

The role of super-shedders in the spread and maintenance of *E. coli* O157 on Scottish cattle farms

L. Matthews^{1*}, G. Gunn², B. Synge², H. Ternent³, I. McKendrick⁴, and M.E.J. Woolhouse¹. ¹Centre for Tropical Veterinary Medicine, University of Edinburgh.

²Scottish Agricultural College (SAC), Inverness. ³Veterinary School, University of Glasgow

⁴Biomathematical and Statistics Scotland (BIOSS), Edinburgh.

Summary

The distribution of prevalences of *Escherichia coli* O157 displays striking heterogeneity across the Scottish cattle population. On approximately 78% of farms in a cross-sectional survey of over 900, no shedding of *E. coli* O157 was detected, but a small proportion, approximately 2% displayed very high prevalences of infection (90-100% of cattle sampled). Using maximum likelihood methods to fit epidemiological models to these data we find that these heterogeneities are best explained when a proportion of animals are assumed to transmit infection at much higher levels than the other animals – the so-called super-shedders.

Introduction

Escherichia coli O157 is an important zoonosis with global distribution. In Scotland, approximately 200 cases of *E. coli* O157 infection in man are reported annually, though much larger outbreaks with significant fatalities have occurred. Cattle are well recognised to be a reservoir for the organism, though the mechanisms by which *E. coli* O157 is maintained in the cattle population are still poorly understood. Studies of *E. coli* O157 prevalence in cattle typically reveal the shedding of the organism to be sporadic and of short duration. However, there exists some evidence to suggest that some cattle may be more persistent shedders, or shed at higher levels than others.

A cross-sectional survey of over 900 farms revealed striking heterogeneities in the distribution of prevalences of *E. coli* O157 across the Scottish cattle population. On approximately 78% of farms in a cross-sectional survey of over 900, no shedding of *E. coli* O157 was detected, but a small proportion, approximately 2% displayed very high prevalences of infection (90-100% of cattle sampled). Here we address the question of whether the few herds displaying high levels of shedding have different epidemiological characteristics, or whether the on-farm transmission dynamics of *E. coli* O157 are such that high prevalences are to be expected occasionally by chance.

There are many potential sources of epidemiological heterogeneity in the Scottish cattle population: some farms may provide a better environment for transmission than others; differing exposure to infection may depend on movement of livestock on and off a farm; some animals may be better transmitters of the organism than others (for example they shed at higher levels – the so-called super-shedders – or for longer periods than others).

Objectives

The objectives of this study are: first, to develop a suite of epidemiological models which incorporate heterogeneities at the animal, farm and metapopulation level; second, to fit these models to the data using maximum likelihood methods; and third,

to draw conclusions as to the likely sources of heterogeneity underpinning the observed distribution of prevalences of *E. coli* O157 on Scottish cattle farms.

Methods

A novel approach is taken, whereby the cross-sectional data are viewed as providing independent snap-shots of a dynamic process. Since group sizes can be small and prevalences of infection low we describe the dynamics in terms of a stochastic SIS process. Thus, within group transmission of infection (occurring at an mean rate β), recovery from infection (which occurs at a mean rate σ), and immigration of infection into the group (which occurs at a rate λ) are probabilistic events. The probability, $p_j(t)$ of there being j animals infected in a group of size N at a time t is given by

$$\frac{dp_j}{dt} = \sigma p_{j+1} + (\lambda + \beta(j-1)/N)(N - (j-1))p_{j-1} - (\sigma + \lambda(N-j) + \beta j(N-j)/N)p_j$$

Heterogeneities are easily incorporated in the model by additionally specifying either

- the proportion of farms with higher transmission or immigration rates plus the ratio specifying how much higher those rates are;
- or the proportion of individuals with higher transmission or longer recovery rates and the ratio specifying how much higher those rates are

In contrast to a deterministic approach, which would provide an expected mean prevalence for a given set of parameter values, the stochastic model leads to an equilibrium distribution of prevalences; this corresponds to equilibrium values of the probabilities, p_j , of there being j infected individuals in the group.

The cross-sectional data are regarded as providing a sample from an equilibrium distribution of prevalences. Alternative models provide different theoretical equilibrium distributions which are fitted to the data using maximum likelihood methods and compared using the likelihood ratio test (in the case of nested models) and AIC (Akaike information criterion) otherwise.

Results

The study results are summarised in the table below. For example, for the super-farm model with increased immigration rates we have listed the best fit values for the transmission rate (0.9), the immigration rate for standard farms (0.006), the proportion of farms which have higher levels of immigration (8%), and the ratio by which the immigration rates on these farms is increased compared to the standard farm (150 fold).

The five alternative models are listed in the table in order of increasing goodness of fit. Moreover, applying the likelihood ratio test or AIC where appropriate, statistically significant increases in likelihood (decreases in negative log likelihood) are achieved as we progress through the models in order of increasing goodness of fit.

So, whilst the models which incorporate heterogeneities at the farm or metapopulation level provide a significantly better fit to the data, they do not provide as good a fit as the model which incorporates heterogeneity at the individual level. Moreover, allowing a level of clustering in the distribution of the super-shedders

across the farm population rather than a purely random distribution further improves the model fit.

Model	No heterogeneities	Super-farm with higher immigration rates	Super-farms with higher transmission rates	Super-shedders with higher transmission rates	Clustered super-shedders
Transmission rate (β)	1.1 (1.0-1.2)	0.9 (0.8-1.1)	0.9 (0.8-1.2)	0.65 (0.06-0.08)	0.65 (0.06-0.08)
Immigration rate (λ)	0.006 (0.003-0.009)	0.006 (0.003-0.009)	0.006 (0.003-0.009)	0.008 (0.006-0.01)	0.008 (0.006-0.01)
% super-farms		8% (6-11)	11% (7-15)		
super-farm ratio		150 (130-190)	3 (2-5)		
% super-shedders				4% (3-6)	5% (3-8)
super-shedder ratio				50 (30-60)	30 (20-40)
(Neg log) Likelihood	1126	1069	1067	1058	1053

Discussion

Our result suggest that there are heterogeneities in the Scottish cattle population which lead to a highly skew distribution of prevalences – the majority of farms at any one time have no shedding animals, whilst a few possess high prevalences of shedding animals. The best fit to the observed data is obtained from a dynamic on-farm transmission model which incorporates heterogeneity at the individual level whereby a proportion of animals are assumed to have much higher transmission rates than the others. Our model suggests that these individuals (denoted super-shedders) which comprise roughly a twentieth of the population have transmission rates several orders of magnitude higher than the non-super-shedders.

Further experimental evidence of the existence of these super-shedders is emerging; the identification of risk factors for super-shedding will be key to the targeting of these individuals as a means of controlling the spread of *E. coli* O157.

Acknowledgements

Financial support for this work was provided by the Wellcome Trust (LM).