

Non-genetic correlations between production traits and litter traits in pigs are important to account for

T. Ostersen^{1*}, M. Henryon¹, X. Guo^{1,2}, O.F. Christensen², G. Su² and A.C. Sørensen¹

¹ Danish Pig Research Centre, Danish Agriculture & Food Council, 1609 Copenhagen V, Denmark; ² Center for Quantitative Genetics and Genomics, Aarhus University, 8830 Tjele, Denmark; * tao@lf.dk

Abstract

We hypothesized that (i) non-genetic correlations between litter size (LS) and daily gain are different from zero, and (ii) modelling non-genetic correlations changes estimates of genetic (co)variances of the traits compared to when they are ignored. We tested these hypotheses using a new multi-trait model that separates genetic and non-genetic correlations by fitting correlations between the residual effects for one trait and non-residual effects of the other trait. We estimated variances by fitting models with and without these non-genetic correlations to LS and daily gain in DanBred Landrace pigs. We found that all non-genetic correlations between LS and daily gain were significantly different from zero. Estimates of genetic (co)variances changed significantly when non-genetic correlations were ignored. Clearly, multi-trait models with production traits and litter size traits should include correlations between residual effects for one trait and non-residual effects of the other trait to separate genetic correlations from non-genetic correlations.

Introduction

Multi-trait models are state-of-the-art for breeding value estimation (Mrode, 2014, p. 70). They provide more unbiased and accurate predictions than single-trait models when non-genetic correlations among traits differ from genetic correlations (Mrode, 2014, p. 70). Modelling non-genetic correlations is clearly important for prediction. However, some of these correlations are often neglected because the model framework has not been developed. An example is when a single phenotype is available for a trait(s) and multiple phenotypes are available for another trait(s). In pig breeding, daily gain and litter size are examples of such traits. Daily gain is often recorded as a single phenotype per animal. Litter size is recorded for each farrowing and each sow can have multiple phenotypes for litter size. A further complication is that two possible non-genetic correlations may exist. First, an animal's record of daily gain can be affected by the litter size in its birth litter. Second, a female's daily gain can affect its non-genetic permanent ability to produce large litters. Standard multi-trait models cannot cater for correlations between the residual effect for one trait and a non-genetic non-residual effect for another trait. The challenge is that neglecting them risks that they are captured by genetic correlations during REML estimation (Walter and Mao, 1985). Clearly, a new model framework that separates non-genetic correlations from genetic correlations is needed. We propose such a model and in doing so, hypothesize that (i) the non-genetic correlations between litter size and daily gain are different from zero, and (ii) modelling non-genetic correlations changes the estimates of genetic (co-)variances of the traits compared to when they are ignored.

Materials & Methods

Experimental design. We estimated (co)variances using two trivariate models fitted to litter size at day 5 (LS), average daily gain until 30kg (G1) and average daily gain from 30 to 100 kg

(G2) in DanBred Landrace pigs. The two models were with and without non-genetic correlations between LS and the production traits.

Data. There were 93,408 records of LS from 45,973 purebred Landrace sows and 64% of the litters were crossbred with DanBred Yorkshire sire and the others were purebred litters. There were 203,411 records of G1 (g/day) and 212,191 records of G2 (g/day) all from purebred DanBred Landrace animals (see Table 1). The pedigree included 237,500 purebred Landrace animals. The data we used was from litters born between January 2016 and April 2021.

Table 1. Descriptive Statistics.

Trait ¹	No. Obs	Mean	SD	Min	Max
LS	93,408	14.2	4.1	0	32
G1	203,411	382	37	295	579
G2	212,191	1,080	109	680	1,609

¹ LS: litter size at day 5; G1: average daily gain (g/day) from birth to 30 kg; G2: average daily gain (g/day) from 30 kg to 100 kg

Model. The tri-variate model to estimate variance and covariance components was:

$$\begin{bmatrix} \mathbf{y}_{LS} \\ \mathbf{y}_{G1} \\ \mathbf{y}_{G2} \end{bmatrix} = \begin{bmatrix} \mathbf{X}_{bLS} & 0 & 0 \\ 0 & \mathbf{X}_{bG1} & 0 \\ 0 & 0 & \mathbf{X}_{bG2} \end{bmatrix} \begin{bmatrix} \mathbf{b}_{LS} \\ \mathbf{b}_{G1} \\ \mathbf{b}_{G2} \end{bmatrix} + \begin{bmatrix} \mathbf{Z}_{aLS} & 0 & 0 \\ 0 & \mathbf{Z}_{aG1} & 0 \\ 0 & 0 & \mathbf{Z}_{aG2} \end{bmatrix} \begin{bmatrix} \mathbf{a}_{LS} \\ \mathbf{a}_{G1} \\ \mathbf{a}_{G2} \end{bmatrix} + \begin{bmatrix} \mathbf{Z}_{cLSp} & \mathbf{Z}_{cLSc} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \mathbf{Z}_{cG2} \end{bmatrix} \begin{bmatrix} \mathbf{c}_{LSp} \\ \mathbf{c}_{LSc} \\ \mathbf{c}_{G2} \end{bmatrix} + \begin{bmatrix} \mathbf{Z}_{pLS} & 0 & 0 \\ 0 & \mathbf{Z}_{eG1} & 0 \\ 0 & 0 & \mathbf{Z}_{eG2} \end{bmatrix} \begin{bmatrix} \mathbf{p}_{LS} \\ \mathbf{e}_{G1} \\ \mathbf{e}_{G2} \end{bmatrix} + \begin{bmatrix} \mathbf{Z}_{eLS} & 0 & 0 \\ 0 & \mathbf{Z}_{lG1} & 0 \\ 0 & 0 & \mathbf{Z}_{lG2} \end{bmatrix} \begin{bmatrix} \mathbf{e}_{LS} \\ \mathbf{l}_{G1} \\ \mathbf{l}_{G2} \end{bmatrix},$$

