

## **Bayesian estimation of direct and correlated response to genetic selection on linear or ratio expressions of feed efficiency in pigs**

*M. Shirali<sup>1</sup>, P.F.Varley<sup>2</sup> & J. Jensen<sup>1</sup>*

*<sup>1</sup>Center for Quantitative Genetics and Genomics, Department of Molecular Biology and Genetics, Aarhus University, DK-8830 Tjele, Denmark  
[Mahmoud.Shirali@mbg.au.dk](mailto:Mahmoud.Shirali@mbg.au.dk) (Corresponding Author)*

*<sup>2</sup>Hermitage Genetics, Kilkenny, Ireland*

### **Summary**

In this study a Bayesian method for estimating genetic parameters for linear (residual feed intake, RFI) and ratio trait (feed conversion ratio, FCR) of feed efficiency in farm animals is presented. The Bayesian method integrates over all unknown model parameters including “fixed” and random effects and properly handles ratio traits that do not have standard distributions. A Bayesian procedure for the analysis of response to selection on linear versus ratio traits was developed and applied in pigs to examine different selection criteria for feed efficiency. The Bayesian methodology allowed prediction of breeding values for ratio and linear definitions of feed efficiency from the same multi-variate model for the traits measured, without the need for approximations to handle ratio traits. RFI was defined directly in the model without the need for a two-step estimation procedure. Two different definitions of RFI, differing in whether genetic or phenotypic correlations between RFI and production traits are constrained to be zero, were extended to Bayesian analysis. Due to unknown distributions the genetic parameters of FCR cannot be directly estimated. Instead we developed the posterior multivariate distribution of additive genetic (co)variance available for selection. The example shows that inference based on this measure is very similar to the estimates of additive genetic (co)variance and, therefore, can be used as well to investigate the posterior distribution of additive genetic variance in the ratio trait FCR. Finally we explored the posterior distribution of the genetic superiority of the selected group when selecting on different definitions of feed efficiency or on production traits. The current study shows that direct selection against FCR results in unexpected selection pressure on its component traits and also on lean meat percentage (LMP). However, direct selection for low genetic RFI allows for selection on the proportion of average daily feed intake (ADFI) that is independent of production. Selection for improved feed efficiency is likely best achieved through multiple-trait selection on genetic RFI and production traits.

*Keywords: Bayesian analysis, feed efficiency, genetic parameters, response to selection*

### **Introduction**

Classically feed efficiency has been defined as output over input such as milk or milk component yield over dry matter consumption in dairy industry. In breeding programs, the inverse of feed efficiency is mainly used, which is known as FCR; for example in growing animals it is normally defined as average daily feed intake (ADFI) over average daily body weight gain (ADG). Gunsett (1984) reported that ratio traits have a distribution deviating from normality if their component traits have normal distributions. Furthermore, means and

(co)variances of their component traits influence the distribution of the ratio between the traits. This is valid for FCR which is usually assumed as ratio of ADFI and ADG.

To circumvent the problems of ratio traits, RFI has been proposed by Koch *et al.* (1963) as a better measure to determine animal's feed efficiency. RFI is a partial component of feed efficiency indicating/measuring the proportion of feed intake that is independent of performance traits and maintenance. If the phenotypic and genetic (co)variances are known for component traits of RFI (e.g. ADFI and ADG) then phenotypic and genetic RFI can be computed using partial phenotypic or genetic regression coefficients of ADFI on production traits following Kennedy *et al.* (1984).

The Bayesian methodology as illustrated by Sorensen *et al.* (1994) provides the marginal posterior distributions for any parameter in the model given the data. Inferences about breeding values are made using the marginal posterior distribution of the vector of breeding values. If non-informative priors are used these distributions take into account the fact that other parameters such as the variance components are being inferred from the data so that proper probability statements on the genetic response can be made. Bayesian methods, therefore, ensures that uncertainties about fixed effects and variance components are taken into account. Marginal posterior distribution of response to selection or genetic superiority of a selected group can be obtained as averaging predicted breeding values obtained by mixed model techniques shown by Sorensen *et al.* (1994).

