from water sources supplied with farm runoffs [40]. Reducing level of ECO157 in faeces and/or length of faecal shedding in cattle directly contributes to decreasing human exposure and thereby is the goal of every control programme [14,41]. However, despite copious research efforts, it has not been possible to reduce faecal shedding prevalence and levels to levels that prevent human exposure and illness. One of the major hindrances is the lack of understanding of the faecal shedding pattern during colonization, the process which makes detection, and thus control, difficult.

The patterns and duration of faecal shedding reported in literature for ECO157 in cattle vary considerably. The herd level shedding prevalence of ECO157 in cattle varies between 8% and 87% [17,25] and the prevalence in a positive herd may vary between 1.7% and 20% [25]. The observed variation in within and between herd prevalence may be explained partly by the difference in the duration of shedding [24]. Variability in the pattern and length of shedding in turn may be governed by a multitude of host, pathogen and environment-related factors such as cattle production type, animal age, colonizing strain of ECO157 and the exposure dose used in the study. Ages of cattle and exposure doses have been shown to be important determinants of ECO157 faecal shedding pattern and duration [5]. Cattle exposed to high doses and young calves tend to shed at a higher level and for a longer period of time [5]. The strain effect on the shedding pattern and duration has been suggested from an inoculation study using a mixture of five strains [3], but it has not been investigated using single strains under controlled experimental conditions which eliminate the possibility of strain interaction within a host. Furthermore, we are not aware of any study that quantitatively describes and compares strain-specific shedding pattern and duration of cattle colonization with ECO157.

For an intermittently shed pathogen, such as ECO157, the duration of colonization comprises one or more periods of active shedding followed by transient non-shedding. Breaking down the length of colonization into its components: shedding and non-shedding (or below detection level shedding) periods and describing them quantitatively is desirable for better detection and control of ECO157 in cattle, but it is constrained by the complexity of shedding behaviour and financial limitations to conduct a large experimental study with intensive sample collection. We hypothesize that the pattern of faecal shedding in cattle is affected by the strain of ECO157 colonizing the host. Thus, there is a need for statistical comparison of strain-specific shedding patterns and durations in terms of the length of shedding and non-shedding episodes and the total duration of colonization for different strains. This can be achieved by employing multistate Markov chain (MC) modelling [19,29].

An MC model is defined by a set of states and transition intensities to describe movement of an individual between these states in continuous time [29]. The movement between states is governed by transition intensities which may depend on time and a set of independent or time-dependent explanatory variables [20]. While recurrent events survival analysis may also be used to analyse multistate models with recurrent events, the MC model has the advantage in terms of bringing out important biological insights which may be ignored in the former method [30]. In an MC model, each censored individual contributes with more information to the model than it can contribute to survival analysis with a single event end-point (that the individual did not reach) because prior state transition experiences of these subjects add useful information to the process. The most important limitation of the survival analysis is its inability to make predictions beyond observed data, whereas multistate MC model can be used to predict beyond untested circumstances [18]. Furthermore, the MC model presents a flexible tool for the study of covariate effect on the various transition rates.

The objective of this study was to use experimental inoculation data on ECO157 faecal shedding in cattle to compare three ECO157 strains by (i) describing and comparing strain-specific frequency and level (CFU/sample) of faecal shedding, and (ii) estimating and comparing the duration of shedding and non-shedding episodes as well as the total length of colonization, using a multistate MC model.

2. Materials and methods

2.1. Description of experimental setup

2.1.1. Experimental animals—Twenty-four, 4–5-month-old Holstein steers, weighing between 127 and 191 kg were purchased from a commercial livestock dealer in Wisconsin. The steers were randomly ear-tagged (with numbers: 1–24), divided into four groups of six, and housed in four separate pens in two rooms (A and B; each with two pens separated by a corridor to prevent contact between groups) at the Livestock Laboratory of the University of Wisconsin-Madison. On arrival to the Laboratory, the steers were vaccinated with Bovi-Shield Gold® (Pfizer, Inc., New York, NY) and Vision 7® (Intervet, Inc., Summit, NJ) in accordance with the label directions. The steers also received a single dose of Ceftiofur (3.0 mg CE/lb Excede®, Pfizer, Inc., New York, NY) to prevent shipping fever. Steers 1 through 12 were housed in room A (with steers 1–6 in pen 1, and steers 7 through 12 in pen 2). Steers 13 through 18 were housed in pen 3 and steers 19 through 24 were housed in pen 4 in room B. The rooms had slatted floor, controlled temperature (20°C) with a positive pressure ventilation system, and were equipped with headlocks and waste flushing system. Only the animal caretakers and the investigators had access to the rooms. The pens were cleaned two times a day by flushing water over the pen floor. Steers were provided with a ration consisting of alfalfa, shelled corn, chopped hay, and a standard concentrate supplement once a day in accordance with the National Research Council's specifications to attain an average weight gains of 0.7 kg/day [31]. The ration was offered in bins and water was provided *ad libitum* using water flow cups. Both feed and water were accessible through headlocks. The steers were monitored daily for their health status prior to the experimental challenge and three times per week after the experimental challenge. Also, for 20 days prior to experimental inoculation, steers were monitored by obtaining and testing recto-anal mucosal swabs (RAMS; described below) once every 2 days to ensure the absence of wild-type ECO157. All procedures were approved by the University of Wisconsin-Madison School of Veterinary Medicine IACUC (protocol number A01388-0-03-09).

2.1.2. Animal inoculation and ECO157 strains—Experimental inoculations of steers in the study with three different ECO157 strains (FRIK 47, FRIK 1641 and FRIK 2533) were conducted simultaneously. For brevity, hereafter, we will refer to these three strains as S1, S2 and S3, respectively. The 12 steers in room A were inoculated with S1, and the six steers each in pens 3 and 4 in room B were inoculated with S2 and S3, respectively. Within each group the steers were inoculated with 10^6 CFUs of the respective strain of ECO157 *per os* in drinking water. Inoculation was performed by offering 1 ml of inoculum diluted in approximately 10 ml water in a clean and empty water cup to individual steers in each group. To ensure complete uptake of the inoculums, steers were restricted from drinking water for 12 h before inoculation. Investigators observed the steers while they consumed the entire inoculums. Additional fresh water was provided in the same water cup to ensure that any remaining inoculum was consumed.

The three strains of ECO157 used in these inoculation experiments represent different ecological histories in terms of their source of identification and isolation. Strain S1 was originally isolated from an infected human patient [32], S2 from a raccoon [37] and S3 from a cattle farm. These strains also have different genotypes as indicated by different genetic

marker profiles [16]. Most notably, strains S2 and S3 lacked *stx1* and *hly₉₃₃* genetic markers, respectively.

