

Simulation modelling of *Mycobacterium avium* subspecies *paratuberculosis* in dairy systems – Final report

I.J. McKendrick, J. Woods, A. Greig, M.R. Hutchings

This is a report for Objective 4 of the SEERAD Flexible Fund (SAC/827/97 – 661601) entitled Role of wildlife in the epidemiology of paratuberculosis of farmed ruminants.

ParaTB FF Report: Objective 4

2. Executive Summary

Mathematical modelling of epidemic systems is a useful tool to explore infection systems with properties which are difficult to observe in the field. Johne's disease, having a long and variable incubation period, during which animals may actively shed bacteria, falls into this category. However, difficulties in observing the progression of Johne's disease in animals result in much uncertainty about the values of key epidemiological parameters.

Johne's disease has a number of properties which would suggest that standard epidemic models would be likely to give inaccurate or even misleading results. The long time period typically seen between infection and signs of clinical disease will be poorly modelled by the standard exponential distribution, while the role of environmental contamination is of vital importance in maintaining and increasing on-farm prevalence.

The primary goal of the objective was to create a mathematical model which describes the important epidemiological features of paratuberculosis infection in an appropriate fashion. This goal was met through the creation of an integrated suite of stochastic models describing the management of a typical Scottish dairy farm, the infection dynamics of paratuberculosis in cattle, the growth and decay of environmental infection, the infection dynamics of paratuberculosis in rabbits, the ability of faecal culture and ELISA tests to detect infection and the effects of vaccination on paratuberculosis infection in cattle.

Uncertainty in parameter estimates was handled through the use of Latin Hypercube Sampling to generate random combinations of plausible parameters. A new methodology was used to weight the outcomes of the resulting simulations to observed field data, and the resulting data were analysed using cutting edge statistical methods.

Many control options were evaluated using this model. Test and cull control policies based on ELISA testing appear to be only slightly less effective than those based on faecal culture testing, an important result, given the cheaper nature of the former tests. Vaccination alone will generate a drop in average prevalence, but will not suffice to reduce the average prevalence to a low level. Management changes alone will generate a drop in average prevalence, but will not suffice to reduce the average prevalence to a low level. ELISA or faecal culture testing and culling will reduce the average prevalence to a low level, but a combination of vaccination or test and cull with management changes is required to reduce the average prevalence to a negligible level. The model suggested that the absence or presence of shedding rabbits is likely to have no appreciable effect on the relative efficacy of the available control options, but this conclusion is extremely sensitive to model assumptions about the avoidance or otherwise of rabbit faeces by cattle. It is important that the nature of rabbit-cattle interactions be thoroughly investigated: in the meantime it is necessary to treat the conclusions of the model with caution where they relate to rabbit control. Most of the control measures exhibit high variability in outcome. Only options involving test and cull do not present a large proportion of scenarios in which the prevalence increases, and only those involving test and cull and management changes can guarantee to reduce the prevalence to negligible levels within 10 years.

3. Introduction to the Project and Background Information

3.1 Overall aims

The aim of this project is to develop an appropriate mathematical model to describe the epidemiology of paratuberculosis infection in cattle which will allow the evaluation of those control methods most frequently proposed in practice. No published models as yet achieve this objective. Once this model has been developed, it will be integrated into a wider model to explore the effect of paratuberculosis infection in rabbits on the epidemiology of the disease in cattle and on the appropriateness of various control strategies.

3.2 Staff Employed on Project

Dr Iain J. McKendrick (Band 5 SPD)
Miss Joanna Wood (Band 6)

3.3 Timetable and Milestones

Task 4a	BioSS	Develop initial model of paratuberculosis dynamics within a herd of cattle
Task 4b	BioSS	Select appropriate mathematical framework for further development of model to incorporate infection in rabbits
Task 4c	BioSS	Develop model of paratuberculosis dynamics within a system comprising cattle and infected rabbits
Task 4d	BioSS	Revise model to incorporate at least one further host species
Task 4e	BioSS	Use the final model to investigate the theoretical efficacy of various control strategies

Initial Timetable:

Task	1998				1999				2000			
	JFM	AMJ	JAS	OND	JFM	AMJ	JAS	OND	JFM	AMJ	JAS	OND
4a	XXX	XXX	XXX									
4b				XXX								
4c					XXX	XXX	XXX	XXX				
4d									XXX	XXX		
4e											XXX	XXX

Revised Timetable (October 2000):

Task	2000				2001				2002			
	JFM	AMJ	JAS	OND	JFM	AMJ	JAS	OND	JFM	AMJ	JAS	OND
4a				XXX								
4b				XXX								
4c					XXX	XXX						
4d							XXX					
4e								XXX	XXX			

4. Results

4.1.1 Materials and Methods

Understanding the epidemiology of paratuberculosis is non-trivial. The infection has properties which make the construction of a coherent picture of an outbreak or experiment difficult, and which will invalidate many of the standard assumptions of epidemic modelling. In particular, the infection:

- Operates over a long timescale. There is potentially a very long gap between animals being infected and showing clinical signs, and a long gap even between the infection time and the time at which appreciable shedding begins. It is therefore sensible explicitly to model the process of infection in the animal.
- Is likely to be very variable in its impact on different animals. Analysis of experimental data shows that the time between infection and the development of clinical signs is not just long, but also highly variable. Stochasticity is important for two reasons. Firstly, the system under study is highly non-linear, and the study of a deterministic version of the process will not necessarily reflect even the average behaviour of the equivalent stochastic system. Secondly, the stochastic effect of the long and variable transition distribution is likely to be amplified by the interplay between infected animals and the farm management system. Many infected animals in a dairy herd will never develop clinical signs and (in the normal course of affairs) be identified as infected, simply because they will have been removed from the herd for other reasons before any such signs would become evident. It is possible that the differences in prevalence which are seen between dairy and beef farms may be at least partially explained by the differences in management system seen on each class of farm. In this context, the development of a

realistic model for the management system and the use of realistic transition distributions are essential to the production of a realistic model, since otherwise a model is unlikely to realistically describe the proportion of infected animals which ultimately contribute to further infection in the herd. The time each animal spends in the intermediate class(es) between infection and developing clinical signs is likely to be long, have a non-zero mode and be highly variable.

- Is likely to be very variable in its impact both within and between different herd realisations. As discussed above, the epidemiological history of individual infected animals are likely to be very variable, and this variability is likely to be further accentuated by the interplay between animal and management system. Where the system has not reached a non-trivial equilibrium distribution (ie, it has non-zero numbers of both susceptible and infected animals), this is likely to result in a quasi-equilibrium distribution with a high variance. Not only is it essential that this variability is fully integrated into the model through the use of an appropriate stochastic simulation, coupled with a suitably high number of replications (runs with the same initial conditions and parameters, but evolving with a different random epidemiological history), but it is vital that the output be analysed using appropriate statistical methods which allow for the different sources of variability at the within and between realisation strata.
- Is difficult to observe. A further by-product of the long transition distributions and possible environmental routes of infection is the difficulty scientists find in inferring pathways of infection, and hence information about the epidemiology of the infection. Hence, it is unclear whether or not adults can be infected, and what the likely balance is between direct and environmental infection. This project assumes that in the absence of clear consensus, it is necessary to allow animals to become infected via all plausible routes.
- Has properties which are non-stationary with respect to time. Once an animal is infected, shedding increases as a function of time, increasing the potential for both direct and indirect infection of other animals. This project attempts to use longitudinal shedding data and expert opinion to model the growth of shedding in each animal, allowing the parameterisation of average shedding levels to be consistent across different stages of infection.
- Can pass from dam to calf. There is clear evidence that animals can be infected *in-utero* and are at higher risk as young calves when suckling from an infected dam. Hence, the system is non-homogeneous, in that calves are exposed to different risks of infection depending on whether the dam is infectious. The proportion of calves affected in this way is likely to be large, and control measures which are efficacious in a homogeneous epidemic are likely to be less effective when applied to a highly inhomogeneous system. Conversely, some strategies, such as slaughtering on the basis of cosanguinity, only make sense when this vertical transmission is modelled explicitly. The model described in this report models the population structure explicitly, defining appropriate risk levels for the offspring of infected dams and allowing testing and culling strategies based on cosanguinity to be tested.
- Has an epidemiology which involves unquantified relationships with the environment and wildlife reservoirs such as rabbits. Earlier research has indicated that *M. a. paratuberculosis* can be isolated from rabbits on Johne's infected farms in Scotland (Greig *et al*, 1999). Other research has indicated that the organism can survive on pasture for an appreciable length of time (Lovell *et al*, 1944). Advice to farmers to reduce the risk of infection through the sequential use of calving pens (Wren, 1998) or the practice of destocking pasture on infected farms for a number of years (Bölske *et al*, 1999) clearly imply that it is recognised that there is an appreciable risk of mechanical transmission of infection, probably via the faecal-oral route, from a local environment contaminated with bacteria. Unfortunately, the nature of these interactions make them both difficult to quantify and also potentially highly variable both between different farms and on the same farm in different places or at different times. Nevertheless, the model described in this report will explicitly model a relationship between cattle, the local environment and the rabbit population. It is believed that the methodology described below will remove many of the problems associated with uncertainty and variability in model parameters.

Issues of uncertainty and variability are fundamental to the project. The model of Collins and Morgan (1991) is in many respects a theoretical exercise which aims only to explore, in rough

terms, the likely epidemiology of Johne's disease in cattle. Hence, parameterisation is not really an issue in this paper. Humphry *et al* (2001), by contrast, go to considerable lengths to produce a (single) set of parameters which can then be used in their model. The uncertainty which attaches to many areas of paratuberculosis epidemiology means that the value of many of the parameters required in a model will be, at best, poorly estimated; at worst, totally unknown. In this context, the desire to identify a canonical parameterisation is unwise, since inferences drawn from a single choice of parameters will be of dubious validity. The situation is complicated further by what is likely to be a large between farm variability in many of the parameters. Parameters which describe such processes as the death rate of *M. a. paratuberculosis* in pasture, or the interactions between levels of bacterial contamination with animal-level force of infection, are likely to vary greatly between different farms, depending on difficult to quantify factors such as soil acidity, micro-climate and the details of management practice on the farm. Even if it were possible to identify a 'typical' value for each parameter, there is no guarantee that the combination of parameter values thus selected would also be in any way typical. Even if it were possible to identify a 'typical' parameter set, inferences about the suitability of different control measures based only on a review of this properties of this system would be fatally flawed. There is no reason to believe that farms, subject to different local epidemiologies, will all behave in an identical fashion when these epidemiologies are disrupted through the introduction of some control policy.

