

Estimation of hip dysplasia heritability using a Bayesian animal model in Rottweiler and Golden Retriever dogs registered in the Kennel Club of Perú (2014-2024)

María José Pastor-Talavera¹ and Jorge Luis Vilela-Velarde^{*}

¹Facultad de Medicina Veterinaria y Zootecnia, Universidad Científica del Sur, Lima, Perú.

***Corresponding author**

jvilela@cientifica.edu.pe

Received: 01.05.2026

Accepted: 14.09.2026

Published: 01.10.2026

ABSTRACT. Canine hip dysplasia (CHD) is a prevalent hereditary orthopedic disorder in which phenotypic selection often fails to identify silent genetic carriers, leading to stagnant genetic trends. This study aimed to obtain heritability estimates (h^2) of CHD in Rottweiler and Golden Retriever populations in Peru to evaluate the potential for genetic improvement within the national registry. A Bayesian threshold animal model was applied to the pedigree and radiographic records of 383 dogs (146 Rottweilers and 237 Golden Retrievers) evaluated between 2014 and 2024 by the Kennel Club del Perú. The h^2 (mean $\pm$ standard deviation) for CHD status (binary trait) was 0.60 ± 0.15 for Rottweilers and 0.65 ± 0.14 for Golden Retrievers. These high h^2 values indicate that a substantial proportion of phenotypic variation is driven by additive genetic effects, exceeding estimates reported in several international populations. However, whether the observed lack of phenotypic progress reflects limitations of mass selection or other factors remains to be determined through formal genetic trend analyses. Future studies should evaluate whether implementing selection programs based on estimated breeding values could accelerate the reduction in CHD prevalence and improve the long-term welfare of these breeds in Peru.

Keywords: Phenotypic selection; Estimated breeding values; Hereditary orthopedic disorders; Genetic parameters

INTRODUCTION

Canine hip dysplasia (CHD) is one of the most prevalent hereditary orthopedic disorders affecting medium- and large-breed dogs worldwide (Pinto Coelho Ribeiro Jardim *et al.*, 2025; Jacobsen *et al.*, 2024). The disease is characterized by joint laxity and incongruence between the femoral head and the acetabulum during skeletal development, which may lead to progressive osteoarthritis, chronic pain, and impaired mobility. The etiology of CHD is complex and multifactorial, involving the interaction between environmental factors such as growth rate, nutrition, and physical activity and a polygenic background that contributes to susceptibility to the disorder (Santana *et al.*, 2021; Mikkola *et al.*, 2021). It is well established that CHD is a polygenic condition; therefore, the implementation of structured breeding programs aimed at reducing its incidence can be beneficial, as addressing only environmental factors may not be sufficient to control the disease (Binversie *et al.*, 2022; Lewis *et al.*, 2013; Jacobsen *et al.*, 2024).

Due to its high prevalence and the limited effectiveness of therapeutic interventions, genetic selection has been considered the most effective long-term strategy to reduce CHD occurrence and improve canine welfare (Lewis *et al.*, 2013). Breeding decisions have traditionally been based on phenotypic evaluation through radiographic screening systems, such as the classification established by the Fédération Cynologique Internationale (FCI, 2022). However, selection based solely on phenotypic character-

istics may result in slow genetic progress because the phenotype observed reflects both genetic and environmental influences, which can obscure an individual's breeding value (Soo & Worth, 2015; Lewis *et al.*, 2013). While current diagnostic and treatment approaches primarily focus on managing the disease in individual animals, genetic evaluation and population-based analyses seek to address the problem at the population level, offering a more sustainable strategy for reducing the prevalence of CHD in future generations (Lucaci *et al.*, 2021; Soo & Worth, 2015).

