Disease transmission potential associated with animal movement within the Scottish sheep industry

Victoriya Volkova, Richard Howey, Nicholas J Savill and Mark E J Woolhouse

Epidemiology Group, Centre for Infectious Diseases, University of Edinburgh, UK

Introduction

Animal movement is a major route for the spread of livestock infectious diseases (Fèvre *et al.* 2006). Understanding the structure of the livestock industry's movement networks and the potential for the transmission of infections within them is crucial for efficient control of livestock diseases, both epidemic and endemic. The analyses can be conducted over time scales relevant to specific infections. In this study, we considered the basic reproduction number, R_0, of infectious disease transmitted by animal movements among Scottish sheep holdings for four one-year intervals (and therefore relevant to chronic infections) from July 2003 to June 2007. Building on earlier work by Woolhouse *et al.* (2005), we evaluated the contribution of network structure, including that of higher order moments of the contact network, to R_0.

Materials and Methods

Data: Scottish Animal Movement System records for 2003 to 2007 were processed in Python programming environment and then transferred to SAS® 9.1.3 software for Windows (SAS Institute Inc.). Scottish sheep holdings were analyzed as a closed network with only movements between holdings within Scotland considered. All movements between farm holdings were included, whether direct or via a livestock market. Movements to slaughter, to shows, and those with incomplete histories (no departure or final destination holding) were excluded. **Analysis:** Each year's movement data can be represented by a contact matrix, **A**, where a(i,j) (the element in ith row and jth column) is 1 if there is a movement from farm j to farm i and 0 otherwise. The expected number of secondary infections from an individual infected holding is proportional to the dominant eigenvalue of the contact matrix. Eigenvalues were calculated using the ARPACK FORTRAN77 code libraries which are specially written to solve large-scale eigenvalue problems, in particular for structured or sparse matrices (Lehoucq *et al.* 1997). Each contact matrix had a dimension of over 15,000 (Table 1, Column 2) with around 70,000 (<3%) non-zero entries. Using a Pentium4 CPU, the dominant eigenvalues were calculated by the ARPACK subroutines in around 10 seconds. The contributions to R_0 from only first order moments (proportional to mean number of contacts) and from only first and second order moments (proportional to the square root of the mean cross-product of in-contacts and out-contacts) were calculated using SAS 9.1.3 software.

Table 1. Network statistics for movements of sheep among the population of Scottish sheep holdings 07/2003-06/2007

| Year	No. farms	Mean contacts	Variance in-contacts	Variance out-contacts	Square root mean cross-product in-contacts*out-contacts	Dominant eigenvalue
07/2003-06/2004	15788	4.6	1302.5	48.4	6.4	8.3
07/2004-06/2005	15314	4.7	1289.5	50.1	6.5	6.2
07/2005-06/2006	15762	4.4	826.6	44.7	5.8	7.6
07/2006-06/2007	15750	4.3	663.2	46.4	6.2	7.6

Chart 1. Distribution of in-contacts for Scottish sheep holdings 07/2006-06/2007
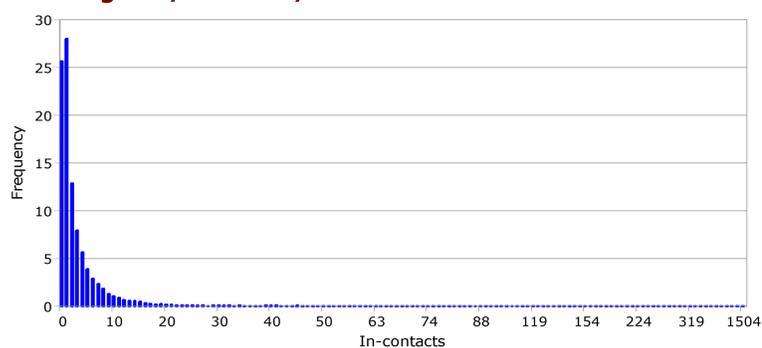
Chart 2. Distribution of out-contacts for Scottish sheep holdings 07/2006-06/2007
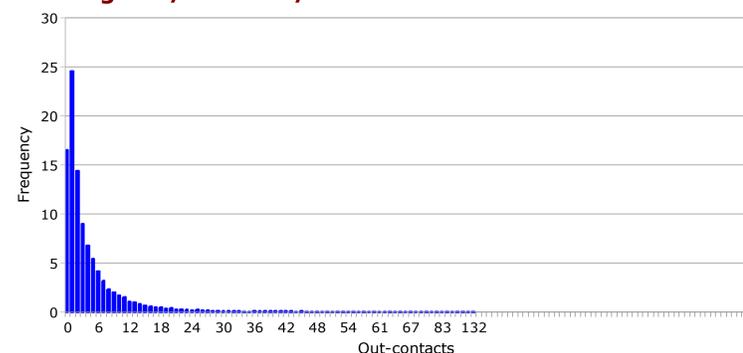
Chart 3. Co-distribution of in- and out-contacts (transformed as log$_e$ [x+1]) for Scottish sheep holdings 07/2006-06/2007
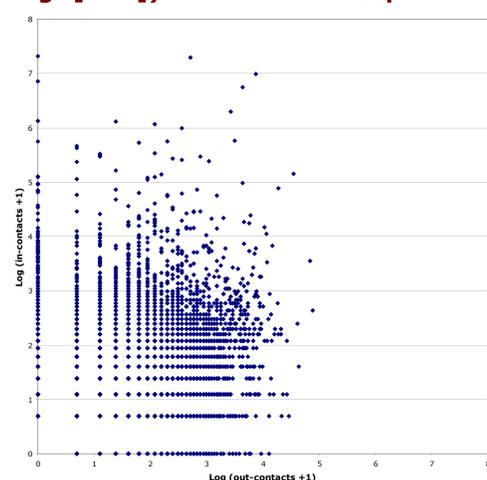
Results and Discussion