2.1.3. Sampling, isolation, enumeration and PCR confirmation—The RAMS samples were collected from all steers 1 day after inoculation and every alternate day thereafter through day 30 post inoculation using sterile cotton tipped applicators (6 inch long). The applicator was inserted 1–2 inches into the rectum and rubbed three to four times along the mucosal surface using rapid motion. All RAMS samples were stored in 5 ml of modified *E. coli* broth (mEC; Remel, Lenexa, KS) and novobiocin (0.02 mg/ml; mECnov; novobiocin by Remel, Lenexa, KS) and processed within 1 h. The RAMS samples were tested for the presence of ECO157 by direct plating of samples on Sorbitol MacConkey agar supplemented with 0.05 mg/l cefixime and 2.5 mg/l potassium tellurite (Cefixime tellurite-Sorbitol MacConkey (CT-SMAC), BD, Sparks, MD). Plates were incubated at 37°C for 20–24 h and suspect colonies of ECO157 were presumptively confirmed by the latex agglutination test (Remel, Lenexa, KS). For each potentially positive sample, up to five randomly selected suspect colonies on CT-SMAC agar and those confirmed by latex agglutination were selected, and sub-cultured in Luria-Bertani (LB) broth (BD Difco™ LB Broth, Miller, Sparks, MD) for 3 h and frozen in 15% glycerol-LB media at –80°C for later confirmation by PCR. Enumeration of ECO157 per RAMS was performed for all positive samples. Exact enumeration was recorded if the RAMS sample had ECO157 level below or slightly over 10⁵ CFU. Four RAMS samples had extremely large ECO157 levels (i.e. much higher than the upper enumeration limit). These RAMS samples were therefore, assigned a number that was greater than our upper enumeration limit by first-order of magnitude (i.e. 10⁶ CFU). The minimum detection limit of the direct plating procedure was <10 CFUs of ECO157/RAMS sample. Template DNA, extracted using the boiled colony method, was used in a realtime PCR as reported by Gonzales *et al.* [16]. PCR confirmation was done using a panel of 28 genetic markers on the OpenArray® System from Applied Biosystems by Life Technologies [16].

2.2. Statistical analyses

The overall and group specific numbers of sampling days when RAMS samples were found positive for the three ECO157 inoculation strains were estimated and compared by computing the medians and interquartile ranges (IQR). The proportion of animals shedding the inoculated strain at any sampling day was calculated by dividing the number of animals shedding the inoculated strain by the total number of animals initially inoculated with that ECO157 strain. Group specific time series of the proportion of steers shedding an ECO157 strain was visualized graphically. The overall and strain-specific average level (CFU/sample) of ECO157 shed by an individual steer was calculated by considering only the positive samples. From this, the overall and the strain-specific distribution of ECO157 shedding level/sample was calculated and represented as the probability distribution function. The median, IQR, and 95% confidence interval (CI) of the ECO157 shedding level/sample/day for all steers were calculated and represented as a time series plot using box plots. The difference in the ECO157 faecal shedding level between the inoculation strain groups was tested by using the nonparametric Wilcoxon-rank-sum test for two samples. The difference in the proportion of animals starting to shed following inoculation with different strains was assessed by comparing the number of animals shedding on the day 2 using a Chi-square test. To quantify the duration of shedding and colonization and assess whether strain effect produced a statistically significantly different dynamics of faecal shedding, a multistate MC model was developed as explained in the next section. Strain S3 was not evaluated in the multistate MC model because this strain was detected only on one occasion post inoculation and in only two steers. Thus, obviously this strain had a markedly different shedding pattern and a shorter duration of shedding compared to the other two

strains. All statistical analyses were performed in R version 2.13.1 (R Development Core Team, 2011). In all analyses of statistical significance, a P -value of 0.05 was used.

2.2.1. Multistate Markov model structure and analysis—The longitudinal faecal shedding data from our ECO157 inoculation study provided information on ECO157 presence/absence and concentration (CFU/RAMS) in faeces on any given day and the pattern of shedding over a period of 30 days. Several pertinent facts on the faecal-shedding pattern were observed, which complemented with the existing understanding of the intermittent nature of ECO157 faecal shedding, formed the basis for constructing the multistate MC model. Some animals started to shed from the next day after inoculation; some did not start shedding until several days after inoculation while a few others were never observed to shed. The animals that started to shed stopped faecal shedding after a few days. A few of the animals that stopped shedding were never detected to shed again while others resumed shedding. Those that resumed shedding went through several episodes of shedding and non-shedding until the end of the study. All of the foregoing observations were used to determine the possible states and transitions in constructing the faecal shedding multistate MC model. A four state transition MC model, with states for latency, shedding, non-shedding, and recovery was developed to describe the observed faecal shedding patterns. The information on the ECO157 presence or absence in a RAMS sample and the time of sample collection since challenge were used to classify animals into one of the four states under the following conditions:

- State 1 = latency: The day of inoculation and negative samples immediately after challenge. If the animals remained negative on 4 consecutive sampling days, they were considered recovered and entered state 4. This transition was allowed because not all exposed animals became colonized by ECO157 as is evident from Table 1.
- State 2 = shedding: Any positive sample during the study period.
- State 3 = intermittent non-shedding: Negative samples following state 2 over a period of less than 7 days or (considering that samples were collected every other day) over 3 consecutive sampling occasions.
- State 4 = Recovery: Negative samples over a period of more than 7 days or over 4 consecutive sampling occasions. Four or more consecutive negative samples following state 2 was considered indicative of clearance of ECO157 from the animal's recto-anal junction; and likewise, this condition following state 1 indicated that the animal never became colonized after exposure (inoculation). This condition was applied to prevent false classification of animals in states 1 and 3 as recovered (state 4) and is consistent with criteria used by several previously published studies [22,36].
- 99 = censored state: Negative samples at the end of the study for which, because of termination of the study, it was not clear whether an animal is in the intermittent non-shedding state (state 3) or in the recovered state (state 4). Thus, by censoring, we mean that the animal could only be in one of the two possible censor states, state 3 or state 4. The reader should note that unlike the usual meaning of a censored event time in the survival analysis, here it is the state (and not the event time) that is censored [20].

A flow diagram of the multistate MC model with three transient states ($i = 1, 2$ and 3) and one absorbing state ($i = 4$) is presented in Figure 1. The arrows in the diagram indicate possible transitions between states and the coefficients (q_{ij}) represent the corresponding transition intensities between states.

The developed multistate MC model was used similarly as in [19], to quantify and compare faecal shedding patterns of strains S1 and S2 in terms of transition intensities among states, and the duration of average one-time and total time of stay in the shedding and non-shedding states. The exact transition times were unknown and they were modelled as such. However, observation times were considered and modelled as fixed and sampling times were considered to be non-informative. The model was analysed using the *msm* package in R. Theoretical and methodological details about the implementation and execution of multistate MC models using the *msm* package in R can be found elsewhere [20]. The multistate MC model was implemented under three specific assumptions about the ECO157 faecal shedding process: (i) transition times are independent of the faecal shedding processes before reaching the current state, (ii) transition intensities are homogeneous across steers inoculated with the same strain and (iii) transition intensities are homogeneous through time.