A methodological approach which avoids many of the pitfalls described above is that of Latin Hypercube Sampling (LHS), (Vose, 2001). Where a parameter is unknown or uncertain, expert opinion and published estimates of the value will be used to define an appropriate range and distribution for the value. These distributions will summarize two different sources of variability, namely genuine between farm variability and our uncertainty in assessing the true values of these parameters. We wish to integrate information over these distributions, since in the former case, it reflects true variability which we wish to factor into our model, and in the latter case because we have no choice. Latin Hypercube sampling is used to sample parameter combinations since the parameter space which has to be sampled is high-dimensional, and the number of parameter combinations has to be limited due to the model running time. This sampling method generates a random sample from the parameter space, but also ensures that the distributions of the parameters are reproduced with great efficiency through its use of stratified sampling without replacement (Iman and Conover, 1980). A further technique, proposed by Iman and Conover (1982), was used which induces a specified rank correlation between parameters. In this project, these correlations were specified to be zero.

In the first instance, it is proposed to sample 2075 parameter combinations. The set of parameter combinations selected is defined as Θ . The resulting realisations of the process where it is simulated without control measures are representative of the parameter space chosen, but are unlikely to be representative of the pattern of infections seen in the field. There is no reason why this should be so: the proposal distributions for many of the parameters contain a substantial element of uncertainty, and hence will have generated many parameter values which are atypical of those actually present in the field. Similarly, because of our lack of knowledge, realisations of particular parameters are likely to have been combined into combinations with a low probability of occurring in practice. This problem can be overcome through the use of a reference dataset, summarising the distribution of prevalences seen in the field.

The best dataset identified for this purpose is that presented in Boelaert *et al*, 1999, describing the Belgian cattle population. In this paper, the authors present the quartiles of the within-herd prevalence distribution for (amongst others) a sample of dairy farms. In this study, each target farm had all adult (> 24 month) cattle sampled. Hence, there is no sampling variability in the distributions presented in the paper. However, the lack of sensitivity of ELISA methods against *M. a. paratuberculosis* will introduce test variability into the sampling distribution. No comparable datasets for the UK prevalence were identified. Work such as that of Cetinkaya *et al* (1994) or Cetinkaya *et al* (1998) were not felt to be useful for this exercise: the sampling scheme in these studies was self selecting, with a commensurate potential for large bias.

To make use of the reference data to quantify the relevance of a particular model realisation, the program is required to model the testing of each infected animal in turn (we assume that the specificity of the test is sufficiently close to 1 that there is no need to model the testing of non-infected animals). Given the model of the test which is being used, it is possible to calculate the mean and variance of the number of animals which would be identified as positive. However, the exact distribution is more difficult to determine numerically. For large mean values, the Central Limit Theorem would justify the use of a normal approximation to the true distribution. Given that the objective of this exercise is to give an approximate weighting to different simulation outcomes, it is felt that it will be reasonable to use the normal approximation regardless of the size of the expected value.

The fractional sero-prevalence is defined as the ratio of the number of animals observed as positive to the total number in the herd. Boelaert *et al* (1999) allow us to define q_0, q_1, q_2, q_3, q_4 , the quartiles of the observed within-herd sero-prevalence distribution. q_0 , the minimum sero-prevalence reported on a positive farm, is, however, not especially helpful, since it almost certainly reflects the observation of a single positive animal on a farm of size between 118 and 133 animals. For the purposes of this exercise, we therefore take q_0 to be $1/N_a$, the minimum fractional sero-prevalence possible given the size of the simulated herd, N_a . The variability in the sero-prevalence distribution contains an element caused by the variation in herd size (and hence in the denominator of the expression). However, it is hoped that the large, quartile-based group sizes in the weighting scheme will ensure that most of the differences in fractional sero-prevalences between different weighting groups will reflect real differences in prevalence. From each simulated replicate of an infected herd with parameters θ , the probability of the observed sero-prevalence falling into each observed quartile is calculated. These are used to calculate weights for each combination of parameters θ . The weights across the parameter space sum to one, and each combination of parameters θ is given a weight commensurate with the observed distribution of prevalences and the number of other simulation runs which gave rise to similar results. Conceptually, where parameter sets give rise to results which are similar to those generated by many other parameter sets, the weighting will be low, whereas a combination which gives rise to results which are seen in the field, but which are rarely generated by the model will be given a high weighting. Parameter combinations which consistently generate prevalences greater than those reported from the field will be given zero weights. It may also be appropriate to analyse the simulation results using weighting schemes which concentrate on the lowest and highest quartiles, to explore the use of control schemes in lightly and heavily infected herds.

Hence, the simulation process is carried out over the set of parameter combinations Θ , and the resulting model sero-prevalences are used to derive weights for each run. Results from simulations across an arbitrary parameter space, once weighted, should be representative of the population of farms included in the survey. Once the weights have been calculated, the simulation model can be used to generate summary prevalence information for various control methods. These prevalences are analysed relative to their baseline prevalence pre-control using appropriate statistical methods, taking due account of the assigned weights. These involve the use of weighted Generalised Linear Mixed Models (Brown and Prescott, 1999), with the variance components in the mixed model being used to separate variability due to intrinsic simulation variability and that due to parameter uncertainty. Where appropriate, the comparison of different control methods is also carried out using weighting schemes which emphasise either low or high prevalence farms.

The output from the baseline simulations is also analysed, without the use of weights, to explore the sensitivity of the model to variability in the parameter estimates. This process proceeds through the mapping of variables into factors with levels containing equal numbers of observations and the use of linear modeling techniques to explore the ability of different factors to explain the observed variability in the data.

The overall Johne's model has 5 separate components, each of which has its own mathematical characteristics. These mathematical structures have been chosen to be the

simplest which will meet the purposes of the study while being capable of being linked together. These are the

- Dairy Herd Management Model
- Cattle Infection Model
- Environment Infection Model
- Wildlife Reservoir Infection Model
- Animal Testing Model.

Several of the model components have additional functionalities built in to allow the modeling of specific control methods. These will be described in more detail below.

4.1.2. Task 4a: Develop initial model of paratuberculosis dynamics within a herd of cattle.

It is useful to consider this model as comprising two sub-components, namely a model for cattle management and a model for cattle infection. It is also appropriate to consider the model of environmental infection as being a necessary component of this milestone.

Dairy Herd Model

Information about individual cattle is stored within a computer program, defining age and management status. Transitions between different management states are specified by random variables or fixed parameters, depending on the nature of the transition. This detailed model is necessary because it allows records to be kept of dam-offspring relationships and hence the modeling of vertical transmission of infection to the offspring of infected animals and of plausible control methods such as slaughter of offspring of infected animals. Animal management status will be updated for each animal on a discrete-time basis, with a time step of one week. The herd is assumed to be closed (i.e., animals are not introduced to the farm), since control methods which are found to be appropriate in the closed case will still be useful in the open herd situation when coupled with sensible purchasing strategies to minimize the risk of importing infection into the herd. The converse is not true, since the long-term epidemiological state of an open herd will be critically driven by the level of infection imported into the herd. It will be assumed that a single infected animal is imported into the farm, in order to introduce the infection, after which the epidemic will be allowed to progress to equilibrium.

It is assumed that the model farm follows a spread calving pattern rather than either of the likely seasonal calving strategies. Statistical evidence (Anon, 2000) suggests that a herd size with $N=90$ milking animals can be used as being typical of the Scottish herd population as a whole.

It is assumed that animals remain pregnant for 10 months, following which they are left empty for 4 months (Castle and Watkins, 1984) before becoming pregnant again. Reproductive failure is not modeled in this exercise, with culling due to reproductive failure being absorbed into higher culling rates for pregnant animals. The culling strategy on a farm is likely to have a large impact on the development of paratuberculosis in the herd. Unfortunately, there is no clear picture of the typical time at which animals are assumed to have reached the end of their productive lives and culled. This is therefore treated as a variable parameter in the model, taking values between 5 and 12 lactations.

It is assumed that 50% of births are male. Furthermore, it is assumed that there is a zero probability of twin calves being born. All male calves are culled from the farm sufficiently soon after birth for them to have no impact on the local epidemiology of the disease. Furthermore, it is assumed that 8% of female calves are born dead or abnormal (Etgen and Reaves, 1978).

Involuntary culling rates are assumed to vary with respect to age and pregnancy status. For calves up to age 6 months we assume a mortality rate of μ_c of 0.00855 per animal per month,

which corresponds to a total mortality of 5% over this period as specified by the Milk Development Council (Anon, 1998). Heifers aged 6 months to 16 months are assigned an instantaneous mortality rate μ_h of 0.001122 per animal per month as derived from Etgen and Reaves (1978). In adult animals, mortality is described by a simple memoryless model with appropriate piecewise constant mortality rates, coupled with a final cull when animals reach a critical lactation number. The standard mortality rate is calculated using data from the Economic Report on Scottish Agriculture (Anon, 2000), but the value depends on the lactation number at which all animals are culled. For each farm, the annual non-age-related cull fraction q is calculated and used as the basis for other mortalities. The instantaneous standard mortality applying to each adult, non-pregnant animal is estimated as $-\ln(1-0.708q)/12$ per animal per month, derived from data collected by a Kingshay Farming Trust survey (Anon, Unspecified). A higher total instantaneous mortality for pregnant animals is estimated of $-\ln(1-0.708q)/12-2704.\ln(1-0.292q)/(12.1920)$ per animal per month.

It is assumed that the farmer wishes to keep his spread calving pattern intact, and will select a replacement animal from the pool of animals aged between 14 and 16 months, aiming for her to calf around the 24 month period. The model selects an animal randomly from the pool of available animals in the correct age range. It is recognized that the choice of an individual animal will not be random, but will depend on breeding merit, so in this situation the choice of a random animal reflects the model's lack of knowledge of this aspect of the herd, rather than a real stochastic effect. Animals which do not enter the herd by age 16 months are culled (Castle and Watkins, 1984, and Etgen and Reaves, 1978).

Cattle Infection Model

Information about individual cattle is stored within the program, defining the infection status. Transitions between different infection states are specified by random variables derived from the appropriate statistical distribution. Animal infection status is updated for each animal on a discrete-time basis, with a time step of one week.

It is assumed that cattle become infected with *M. a. paratuberculosis* through one of three routes:

- 1/ Infection *in-utero* (or immediate post-natal) from an infected dam.
- 2/ Animal to animal contact with an infectious animal.
- 3/ Contact with a contaminated environment.

