Animal (2015), **9:3**, pp 380–387 © The Animal Consortium 2014
doi:10.1017/S1751731114002717

animal

Heterogeneous variances and genetics by environment interactions in genetic evaluation of crossbred lambs

G. C. Márquez[1a], W. Haresign[2], M. H. Davies[3], R. Roehe[4], L. Bünger[4], G. Simm[4] and R. M. Lewis[1†]

[1]*Department of Animal and Poultry Sciences, Virginia Tech, Blacksburg, VA 24060, USA;* [2]*Institute of Biological, Environmental, and Rural Sciences, Aberystwyth University, Aberystwyth, SY23 3FG, UK;* [3]*ADAS Rosemaund, Preston Wynne, HR1 3PG, UK;* [4]*Animal & Veterinary Sciences Group, SRUC, Edinburgh, EH9 3JG, UK*

(Received 7 December 2013; Accepted 8 September 2014; First published online 19 November 2014)

Accounting for environmental heteroscedasticity and genetics by environment interaction (G × E) in genetic evaluation is important because animals may not perform predictably across environments. The objectives of this study were to evaluate the presence and consequences of heteroscedasticity and G × E on genetic evaluation. The population considered was crossbred lambs sired by terminal sires and reared under commercial conditions in the UK. Data on 6325 lambs sired by Charollais, Suffolk and Texel rams were obtained. The experiment was conducted between 1999 and 2002 on three farms located in England, Scotland and Wales. There were 2322, 2137 and 1866 lambs in England, Scotland and Wales, respectively. A total of 89 sires were mated to 1984 ewes of two types (Welsh and Scottish Mules). Most rams were used for two breeding seasons with some rotated among farms to create genetic links. Lambs were reared on pasture and had their parentage, birth, 5 week, 10 week, and slaughter weights recorded. Lambs were slaughtered at a constant fatness, at which they were ultrasonically scanned for fat and muscle depth. Heteroscedasticity was evaluated in two ways. First, data were separated into three subsets by farm. Within-farm variance component estimates were then compared with those derived from the complete data (Model 1). Second, the combined data were fitted, but with a heterogeneous (by farm) environmental variance structure (Model 2). To investigate G × E, a model with a random farm by sire (F × S) interaction was used (Model 3). The ratio of the F × S variance to total variance was a measure of the level of G × E in the population. Heterogeneity in environmental variability across farm was identified for all traits ($P < 0.01$). Rank correlations of sire estimated breeding value between farms differed for Model 1 for all traits. However, sires ranked similarly (rank correlation of 0.99) for weight traits with Model 2, but less so for ultrasonic measures. Including the F × S interaction (Model 3) improved model fit for all traits. However, the F × S term explained a small proportion of variation in weights (<2%) although more in ultrasonic traits (at least 10%). In conclusion, heteroscedasticity and G × E were not large for these data, and can be ignored in genetic evaluation of weight but, perhaps, not ultrasonic traits. Still, before incorporating heteroscedasticity and G × E into routine evaluations of even ultrasonic traits, their consequences on selection response in the breeding goal should be evaluated.

Keywords: crossbred lambs, genetics by environment interaction, heterogeneous variances, sheep

Implications

Genetics by environment interaction (G × E) and heterogeneous environmental variances may impact genetic evaluation. Where appreciable, sheep reared in different environments may not perform predictably. Different variances across environments were found, with G × E more pronounced for ultrasonic than for weights traits up to slaughter. Still, their impacts were generally small. Genetic evaluation aims to assist livestock industries to achieve defined breeding goals; environmental heterogeneity and G × E can slow progress toward that aim. Although incorporating heteroscedasticity and G × E into genetic evaluation of ultrasonic traits may be justified, the utility of doing so must be considered within the framework of industry breeding goals.

Introduction

An animal's phenotype reflects a combination of its genetics and environment. Selection often takes place among animals that are reared in different climatic and husbandry conditions, and animals (and their progeny) may not perform

[a] Present address: ABS Global Inc., DeForest, WI 53532, USA.
[†] Present address: Animal Science Department, University of Nebraska, Lincoln, NE 68583, USA. E-mail: rlewis5@unl.edu

uniformly across them. Nonetheless, genetic evaluation programmes often assume that animals will perform consistently across environments, and that variability in performance in different environments will be similar. A wealth of evidence has shown that is not the case, and that ignoring such effects had unfavorable consequences on genetic evaluation schemes (Robert-Granè *et al*., 1999; Mulder and Bijma, 2005).

Differences in phenotypic variances across flocks can arise from differences in production conditions such as management, nutrition and climate. Such environmental heteroscedasticity (sub-populations with different environmental variances) has been found in several livestock species for a multitude of traits (SanCristobal-Gaudy *et al*., 2001; Rowe *et al*., 2006; Nakaoka *et al*., 2007). Variable performance levels across flocks can also arise from sensitivities of genotypes to their environmental circumstances. Such genotype by environment interactions (G × E) have been observed in sheep and other species (e.g. Maniatis and Pollott, 2002; Pollott and Greeff, 2004; Steinheim *et al*., 2008).

Ignoring environmental heteroscedasticity and G × E can hinder the robustness of genetic evaluations. Accuracy of selection can be affected, leading to decreases in genetic response (Mulder and Bijma, 2005). Variance components may be poorly estimated and estimated breeding value (EBV) biased, leading to re-rankings of animals (Hill, 1984; Garrick and Van Vleck, 1987). These effects often were greater when animals were selected on EBV derived from individual phenotypes, which remains the norm in livestock species, rather than on family mean performance (Hill and Zhang, 2004).