Assumption (i) is an inherent property of Markov models which could not be assessed due to lack of data about the exact time of transitions [23]. Assumption (ii) was assessed by including the 'ECO157 strain' variable as a covariate in the saturated model (H_{s1} , where the subscript 's' denotes saturated and '1' denotes the model number) and comparing the resulting model fit to the fit for the model without the covariate (H_0). In the saturated model (H_{s1}), the effect of ECO157 strain (covariate) in all of the five permitted transitions was assessed. Therefore, the full model had a total of 10 parameters, 5 corresponding to baseline transition intensities and 5 different regression coefficients for the covariate effect. Several reduced versions of the full model were assessed by constraining selected transition intensities and to identify the most parsimonious covariate model. In doing so, the following biologically meaningful relationships between transition intensities that could generate a better model fit were considered: (a) in H_{s2} , we tested if the ECO157 strain effect was equal on the progression (i.e. $q_{12} = q_{23}$) and recovery (i.e. $q_{14} = q_{24}$) transitions; (b) in H_{s3} , we tested if the effect of ECO157 strain was the same only for transitions to recovery (i.e. only $q_{14} = q_{24}$); (c) in H_{s4} , we tested if the strain effect was the same only for the progression from state 2 to state 3 and regression from state 3 to state 2 (i.e. only $q_{23} = q_{32}$); (d) in H_{s5} , we tested if the effect of strain on all transition intensities was the same (i.e. $q_{12} = q_{14} = q_{23} = q_{24} = q_{32}$); and (e) in H_{s6} , we tested if all the transition intensities were of equal magnitude with transition intensities for recovery acting in the opposite direction to those for progression and regression (i.e. $q_{12} = -q_{14} = q_{23} = -q_{24} = q_{32}$). The best covariate model was selected by comparing the likelihood ratios (LR) test statistics. Strain-specific differences in faecal-shedding patterns were assessed by comparing several measures including: (1) hazard ratios (HRs), (2) ratios of transition intensities for states having more than one transition out of it, (3) duration of stay in each state (i.e. sojourn and total time resulting from recurrent visit to the state) and (4) the probability for not recovering from each state (i.e. the survival rate in the jargon of survival analysis). Assumption (iii) was assessed by fitting a time-inhomogeneous model that allows time to be piece-wise constant [20] and comparing the fit of this model to the time-homogeneous model. In the model with the strain effect, strain (S1) was used as the reference group.

The internal validity of the model was assessed by comparing the observed proportion of steers in different states over time with the model predicted mean prevalences in different states and constructing the 95% CIs around the mean prevalences, using bootstrapping methods. To calculate the CIs, 1000 simulations of random vectors were generated from the asymptotic multivariate normal distribution derived from the maximum likelihood estimates (and covariance matrix) of the log transition intensities followed by the calculation of the covariate effects and the expected prevalences for each replicate [15]. The external validity of the model was assessed for strain S1 by using an independent data set on faecal shedding described for the control group of calves in a recently published study [22].

3. Results

3.1. Descriptive statistics

The data set consisted of 15 RAMS samples for each of the 24 steers over the 30 day study period with a total of 360 RAMS samples. In the S1 strain group, 8 out of 12 (67 %) animals were positive while in the S2 and S3 strain groups only 2 out of 6 animals (33%) were positive on the day following inoculation (i.e. day 2). However, the difference in proportions positive on day 2 between groups was not statistically significant. While the time series of the proportion of positive animals in the S1 and S2 strain groups followed an irregular pattern over the study period, there were no positive animals in the S3 group after day 2 (Figure 2(a)). The median numbers (and ranges) of positive samples detected per steer inoculated with strains S1, S2 and S3 were 8 (0–13), 7 (0–10) and 0 (0–1), respectively. Then obviously, the number of positive samples for steers in the S1 and S2 strain groups were significantly higher than for steers in the S3 strain group (P -values = 0.005 and 0.01, respectively). The median shedding level per steer was $10^{2.98}$ CFUs (IQR = $10^{1.92} - 10^{4.29}$ CFUs). Illustrative time series plots are presented in Figure 2(b–d) depicting variation in the overall and strain-specific shedding levels during the study period, respectively. The median shedding levels in the S1 ($10^{2.97}$ CFUs) and S2 (10^3 CFUs) strain groups were significantly greater than those for the S3 strain group in which the maximum shedding level was $10^{1.25}$ CFUs (P -values 0.04 and 0.03, respectively); here comparison to the higher of the two shedding loads was deemed more informative than comparison to the median because there were only two observations for shedding level in S3 group). However, the median shedding levels of strains S1 and S2 were similar. The distribution of all detected ECO157 shedding levels was bimodal (Figure 3). Approximately, 60% of samples had less than 5000 CFUs of ECO157 and only about 25% of positive samples had levels above 10^4 CFU. However, almost all steers shedding ECO157 were detected to shed at high levels at some point during their colonization with the exception of a few.

3.2. Multistate MC model

The portion of the data set for strains S1 and S2 contained 288 observations with 80 positive samples (categorized as the shedding state or state 2). The negative samples ($n = 208$) were categorized into different states as follows: 19 observation as latent, 5 as intermittent non-shedding and 183 as recovered. The remaining single data point was censored (i.e. the steer could have been either in the intermittent non-shedding or recovered state). Table 1 shows the number of transition between model states. The best covariate model was the H_{s6} model (with constraints $q_{12} = -q_{14} = q_{23} = -q_{24} = q_{32}$) which had a borderline significantly better fit ($df = 6$, LR P -value = 0.13) than the model without the covariate. Considering the biological relevance of the strain effect on the dynamics of faecal shedding, model H_{s6} was examined further. For this model, the comparison between the observed and predicted prevalences including their 95% CI for steers in different states indicated a very good fit of the model to the data (Figure 4). It was concluded that the data supported the internal validity of the model. The fitted time-inhomogeneous model did not suggest any evidence of transition intensities changing over time, therefore, the third model assumption of homogeneous transition intensities through time was considered to hold.