Animals, once infected, enter a sub-clinical state. After a random time, modeled using an appropriate gamma distribution, the animal makes a transition to a clinical state, at which time it sheds bacteria at a high level α_{\max} . While in the sub-clinical state, it progresses through three sub-states, exhibiting progressively higher levels of shedding. Once in the clinical state, after an exponentially distributed period of time of mean $1/\mu_p$ the animal is culled. The shedding rate for infected calves is assumed to be zero, i.e. the model assumes calves will have become adults before shedding starts. The shedding rates are parameterised assuming an underlying exponential growth process fitting the data presented in Whitlock *et al* (1994). The model assumes that at the end of the latent period, each infected animal reaches the critical final shedding level. Where the latent period has been long, this implies that the build-up of shedding has been slow. This final shedding level α_{\max} is the key shedding parameter, which is explored using Latin Hypercube Sampling.

It is assumed that the relative risk of the different infection routes varies depending on the age of the animal. It is assumed that young calves (< 6 months of age) have a different, and probably higher, risk of being infected when in a contaminated environment.

Infection via the first route is simple to model. Where a calf is born to an infected dam, and the calf is retained in the herd, the calf is deemed to be sub-clinically infected with probability $P_{d \rightarrow c}$. The time which the animal remains in the sub-clinical class is modeled using a gamma distribution with a smaller mean than that for animals infected later in life.

Animal to animal infection can, in principle, be modeled using standard epidemic model simulation techniques to determine the time at which, everything else remaining equal, a given animal will be infected by another animal. However, the modeling of environmentally mediated infection is more challenging, and it is convenient to absorb the simulation of direct animal to animal infection events into the more general process of simulating all infection events.

Environmental Infection

Environmental contamination with *M. a. paratuberculosis* is modeled using a deterministic ordinary differential equation which is linear with respect to the infected cattle populations, infected rabbit population and any other infected wildlife population and which has a linear removal rate. Conditional on the state of the stochastic process, this system is therefore locally a linear ODE with constant coefficients which can be solved analytically.

The force of infection driven by a contaminated environment is an abstract concept designed to allow us to explore the potential effect of environmental persistence on the survival and prevalence of Johne's disease in a dairy herd. The random function $\mathbf{c}(t)$ summarizes the quantity of infective material present on the farm. The infective impact which this might have on individual animals is difficult to model on the micro-level, depending as it will on the distribution of infection on the farm, the feeding and contact patterns of the animals involved and the nature of any dose-response relationship which might be present. In so far as these factors depend on, as yet, unknown scientific results (dose-responses) or on an extremely detailed model of animal behaviour which will depend critically on the situation on a specific farm, it is more useful to treat the problem at a more abstract level. Assuming that a given level of contamination $\mathbf{c}(t)$ will have a specific and unique impact on the force of infection of each animal, and that this relationship is summarised via an arbitrary function $f(\mathbf{c})$, we can make certain statements about the function f .

1. $f(0)=0$. Where no contamination is present, there should be no impact on the total force of infection.
2. $f'(\mathbf{c}) \geq 0 \quad \forall \mathbf{c} \geq 0$. The first derivative with respect to time of f must be non-negative for all values of \mathbf{c} , ensuring that f is monotonically increasing for all values of \mathbf{c} . Intuitively, higher values of \mathbf{c} must equate to a non-decreasing contribution to the total force of infection.

It is possible to add a further restriction on f , although this restriction is more speculative. It is, however, very important in ensuring that the resulting model gives rise to results in line with those seen in the field.

3. $\lim_{\mathbf{c} \rightarrow \infty} f(\mathbf{c}) < \infty$. The force of infection does not increase without bound as \mathbf{c} increases.

Without loss of generality, we can therefore scale $f(\mathbf{c})$ such that

$$\lim_{\mathbf{c} \rightarrow \infty} f(\mathbf{c}) = 1.$$

In intuitive terms, after a certain level is reached, further increases in the levels of contamination are unlikely to produce appreciable increases in the rates at which animals become infected. In practical terms, this may relate to the nature of interaction between the organism and the animal: e.g. the number of occasions the animal is likely to graze in a contaminated area, or the speed at which it ingests contaminated pasture.

For simulation purposes, a particular choice of f is used. It is felt that, since the parameterization of the link function is by far the most speculative component of the project, it is worthwhile to use a simple link function to facilitate the evaluation of a wider range of parameter combinations using Latin Hypercube Sampling (LHS). It was therefore decided to use a piecewise linear link function of the form

$$f(c) = \begin{cases} 0 & \text{if } c < c_{\min} \\ \frac{c - c_{\min}}{c_{\max} - c_{\min}} & \text{if } c_{\min} \leq c \leq c_{\max} \\ 1 & \text{if } c_{\max} < c \end{cases} .$$

The function is fully parameterised by two quantities, c^{\min} and c^{\max} which denote the points in the function with discontinuities in the first derivative. These are drawn from an appropriate LHS distribution. $f_c(c)$, the link function for environment to calf infection and $f_a(c)$, the link function for environment to adult cattle infection are defined in this manner.

The model describes the properties of a single environmental pool of infection, where new infection arises from shedding from both cattle and rabbits. There is now some evidence that cattle actively avoid contact with cattle faeces but avoid contact with rabbit faeces to a much lesser degree. It was not felt appropriate to model the rabbit-derived and cattle-derived pools of infection separately, although the mathematical structure would have allowed this, because of the dearth of empirical evidence for parameterization of the resulting model. The choice of parameters is likely to be of critical importance in determining the model epidemiology, but there was no suitable data to inform the choice of 'rabbit-faeces' parameters relative to 'cattle-faeces' values. The choice of a single infection pool formulation for the model does, however, have an effect on the conclusions of the study, in that it restricts the range of outcomes arising from rabbit infection. A model in which cattle have a higher exposure to rabbit faeces than cattle faeces would obviously tend to give greater weight to the nature of the infection in the rabbit population.

Susceptible animals are classed in two categories: calves under the age of 6 months and adult animals. This allows calves to have different and potentially higher infection rates. The infection process in cattle is defined by the following instantaneous transition probabilities:

Prob(susceptible adult at time $t \rightarrow$ sub-clinically infected adult at time $t+dt$) =

$$\frac{\beta_d}{N_a} Z_a(t) dt + \frac{\beta_{ai}}{N_a} f_a(c(t)) dt$$

Prob(susceptible calf at time $t \rightarrow$ sub-clinically infected calf at time $t+dt$) = $\frac{\beta_{ci}}{N_c} f_c(c(t)) dt$.

The times for which calves and adults remain in the sub-clinically infected class are modeled as Gamma random variables, where the mean is smaller for animals infected as calves.

Prob(clinically infected animal at time $t \rightarrow$ culled animal at time $t+dt$) = $\mu_p dt$.

N_a and N_c are the total number of adults and calves respectively. A true-mass action formulation is used to define these equations. Note that it is assumed that adults may be infected both through environmental contamination and through direct animal to animal contact. It is assumed that calves may only be infected through environmental contamination. Calf infection from the dam is already modeled using a different mechanism, and it is assumed that calves will have little direct contact with other adult animals, while a negligible number of calves will themselves develop clinical Johne's disease during this critical six month period. The use of a different function f_c for calves allows these animals to be exposed to a higher risk of infection than adults when exposed to the same level of environmental contamination.

The repeated calculation of the exact time to the next infection for each animal is likely to prove computationally expensive. A small extra assumption has the potential to allow much more efficient computation of the model. The standard assumption that animals start to shed immediately on becoming infected by the organism is unlikely to be strictly true, especially for a slowly reproducing organism such as *M. a. paratuberculosis*. The model assumes that sub-

clinically infected animals will only become infectious at the end of the month during which they have been infected, dramatically improving the efficiency of the simulations program.

Once an animal has entered the class of advanced clinical disease, it is assumed that the disease will quickly advance to the stage of death or (more likely) culling. This is largely driven by the speed of response of the farmer and veterinarian. It is assumed that the animal is culled on average 1 week after entering this state. Hence the instantaneous mortality is 4 per animal per month.

The model is parameterised using the highest quality data or information available. Where data is available, either from other objectives in this Flexible Fund project, from other SAC or MRI experiments, or from published papers, they has been reanalysed using appropriate statistical methods to generate estimates of the values of parameters and of the variability of these estimates. Where detailed data are not available, summary statistics have been collated from the published literature, and pooled to give overall estimates. Where no quantitative information is available, qualitative or semi-quantitative statements in the published literature have been used to generate parameter estimates.

4.1.3. Task 4b: Select appropriate mathematical framework for further development of model to incorporate infection in rabbits.

The infected rabbit population should be modeled as a memoryless stochastic process with linear removal rate and an infection rate which is linear with respect to the susceptible rabbit population but non-linear (via an arbitrary non-linear function) with respect to the environmental contamination.

The amount of environmental contamination is specified by the random function $\mathbf{c}(t)$. The evolution of this function is specified by the ordinary differential equation

$$\frac{d\mathbf{c}}{dt} = \alpha_{cs} \mathbf{Y}_c(t) + \alpha_{as} \mathbf{Y}_a(t) + \alpha_{ac} \mathbf{Z}_a(t) + \alpha_r \mathbf{Z}_r(t) + \alpha_f \mathbf{Z}_f(t) - \delta \mathbf{c}$$

where $\mathbf{Y}_c(t)$ are the number of calves infected with paratuberculosis, $\mathbf{Y}_a(t)$ are the number of adult cattle infected with paratuberculosis subclinically, $\mathbf{Z}_a(t)$ are the number of adult cattle infected with clinical paratuberculosis, $\mathbf{Z}_r(t)$ are the number of rabbits infected with paratuberculosis, and $\mathbf{Z}_f(t)$ are the number of foxes infected with paratuberculosis. Each of $\mathbf{Y}_c(t)$, $\mathbf{Y}_a(t)$, $\mathbf{Z}_a(t)$, $\mathbf{Z}_r(t)$, and $\mathbf{Z}_f(t)$ are stochastic random variables, so that $\mathbf{c}(t)$ itself is a random function. The α terms denote the shedding rates of different categories of infected animal. δ denotes the decay rate of the organism in the environment. There is no birth-rate term in the equation since the bacterium is an obligate parasite. However, conditioning on particular values of these random variables, the ODE simplifies to

$$\frac{d\mathbf{c}}{dt} = A - \delta \mathbf{c}$$

with A constant, which has the solution

$$\mathbf{c}(t) = \mathbf{c}(0)e^{-\delta t} + \frac{A}{\delta}(1 - e^{-\delta t})$$

where $\mathbf{c}(0)$ is the initial value of the function. Hence, between discrete stochastic events which change the values of $\mathbf{Y}_c(t)$, $\mathbf{Y}_a(t)$, $\mathbf{Z}_a(t)$, $\mathbf{Z}_r(t)$, and $\mathbf{Z}_f(t)$, $\mathbf{c}(t)$ will evolve deterministically in line with the function. Hence, within the simulation model, given that we know the value of $\mathbf{c}(t)$ when an event occurs which changes any of the above random

variables, we therefore know the value of $\mathbf{c}(t)$, $Y_c(t)$, $Y_a(t)$, $Z_a(t)$, $Z_r(t)$, and $Z_f(t)$ for all future times up until the next stochastic event. This is sufficient information for the simulation process to be fully defined, given that we know the initial state of the system. This framework allows the rabbit infection dynamics to be modeled stochastically, while modeling the effect of rabbit infection on the level of environmental contamination in a computationally efficient manner which integrates fully with the cattle model.

