Advanced survival models for risk-factor analysis in scrapie.

Fabien Corbière, Francis Barillet, Olivier Andréoletti, Francis Fidelle,
Nathalie Laphitz-Bordet, François Schelcher, Pierre Joly

▶ To cite this version:

1   **Advanced survival models for risk factors analysis in scrapie.**

2

3   **Fabien CORBIERE[1,5*], Francis BARILLET[2], Olivier ANDREOLETTI[1], Francis**

4   **FIDELLE[3], Nathalie LAPHITZ-BORDET[4], François SCHELCHER[1], Pierre JOLY[5]**

5

6   1 – UMR Interactions Hôtes Agents Pathogènes, Ecole Nationale Vétérinaire de Toulouse, 23

7   chemin des Capelles, 31076 Toulouse Cedex, France

8   2 - Station d'amélioration des animaux, Institut National de la Recherche Agronomique, BP

9   27, 31326 Castanet-Tolosan cedex, France

10   3 – Centre Départemental d'Elevage Ovin, quartier Ahetzia 64130 ORDIARP, France

11   4 – Direction Départementale des Services Vétérinaires, cours Lyautey, 64000 PAU, France

12   5 – EMI 0338 (Biostatistique), Institut de Santé Publique Et Développement, Université

13   Victor Segalen Bordeaux 2, 146 rue Léo Saignat, 33076 Bordeaux Cedex, France

14

15   *** Corresponding author:**

16   Fabien Corbière,

17   EMI 0338 Biostatistique, Institut de Santé Publique Et Développement,

18   Université Victor Segalen Bordeaux 2

19   146 rue Léo Saignat, 33076 Bordeaux Cedex, France

20   (**fabien.corbiere@isped.u-bordeaux2.fr**)   **Tel (33) 5 57 57 45 79; Fax (33) 5 56 24 00 81**

21

22   **Running title:** PrP genotype influence on scrapie infection risk and incubation period.

23   Number of words in the summary: 227

24   Number of words in the main text: 4985

25   Number of tables: 6 - Number of figures: 2

26   **Summary**

27   Because of the confounding effects of long incubation duration and flock management,

28   accurate epidemiological studies of scrapie outbreaks are difficult to carry out. In this study,

29   641 Manech Red Faced sheep from 6 scrapie affected field flocks in Pyrenees Atlantiques,

30   France, were monitored for clinical scrapie over a 6 to 9 year period. Over this period, 170

31   scrapie clinical cases were recorded and half of the culled animals were submitted for post

32   mortem TSE diagnosis to assess their infectious status. Collected data were analyzed using a

33   "mixture cure model" approach, which allowed for the discriminating effect of PRP genotype

34   and flock origin on incidence and incubation period. Simulations were performed to evaluate

35   the applicability of such a statistical model to the collected data. As expected, ARR

36   heterozygote sheep were less at risk of becoming infected than ARQ/ARQ individuals and

37   had a delayed age at clinical onset. Conversely, when compared to ARQ/ARQ, the VRQ

38   haplotype was associated with an increased infection risk, but not a shorter incubation period.

39   Considering the flock effect, we observed that a high incidence rate was not associated with

40   shorter incubation periods and that incubation period could be significantly different in flocks

41   harboring similar infection risks. These results strongly support that other parameters like the

42   nature of the agent or flock management could interfere with epidemiological dynamics of the

43   infection in scrapie affected flocks.

44

45   **Introduction**

46   Transmissible spongiform encephalopathies (TSE) are neurodegenerative disorders occurring

47   primarily in sheep (scrapie), cattle (bovine spongiform encephalopathy - BSE), or humans

48   (Creutzfeldt-Jakob disease - CJD). They all share similar characteristics including long

49   incubation periods, a progressive pattern of disease and a clinical course resulting in death

50   (Fraser, 1976). Under natural conditions, it is considered that scrapie infection in sheep occurs

51   mainly after an early oral contamination (around birth) (Andreoletti *et al.*, 2000; Heggebo *et*

52   *al.*, 2000). Scrapie clinical onset generally occurs between 2 and 7 years (Detwiler & Baylis,

53   2003).

54   In prion diseases, the accumulation of an abnormal isoform (PrPsc) of a normal cellular

55   protein (PrPc) in tissues from infected individuals is currently considered as a disease

56   hallmark. Most of the diagnostic tests currently available are based on biochemical detection

57   of the abnormal protein (McKinley *et al.*, 1983; Race *et al.*, 2001). However post-mortem

58   tests, as carried out in the current European surveillance program (rapid test on the obex), are

59   reliable only for detection of infected animals in the second half of incubation period.

60   Because of long TSE incubation periods, data analysis is difficult without reference to flock

61   demography and management. Indeed infected individuals could be culled or die from other

62   causes before the clinical onset (intercurrent diseases – economical reasons). In this situation

63   no reliable information about their infectious status will be available (Begara-McGorum *et al.*,

64   2000; Ryder *et al.*, 2001; Thorgeirsdottir *et al.*, 2002; Billinis *et al.*, 2004).

65   Evaluation of genetic and environmental risk factors in scrapie have been conducted mainly

66   using case-control designs in which a set of affected animals is compared to their healthy

67   flock-mates or to a reference population (Hunter *et al.*, 1997; Thorgeirsdottir *et al.*, 1999;

68   Tranulis *et al.*, 1999; O'Doherty *et al.*, 2002; Acin *et al.*, 2004; Baylis *et al.*, 2004; Billinis *et*

69   *al.*, 2004; Tongue *et al.*, 2006). Such approaches revealed that TSE susceptibility in sheep is

70   controlled mainly by polymorphisms at codons 136 (T,V,A), 154 (R,H) and 171 (R,Q,H,K) of

71   the PRP gene (Clouscard *et al.*, 1995; Hunter *et al.*, 1996). $V^{136}R^{154}Q^{171}$/VRQ, ARQ/VRQ

72   and ARQ/ARQ animals are usually considered the most susceptible to scrapie, whereas

73   homozygous or heterozygous AHQ and heterozygous ARR animals only show a marginal

74   susceptibility, ARR/ARR sheep being considered to be fully clinically resistant (Detwiler &

75   Baylis, 2003).

76   Surveys based on long term individual monitoring of an exposed population are less subject to

77   sampling bias (Tongue *et al.*, 2006). Consequently, they should be considered as more

78   relevant than case-control or cross-sectional studies for an accurate evaluation of the effect of

79   environmental or genetic factors on infection rate and incubation period.