The mean sojourn time (duration of one-time occupancy of a state) and the expected total length of stay in the transient states are shown in Table 2. While the two strains had similar sojourn times, the total lengths of stay in the shedding and non-shedding states for the S1 strain were on average almost four times greater than for the S2 strain. Such a difference in the total duration of stays could be explained by more frequent re-entry of animals colonized with the S1 strain into the shedding and non-shedding states as compared to those colonized with the S2 strain. Indeed, Figure 5 shows that steers inoculated with the strain S1 were on

average less likely to recover from the latent or shedding states compared to the steers inoculated with strain S2. This is also supported by on average greater HR for moving from latent and shedding states to the recovered state in the S2 group compared to S1 group (Figure 6). It should be noted here that CIs of the estimated HRs crossed the value of 1 by a narrow margin indicating a lack of conclusive evidence of the effect of strains on the transition intensities. The average probability for not entering the recovered state from all transient states against time was greater for S1 compared to S2 group (Figure 7). All together, these results indicate that colonization by the strain S1 tends to last longer than by strain S2 and explain the estimated longer average total times spent in the shedding and non-shedding states for S1 compared to S2 (Table 2).

The external validation with an independent data set from [22] for the strain S1 predicted the average sojourn and total time in state 2 (the shedding state) of 4 days (95% CI = 0.1, 24.3) and 11.4 days (95% CI = 3.9, 31.1), respectively, which are very close to the model predicted estimates here (4.3 and 14.6 days respectively). Predictions of the average sojourn and total times in the non-shedding state of 0.2 days (95% CI = 0, 32) and 0.46 days (95% CI = 0.07, 2.7), estimated for the independent data, also had reasonably good fit to the model predictions of 0.4 and 1.2 days, respectively.

4. Discussion

In this study, we evaluated the effect of ECO157 strain on the ECO157 faecal shedding dynamics during colonization in cattle using experimental inoculation data. We found that one strain (S3) was substantially different from the other two strains in that it was detected only once during the entire study period and at 10^4 CFU/RAMS lower levels. This is in agreement with the fact that it lacked the whole plasmid pO157, which among other genetic material carries an important virulence marker *hly₉₃₃* (enterohemolysin) held responsible for survival of ECO157 in the cattle host [26,27]. While shedding frequencies and levels were similar between S1 and S2, they appeared to have different shedding patterns. Steers colonized by S1 appeared to make recurrent visits between shedding and non-shedding more often than steers colonized by S2 although the sojourn times of the two strain groups were comparable. Thus, the total duration of stay in the shedding state (state 2) for steers in the S1 group was on average almost four times longer compared to steers in the S2 group. In the sections that follow, we elaborate on the findings and discuss the implications, merits and limitations of the study.

The period of ECO157 colonization in cattle has been recognized to exhibit an intermittent shedding pattern, with periods of non-shedding (or shedding at undetectable levels) interspersed within shedding periods [2,28]. Nevertheless, shedding and colonization have been traditionally used interchangeably without firm distinction between the two terms. An important achievement of this study was, therefore, in the ability to quantitatively partition the total colonization time of ECO157 into shedding and non-shedding periods and to study the patterns of faecal shedding. Evidence of the strain effect on the pattern of ECO157 faecal shedding was found in this study. While the strain effect model had only a borderline better fit compared to the model without the strain effect (P -value = 0.13), its internal and external validation supported its validity. Furthermore, CIs of the estimated HRs and ratios of transition intensities crossed the value of one, i.e. zero if log transformation is considered, (which would indicate absence of the strain effect) by only a small margin. Thus, it seems plausible that the borderline significant fit of the strain effect model was a consequence of the considerably small sample size rather than there not being a difference between strains S1 and S2. The strain effect model indicated that strains S1 and S2 have similar sojourn times in the states of faecal shedding. However, strain S1 had on average four times longer total length of stay in the shedding and non-shedding states during colonization compared to

strain S2. This can be explained by two biological phenomena. First, steers colonized with strain S1 were on average more likely to move to the shedding state as indicated by the HRs in Figure 6 and transition intensity ratios in Figure 5. Second, the HRs and transition intensity ratios also indicated that steers colonized with strain S1 were more likely to make recurrent visits between the shedding and non-shedding states before final clearance of ECO157 from faeces compared to those colonized with strain S2. The average duration of colonization, obtained by adding the average length of time spent in the latent, shedding and non-shedding states, were 16 (95% CI = 6, 38) days and 5 (95% CI = 1, 17) days for strains S1 and S2. Previously published literature have reported average duration of faecal shedding of ECO157 (strains: 86–24 Nal-R and FRIK 1275) in cattle to vary between 17 and over 43 days [34,38], which is consistent with the average colonization time for S1 estimated in this study. While it would be interesting to compare the genetic markers of strains 86–24 Nal-R and FRIK 1275 with S1, we were not able to determine genetic markers for the former two strains from the published literature.

The substantially longer total duration of shedding of strain S1 compared to S2 has important implications on the spread and control of ECO157 within the cattle population. It is reasonable to assume that longer total shedding time provides longer opportunity for the pathogen spread to the susceptible herd members. Given an equal or better colonization ability compared to the other strains, we expect that a strain with the longer total shedding period will be more prevalent and more adapted to the cattle population. On the other hand, its longer total shedding time may also provide a greater opportunity for its detection and may require a less intensive sampling to detect all positive animals.

Apart from the study presented here, we are aware of only two other studies which tried to separate and quantify the shedding and non-shedding states, using an MC approach for *Listeria monocytogenes* in cattle [18] and *Salmonella* in pigs [19]. That being said, in Ivanek *et al.* [18], intermittent shedding was studied in the naturally infected calves and thus the estimated duration of shedding and non-shedding episodes cannot be used to determine the duration of infection following exposure mainly for two reasons: (i) because the time when calves became colonized was unknown and (ii) the possibility of re-infection(s) was not considered. In that sense, even the shedding and non-shedding durations do not represent the shedding behaviour within one colonization cycle. We used the experimental data with known day of challenge and used the literature supported decision criteria to rule out the possibility of shedding due to re-infection. Thus, the study presented here represents one colonization cycle and demonstrates how multistate MC modelling approach could be used to describe intermittent shedding behaviour in terms of duration of shedding and non-shedding episodes as well as the total colonization time influenced by the ECO157 strain covariate.