4.1.4. Task 4c: Develop model of paratuberculosis dynamics within a system comprising cattle and infected rabbits.

The rabbit population is modeled as a stochastic process with the following instantaneous transition probabilities:

$$\text{Prob}(\text{susceptible rabbit at time } t \rightarrow \text{infected rabbit at time } t+dt) = \frac{\beta_r}{N_r} f_r(\mathbf{c}(t)) dt$$

$$\text{Prob}(\text{infected rabbit at time } t \rightarrow \text{susceptible rabbit at time } t+dt) = \mu_r dt.$$

These equations should not be interpreted as implying that rabbits recover from *M. a. paratuberculosis* infection. It is assumed that the rabbit population is in equilibrium, with births equally balanced by deaths, so the apparent transition from infected to susceptible rabbit equates to the death through natural causes of an infected rabbit, at the same rate as susceptible rabbits, and, on average, this death being balanced by the introduction and survival in the population of a young, uninfected rabbit. This is a simple model of rabbit population dynamics, but since we have only an abstract concept of the interaction of the rabbit population with the environment and cattle population, it seems unnecessary to incorporate over-much detail about the actual rabbit population. It would only be seriously inappropriate if the rabbit population was at serious risk of dying out spontaneously due to a particularly high death rate relative to the natural birth rate. Should the death rate increase in such a way as part of a system of control measures for Johne's disease, it can easily be modeled in this mathematical structure by reducing the size of N_r , the equilibrium rabbit population.

Rabbits become infected only through contact with a contaminated environment, and immediately pass into a state of maximal shedding. *A priori*, this may be biologically unrealistic, but our lack of knowledge about the *in vivo* dynamics of *M. a. paratuberculosis* indicates the use of simple modeling assumptions. This lack of knowledge also complicates estimation of the shedding rate in rabbits. SAC field data from positive rabbits was analysed to provide an estimate of the mean shedding per gram of faeces. However, nothing is known of the infection status of these animals other than that they are indeed infected. The extremely small sample may therefore be unrepresentative of shedding in the infected population as a whole. To reflect this uncertainty, the estimated log mean shedding is used as the mean of the hypercube distribution, but the standard deviation of the LHS is greatly inflated relative to the standard error of the estimated log mean.

It is appropriate to use a discrete approximation to the stochastic process in the rabbit population, since it is unlikely that animals will immediately become infectious on infection, and it will dramatically increase the efficiency of the model. In a (short) time interval of length Δt , the probability of an individual susceptible becoming infected is approximately

$$1 - \exp\left\{-\frac{\beta_r}{N_r} \int_0^{\Delta t} f_r(\mathbf{c}(t)) dt\right\},$$

while the probability of an individual infective being removed is approximately $1 - e^{-\mu_r T}$. In a short time interval, we can assume that there is a negligible chance of an animal becoming infected and then being replaced. Hence, if at time t there are $Z_r(t)$ infected rabbits and therefore $N_r(t) - Z_r(t)$ susceptibles, at time $t + \Delta t$ we can state that

there have been approximately $P(\Delta t)$ removals and $I(\Delta t)$ infections where $P(\Delta t)$ is a Binomial random variable with parameters $Z_r(t)$ and $1 - e^{-\mu_r T}$ and $I(\Delta t)$ is a Binomial random variable with parameters $N_r(t) \cdot Z_r(t)$ and $1 - \exp\left\{-\frac{\beta_r}{N_r} \int_0^{\Delta t} f_r(c(t)) dt\right\}$. The value of $Z_r(t + \Delta t)$ can then be updated accordingly.

4.1.5. Task 4d: Revise model to incorporate at least one further host species.

It would be simple to use a version of the framework developed to model rabbit infection dynamics to describe any other reservoir population. However, the only species which consistently exhibited high levels of *M. a. paratuberculosis* in the field survey carried out under Objective 2 was the fox. It is likely that cattle have a strong aversion to fox faeces, which will promote avoidance strategies among cattle. It is therefore unlikely that contaminated fox faeces will present much risk of infecting either cattle or rabbits and hence will have a minimal effect on the epidemiology in these species. In addition, the distribution of the infection *in vivo* in the fox suggested that it is likely that infection in this species is mechanical. Hence, there is no evidence of any multiplication of the bacteria *in vivo* in the fox. In this context, there are no grounds for including foxes in the model since they will have a negligible or zero influence on the epidemiology in other animals. It was therefore concluded, after consultation with the rest of the Flexible Fund group, that there was no scientific rationale for a fox model to be implemented.

4.1.6. Task 4e: Use the final model to investigate the theoretical efficacy of various control strategies.

The following control methods are modeled within the program:

1. Slaughter of clinically infected animals. The individual-based simulation model for the cattle population allows the removal of specific animals. Once a clinically infected animal has been identified, it can be summarily removed.
2. Slaughter of the dams, siblings and offspring of identified infected animals. The individual-based simulation model for the cattle population allows the identification and removal of specific animals related to identified clinical cases. Once these animals have been identified, they can be summarily removed.
3. The testing by faecal culture and conditional slaughter of the dams, siblings and offspring of identified infected animals. Once infection is detected in the herd through identification of a clinical case, related animals can be tested. The program will model the test procedure on each infected animal. It is assumed that the probability of detection of animals by faecal culture will depend on the mean number of bacteria shed in the faeces with additional Poisson variability.
4. The testing by ELISA and conditional slaughter of the dams, siblings and offspring of identified infected animals. Testing by ELISA is modeled with a sero-conversion process which applies independently to each animal. We assume that a proportion of animals (87%) sero-convert sufficiently at some fraction of their latent period to allow detection by ELISA. The fractional time at which the event occurs is modeled by a Beta distribution with parameters derived from Collins (1996), where estimates are provided of the proportions of animals in stage 2 and 3 of infection which will be detected by ELISA.
5. The annual faecal testing and conditional slaughter of all animals. Once infection is detected in the herd through identification of a clinical case, all animals can be tested annually. The details of the test model are as given in option 3 above.
6. The annual ELISA testing and conditional slaughter of all animals. Once infection is detected in the herd through identification of a clinical case, all animals can be tested annually. The details of the test model are as given in option 4 above.
7. Husbandry measures to reduce calf exposure. Many different options are available to reduce the exposure of calves to infection. The effect of these measures is not easily

quantified. However, the effect of such measures can be modeled by varying the dam-calf contact parameter.

8. Husbandry methods to reduce cattle exposure. Several different options are available to reduce the exposure of cattle to infection, such as the provision of water rather than depending on natural water sources. The effect of these measures is not easily quantified. However, the effect of such measures can be modeled by reducing the cattle infection parameter.
9. Vaccination. It is recommended that animals be vaccinated within the first month of life. This is modeled by each animal being vaccinated at birth, and hence moved to the vaccinated class. Once in the vaccinated class, infection, if it occurs, will progress at a slower rate than in unvaccinated animals, i.e., the time spent in the pre-advanced clinical states will have a higher mean. The vaccination model is parameterised using data from van Schaik *et al* (1996) in the absence of any published data describing the efficacy of the Weybridge strain vaccine in the UK. It is assumed that serological testing will not be carried out on animals in a herd with a vaccination policy, so there is no need to model the effect of vaccination on serological testing.
10. One-off cull of wildlife reservoir. This can easily be modeled by setting a cull probability p_c and removing a binomially distributed random subset of the currently infected population of rabbits. This does not imply that only infected rabbits are culled, but rather that the population is quickly re-established to equilibrium levels with susceptible animals.
11. Long-term reduction in level of wildlife reservoir. This can easily be modeled by changing the equilibrium population size. A lower population size will result in a smaller average size of infected population, and hence a reduction in environmental shedding.

In practice, combinations of these control options are generally recommended. The following combinations are evaluated.

12. Management, Low. Options 1, 7, 8, 10. This strategy combines all the perceived lower impact management options.
13. Management, High. Options 2, 7, 8, 11. This strategy combines all the perceived higher impact management options.
14. Wisconsin 'Standard' plan: Options 1, 6, 7, 8. This strategy is designed to be comparable to a basic integrated strategy proposed by the University of Wisconsin.
15. Wisconsin 'Accelerated' plan: Options 1, 5, 6, 7, 8. This strategy is designed to be comparable to a more intensive integrated strategy proposed by the University of Wisconsin.
16. Wisconsin 'Ultimate' plan: Options 2, 5, 6, 7, 8. This strategy is designed to be comparable to the most intensive integrated strategy proposed by the University of Wisconsin.
17. Vaccination program: Options 1, 7, 8, 9. This strategy combines vaccination with husbandry measures in an alternative integrated strategy.

Throughout the list of options, with the exception of option 0, where no efforts are made to control the disease, all options incorporate option 1, i.e., all animals with advanced clinical disease will be removed from the herd.

The resulting datasets are analysed using a weighted generalised linear mixed model. The model is a generalised linear model to allow for the binomial nature of the data; the low prevalences seen in some realisations ensure that approximations which neglect the error distribution will give rise to biased results. The model is a mixed model to allow for the variance structure which is present in the data; replicates within a given parameter combination represent pure (stochastic) variability, while variability between different parameter combinations is a reflection of between farm variability and our own uncertainty in choosing parameter values. Parameter combination is fitted as a random effect. The model is weighted to allow the analysis to give more weight to those realisations which give rise to prevalences which are typical of those seen in the field.

It was found that even after 30 years of infection, farms had not yet reached a quasi-equilibrium, and hence the prevalence was still increasing, albeit at a very low rate. Hence

the effect of a given control method must be compared to the baseline figure of the likely 10-year increase in the absence of any control measures.