In the United Kingdom, 70% of the lamb crop has had terminal sire breeding, with Charollais, Suffolk and Texel the predominant breeds used (Pollott and Stone, 2004). Environments in which lambs were reared also differ. By performance testing terminal sire rams in several environments, the extent and consequence of heteroscedasticity and G × E on genetic evaluation can be examined. Such were the objectives of this study using a population of terminal sire cross-lambs reared under commercial conditions.

Material and methods

Animal care and use

The Animal Experiment Committees at the Institute of Biological Environmental and Rural Sciences, the Scottish Agricultural College (SAC) and ADAS UK Ltd (ADAS) approved all procedures and protocols used in the experiment.

Animal resources

Data on 6325 crossbred lambs sired by Charollais, Suffolk and Texel rams were obtained. There were a total of 89 rams, which came from their breed's sire referencing schemes. These were cooperative breeding schemes where reference rams were shared among flocks to create connectedness and facilitate within-breed genetic evaluation. The rams were selected according to a lean growth index designed to increase carcass lean growth, while constraining fat growth at a constant age end point (Simm and Dingwall, 1989). Sires were chosen from the top and bottom 5% of available rams based on index score and categorized as 'high' or 'low' lean growth index. High *v.* low index rams differed in their EBV when evaluated at ~21 weeks of age. In high index rams, live weight EBV were 6.6 ± 0.5 kg greater, ultrasonic muscle depth (UMD) EBV were 2.3 ± 0.2 mm thicker, and ultrasonic fat depth EBV were 0.49 ± 0.12 mm thinner, than in low index rams (Márquez *et al*., 2012).

Lambs in this study came from mating of the terminal sires to Scottish or Welsh Mules. The Mule ewes were developed from the matings of Bluefaced Leicester rams with Scottish Blackface and (Welsh) Hardy Speckled Face ewes (van Heelsum *et al*., 2003; Mekkawy *et al*., 2009). Matings between Mule ewes and terminal sires took place between 1999 and 2002 on three farms in the United Kingdom (one each in England, Scotland and Wales). Most sires were used for two breeding seasons and were physically moved between farms to create genetic links among farms and years (Márquez *et al*., 2012 and 2013). Matings were designed so that the number of rams from high and low index categories, and from the three breeds, were balanced across farms, years and ewe breeds.

At birth, lamb parentage and weight (BWT) were recorded. Mule ewes were turned out to pasture within 48 h of lambing with at most two lambs. Excess lambs were fostered to other ewes. Singletons and twins were grazed separately. Lamb's weights were further recorded at ~5 week (5WT) and 10 week (10WT) of age.

Once lambs were ~10 weeks old they were evaluated subjectively for finishing condition every 2 weeks. This entailed lambs being restrained and assessed for fatness by palpation of the vertebral process and ribs. The fatness score ranged from 1 (devoid) to 5 (extreme), with L and H indicating 'low' and 'high' condition within a score, respectively. They were slaughtered once reaching a target finished condition of 3L fat score, which corresponded to ~11% subcutaneous fat (Kempster *et al*., 1986). Lambs were finished to a constant fatness so they could be compared at equitable levels of physiological maturity. Upon finishing, lambs' weights, henceforth referred to as slaughter weight (SWT), were obtained. The lambs were also ultrasonically scanned for muscle and fat depth. Their UMD was measured at the deepest point of the eye muscle (*longissimus lumborum*) at the third lumbar vertebra. Ultrasonic fat depth was measured at the same location and at 1 and 2 cm lateral to it and averaged. When finished, lambs were processed at a commercial abattoir. Further details of design and husbandry were provided by Márquez *et al*. (2012 and 2013).

Genetic groups

A pedigree was assembled, which consisted of 1 325 736 animals. There were six distinct (unrelated) breed types in the pedigree. Unknown parents for each breed were fitted as a genetic group: one for each terminal sire breed (the sires of the lambs), one for each Mule ewe breed types (the dams of the lambs) and one for the Bluefaced Leicester (the maternal

grandsires of lambs). Across breeds the unknown parents were unrelated justifying their fit as separate genetic groups. In addition, by fitting groups, differences in genetic means among breeds were accounted for, thereby reducing bias in the evaluation (Van Vleck, 1990).

Heterosis effects could not be explicitly fit in the analyses as performance and pedigree data on the hill breeds used to establish the crosses were unavailable. However, the combination of breed types (½ terminal sire breed, ¼ hill breed, ¼ Bluefaced Leicester) was consistent for all lambs and therefore the expected levels of heterozgosity. Furthermore, by fitting genetic groups in the analyses, lamb EBV were adjusted for mean differences in parental breeds. All analyses in this study were performed using ASReml (Gilmour *et al.*, 2009).

Heteroscedasticity

The traits investigated were BWT, 5WT, 10WT, SWT, UMD and log transformed ultrasonic fat depth (logUFD). Ultrasonic fat depth was transformed to approximate normality. Analyses of the effects of index selection on these traits have been reported previously (Márquez *et al.*, 2012 and 2013).