where \mathbf{y}_{LS} , \mathbf{y}_{G1} and \mathbf{y}_{G2} are vectors of records for LS, G1 and G2; \mathbf{b}_{LS} is a vector of fixed effects for LS including herd-year, year-month-litterbreed, year-parity, and first and second order regressions on the sow's age at first farrowing as well as time interval between farrowings; \mathbf{b}_{G1} is a vector of fixed effects for G1 including herd-year-month, sex and birth-parity as well as a regression on the age of the sow at farrowing; \mathbf{b}_{G2} is a vector of fixed effects for G2 including herd-year-stable, sex, and birth-parity. The vectors \mathbf{a}_{LS} , \mathbf{a}_{G1} and \mathbf{a}_{G2} contain the genetic effects of the sow on LS, the genetic effect of the pig on G1 and the genetic effect of the pig on G2. The vectors \mathbf{c}_{LSp} and \mathbf{c}_{LSc} contain non-genetic random service sire effects on LS for purebred and crossbred litters. The vector \mathbf{c}_{G2} contains random effects of pen on G2. The vector \mathbf{p}_{LS} contains permanent non-genetic effects of the sow on LS, whereas the vectors \mathbf{l}_{G1} and \mathbf{l}_{G2} contained non-genetic random effects of birth-litter. Vectors \mathbf{e}_{G1} and \mathbf{e}_{G2} contain the residuals for G1 and G2 following the individual animal, whereas \mathbf{e}_{LS} contains the residual for LS following the individual litter. All \mathbf{X} and \mathbf{Z} matrices are design matrices. The random effects are distributed as:

$$\begin{bmatrix} \mathbf{a}_{LS} \\ \mathbf{a}_{G1} \\ \mathbf{a}_{G2} \end{bmatrix} \sim MVN(0, \mathbf{G}_0 \otimes \mathbf{A}); \mathbf{c}_{LSp} \sim N(0, \mathbf{I}\sigma_{cLSp}^2); \mathbf{c}_{LSc} \sim N(0, \mathbf{I}\sigma_{cLSc}^2);$$

$$c_{G2} \sim N(0, \mathbf{I}\sigma_{cG2}^2); \begin{bmatrix} \mathbf{p}_{LS} \\ \mathbf{e}_{G1} \\ \mathbf{e}_{G2} \end{bmatrix} \sim MVN(0, \mathbf{R}\mathbf{P}_0 \otimes \mathbf{I}); \begin{bmatrix} \mathbf{e}_{LS} \\ \mathbf{l}_{G1} \\ \mathbf{l}_{G2} \end{bmatrix} \sim MVN(0, \mathbf{R}\mathbf{L}_0 \otimes \mathbf{I}),$$

where \mathbf{G}_0 is a 3x3 additive-genetic covariance matrix; $\mathbf{R}\mathbf{P}_0$ is a 3x3 non-genetic covariance matrix related to animals; $\mathbf{R}\mathbf{L}_0$ is a 3 x 3 non-genetic covariance matrix related to litters; \mathbf{I} is an identity matrix and \mathbf{A} is the numerator relationship matrix.

The second model ignoring non-genetic correlations between production traits and litter size is as above, but with the off-diagonal elements of the first row and column of RP_0 and RL_0 equal to zero. (Co-)variance components in the above models were estimated with REML using DMU (Madsen & Jensen, 2013)

Results

Non-genetic correlations. All non-genetic correlations between LS and the two growth traits were significantly different from zero (Table 2). The correlations between residuals for the two growth traits and the permanent non-genetic effect of the sow on LS were favourable 0.21 and 0.12 for G1 and G2, whereas the correlations between the residual for LS and the non-genetic birth litter effects were unfavourable -0.29 and -0.11 for G1 and G2.

Table 2. Non-genetic variances for residual, permanent non-genetic sow effect and birth-litter effect (correlations above diagonal).