The three strains used in this study had different ECO157 ecological histories. Namely, strain S3 was isolated from a cattle farm (C.W. Kaspar, unpublished data), strain S2 was isolated from a raccoon [37] and strain S1 is a fully sequenced human pathogen 933W (ATCC43895) originally isolated from an infected human patient [32]. The three strains also had different genotypes as indicated by the pulse field gel electrophoretic patterns [37,38] and genetic marker profiles [16]. In particular, compared to strain S1, strain S3 lacked genetic markers O11-C, R4-N and R5-N as well as the whole plasmid pO157, including the genetic marker *hly*₉₃₃, while strain S2 lacked two markers, *stx1* and R4-N [16]. The inability of S3 to successfully colonize steers after inoculation in contrast to the other two strains and the difference in the shedding patterns observed between strains S1 and S2 may thus be explained by the differences in their genotypes. Indeed, it has been reported that ECO157 strains with different virulence factors, and their combinations, have different colonization abilities in cattle and are also relevant from the human health point of view. In particular

stx1 and *stx2* are known to be part of the genome of lambdoid prophages [13] responsible for production of cytotoxin that contributes to diarrhoea observed in infected humans [21]. On the other hand *hly₉₃₃* is a plasmid encoded virulence factor responsible for cytotoxin known to cause lysis of red blood cells [12]. These findings obviously lead to another interesting question: could the observed variations in faecal shedding prevalence under field conditions in otherwise similar herds partly be explained by the differences in the genetic markers of the colonizing strains? It is not a purely speculative statement, because, at least one experimental study involving inoculation of calves with a mixture of five ECO157 strains documented the predominance of only a few strains in the faeces [3]. Another field study reported differences in the restriction endonuclease digestion profiles among ECO157 positive cattle farms [10]. This warrants an independent field study to investigate the association between specific genetic marker combinations and shedding prevalence, as was also suggested by at least one other recently published study [8]. The ECO157 shedding levels per sample/day had a bimodal distribution with the majority of samples sharing low shedding levels (<5000 CFUs) and only a small proportion (<20%) having high shedding levels (>50,000 CFUs) (Figure 3). The proportion of RAMS under high and low shedding level category observed in this study was consistent with the previously published study that reported shedding levels of less than 5000 CFUs/g of faeces in 53% of samples and more than 50,000 CFUs/g of faeces in 27% of the samples [35]. The occurrence of the bimodal shedding pattern has been reported and was attributed to the presence of a few high level shedders within the population [4]. However, a close examination of the individual steer-shedding level data revealed that almost all steers shed at a high level (>10,000) at some point during the course of colonization, and the high shedding level was not always observed in the first shedding episode; rather, occasionally, the high shedding levels were present in the subsequent episodes of shedding (i.e. after a transient period of non-shedding). Therefore, the observed bimodal shedding level pattern in this study essentially reflects shedding at two different rates (high and low) during the course of colonization by ECO157 rather than the existence of two distinct cattle sub-populations adapted to shedding at different rates.

The findings in this study depend on several important assumptions that were made in the development of the model in addition to the underlying multistate MC model assumptions. The assumptions that the steers were correctly classified into shedding and non-shedding states and that the model structure in terms of states considered and the transitions permitted between states were valid. The ability to classify the individuals correctly into shedding and non-shedding states was related to the performance of the diagnostic test. The data used in this study were generated using direct plating on CT-SMAC to detect the presence of ECO157 in RAMS which is a standard method [11]. While we acknowledge that faecal shedding may have been under detected if the shedding level was below 10 CFUs/RAMS, overall we do not expect misclassification of steers to be a major problem in this study. Intermittent shedding behaviour of ECO157 among cattle is a widely accepted phenomenon [1,33]; therefore, we believe that the model structure used in this study closely reflected the biological process of faecal shedding in cattle.

In conclusion, we have quantitatively described the shedding pattern of ECO157 and compared the potential influence of colonizing strain on the shedding dynamics. The results indicated that faecal shedding associated with different ECO157 strains varies with respect to shedding frequency, levels, patterns and durations. Since the three strains evaluated in this study had different genotypes, our findings support previous reports that specific genetic (virulence) markers play an important role in the ability of ECO157 to colonize and persist in cattle [3,7]. The parameter estimates obtained from this study will be useful in mathematical modelling to advance our understanding of ECO157 transmission dynamics.

Collectively, these findings will support efforts to control ECO157 in cattle with the ultimate goal of protecting human health.

Acknowledgments

This work was supported by the National Science Foundation grant NSF-EF-0913367 to R.I. funded under the American Recovery and Reinvestment Act of 2009. Carroll's work was supported by a grant from the National Cancer Institute R37-CA057030. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. The authors are thankful to two unknown reviewers for their thoughtful comments and suggestions that improved the manuscript.

References

1. Bach SJ, McAllister TA, Veira DM, Gannon VPI, Holley RA. Transmission and control of *Escherichia coli* O157:H7 – a review. *Can J Anim Sci.* 2002; 82:475–490.
2. Besser TE, Hancock DD, Pritchett LC, McRae EM, Rice DH, Tarr PI. Duration of detection of fecal excretion of *Escherichia coli* O157:H7 in cattle. *J Infect Dis.* 1997; 175:726–729. [PubMed: 9041355]
3. Brown CA, Harmon BG, Zhao T, Doyle MP. Experimental *Escherichia coli* O157:H7 carriage in calves. *Appl Environ Microbiol.* 1997; 63:27–32. [PubMed: 8979335]
4. Chase-Topping ME, McKendrick IJ, Pearce MC, MacDonald P, Matthews L, Halliday J, Allison L, Fenlon D, Low JC, Gunn G, Woolhouse MEJ. Risk factors for the presence of high-level shedders of *Escherichia coli* O157 on Scottish farms. *J Clin Microbiol.* 2007; 45:1594–1603. [PubMed: 17360845]
5. Cray WC Jr, Moon HW. Experimental infection of calves and adult cattle with *Escherichia coli* O157:H7. *Appl Environ Microbiol.* 1995; 61:1586–1590. [PubMed: 7747972]
6. Crump JA, Sulka AC, Langer AJ, Schaben C, Crielly AS, Gage R, Baysinger M, Moll M, Withers G, Toney DM, Hunter SB, Hoekstra RM, Wong SK, Griffin PM, Van Gilder TJ. An outbreak of *Escherichia coli* O157:H7 infections among visitors to a dairy farm. *N Engl J Med.* 2002; 347:555–560. [PubMed: 12192014]
7. Dopfer D, Geue L, de Bree J, de Jong MC. Dynamics of verotoxin-producing *Escherichia coli* isolated from German beef cattle between birth and slaughter. *Prev Vet Med.* 2006; 73:229–240. [PubMed: 16488028]
8. Dopfer D, Geue L, Schares S, Mintel B, Hoffmann B, Fischer EA. Dynamics of shiga-toxin producing *Escherichia coli* (STEC) and their virulence factors in cattle. *Prev Vet Med.* 2011; 103:22–30. [PubMed: 21920615]
9. Elder RO, Keen JE, Siragusa GR, Barkocy-Gallagher GA, Koohmarale M, Laegreid WW. Correlation of enterohemorrhagic *Escherichia coli* O157 prevalence in feces, hides, and carcasses of beef cattle during processing. *Proc Natl Acad Sci.* 2000; 97:2999–3003. [PubMed: 10725380]
10. Faith NG, Shere JA, Brosch R, Arnold KW, Ansay SE, Lee MS, Luchansky JB, Kaspar CW. Prevalence and clonal nature of *Escherichia coli* O157:H7 on dairy farms in Wisconsin. *Appl Environ Microbiol.* 1996; 62:1519–1525. [PubMed: 8633851]
11. Fox JT, Renter DG, Sanderson MW, Thomson DU, Lechtenberg KF, Nagaraja TG. Evaluation of culture methods to identify bovine feces with high concentrations of *Escherichia coli* O157. *Appl Environ Microbiol.* 2007; 73:5253–5260. [PubMed: 17574994]
12. Fratamico PM, Bagi LK, Pepe T. A multiplex polymerase chain reaction assay for rapid detection and identification of *Escherichia coli* O157:H7 in foods and bovine feces. *J Food Prot.* 2000; 63:1032–1037. [PubMed: 10945576]
13. Friedman DI, Court DL. Bacteriophage lambda: Alive and well and still doing its thing. *Curr Opin Microbiol.* 2001; 4:201–207. [PubMed: 11282477]
14. Gautam R, Bani-Yaghoob M, Neill WH, Dopfer D, Kaspar C, Ivanek R. Modeling the effect of seasonal variation in ambient temperature on the transmission dynamics of a pathogen with a free-living stage: Example of *Escherichia coli* O157:H7 in a dairy herd. *Prev Vet Med.* 2011; 102:10–21. [PubMed: 21764472]

