# Using genotype probabilities in survival analysis: a scrapie case 

Zulma Vitezica, Jean Michel J. M. Elsen, Rachel Rupp, Clara Díaz

## To cite this version:

Zulma Vitezica, Jean Michel J. M. Elsen, Rachel Rupp, Clara Díaz. Using genotype probabilities in survival analysis: a scrapie case. Genetics Selection Evolution, 2005, 37 (4), pp.403-415. 10.1051/gse:2005008 . hal-02677956

## HAL Id: hal-02677956 <br> https://hal.inrae.fr/hal-02677956

Submitted on 31 May 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Copyright

# Using genotype probabilities in survival analysis: a scrapie case 

Zulma G. Vitezic a*, Jean-Michel ElSEN ${ }^{\text {a }}$, Rachel RUPP ${ }^{\text {a }}$, Clara Díaz ${ }^{\text {b }}$<br>${ }^{\text {a }}$ Station d'amélioration génétique des animaux, Institut national de la recherche agronomique, 31326 Castanet Tolosan, France<br>${ }^{\mathrm{b}}$ Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria, Dpto. Mejora Genética Animal, Ctra. de la Coruña km 7.5, 28040 Madrid, Spain

(Received 5 August 2004; accepted 11 January 2005)


#### Abstract

The objective was to evaluate the potential use of genotype probabilities to handle records of non-genotyped animals in the context of survival analysis. To do so, the risks associated with the PrP genotype and other transmission factors in relation to clinical scrapie were estimated. Data from 4049 Romanov sheep affected by natural scrapie were analyzed using survival analysis techniques. The original data set included 1310 animals with missing genotypes; five of those had uncensored records. Different missing genotype-information patterns were simulated for uncensored and censored records. Three strategies differing in the way genotype information was handled were tested. Firstly, records with unknown genotypes were discarded (P1); secondly, those records were grouped in an unknown class (P2). Finally the probabilities of genotypes were assigned (P3). Whatever the strategy, the ranking of relative risks for the most susceptible genotypes (VRQ-VRQ, ARQ-VRQ and ARQ-ARQ) was similar even when the non-genotyped animals were not a negligible part of uncensored records. However, P3 had a more efficient way of handling missing genotype information. As compared to P1, either P2 or P3 avoided discarding the records of non-genotyped animals; however, P3 eliminated the unknown class and the risk associated with this group. Genotype probabilities were shown to be a useful technique to handle records of individuals with unknown genotype.


genotype probabilities / survival analysis / PrP genotypes / scrapie

## 1. INTRODUCTION

Animal health is a concern in any production system. Animal diseases have an economic impact because they may affect the level of production, shorten the length of productive life of animals and be the cause of discarding the animal products. Nowadays, there is an increasing interest in improving genetic

[^0]resistance to diseases. In many cases, this interest is mainly due to the existence of links between the animal's diseases and human health.

Evidence of genes influencing disease resistance exists [2]. Scrapie is one of several diseases known as Transmissible Spongiforme Encephalopathies (TSE). TSE may affect animals and humans. Resistance/susceptibility of sheep to scrapie is largely under the control of the PrP gene. In sheep, several point mutations at codons 136 (T, A or V), 154 (R or H) and 171 (R, Q, H, K) have been associated with resistance-susceptibility to scrapie. In general, the VRQ allele is related to scrapie susceptibility and ARR confers resistance, with an intermediate situation for ARQ, AHQ and ARH. Whilst the PrP gene largely controls susceptibility to scrapie (i.e. Hunter et al., [9]), other genes have also been detected in relation to susceptibility-resistance to scrapie [14].

Genotyping provides a powerful tool in relation to breeding for genetic resistance. However, in most commercial populations the information of genotypes is often incomplete due to a lack of genotyping a large proportion of individuals in the population. Exclusion of non-genotyped individuals from the data analysis may affect the estimates of correction factors included in the model [15]. Therefore, any tool to handle records of non-genotyped individuals is desired. Genotype probabilities have been used in the linear context to estimate the effect of major genes [11] and avoid bias in the estimate of the effect of QTL in populations where genotyping was not complete $[8,13]$. Meuwissen and Goddard [13] pointed out the usefulness of genotype probabilities to avoid discarding information and to correct for the effect of selection. However, genotype probabilities have not been largely explored in the context of survival analysis. This appears to be an appealing strategy to be used to estimate risk factors associated with the different genotypes or to identify transmission factors [3].

