

EL USO DE LA PREDICCIÓN DE GENOTIPOS EN EL ANÁLISIS DE SUPERVIVENCIA: UN CASO DE SCRAPIE

C. Díaz¹, Z. Vitezica², R. Rupp², J. M. Elsen².

¹ Dpto de Mejora Genética Animal. INIA. Madrid

² Station d'Amélioration Génétique des Animaux. INRA – Toulouse. France.

INTRODUCCION

El *scrapie* es una enfermedad neurodegenerativa de los ovinos, que pertenece al grupo de las Encefalopatías Espongiformes Transmisibles (EET). Hasta la fecha se sabe que en gran parte la resistencia o susceptibilidad a la enfermedad depende de mutaciones puntuales que afectan a los codones 136, 154 y 171 de la ORF del gen PrP (ej. Elsen y col. 1999).

El análisis de supervivencia ofrece un marco apropiado para el análisis de resistencia-susceptibilidad a enfermedades (Elsen y col., 1999). Sin embargo, la existencia de información faltante, en nuestro caso los genotipos de los individuos, se traduce en una pérdida de datos para la estimación de los factores de riesgo.

Hasta la fecha existen varios algoritmos para calcular las funciones de probabilidad de genotipos utilizando la información del pedigrí y el genotipo conocido en algunos individuos. Bajo el modelo monogénico la verosimilitud de un pedigree puede calcularse eficientemente utilizando el algoritmo de "peeling" (Elston y Stewart, 1971). El peeling consiste en un eficiente reordenamiento de las adiciones y multiplicaciones necesarias para calcular la verosimilitud del pedigree. Fernando y col. (1993) y Janss y col. (1995) presentaron un algoritmo iterativo (peeling iterativo) que permite calcular las probabilidades de los genotipos de los miembros de un pedigree.

El objetivo de este trabajo es comparar dos estrategias para la inclusión de la información genotípica en la estimación del riesgo.

MATERIALES Y METODOS

Datos: Los datos se han obtenido de un rebaño experimental (Langlade) del INRA donde se produjo un brote de *scrapie* en Abril de 1993. La infección se mantiene de manera natural en la finca desde entonces. El archivo de datos contiene 4049 registros de animales Romanov nacidos entre 1983 y 2002. Se puso como condición que los animales nacidos con anterioridad a 1993 estuvieran presentes en el momento del brote infeccioso. Del total de animales, cuatrocientos cuarenta y siete murieron de *scrapie*. Un animal fue asignado al grupo de *scrapie* si presentó los síntomas típicos de la enfermedad y el resultado de la histología fue positivo. Los animales que no murieron antes de marzo del 2002 o murieron por otra causa fueron censurados asignándoseles dicha fecha, como fecha de censura. Los animales han estado involucrados en distintos protocolos experimentales, por lo cual se definió un efecto de cuatro clases para tomar en cuenta su posible efecto en el riesgo.

Estimación de genotipos: De los 4049 animales en datos, 2656 tenían un genotipo conocido. En el primer análisis, los individuos con genotipo faltante fueron asignados a un grupo de desconocidos. Posteriormente, se procedió a estimar los genotipos de los individuos desconocidos utilizando el algoritmo de "peeling iterativo" (Fernando y col., 1993). Respecto de un individuo i con genotipo l_i , el algoritmo particiona el pedigree en dos partes. La porción anterior al individuo i incluye sus padres y sus hermanos completos, mientras que la posterior contiene la información proveniente de sus apareamientos y sus descendientes. La probabilidad de que i tenga genotipo l_i dada la información fenotípica del resto de los miembros del pedigrí es igual a:

$$P(l_i/\mathbf{y}) = \frac{P(\mathbf{y}_a, l_i)P(y_i/l_i)P(\mathbf{y}_p/l_i)}{L}$$