4.1.7. Discussion

Model Incorporating Rabbit Shedding

The results are presented in Figure 4.1 in relative terms, i.e. the reduction in prevalence as a fraction of the initial prevalence. The control options are listed in order of decreasing mean efficacy.

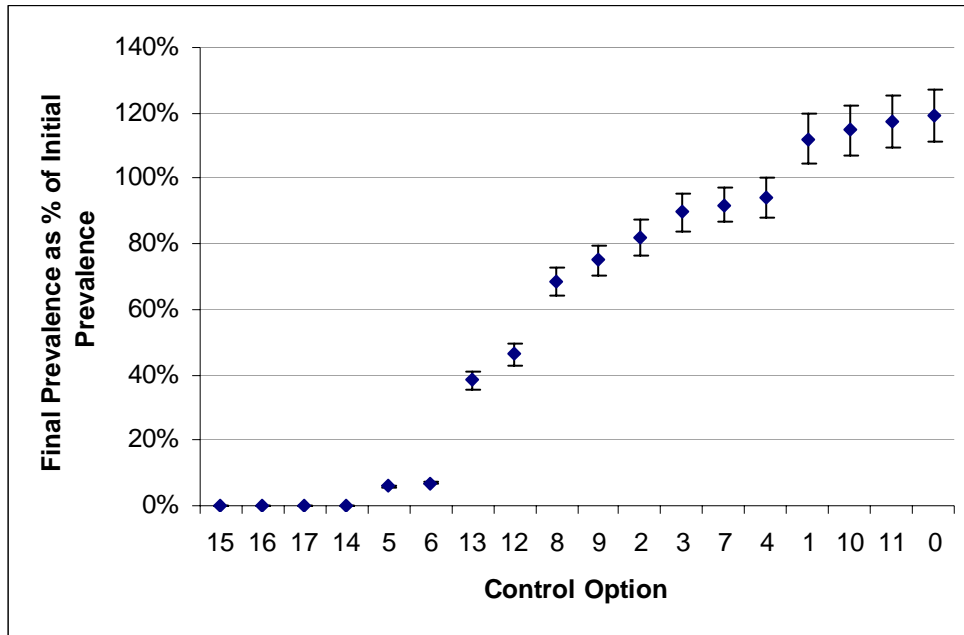


Figure 4.1. Ultimate Relative Prevalences for Different Control Methods Assuming Rabbit Shedding.

The results presented for control options 7 and 8 are calculated assuming a control efficacy of 90%, i.e. the parameter is reduced to one tenth of the baseline value. In practice, it will be impossible to judge exactly how effective a control option will be in purely quantitative terms, so the effects of progressively greater reductions were estimated from the simulations.

Considering the calf-dam post-partum infection effect shown in Figure 4.2, it appears that there is a reasonable return (in terms of reduction in prevalence) over the entire range of likely reductions, although there may be some evidence of diminishing returns for high reductions.

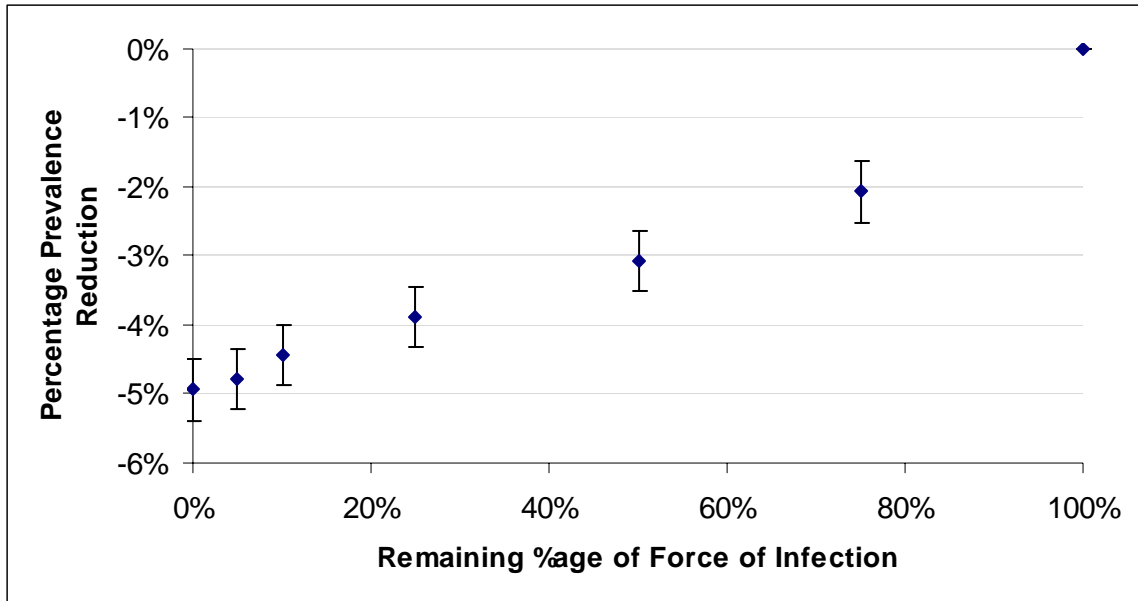


Figure 4.2. Absolute Reductions in Prevalence from Changes in Calf Husbandry Applied at Various Levels of Efficacy.

Considering the environmental infection parameters in Figure 4.3, a similar pattern may be observed. Benefits can be seen in terms of lower prevalences throughout the range of possible reductions, but there is definite evidence of reducing returns as the parameters become close to zero.

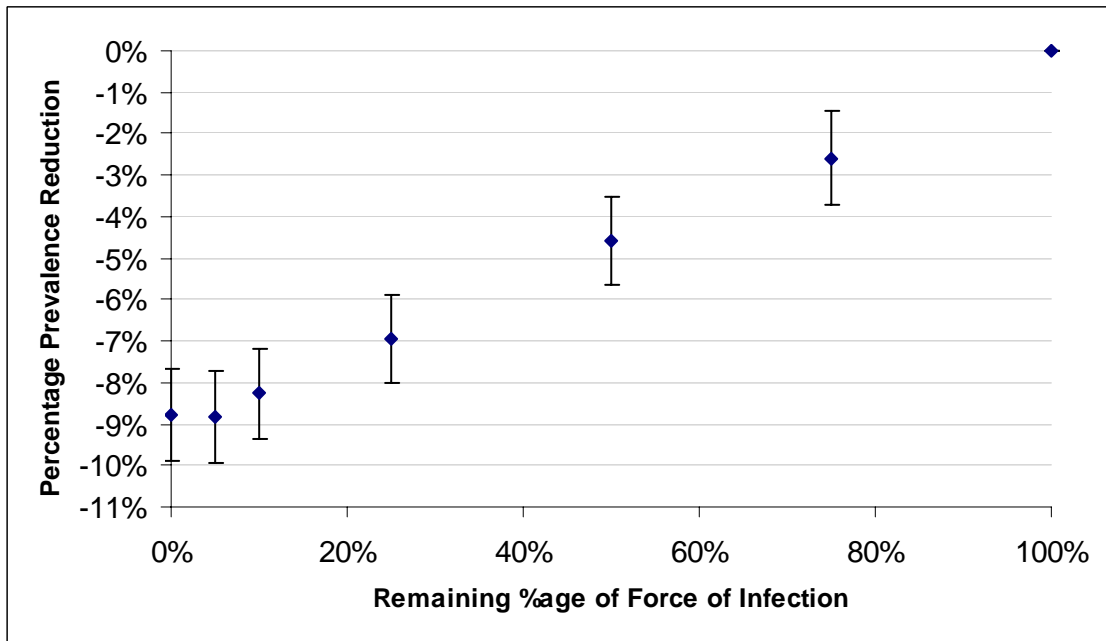


Figure 4.3. Absolute Reductions in Prevalence from Changes in Cattle Husbandry Applied at Various Levels of Efficacy.

Options 1, 10 and 11 all exhibited a point estimate of a reduction in prevalence, but these changes were not statistically significantly different to those seen in the baseline, non-control option.

It is clear that both control measures involving rabbits have only a small effect on the overall prevalence, although, in addressing a route of infection which will not be directly affected by any of the other control methods, they may have an important role in a portmanteau strategy.

In particular, long-term eradication is likely to be impossible in the absence of a rabbit control strategy. It is planned that this issue will be investigated in further work using the model system. There was no apparent difference in efficacy between the two rabbit control strategies ($p=0.8$). Higher reductions in rabbit numbers may give rise to higher drops in prevalence, but such strategies are likely to be expensive. It should be noted that this result is likely to be highly sensitive to the modelling assumption of a single pool of infection, with no differential exposure to rabbit and cattle-derived environmental infection arising from cattle faecal avoidance strategies. If cattle do avoid cattle and rabbit faeces to different degrees, and hence create a differential exposure to the different sources of infection, the results relating to rabbit control would be likely to change radically. Hence it is important to quantify the difference in cattle behaviour towards these different types of faeces. Since the completion of the modelling studies, experiments have been conducted that confirm that grazing cattle do not avoid rabbit faeces. That cattle strongly avoid their own faeces but don't avoid rabbit faeces will have a major impact on the force of infection applying to cattle from environmental distributions of paratuberculosis. In this context, it is reasonable to treat the conclusions of the model with caution where they apply to rabbit control methods.

Removing animals more quickly from the herd once clinical infection was identified is likely to reduce prevalence by a small amount (although this is not statistically significant), at a negligible cost. Removing relatives at the same time will reduce the prevalence by around 6% in absolute terms, which corresponds to almost 20% of the infected population. This is the most effective of all the 'non-testing' strategies, but also the most expensive in terms of opportunity cost. Testing related animals and then removing those which are positive seems likely to be a more cost-effective strategy. The differences between these strategic text and cull strategies and the strategic cull strategy are not statistically significant in relative terms, but the data would suggest that the strategic testing strategies are less effective than the strategic cull strategy, for obvious reasons. There is no statistically significant difference between the outcomes of the strategic testing policies based on faecal and ELISA testing. However, the point estimates of the effects would indicate that the ELISA testing strategy is less effective.

As discussed earlier, improving animal husbandry to reduce the risk of infection on-farm allow respectable reductions in prevalence, and are likely to be worthwhile whatever the size of reduction in infection potential. They are therefore an invaluable component of any portmanteau strategy.

Vaccination alone is unlikely to produce an appreciable reduction in prevalence; the simulation exercise suggests that reductions in prevalence of around 7% may be achieved: this is equivalent to a reduction to around 75% of baseline. These results suggest that the benefits of vaccination may arise largely from the postponement of disease and hence economic loss to later in an animal's productive lifetime, these effects being associated with a reduction in the incidence of *clinical* disease. However, vaccination as part of an integrated management strategy can produce a reduction in prevalence equivalent to that seen in the best alternative portmanteau strategies.