80   'Cure models' are part of the mixture models family (Bohning & Seidel, 2003). In 'mixture

81   cure models' it is considered that the studied population is a mixture of susceptible (i.e. that

82   may undergo the event of interest) and non-susceptible individuals (i.e. that will never

83   undergo the event of interest) (Farewell, 1982). Unlike classical survival analysis, they allow

84   a separate estimation of covariate effects on the incidence and incubation length. Cure models

85   also allow the estimation of the proportion of healthy (or conversely infected) individuals in a

86   population, including individuals that did not last the total length of the study (Lam *et al.*,

87   2005)

88   In this study we propose a model based on 'mixture cure models' approach for scrapie

89   epidemiological analysis. Robustness and reliability boundaries of the model were assessed

90   by simulations before analyzing data collected over 6 to 9 years in 6 naturally scrapie affected

91   flocks in Pyrenees Atlantiques, France.

92

93   **Material and Method**

94   **Model design**

95   In the model, death from scrapie is considered as the event of interest. Clinical scrapie cases

96   are considered to be scrapie-infected and have uncensored observations on their lifetime. For

97   apparently healthy sheep that are removed from the flock (right-censored records) it is

98   unknown whether they are scrapie-free or infected but removed prior to onset of clinical

99   scrapie. For each right censored animal, the model computes the probability to be scrapie-

100   infected, given its age and covariates information.

4

101   The model assumes that:

102   i) most or all deaths from scrapie occurs during a determined age-period ,

103   ii) monitoring is long enough for clinical onset to have appeared in most of the infected

104   animals,

105   ii) the longer an animal lives, the lower the probability of it being scrapie infected. Animals

106   which live longer than the last observed scrapie clinical case had an extremely low probability

107   (if not zero) of incubating scrapie.

108

109   If U is the indicator denoting SI animals (i.e. $U=1$ if the animal is scrapie-infected and

110    $U=0$ if non infected), and T is a non-negative random variable denoting the failure time of

111   interest, defined only if $U=1$, the mixture cure model is given as follows:

$$S(t|x,z) = \pi(z)\ S(t|U=1,x) + (1-\pi(z))$$

113   where $S(t|x,z)$ is the unconditional (marginal) survival function of T for the entire population

114   under study (that is SF and SI groups), and $\pi(z)=P(U=1|z)$ is the probability of being infected

115   given a covariate vector $z = (z_1,...,z_q)'$. $S(t|U=1,x) = P(T>t|U=1,x)$ is the conditional survival

116   function for SI animals given a covariate vector $x = (x_1,...,x_m)'$ (it may include the same

117   covariate as z). The use of conditional attached to this function is to stress that the distribution

118   of time refers not to the whole group of animals but only to the animals that are in the SI

119   group. Note that $S(t|x,z)\rightarrow 1-\pi(z)$ as $t \rightarrow \infty$, where $1-\pi(z)$ represents the proportion of non

120   infected animals. When $\pi = 1$, that is when no scrapie-free portion is assumed, the model

121   reduces to the traditional survival model. Whether the inclusion of a proportion of scrapie-free

122   animals leads to a significantly better fit to the data than a traditional survival model with no

123   scrapie-free animals can be tested by the deviance test statistics proposed by (Maller & Zhou,

124   1996).

125

126   Various parametric and semiparametric approaches have been proposed for mixture cure

127   models (Peng & Carriere, 2002; Lam *et al.*, 2005). For modelling the influence of exploratory

128   variables on the incidence, a logistic regression model is usually chosen (Kuk & Chen, 1992;

129   Peng & Dear, 2000) :

$$\pi(z) = p(U=1|z)=\exp(\beta' z)/(1+\exp(\beta' z))$$

131   where β is the vector of regression parameters associated to z and contains an intercept. The

132   conditional survival function of infected animals is modelled through the semiparametric Cox

133   proportional hazards model (Cox, 1972), which is given by :

$$S(t|U = 1, x) = S_0(t|U = 1)^{\exp(\gamma' x)}$$

135   where γ is the vector of regression parameters associated to x and $S_0(t|U=1)$ is the baseline

136   conditional survival function, which is left unspecified.

137   Through the vectors of regression parameters β and γ the mixture survival model is able to

138   separate the covariates effect on the incidence and the latency. An estimate of the true

139   proportion of SI animals SI_{pop}, given z and x, is provided by averaging the individual

140   probabilities $P_i(U=1|z_i,x_i)$:

$$SI_{pop} = \frac{1}{N}\sum_i^N P_i(U=1|z_i,x_i) = \frac{1}{N}\sum_i^N \left[\delta_i + (1-\delta_i)\frac{\pi_i(z_i)S(t_i|U=1,x_i)}{1-\pi_i(z_i)+\pi_i(z_i)S(t_i|U=1,x_i)} \right]$$

142   where δ_i is the censoring indicator with 1 if t_i is uncensored and 0 otherwise. Obviously if $\delta_i =$

143   1, then $P_i(U=1)=1$. When $\delta_i =0$, then $P_i(U=1|z_i, x_i)$ will depend on the survival length and will

144   drop to zero as $t \to \infty$. Note that the better the model fits the data, through covariate vectors z'

145   and x', the more accurate is the estimation of the proportion of SI animals.

146

147   **Simulation studies**

148   Simulations were conducted to test i) the ability of the model to estimate the proportion of

149   scrapie-infected animals and to discriminate covariate effects on the infection risk and

150   incubation duration and ii) the effect of individual monitoring length on model outputs. We

151   assumed that non infected animals (named Scrapie Free animals SF) would never die from

152   scrapie. Consequently observations realized on SF animals were right censored. Scrapie

153   infected animals (named Scrapie-Infected SI), either died from clinical scrapie (uncensored

154   records) or were eliminated from the flock before clinical onset (right-censored). Simulations

155   were performed using (i) genetic and biological parameters (infection rates, ages at clinical

156   death, and flock demography) described in scrapie outbreaks or already used in mathematical

157   modelling (simulation design 1) (Matthews *et al.*, 2001; Hagenaars *et al.*, 2003; Hopp *et al.*,

158   2003; Baylis *et al.*, 2004; Eglin *et al.*, 2005; Touzeau *et al.*, 2005) and (ii) ages at death from

159   scrapie reflecting observation realized in our data set (simulation design 2).

160   The capacity of the model to separate covariate effects on incidence and incubation length

161   (age at death) was assessed by generating two independent binary covariates, one (Z_1)

162   affecting only the incidence and the other (Z_2) affecting only the incubation duration. The

163   incidence is given the logistic form $\pi(z_1,z_2)=\exp(\beta_0+\beta_1Z_1+\beta_2Z_2)/(1+\exp(\beta_0+\beta_1Z_1+\beta_2Z_2))$, where

164   β_1 and β_2 are the effects of covariates Z_1 and Z_2 on the proportion of infected individuals,

165   respectively. Since Z_2 should have no effect on the incidence, β_2 was set to 0. Thus, the

166   proportion of infected animals is $\pi(z_1=0,z_2)=\exp(\beta_0)/(1+\exp(\beta_0))$ for animals sharing $Z_1=0$ and

167    $\pi(z_1=1,z_2)=\exp(\beta_0+\beta_1)/(1+\exp(\beta_0+\beta_1))$ for those sharing $Z_1=1$. We set $\beta_0=$ -0.5 and $\beta_1=$ -1, so

168   that the corresponding proportions of infected sheep are 37.7 % (animals with $Z_1=$ 0) and

169   18.22 % (animals with $Z_1=1$), respectively.