The aims of this study were 1) to derive methods for Bayesian prediction of breeding values for both phenotypic and genetic RFI as well as for FCR without invoking unrealistic distributional assumptions for FCR; 2) estimate genetic parameters for production traits of growth, feed intake, and lean meat production along with derived traits of RFI and FCR; 3) to derive Bayesian estimates of direct and correlated responses to selection on feed efficiency measured either as RFI or FCR.

## Material and methods

### Data

In this study, 3027 MaxGro pigs (2621 boars, 406 gilts) from Hermitage Genetics (Kilkenny, Ireland) were used which went through feed intake recording from start body weight (BW) of 52 kg (11, standard deviation (SD)) until end BW of 110 kg (11, SD). The ADFI was calculated as total feed intake divided by number of days in the entire test period. The ADG was calculated as total body weight gain in the test period divided by number of days on test. Lean meat percentage (LMP) was predicted using a transformation of fat layer and muscle depths between 3<sup>rd</sup> and 4<sup>th</sup> last ribs from ultrasound pictures taken at the end of the test period. Pedigree information was available for at least 4 generations back, comprising a total of 6237 animals.

### Statistical model

A trivariate analysis was used for ADFI, ADG and LMP traits using the following models:

- (1)
- (2)
- (3)

where , and are vectors of phenotypic records for ADFI, ADG and LMP, respectively; vectors , and contain "fixed" effects of year-quarter, gender and parity for traits ADFI, ADG

and LMP, respectively; and are “fixed” regressions on start BW for ADFI and ADG, respectively; is “fixed” regressions on end BW for LMP; , and , and , and are vectors of animal random additive genetic and pen effects for ADFI, ADG, and LMP, respectively; and , , and are vectors of residuals for ADFI, ADG and LMP, respectively. Matrices are design matrices for year-quarter, gender, and parity effects; is a vector of start body weights for each animal and is a vector of end body weights. and are the corresponding design matrices for additive genetic animal effects ( , , and ) and permanent environmental effect of pen ( , , and ) for the three traits, respectively. Prior distributions for all random vectors in the models were multivariate normal distributions with mean zero and: , where is a  $3 \times 3$  matrix of residual (co)variances, , where is the additive genetic relationship matrix and is a  $3 \times 3$  matrix of additive genetic (co)variances; and where is a  $3 \times 3$  matrix of pen (co)variances. The prior distributions for all covariance matrices ( , , and ) were inverse Wishart distributions with prior degree of belief corresponding to flat priors. Similarly, the prior probability distributions for all “fixed” location parameters were taken as flat priors.

A Bayesian estimation method using Gibbs sampling was used to obtain posterior distributions for all parameters included in the trivariate analysis combining models 1 to 3, including the matrices of variances and covariances. The Gibbs sampler was run for 1.1 M rounds with first 100 K rounds were considered as burn-in and after the burn-in every 250th sample was saved for posterior analysis. The RJMC module in the DMU software package of Madsen & Jensen (2014) was used for the analysis.

### Posterior distribution of different definitions of RFI

RFI was defined in two different ways: 1) phenotypic RFI ( $RFI_p$ ) by using phenotypic partial regression coefficients; and 2) genetic RFI ( $RFI_g$ ) conditioning breeding values of ADFI by breeding values for ADG and LMP using genetic partial regression coefficients. For calculation the partial regression coefficients ( ) for ADG and LMP were computed from the genetic (co)variance matrix while for the partial phenotypic regression coefficients ( ) was used obtain from the phenotypic (co)variance matrix. Within a posterior sample both types of RFI involve conditional normal distributions and the derivations are, therefore, straightforward, with the phenotypic and the genetic (co)variance matrices which are subdivided in and , where diagonal of matrices are variances and the off-diagonals are covariances.

The Bayesian estimations of partial phenotypic ( ) and genetic ( ) regressions coefficients were obtained as:

which are  $2 \times 1$  vector-valued functions that are obtained in each sample from the Gibbs output. The and are  $2 \times 2$  matrices of phenotypic and genetic (co)variance matrices for production traits of ADG and LMP from and , respectively. The contain phenotypic and genetic covariances between production traits ADG and LMP with ADFI.