15. Gentleman RC, Lawless JF, Lindsey JC, Yan P. Multi-state Markov models for analysing incomplete disease history data with illustrations for HIV disease. *Stat Med*. 1994; 13:805–821. [PubMed: 7914028]
16. Gonzales TK, Kulow M, Park D, Kaspar CW, Anklam KS, Pertzborn KM, Kerrish KD, Ivanek R, Dopfer D. A high-throughput open-array qPCR gene panel to identify, virulotype, and subtype O157 and non-O157 enterohemorrhagic *Escherichia coli*. *Mol Cell Probes*. 2011; 25:222–230. [PubMed: 21925264]
17. Hancock DD, Besser TE, Kinsel ML, Tarr PI, Rice DH, Paros MG. The prevalence of *Escherichia coli* O157:H7 in dairy and beef cattle in Washington State. *Epidemiol Infect*. 1994; 113:199–207. [PubMed: 7925659]
18. Ivanek R, Grohn YT, Ho AJ, Wiedmann M. Markov chain approach to analyze the dynamics of pathogen fecal shedding – example of *Listeria monocytogenes* shedding in a herd of dairy cattle. *J Theor Biol*. 2007; 245:44–58. [PubMed: 17092523]
19. Ivanek R, Osterberg J, Gautam R, Sternberg RL. Salmonella fecal shedding and immune responses are dose- and serotype-dependent in pigs. *PLoS One*. 2012; 7:e34660. [PubMed: 22523553]
20. Jackson CH. Multi-state models for panel data: The msm package for R. *J Stat Software*. 2011; 38:1–28. Available at <http://www.jstatsoft.org/v38/i08/paper>.
21. Jackson MP. Structure-function analyses of Shiga toxin and the Shiga-like toxins. *Microb Pathog*. 1990; 8:235–242. [PubMed: 2200941]
22. Jeong KC, Kang MY, Kang J, Baumler DJ, Kaspar CW. Reduction of *Escherichia coli* O157:H7 shedding in cattle by addition of chitosan microparticles to feed. *Appl Environ Microbiol*. 77:2611–2616. [PubMed: 21335379]
23. Kay R. A Markov model for analysing cancer markers and disease states in survival studies. *Biometrics*. 1986; 42:855–865. [PubMed: 2434150]
24. Khaita ML, Smith DR, Stoner JA, Parkhurst AM, Hinkley S, Klopfenstein TJ, Moxley RA. Incidence, duration, and prevalence of *Escherichia coli* O157:H7 fecal shedding by feedlot cattle during the finishing period. *J Food Prot*. 2003; 66:1972–1977. [PubMed: 14627271]
25. Laegreid WW, Elder RO, Keen JE. Prevalence of *Escherichia coli* O157:H7 in range beef calves at weaning. *Epidemiol Infect*. 1999; 123:291–298. [PubMed: 10579450]
26. Lim JY, Sheng H, Seo KS, Park YH, Hovde CJ. Characterization of an *Escherichia coli* O157:H7 plasmid O157 deletion mutant and its survival and persistence in cattle. *Appl Environ Microbiol*. 2007; 73:2037–2047. [PubMed: 17277224]
27. Lim JY, Yoon J, Hovde CJ. A brief overview of *Escherichia coli* O157:H7 and its plasmid O157. *J Microbiol Biotechnol*. 20:5–14. [PubMed: 20134227]
28. McGee P, Scott L, Sheridan JJ, Earley B, Leonard N. Horizontal transmission of *Escherichia coli* O157:H7 during cattle housing. *J Food Prot*. 2004; 67:2651–2656. [PubMed: 15633666]
29. McKay JR, Franklin TR, Patapis N, Lynch KG. Conceptual, methodological, and analytical issues in the study of relapse. *Clin Psychol Rev*. 2006; 26:109–127. [PubMed: 16371242]
30. Meira-Machado L, Una-Alvarez JD, Cadarso-Suarez C, Andersen P. Multi-state models for the analysis of time-to-event data. *Stat Methods Med Res*. 2009; 18:195–222. [PubMed: 18562394]
31. NRC. Nutrient Requirement of Dairy Cattle. 7. National Academies Press; Washington, DC: 2001. Revised ed
32. Perna NT, Plunkett G, Burland V, Mau B, Glasner JD, Rose DJ, Mayhew GF, Evans PS, Gregor J, Kirkpatrick HA, Posfai G, Hackett J, Klink S, Boutin A, Shao Y, Miller L, Grotbeck EJ, Davis NW, Lim A, Dimalanta ET, Potamousis KD, Apodaca J, Anantharaman TS, Lin J, Yen G, Schwartz DC, Welch RA, Blattner FR. Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7. *Nature*. 2001; 409:529–533. [PubMed: 11206551]
33. Robinson SE, Wright EJ, Hart CA, Bennett M, French NP. Intermittent and persistent shedding of *Escherichia coli* O157 in cohorts of naturally infected calves. *J Appl Microbiol*. 2004; 97:1045–1053. [PubMed: 15479421]
34. Sanderson MW, Besser TE, Gay JM, Gay CC, Hancock DD. Fecal *Escherichia coli* O157:H7 shedding patterns of orally inoculated calves. *Vet Microbiol*. 1999; 69:199–205. [PubMed: 10512044]