170   The lognormal distribution was used as the distribution function for life expectancies of

171   infected animals, with survival function $S(t|U=1)=1-\Phi[(\ln t-\mu-\gamma_1Z_1-\gamma_2Z_2)/\sigma]$, where Φ is the

172   distribution function of the standard normal law. Unlike for the incidence part, γ_1 was set to 0,

173   because Z_1 should have no effect on the latency, whatever the value of Z_2. In simulation

174   design, 1, the scale (μ) and shape (σ) parameters for the lognormal distribution function were

7

175 set to 1.2 and 0.4 respectively. In the absence of censoring, events of interest (scrapie deaths)

176 were allowed to occur at median age 3.3 years old (interquartile range = 2.5 - 4.2,) for

177 individuals with $Z_2=0$. We set $\gamma_2 = 0.3$ so that infected individuals with $Z_2=1$ would die later,

178 at median age 4.4 years old (interquartile range = 3.4 – 5.7). In simulation design 2, we set μ

179 = ln 2, $\sigma = 0.5$, $\gamma_1 = 0$ and $\gamma_2 = 0.375$, so that median ages at clinical onset were 2.00

180 (interquartile range = 1.5 – 2.8) and 3.00 years old (interquartile range = 2.1 – 4.2) (figure 1a).

181 Monitoring length were generated according to the Weibull distribution $W(\lambda,\rho)$ with shape (λ)

182 and scale (ρ) parameters. Thirteen different scenarios were investigated with median life

183 expectancy (meaning monitoring length) ranging from 2.5 years to 9.5 years old. These

184 scenarios covered a large panel of flock management policies and demography. For each

185 scenario and simulation design, 500 independent data sets, each constituted with 500

186 individuals, were generated and submitted to model analysis. The absolute biases $(B(\hat{c})= \Sigma_i$

187 $(\hat{c}_i-c_0)/500)$ and mean squared errors $(MSE(\hat{c})= \Sigma_i\ (\hat{c}_i-c_0)^2/500)$, where $\hat{c}_i$ are the estimates of

188 $c_0,$ were computed for the 5 parameters estimates.

189 Sample generation and model computations were performed using SAS software (SAS-PC

190 system®, Version 8.2 for Windows, SAS Institute Inc., Cary, NC, 1999-2001).

191

192 **Flocks**

193 Investigations were carried out on 6 scrapie naturally affected dairy flocks, bred by private

194 farmers, in Pyrenees Atlantiques, France. These flocks had been involved in a long-term

195 scrapie epidemiology research project since 1994 (flock C) and 1998 (other 5 flocks). Sheep

196 were all Manech Red Faced pure-breed. Table 1 shows, for each flock, the average size and

197 the estimated year of first occurrence of scrapie. The high incidences of scrapie clinical

198 suspicions (confirmed or not) in ewes born prior to enrolment in the research project, suggest

199 different, but whatever high, infection pressures (table 1).

200   At the time of enrolment, PrP polymorphism at codons 136, 154 and 171 of the PRP gene was

201   determined for all present sheep, including breeding males and females and replacement

202   lambs. Birth date, pedigree, date and cause of death or removal from the flock were

203   systematically recorded for all the animals up to January 2006. No valuable information could

204   be collected for male and female lambs slaughtered between 1 to 3 months old.

205   To comply with the requirement of adequate monitoring length and the provision of high

206   quality data (including reliable diagnosis) only a few birth cohorts were considered within

207   each flock for the analysis. Birth cohorts were selected in which all scrapie clinical suspicions

208   were confirmed by histopathology and complete PrP genotype profiles were available.

209   The data set submitted for analysis consisted of the 1998 birth cohort (born between October

210   and December 1997) for flock A, the 1999 birth cohort (born between October and December

211   1998) for flocks B, D, E and F and the 1995, 1996, 1997 birth cohorts in flock C (animals

212   born in November and December 1994, 1995 and 1996, respectively) (table 2). Only

213   homebred animals were included in the analysis, while purchased sheep (n=10) were not

214   considered. Finally, our sample comprised 641 sheep including 170 scrapie clinical cases.

215   Apparently healthy culled sheep were submitted for PrPsc detection on obex and palatine

216   tonsils by ELISA (Platelia BSE detection Kit, Biorad) and immunohistochemistry

217   (Andreoletti *et al.*, 2002). All the sheep included in our initial sample were not examined,

218   either because of failure to track these animals during the elimination process, or because they

219   were still alive at the time of writing (January 2006). From these last animals palatine tonsils

220   biopsies were performed each year from 2002 to 2005 (inclusive) for PrPsc

221   immunohistochemistry detection (Andreoletti *et al.*, 2002). To account for missing

222   information, a one-side 95 percent confidence interval for the true proportion of infected

223   animals in each PRP genotype group was calculated using the hypergeometric law.

224   In the mixture cure model analysis, PRP genotype and the flock were used as covariates and

225   age at death from clinical scrapie was considered as the survival measurement. Ninety-five

226   percent confidence intervals for adjusted odds ratios (OR) from the logistic part and adjusted

227   relative risks (RR) from the Cox PH part of the mixture cure model were computed using the

228   bias corrected, accelerated bootstrap method (Davison & Hinkley, 1997).

229   **Results**

230   **Simulations**

231   As expected, the longer an animal lived (long individual monitoring), the smaller were the

232   mean absolute biases and mean squared errors (MSE, not shown) for the different estimates

233   (figure 1B and 1C). Biases and MSE were acceptable for inference when the monitoring time

234   was longer than the median (theoretical) incubation duration. For the first set of simulated

235   incubation times (median age at death 3.3 and 4.45 years old), the estimates of the proportion

236   of infected animals (β_0) and the effects of Z_2 on the incidence (β_2) were highly biased

237   (absolute bias over 0.1) for median monitoring times less than 5.5 years (figure 1B). For the

238   second set of simulated incubation durations (median age at death 2 and 3.3 years old) similar

239   results were obtained for median monitoring times under 3.5 years (figure 1C). In our

240   population median age at death in scrapie affected animals was 1.7 years (figure 1A and table

241   3) and the median monitoring time was 4.90 years (table 4). Under these conditions, the

242   models outputs were expected to be relevant in the analysis our data set.

243

244   **PrP genetic structure of the studied cohorts**

245   Amongst the studied cohorts, the ARQ and ARR alleles were dominant while AHQ and VRQ

246   alleles were rare (table 2). The sample's genetic structure is consistent with previously

247   reported PRP polymorphism frequencies in the Manech Red Faced breed (Palhiere *et al.*,

248   2002).

