

THE IMPACT OF FITTING INCORRECT MODELS ON THE PARTITIONING OF GENETIC VARIANCE COMPONENTS FOR BINOMIAL LAMB TRAITS

P.R. Amer K. Stachowicz, P. Pierard and G.M. Jenkins

AbacusBio Limited, Dunedin, New Zealand

SUMMARY

This paper presents a simulation study that shows that failure to fit dam permanent environmental effects in variance component estimation of lamb survival results in an upward bias in the estimate of maternal genetic variance. In contrast, fitting litter effects has little impact on variance component estimation. These results have implications for sheep genetic evaluation of lamb survival, and values for the direct and maternal heritability of lamb survival at the lower end of the range of those published based on alternative models should be used in national genetic evaluations.

INTRODUCTION

Lamb survival is lowly heritable, with many sources of variation affecting phenotype. These sources of variation include environmental effects, dam lifetime permanent environmental effects, litter effect, maternal genetic effects and direct genetic effects. Concerns have been raised that data structure may be insufficient for genetic variance component estimation software to partition maternal genetic effects from dam permanent environmental effects, meaning that some of the variance explained by maternal genetics may be incorrectly assigned to dam permanent environmental effects. This could lead to the, perhaps false, assumption that the maternal genetic effect is not significant. The opposite could also be true, and where permanent environmental effects are not accounted for, the maternal genetic effect could be artificially inflated. The aim of this study was to determine whether the data structures typical in New Zealand sheep flocks are sufficient for partitioning maternal genetic variance from environmental litter variance and dam lifetime permanent environmental variance. The distribution of the trait was also tested to determine if there was an effect on variance component estimation due to the trait being binary instead of continuous, and when binary, for dependence on incidence.

MATERIALS AND METHODS

This study was based on simulation of a population of animals born during 1995-2001. The population structure and relationship between individuals is set based on a real pedigree from a single sheep flock to reflect industry data, so as to be typical of that which would be used for variance component estimation in practice. The proportions of ewes lambing singles, twins, triplets and greater litter sizes were .15, .68, .16 and .01 respectively. There were 103 sires with at least progeny 30 progeny, and a total of 20,107 lambs born across the seven years. There was excellent connectedness across years due to several rams having large numbers of progeny across many years, and a further 18 rams with progeny across either 2 or 3 years (see Amer and Jopson, 2003, for further details on the pedigree structure).. Animals from the first generation (i.e. unknown parentage) were attributed random true breeding values, as sampled from a normal distribution with a mean of 0 and genetic variances (direct and maternal) equal to the respective heritabilities (total phenotypic variance was assumed to be 1). As lamb survival is a binary trait but phenotypes were sampled from a normal distribution, a threshold (described below) was set to determine survival on the binary scale. The lamb survival phenotypes of individual progeny of known parents identified in the existing pedigree in subsequent generations were derived using methods described below.

Five random effects were simulated as contributing to the continuous lamb survival phenotype (aphen): the direct genetic effect, maternal genetic effect, dam permanent environmental effect, litter environmental effect and a residual effect. All founders in the flock pedigree were assigned values for genetic parameters, and then the simulation iterated down through generations within the pedigree simulating genotypes and phenotypes for all animals present, based on values previously simulated for their parents. To preserve pedigree integrity, no removal of animals as parents was undertaken, even when their simulated phenotype was “dead”.

Two different binary (alive or dead) traits were derived from the continuous phenotype – one to represent high and one to represent low survival rates. To simulate a high survival binary trait, a threshold of 1.4 was defined to set survival to either dead (>1.4) or alive (≤ 1.4) – equating to approximately 8% of lambs in the population dying. A threshold of 1 was used to simulate a low survival binary trait - corresponding to approximately 15% lamb death. Expected values for heritabilities of binary traits when estimated using normal linear models was derived based on Dempster and Lerner (1950).

Six scenarios with varying levels of simulated variance components were used to simulate each of the lamb survival traits (Table 1). HIGH (Table 1) reflects a situation where there are larger variance components in the lamb survival trait. LOW (Table 1) is perhaps more reflective of the real world scenario, where direct and maternal heritabilities are low and dam permanent environmental effect is low and equivalent to direct and maternal heritabilities. LOW1-4 (Table 1) were variations of LOW in that the same amount of genetic variation was present within the trait with each scenario (exception - LOW3), but the variance was distributed among the parameters differently - LOW1 (no litter effect), LOW2 (no permanent environmental effect), LOW3 (no litter or permanent environmental effect) and LOW4 (no maternal genetic effect). Each scenario was simulated 20 times and results averaged.

For each trait (continuous, low survival, high survival) and each simulation scenario (HIGH, LOW, LOW1, LOW2, LOW3 and LOW4) five different models were fitted in ASREML (Gilmour et al, 2008) (Table 2). Although lamb survival is a binary trait, the continuous version was also analysed to determine if the efficiency of estimation and partition of variance components was influenced by the distributional properties of the trait.