35. Sanderson MW, Sargeant JM, Nagaraja TG. Effect of pooling bovine fecal samples on the sensitivity of detection of *E. coli* O157:H7. *Vet Microbiol.* 2005; 110:125–130. [PubMed: 16140477]
36. Schouten JM, Graat EA, Frankena K, Van Zijderveld F, De Jong MC. Transmission and quantification of verocytotoxin-producing *Escherichia coli* O157 in dairy cattle and calves. *Epidemiol Infect.* 2009; 137:114–123. [PubMed: 18346284]
37. Shere JA, Bartlett KJ, Kaspar CW. Longitudinal study of *Escherichia coli* O157:H7 dissemination on four dairy farms in Wisconsin. *Appl Environ Microbiol.* 1998; 64:1390–1399. [PubMed: 9546176]
38. Shere JA, Kaspar CW, Bartlett KJ, Linden SE, Norell B, Francey S, Schaefer DM. Shedding of *Escherichia coli* O157:H7 in dairy cattle housed in a confined environment following waterborne inoculation. *Appl Environ Microbiol.* 2002; 68:1947–1954. [PubMed: 11916717]
39. Todar, K. [accessed 20 July 2011] Pathogenic *E. coli* in Todar's Online Textbook of Bacteriology. 2011. Available at <http://www.textbookofbacteriology.net/e.coli.html>
40. Valcour JE, Michel P, McEwen SA, Wilson JB. Associations between indicators of livestock farming intensity and incidence of human Shiga toxin-producing *Escherichia coli* infection. *Emerg Infect Dis.* 2002; 8:252–257. [PubMed: 11927021]
41. Vilde DA, Larzabal M, Garbaccio S, Gammella M, Rabinovitz BC, Elizondo AM, Cantet RJ, Delgado F, Meikle V, Cataldi A, Mercado EC. Reduced faecal shedding of *Escherichia coli* O157:H7 in cattle following systemic vaccination with gamma-intimin C and EspB proteins. *Vaccine.* 2011; 29:3962–3968. [PubMed: 21477674]

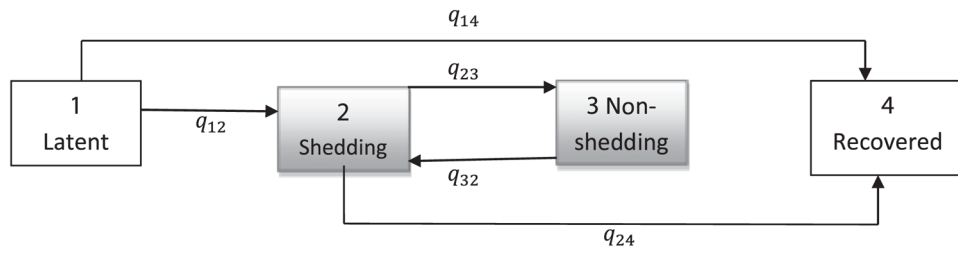


Figure 1.

A flow diagram of a multistate model describing the states and allowed transitions for faecal shedding of ECO157. The four possible states ($i = 1,2,3,4$) correspond to the latent, shedding, non-shedding and recovered states. Coefficients (q_{ij}) represent the instantaneous transition intensities for a steer from state to state.

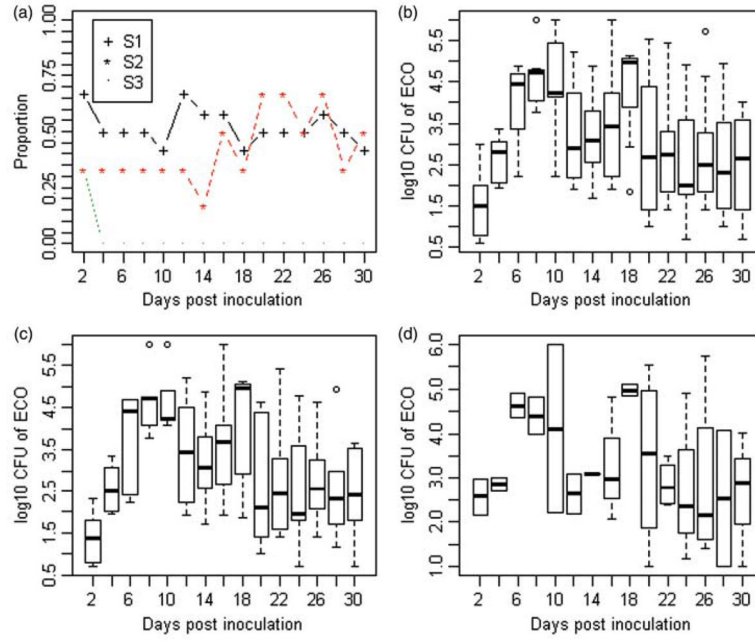


Figure 2.

Time series of the proportion of steers shedding ECO157 post inoculation by strain (a), and box-plots showing the detected ECO157 levels (CFUs) per day among all three strain groups of steers (b), and separately among S1 (c) and S2 strain groups (d). In boxplots in b–d, upper and lower ends of the box represent the IQR and the thick horizontal line inside the box indicates the median, while the whiskers and circles indicate 95% CIs and minimum/maximum values, respectively.

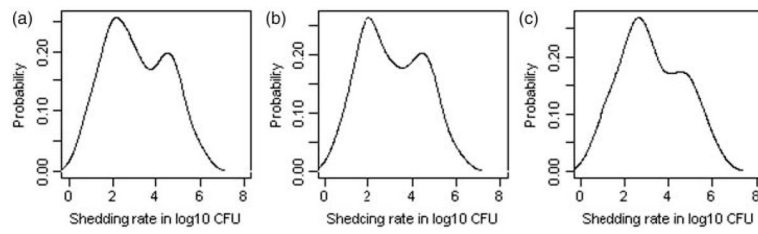


Figure 3. The distribution of ECO157 levels (CFU) detected in RAMS samples in all strain groups (a), S1 strain group (b) and S2 strain group (c).

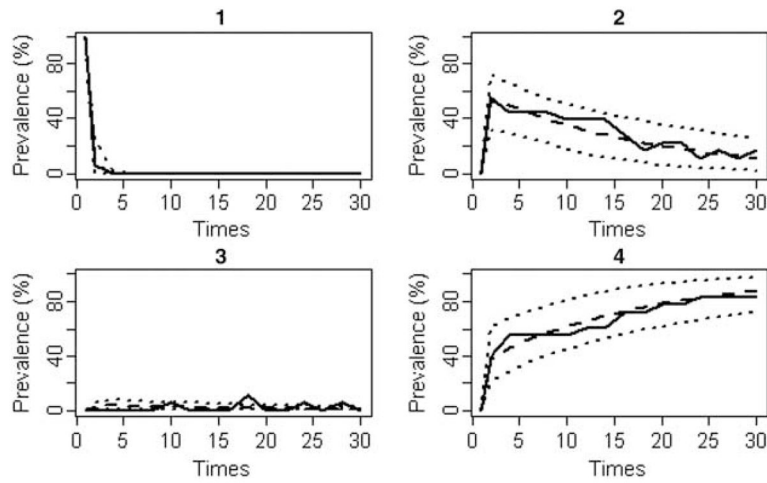


Figure 4. Observed versus model predicted prevalences with 95% CI of steers in different states of faecal shedding post inoculation. The solid line represents the observed prevalence, dashed line is the mean expected prevalence and the two dotted lines represent upper and lower bounds of the 95% CI.