249   Because of low numbers of individuals in some PrP genotypes, animals were grouped

250   according to their level of susceptibility to classical scrapie (Defra, 2003). ARQ/ARQ,

251   AHQ/AHQ and AHQ/ARQ sheep were considered in a medium risk group named S/S

252   (n=343). ARR/ARQ, ARR/AHQ and ARR/VRQ animals were included in a low risk group

253   named R/S (n=212). ARQ/VRQ, AHQ/VRQ and VRQ/VRQ animals were included in a

254   single high risk group named VRQ/x (n=49) (table 3). Considering these PrP genotype

255   groups, the genetic structure was not statistically different in the 6 selected flocks (Chi-square

256   test with 15 degrees of freedom χ^2= 15.95, p=0.38).

257   **Scrapie clinical cases**

258   Clinical scrapie cases were mainly observed in ARQ/ARQ (n=131 – 77.06%), and ARQ/VRQ

259   (n=28 – 16.47%) genotypes, while heterozygote ARR were poorly affected (R/S sheep n= 9 –

260   5.29%) (table 2 and 3). High incidences in susceptible PrP genotypes ARQ/ARQ and

261   ARQ/VRQ suggested a high infection pressure. No clinical cases were observed, neither in

262   ARR/ARR (n=37), nor in ARR/VRQ (n=3), ARR/ARH (n=3) or AHQ/VRQ (n=3) animals.

263   However the number of animals with these three last genotypes was too low to draw any

264   conclusion.

265   The Kaplan-Meier plots of the survival distribution functions for scrapie clinical occurrence

266   indicate the absence of new scrapie cases after 5.54 years, whatever the genotype group

267   considered (figure 2). This lack of new clinical cases fulfils a basic requirement allowing the

268   application of the Cox PH mixture cure model, i.e. (i) most or all death due to scrapie

269   occurred in a defined age period (ii) the monitoring length was sufficient to allow almost all

270   infected animals to show clinical signs.

271

272   **Monitoring length in clinically healthy sheep**

273   A large number (n=264 – 60.83%) of clinically healthy animals were eliminated by breeders

274   for husbandry reasons at younger ages than the last observed clinical case (table 4). However,

275   94.24 % of these symptomless sheep lived older than the median age at clinical onset (1.70

276   years old) and 72.12 % reached 3.57 years old, which represented the 90th percentile of the

277   age distribution of scrapie clinical cases.

278   Because of the implementation of French TSE legislation at the beginning of 2003, breeders

279   had to remove VRQ carrier animals from scrapie affected flocks. Consequently VRQ animals,

280   mainly ARR/VRQ, were eliminated earlier than expected and had a statistically shorter

281   follow-up than R/S animals (analyse of variance, $F_{431,2} = 13.13$, $p<10^{-4}$).

282

283   **Active detection of subclinically infected sheep**

284   From the 471 clinically healthy sheep (with no clinical scrapie) eliminated during the study,

285   220 (46.71 %) were submitted to post mortem for PrPsc detection test (mean age 4.72 years

286   old – 95% confidence interval = 1.87- 7.57). Sampled animals represented 70% of VRQ/x and

287   72.04 % of S/S but only 24.13 % of R/S and 13.51 % of R/R (table 5).

288   From the tested animals only four ARQ/ARQ sheep (aged 1.89, 2.25, 2.32 and 4.39 years),

289   and one ARQ/VRQ sheep (3.04 years) were found positive. Fifty nine animals initially

290   included in the study were still alive at the time of writing (5 ARR/ARR, 42 ARR/ARQ and

291   12 ARQ/ARQ). No PrPsc was detected in any of these animals using tonsil biopsy. According

292   to the hypergeometric law, the one sided 95% confidence interval for the number of

293   apparently healthy but infected animals eliminated from the flocks was 5 to 25 (table 5).

294

295   **Results from the mixture cure model analysis**

296   The deviance statistic test ($\chi^2_{01}=30.93$, $p<10^{-4}$) indicated that incorporating a scrapie-free

297   fraction provided a better fit to the data than the traditional Cox proportional hazards model

298   and that estimated effects were more relevant.

299   **Proportion of infected animals**

300   ARR/ARR animals were not included in the analysis, because there were no confirmed

301   clinical cases of this genotype. According to the proposed model, the predicted number of

302   subclinically infected animals was 20.75 (respectively 3.65 in the R/S group, 14.83 in the S/S

303 group and 3.22 in the VRQ carrier group, point estimate minus number of clinical cases)

304 (table 5). Strikingly these predicted numbers were in close agreement with those obtained by

305 the active detection of subclinical cases.

Genotype influence on incidence and incubation duration

307 Results from the Cox mixture cure model indicated that ARR heterozygote animals were at

308 lower risk of infection than S/S animals (table 6). Conversely VRQ allele carriers (excluding

309 ARR/VRQ) were at higher risk of being infected. Age at clinical onset was significantly

310 delayed in R/S animals when compared to S/S animals. No significant difference was found

311 between S/S and VRQ allele carrier (table 6).

Flock effect on incidence and incubation length

313 Because of insufficient numbers of ARQ/VRQ or ARR/ARQ clinical scrapie cases in some

314 flocks the analysis was restricted the ARQ/ARQ animals (table 6). The risk for ARQ/ARQ

315 animals of being infected was significantly lower in flock A, D and E than in flock C (table

316 6). No significant difference was observed between flock C, flock B and F. Interestingly age

317 at clinical onset was significantly shortened in some flocks with a lower infection risk (flock

318 D compared to flock C). Meanwhile, age at clinical onset could be significantly different in

319 flocks harbouring a similar infection risk: infected animals from flock F had delayed clinical

320 onset compared to flock C while in flock B it was shortened. Taken together, these results

321 suggest strongly that age at clinical onset and infection risk are not associated parameters.

322

Discussion

Working hypothesis

325 Only a few studies of scrapie outbreaks, based on longitudinal monitoring, have been

326 published (Elsen *et al.*, 1999; Diaz *et al.*, 2005). Using survival analysis, they aimed at

327 determining the influence of PrP genotype, rearing type and dam clinical status on individual

328 risk of developing clinical scrapie. Survival analysis assumes intrinsically that if, in an

329 exposed population, a long enough and complete surveillance of individuals is possible, each

330 would experience the event of interest i.e. clinical scrapie. For TSE, this hypothesis is

331 obviously inadequate. Indeed, under natural exposure, ARR/ARR animals and a large

332 proportion of susceptible genotype animals will remain uninfected (Elsen *et al.*, 1999).

333 Because a mixed population of susceptible and non-susceptible individuals is considered,

334 'mixture cure models' appear to be an attractive approach for scrapie epidemiological

335 analysis. However fulfilment of several conditions is compulsory for their sound application.