Recent advances in quantitative genetics emphasize the importance of estimating genetic parameters such as heritability in order to optimize breeding strategies. Heritability and genetic variation of a trait provide essential information for breeding programs as the response to selection is directly proportional to the heritability when selection is based on individual phenotypic values (Strandberg *et al.*, 2023). Studies using pedigree-based genetic evaluations have shown that genetic parameter estimates for CHD may vary across breeds, populations, countries, scoring systems, and statistical models, limiting the direct extrapolation of results between populations (Wang *et al.*, 2020).

In Peru, despite the existence of pedigree records and radiographic evaluations in purebred dogs, there is limited information regarding the genetic parameters associated with CHD. Although CHD is recognized as a developmental and heritable condition and family history

is already considered a risk factor in screening programs, population-specific estimates of genetic parameters relevant for breeding programs remain limited, highlighting the need for further studies in this field (Jacobsen *et al.*, 2024). Rottweilers and Golden Retrievers are among the breeds reported to be affected by CHD (Binversie *et al.*, 2022). Therefore, the aim of this study was to estimate the heritability of CHD in Rottweiler and Golden Retriever dogs registered in the Kennel Club del Perú (KCP) between 2014 and 2024. We hypothesized that CHD in Peruvian Rottweilers and Golden Retrievers exhibits moderate to high heritability, or a detectable additive genetic signal, supporting the future development of estimated breeding value (EBV)-based genetic evaluation within the national registry.

MATERIALS AND METHODS

Data source and phenotypic classification

Pedigree and CHD scoring records covering the period from 2014 to 2024 (year of radiographic evaluation) were provided by the KCP. The database included information on breed, sex, age at evaluation, and hip score according to the FCI grading system of Rottweilers ($n = 146$) and Golden Retrievers ($n = 237$). The official dog-level FCI grade recorded by KCP was used as the phenotype for all analyses. According to the KCP protocol, when both hips were evaluated, the final grade corresponded to the more severely affected hip. For example, if the left hip was graded B and the right hip was graded C, the dog was assigned grade C. This conservative approach ensures that the phenotype reflects the worst hip conformation observed, which is relevant for breeding decisions aimed at eliminating dysplasia. The dataset included 131 and 84 sires and 185 and 178 dams for Rottweilers and Golden Retrievers, respectively, with age at evaluation ranging from 7 to 83 months (median: 27 months; first quartile – third quartile: 21–40) for Golden Retrievers and from 10 to 91 months (median: 25 months; first quartile – third quartile: 20–33) for Rottweilers. Hip scores were categorized into five ordinal levels: A (0 = negative), B (1 = unilateral without bony deformation), C (2 = bilateral with or without mild bony deformation), D (3 = bilateral with moderate bony deformation), and E (4 = bilateral with severe bony deformation). For binary analysis, dogs scored A were classified as non-dysplastic (negative), while those with scores B through E were considered dysplastic.

Estimation of variance components and heritability

A Bayesian threshold model approach was employed to estimate variance components and heritability for hip dysplasia using a univariate animal model. Separate univariate Bayesian threshold animal models were fitted for each breed and trait definition. Two trait definitions were analyzed: (1) a binary trait (presence/absence of dysplasia) and

(2) an ordinal categorical trait with five ordered categories corresponding to the FCI grades (FCI, 2022). For the binary trait definition, the sequential justification for the dichotomization was as follows. First, the FCI scoring system classifies hip joints into five grades (A through E), where grade A represents excellent hip conformation without any signs of dysplasia, and grades B through E represent varying degrees of hip dysplasia, ranging from mild (grade B) to severe (grade E). Second, from a genetic and clinical perspective, even mild forms of hip dysplasia (grades B and C) have a heritable component and can progress to osteoarthritis over time, making them relevant for breeding selection decisions. Therefore, the cut-off point was established between grades A and B: animals with FCI grade A were classified as phenotypically unaffected (coded as 0), while those with grades B, C, D, or E were classified as affected (coded as 1). This threshold ensures that all degrees of dysplasia, including mild forms, are captured in the binary analysis, maximizing the genetic variation available for estimation and aligning with breeding goals aimed at complete elimination of the condition.