A sample from the posterior distribution of breeding values for phenotypic ( ) and genetic ( ) RFI is:

The breeding values for RFI estimates can be obtained simultaneously for all animals by conditional distribution of the breeding values for ADFI ( ) given breeding values of ADG ( ) and LMP ( ) by either using phenotypic ( ) or genetic ( ) partial regression coefficients.

The distributions of different RFI definitions were obtained as the distribution of

ADFI conditional on all other model parameters and on ADG and LMP. The corresponding variances and covariances can be obtained using following equations: (7) (8)

where and are genetic or phenotypic (co)variances, respectively. with and are the phenotypic partial regression coefficients from and and are the genetic regression coefficients from for ADG and LMP, respectively.

### Posterior distribution of FCR

The FCR is a ratio between two normally distributed and usually correlated traits, and therefore has a distribution which depends on the means of the two traits involved in the ratio as well as their (co)variance. Therefore, the breeding value for FCR is dependent on the “fixed” location parameters since it depends on the mean of ADFI ( $\mu_{ADFI}$ ) and ADG ( $\mu_{ADG}$ ). Following Gunsett (1984), the breeding value for FCR ( $\mu_{FCR}$ ) can be calculated from underlying parameters using the following equation for a given sample:

where estimate of  $\mu_{ADFI}$  can be obtained from the model for ADFI (1) as the sum of average of each “fixed” effects, year-quarter, gender, and parity, in addition to the population average start BW. Similarly, the estimate of  $\mu_{ADG}$  can be obtained. In this way inaccuracy of computing the mean is taken into account when deriving the posterior distribution of breeding values etc. for FCR.

### Genetic (co)variance available for selection

Since genetic selection is usually carried out in the same age group the amount of genetic (co)variance available for selection at a given time point is:

where  $\Sigma_{G|T}$  is the distribution of genetic (co)variance available for selection after integrating over the genetic trend ( $T$ ), and  $\mathbf{A}$  is an incidence matrix relating the additive genetic values ( $G$ ) of individuals to yearly batches. This derivation is the extension of Sorensen et al. (2001) to a multivariate setting.

The superiority of selected group of animals is defined as the mean of a selected group of animals after truncation selection on breeding values conditional on the genetic trend. The Bayesian measure of superiority of a selected group is estimated as the difference between mean of breeding values in the selected population compared to the population corrected for the genetic trend. This will yield an expression of the superiority of the selected group in every sample from the posterior distribution depending on the selection rule. Three scenarios were developed to compare the effect of selection for different definitions of feed efficiency with each other. The number of individuals ranked for analysis was decided based on truncation selection percentage, ranging from the top 5 to the top 30% of animals, but here only results of truncation selection of selecting the top 10% are presented since the results and conclusion were consistent across different truncation selection percentages.

## Results and discussion

The Bayesian estimation of responses to selection for feed efficiency traits are presented in Table 1 and the genetic correlation among feed efficiency and production traits are presented in Table 2. Bayesian analysis suggests that direct selection for low genetic RFI does not have correlated response on production traits in the breeding program, since the model ensured no genetic correlation between these traits. The existence of correlated responses on production traits from direct selection for low phenotypic RFI is due to existence of genetic correlation between these traits which are caused by using phenotypic partial regression coefficients that ensures the phenotypic correlation between phenotypic RFI and production traits to be zero. Selection for low ratio trait FCR results in unexpected selection pressure on feed intake and production traits (e.g. LMP) in the breeding program. This unexpected selection pressure can be explained by genetic correlations among ADFI, ADG, and LMP. A large reduction in ADFI, which is the numerator component of FCR, may be due to the large positive posterior mean of the genetic correlation between ADFI and ADG. The low posterior mean of the genetic correlation between FCR and ADG indicates a lower change in ADG due to selection for FCR, while a large negative posterior mean of the genetic correlation between FCR and LMP explains the large indirect genetic response on LMP from selection for low FCR. The substantial reduction in ADFI through direct selection on FCR could also be due to the correlated response on LMP, as increased leanness is one of the underlying biological reasons for improved feed efficiency. Therefore, selection for FCR is not an efficient strategy as the improvement in this trait can be due to improvement in leanness rather than improvement in efficiency of nutrient utilization per se. Gunsett (1984) also showed the unexpected selection pressure on component traits of ratio traits. Therefore, truncation point selection as used for normally distributed variables cannot be used.