Within farm. Heteroscedasticity owing to farm was tested by creating three subsets of data based on where lambs were born and reared. There were 2322, 2137 and 1866 lambs born in England, Scotland and Wales, respectively. The model fitted was:

$$y_i = X_i\beta_i + Z_{a_i}a_i + Z_{d_i}d_i + e_i \qquad \text{(Model 1)}$$

where y_i is the vector of observations, β_i the vector of fixed effects coefficients, a_i the vector of genetic animal effects, d_i the vector of rearing dam effects and e_i the vector of random residual effects. The X_i, Z_{ai} and Z_{di} matrices were incidence matrices relating to observations in β_i, a_i and d_i, respectively. The *i* subscript referred to data from each of the three farms. Fixed effects were an overall mean, lamb sex (ewe or wether), age of dam (2 to 5 years) and birth year (2000 to 2003). For all traits except BWT, a birth-rearing rank effect was fitted with four categories: single born/single reared, twin or more born/single reared, single or twin born/twin reared and triplet born/twin reared. For BWT, birth rank (single, twin or triplet) was fitted. Covariates for all traits except SWT and UMD were age at measurement. For SWT and UMD, the covariate was estimated subcutaneous fat per cent at slaughter. Fat score was transformed to subcutaneous fat per cent according to Kempster *et al.* (1986).

The (co)variance structure of Model 1 was:

$$\text{var}\begin{bmatrix} a_i \\ d_i \\ e_i \end{bmatrix} = \begin{bmatrix} \mathbf{A}\sigma^2_{a_i} & 0 & 0 \\ 0 & \mathbf{I}\sigma^2_{d_i} & 0 \\ 0 & 0 & \mathbf{I}\sigma^2_{e_i} \end{bmatrix}$$

where **A** is the numerator relationship matrix among animals in the pedigree and **I** the identity matrix of appropriate dimensions, σ^2_{ai} the additive genetic variance, σ^2_{di} the environmental rearing dam variance and σ^2_{ei} the residual environmental variance. Genetic groups were considered in **A**. Since the data were on crossbred animals, estimates of genetic variance were possibly increased by dominance effects. However, as noted earlier, it was presumed that heterotic effects were consistent among lambs in these data. Heritabilities were estimated within farm as the ratio of genetic variance to the sum of the total variances (i.e. $h^2_i = \sigma^2_{a_i}/(\sigma^2_{a_i} + \sigma^2_{d_i} + \sigma^2_{e_i})$).

A likelihood ratio test revealed that rearing dam did not explain substantial variation in slaughter traits (SWT, UMD, logUFD; $P > 0.2$), and therefore the rearing dam random effect was omitted for these traits. A maternal additive effect could not be fitted because of the lack of pedigree information on Scottish Blackface and Hardy Specked Face hill breeds, the dam breeds of the Mule ewes.

For each trait, log likelihoods for data from each farm were obtained. These were independent samples, and therefore the log likelihoods were summed and compared against a model fitted to the combined data. In the combined model, additional effects of farm and farm by birth year interaction were included. In the absence of heteroscedasticity, the sum of the log likelihoods from the independent samples and the log likelihood from the combined data would be expected to be equal. **A** likelihood ratio test with 2 d.f. was used to test whether the sum of the log likelihoods from the independent samples differed from the log likelihood from the combined data. Rank correlations of EBV from the combined and within-farm data were obtained to investigate any consequences of variance heterogeneity. Some sires did not have progeny on all farms. For those that did, re-rankings of sires were investigated, and correlations between EBV in the different farms were obtained.

Across farm. The second method to test variance heterogeneity was by fitting heterogeneous residual (farm) variances (Model 2). In this model, the combined data were used, but separate residual variances were estimated for each farm. The fixed effects of Model 1, in addition to farm, and farm by year interaction, were fitted to all the data with a modified (co)variance structure. The (co)variance matrix remained the same as in Model 1, except:

$$\text{var}\begin{bmatrix} a \\ d \\ e_1 \\ e_2 \\ e_3 \end{bmatrix} = \begin{bmatrix} \mathbf{A}\sigma^2_a & 0 & 0 & 0 & 0 \\ 0 & \mathbf{I}\sigma^2_d & 0 & 0 & 0 \\ 0 & 0 & \mathbf{I}\sigma^2_{e_1} & 0 & 0 \\ 0 & 0 & 0 & \mathbf{I}\sigma^2_{e_2} & 0 \\ 0 & 0 & 0 & 0 & \mathbf{I}\sigma^2_{e_3} \end{bmatrix} \qquad \text{(Model 2)}$$

where $\sigma^2_{e_i}$ ($i = 1, 2, 3$) is the residual variance of farm *i*. Within-farm heritabilities for this model were calculated as $h^2_i = \sigma^2_a/(\sigma^2_a + \sigma^2_d + \sigma^2_{e_i})$.

The log likelihood for this model was obtained for each trait, and was tested against a null model with a single residual variance component, with a likelihood ratio test with 2 d.f. The consequences of heteroscedasticity were investigated by obtaining rank correlation of EBV calculated assuming either heterogeneous or homogeneous environmental variances.