Survival analysis provides an adequate framework to analyze resistancesusceptibility to scrapie [1, 6]. Survival analysis has been used to analyze scrapie genetic resistance in the INRA Langlade flock $[3,6]$. These two studies handled missing genotypes differently. While Elsen et al. [6] discarded information of non-genotyped individuals, Díaz et al. [3] included this information in the analyses. Díaz et al. [3] compared two different approaches to account for the unknown information on the PrP genotype in order to estimate the risk associated with it. Firstly, an unknown class including individuals with unknown genotype was generated. Secondly, genotype probabilities were used. The results were similar under both strategies. The authors argued that the results were similar because most uncensored records had genotype information. However, this is not the general situation and, even in the scrapie case where

PrP genotyping is largely done, the information on genotypes is often incomplete. The objective of this work was to evaluate the potential use of genotype probabilities to handle records of non-genotyped individuals in the estimation of risk associated with PrP genotypes and other transmission factors.

## 2. MATERIALS AND METHODS

### 2.1. Data

Data were obtained from an INRA experimental flock ("Langlade") located near Toulouse (France) and first described by Elsen et al. [6]. The present study includes data from animals of the Romanov breed. In August 1993, the flock suffered an outbreak of scrapie and the natural selection against the susceptible genotypes drastically changed the distribution of the different genotypes [6]. Since then, appropriate matings of PrP-genotyped animals (i.e. between susceptible animals) were carried out in order to maintain a naturally scrapie infected flock. The flock was genetically closed between 1979 and 1996. In 1997 and 2001 several animals were brought into the flock from another INRA experimental farm. PrP genotyping started at the beginning of the scrapie outbreak in 1993 and has been systematically done since 1994.

The final dataset for survival analysis consisted of 4049 animals that were living in the Langlade flock between the 1st of April 1993 and the 4th of March 2002. Among those, 447 were uncensored, i.e. died of scrapie during the period analyzed. The animals were classified as 'died of scrapie' when they showed clinical signs confirmed by positive histology. For the survival analysis, a record was considered as censored when the animal either died of other causes or did not die before March 2002. There were 1310 animals with unknown genotype. Among them, 5 died of scrapie (and were therefore uncensored). The probabilities of genotypes were computed taking into account the pedigree information available, which included records registered from 1983. The method for calculating genotype probabilities is described below.

### 2.2. Estimation of PrP probabilities

Four alleles of the PrP gene were found in the Langlade flock: ARR, ARQ, AHQ and VRQ with ten resulting PrP genotypes. For genotyped individuals, the genotype probability was equal to unity for the observed $\operatorname{PrP}$ genotype and zero for all other possible genotypes. The genotypes of individuals that had not been typed but whose parents were homozygous were reconstructed and
considered as known. For non-genotyped individuals, these probabilities were estimated using an iterative peeling approach as in Janss et al. [10]. The peeling equations were as in Fernando et al. [7]. Genotype probabilities were computed using the pedigree data and available genotype information. To avoid over/underflows, a log form of Fernando et al.'s equations was implemented. The iterative process was repeated until the absolute difference of the estimated genotype probability between two iterations was less than $10^{-3}$ for all individuals.

### 2.3. Strategies to be compared and simulation scenarios

Three approaches relative to the treatment of missing genotype information were compared. Firstly, the non-genotyped animals were excluded from the data analysis (P1). Secondly, a specific 'unknown' group of animals was created to include non-genotyped animals (P2). Thirdly, estimates of genotype probabilities were included in the analysis (P3). The P2 and P3 were defined to avoid discarding information available to estimate transmission factors in the model.