Table 1 Alternative sets of variance component parameters used as simulation inputs

	Scenario					
	HIGH	LOW	LOW1	LOW2	LOW3	LOW4
h^2_{direct}	0.2	0.05	0.05	0.05	0.05	0.05
$h^2_{maternal}$	0.15	0.05	0.05	0.05	0.1	0
$damPE (c^2)$	0.1	0.05	0.1	0	0	0.1
$litter (l)^2$	0.05	0.05	0	0.1	0	0.05

RESULTS AND DISCUSSION

When dam permanent environmental effect (damPE) is present in the dataset and accounted for in the genetic analysis, the estimate of maternal heritability of lamb survival is what would be expected based on the known (simulated) values in the dataset (Table 3). That is, the data structure is sufficient for the partitioning of variance between maternal genetics and dam permanent environmental effects. Conversely, when dam permanent environmental effects are present in the data but are not accounted for in the statistical model fitted, then the maternal heritability is artificially inflated (Table 3). This is the case whether a continuous lamb survival phenotype

(0.091 when damPE is not fitted versus 0.049 when damPE is fitted), low survival phenotype (0.039 versus 0.021) or high survival phenotype (0.029 versus 0.016) is simulated. These results were consistent with results from the model when applied to the HIGH variance component set, whereby all components (direct genetics, maternal genetics, dam permanent environmental effects and litter effects) are simulated to be present within the population (Table 4). In this scenario, failure to account for dam permanent environmental effect resulted in a 35-55% increase in maternal heritability (Table 4).

Table 2 Mixed models fitted - models 1-4 are fitted for the binary and continuous traits, while the probit model (5) is only fitted for the binary traits

Model	Equations
1	cSurv/lSurv/hSurv = byr + sex + br + aod + animal + dam + dampe + litter
2	cSurv/lSurv/hSurv = byr + sex + br + aod + animal + dam + dampe
3	cSurv/lSurv/hSurv = byr + sex + br + aod + animal + dam + litter
4	cSurv/lSurv/hSurv = byr + sex + br + aod + animal + dam
5 (probit)	lSurv/hSurv (probit) = byr + sex + br + aod + animal + dam + dampe + litter

cSurv = continuous lamb phenotype, lSurv = low incidence lamb survival binary phenotype, hSurv = high incidence lamb survival binary phenotype, byr = birth year, br = birth rank, animal = animal's unique identifier (relationships included by fitting the A matrix, which describes pedigree relationships, dam = dam uid, dampe = permanent environmental effect, litter = litter effect.

Table 3 Effect of not accounting for permanent environmental effect (damPE) when it is present in the data (low survival variance assumed) – expected based on simulated values and estimated based on variance component estimation

TRAIT	h ² direct		h ² maternal		damPE	
	Expected	estimated	expected	estimated	expected	estimated
<i>Dam permanent environmental effect simulated and fitted</i>						
Continuous	0.05	0.055	0.05	0.049	0.1	0.042
Low survival	0.021	0.025	0.021	0.021	0.1	0.021
High survival	0.015	0.015	0.015	0.016	0.1	0.017
<i>Dam permanent environmental effect simulated but not fitted</i>						
Continuous	0.05	0.054	0.05	0.091	0.1	n/a
Low survival	0.021	0.025	0.021	0.039	0.1	n/a
High survival	0.015	0.016	0.015	0.029	0.1	n/a

Table 4 Effect of not accounting for either permanent environmental effect (damPE) or litter effect when they are present in the data (high survival variance assumed) – expected (exp) based on simulated values and estimated (est) based on variance component estimation

TRAIT	h ² direct		h ² maternal		damPE		litter	
	exp	est	exp	est	exp	est	exp	est
<i>All terms simulated, all terms fitted</i>								
Continuous	0.2	0.203	0.15	0.153	0.1	0.047	0.05	0.010
Low survival	0.086	0.097	0.064	0.065	0.1	0.033	0.05	0.009
High survival	0.061	0.072	0.046	0.050	0.1	0.030	0.05	0.013
<i>All terms simulated, litter not fitted</i>								
Continuous	0.2	0.203	0.15	0.153	0.1	0.050	0.05	n/a
Low survival	0.086	0.097	0.064	0.065	0.1	0.035	0.05	n/a
High survival	0.061	0.072	0.046	0.050	0.1	0.033	0.05	n/a
<i>All terms simulated, dam permanent environmental effect not fitted</i>								
Continuous	0.2	0.193	0.15	0.206	0.1	n/a	0.05	0.013
Low survival	0.086	0.094	0.064	0.099	0.1	n/a	0.05	0.013
High survival	0.061	0.071	0.046	0.078	0.1	n/a	0.05	0.016

Additional simulations (results not shown) revealed that when variance components were not simulated, they were correctly predicted as being zero. Furthermore, when a generalised linear model approach was used to account for binomially distributed data, the only difference in results was as could be predicted using the transformation described by Dempster and Lerner (1950).

Everett Hincks et al (2014) published variance component estimates for NZ sheep using a range of variance models. One of their models included dam permanent environmental effects, and this model gave the lowest estimate of the maternal genetic variance of lamb survival. We conclude that typical sheep pedigree structures in New Zealand sheep populations are sufficient for disentanglement of multiple variance components, and failure to fit an existing random effect in variance component estimation procedures can result in inflation of other variance components.

REFERENCES

- Amer, P.R. and Jopson, N.B. (2003). *Proc. NZ Soc. Anim. Prod.*, **63**: 201.
 Dempster, E.R. and Lerner, I.M. (1950). *Genetics*, **35**: 212.
 Everett Hincks, J.M., Mathias-Davis, H.C., Greer, G.J., Auvray, B.A. and Dodds, K.G. (2014). *J. Anim. Sci.*, **92(7)**: 2885.
 Gilmour, A.R., Gogel, B.J., Cullis, B.R., and Thompson, R. (2009). ASReml User Guide Release 3.0. VSN International Ltd, Hemel Hempstead, HP1 1ES, UK.