Genotype by environment interaction

To investigate the presence of G × E, an animal model was fitted with a random farm by sire (F × S) interaction term (Model 3). Fixed effects were the same as in Model 1. Random effects were animal, farm, F × S and a random residual. A random rearing dam was fitted for BWT, 5WT and 10WT. The (co)variance structure for this model was:

$$\text{var}\begin{bmatrix} a \\ f \\ fxs \\ d \\ e \end{bmatrix} = \begin{bmatrix} \mathbf{A}\sigma_a^2 & 0 & 0 & 0 & 0 \\ 0 & \mathbf{I}\sigma_f^2 & 0 & 0 & 0 \\ 0 & 0 & \mathbf{I}\sigma_{fxs}^2 & 0 & 0 \\ 0 & 0 & 0 & \mathbf{I}\sigma_d^2 & 0 \\ 0 & 0 & 0 & 0 & \mathbf{I}\sigma_e^2 \end{bmatrix} \quad \text{(Model 3)}$$

where **A** is the numerator relationship matrix, σ_a^2, σ_f^2 and $\sigma_{f\times s}^2$ the variance components associated with animal (additive genetic), farm and F × S, respectively. Other variance components were defined as in Model 1 and Model 2. The F × S interaction component would indicate the amount of G × E in a population (Dickerson, 1962). To test for its significance, a likelihood ratio test was performed by comparing it to a model without the random F × S interaction term. The ratio of F × S to total variance was calculated to quantify the extent of G × E in the population. The heritability was calculated as the ratio of genetic variance to total variance.

To investigate whether any G × E was caused by heterogeneous phenotypic variances, traits were standardized to their within-farm variance, and Model 3 was again fitted. Large differences in variance component estimates, and re-ranking of sires in standardized as compared with unstandardized data, would indicate the importance of variance heterogeneity.

Connectedness

In order to avoid biases in our EBV, the study was designed to establish sound genetic links, or connectedness, among farm locations within and across terminal sire breeds and index categories. The sufficiency of the design was explored by quantifying the strength of connections using prediction error correlations (Lewis *et al.*, 2005; Kuehn *et al.*, 2007 and 2008). Using 5WT as the example trait, and a heritability of 0.20, connectedness correlations were derived among farms and breed-index categories. The mixed linear animal model fitted included farm-year combination, sex-birth rearing type combination and age of dam as fixed effects.

Results

Summary statistics for BWT, 5WT, 10WT, SWT, UMD and logUFD are provided in Table 1 relative to sire breed. As reported previously (Márquez *et al.*, 2012 and 2013), weights and ultrasound measures differed with respect to sire breed, although changes in means were generally proportional to changes in s.d. (similar CV across breeds).

Within farm

When the data were separated by farm, likelihood ratio tests indicated the presence of heterogeneity in the environmental variance for all traits ($P < 0.01$). However, the estimates of total variance and heritability were similar for the combined data, and for within each subset of farm data (Table 2).

Rank correlations between lamb EBV with the full data and farm subsets ranged from: 0.77 to 0.81 for BWT; 0.55 to 0.93 for 5WT; 0.57 to 0.74 for 10WT; 0.71 to 0.82 for SWT; 0.70 to 0.83 for UMD and 0.76 to 0.95 for logUFD. The rank correlations estimated within a particular farm were not consistently higher or lower than those in the other farms, nor were there clear patterns among correlations within farms. The rank correlations among lamb EBV were higher than those among sire EBV, reflecting the fewer numbers of sires than lambs on individual farms (results not shown).

Across farm

Allowing for heterogeneous environmental variances among farms (Model 2) provided a better fit to the data for all traits ($P < 0.01$). However, when comparing the genetic variances and heritabilities obtained from models with heterogeneous *v.* homogenous variance structures, they were within the standard error for most traits (except SWT and UMD) (Table 3).

Rank correlations between EBV obtained from the homogenous and heterogeneous variance models were 0.99 for all

Table 1 *Summary statistics for birth, 5 week, 10 week and slaughter weights, and for ultrasonic muscle (UMD) and log transformed fat depths (logUFD) by sire breed*

Trait	Mean	s.d.	CV (%)	Minimum	Maximum
Birth weight (kg)					
Charollais	4.7	0.93	19.6	2.0	8.3
Suffolk	4.8	0.94	19.6	2.2	8.5
Texel	4.7	0.96	20.3	2.0	8.2
5-week weight (kg)					
Charollais	16.3	3.69	22.6	5.8	31.5
Suffolk	16.9	3.68	21.8	5.5	28.8
Texel	16.6	3.85	23.2	5.5	29.5
10-week weight (kg)					
Charollais	26.3	5.36	20.4	7.6	44.2
Suffolk	26.9	5.04	18.8	11.3	43.0
Texel	26.4	5.32	20.1	9.0	44.3
Slaughter weight (kg)					
Charollais	42.2	4.62	11.0	29.0	62.0
Suffolk	42.5	4.68	11.0	29.8	61.0
Texel	40.7	4.43	10.9	28.0	59.2
UMD (mm)					
Charollais	24.8	2.20	8.9	17.5	33.0
Suffolk	24.6	2.19	8.9	18.3	32.3
Texel	24.9	2.25	9.1	17.0	36.2
logUFD (mm)					
Charollais	1.4	0.31	22.6	0.2	2.4
Suffolk	1.3	0.29	22.2	0.4	2.2
Texel	1.3	0.30	22.8	0.1	2.5

Table 2 *Estimates of genetic and environmental variance and heritability for growth and slaughter traits in sheep. Combined model includes all data, and country subsets includes data only from farm in that country*