Based on previous results [3], two situations concerning the genotyped animals of uncensored ( U ) and censored ( C ) data were simulated. For each situation ( U and C ), the genotypes of animals were discarded in different proportions from the data file. As a result, several datasets with different proportions of uncensored and censored records with unknown genotypes were generated. For uncensored data, two scenarios were considered: $25 \%$ (U-25) and $50 \%$ (U-50) of scrapie animals were assumed to have an unknown genotype. Correspondingly, in censored data, 25\% (C-25) and 50\% (C-50) of the animal genotypes were assumed unknown. In each scenario, the simulated loss of genotyping was randomly performed taking into account the genotype of the animals, in a stratified manner such as the loss of information involved all the classes of genotypes. Therefore, the number of animals with known genotypes for each class decreased relative to the original data set. The initial scenario corresponded to the original situation with only five unknown-genotype individuals in uncensored data (O). In Table I, the distribution of genotypes for censored and uncensored records are shown under each scenario.

### 2.4. Models

Survival analysis techniques were used. Failure time ( $t$ ) was expressed as the age of the animal at the scrapie diagnosis. This analysis models the hazard of

Table I. Distribution of genotypes under each scenario (25\% and 50\% of genotypes were missing) and group of data: uncensored (U) or censored (C).

| Genotypes | Original |  | $\mathbf{2 5 \%}$ |  | $\mathbf{5 0 \%}$ |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | U | C | U | C | U | C |
| Unknown | 5 | 1305 | 116 | 1879 | 227 | 2453 |
| AHQ-AHQ | 1 | 36 | 1 | 25 | 1 | 18 |
| AHQ-VRQ | 2 | 258 | 1 | 195 | 0 | 132 |
| ARQ-AHQ | 6 | 118 | 3 | 91 | 2 | 60 |
| ARQ-ARQ | 86 | 186 | 61 | 145 | 39 | 87 |
| ARQ-VRQ | 227 | 320 | 170 | 249 | 116 | 162 |
| ARR-AHQ | 0 | 96 | 0 | 75 | 0 | 47 |
| ARR-ARQ | 2 | 174 | 2 | 125 | 1 | 85 |
| ARR-ARR | 0 | 333 | 0 | 235 | 0 | 168 |
| ARR-VRQ | 6 | 451 | 3 | 334 | 2 | 231 |
| VRQ-VRQ | 112 | 325 | 90 | 249 | 59 | 159 |

an animal to be affected by scrapie at time $t$ provided that it did not show signs of scrapie till that moment. Therefore, it describes the rate at which animals are showing signs of scrapie over time. A Cox proportional hazards model was considered in the analyses. Under this model, the hazard of an animal to die of scrapie at time $t$ is written as the product of a specified baseline $\lambda_{0}$ and a set of explanatory variables or stress factors $e^{x^{\prime} \beta}$ modifying the baseline.

Two models, similar to those used in Díaz et al. [3], were run. Both models had a common part, including significant transmission factors. The flock effect $F_{j}$ is the effect of the experimental group in which the animal $i$ th is included. This effect was divided into five groups ( $j=1,5$ ); animals coming from outside Langlade ( $n=33$ ), animals from the main flock ( $n=3820$ ), animals experimentally infected with Teladorsagia Circumcinta ( $n=100$ ), animals involved in grazing experimentally infected pasture ( $n=15$ ) and animals in a special protocol where animals are left to older ages $(n=81)$. $I_{k}$ is a time dependent effect which is the combination of individual age and the level of infection challenge assuming that changes occur at the beginning of each lambing season $(\pi)$. Sheep were classified into three groups of age: $0-24(n=3807),>24-36(n=99),>36$ months $(n=143) . S x_{l}$ is the effect of sex, males $(n=1785)$ and females $(n=2264)$. The effect $R_{o}(o=1,4)$ is the combination between rearing type (maternal rearing or artificial rearing) and dam's disease status (scrapie or non-scrapie dams).

Two models were investigated that differed in the way that $\operatorname{PrP}$ genotype information was handled. Firstly, survival analysis was performed as

$$
\lambda_{i}(t)=\lambda_{0}(t) \exp \left\{F_{j}+I_{k}(\pi)+S x_{l}+\operatorname{Pr} P_{m}+R_{o}\right\}
$$

where $\lambda_{i}(t)$ represents the hazard of an animal to become a scrapie-affected animal at age $t, \lambda_{0}$ is the baseline hazard function representing the average risk of the population and $\operatorname{Pr} P_{m}$ is the effect of the animal's genotype (known or unknown). Under P1, this model included ten classes for the effect of the PrP genotype of the animals. An additional class was included to account for non-genotyped animals under P2.