Options 12 and 13 were evaluated to explore the efficacy of strategies which relied only on management changes on the farm, coupled with the removal of animals which were likely to be infected. These portmanteau strategies are more effective than any of their individual components, and do give rise to reductions in prevalence which approximate to half of the initial level, but they are unable to drive the infection to very low levels. Option 13 was designed to incorporate the more expensive (and probably more effective) management options relative to those used in option 12. This class of portmanteau strategy can reduce the prevalence to just under 50% of its previous value, but further reductions require on-going testing strategies.

Options 5 and 6 (testing of all animals each year) are the only individual strategies which are capable of producing a large reduction in prevalence. Each strategy is capable of removing almost 95% of the initial prevalence within 10 years. The point estimate of the mean prevalence reduction achievable under ELISA sampling is smaller than that achievable with faecal testing, but there is no statistically significant difference between the effects of the

ELISA and faecal culture testing regimens. Hence, the ELISA test is virtually as effective in reducing the prevalence, as well as being cheaper. A similarly non-significant difference is visible when comparing options 3 and 4. Investigation of the effects of longer waiting times for faecal culture results suggests that the comparative advantages of the two tests are different: the ELISA test has a different (and generally lower) sensitivity, but this is offset by a faster turn-around time. Where faecal culture is used, it is therefore essential that positive results are returned to the farmer and acted upon quickly. These results indicate that any test with higher sensitivity than ELISA testing and a turn-around time closer to that of ELISA would have enormous potential to control the disease.

Evaluation of the three 'Wisconsin' portmanteau strategies (14, 15 and 16) shows that all three are very effective in driving the prevalence to very low levels (or zero) by the end of the 10 year period. Presumably the more comprehensive strategies achieve this objective more quickly, and hence do provide some excess economic benefit to weigh against their greater cost. This would involve a more detailed cost-benefit analysis than is included in this model. Given that these options are all very largely removing all of the infected animals from the herd most of the time, it is even more essential that comparisons are based on the relative drop in prevalence, since the absolute drop will be largely determined by the size of the initial prevalence and hence the availability of infected animals to be removed. Analysing the relative prevalences, the prevalence achieved under Option 15 is significantly lower than that achieved under Option 14 ($p < 0.001$). This reduction reflects the efficacy of using both faeces and ELISA testing on the same herd at the same time. There are no statistically significant differences between Options 15 and 16. The vaccination portmanteau strategy can also reduce the prevalence to a low level, and may present cost-benefit advantages when compared with the test and cull strategies. It appears to reduce the mean relative prevalence to a value similar to that achieved under the Standard Wisconsin plan; however, the Wisconsin 'Accelerated' and 'Ultimate' plans do give rise to statistically significantly lower mean prevalences ($p = 0.02$). In practice, it is doubtful whether this formally statistically significant difference would be of any real veterinary significance.

It is instructive to examine the results of the statistical analyses in more detail. The analysis of the weighted simulation results gives rise to the means and standard errors presented in the tables above, but the GLMM approach can be used to estimate the prevalence for each scenario both before and after the control method has been implemented. These latter values can then be subtracted from the former to give a point estimate of the effect of the control strategy on each scenario. The distribution of these differences can be examined visually and summarised using appropriate distributions. The distributions which are generated are frequently bimodal. This represents a dichotomy in the parameter space between those farms with parameter combinations for which the specified control option is effective in reducing the prevalence (negative effects) and those for which it has no effect, or potentially generates a higher prevalence. This may seem counterintuitive, and so deserves further discussion.

The model incorporates several different routes by which animals might become infected, with different routes being important at different ages. In addition, animals infected as calves have a different clinical history to those infected as adults. Given particular combinations of parameters, it is possible for a control option which reduces the risk of infection in a young animal to have a counterproductive effect at the herd level, if the consequence of a reduction in calf infection is merely the creation of a larger pool of adult susceptibles, where the consequence of an adult infection ensuing are worse (i.e., perhaps generating a higher mean volume of bacteria shed) than of a calf infection. There is evidence of counterproductivity in both control option 7 and 8 as the efficacy of the projected husbandry improvements approached 100%. Examination of the parameter space indicates that the scenarios which exhibit this behaviour are those with a high prevalence and hence a high level of potential animal to animal or environment to animal infection. This is consistent with the rationale given above. Reductions in the dam to calf infection rate will obviously increase the number of adult susceptibles as will any reduction in the direct environment to animal infection rate (which also affects the environment to calf infection rate). However, in both cases there is a large pool of existing infectives which generate a high chance of adult animals becoming infected. Such counterintuitive effects should not be unexpected in complex non-linear

systems such as that under investigation. A number of similar effects have been reported in the literature, most notably that involving rubella immunisation for children, where the initial effect of the vaccination campaign is to increase the risk of pregnant women being infected with German measles. It is precisely because the effect of such non-linear interactions are so difficult to conceptualise that the methodology proposed in this study is appropriate. Allowing for variability and uncertainty in the parameters allows us to explore the full variety of outcomes that are likely to arise from the system. Merely using the mean value of each parameter, as is frequently the case in epidemiological simulations, would only ever present a single mode of behaviour, and in the current study this would be actively misleading.

As a consequence of this behaviour, the summary statistics of the mean and standard errors of the changes in prevalence are not strictly appropriate in those cases where the results are bimodal. However, they are likely to be more than adequate for the purpose of reviewing the mean change in prevalence in this study. It should be noted, however, that cost-benefit analyses, where the calculation of financial benefits may magnify the variability seen in the bimodal prevalence distributions should clearly be carried out on a scenario by scenario basis, rather than being applied to notional mean effects. A mixture model consisting of two normal distributions with distinct means and variances was fitted to each set of scenario effects for each control option. It should be noted that those control options with a strong effect did not exhibit the bimodal behaviour described above: the bimodality being swamped by the strength of the intervention. These mixture models were then used to identify the median and 95% points of the full distribution of outcomes. These are plotted in Figure 4.4. This plot illustrates the range of outcomes seen in individual scenarios for each control method.

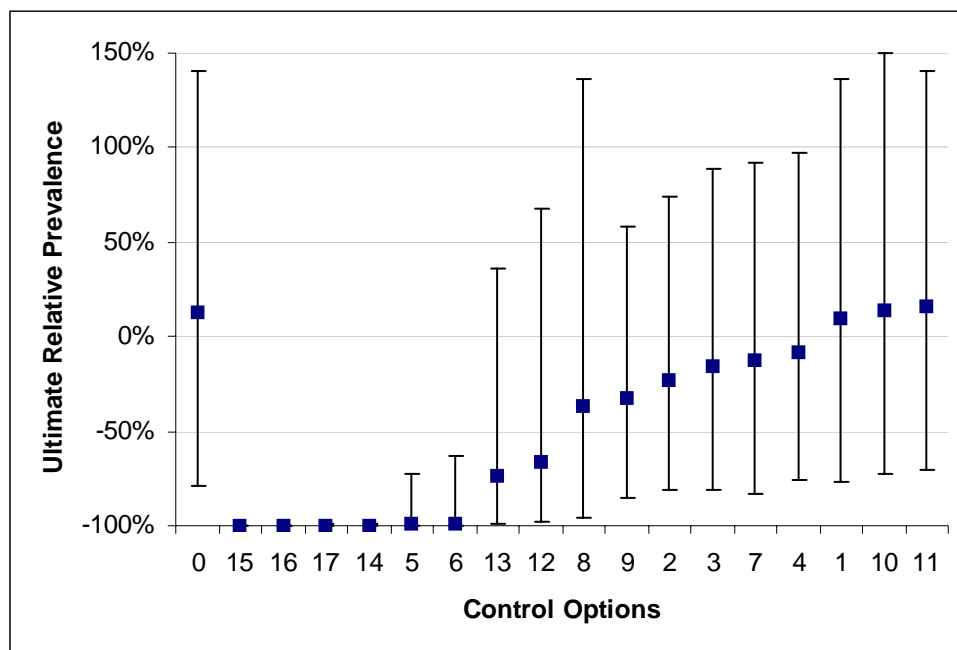


Figure 4.4. Median and 95% Points of the Effect Distribution for Different Control Options.

The ordering of effects is identical to that arising from considerations of the estimated mean effect, confirming that the arguments considered earlier retain their validity. What is of particular interest, however, is to see the variability in outcome accruing to different scenarios under different control options. It is particularly noticeable that all options other than those involving test and cull give rise to a sizeable number of outcomes for which the final prevalence is higher than the initial value. This confirms the frequently heard argument that testing is a *sine qua non* for successful reduction of paratuberculosis prevalence. Even test and cull strategies alone, although exhibiting very low final prevalences in the bulk of scenarios, exhibit a large tail of outcomes in which the prevalence, although reduced, remains at far from negligible levels. Only test and cull coupled with management changes appears to guarantee very low levels of prevalence under all scenarios.

It should be stressed that the parameter distributions used in this exercise incorporated elements both of farm to farm variability and of uncertainty. It is possible that more precise determination of some parameter values would suffice to restrict the model to regions of the parameter space in which counterproductive or low efficacy effects do not occur. Nevertheless, given the current state of knowledge about the likely values of key epidemiological parameters, the existence of these counterintuitive and/or undesirable scenarios cannot be ruled out. All of the results presented in this report of course depend on the assumptions built into the herd and disease models, and should be interpreted in the light of those qualitative and quantitative assumptions.

Model With Rabbit Shedding, High and Low Prevalence Scenarios

The analyses presented in this report are produced by weighting simulation results within a statistical analysis. The results discussed in the previous section were produced by weighting the simulation results with respect to an observed epidemiological distribution. However, reweighting the results relative to those farms with (observed) higher or lower prevalences can allow the analysis to explore the relative efficacy of different strategies on high and low prevalence farms. For this purpose, the simulation results are reweighted relative to the highest and lowest quartiles of the observed prevalence distribution.

The effect of different prevalence levels on the efficacy of the control strategies can be well summarised by plotting the observed values in a scatterplot as in Figure 4.5. In each case, the paired prevalence values bracket the equivalent estimated value in Figure 4.1.

