Use of threshold and linear models to estimate variance components and breeding values for disease resistance in Italian heavy pigs

M. Cappelloni\textsuperscript{a,} \textsuperscript{b}, M. Gallo\textsuperscript{a} and A Cesaranib,c

\textsuperscript{a}Associazione Nazionale Allevatori Suini (ANAS), Roma, Italy; \textsuperscript{b}Dipartimento di Agraria, Università degli Studi di Sassari, Sassari, Italy; \textsuperscript{c}Animal and Dairy Science Department, University of Georgia, Athens, GA, USA

ABSTRACT
The Italian National Pig Breeders Association (ANAS) manages the breeding programs of the Italian Large White (ILW), Landrace (IL), and Duroc (ID) breeds, mainly oriented to the production of PDO hams. ANAS evaluates the inclusion of genetic resistance in its breeding scheme. This study aimed to estimate variance components and breeding values (EBV) using threshold (TM) and linear (LM) models. During the sib test performed at the genetic station of ANAS from 1997 to 2021, 9,595 (respiratory diseases) and 12,046 (enteritis) diagnoses were collected by the veterinarians. The trait was recorded as a dichotomous variable: affected animals with 1, whereas healthy with 2. A multi-breed model was applied with the breed, sex, and farm sector as fixed effects and litter, animal, and residual as random effects. The same model was also applied within each breed, removing the breed effect. The limited data size within single breed did not allow to estimate accurate variance components. With the multi-breed model, low heritabilities were estimated for respiratory (0.09) and enteritis (0.15), with TM and LM leading to the same values. The multibreed model led to more precise variance components estimation and to EBV very close to the ones from the single breed analyses. Pearson and Spearman rank correlations between EBV estimated with TM and LM in the multibreed scenario were $\geq0.97$. Results of this study demonstrated the feasibility of including disease resistance among the breeding goals of ANAS and that both TM and LM could be used to reach this objective.

HIGHLIGHTS
- Genetic selection for disease resistance in heavy pigs is possible
- Low heritabilities estimated for enteritis and respiratory disorders
- Threshold and linear models gave similar results

Introduction
The Italian National Pig Breeders Association (ANAS) is the official Italian Pig Breed Association. It runs several authorised pure breeding programs in Italy that have changed the genome of some Italian pig breeds (Schiavo et al. 2016). Italian swine herdbook started in 1962, but in 1988 ANAS, upon request of the Ministry of Agriculture, implemented a new breeding scheme. The first goal of this project was to improve the production efficiency of typical Italian hams, which requires heavy pigs (slaughtered at 160–170 kg live weight at a minimum age of nine months), with carcasses and meat particularly suitable for curing.

ANAS manages the breeding programs of the Italian Large White (ILW), Landrace (IL), and Duroc (ID) pig breeds, which are the reference breeds for the main Italian Protected Denomination of Origin (PDO). ANAS breeding program for heavy pigs is based on the sibling genetic evaluation (i.e. Sib Test) in the ANAS testing station, where for each candidate young boar, a group of three full sibs (both castrated male and females) is raised with a “quasi ad libitum” feeding system (Kielanowski 1965). The health status of all animals entering the testing station of ANAS is monitored to assess the resistance to diseases and environmental stress in intensive farms.

Currently, ANAS participates in the SUIS 2.0 project – “Italian Sustainable Pig Farm”. Among the objectives of the project, ANAS is committed to reducing morbidity, mortality, and losses, to improving the health of the animals, and to reducing the use of antimicrobials. Gol et al. (2015) pointed out that respiratory...
disorders (RD) and enteric diseases are causing substantial economic losses in the pig industry worldwide. RD represent one of the major concerns for swine producers, also because it can increase the susceptibility to other diseases (Okamura et al. 2012). Thus, the aim of the present study was to evaluate the possibility of developing a new genetic index for disease resistance in pigs, both by threshold and linear models applied to single and multi-breed analyses. Disease resistance was registered as dichotomous traits, i.e. a trait with just two levels (affected or unaffected). The use of threshold models with dichotomous traits, especially with just two levels as in this case, is theoretically suggested because they better account for the probabilities structure of categorical data (Gianola 1982).