*Table 1. Posterior means of direct (bold numbers) and correlated (un-bold numbers in a row) additive genetic superiority of the selected group, when the top 10% population were selected for single traits of feed efficiency, are presented in each row along with their posterior standard deviation in parentheses.*

Scenario <sup>1</sup>	Genetic superiority of the selected group					
	, kg/d	, kg/d	FCR, kg/kg	ADFI, kg/d	ADG, kg/d	LMP, %
	<b>-0.161</b>	-0.158	-0.155	-0.151	0.004	0.015
	(0.013)	(0.013)	(0.013)	(0.017)	(0.004)	(0.100)
	-0.149	<b>-0.170</b>	-0.147	-0.233	-0.035	0.160
	(0.016)	(0.011)	(0.014)	(0.020)	(0.010)	(0.161)
FCR	-0.144	-0.145	<b>-0.174</b>	-0.176	0.002	0.978
	(0.015)	(0.013)	(0.013)	(0.025)	(0.011)	(0.183)

<sup>1</sup> = residual feed intake estimated from partial genetic coefficients; = residual feed intake estimated from partial phenotypic coefficients; FCR = feed conversion ratio; ADFI = average daily feed intake during the test period; ADG = average daily gain during the test period; LMP = lean meat percentage at end of the test

*Table 2. Posterior means of genetic correlations and standard deviation in parentheses.*

Traits <sup>1</sup>	ADFI	ADG	LMP
	0.92 (0.04)		
ADFI	0.51 (0.05)	0.77 (0.04)	
ADG	0.00 (0.00)	0.35 (0.10)	0.82 (0.04)
LMP	0.00 (0.00)	-0.06 (0.08)	-0.39 (0.08)    -0.17 (0.10)

FCR	0.89 (0.04)	0.82 (0.03)	0.52 (0.08)	-0.07 (0.10)	-0.40 (0.08)
-----	-------------	-------------	-------------	--------------	--------------

---

<sup>1</sup> = residual feed intake estimated from partial genetic coefficients; = residual feed intake estimated from partial phenotypic coefficients; FCR = feed conversion ratio; ADFI = average daily feed intake during the test period; ADG = average daily gain during the test period; LMP = lean meat percentage at end of the test

## List of References

- Gunsett, F. C., 1984. Linear Index Selection to Improve Traits Defined as Ratios. *J. Anim. Sci.* 59:1185-1193.
- Koch, R. M., L. A. Swiger, D. Chambers & K. E. Gregory, 1963. Efficiency of Feed Use in Beef Cattle. *Journal of Animal Science* 22:486-494.
- Kennedy, B. W., J. H. van der Werf & T. H. Meuwissen, 1993. Genetic and statistical properties of residual feed intake. *Journal of Animal Science* 71:3239-3250.
- Shirali M., A. B. Strathe, T. Mark, B. Nielsen & J. Jensen, 2017. Joint analysis of longitudinal feed intake and single recorded production traits in pigs using a novel Horizontal model. *J. Anim. Sci.* 95:1–14.
- Sorensen, D., C.S. Wang, J. Jensen & D. Gianola, 1994. Bayesian analysis of genetic change due to selection using Gibbs sampling. *Genet. Sel. Evol.* 26:333-360.
- Sorensen D., R. Fernando & D. Gianola, 2001. Inferring the trajectory of genetic variance in the course of artificial selection. *Genet. Res.* 77(1):83-94.
- Madsen, P. & J. Jensen, 2014. An User's guide to DMU: a package for analysing multivariate mixed models. Version 6, release 5.2. Aarhus University, Foulum, Denmark.