The annual mean number of contacts for the 15,000+ Scottish sheep holdings was consistently in the range 4.3 to 4.7 for the 4 years investigated (Table 1, Column 3). Here, we evaluate contact rate in terms of numbers of farm contacts (ignoring the frequency of contacts with a given farm within the year and the absolute numbers of sheep moved). Since movements to other kinds of holding or outside Scotland are ignored, these values represent both the mean number of in-contacts (farms sheep were received from) and out-contacts (farms sheep were sent to).

The mean values mask very high variances in in-contacts (Chart 1; Table 1, Column 4) and out-contacts (Chart 2; Table 1, Column 5). Both distributions are highly over-dispersed, i.e. most holdings send sheep to or receive sheep from few or no other holdings, but a small fraction send or receive from a large number of other holdings. The variance is particularly high for in-contacts, with instances of holdings (likely to be sheep dealers) receiving animals from over 1000 other holdings. High variances in contact rates can significantly increase R_0 (Woolhouse *et al.* 1997) but only if in-contact rate and out-contact rate are positively correlated. For these data the correlation is weakly positive, ranging from +0.07 to +0.11 over the 4 years of interest (Chart 3). The net effect of these second order moments is to increase R_0 by 32 to 44% (Table 1, Column 6).

Higher order effects may also influence R_0. Investigation of these requires contact rates for a whole population (not a sample) and calculation of the dominant eigenvalue of the contact matrix. This has previously only been done for small populations (e.g. Woolhouse *et al.* 1991) but the ARPACK software allows the calculation to be performed for much larger populations. Here, higher order effects act further to increase R_0 by up to 31% (Table 1, Column 7). Interestingly, higher order effects acted to decrease R_0 slightly in 2004-05. We conjecture that this reflects anomalous sheep trading patterns during that year.

In conclusion, this study demonstrates that: (i) the detailed structure of the network of movements of sheep between Scottish holdings does influence the potential for disease transmission; (ii) higher order effects not apparent from simple summary statistics (means, variances and covariances) can make a significant contribution; (iii) this contribution can show significant year-to-year variability.

References

Fèvre EM, Bronsvoort BMC, Hamilton KA, Cleaveland S, 2006. Animal movements and the spread of infectious diseases. Trends in Microbiol. 14:125-131.

Lehoucq RB, Sorensen DC Yang C, October 1997. {ARPACK} users' guide: Solution of large scale eigenvalue problems with implicitly restarted Arnoldi methods. Technical Report from http://www.caam.rice.edu/software/ARPACK/, Computational and Applied Mathematics, Rice University.

SAS Institute Inc., 2000-2004. SAS 9.1.3 Help and Documentation, Cary, NC, USA.

Woolhouse MEJ, Watts CH, Chandiwana SK, 1991. Heterogeneities in transmission rates and the epidemiology of schistosome infection. Proc Biol Sci. 245:109-14.

Woolhouse, MEJ, Dye C, Etard JF, Smith T, Charlwood JD, Garnett GP, Hagan P, Hii JL, Ndhlovu PD, Quinnell RJ, Watts CH, Chandiwana SK, Anderson RM, 1997. Heterogeneities in the transmission of infectious agents: implications for the design of control programs. Proc Natl Acad Sci USA. 94:338-42.

Woolhouse MEJ, Shaw DJ, Matthews L, Liu W-C, Mellor DJ, Thomas MR, 2005. Epidemiological implications of the contact network structure for cattle farms and the 20-80 rule. Biol Lett. 1:350-352.