Materials and methods

During the Sib Tests carried out from 1997 to 2021, 21,641 records of health status as binary traits (1 for affected and 2 for healthy animals) were collected by the ANAS veterinary, who visited the testing station every day. The phenotypes recording was based on the experience of the veterinary, who used modular, industrial handheld instruments (different software updates of PSION Workabout across the years) to record the animal, the type of disease, and the treatment given. These subjectively collected data are transferred to ANAS databases. This study focussed on two problems: respiratory disease, inflammations of any section of the respiratory system, and enteritis, inflammations of the gastrointestinal tract mainly causing diarrhoea.

Table 1 shows the size of the analysed dataset. For respiratory diseases (N = 9,595), 52% of observations were from ILW, 29% from ID, and 19% from IL; for enteritis (N = 12,046), 69% of observations were from ILW, 33% from ID, and 24% from IL. On average, 36% suffered respiratory diseases, and 35% had enteric disorders.

**Table 1.** Size of the investigated datasets.

|          | Respiratory | Enteritis |
|----------|-------------|-----------|
|          | Healthy     | Affected  | Total | Healthy | Affected | Total |
| Large white | 3,187      | 1,839    | 5,026 | 4,707   | 1,885    | 6,592 |
| Landrace  | 996        | 801      | 1,797 | 1,606   | 703      | 2,309 |
| Duroc     | 1,937      | 835      | 2,722 | 1,576   | 1,569    | 3,145 |
| Total     | 6,120      | 3,475    | 9,595 | 7,889   | 4,157    | 12,046 |

To estimate variance components, the following multi-breed animal model was used for both threshold and linear analyses:

\[ y_{ijk\text{on}} = BRD_i + SEX_k + CG_j + LIT_o + a_n + e_m \]  

(1)

where: \( y_{ijk\text{on}} \) was the analysed phenotype (i.e. respiratory and enteritis diseases); BRD was the fixed effect of \( i \)th breed (3 levels); SEX was the fixed effect of the \( k \)th sex of the animal (2 levels); CG was the fixed effect of the \( j \)th contemporary group (i.e. combination of year, month, farm of origin, and pen); LIT was the random effect of the \( n \)th litter distributed as N(0, \( \sigma_{LIT}^2 \)) where \( \sigma_{LIT}^2 \) was the variance of the effect, and \( I \) was the identity matrix; \( a_n \) was the random effect of the \( n \)th animal, distributed as N(0, \( \sigma_a^2 \)) where \( \sigma_a^2 \) was the additive genetic variance, and \( A \) was the pedigree relationship matrix, which was built tracking back the pedigree for three generations (total of 25,692 animals for enteritis and 21,505 for respiratory, respectively); \( e_m \) was the random residual distributed as N(0, \( \sigma_e^2 \)) where \( \sigma_e^2 \) is the residual variance. The same model (1) was repeated within each breed by removing the fixed effect of the breed.

The heritability (\( h^2 \)) of the trait was computed using the following equation:

\[ h^2 = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_{LIT}^2 + \sigma_e^2} \]

with the symbols having the same meaning above-mentioned. In the TM, heritability on the liability scale was transformed in the observed scale (i.e. 0–1) using the equation proposed by Dempster and Lerner (1950) and by Wray and Visscher (2015): heritability on the observed scale is a function of heritability on the liability scale and the incidence of the diseases.

Variance components were estimated via a Gibbs sampling method with 100,000 rounds, 10,000 iterations discarded as burn-in, thin-in of 10 (i.e. saving 1 sample every 10 iterations). The softwares used were `thrgibbs1f90` and `gibbs2f90` for TM and LM models, respectively; posterior means (including confidence interval at 95%, CI95%) for all the parameters were calculated using `postgibbsf90`. All routines are implemented in the BLUPF90 suite (Miszta et al. 2014). These variance components were used to estimate the breeding values within each model. EBVs from multi and single-breed analyses were compared to assess differences between these two scenarios. In the multi-breed analysis, EBV from TM and LM of all animals included in the analyses were compared using Pearson and Spearman rank correlations. Moreover,
the EBV from the multibreed were then standardised within a breed and compared between healthy and affected. The differences between the two groups were tested using the honestly significant difference test (Tukey HSD), and significance was declared for $P < .05$.

**Results and discussion**

Binary traits, such as diseases, can be analysed using linear and threshold models. Several studies investigated their differences (e.g. Varona et al. 1999; Casellas et al. 2007; Silvestre et al. 2007). Pioneer studies involving simulated datasets demonstrated the advantages of using threshold over linear models (Meijering and Gianola, 1985; Hoeschele, 1988). Thus, in the present study, results about variance components and EBV were compared between the two models to point out any differences in choosing one of them.