In the second model, the hazard was modeled similarly except for the genotype effect. To take into account genotype probabilities for non-genotyped animals, each individual was assigned a corresponding vector of probabilities:

$$
\lambda_{i}(t)=\lambda_{0}(t) \exp \left\{F_{j}+I_{k}(\pi)+S x_{l}+\sum_{m=1}^{10} b_{m} x_{i m}+R_{o}\right\} .
$$

Probabilities $x_{i m}$ of the $m$ possible genotypes were estimated for each animal $i$, with an effect $b_{m}$ which is the regression coefficient that represents the effect of the $m$ th genotype on the hazard. This model was used in P3.

The analyses were performed using the software package Survival Kit V3. 12 [5].

## 3. RESULTS

### 3.1. Original scenario

Relative risks associated with the PrP-genotype effect for P1, P2 and P3 are presented in Table II. Under P3 the relative risk of the $m$ th genotype was calculated from the estimate of $b_{m}$, as $\exp \left(b_{m} x_{m}\right)$ for $x_{m}$ equal to one. The ranking of VRQ-VRQ, ARQ-VRQ and ARQ-ARQ genotypes was similar for the three strategies. Whatever the strategy, the VRQ-VRQ genotype had a risk about three times higher than the heterozygote ARQ-VRQ risk, and about six times higher than the homozygote ARQ-ARQ risk. Identical ranking for the most susceptible genotypes is described in Elsen et al. [6], for this population.

Figure 1 represents the fraction of individuals still alive $t$ days after birth for each group of genotypes under the P1 (a), P2 (b) and P3 (c) strategies, respectively. The survival function of an individual with a specific genotype was estimated from the cumulative baseline hazard function [12]. Figure 1 provides an illustration of the average differences in age at scrapie signs for each

Table II. Relative risk of genotypes to ARQ-VRQ genotype with different approaches to treat missing genotype information: non-genotyped individuals were eliminated (P1), non-genotyped individuals formed the unknown group (P2), the probabilities of genotypes were included (P3).

| Genotypes | P1 | P2 | P3 |
| :---: | :---: | :---: | :---: |
| Unknown |  | 0.179 |  |
| AHQ-AHQ | 0.033 | 0.034 | 0.034 |
| AHQ-VRQ | 0.020 | 0.020 | 0.011 |
| ARQ-AHQ | 0.028 | 0.029 | 0.028 |
| ARQ-ARQ | 0.521 | 0.523 | 0.528 |
| ARQ-VRQ | 1 | 1 | 1 |
| ARR-AHQ | 0 | 0 | 0.002 |
| ARR-ARQ | 0.014 | 0.016 | 0.015 |
| ARR-ARR | 0 | 0 | 0 |
| ARR-VRQ | 0.030 | 0.031 | 0.030 |
| VRQ-VRQ | 2.874 | 2.876 | 2.936 |

genotype. The estimates of the survival rate for each genotype were different, being lower under the P1 and P2 strategies. Under P1 and P2, $90 \%$ of VRQ-VRQ animals showed scrapie signs before 900 days of age while only $2 \%$ for ARR-ARQ showed signs before that age. However, under P3, only $37 \%$ of VRQ-VRQ animals showed scrapie signs 900 days while $0.2 \%$ was obtained for ARR-VRQ.

### 3.2. Simulated scenarios

The effect of discarding non-genotyped individuals on the relative risks associated with the genotypes was studied. The results from the simulation were compared to the original situation (O).

Under P2, the results for P2-U-25 and P2-U-50 showed that the risk of the unknown group was not negligible and increased with the number of individuals with uncensored records included in this group (Tab. III). The unknown class changed from the fourth position in P2-O to the second position in P2-U-25 and P2-U-50. Thus, the risk associated with the unknown group was intermediate between $V R Q-V R Q$ and $A R Q-V R Q$. Ignoring the unknown group, the ranking of VRQ-VRQ, ARQ-VRQ and ARQ-ARQ genotypes listed in decreasing value of risk was similar among P2-O, P2-U-25 and P2-U-50. Identical results were found for P1.