336 Such constraints require hypotheses about scrapie pathogenesis and biology. Amongst the

337 basic hypotheses we considered were that animals born in an infected flock, if not infected in

338 their early life, would remain negative. Under natural exposure conditions, scrapie

339 contamination is considered to occur preferentially around birth (Andreoletti *et al.*, 2000;

340 Heggebo *et al.*, 2000; van Keulen *et al.*, 2000). We therefore hypothesized that age at death

341 from scrapie (clinical stage) was a relevant approximation of the incubation period. Moreover

342 animal susceptibility seems to decrease dramatically with age (Hourrigan *et al.*, 1979;

343 Andreoletti *et al.*, 2000). Clinical cases have been reported in young and adult susceptible

344 animals introduced to infected flocks (Hourrigan *et al.*, 1979; Ryder *et al.*, 2004) but the

345 importance, relative to the neonatal contamination, of such lateral transmission in adult sheep

346 could not be estimated. The other main hypothesis we made was that a very few (if not zero)

347 infected individuals would be alive at the end of the study. The observed survival distribution

348 plots were consistent with this hypothesis. However, existence of long term subclinical

349 carriers remains a major question of scrapie epidemiology. Currently, it is impossible to

350 assume that an apparently healthy animal (whatever the test used to establish the infectious

351 status) is not incubating scrapie. Recent description of atypical cases or Nor98 cases in old

352 and apparently healthy animals, and difficulties in assessing the diagnosis, sustain the long

353 term subclinical carriers' hypothesis (Benestad *et al.*, 2003; Le Dur *et al.*, 2005). However,

354 atypical scrapie occurs at a very low detected prevalence level (3 to 11 cases per 10000

355 examined) and in most cases, only one to three cases could be detected in stamped-out

356 affected flocks (De Bosschere *et al.*, 2004; Onnasch *et al.*, 2004; Orge *et al.*, 2004). This

357 implies that approximately 0.2 to 0.8 sheep could have be infected with atypical scrapie in the

358 considered flocks, which is negligible when compared to the number of classical scrapie cases

359 in the studied cohorts (n=170). At the population level, the influence of an atypical case on

360 the model outputs was considered to be negligible.

361   Finally, even if the hypothesis of some adult lateral transmission and long term subclinical

362 carriers could not definitely be ruled out, both phenomena seemed marginal enough in our

363 population to avoid major transgression from the model application. From simulations, major

364 biases were only observed when the monitoring length was shorter than the median

365 (theoretical) incubation duration. Similar trends were obtained by (Yu *et al.*, 2004) when

366 studying the influence of the follow-up length on the cure fraction estimation for several

367 human cancers. The monitoring length in the studied sheep was long enough to insure small

368 biases for the estimates of PRP genotype and flock effects.

369

370 **Asymptomatic culled animals**

371 The mixture cure model approach allowed the estimation of infected individuals and included

372 those eliminated while incubating the disease. Model outputs and laboratory findings were in

373 close agreement, and indicated a very low number of sheep removed while incubating scrapie.

374 This result is consistent with observations from another study based on longitudinal survey in

375 a Texel flock (Baylis *et al.*, 2002). It contrasts, however, with other publications in which

376 high numbers of scrapie incubating animals were described (Thorgeirsdottir *et al.*, 2002;

377   Billinis *et al.*, 2004). Similarly, the modelling of a scrapie outbreak in a cheviot flock

378   predicted a high ratio of infections to cases (2.2:1) (Matthews *et al.*, 2001).

379   Discrepancies between these results lay certainly in the data collection plan. Studies reporting

380   a high proportion of asymptomatic animals were based on cross-sectional designs with data

381   collected at stamping-out. In our study, most sheep were culled after a long individual

382   monitoring period which allowed scrapie clinical onset. As indicated by our simulations,

383   shorter monitoring lengths, as modelled by (Matthews *et al.*, 2001) (median length 3.00

384   years), would have resulted in the observation of fewer clinical cases and a higher number of

385   subclinical cases.

386

387   **Genetic susceptibility to scrapie and incubation period**

388   The comparison of the fit provided by the mixture cure model and the traditional Cox

389   Proportional model indicated than our approach was more relevant when analysing the PrP

390   genotype and flock effects. According to our results, with ARQ homozygote animals as the

391   baseline, VRQ carriers were at higher risk of infection and ARR heterozygotes at lower risk.

392   This is consistent with most already published studies based on data collected from culled

393   flocks (Thorgeirsdottir *et al.*, 1999; Tranulis *et al.*, 1999; Acin *et al.*, 2004; Billinis *et al.*,

394   2004).

395   Incubation length is a major feature of TSE phenotype. In our population, while clinical signs

396   were delayed in ARR heterozygotes compared to ARQ homozygotes, no difference could be

397   observed between ARQ/VRQ carriers and ARQ homozygotes. Similar phenomenon was

398   observed in an Irish flock (O'Doherty *et al.*, 2002). However it differed from estimations

399   obtained in a French Romanov flock (Elsen *et al.*, 1999) and in a Texel flock (Baylis *et al.*,

400   2002). In both these naturally affected scrapie flocks significant differences in age at death

401   were reported between ARQ/ARQ and ARQ/VRQ.

17

402   In sheep, experimental challenge indicated that incubation period depends both on sheep

403   genotype and TSE isolate. While most scrapie isolates will produce shorter incubation periods

404   in VRQ allele carriers, other TSE agent such as BSE will behave conversely (Foster *et al.*,

405   2001; Jeffrey *et al.*, 2006). In this context, difference in agent (strain) could be a possible

406   explanation for the observed variability.

407   In rodent scrapie models, it has been demonstrated that attack rate and incubation length

408   variations could be observed according to the infectious dose. Low infectious dose could lead

409   to incubation period lengthening and decreased infection efficiency (Kimberlin & Wilesmith,

410   1994; Jacquemot *et al.*, 2005). In natural scrapie there is no available estimation of the actual

411   infectious pressure. Because of differences in the observed incidences, the infection pressure

412   is usually considered to be different according to the cohort considered within a flock or

413   between flocks (Baylis *et al.*, 2002; Touzeau *et al.*, 2005). In our study, scrapie incubation

414   duration appeared not to be associated with the infection rate. Age at clinical onset in

415   ARQ/ARQ infected animals also clearly differed (shorter incubation period in our study) from

416   previously reported in animals bearing the same genotype (Woolhouse *et al.*, 1998; Elsen *et

417   al.*, 1999; O'Doherty *et al.*, 2002; Redman *et al.*, 2002; Baylis *et al.*, 2004). Taken together

418   these observations could suggest that (i) biologically different scrapie agents are involved in

419   the different flocks or that (ii) other factors linked to flock management could influence

420   incidence and incubation.

421   To go further, the evaluation of agent biodiversity in the studied cohorts from these flocks is

422   ongoing through biochemical studies and bioassays. Meanwhile the effect of increasing

423   infectious dose on incubation length in animals bearing similar PRP genotypes and which

424   were orally contaminated at birth is under investigation.

425

426   The mixture cure model presented here has revealed an interesting tool to analyze data

427   collected from longitudinal surveys in naturally affected scrapie flocks. Its main constraint is

428   the requirement of a long enough individual monitoring period. Finally, because such models

429   allow for the combinatory analysis of several covariate effects they should be considered as a

430   potential powerful tool for epidemiological analysis in animal diseases.