	Trait					
	BWT (kg^2)	5WT (kg^2)	10WT (kg^2)	SWT (kg^2)	UMD (mm^2)	logUFD (mm^2)
Genetic variance						
Combined	0.110 ± 0.023	0.69 ± 0.15	1.68 ± 0.36	5.29 ± 0.64	1.33 ± 0.15	0.019 ± 0.003
England	0.094 ± 0.034	0.59 ± 0.24	2.01 ± 0.63	5.86 ± 0.95	1.31 ± 0.23	0.027 ± 0.004
Scotland	0.097 ± 0.033	1.25 ± 0.37	1.81 ± 0.63	6.46 ± 1.18	1.43 ± 0.24	0.015 ± 0.003
Wales	0.094 ± 0.034	0.67 ± 0.26	1.32 ± 0.53	4.39 ± 0.98	1.60 ± 0.29	0.027 ± 0.005
Environmental variance						
Combined	0.27 ± 0.02	3.61 ± 0.12	8.07 ± 0.26	10.67 ± 0.47	2.67 ± 0.11	0.046 ± 0.002
England	0.29 ± 0.01	2.89 ± 0.17	5.84 ± 0.41	7.61 ± 0.66	2.69 ± 0.18	0.035 ± 0.003
Scotland	0.26 ± 0.02	2.54 ± 0.21	5.73 ± 0.41	11.79 ± 0.89	1.96 ± 0.17	0.046 ± 0.003
Wales	0.29 ± 0.02	4.20 ± 0.23	9.73 ± 0.51	11.76 ± 0.81	3.19 ± 0.23	0.046 ± 0.003
Heritability[1]						
Combined	0.22 ± 0.04	0.13 ± 0.03	0.14 ± 0.03	0.33 ± 0.04	0.33 ± 0.04	0.30 ± 0.04
England	0.18 ± 0.06	0.12 ± 0.05	0.19 ± 0.06	0.43 ± 0.06	0.33 ± 0.05	0.43 ± 0.06
Scotland	0.20 ± 0.06	0.26 ± 0.07	0.17 ± 0.06	0.35 ± 0.06	0.42 ± 0.06	0.24 ± 0.05
Wales	0.18 ± 0.05	0.12 ± 0.05	0.10 ± 0.04	0.27 ± 0.06	0.33 ± 0.05	0.38 ± 0.06

BWT = birth weight; 5WT = 5-week weight; 10WT = 10-week weight; SWT = slaughter weight; UMD = ultrasonic muscle depth; logUFD = log ultrasonic fat depth.
[1]Heritabilities are without units.

Table 3 *Genetic and environmental variances and heritabilities for homogeneous and heterogeneous variance models for growth and slaughter traits*

	BWT (kg^2)	5WT (kg^2)	10WT (kg^2)	SWT (kg^2)	UMD (mm^2)	logUFD (mm^2)
Genetic variance						
HOM	0.12 ± 0.03	0.91 ± 0.18	2.11 ± 0.41	6.01 ± 0.67	1.50 ± 0.16	0.024 ± 0.003
HET	0.13 ± 0.02	0.94 ± 0.19	2.14 ± 0.42	6.00 ± 0.67	1.34 ± 0.15	0.020 ± 0.003
Environmental variance						
HOM	0.27 ± 0.01	3.16 ± 0.12	6.87 ± 0.26	10.22 ± 0.49	2.58 ± 0.12	0.004 ± 0.002
England	0.28 ± 0.02	2.88 ± 0.16	5.85 ± 0.32	12.44 ± 0.66	2.02 ± 0.13	0.005 ± 0.002
Scotland	0.24 ± 0.02	2.73 ± 0.15	5.98 ± 0.32	7.72 ± 0.53	2.69 ± 0.14	0.004 ± 0.002
Wales	0.30 ± 0.02	3.96 ± 0.19	9.03 ± 0.43	10.82 ± 0.64	3.41 ± 0.17	0.053 ± 0.003
Heritability[1]						
HOM	0.24 ± 0.04	0.17 ± 0.03	0.18 ± 0.03	0.37 ± 0.04	0.36 ± 0.03	0.34 ± 0.04
England	0.24 ± 0.04	0.19 ± 0.04	0.20 ± 0.04	0.33 ± 0.03	0.39 ± 0.04	0.30 ± 0.04
Scotland	0.26 ± 0.05	0.20 ± 0.04	0.20 ± 0.04	0.44 ± 0.04	0.33 ± 0.03	0.33 ± 0.04
Wales	0.23 ± 0.04	0.16 ± 0.03	0.15 ± 0.03	0.35 ± 0.04	0.28 ± 0.03	0.27 ± 0.03

BWT = birth weight; 5WT = 5-week weight; 10WT = 10-week weight; SWT = slaughter weight; UMD = ultrasonic muscle depth; logUFD = log ultrasonic fat depth; HOM = homogeneous variances model; HET = heterogeneous variances model.
[1]Heritabilities are without units.

weight traits (both animals and sires), and 0.88 and 0.84 for UMD and logUFD, respectively, among sires. These results indicate that re-ranking only would be observed for ultrasonic traits, although they would not be substantial. The across farm estimates of heritabilities were similar to the within-farm heritabilities of Model 1.

Genotype by environment interaction

For all traits, including a random F × S interaction in the model resulted in a better fit ($P < 0.001$, except $P = 0.02$ for SWT). Heritabilities were similar to those estimated in Models 1 and 2. The proportion of the F × S variance to total variance was small for weight traits, but more pronounced for ultrasonic measures (Table 4). Standardizing traits to a common within-farm variance did not have an effect on variance components or rankings (results not shown).