Low heritabilities were estimated for both investigated diseases (Table 2). Overall, lower estimates were found for respiratory disease compared to enteritis. The limited size of the datasets recorded within breed led to larger standard deviations than the multi-breed analysis. In particular, the CI95% of the heritability estimated for respiratory disorders, with both LM and TM, in Duroc and Landrace included the zero. Also, the CI95% of $h^2$ estimated for enteritis in Landrace included zero. In the single breed analysis, estimates from TM (transformed in the observed scale 0–1 to be comparable with linear models) were larger (with also larger standard deviations) than those from LM.

As far as the multi-breed analysis, the same values were observed within disease for TM and LM. The multi-breed analysis was adopted to consider a larger number of observations and to have breeding values estimated for all breeds in a single run. Moreover, since the three investigated breeds are raised together in the same testing station, but they are closed populations not connected by pedigree, the use of a multi-breed model with the breed as a fixed effect allowed more precise estimates. Table 3 shows the comparison between BV estimated using the LM in single and multi-breed analyses. For the abovementioned reasons (i.e. breeds not connected by pedigree and use of the breed as a fixed effect), EBVs from single and multi-breed analyses were the same (Table 3).

Heritabilities estimated in the present study are of the same magnitude as those reported in the literature for diseases in pigs (e.g. Henryon et al. 2001) and other livestock species (Snowder et al. 2006; Schneider et al. 2010). Henryon et al. (2001) reported heritabilities of 0.12 and 0.16 for respiratory diseases and diarrhoea. Heritability of 0.29 was estimated for atrophic rhinitis (i.e. one respiratory disorder) by Okamura et al. (2012) in Landrace pigs. As far as the other livestock species were concerned, Schneider et al. (2010) estimated the genetic parameters for incidence of bovine respiratory disease, and they reported estimates of $0.10 \pm 0.11$ and $0.02 \pm 0.06$ in Angus and Simmental preweaned calves, respectively. Snowder et al. (2006) reported heritability estimates for resistance to RD in 9 cattle breeds ranging from 0.04 to 0.08 $\pm 0.01$. Moreover, the heritabilities estimated in the present study were similar to those reported in other livestock species for functional traits, usually lower than production traits (Sewalem et al. 2011; Martin et al. 2018; Cesaran et al. 2020). As far as the comparison between TM and LN models was concerned, Silvestre et al. (2007) analysed the hip dysplasia genetic parameters in Estrela Mountain Dog comparing these two models. These authors reported indistinguishable heritability, repeatability, and genetic trends between the

**Table 2.** Heritability was estimated using linear and threshold (transformed in the observed scale 0–1) models for respiratory and enteritis in both single and multi-breed analysis.

| Disease  | Breed    | Linear model | Threshold model |
|---------|----------|--------------|-----------------|
|         |          | $h^2$ (SD) CI95% | $h^2$ (SD) CI95% |
| Respiratory | Duroc    | 0.04 (0.05) 0.00–0.16 | 0.09 (0.07) 0.00–0.20 |
|          | Large white | 0.12 (0.04) 0.06–0.20 | 0.14 (0.05) 0.06–0.23 |
|          | Landrace  | 0.07 (0.06) 0.00–0.19 | 0.13 (0.12) 0.00–0.42 |
|          | Multibreed | 0.09 (0.03) 0.02–0.15 | 0.09 (0.04) 0.03–0.16 |
| Enteritis | Duroc    | 0.22 (0.06) 0.10–0.34 | 0.24 (0.07) 0.10–0.42 |
|          | Large white | 0.16 (0.05) 0.08–0.25 | 0.15 (0.04) 0.08–0.23 |
|          | Landrace  | 0.08 (0.06) 0.01–0.21 | 0.10 (0.09) 0.00–0.27 |
|          | Multibreed | 0.15 (0.03) 0.09–0.22 | 0.15 (0.03) 0.10–0.20 |

$h^2$: heritability; SD: standard deviation; CI95%: confidence interval at 95%.