431

References

Acin, C., Martin-Burriel, I., Goldmann, W., Lyahyai, J., Monzon, M., Bolea, R., Smith, A., Rodellar, C., Badiola, J. J. & Zaragoza, P. (2004). Prion protein gene polymorphisms in healthy and scrapie-affected Spanish sheep. *J Gen Virol* **85**, 2103-2110.

Andreoletti, O., Berthon, P., Levavasseur, E., Marc, D., Lantier, F., Monks, E., Elsen, J. M. & Schelcher, F. (2002). Phenotyping of protein-prion (PrPsc)-accumulating cells in lymphoid and neural tissues of naturally scrapie-affected sheep by double-labeling immunohistochemistry. *J Histochem Cytochem* **50**, 1357-1370.

Andreoletti, O., Berthon, P., Marc, D., Sarradin, P., Grosclaude, J., van Keulen, L., Schelcher, F., Elsen, J. M. & Lantier, F. (2000). Early accumulation of PrP(Sc) in gut-associated lymphoid and nervous tissues of susceptible sheep from a Romanov flock with natural scrapie. *J Gen Virol* **81**, 3115-3126.

Baylis, M., Chihota, C., Stevenson, E., Goldmann, W., Smith, A., Sivam, K., Tongue, S. & Gravenor, M. B. (2004). Risk of scrapie in British sheep of different prion protein genotype. *J Gen Virol* **85**, 2735-2740.

Baylis, M., Goldmann, W., Houston, F., Cairns, D., Chong, A., Ross, A., Smith, A., Hunter, N. & McLean, A. R. (2002). Scrapie epidemic in a fully PrP-genotyped sheep flock. *J Gen Virol* **83**, 2907-2914.

Begara-McGorum, I., Clark, A. M., Martin, S. & Jeffrey, M. (2000). Prevalence of vacuolar lesions consistent with scrapie in the brains of healthy cull sheep of the Shetland Islands. *Vet Rec* **147**, 439-441.

Benestad, S. L., Sarradin, P., Thu, B., Schonheit, J., Tranulis, M. A. & Bratberg, B. (2003). Cases of scrapie with unusual features in Norway and designation of a new type, Nor98. *Vet Rec* **153**, 202-208.

Billinis, C., Psychas, V., Leontides, L., Spyrou, V., Argyroudis, S., Vlemmas, I., Leontides, S., Sklaviadis, T. & Papadopoulos, O. (2004). Prion protein gene polymorphisms in healthy and scrapie-affected sheep in Greece. *J Gen Virol* **85**, 547-554.

Bohning, D. & Seidel, W. (2003). Editorial: recent developments in mixture models. *Comput Stat Data An* **41**, 349-357.

Clouscard, C., Beaudry, P., Elsen, J. M., Milan, D., Dussaucy, M., Bounneau, C., Schelcher, F., Chatelain, J., Launay, J. M. & Laplanche, J. L. (1995). Different allelic effects of the codons 136 and 171 of the prion protein gene in sheep with natural scrapie. *J Gen Virol* **76 (Pt 8)**, 2097-2101.

Cox, D. R. (1972). Regression Models and Life-Tables. *J R Stat Soc Ser B Met* **34**, 187-220.

Davison, A. C. & Hinkley, D. V. (1997). *Bootstrap methods and their application.* New-York: Cambridge University Press.

De Bosschere, H., Roels, S., Benestad, S. L. & Vanopdenbosch, E. (2004). Scrapie case similar to Nor98 diagnosed in Belgium via active surveillance. *Vet Rec* **155**, 707-708.

Defra (2003). National Scrapie Plan for Great Britain. Department for food, Environment and rural Affairs: http://www.defra.gov.uk/animalh/bse/othertses/scrapie/nsp/publicatsrpts/tables.htm.

Detwiler, L. A. & Baylis, M. (2003). The epidemiology of scrapie. *Rev Sci Tech* **22**, 121-143.

Diaz, C., Vitezica, Z. G., Rupp, R., Andreoletti, O. & Elsen, J. M. (2005). Polygenic variation and transmission factors involved in the resistance/susceptibility to scrapie in a Romanov flock. *J Gen Virol* **86**, 849-857.

Eglin, R. D., Warner, R., Gubbins, S., Sivam, S. K. & Dawson, M. (2005). Frequencies of PrP genotypes in 38 breeds of sheep sampled in the National Scrapie Plan for Great Britain. *Vet Rec* **156**, 433-437.

Elsen, J. M., Amigues, Y., Schelcher, F., Ducrocq, V., Andreoletti, O., Eychenne, F., Khang, J. V., Poivey, J. P., Lantier, F. & Laplanche, J. L. (1999). Genetic susceptibility and transmission factors in scrapie: detailed analysis of an epidemic in a closed flock of Romanov. *Arch Virol* **144**, 431-445.

Farewell, V. T. (1982). The use of mixture models for the analysis of survival data with long-term survivors. *Biometrics* **38**, 1041-1046.

Foster, J. D., Parnham, D., Chong, A., Goldmann, W. & Hunter, N. (2001). Clinical signs, histopathology and genetics of experimental transmission of BSE and natural scrapie to sheep and goats. *Vet Rec* **148**, 165-171.

Fraser, H. (1976). The pathology of a natural and experimental scrapie. *Front Biol* **44**, 267-305.

Hagenaars, T. J., Donnelly, C. A., Ferguson, N. M. & Anderson, R. M. (2003). Dynamics of a scrapie outbreak in a flock of Romanov sheep--estimation of transmission parameters. *Epidemiol Infect* **131**, 1015-1022.

Heggebo, R., Press, C. M., Gunnes, G., Lie, K. I., Tranulis, M. A., Ulvund, M., Groschup, M. H. & Landsverk, T. (2000). Distribution of prion protein in the ileal Peyer's patch of scrapie-free lambs and lambs naturally and experimentally exposed to the scrapie agent. *J Gen Virol* **81**, 2327-2337.

Hopp, P., Webb, C. R. & Jarp, J. (2003). Monte Carlo simulation of surveillance strategies for scrapie in Norwegian sheep. *Prev Vet Med* **61**, 103-125.

Hourrigan, J., Klingsporn, A. L., Clark, W. W. & de Camp, M. (1979). Epidemiology of scrapie in the Unided States. In *Slow Transmissible Dideases of the Nervous System*, pp. 331-356. Edited by S. B. Prusiner & W. J. Hadlow: New York : Academic Press.

Hunter, N., Goldmann, W., Foster, J. D., Cairns, D. & Smith, G. (1997). Natural scrapie and PrP genotype: case-control studies in British sheep. *Vet Rec* **141**, 137-140.