Figure 1. Survivor curves of different genotypes when non-genotyped individuals were eliminated (a), non-genotyped individuals formed the unknown group (b), and the probabilities of genotypes were included (c). The genotypes are listed and drawn in decreasing survival rate, ignoring the unknown group.

Under P3, the relative risk effects associated with the PrP genotypes were computed for the different scenarios: P3-O, P3-U-25 and P3-U-50 (Tab. IV). Similar patterns were observed for P3-O and the two simulated scenarios. The highest risk groups consisted of animals carrying VRQ-VRQ, ARQ-VRQ

Table III. Relative risk of genotypes to the ARQ-VRQ genotype under P2 with the original situation (P2-O) and when 25\% (P2-U-25) and 50\% (P2-U-50) of genotyping were missing in uncensored data.

| Genotypes | P2-O | P2-U-25 | P2-U-50 |
| :---: | :---: | :---: | :---: |
| Unknown | 0.179 | 1.674 | 2.547 |
| AHQ-AHQ | 0.034 | 0.048 | 0.066 |
| AHQ-VRQ | 0.020 | 0.020 | 0 |
| ARQ-AHQ | 0.029 | 0.021 | 0.020 |
| ARQ-ARQ | 0.523 | 0.545 | 0.530 |
| ARQ-VRQ | 1 | 1 | 1 |
| ARR-AHQ | 0 | 0 | 0 |
| ARR-ARQ | 0.016 | 0.016 | 0.016 |
| ARR-ARR | 0 | 0 | 0 |
| ARR-VRQ | 0.031 | 0.019 | 0.017 |
| VRQ-VRQ | 2.876 | 2.992 | 3.375 |

Table IV. Relative risk of genotypes to the ARQ-VRQ genotype under P3 with the original situation (P3-O) and when 25\% (P3-U-25) and 50\% (P3-U-50) of genotyping were missing in uncensored data.

| Genotypes | P3-O | P3-U-25 | P3-U-50 |
| :---: | :---: | :---: | :---: |
| AHQ-AHQ | 0.034 | 0.047 | 0.057 |
| AHQ-VRQ | 0.011 | 0.031 | 0.063 |
| ARQ-AHQ | 0.028 | 0.048 | 0.067 |
| ARQ-ARQ | 0.528 | 0.549 | 0.534 |
| ARQ-VRQ | 1 | 1 | 1 |
| ARR-AHQ | 0.002 | 0.011 | 0.019 |
| ARR-ARQ | 0.015 | 0.060 | 0.110 |
| ARR-ARR | 0 | 0.011 | 0.027 |
| ARR-VRQ | 0.030 | 0.066 | 0.011 |
| VRQ-VRQ | 2.936 | 2.781 | 2.874 |

and ARQ-ARQ genotypes. Estimated probabilities tended to keep the ranking of these genotypes in terms of risk to the original order even when the non-genotyped animals were not a negligible part of uncensored records. Using genotype probabilities, the genotypes such as ARR-ARR and ARR-AHQ, that previously did not appear at risk, showed a small relative risk (Tabs. III and IV).

Different missing scenarios in censored data were simulated and also compared to the original situation. Under P1, P2 and P3, when the number of nongenotyped individuals with censored data increased, the risk ranking among the most susceptible genotypes: VRQ-VRQ, ARQ-VRQ, and ARQ-ARQ, was similar to the original scenario. Under P2, with the increase of the missing genotype information, the risk for an unknown group became almost zero. The risk changed from 0.179 with 1305 non-genotyped censored records in the original situation to 0.005 when $50 \%$ of genotypes were missing in censored data (2453 records).