For each individual, an unobserved latent liability l was modeled as

$$l = X_b + Z_a + e$$

where b denotes fixed effects, a the additive genetic animal effects, and e the residual effects. The observed binary or ordinal CHD phenotype was linked to the latent liability through threshold parameters. Incidence matrices X and Z link the fixed and random effects to the observations. Four models were evaluated for each breed and trait type. Model 1: null model (no fixed effects); Model 2: including sex as a fixed effect; Model 3: including age as a fixed effect; and Model 4: including both sex and age as fixed effects. Age at radiographic evaluation was included as a continuous covariate (in months). Variance components and heritability

$$h^2 = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_e^2}$$

were estimated using Markov Chain Monte Carlo (MCMC) methods via Gibbs sampling. Models were fitted in R (version 4.5.2) using the MCMCglmm package (Hadfield, 2010). Three independent MCMC chains were run for each model with different random seeds to assess convergence using the Gelman-Rubin diagnostic. For binary trait analysis, a total of 1,200,000 iterations were run, with the first 200,000 discarded as burn-in, and samples collected every 100 iterations (thinning interval = 100), resulting in 10,000 retained samples per chain for posterior inference. For ordinal trait analyses, longer chains were required to ensure adequate mixing: 3,000,000 iterations were run with 500,000 discarded as burn-in and a thinning interval of 1,000, yielding 2,500 retained samples per chain. Burn-in and thinning intervals were determined based on visual inspection of trace plots and autocorrelation analysis to

ensure adequate mixing and independence of samples

Two trait definitions were analyzed, each requiring specific prior specifications according to their distributional assumptions. For the binary trait (affected vs. unaffected), a threshold model with a probit link function was fitted using the family = "categorical" option. The residual variance on the underlying liability scale was fixed at ($\sigma_e^2 = 1$), which is a mathematical requirement for model identifiability in binary threshold models (Dreher et al., 2019; Sorensen et al., 1995). A moderately informative inverse-Wishart prior was specified for the additive genetic variance with scale parameter $V = 1$ and degrees of freedom $\nu = 5$. This specification provides sufficient regularization to ensure model convergence while allowing the data to substantially influence the posterior estimates (Hadfield, 2010). Heritability on the liability scale was calculated for each MCMC iteration as

$$h^2 = \frac{\sigma_a^2}{\sigma_a^2 + 1}$$

and posterior means and credible intervals were derived from the resulting distribution of h^2 values across all retained samples. Heritability on the liability scale was calculated for each retained MCMC iteration, and the posterior means and 95% credible intervals reported in the tables were then derived from this distribution of h^2 values. It is important to note that because this transformation is non-linear, the posterior mean of h^2 is not equal to the formula applied to the posterior mean of (a consequence of Jensen's inequality). This is a standard property of Bayesian inference and does not indicate any inconsistency between the reported variance components and heritability estimates.

For the ordinal trait, a threshold model was fitted using the family = "ordinal" option. A moderately informative inverse-Wishart prior was specified for the additive genetic variance with $V = 1$ and $\nu = 10$, consistent with the binary trait specification but providing additional regularization appropriate for the ordinal structure. The residual variance was not fixed, as the ordinal threshold model in MCMCglmm estimates this parameter internally alongside the threshold cut-points (Hadfield, 2010). Heritability on the liability scale was calculated for each MCMC iteration as

$$h^2 = \frac{\sigma_a^2}{\sigma_a^2 + 1}$$

using the same formula as for the binary trait, since the residual variance on the underlying probit liability scale equals 1 for both model formulations (Sorensen et al., 1995; Dreher et al., 2019). To assess the influence of prior specification on posterior estimates, a sensitivity analysis was conducted using alternative priors with $\nu = 2$, $\nu = 10$, and $\nu = 1000$ for binary traits, and $\nu = 5$, $\nu = 20$, and $\nu = 50$ for ordinal traits. Posterior heritability estimates were compared across prior specifications to evaluate robustness. Initial trials with more extreme specifications ($\nu =$

0.002 and $\nu = 100$) were also explored but yielded poor convergence diagnostics and were therefore excluded from the final analysis.