Connectedness

Among farm locations, connectedness correlations were between 0.61 and 0.67. Between the high and low index category within a breed, these correlations ranged from 0.44 for the Suffolk to 0.53 for the Charollais. Values between breeds were only slightly lower (0.40). Correlations of 0.10 and above were shown to be indicative of strong connectedness (Kuehn *et al.*, 2008). Although there were only 8 sires shared between Wales and Scotland, 14 between

Table 4 *Variance components estimates for the genetics by environment interaction models for growth and slaughter traits*

	BWT (kg^2)	5WT (kg^2)	10WT (kg^2)	SWT (kg^2)	UMD (mm^2)	logUFD (mm^2)
Genetic variance	0.18 ± 0.03	1.02 ± 0.02	2.31 ± 0.53	6.60 ± 0.74	1.41 ± 0.20	0.026 ± 0.003
F × S variance	0.009 ± 0.004	0.09 ± 0.04	0.19 ± 0.09	0.22 ± 0.14	0.47 ± 0.11	0.013 ± 0.002
Heritability[1]	0.30 ± 0.05	0.15 ± 0.05	0.15 ± 0.05	0.37 ± 0.04	0.30 ± 0.04	0.28 ± 0.05
G × E[1,2]	0.015 ± 0.007	0.013 ± 0.007	0.012 ± 0.007	0.012 ± 0.008	0.10 ± 0.02	0.13 ± 0.03

BWT = birth weight; 5WT = 5-week weight; 10WT = 10-week weight; SWT = slaughter weight; UMD = ultrasonic muscle depth; logUFD = log ultrasonic fat depth; F × S = sire by farm interaction; G × E = genetics by environment interaction.
[1]Heritability and G × E are without units.
[2]G × E defined as F × S variance as a proportion of total variance.

Wales and England, and 13 between Scotland and England, the rotation of rams among farms generated the well-connected design intended.

Discussion

Variance heterogeneity

Heteroscedasticity was present in this population, especially for ultrasonic traits. In the combined data, the additive genetic variance was similar to that estimated within farms (Model 1). These estimates changed little when fitting Model 2. Such was the case even when a homogeneous farm variance was assumed.

For both weight and ultrasound traits, accounting for heterogeneous variances improved model fit. However, for the weight traits, rank correlations between EBV obtained with homogenous and heterogeneous variances were near one. This suggested that any consequences of heteroscedasticity were not pronounced for weight traits, in agreement with previous results (Canavesi *et al*., 1995). Sire re-ranking was more evident for UMD and logUFD, suggesting heteroscedasticity would have a greater effect on the genetic evaluation of ultrasound traits.

Ignoring heterogeneous variances in genetic evaluation has risks. As observed in this study, animals may be incorrectly ranked resulting in lower selection response. Accuracies of EBV may also be affected. By fitting a heterogeneous variance model, EBV would be scaled, lessening the impact of inaccuracies in the estimation (Gianola, 1986). Given the presence of heterogeneous variances, several livestock breeds have developed genetic evaluation models that account for heteroscedasticity (Wiggans and VanRaden, 1991; Nakaoka *et al*., 2007).

An effective way to mediate bias in EBV owing to heterogeneous variances would be to test progeny in different environments. In progeny testing of dairy cattle, ranking of bulls was not greatly affected by heteroscedasticity when their daughters were randomly distributed among farms with high and low variances (Winkelman and Schaeffer, 1988). Sire referencing schemes, such as those from which the rams used in this study were drawn, provide another way of distributing genetics of sires to many flocks. It has been reported that assumptions of homogeneity may not lead to substantial decreases in selection response when heritabilities are higher in more variable populations (Garrick and Van Vleck, 1987). No such pattern was found in these data.

Evidence for heterogeneity of variances within individual sheep breeds has been reported. SanCristobal-Gaudy *et al*. (2001) found that selecting for increased litter size led to increases in variability of the trait, and that using a heterogeneous variance model resulted in increased selection response. In a study comparing different breeds, Tosh and Kemp (1994) found variable estimates of heritability for weights up to 100 days in three breeds (Hampshire, Polled Dorset and Romanov). They also report heterogeneous breed variances, and suggested accounting for breed specific variance estimates may be necessary when comparing different breeds in an across-breeds genetic evaluation.

Genetics by environment interactions

The ratio of F × S to total variance was shown to be indicative of the presence and influence of G × E within a population (Dickerson, 1962; Meyer, 1987). For weight traits, F × S explained ~1% of the total variation. For ultrasonic traits, this percentage was greater (10 to 13%), indicating that G × E has a larger influence on body composition traits. For weight traits, our results were similar to Maniatis and Pollott (2002), also in sheep; however, they reported a lower proportion of variance owing to F × S in ultrasonic traits than in the current study.

In our case, including the F × S effect in the analyses decreased estimates of heritability. Such was also the case for Maniatis and Pollott (2002). Here, as in their study, ignoring F × S may have inflated estimates of additive genetic variance. They hypothesized that some of the additive genetic variance was being partitioned into the F × S variance component, yielding downwardly biased heritabilities. Shrunk additive genetic variances were also found by Hagger (1998) for ADG in sheep when fitting an F × S effect. Therefore, levels of G × E in production traits appear to be low but real in sheep populations.

Misztal (1990) suggested that an explanation for a significant F × S interaction was poor representation of sires across flocks, where genetic evaluations were more severely regressed. In our study, sires were well represented across flocks, with a proportion of sires having progeny in two of the three farms. The connectedness among farms was also strong. Another reason for the F × S interaction may be preferential treatment of some half-sib groups (Meyer, 1987). However, given the design of this experiment, with

management intentionally standardized across farms, such would not be anticipated.