L es la verosimilitud del pedigrí, igual a $L = \prod_{i_i} P(\mathbf{y}_a, l_i)P(y_i/l_i)P(\mathbf{y}_p/l_i)$ siendo $P(\mathbf{y}_a, l_i)$ la probabilidad conjunta de los fenotipos de los miembros anteriores a i y su genotipo l_i , y $P(\mathbf{y}_p/l_i)$ la probabilidad condicional de los fenotipos de los miembros posteriores a i , dado que i tiene genotipo l_i . La probabilidad $P(y_i/l_i)$ es la penetrancia. Estas probabilidades definen las ecuaciones de peeling anterior y posterior, respectivamente. Para calcular la probabilidad de los genotipos, ambas ecuaciones fueron empleadas iterativamente, hasta que la diferencia absoluta de las probabilidades estimadas entre dos iteraciones fue menor a 0.001 para cada uno de los animales.

Modelo de análisis: el tiempo de supervivencia se expresó como la edad de los animales al diagnóstico de *scrapie*. Así

$$\lambda_i(t) = \lambda_0 \exp\{F_j + ALI_k(\pi) + Sx_m + PrP_n + rtds_o\}$$

donde, $\lambda_i(t)$ representa el riesgo de un individuo a morir de *scrapie* en el tiempo t ; λ_0 es el riesgo medio de la población; F_j es un efecto clase que representa el efecto que sobre el riesgo presentan los distintos grupos experimentales; ALI_k es el efecto de la interacción de la edad del individuo con el nivel de infección asumiendo que éste cambia al inicio de cada paridera (\square). Los animales se dividieron en tres grupos de edad, de hasta dos años, tres años y mayores de tres años; Sx_m es el sexo de cada individuo; PrP_n es el efecto del genotipo de cada animal o desconocido en su defecto y $rtds_o$ es el efecto de la interacción del tipo de cria (materna y artificial) con el estado de la enfermedad de la madre (*scrapie* o *no_scrapie*). En el segundo modelo, el riesgo se modeliza de la misma forma excepto para el genotipo. Para incluir dicha información se genera un vector de probabilidades, que representa la probabilidad que cada individuo tiene de llevar cada genotipo.

donde b representa la regresión de la probabilidad del genotipo l del individuo i sobre el riesgo del individuo a la edad t ; x_{il} es la probabilidad del individuo i de tener el genotipo l ; ng es el número de genotipos probables (diez). El Survival Kit fue utilizado para realizar los análisis (Ducrocq y Solkner, 1998)

RESULTADOS Y DISCUSION

En la Figura 1 se presentan los riesgos estimados para cada genotipo para el modelo donde se utilizan las probabilidades de genotipos y donde se asignan los genotipos desconocidos a una clase desconocida. Si se comparan las barras, se puede observar que los individuos de genotipo VRQ-VRQ presentan un riesgo unas tres veces mayor que el heterocigoto (ARQ-VRQ) y seis veces mayor que el homocigoto ARQ-ARQ. Resultados similares han sido descritos en esta misma población (Elsen y col., 1999). Dicha relación de riesgos parece ser similar entre las dos estrategias (Figura 1). Sin embargo, cuando se utiliza la asignación de los genotipos desconocidos a una misma clase, esta clase pasa a ser el cuarto grupo en riesgo, denotando que existe una heterogeneidad de individuos en dicho grupo. Al introducir la estimación de probabilidades de genotipos, los riesgos relativos atribuidos a cada genotipo estimado son de una magnitud similar a los correspondientes a los genotipos conocidos (Figura 1). En la Figura 2 se presentan los riesgos relativos al riesgo del genotipo ARQ-VRQ estimados para cada clase de

genotipos en función de la probabilidad de genotipo estimado. Asumiendo como extremos el caso $P_b=1.0$, sin incertidumbre respecto de los genotipos, y la situación con total desconocimiento ($P_b=0.0$). En dicha figura se puede observar que a partir de probabilidades del orden de 0.3, se pueden determinar diferentes riesgos para cada genotipo.

Figura 1. Riesgos relativos al genotipo ARQ-VRQ en función del genotipo y de la probabilidad de genotipo.