Convergence and mixing were assessed for all main parameters, including additive genetic variance, heritability, fixed effects (sex and age), threshold parameters (for the ordinal model), and deviance, using multiple diagnostic criteria. Visual inspection of trace plots from the three independent chains was performed to assess mixing and stationarity. The potential scale reduction factor (Gelman-Rubin diagnostic, $\hat{R}$) was calculated for all variance components, with values < 1.05 indicating adequate convergence between chains. In addition, autocorrelation at lag 10 was analyzed for both breeds and trait definitions, with values below 0.20 indicating minimal dependence between retained samples and adequate chain mixing. To confirm that the posterior distributions were based on a sufficient number of effectively independent samples, the effective sample size (ESS) was calculated, with values exceeding 100 considered optimal, for all variance components across all models.

Posterior means, standard deviations, and 95% highest posterior density (HPD) credible intervals were calculated for all parameters from the combined posterior samples across the three chains. Candidate models were compared using the Deviance Information Criterion (DIC), interpreted as a criterion balancing model fit and effective complexity. Models with lower DIC were preferred, and Δ DIC was reported to evaluate the strength of model support (Spiegelhalter et al., 2002). Following the guidelines proposed by Spiegelhalter et al. (2002), differences greater than 7 units indicate strong evidence in favor of the model with the lower DIC, differences between 3 and 7 units represent moderate evidence, and differences below 3 units suggest that the models are essentially equivalent in terms of fit. When competing models had Δ DIC < 3 , the principle of parsimony was applied to select the simpler model.

The complete MCMCglmm model calls, including all fixed and random effect specifications, residual structures, family distributions, prior specifications, and MCMC parameters, are provided in Supplementary Material to ensure full reproducibility of the analyses. Prior sensitivity analysis, autocorrelations, and traces are also shown.

RESULTS

Pedigree structure and completeness

The pedigrees comprised 1,919 Golden Retrievers (84 sires, 178 dams) and 1,781 Rottweilers (131 sires, 185 dams). Pedigree completeness was 95.21% and 88.10%, respectively, with 92 and 212 founders across eight and four generations. The mean inbreeding coefficient was $< 2\%$ in both breeds. These characteristics indicate adequate pedigree connectivity for reliable Bayesian estimation of genetic parameters (Table 1).

Table 1.
Pedigree structure of Rottweiler and Golden Retriever populations registered in the Kennel Club del Perú (2014–2024).

Parameter	Rottweiler	Golden Retriever
Total animals	1,781	1,919
Unique sires	131	84
Unique dams	185	178
Animals with unknown sex	1,465	1,657
Founders (≥ 1 unknown parent)	212	92
Animals with both parents known (%)	1,569 (88.10%)	1,827 (95.21%)
Number of generations	4	8
Mean inbreeding coefficient	0.0106	0.0175
Range of inbreeding coefficients	0 to 0.25	0 to 0.3125

MCMC convergence diagnostics

Convergence was evaluated using three independent MCMC chains for each model to ensure a robust assessment. Visual inspection of the trace plots indicated stable mixing for the additive genetic variance and heritability, with no apparent trends or multimodality. To quantify convergence between chains, the potential scale reduction factor ($\hat{R}$) was calculated. All selected models yielded $\hat{R}$ values below 1.05, confirming that the chains explored the same parameter space. Autocorrelation decayed rapidly across all parameters. By retaining samples at intervals of 100 iterations for the binary models and 1,000 iterations

for the ordinal models, we minimised the dependence between successive draws. As a result, the ESS exceeded 100 for all variance components in every model and reached over 1,000 for the ordinal analyses. These results collectively indicate that the MCMC chains converged reliably and that the posterior estimates of heritability and variance components are reproducible.