**Table 3.** Comparison between breeding values estimated using the linear model in single and multi-breed analyses.

| Breed | Scenario | Enteritis | Respiratory |
|-------|----------|----------|-------------|
|       |          | Min Mean ± SD Max | Correlation | Min Mean ± SD Max | Correlation |
| Duroc | Single-breed | $-0.33$ $0.03 \pm 0.10$ 0.45 | 0.99 | $-0.05$ $0.01 \pm 0.02$ 0.08 | 0.99 |
|       | Multi-breed | $-0.25$ $0.03 \pm 0.08$ 0.39 | 0.99 | $-0.11$ $0.01 \pm 0.04$ 0.15 | 0.99 |
| Large white | Single-breed | $-0.28$ $0.01 \pm 0.07$ 0.26 | 1.00 | $-0.20$ $0.00 \pm 0.05$ 0.19 | 1.00 |
|       | Multi-breed | $-0.28$ $0.01 \pm 0.06$ 0.25 | 1.00 | $-0.16$ $0.00 \pm 0.04$ 0.14 | 1.00 |
| Landrace | Single-breed | $-0.13$ $0.01 \pm 0.04$ 0.13 | 0.99 | $-0.08$ $0.01 \pm 0.03$ 0.14 | 0.99 |
|       | Multi-breed | $-0.22$ $0.01 \pm 0.06$ 0.20 | 0.99 | $-0.10$ $0.01 \pm 0.04$ 0.17 | 0.99 |
two models. Carlén et al. (2006) studied which model – among survival, threshold, and linear – would result in a more precise genetic evaluation for mastitis in dairy cattle. They estimated greater heritabilities using TM compared to LM. However, when transformed to the liability scale, the h² estimated using the LM was close to the one estimated using the TM.

The EBV estimated using threshold and linear models in the multi-breed analysis were compared to identify potential differences in using one of them. Pearson and Spearman rank correlations between EBV of all animals considered in the analysis in the two models were as large as 0.97 (Table 4). Thus, the use of both models would lead to almost the same EBVs and rank of the animals. Results about the comparison of breeding values estimated using TM or LM models are discordant. Previous studies already reported very high correlations between EBV estimated using these two models (Weller and Ron 1992; Boettcher et al. 1999). Silvestre et al. (2007) found important differences between the ranking lists in the two models, and they suggested the use of TM for the genetic evaluation of hip dysplasia in dogs. On the contrary, Carlén et al. (2006) reported a correlation close to the unity between EBV estimated with TM and LM model.

EBV estimated using the multi-breed analysis were standardised within each breed and compared between healthy and affected animals (Table 5). As expected, because 1 was assigned to the affected animals and 2 to the healthy ones, the average EBV for the affected animals was lower (P < .01) than healthy animals.

Conclusions

Results of the present study demonstrated the possibility of including disease resistance among the breeding goals of ANAS. It has been shown that multi-breed analysis allows estimating more precise heritabilities without changing the breeding values or the rank of the animals. TM and LM can be used indifferently within the multi-breed analysis to estimate variance components and breeding values for disease resistance in Italian heavy pigs. Thus, since LM is easier and faster to be implemented and the heritability estimated using this model is already in the observed scale, we suggest using the multibreed linear model to implement disease resistance among the breeding goals of the Italian pigs breeding scheme.

Acknowledgments

The authors would like to acknowledge the Italian National Pig Breeders Association (ANAS) for providing the datasets used under agreement for this study.

Ethical approval

Animal Care and Ethical approvals were not needed as data were obtained from pre-existing databases.

Disclosure statement

The authors declare no conflicts of interest.

Funding

This work was financially supported by the project SUIS 2.0 (co-funded by the European Agricultural Fund for Rural Development of the European Union and by the Italian Ministry of Agriculture, Food and Forestry—MiPAAF).

ORCID

M. Cappelloni http://orcid.org/0000-0001-7742-8644
A Cesarani http://orcid.org/0000-0003-4637-8669

Data availability statement

Data is subject to third-party restrictions. The data supporting this study belong to ANAS, and restrictions apply to the
availability of these data, which were used under agreement for this study.

References

Boettcher PJ, Jairath LK, Dekkers JCM. 1999. Comparison of methods for genetic evaluation of sires for survival of their daughters in the first three lactations. J Dairy Sci. 82(5):1034–1044.

Carlén E, Emanuelson U, Strandberg E. 2006. Genetic evaluation of mastitis in dairy cattle using linear models, threshold models, and survival analysis: a simulation study. J Dairy Sci. 89(10):4049–4057.

