

How to take into account the flock size in the analysis of epidemiological study of scrapie infection

Sandrine PHILIPPE¹, Didier CALAVAS¹, Christian DUCROT² & Pascal ROY³

1. AFSSA, Unité Epidémiologie, 31, avenue Tony Garnier, Lyon, France
2. INRA, Unité d'épidémiologie animale, Saint Genès-Champagnelle, France
3. Université Claude Bernard - Lyon 1, UMR 5558 du CNRS, Villeurbanne, France

Scrapie in sheep and goats is a natural and contagious disease, clinically described as soon as the 18th century and ever since then enzootic in many countries. Scrapie as bovine spongiform encephalopathy (BSE) in cattle or Creutzfeldt-Jakob disease in humans are all fatal diseases related to transmissible spongiform encephalopathies (TSE). In comparison with BSE, scrapie is considered harmless to humans, although the similarity between BSE in sheep and natural scrapie needs, if not a total eradication of this, at least to control its spread. The causal process relevant to understand the epidemiology of scrapie represents a complex interplay between factors acting on the flock and on animal level. Although, despite the major advances in the knowledge of the disease, the transmission process, especially the routes of transmission of scrapie within and between flocks remains insufficiently known to control the spread of the disease.

At animal level, the genetic susceptibility to scrapie is a major factor in the transmission process. The demographic factors: age, breed, geographic location need as well to be taken into account as potential biases.

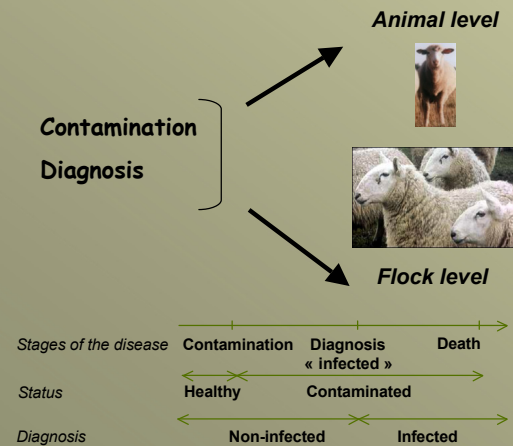
In addition, the flock size is assumed to play an essential and complex role in the transmission process. It has to be included in the analysis of epidemiological studies. For example, in a case-control study on scrapie risk in Norwegian flocks, the flock size has been included in the analysis as a covariate in a logistic regression (P. Hopp et al., *Prev. Vet.Med.*, 2001, 1622, 183-198).



Main purpose

To build a statistical model that allows to take into account the flock size in the analysis of epidemiological studies on scrapie transmission between flocks. Results are compared with those obtained using a logistic model including flock size as a covariate in a Case-control study.

Characteristics of the disease



Relationship between the risk of infection at animal level and at flock level

X_i denotes binary indicator contamination status, and $p_i = \Pr(X_i = 1)$
 Y_i denotes binary indicator infection diagnosis, and $p'_i = \Pr(Y_i = 1 | X_i = 1)$
 Probability for the i^{th} animal of the flock to be diagnosed as "infected"
 $\Pr(Y_i = 1) = \Pr(Y_i = 1 | X_i = 0) \Pr(X_i = 0) + \Pr(Y_i = 1 | X_i = 1) \Pr(X_i = 1)$
 Assuming that there is no "false positive",
 $\Pr(Y_i = 1) = p_i p'_i$

Lets, the i^{th} animal of a flock

For a sheep flock of k animals,

Y denote binary indicator infection diagnosis of the flock,
 Probability for a sheep to be diagnosed as "non-infected": $\Pr(Y = 0 | k) = \Pr(\prod_{i=1}^k Y_i = 0)$
 Assuming the independence the animals of the flock, $\Pr(Y = 0 | k) = \prod_{i=1}^k \Pr(Y_i = 0) = \prod_{i=1}^k (1 - p_i p'_i)$
 Since all animals of the flocks have the same probability ρ to be contaminated, i.e. $\forall i \in \{1, \dots, k\}, p_i = \rho$, and all animals contaminated of the flock have the same probability ρ' to be diagnosed as "infected", i.e. $\forall i \in \{1, \dots, k\}, p'_i = \rho'$,
 $\Pr(Y = 0 | k) = (1 - \rho \rho')^k$ or $\Pr(Y = 0 | k) = (1 - \theta)^k$, when ρ and ρ' are confused, $\theta = \rho \rho'$

Case-control study

The study was designed as a case-control study stratified on the geographical localisation and the breed of the flocks, with four controls for every case of each stratum. The unit of concern was the sheep flock. 444 flocks were included. Case flocks were flocks where at least one animal was suspected as infected through the scrapie surveillance network between January 1996 and July 1999, with pathological confirmation of the disease. Control flocks were selected among sheep flocks without Scrapie suspicion at the time of the study. The main putative risk factors were investigated.

Regression model

The corresponding regression model introduces the minus logarithm of the flock size as an offset
 $\log(-\log(1 - \Pr(Y = 1 | k))) = \log(-\log(1 - \theta)) + \log(k)$

Results

Risk factors	Logit link			complementary log-log link					
	Adjusted on stratum	Adjusted on stratum and flock size		Adjusted on stratum					
	OR	CI 95%	s.d.	OR	CI 95%	s.d.	OR	CI 95%	s.d.
Purchase of animals	1.94	[1.04;3.61]	0.61	1.70	[0.90;3.22]	0.55	1.44	[0.82;2.52]	0.41
Temporary contacts between flocks	0.77	[0.44;1.35]	0.22	0.74	[0.42;1.32]	0.22	0.50	[0.49;1.33]	0.21
Temporary stay of animals in other flocks	0.82	[0.34;1.94]	0.36	0.79	[0.33;1.90]	0.35	0.72	[0.33;1.59]	0.29
Temporary introduction of animals from other flocks	0.61	[0.17;2.08]	0.38	0.64	[0.18;2.32]	0.42	0.76	[0.24;2.45]	0.45

Standard deviation (s.d.)

References: S. Shiboski et al., *JID*, 1996, 174(Suppl 2): S188-200; S. Shiboski et al., *J Am Stat Assoc*, 1996, 174(Suppl 2): S188-200

Conclusion

The design of this case-control study justified the use of a complementary log-log link function. Parameter estimates differed from those obtained when a logistic model was fitted introducing the flock size as a covariate.

In this analysis, all animals of a flock were assumed to have the same probability to be contaminated. The next step would be to take into account the difference in disease susceptibility due for example to genetic factors or age.