Population description, model selection and heritability estimates

In Rottweilers, 109 dogs were graded A, 23 B, 10 C, and 4 D, with no dogs graded E. In Golden Retrievers, 150 dogs were graded A, 54 B, 21 C, 10 D, and 2 E. Binary affected

Table 2.
Distribution of Fédération Cynologique Internationale (FCI) hip dysplasia grades in Rottweiler and Golden Retriever dogs evaluated in Peru.

Dysplasia grade ^a	Rottweiler [n (%)]	Golden Retriever [n (%)]
Grade A	109 (74.7)	150 (63.3)
Grade B	23 (15.8)	54 (22.8)
Grade C	10 (6.8)	21 (8.9)
Grade D	4 (2.7)	10 (4.2)
Grade E	0 (0.0)	2 (0.8) ^b
Total	146	237

Note. Hip score: A (0 = Negative), B (1 = Unilateral without bony deformation), C (2 = Bilateral with/without mild bony deformation), D (3 = Bilateral with moderate bony deformation), E (4 = Bilateral with severe bony deformation).

^a Fédération Cynologique Internationale (FCI) grades.

^b For the Golden Retriever breed, the observations with grade E were combined with grade D.

Table 3.

Posterior means and standard deviation of variance components and heritability of hip dysplasia in Rottweiler in Peru.

Parameter	Model 1	Model 2	Model 3	Model 4
Binary traits				
Additive variance (σ_a^2)	1.8123 (1.6471) (0.2658 - 4.6978)	2.099 (3.1241) (0.2654 - 5.4181)	2.0892 (2.0537) (0.2719 - 5.6131)	2.6737 (8.3067) (0.2350 - 6.8519)
Heritability (h^2)	0.5743 (0.1486) (0.2954 - 0.8525)	0.5854 (0.1536) (0.3050 - 0.8822)	0.5978 (0.1529) (0.3162 - 0.8837)	0.6016 (0.1610) (0.3129 - 0.9143)
DIC	160.16	160.16	157.59*	157.6
Ordinal traits				
Additive variance (σ_a^2)	1.6730 (1.0577) (0.4066 - 3.6418)	1.7255 (1.1974) (0.4333 - 3.8149)	1.8158 (1.2775) (0.3797 - 4.1256)	1.9093 (1.6043) (0.4280 - 4.3480)
Heritability (h^2)	0.5856 (0.1162) (0.3634 - 0.8059)	0.5880 (0.1206) (0.3634 - 0.8059)	0.5983 (0.1204) (0.3691 - 0.8332)	0.6044 (0.1223) (0.3760 - 0.8469)
DIC	197.71	197.93	193.48	151.80*

DIC: Deviance Information Criterion. SD and 95% credible intervals are shown in parentheses. Model 1: no effect, Model 2: sex, Model 3: age, Model 4: sex + age. Heritability was calculated for each individual MCMC iteration, and the posterior mean and 95% credible intervals were derived from the resulting distribution. Due to the non-linear nature of this transformation, the posterior mean of h^2 does not equal the formula applied to the posterior mean of additive variance (Jensen's inequality).

* The best model, smallest DIC

Table 4.

Posterior means and standard deviation of variance components and heritability of hip dysplasia in Golden Retriever in Peru.

Parameter	Model 1	Model 2	Model 3	Model 4
Binary traits				
Additive variance (σ_a^2)	2.1555 (1.9334) (0.3025 - 5.4246)	2.3535 (2.3589) (0.2721 - 6.1060)	2.5413 (2.3706) (0.2988 - 6.5553)	3.3574 (4.8429) (0.3153 - 9.1118)
Heritability (h^2)	0.6148 (0.1429) (0.3412 - 0.8796)	0.6268 (0.1454) (0.3488 - 0.8902)	0.6449 (0.1435) (0.3683 - 0.9053)	0.6773 (0.1456) (0.3914 - 0.9405)
DIC	299.97	299.12	298.03*	296.64
Ordinal traits				
Additive variance (σ_a^2)	1.3103 (0.7192) (0.4194 - 5.7634)	1.3708 (0.7949) (0.4177 - 2.8010)	1.5609 (1.3621) (0.4688 - 3.1658)	1.6515 (1.0479) (0.4223 - 3.5452)
Heritability (h^2)	0.5379 (0.1059) (0.3468 - 0.7519)	0.5455 (0.1096) (0.3363 - 0.7540)	0.5720 (0.1107) (0.3597 - 0.7827)	0.5843 (0.1132) (0.3665 - 0.8022)
DIC	416.51	414.99	404.16*	404.42