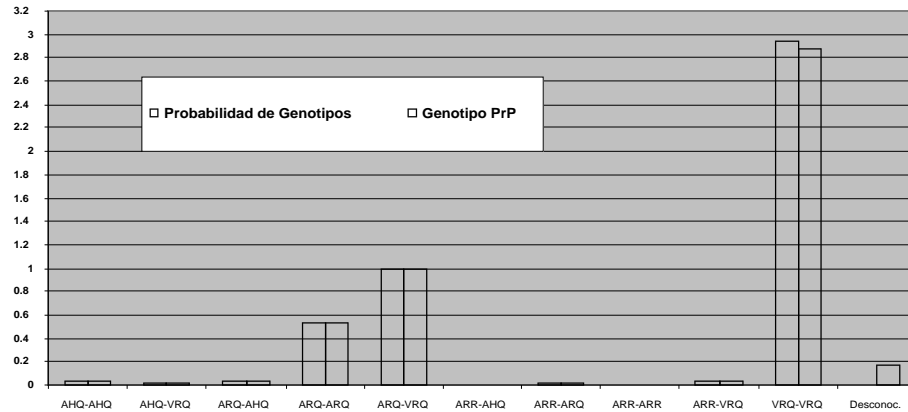
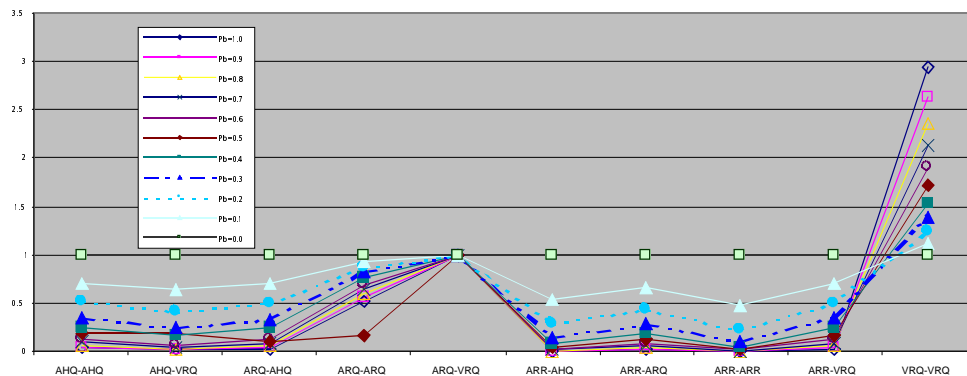


Figura 2. Riesgos relativos a genotipo ARQ-VRQ en función de la incertidumbre en la probabilidad asignada.



Estos resultados sugieren que la estimación de los genotipos podría permitir la discriminación de grupos de riesgo. Sin embargo, la escasa diferencia observada entre ambas estrategias podría atribuirse al número reducido de datos completos en el grupo de animales de genotipo desconocido.

Agradecimientos: Los autores agradecen la colaboración de Daniel Lajous y Francis Eychenne en la obtención de la información. Este trabajo se ha desarrollado parcialmente con la financiación de la OECD Co-operative Research Programme: Biological Resource Management for Sustainable Agriculture Systems.

REFERENCIAS BIBLIOGRAFICAS

- Ducrocq, V., Solkner. 1998. Proc. 6th WCGALP. 27:447-448.
 Elsen, J.M., Amigues, Y., Schelcher, F., Ducrocq, V., Andreoletti, O., Eychenne, F., TienKhang, J.V., Poivey, J.P., Lantier, F., Laplanche, J.L., 1999. Arch.Virol. 144: 431-445.
 Elston, R.C.; Stewart, J. Hum.Hered. 21 (1971) 523-542.
 Fernando, R.L., Stricker, C., Elston, R.C., Theor.Appl.Genet. 87 (1993) 89-93.
 Janss, L.L.G., Thompson, R., van Arendonk, J.A.M., Theor.Appl.Genet. 91 (1995) 1137-1147