The effect of discarding non-genotyped animals on confidence intervals of the relative risks associated with transmission factors was studied. Confidence intervals for the relative risk were found by exponentiating the lower and upper limits [12]. P1, P2 and P3 were compared for the $R_{0}$ effect (a combination between the rearing type and the dam's disease status) when $50 \%$ of the genotypes were missing in uncensored data. As it was expected, P1 showed an impact on confidence intervals of each level of the $R_{0}$ factor. The effect was more important when the amount of discarded information was the largest. Going from P1-O to P1-U-50 resulted in an increase in the confidence interval for the relative risk of the $R_{0}$ effect. For example, the $95 \%$ confidence intervals for $R_{1}$ were $(0.974,1.679)$ and $(0.882,1.906)$ under P1-O to P1-U-50, respectively. However, the corresponding $95 \%$ confidence intervals for $R_{1}$ were ( $0.961,1.655$ ) and $(0.997,1.699)$ under P3-O to P3-U-50, respectively. Using P3, the estimation of other transmission factors is not affected because the non-genotyped individuals are taken into account in the analysis. The P2 strategy provided similar results to P3.

## 4. DISCUSSION

Three approaches to account for the non-genotyped animals to estimate risk associated to genotypes and other transmission factors have been illustrated, in the context of survival analysis. Three strategies to handle missing genotype information (P1, P2 and P3) were compared. Thus, in terms of the amount of information available to estimate the risks associated with different $\operatorname{PrP}$ genotypes, P3 seems to make a more efficient use of all the information available. However, P2 uses the same amount of information that P1 because all non-genotyped animals are included in the unknown class while for P1 they are discarded. Meuwissen and Goddard [13] pointed out the contribution of genotype probabilities to avoid discarding information to estimate the effect of QTL in the population where genotyping was not complete.

Regardless of the strategy, the ranking of relative risks for the most susceptible genotypes (VRQ-VRQ, ARQ-VRQ and ARQ-ARQ) was similar even when the non-genotyped animals were not a negligible part of uncensored records. The amount of uncensored relative to the censored together with the failure time distribution between the censored and uncensored classes will affect the estimation of relative risk [4].

An important issue of survival analysis is the estimation of other transmission factors included in the models. While the exclusion of individuals from the analysis did not seem to affect the ranking among the most susceptible genotypes, discarding non-genotyped individuals from the data showed an impact on the confidence interval of transmission factors. This result is not surprising, although it points out the importance of specific strategies oriented to maintain records of affected animals with unknown genotype in the data set, as P2 and P3 do. These strategies will provide more precise estimates of other transmission factors included in the model.

An advantage of P 3 with respect to P 2 is that it allowed us to include animals with an unknown genotype in the analysis assigning them a probability of carrying each of the tenth possible genotypes and eliminating the uncertainty of the "unknown" group. It was shown that the relative risk associated with the unknown group was not negligible and increased with the number of individuals with uncensored records included in this group. In our population, an increase in the number of non-genotyped individuals is expected to increase the number of individuals carrying susceptible genotypes assigned to the unknown group provided they are the most frequent genotypes among the Romanov, at Langlade. In general, the risk associated with the unknown class would depend on the incidence of scrapie among non-genotyped individuals and the underlying distribution of susceptible genotypes. However, P3 assigned a small risk with genotypes without any clinical sign of scrapie. This effect will also depend upon the structure of the data. In our data, risk to show clinical signs has been found for heterozygotes of ARR with a chance of having ARR/ARR progeny. A way to avoid these effects would also be to condition on the individual's phenotype to assign probabilities.

The effect of P3 on the estimate of the survivor curves has also been pointed out. P3 has been proven to have an impact on the estimates of survival rates for each genotype. The estimates of the survival rate for each genotype were higher under the P3 strategy. The overestimation of P3 on the survivorship was particularly noticeable for ARQ-ARQ, ARQ-VRQ and VRQ-VRQ. The inclusion of probabilities to account for non-genotyped animals (instead of discarding or grouping them) increased the number of animals still alive
"carrying" susceptible genotypes. Thus, the assignment of probabilities affected the estimate of average risk of the population and an other result is that it increased the overall survival rate. Nevertheless, the effect of P3 will very much depend on the data structure. Our result is the consequence of having a large amount of censored records with a censoring time going further than the failure time of scrapie animals and with a large probability of carrying susceptible genotypes. However, relative risks and ranking among genotypes stayed unchanged for all strategies. This may be so because under a semi-parametric approach the estimation of the stress factors and correspondingly the risks do not depend on the baseline [12].