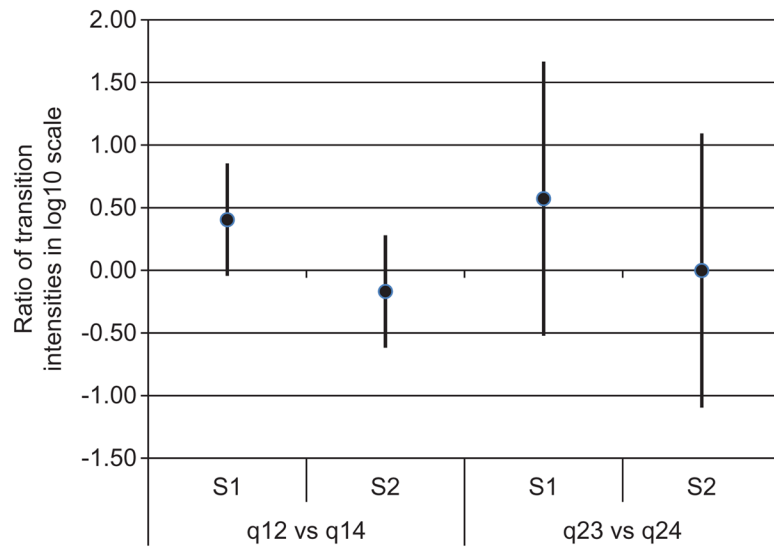


Figure 5. Mean ratios of transition intensities out of a state with the corresponding 95% CIs in log10 scale.

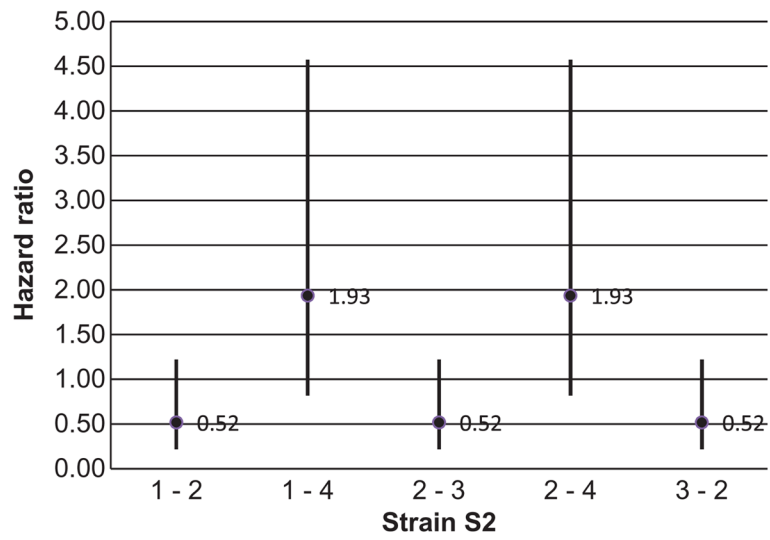


Figure 6. HRs with 95% CIs showing the strain effect (S2 compared to S1) on the model transition intensities. States 1, 2, 3 and 4 correspond to the latent, shedding, non-shedding and recovery states, respectively.

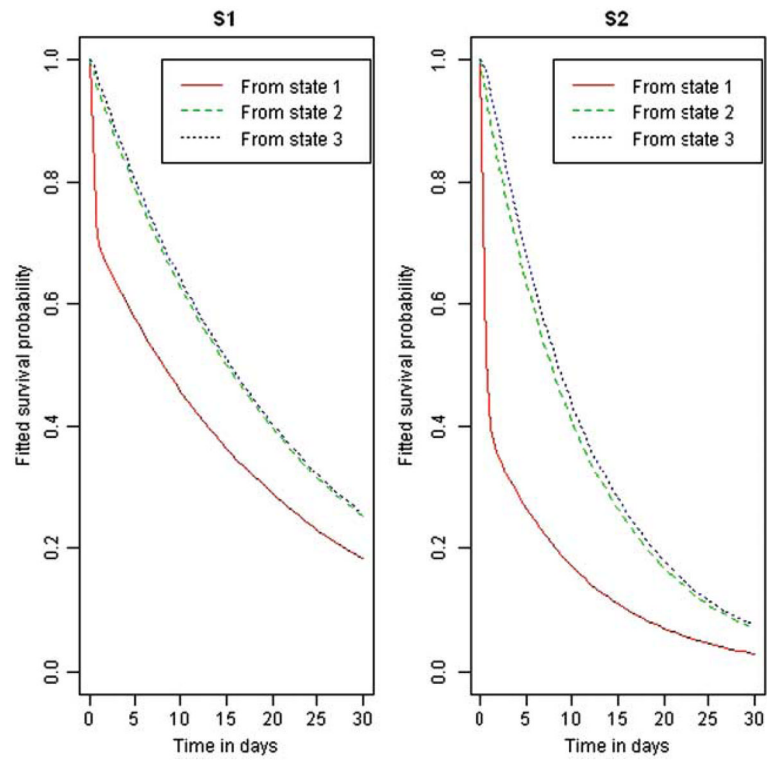


Figure 7.

Expected probability of remaining in the latent (state 1), shedding (state 2) and non-shedding (state 3) states of the faecal shedding model as opposed to entering the recovered state for strains S1 and S2.

Table 1

Summary of the observed number of transitions between states.

From	To			
	1	2	3	4
1	1	11	0	7
2	0	64	5	8
3	0	5	0	0
4	0	0	0	168

Table 2

Average sojourn and total time of stay in each state in days and the associated 95% CIs for ECO157 strains S1 and S2.

Strains	States	Sojourn time (95% CI)	Total time (95% CI)
S1	Latent	0.3 (0.2, 0.6)	0.36 (0.2, 0.6)
	Shedding	4.3 (0.4, 17.2)	14.6 (5.7, 33.2)
	Non-shedding	0.4 (0.03, 4)	1.2 (0.3, 4.4)
S2	Latent	0.4 (0.2, 0.7)	0.32 (0.2, 0.7)
	Shedding	5.3 (0.9, 12.8)	4.2 (0.8, 16.3)
	Non-shedding	0.7 (0.1, 8.2)	0.25 (0.02, 3)

Note: The CIs were calculated by simulating 1000 random vectors from the asymptotic multivariate normal distribution implied by the maximum likelihood estimates (and covariance matrix) of the log transition intensities and covariate (strain) effect and then back transforming the estimates.