Ultrasonic traits had greater indication of heteroscedasticity than weight traits, and also had a higher proportion of variation explained by the F × S interaction. Dickerson (1962) and Canavesi *et al.* (1995) found that F × S interaction may be caused by, or at least inflated by, heterogeneous variances. When variances were standardized across farms, the variance component estimates, and the proportion of F × S interaction variance to total variance, did not change. Notter *et al.* (1992) and Maniatis and Pollott (2002) reported similar results.

Effects on genetic evaluation

Weight at slaughter reflects an animal's growth to a certain end point, such as a target level of fatness. As such, it is a combination of the bone, fat, lean and other tissues deposited in an animal as it grows. Evidence of heterogeneity and G × E was not observed in SWT, or in earlier weights, but it was in ultrasonic traits. Ultrasonic measures were shown to be indicative of fat and lean tissue deposition in an animal (Emenheiser *et al.*, 2010), and therefore can be thought of as components of SWT. Perhaps when considering the components rather than the culmination of growth, heterogeneity and G × E become more apparent. Our findings indicate that accounting for heterogeneity and G × E in genetic evaluation of ultrasonic measures, at least in progeny of terminal sires, will reduce such bias.

In selection regimes, where animals were often reared in environments that differed, ignoring G × E when estimating variance components in genetic evaluation led to reductions in selection response (Garrick and Van Vleck, 1987; Mulder and Bijma, 2006). Mulder and Bijma (2005) found that progeny testing schemes were more robust to G × E than sib-testing schemes: when including information on progeny, in the presence of any G × E, the rate of genetic change was greater. The current data were derived from a progeny testing scheme. It was therefore anticipated that it would have less of an impact of G × E than otherwise.

In the presence of G × E, the breeding objective of selection programmes in different environments may differ. The construction of selection tools may also differ because genetic (co)variances between traits may vary across environments. With the presence of G × E, a way to optimize selection programmes would be to have an overall breeding goal yet test progeny in more than one environment, as was the case in the current study.

Clearly the consequences of heteroscedasticity or G × E on genetic evaluation programmes must be carefully considered before being incorporated into genetic evaluation. The limited extent of environmental heteroscedasticity observed in this study may justify it being ignored even for ultrasonic traits, as re-ranking of sires was trivial. Accounting for any G × E in the genetic evaluation of ultrasonic traits may be more important: the F × S random component explained at least 10% of the variation in these traits. Still, to robustly estimate the F × S effect, the number of offspring per sire needs to be large enough and connectedness among their offspring needs to be sufficient. Such was the case in this study but may not be so in industry breeding schemes.

Even where heteroscedasticity or G × E may be important, incorporating them into genetic evaluation schemes could be complicated. First, environments must be delineated. In the current study this was straightforward; by its design, lambs were reared in three distinct locations within the United Kingdom. However, in genetic evaluation schemes, environments may be less easily distinguished, may overlap and may vary gradually across geographic regions and climates. Furthermore, environmental conditions would not be static over time, even on individual farms.

When deciding whether to incorporate G × E or heterogeneous variances into genetic evaluation, the efficacy of running such evaluations also deserves consideration. When fitting models with more random effects, solutions may be more difficult to obtain. Furthermore, the amount of data in current routine genetic evaluations would be large, with computational time a constraint. Therefore, the costs of accounting for heteroscedasticity and G × E in routine, particularly multivariate, genetic evaluations need to be considered.

Conclusions

The aim of genetic evaluation programmes is to assist livestock industries achieve defined breeding goals. The presence of environmental heterogeneity or G × E may hinder progress towards these goals. However, before incorporating such factors into routine genetic evaluations, their extent and consequence on reaching breeding goals need to be carefully evaluated. In the present study, incorporating such comprehensive statistical models for weight traits was not warranted.

Acknowledgements

The authors would like to acknowledge the Meat and Livestock Commission (now English Lamb and Beef Executive (EBLEX), Quality Meats Scotland (QMS), and Hybu Cig Cymru – Meat Promotion Wales (HCC); the UK Department of Environment, Food and Rural Affairs, The British Charollais Sheep Society (Norfolk, UK), the Suffolk Sheep Society (Ballymena, UK) and the British Texel Sheep Society (Kenilworth, UK) for funding of this study. The authors would also like to thank the technical staff at Aberystwyth University, ADAS Rosemaund and the Scottish Agricultural College (now Scotland's Rural College) for their support of the experimental programme, and E. Stephens for technical editing. The authors acknowledge Advanced Research Computing at Virginia Tech for providing computational resources and technical support that have contributed to the results reported within this paper (http://www.arc.vt.edu).

References

Canavesi F, Schaeffer LR, Burnside EB, Jansen GB and Rozzi P 1995. Sire-by-herd interaction effect when variances across herds are heterogeneous. I. Expected genetic progress. Journal of Animal Breeding and Genetics 112, 95–106.

Dickerson GE 1962. Implications of genetic-environmental interaction in animal breeding. Animal Production 4, 47–63.

Emenheiser JC, Greiner SP, Lewis RM and Notter DR 2010. Validation of live animal ultrasonic measurements of body composition in market lambs. Journal of Animal Science 88, 2932–2939.

Garrick DJ and Van Vleck LD 1987. Aspects of selection for performance in several environments with heterogeneous variances. Journal of Animal Science 65, 409–421.

Gianola D 1986. On selection criteria and estimation of parameters when the variance is heterogeneous. Theoretical and Applied Genetics 72, 671–677.