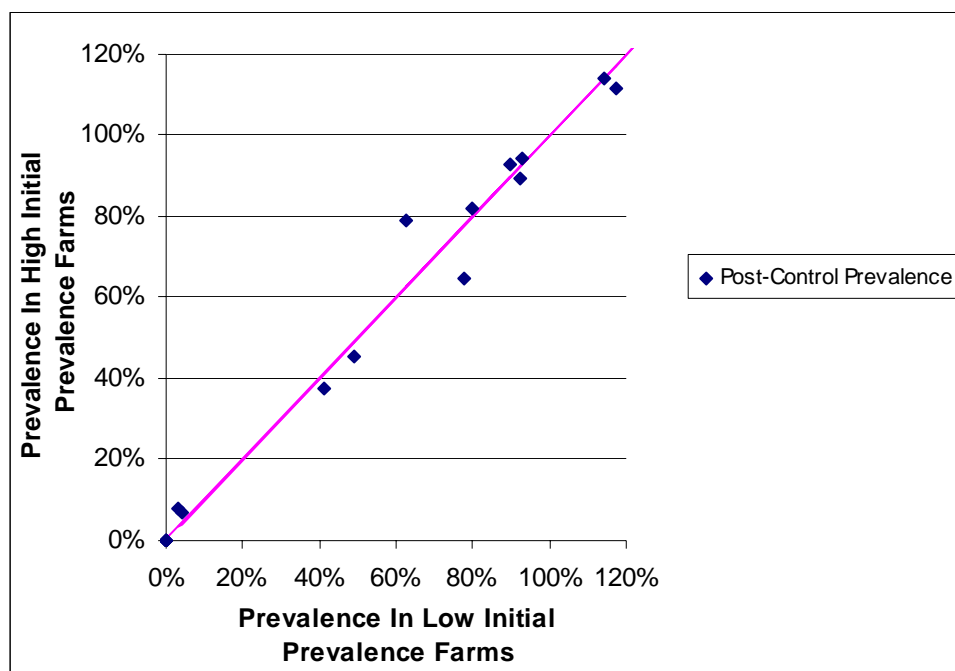


Figure 4.5. Plot of Ultimate Relative Prevalences Under Different Control Schemes on Low and High Prevalence Farms.

The line in Figure 4.5 represents the situation where a control measure is equally efficacious on high and low prevalence farms. It is clear that for most control methods, there is little difference in efficacy on the two classes of farms. The exceptions are control options 8 and 9, which show moderate differences. Of these, only the difference in prevalences arising from option 9 is formally statistically significant ($p=0.046$). Hence the vaccination policy is significantly more effective on low prevalence farms, and improvements in cattle husbandry are likely to be more effective on high prevalence farms. The difference, however statistically significant, is however only of the order of 15-20% of the relative prevalence. Similarly, the test and cull options 5 and 6 are both statistically significantly more effective on low

prevalence farms, but this difference is unlikely to be important in the field. It is probably more relevant to review the rankings of the different control methods in the two situations, as shown in the figures above. Changes in efficacy arising from initial high or low prevalence in the herd leaves the overall rankings of the different options unchanged. Four pairs of strategies swap positions in the list, options 5 and 6, 8 and 9, 3 and 7 and options 10 and 1. Of these, only the change in relative ranking of options 8 and 9 is both statistically well founded and possibly relevant in the field.

Hence, it can be concluded that the rankings defined in the earlier section are broadly applicable to all farms, regardless of whether they initially have a high or low prevalence. This should not, however, be interpreted as a statement that all control options are equally appropriate on all classes of farms. Equivalent effects on the relative prevalence equate to dramatically different effects on the absolute prevalence in farms in the two classes, and in many respects it is the change in absolute prevalence which defines the productivity and economic benefits which follow from a given control measure. The calculation of financial cost-benefit figures for the different control measures is beyond the remit of this project. However, the methodological approach presented here has the potential not only to facilitate the calculation of cost-benefit figures for the 'average' farm, but also to allow the calculation of cost-benefits for farms with high or low initial prevalences.

Model Without Rabbit Shedding

Carrying out the same Latin Hypercube Sampling and weighting procedure on the model with the rabbit components made inactive will have two effects on the results. Most obviously, the simulated epidemiologies will not exhibit any of the effects derived from the presence of infective rabbits. Given the results seen when evaluating the efficacy of rabbit-based control methods, this is not likely to have a large effect. A more subtle effect lies in the nature of the weighting scheme derived from the LHS exercise. The absence of the rabbit-based infection route will force more weight on to alternative routes of infection, and hence change the weighting associated with the parameters driving these alternative routes. Fundamentally, a combination of parameters which is likely to give rise to results comparable to those seen in the field in the presence of rabbit shedding may look less plausible when rabbit shedding is removed from the system (and contrarywise). The relative prevalences are presented in Figure 4.6.

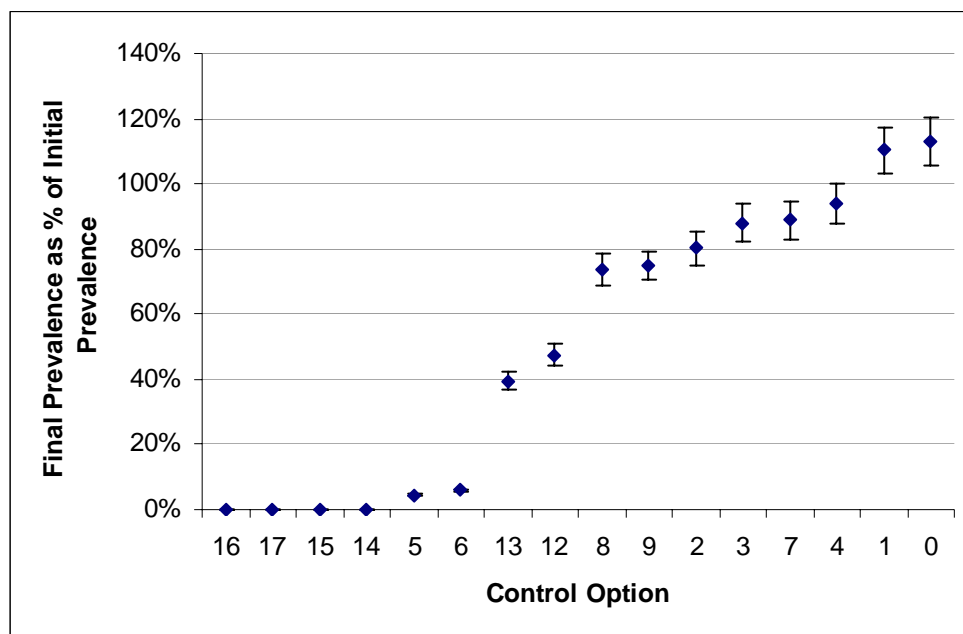


Figure 4.6. Ultimate Relative Prevalences for Different Control Methods Assuming No Rabbit Shedding.

Comparing the ranking of the control measures with those seen in Figure 4.1, there is little change in the relative efficacies of the different control measures. This is exhibited in Figure 4.7, which plots the relative final prevalences with and without rabbits for each of the control measures.

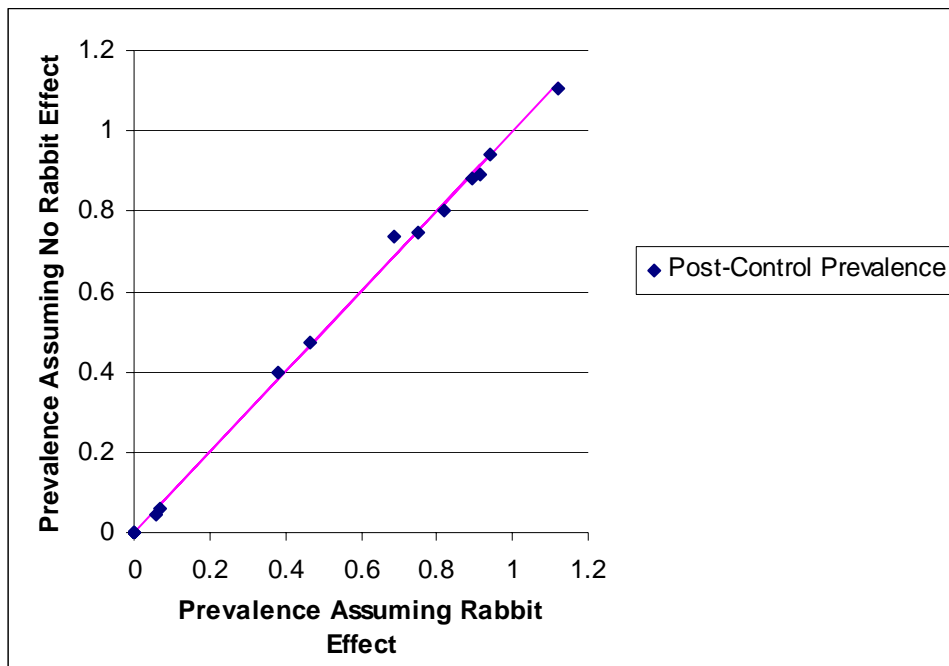


Figure 4.7. Comparison of Ultimate Relative Prevalences Under Different Control Schemes in the Presence and Absence of Rabbit Shedding.

Most of the points in this plot lie close to the equality line which corresponds to an equal drop in prevalence regardless of the presence or absence of rabbits. All but three of the points lie just below the line, representing control measures which are slightly less effective in the presence of rabbits than in the absence of rabbit shedding. Only one of these drops is statistically significant, but it is clear in any case that they have no real biological significance. Interpretation is simple also; it is obvious that where a control measure is fundamentally driven by the removal of infective animals and/or the reduction of contact of animals with environmental contamination, the presence of any additional route of infection, with a commensurate increase in environmental infection, will always reduce the efficacy of the control measure. The exception to this general rule is control option 8: animal husbandry changes which give rise to a lower animal to animal force of infection, and options 12 and 13, the effects of which largely depend on the use of option 8. This effect is not statistically significant, but is likely to be real. The effect is unlikely to be a direct epidemiological effect such as those seen in the other control measures. It is hard to see how the presence of rabbit shedding *per se* could make a policy of reducing other sources of infection more efficacious. Rather, the explanation is more likely to lie in the relative weighting given to different parameter options. *A priori*, it is reasonable to believe that the direct force of infection reduction strategy could be more sensitive to this effect than the other strategies explored in the study. The animal to animal infection rate and the rabbit to animal infection rate are complementary forces, operating on the same animals at the same time. Where some of the observed prevalence can be assigned as being caused by rabbit infection, a smaller amount is available to be assigned to the direct animal to animal force of infection, so the initial weighting of this parameter will assign more weight to lower values. When the direct force of infection is then reduced as part of a control strategy, the lower values associated with this parameter will be less subject to the counterproductive pressures discussed earlier. Hence, overall, control strategy 8 can perform successfully on more farms in the presence of rabbits. It should be stressed, however, that the biological significance of this result is likely to be negligible.