DIC: Deviance Information Criterion. SD and 95% credible intervals are shown in parentheses. Model 1: no effect, Model 2: sex, Model 3: age, Model 4: sex + age. Heritability was calculated for each individual MCMC iteration, and the posterior mean and 95% credible intervals were derived from the resulting distribution. Due to the non-linear nature of this transformation, the posterior mean of h^2 does not equal the formula applied to the posterior mean of additive variance (Jensen's inequality).

* The best model, smallest DIC

status was then defined according to the analytical threshold described in the Methods (Table 2). Using the strict A versus B–E threshold, the affected status was 25.3% in Rottweilers and 36.7% in Golden Retrievers.

Heritability Estimates Posterior means of heritability and additive genetic variance (σ_a^2) are presented for both binary and ordinal analyses across the four models tested for both breeds (Tables 3 and 4).

For Rottweilers, under the binary model, DIC values progressively decreased from Model 1 and Model 2 (DIC = 160.16) to Model 3 (DIC = 157.59) and Model 4 (DIC = 157.60), with Model 3 providing the best fit (Table 3). However, the difference between Model 3 and Model 4 was only 0.01 units, indicating that the inclusion of sex in addition to age provided only a marginal improvement in model fit. For the ordinal trait, Model 4 yielded the lowest DIC (151.80), outperforming Model 3 by 41.68 units, which represents strong evidence in favor of the complete model. This suggests that for the ordinal trait in Rottweilers, both sex and age explain a relevant portion of the phenotypic variation. Consequently, heritability estimates from Model 3 ($h^2 = 0.598$, binary) and Model 4 ($h^2 = 0.604$, ordinal) were selected as the most appropriate for Rottweilers.

Similar trends were observed in Golden Retrievers (Table 4). In the binary analysis, DIC decreased progressively from Model 1 (DIC = 299.97) to Model 4 (DIC = 296.64), with Model 4 providing the best fit. The difference between Model 3 (DIC = 298.03) and Model 4 was 1.39 units, indicating weak evidence and suggesting that both models are essentially equivalent. Similarly, in the ordinal analysis, Model 3 (DIC = 404.16) outperformed Model 4 (DIC = 404.42) by only 0.26 units. Because grade E was rare in Golden Retrievers, grades D and E were combined for ordinal analysis, as described in the Methods section. These results indicate that both sex and age contribute to explaining phenotypic variation in CHD among Golden Retrievers. However, Model 3 (age only) was selected for both trait definitions based on the principle of parsimony. For the binary trait, Models 3 and 4 showed essentially equivalent fit (Δ DIC = 1.39), whereas for the ordinal trait, Model 3 yielded the lowest DIC with a negligible difference from Model 4 (Δ DIC = 0.26). Model 3 also demonstrated superior convergence diagnostics, with R-hat values closer to 1.0, lower autocorrelation at lag 10, and higher effective sample sizes compared to Model 4. Accordingly, the heritability estimates from Model 3 ($h^2 = 0.645$, binary; $h^2 = 0.572$, ordinal) were selected as the most appropriate for this breed, although the credible intervals remained wide.