Gilmour AR, Gogel BJ, Cullis BR and Thompson R 2009. ASReml User Guide Release 3.0. VSN International Ltd, Hemel, Hempstead, UK.

Hagger C 1998. Litter, permanent environmental, ram-flock, and genetic effects on early weight gain of lambs. Journal of Animal Science 76, 452–457.

Hill WG 1984. On selection among groups with heterogeneous variance. Animal Production 39, 473–477.

Hill WG and Zhang XS 2004. Effects on phenotypic variability of directional selection arising through genetic differences in residual variability. Genetical Research 83, 121–132.

Kempster AJ, Cook GL and Grantley-Smith M 1986. National estimates of the body composition of British cattle, sheep and pigs with special reference to trends in fatness. A review. Meat Science 17, 107–138.

Kuehn LA, Lewis RM and Notter DR 2007. Managing the risk of comparing estimated breeding values across flocks or herds through connectedness: a review and application. Genetics Selection Evolution 39, 225–247.

Kuehn LA, Notter DR, Nieuwhof GJ and Lewis RM 2008. Changes in connectedness over time in alternative sheep sire referencing schemes. Journal of Animal Science 86, 536–544.

Lewis RM, Crump RE, Kuehn LA, Simm G and Thompson R 2005. Assessing connectedness in across-flock genetic evaluations. Journal of Animal Science 83 (suppl. 1), 101.

Maniatis N and Pollott GE 2002. Genotype by environment interactions in lamb weight and carcass composition traits. Animal Science 75, 3–14.

Márquez GC, Haresign W, Davies MH, Roehe R, Bünger L, Simm G and Lewis RM 2013. Index selection in terminal sires improves lamb performance at finishing. Journal of Animal Science 91, 38–43.

Márquez GC, Haresign W, Davies MH, Emmans GC, Roehe R, Bünger L, Simm G and Lewis RM 2012. Index selection in terminal sires improves early lamb growth. Journal of Animal Science 90, 142–151.

Mekkawy W, Roehe R, Lewis RM, Davies MH, Bünger L, Simm G and Haresign W 2009. Genetic relationship between longevity and objectively or subjectively assessed performance traits in sheep using linear censored models. Journal of Animal Science 87, 3482–3489.

Meyer K 1987. Estimates of variances due to sire x herd interactions and environmental covariances between paternal half-sibs for first lactation dairy production. Livestock Production Science 17, 95–115.

Misztal I 1990. Restricted maximum likelihood estimation of variance components in animal model using sparse matrix inversion and a supercomputer. Journal of Dairy Science 73, 163–172.

Mulder HA and Bijma P 2005. Effects of genotype x environment interaction on genetic gain in breeding programs. Journal of Animal Science 83, 49–61.

Mulder HA and Bijma P 2006. Benefits of cooperation between breeding programs in the presence of genotype by environment interaction. Journal of Dairy Science 89, 1727–1739.

Nakaoka H, Narita A, Ibi T, Sasae Y, Miyake T, Yamada T and Sasaki Y 2007. Effectiveness of adjusting for heterogeneity of variance in genetic evaluation of Japanese Black cattle. Journal of Animal Science 85, 2429–2436.

Notter DR, Tier B and Meyer K 1992. Sire x herd interactions for weaning weight in beef cattle. Journal of Animal Science 70, 2359–2365.

Pollott GE and Greeff JC 2004. Genotype x environment interactions and genetic parameters for fecal egg count and production traits of Merino sheep. Journal of Animal Science 82, 2840–2851.

Pollott GE and Stone DG 2004. Mating structure of the sheep industry. In The breeding structure of the British sheep industry 2003 (ed. RD Eglin, A Ortiz Pelaez and CJ Cook), pp. 18–22. Department for Environment Food and Rural Affairs, London, UK.

Robert-Granìè C, Bonaôti B, Boichard D and Barbat A 1999. Accounting for variance heterogeneity in French dairy cattle genetic evaluation. Livestock Production Science 60, 343–357.

Rowe SJ, White IMS, Avendaño S and Hill WG 2006. Genetic heterogeneity of residual variance in broiler chickens. Genetics Selection Evolution 38, 617–635.

SanCristobal-Gaudy M, Bodin L, Elsen J-M and Chevalet C 2001. Genetic components of litter size variability in sheep. Genetics Selection Evolution 33, 249–271.

Simm G and Dingwall WS 1989. Selection indices for lean meat production in sheep. Livestock Production Science 21, 223–233.

Steinheim G, Ødegård J, Ådnøy T and Klemetsdal G 2008. Genotype by environment interaction for lamb weaning weight in two Norwegian sheep breeds. Journal of Animal Science 86, 33–39.

Tosh JJ and Kemp RA 1994. Estimation of variance components for lamb weights in three sheep populations. Journal of Animal Science 72, 1184–1190.

van Heelsum AM, Lewis RM, Davies MH and Haresign W 2003. Growth and carcass characteristics in wether lambs of a crossbred dam line. Animal Science 76, 45–53.

Van Vleck LD 1990. Breeding value prediction with maternal genetic groups. Journal of Animal Science 68, 3998–4013.

Wiggans GR and VanRaden PM 1991. Method and effect of adjustment for heterogeneous variance. Journal of Dairy Science 74, 4350–4357.

Winkelman A and Schaeffer LR 1988. Effect of heterogeneity of variance on dairy sire evaluation. Journal of Dairy Science 71, 3033–3039.