## 5. CONCLUSION

In order to deal with the risk analysis, if part of the genotypes is missing, P 2 and P 3 are preferable to P1. But both P2 and P3 have advantages and disadvantages. The relative benefits of the two approaches depend upon the circumstances and the question asked. Both P2 and P3 avoided discarding the records of non-genotyped animals with respect to P1. However, genotype probabilities eliminated the unknown class and the risk associated with this group with respect to P2. Nevertheless the contribution of genotype probabilities to estimate the risk associated with genotypes, will depend upon the structure of the information available to estimate probabilities.

## ACKNOWLEDGEMENTS

The authors wish to thank Dr. Vincent Ducrocq, for helpful comments. Zulma Vitezica was funded by Ministerio de Educación, Ciencia y Técnica de la República Argentina and INRA. Clara Díaz was Collaborator, via a fellowship under OECD Co-operative Research Programme: Biological Resource Management for Sustainable Agriculture Systems.

## REFERENCES

[1] Baylis M., Goldman W., Houston F., Cairns D., Chong A., Ross A., Smith A., Hunter N., McLean A.R., Scrapie endemic in fully PrP-genotyped sheep flock, J. Gen. Virol. 83 (2002) 2907-2914.
[2] Bishop S., Breeding for Disease Resistance: uniting genetics and epidemiology, course notes, 2002.
[3] Díaz C., Vitezica Z.G., Rupp R., Elsen J.M., El uso de la predicción de genotipos en el análisis de supervivencia: un caso de scrapie, ITEA 24 (II) (2003) 456-458.
[4] Ducrocq V., Survival Analysis Applied to Epidemiology, course notes, 2002.
[5] Ducrocq V., Sölkner J., The Survival Kit. A Fortran package for the analysis of survival data, in: Proceedings of the 6th World Congress on Genetics Applied to Livestock Production, January 11-16, 1998, Vol. 27, University of New England, Armidale, pp. 447-448.
[6] Elsen J.M., Amigues Y., Schelcher F., Ducrocq V., Andreoletti O., Eychenne F., Tien Khang J.V., Poivey J.P., Lantier F., Laplanche J.L., Genetic susceptibility and transmission factors in scrapie: detailed analysis of an epidemic in a closed flock of Romanov, Arch. Virol. 144 (1999) 431-445.
[7] Fernando R.L., Stricker C., Elston R.C., An efficient algorithm to compute the posterior genotypic distribution for every member of a pedigree without loops, Theor. Appl. Genet. 87 (1993) 89-93.
[8] Hofer A., Kennedy B.W., Genetic evaluation for a quantitative trait controlled by a polygenes and a major locus with genotypes not or only partially known, Genet. Sel. Evol. 25 (1993) 537-555.
[9] Hunter N., Foster J.D., Goldman W., Stear M.J., Hope J., Bostock C., Natural scrapie in a closed flock of Cheviot sheep occurs only in specific PrP genotypes, Arch. Virol. 141 (1996) 809-824.
[10] Janss L.L.G., van Arendonk J.A.M., van der Werf J.H.J., Computing approximate monogenic model likelihoods in large pedigrees with loops, Genet. Sel. Evol. 27 (1995) 567-579.
[11] Kinghorn B.P., Kennedy B.W., Smith C., A method of screening for genes of major effect, Genetics 134 (1993) 351-360.
[12] Klein J.P., Moeschberger M.L., Survival Analysis: Techniques for Censored and Truncated data, Springer-Verlag, New York Inc., 1997.
[13] Meuwissen T.H.E., Goddard M.E., Estimation of effects of quantitative trait loci in large complex pedigrees, Genetics 146 (1997) 409-416.
[14] Moreno C.R., Lantier I., Lantier F., Andreoletti O., Vaiman D., Sarradin P., Echeynne F., Cribiu E.P., Cosseddu G., Elsen J.M., Transposition to sheep of mouse quantitative trait loci (QTL) influencing susceptibility to prion diseases, in: Proceedings of the 7th World Congress on Genetics Applied to Livestock Production, August 19-23, 2002, Vol. 31, pp. 687-690.
[15] Searle S.R., Linear Models, John Wiley \& Sons, New York Inc., 1971.


[^0]:    * Corresponding author: vitezica@germinal.toulouse.inra.fr