Across both breeds, adding fixed effects generally improved the model fit over the null model, as shown by the lower DIC values (Supplementary Material). However, this improvement often levelled off when both sex and age were included together. Contrary to initial expectations, ordinal models did not consistently yield higher heritability estimates than binary models; the estimates were broadly comparable, with the binary trait producing

slightly higher posterior means in Golden Retrievers. Ultimately, while models incorporating age and/or sex consistently outperformed the intercept-only model, the most complex specification did not always provide a meaningful advantage over simpler alternatives. Consequently, the principle of parsimony, supported by robust convergence diagnostics, guided the final model selection across the different breed–trait combinations.

DISCUSSION

This study estimated breed-specific heritability for binary CHD status and ordinal FCI grade on the latent liability scale using Bayesian threshold animal models. The h^2 for the binary trait (status) were 0.5978 ± 0.1529 for Rottweilers (Table 3) and 0.6449 ± 0.1435 for Golden Retrievers (Table 4), based on the best-supported model. Although the posterior mean estimates suggest moderate-to-high heritability, the relatively large posterior standard deviations, particularly for Rottweilers, indicate substantial uncertainty and should be considered when interpreting these estimates. Additionally, these heritability values are considerably higher than those reported in other international populations, which typically range from 0.11 to 0.45 (Babá *et al.*, 2019; Jardim, 2023; Mäki *et al.*, 2002). These posterior means suggest that additive genetic effects may explain a moderate-to-high proportion of variation in latent CHD liability in the analyzed populations. Additionally, the posterior mean of additive genetic variance increased when fixed effects were accounted for, reflecting a more accurate partitioning of the total phenotypic variance into its genetic and environmental components. This high genetic variance indicates that genetic improvement may be feasible, although the expected response to selection would require additional evaluation of the EBV accuracy, selection intensity, and genetic trends. (Alsada, 2023). These findings are consistent with a detectable additive genetic component, while still allowing for environmental and other sources of variation (Binversie *et al.*, 2022). Differences in h^2 across studies may also reflect variations in population structure, selection intensity, management practices, or other population-specific factors. Whether the comparatively higher estimates observed in the present study are related to these factors remains to be investigated. Importantly, while these high heritability estimates indicate strong potential for rapid genetic gain, they do not imply that Peruvian populations will achieve phenotypic outcomes superior to those of other countries. Rather, implementing structured breeding programs could enable these populations to reach the same phenotypic levels as populations in countries with established CHD reduction programs, potentially in a shorter timeframe due to the greater available additive genetic variance.

The choice of a Bayesian threshold model was particularly appropriate for these data. As noted in similar studies (Babá *et al.*, 2019; Stock *et al.*, 2011), CHD scores are cate-

gorical and do not follow a normal distribution. Because CHD scores are categorical or ordinal, threshold models are more appropriate than Gaussian linear models for representing an underlying liability. The threshold model assumes an underlying continuous liability scale, which better represents the polygenic nature of CHD. Model selection was based on the DIC, which balances model goodness-of-fit against model complexity, with lower values indicating a better-fitting model (Spiegelhalter *et al.*, 2002). A summary of the DIC differences between the competing models for heritability estimation is shown in Table 5. To interpret the magnitude of the DIC differences between competing models, we followed the guidelines proposed by Spiegelhalter *et al.* (2002), where differences > 7 units indicate strong evidence in favor of the model with the lower DIC, differences between 3–7 units represent moderate evidence, and differences < 3 units suggest that the models are essentially equivalent in terms of fit. Among the evaluated models, Model 3 (age) was selected as the best-fitting model for the binary trait in both breeds and the ordinal trait in Golden Retrievers. For the ordinal trait in Rottweilers, Model 4 (sex + age) provided the best fit, suggesting that the inclusion of sex and age better explains the ordered severity of CHD in this breed. The Bayesian framework is a strength of the present study because it incorporates prior distributions and pedigree information when estimating variance components. Nevertheless, the resulting estimates may still be sensitive to the choice of priors and the degree of pedigree connectedness. These findings suggest the possibility that the Peruvian breeding population for these breeds may harbor considerable additive genetic variance that has not yet been fully exploited through current breeding practices, although this interpretation remains hypothetical and would require further validation through the estimation of EBVs, selection response, and formal genetic trend analyses.