The study indicates that the relative efficacies of the different strategies are barely affected by the choice of rabbit shedding/non-shedding assumption. Hence, the relative rankings of the non-vaccination strategies can be used to categorise the efficacy of different control options, since the model suggests that these conclusions are robust against the existence or non-existence of an infection reservoir in the rabbit population. Again, however, it should be noted that since the completion of the modelling studies, experiments have been conducted that confirm that grazing cattle do not avoid rabbit faeces, whereas the model described above assumes equal levels of avoidance for rabbit and cattle faeces.

4.2. N/A

4.3. Summary of Results

A suite of models have been developed which model

- Herd management in a herd of dairy cattle
- Infection dynamics of *M. a. paratuberculosis* in cattle
- Population dynamics of *M. a. paratuberculosis* in the environment
- Infection dynamics of *M. a. paratuberculosis* in rabbits
- Various control options, including testing of animals based on both faecal culture and ELISA methods.

No published model has previously integrated all of these factors.

These models have been linked together and evaluated within a Latin Hypercube Sampling framework which allows the simulation exercise to incorporate uncertainty and variability in model parameters. The results of these simulations were weighted relative to epidemiological field data and analysed using new statistical methods. As a result, the results of the analysis should be representative of the field situation.

The efficacy of the different control measures are well summarised by Figure 4.8. The control methods evaluated exhibit a wide range of effects, ranging from virtual eradication at one extreme to virtual inefficacy at the other.

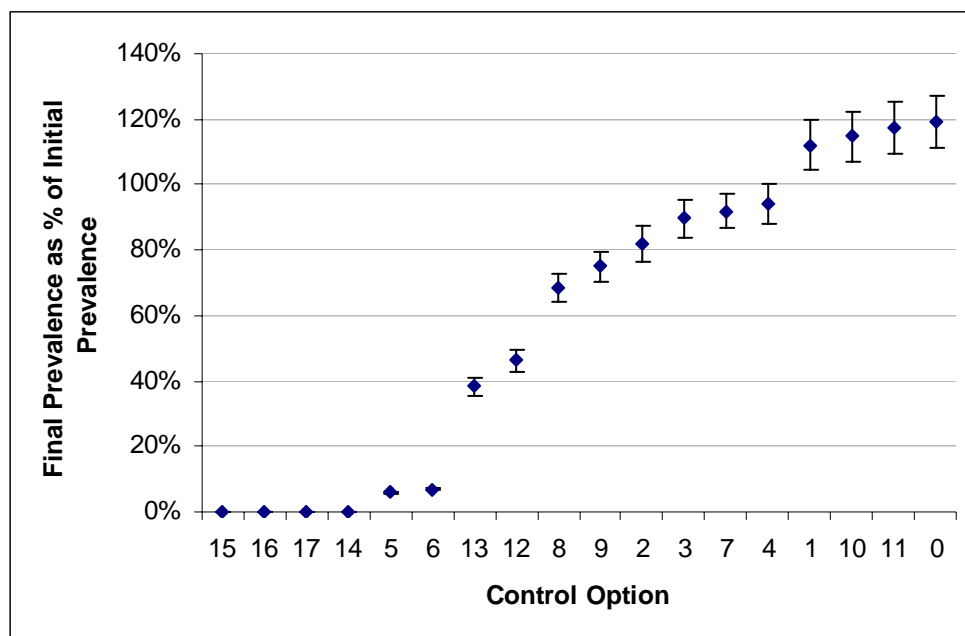


Figure 4.8. Ultimate Relative Prevalences for Different Control Methods Assuming Rabbit Shedding.

5. Conclusions

- Epidemiological models evaluated with the Latin Hypercube Sampling with weighting methodology give rise to data which, when analysed using Generalised Linear Mixed modelling methods, are highly effective in illuminating the properties of the model and the efficacy of different control methods.
- Test and cull control policies based on ELISA testing appear to be only slightly less effective than those based on faecal culture testing, but this difference is not statistically significant.
- Vaccination alone will generate a drop in prevalence, but will not suffice to reduce the prevalence to a low level.
- Management changes alone will generate a drop in prevalence, but will not suffice to reduce the prevalence to a low level.
- ELISA or faecal culture testing and culling will reduce the prevalence to a low level.
- A combination of vaccination with management changes is required to reduce the prevalence to a negligible level.
- Test and cull when coupled with management changes has the potential to reduce the prevalence to a negligible level. Some of the more intensive strategies proposed in the USA appear to make little or no difference at the 10-year level. However, these strategies may be useful in delivering lower prevalences more quickly.
- Control measures targeted at the rabbit population have little effect on the dynamics of the disease. However, this conclusion is highly sensitive to some of the modelling assumptions made in the modelling process. More work is required on the interactions between cattle and faeces in the environment before this issue can be tackled successfully. In the meantime, the conclusions of the model should be treated with caution where they relate to rabbit control.
- Some of the control options exhibit bimodal responses: i.e. some farms exhibit improvements as a consequence of the control strategy and others do not. In fact, some exhibit increases in prevalence as a consequence of strategies designed to reduce infection levels.
- Most of the control measures exhibit high variability in outcome. Only options involving test and cull do not present a large proportion of scenarios in which the prevalence increases, and only those involving test and cull and management changes can guarantee to reduce the prevalence to negligible levels within 10 years.
- There is some evidence that the vaccination strategy is particularly effective on farms with a low initial prevalence, and that perhaps changes in animal husbandry which reduce direct animal to animal infection rates are particularly effective on farms with high initial prevalences. However, the initial prevalence has, at most, a marginal effect on the relative efficacy of the different control measures evaluated.
- The relative efficacies of the control strategies, and hence the points specified above, are insensitive to the assumptions made about rabbit shedding. Hence advice on appropriate control strategies can for the most part be made without reference to the rabbit situation on a given farm.

6. Communicated Outputs

None to date. It is planned to submit a paper describing the model to the *Journal of the Royal Society Series B*.

7. Resources

8. Acknowledgements

Milk Development Council for access to 'Research into Practice' information booklets for dairy farmers. Thanks to Basil Lowman, SAC St Boswells and George Gunn and Michael Pearce of SAC Inverness for opinions about the nature of typical Scottish dairy herds.

9. Annexes

9.1 References

Anonymous, Unspecified, *Longevity- controlling culling to improve herd profitability*, MDC Research Into Practice Publication 51, Milk Development Council, Cirencester.

Anonymous, 1998, *Calf rearing- how to get it right and minimise losses*, MDC Research Into Practice Publication 29, Milk Development Council, Cirencester.

Anonymous, 2000, *Economic Report on Scottish Agriculture 2000 Edition*, Scottish Executive Rural Affairs Department, Edinburgh.

Boelaert, F., Walravens, K., Biront, P., Vermeersch, J.P., Berkvens, D., and Godfroid, J., 1999, Prevalence of Paratuberculosis (Johne's Disease) in the Belgian Cattle Population, *Proceedings of the 6th International Colloquium on Paratuberculosis*, 76-88.

Bölske, G., Sternberg, S., Visker, D. and Larsson, B., 1999, Use of culture and serology in tracing and screening for paratuberculosis in Sweden, *Proceedings of the 6th International Colloquium on Paratuberculosis*, 349-352.

Brown, H. & R. Prescott, 1999, *Applied Mixed Models in Medicine*. Chichester: Wiley.

Castle, M.E. and Watkins, P., 1984, *Modern Milk Production*, 2nd edition, Faber and Faber, London.

Cetinkaya, B., Egan, K. and Morgan, K.L., 1994, A practice-based survey of the frequency of Johnes disease in south-west England, *Veterinary Record*, **134**, 494-497.

Cetinkaya, B., Erdogan, H.M. and Morgan, K.L., 1998, Prevalence, incidence and geographical distribution of Johnes's disease in cattle in England and the Welsh borders, *Veterinary Record*, **143**, 265-269.

Collins, M. T., 1996, Diagnosis of paratuberculosis, *Veterinary Clinics of North America: Food Animal Practice*, **12**, 357-371.

Collins, M.T. and Morgan, I.R., 1991, Epidemiological model of paratuberculosis in dairy cattle, *Preventive Veterinary Medicine*, **11**, 131-146.

Etgen, W.M. and Reaves, P.M., 1978, *Dairy Cattle Feeding and Management*, 6th edition, Wiley, New York.

Greig, A., Stevenson, K., Henderson, D., Perez, V., Hughes, V., Pavlik, I., Hines, M. E., McKendrick, I.J., and Sharp, J.M., 1999, An epidemiological study of paratuberculosis in wild rabbits in Scotland, *Journal of Clinical Microbiology*, **37**, 1746-1751.

Humphry, R., Stott, A.W., Jones, G.M., Gunn, G.J., 2001, An economic evaluation of Johnes's disease (paratuberculosis) in the beef herd using the Markov Chain model, in *Assessment of surveillance and control of Johnes's disease in farm animals in GB*, SAC Veterinary Science Division, <http://www.defra.gov.uk/animalh/diseases/sac2.PDF>.

Iman, R., Conover, W., 1980, Small Sample Sensitivity Analysis Techniques for Computer Models, with an Application to Risk Assessment, *Commun. Statist.-Theor. Meth.* **A9**(17), 1749-1874.

Iman, R. and Conover, W., 1982, A Distribution-free approach to inducing rank order correlation among input variables, *Commun. Statist.-Simula. Computa.* **11**(3), 311-334.

Lovell, R., Levi, M. and Francis, J., 1944, Studies on the survival of Johnes's bacilli, *J. Comp. Path.*, **54**, 120-129.

Van Schaik, Kalis, C.H.J., Benedictus, G., Dijkhuizen and A.A., Huirne, R.B.M., 1996, Cost-benefit analysis of vaccination against paratuberculosis in dairy cattle, *Veterinary Record*, **139**, 624-627.

Vose, D., 2001, *Risk Analysis: A Quantitative Guide*, 2nd Edition, John Wiley and Sons, Chichester.

Whitlock, R.H., Hutchinson, L.T., Sweeney, R.W., Spencer, P.A., Rosenberger, A.E. and van Buskirk, M.A., 1994, Pattern of detection of *M. a. paratuberculosis* infected cattle in ten dairy herds cultured every six months for four years, *Proceedings of the 4th International Colloquium on Paratuberculosis*, 47-53.

Wren, G., 1998, Johne's disease control strategies, *Bovine Veterinarian*, May-June, 1998.