	RP_0^1			RL_0^2		
	LS	G1	G2	LS	G1	G2
LS	1.15	0.21	0.12	11.5	-0.29	-0.11
G1	5.3	557	0.06	-12	148	0.28
G2	8.4	98	4425	-7.6	71	421

¹ RP_0 contained variances for the animal-level non-genetic variances – i.e. permanent non-genetic sow effect for litter size and residual for the two growth traits. Standard errors of correlations were within the range 0.01 to 0.02, standard errors of diagonal elements were 0.07, 5 and 37

² RL_0 contained variances for the litter-level non-genetic variances – i.e. residual for litter size and birth-litter variance for the two growth traits. Standard errors of correlations were within the range 0.01 to 0.02, standard errors of diagonal elements were 0.07, 3 and 14.

Genetic variances. The genetic variances changed significantly when the non-genetic correlations were ignored (Table 3). The genetic variance of LS increased from 1.03 to 1.30 and the genetic correlations between LS and the two growth traits weakened. The correlation between LS and G1 changed from -0.11 to -0.04, and the correlation between LS and G2 changed from -0.24 to -0.16.

Table 3. Genetic variances for models with and without non-genetic correlations between litter size and production traits (correlations above diagonal).

	With non-genetic correlations ¹			Without non-genetic correlations ¹		
	LS	G1	G2	LS	G1	G2
LS	1.03	-0.11	-0.24	1.30	-0.04	-0.16
G1	-1.8	234	0.37	-0.74	230	0.37
G2	-13	299	2787	-9.5	292	2774

¹ Standard errors of correlations were within the range 0.02 to 0.04, standard errors of diagonal elements were 0.1, 10 and 80 for LS, G1 and G2.

Discussion

Our findings supported our hypothesis that non-genetic correlations between litter size and daily gain are significant. We found significant and favourable correlations between the residuals for the two growth traits and the non-genetic permanent effect of the sow on LS. We found significant and unfavourable correlations between the residual for LS and the non-genetic birth litter effects on the two growth traits. Our findings also supported our second hypothesis that modelling these correlations changes the estimates of genetic variances compared to when they

are ignored. We found significant changes in both the genetic variance of LS and the genetic correlations between LS and the two growth traits. Clearly, multi-trait models with production traits and litter size traits should include correlations between residual effects for one trait and non-residual effects of the other trait to separate genetic correlations from non-genetic correlations.

The change in genetic variances and correlations was likely caused by both the non-genetic correlations on animal-level as well as on litter-level. Models, where only litter-level non-genetic correlations were ignored, caused inflated genetic variance for LS as well as very strong negative genetic correlations between LS and the two growth traits (results not shown). Oppositely, models, where only animal-level non-genetic correlations were ignored, did not cause inflated genetic variance for LS, but instead genetic correlations between LS and the two growth traits became closer to zero (results now shown). So, it seems that modelling both animal-level and litter-level non-genetic correlations are of importance for these traits.

The multi-trait model that we developed is novel because it handles non-genetic correlations between traits, where the levels of observation are different. We fitted our model to litter size, measured at the litter level, and daily gain, measured at the individual level. Not only is the model relevant for these traits, it may also be relevant for other traits on the same levels, for other mammalian species, where litter size is larger than one (Mínguez *et al*, 2015), and for traits on different levels. Examples of traits on different levels include non-genetic correlations between group recorded traits and individually recorded traits (Ma *et al*, 2021) and between longitudinally recorded traits and individually recorded traits (Shirali *et al*, 2017). So, our model is not only relevant to litter size and production traits in pigs, but also litter traits and individual traits in other species. It is also relevant for other scenarios, where the level of observation differs among traits analysed.

References

- Madsen, P. and Jensen, J. (2013) A user's guide to DMU, version 6, release 5.2. Center for Quantitative Genetics and Genomics Dept. of Molecular Biology and Genetics, University of Aarhus.
- Ma, X., Christensen, O.F., Gao, H., Huang, R., Nielsen, B. et al (2021). *Heredity*, 126(1):206-217. <https://doi.org/10.1038/s41437-020-0339-3>
- Mínguez, C., Sanchez, J.P., El Nagar, A.G, Ragab, M. and Baselga, M. (2015) *J. Anim. Breed. Genet.* 133 (2016) 303–315. <https://doi.org/10.1111/jbg.12197>
- Mrode, R.A. (2014) *Linear Models for the Prediction of Animal Breeding Values*. CABI, Oxfordshire, UK
- Shirali, M., Strathe, A. B., Mark, T., Nielsen, B. and Jensen, J. (2017). *Journal of animal science*, 95(3), 1050-1062. <https://doi.org/10.2527/jas.2016.0606>
- Walter, J.P. and Mao, I.L. (1985) *J Dairy Sci* 68:91-98. [https://doi.org/10.3168/jds.S0022-0302\(85\)80802-X](https://doi.org/10.3168/jds.S0022-0302(85)80802-X)