Casellas J, Caja G, Ferret A, Piedrafita J. 2007. Analysis of litter size and days to lambing in the Ripollesa ewe. I. Comparison of models with linear and threshold approaches. J Anim Sci. 85(3):618–624.

Cesarani A, Gaspa G, Masuda Y, Degano L, Vicario D, Lourenco DA, Macciotta NPP. 2020. Variance components using genomic information for 2 functional traits in Italian Simmental cattle: calving interval and lactation persistency. J Dairy Sci. 103(6):5227–5233.

Dempster ER, Lerner IM. 1950. Heritability of threshold characters. Genetics. 35(2):212–236.

Gianola D. 1982. Theory and analysis of threshold characters. J Anim Sci. 54(5):1079–1096.

Gol S, Estany J, Fraile LJ, Pena RN. 2015. Expression profiling of the GBP1 gene as a candidate gene for porcine reproductive and respiratory syndrome resistance. Anim Genet. 46(6):599–606.

Henryon M, Berg P, Jensen J, Andersen S. 2001. Genetic variation for resistance to clinical and subclinical diseases exists in growing pigs. Anim Sci. 73(3):375–387.

Hoeschele I. 1988. Comparison of “maximum a posteriori estimation” and “quasi best linear unbiased estimation” with threshold characters. J Anim Breed Genet. 105: 337–361.

Kielanowski J. 1965. Estimates of the energy cost of protein deposition in growing animals. Energy Metabolism 3: 13–20.

Martin P, Barkema HW, Brito LF, Narayana SG, Miglior F. 2018. Symposium review: novel strategies to genetically improve mastitis resistance in dairy cattle. J Dairy Sci. 101(3):2724–2736.

Meijering A, Gianola, D. 1985. Linear versus nonlinear methods of sire evaluation for categorical traits: a simulation study. Genet Sel Evol. 17(1):115s132.

Misztal I, Tsuruta S, Lourenco D, Aguilar I, Legarra A, Vitezica Z. 2014. Manual for BLUPF90 family of programs. Retrieved from http://nce.ads.uga.edu/wiki/lib/exe/fetch.php?media=blupf.90_all7.pdf

Okamura T, Onodera W, Tayama T, Kadowaki H, Kojima-Shibata C, Suzuki E, Uemoto Y, Mikawa S, Hayashi T, Awata T, et al. 2012. A genome-wide scan for quantitative trait loci affecting respiratory disease and immune capacity in Landrace pigs. Anim Genet. 43(6):721–729.

Schiavo G, Galimberti G, Calò DG, Samoré AB, Bertolini F, Russo V, Gallo M, Buttazzoni L, Fontanesi L. 2016. Twenty years of artificial directional selection have shaped the genome of the Italian Large White pig breed. Anim Genet. 47(2):181–191.

Schneider MJ, Tait RG, Jr Ruble MV, Busby WD, Reecy JM. 2010. Evaluation of fixed sources of variation and estimation of genetic parameters for incidence of bovine respiratory disease in preweaned calves and feedlot cattle. J Anim Sci. 88(4):1220–1228.

Sewalem A, Miglior F, Kistemaker GJ. 2011. Genetic parameters of milking temperament and milking speed in Canadian Holsteins. J Dairy Sci. 94(1):512–516.

Silvestre AM, Ginja MMD, Ferreira AJA, Colaco J. 2007. Comparison of estimates of hip dysplasia genetic parameters in Estrela Mountain Dog using linear and threshold models. J Anim Sci. 85(8):1880–1884.

Snowder GD, Van Vleck LD, Cundiff LV, Bennett GL. 2006. Bovine respiratory disease in feedlot cattle: environmental, genetic, and economic factors. J Anim Sci. 84(8):1999–2008.

Varona L, Misztal I, Bertrand JK. 1999. Threshold-linear versus linear-linear analysis of birth weight and calving ease using an animal model: II. Comparison of models. J Anim Sci. 77(8):2003–2007.

Weller JI, Ron M. 1992. Genetic analysis of fertility traits in Israeli Holsteins by linear and threshold models. J Dairy Sci. 75(9):2541–2548.

Wray NR, Visscher PM. 2015. Quantitative genetics of disease traits. J Anim Breed Genet. 132(2):198–203.