Hunter, N., Foster, J. D., Goldmann, W., Stear, M. J., Hope, J. & Bostock, C. (1996). Natural scrapie in a closed flock of Cheviot sheep occurs only in specific PrP genotypes. *Arch Virol* **141**, 809-824.

Jacquemot, C., Cuche, C., Dormont, D. & Lazarini, F. (2005). High incidence of scrapie induced by repeated injections of subinfectious prion doses. *J Virol* **79**, 8904-8908.

Jeffrey, M., Gonzalez, L., Chong, A., Foster, J., Goldmann, W., Hunter, N. & Martin, S. (2006). Ovine Infection with the Agents of Scrapie (CH1641 Isolate) and Bovine Spongiform Encephalopathy: Immunochemical Similarities can be Resolved by Immunohistochemistry. *J Comp Pathol* **134**, 17-29.

Kimberlin, R. H. & Wilesmith, J. W. (1994). Bovine spongiform encephalopathy. Epidemiology, low dose exposure and risks. *Ann N Y Acad Sci* **724**, 210-220.

Kuk, A. Y. C. & Chen, C.-H. (1992). A Mixture Model Combining Logistic Regression with Proportional Hazards Regression. *Biometrika* **79**, 531-541.

Lam, K. F., Fong, D. Y. & Tang, O. Y. (2005). Estimating the proportion of cured patients in a censored sample. *Stat Med* **24**, 1865-1879.

Le Dur, A., Beringue, V., Andreoletti, O., Reine, F., Lai, T. L., Baron, T., Bratberg, B., Vilotte, J. L., Sarradin, P., Benestad, S. L. & Laude, H. (2005). A newly identified type of scrapie agent can naturally infect sheep with resistant PrP genotypes. *Proc Natl Acad Sci U S A* **102**, 16031-16036.

Maller, R. A. & Zhou, X. (1996). *Survival analysis with long-term survivors*. New York: Wiley.

Matthews, L., Coen, P. G., Foster, J. D., Hunter, N. & Woolhouse, M. E. (2001). Population dynamics of a scrapie outbreak. *Arch Virol* **146**, 1173-1186.

McKinley, M. P., Bolton, D. C. & Prusiner, S. B. (1983). A protease-resistant protein is a structural component of the scrapie prion. *Cell* **35**, 57-62.

O'Doherty, E., Healy, A., Aherne, M., Hanrahan, J. P., Weavers, E., Doherty, M., Roche, J. F., Gunn, M. & Sweeney, T. (2002). Prion protein (PrP) gene polymorphisms associated with natural scrapie cases and their flock-mates in Ireland. *Res Vet Sci* **73**, 243-250.

Onnasch, H., Gunn, H. M., Bradshaw, B. J., Benestad, S. L. & Bassett, H. F. (2004). Two Irish cases of scrapie resembling Nor98. *Vet Rec* **155**, 636-637.

Orge, L., Galo, A., Machado, C., Lima, C., Ochoa, C., Silva, J., Ramos, M. & Simas, J. P. (2004). Identification of putative atypical scrapie in sheep in Portugal. *J Gen Virol* **85**, 3487-3491.

Palhiere, I., François, D., Elsen, J. M., Barillet, F., Amigues, Y., Perrret, G. & Bouix, J. (2002). Alleles frequencies of the PrP gene in 29 french sheep breeds. Possible use in selection for resistance to scrapie. In *7th WCGALP, session 13*, pp. 13-16. Montpellier, France.

Peng, Y. & Dear, K. B. (2000). A nonparametric mixture model for cure rate estimation. *Biometrics* **56**, 237-243.

Peng, Y. & Carriere, K. C. (2002). An empirical Comparison of Parametric and semiparametric cure models. *Biom J* **8**, 1002-1014.

Race, R., Raines, A., Raymond, G. J., Caughey, B. & Chesebro, B. (2001). Long-term subclinical carrier state precedes scrapie replication and adaptation in a resistant species: analogies to bovine spongiform encephalopathy and variant Creutzfeldt-Jakob disease in humans. *J Virol* **75**, 10106-10112.

Redman, C. A., Coen, P. G., Matthews, L., Lewis, R. M., Dingwall, W. S., Foster, J. D., Chase-Topping, M. E., Hunter, N. & Woolhouse, M. E. (2002). Comparative epidemiology of scrapie outbreaks in individual sheep flocks. *Epidemiol Infect* **128**, 513-521.

Ryder, S., Dexter, G., Bellworthy, S. & Tongue, S. (2004). Demonstration of lateral transmission of scrapie between sheep kept under natural conditions using lymphoid tissue biopsy. *Res Vet Sci* **76**, 211-217.

Ryder, S. J., Spencer, Y. I., Bellerby, P. J. & March, S. A. (2001). Immunohistochemical detection of PrP in the medulla oblongata of sheep: the spectrum of staining in normal and scrapie-affected sheep. *Vet Rec* **148**, 7-13.

Thorgeirsdottir, S., Sigurdarson, S., Thorisson, H. M., Georgsson, G. & Palsdottir, A. (1999). PrP gene polymorphism and natural scrapie in Icelandic sheep. *J Gen Virol* **80 (Pt 9)**, 2527-2534.

Thorgeirsdottir, S., Georgsson, G., Reynisson, E., Sigurdarson, S. & Palsdottir, A. (2002). Search for healthy carriers of scrapie: an assessment of subclinical infection of sheep in an Icelandic scrapie flock by three diagnostic methods and correlation with PrP genotypes. *Arch Virol* **147**, 709-722.

Tongue, S. C., Pfeiffer, D. U., Warner, R., Elliott, H. & Del Rio Vilas, V. (2006). Estimation of the relative risk of developing clinical scrapie: the role of prion protein (PrP) genotype and selection bias. *Vet Rec* **158**, 43-50.

Touzeau, S., Chase-Topping, M. E., Matthews, L., Lajous, D., Eychenne, F., Hunter, N., Foster, J. D., Simm, G., Elsen, J. M. & Woolhouse, M. E. (2005). Modelling the spread of scrapie in a sheep flock: evidence for increased transmission during lambing seasons. *Arch Virol*.

Tranulis, M. A., Osland, A., Bratberg, B. & Ulvund, M. J. (1999). Prion protein gene polymorphisms in sheep with natural scrapie and healthy controls in Norway. *J Gen Virol* **80 (Pt 4)**, 1073-1077.

van Keulen, L. J., Schreuder, B. E., Vromans, M. E., Langeveld, J. P. & Smits, M. A. (2000). Pathogenesis of natural scrapie in sheep. *Arch Virol Suppl*, 57-71.

Woolhouse, M. E., Stringer, S. M., Matthews, L., Hunter, N. & Anderson, R. M. (1998). Epidemiology and control of scrapie within a sheep flock. *Proc R Soc Lond B Biol Sci* **265**, 1205-1210.

Yu, B., Tiwari, R. C., Cronin, K. A. & Feuer, E. J. (2004). Cure fraction estimation from the mixture cure models for grouped survival data. *Stat Med* **23**, 1733-1747.

Table 1: size and within flock scrapie history for the 6 selected flocks.

	Flock					
	A	B	C	D	E	F
Flock size (adult ewes) †	321	331	302	463	396	380
Estimated start of scrapie outbreak	1992	1992	1993	1995	1996	1996
Number of scrapie clinical cases in ewes born before the studied cohorts of birth	135	70	96	150	90	190

† flock size at lambing of the studied cohorts of birth.

Table 2: PrP genotype distribution and number of clinical scrapie cases in the studied birth cohorts from 6 (A to F) naturally scrapie affected flocks.

Group	PrP genotype	A 1998	B 1999	C 1995	C 1996	C 1997	D 1999	E 1999	F 1999	Total	Number of clinical cases
S/S	ARQ/ARQ	43	42	46	43	25	45	53	34	331	131
	AHQ/ARQ	2	-	3	3	-	3	1	-	12	1
	ARQ/VRQ	8	5	3	-	6	6	10	9	47	28
VRQ/x	AHQ/VRQ	-	-	1			-	-	-	1	0
	VRQ/VRQ	-	-	-	1	-	-	-	-	1	1
	ARR/ARQ	28	24	18	15	29	33	34	25	206	9
R/S	ARR/AHQ	-	-	-	-	3	-	-	-	3	0
	ARR/VRQ	1	-	-	-	-	1	1	-	3	0
R/R	ARR/ARR	7	-	-	4	7	7	10	2	37	0
	Total	89	71	71	66	70	95	109	70	641	170

25

Table 3: Clinical scrapie cases (numbers and percentages) per genotype group and median age at death (clinical scrapie) in the studied birth cohorts.

Flock	PrP genotype group			
	R/R	R/S	S/S	VRQ/x
A	7 7.90 %	29 32.58%	45 50.56%	8 8.99%
B	0 0.00%	24 33.80%	42 59.15%	5 7.04%
C	11 5.31%	65 31.40%	120 57.97%	11 5.31%
D	7 7.37%	34 35.79%	48 50.53%	6 6.32%
E	10 9.17%	35 32.11%	54 49.54%	10 9.17%
F	2 2.86%	25 35.71%	34 48.57%	9 12.86%
Total	37 5.77%	212 33.07%	343 54.51%	49 7.64%
Clinical cases	0	9 4.25%	132 38.48%	29 59.18%
Median age at death (years) [min – max]	-	3.61 [3.26 – 4.13]	1.62 [0.77 – 5.54]	2.05 [1.14 -4.76]

Table 4: Monitoring duration (years) in apparently healthy sheep from the studied birth cohorts.

PrP genotype Group	N	monitoring duration (years)			
		25^e centile	Median	75^e centile	Maximum
R/S	203	3.58	5.59	7.09	9.09
S/S	211	3.47	4.69	5.64	8.04
VRQ/x	20	2.51	3.90	4.43	8.07
Total	434	3.50	4.90	6.36	9.09

Table 5: Number and percentage of infected animals by PrP genotype group according to clinical cases and laboratory tests findings in clinically healthy animals and the logistic Cox PH mixture cure model in the studied birth cohorts.

PRP genotype group	Clinical cases	Clinically healthy sheep Positive / tested (total)	Estimated number (%) of infected sheep	
			Clinical + subclinical cases (95 % CI)	logistic-CoxPH mixture cure model (point estimate)
R/R	0	0/5 (37)	-	-
R/S	9	0/49 (203)	9 – 20 (4.24 – 9.43)	12.65 (5.97)
S/S	132	4/152 (211)	136 – 142 (39.35 – 41.40)	146.83 (42.81)
VRQ/x	29	1/14 (20)	30 – 33 (61.22 – 67.34)	31.22 (63.71)
Total	170	5/220 (471)	175 - 195 (28.97 – 32.28)	190.70 (31.58)

† the lower bound is the observed proportion of scrapie infected sheep based on clinical survey and laboratory tests, the upper bound is derived from the hypergeometric law to account for clinically healthy sheep without laboratory test (n = 251). The large confidence interval for R/S animals was due to the low proportion of sampled animals.

Table 6: Effects of PRP genotype and flock origin on incidence and incubation duration in the studied birth cohorts from 6 scrapie infected flocks according to the logistic-Cox PH mixture cure model.

Variable	Incidence			Incubation duration		
	OR	CI$_{95\%}$*	p†	HR	CI$_{95\%}$*	p†
Genotype *vs. S/S*						
R/S	0.08	0.04 – 0.14	<0.0001	0.28	0.14 – 0.56	0.0003
VRQ/x	2.71	1.41 – 5.21	0.003	0.84	0.55 – 1.28	0.4117
Flock *vs. flock C* (ARQ/ARQ animals only)						
A	0.35	0.17 – 0.72	0.0045	1.75	0.95 – 3.22	0.0723
B	0.72	0.36 – 1.46	0.3695	2.49	1.46 – 4.25	0.0008
D	0.40	0.20 – 0.81	0.0113	1.92	1.08 – 3.43	0.0264
E	0.20	0.09 – 0.43	<0.0001	1.33	0.67 – 2.61	0.4144
F	1.74	0.76 – 3.89	0.1796	0.41	0.23 – 0.72	0.0021

OR = adjusted odds ratio; HR= adjusted hazard risk; *CI$_{95\%}$ = 95 % confidence interval for OR and HR estimated by the bias corrected accelerated bootstrap method over 5000 re-samples. † Level of significance of the variable (p value). The S/S genotype group and flock C were used as baseline. The influence of flock could not be investigated in R/S and VRQ allele carriers groups, due to few clinical cases in some flocks.

Captions to figures

Figure 1: (A) simulation design for incubation lengths. Survival distribution functions for simulation design 1, (black dotted lines, median age at onset 3.3 and 4.45 years) and simulation design 2 (grey dotted lines, median age at onset 2 and 3 years) ; survival distribution function for scrapie clinical cases in the studied population (solid black line). **(B and C)** Absolute biases of the estimates for each monitoring scenario (B, simulation design 1 and C simulation design 2) using the mixture cure model: white circle (β_0), proportion of infected individuals; white square (β_1), effect of covariate Z_1 on the incidence; white triangle (β_2), effect the Z_2 on the incidence; black square (γ_1), effect of covariate Z_1 on latency, black triangle (γ_2), effect of covariate Z_2 on latency.

Figure 2: Kaplan-Meier survival distribution functions for deaths from scrapie in the studied birth cohorts from 6 flocks of Manech Red Faced sheep. Overall sample (solid line, n=604); R/S genotype group (dotted line, n=212); S/S genotype group (medium dashed line, n=343); VRQ allele carrier group (long dashed line, n=49). Survival times are measured from date of birth. Censored events (i.e. culling or deaths from causes other than scrapie) are shown as circles. The last observed clinical case (5.54 years old) is indicated by the black arrow.