Despite this high heritability, preliminary observations of phenotypic prevalence in the KCP registry suggest that the proportion of dogs diagnosed with CHD has not substantially decreased between 2014 and 2024. However, formal genetic trend analyses based on estimated breeding values were not conducted in this study, and therefore,

whether actual genetic progress has occurred over time remains to be determined. In Peru, breeding decisions have traditionally relied primarily on individual phenotypic radiographic scores, whereas the use of genetic evaluation methods that incorporate pedigree relationships or estimated breeding values (EBVs) remains limited. In general, selection based solely on individual phenotypes may be less accurate than selection based on EBVs because EBVs incorporate information from relatives and pedigree relationships (Soo & Worth, 2015). Selection based only on individual phenotypic measurements generally results in slower genetic improvement when compared with strategies that incorporate familial information, since hip conformation scores are strongly correlated with those of parents, grandparents, and siblings (Bell, 2018).

The high h^2 estimates obtained in this study suggest that substantial additive genetic variation exists within the analyzed populations. However, phenotypic selection alone may be insufficient because individual radiographic scores cannot fully capture underlying genetic merit. Phenotypically normal dogs may still have unfavorable breeding values because individual radiographic scores are imperfect predictors of transmissible additive genetic risk (Woody, 2025). As argued by Wang *et al.* (2020), dogs with normal radiographic scores may still have unfavorable EBVs and contribute additive genetic risk to the next generation, highlighting a limitation of selection strategies based solely on pass–fail radiographic screening. This illustrates that for polygenic traits, selection based only on observed phenotype may be less accurate than selection based on estimated breeding values.

Furthermore, the reliance on the FCI grading system as a binary model used in this study (comparing Grade A vs. Grades B–E) provides a practical threshold for breeders. This threshold ensures that all degrees of dysplasia, including mild forms, are captured in the binary analysis, maximizing the genetic variation available for estimation and aligning with breeding goals aimed at the complete elimination of the condition. Although collapsing ordinal categories into a binary trait may reduce phenotypic resolution, this approach is commonly used in genetic studies of CHD because it reflects the practical pass–fail criteria

Table 5.
Summary of DIC differences between competing models for heritability estimation of canine hip dysplasia.

Breed	Trait	Best Model	DIC	Alternative Model	DIC	Δ DIC	Evidence strength
Rottweiler	Binary	Model 3 (age)	157.59	Model 4 (sex + age)	157.60	0.01	Weak
	Ordinal	Model 4 (sex + age)	151.80	Model 3 (age)	193.48	41.68	Strong
Golden Retriever	Binary	Model 3 (age)	298.03	Model 4 (sex + age)	296.64	1.39	Weak
	Ordinal	Model 3 (age)	404.16	Model 4 (sex + age)	404.42	0.26	Weak

Δ DIC = difference in DIC between the best model and the second-best model. Evidence strength interpreted following Spiegelhalter *et al.* (2002): < 3 = weak; 3–7 = moderate; > 7 = strong.

applied in many breeding programs. Similar binary classification approaches have been employed in other studies, in which grade A animals are considered unaffected, while grades B-E are grouped as affected categories to facilitate genetic analyses and selection decisions (Mostafa *et al.*, 2022). However, the absence of a clear downward trend in phenotypic prevalence suggests that the current level of selection pressure may remain insufficient, although this interpretation should be confirmed through formal genetic trend analyses.

In some breeding systems, dogs classified as FCI grade B are considered clinically normal or near-normal and are therefore allowed to reproduce. However, whether grade B animals contribute to future CHD risk should be evaluated through EBVs or longitudinal selection analyses rather than inferred from the present variance-component model (Tomé *et al.*, 2026). Breed-specific differences in hip joint morphology and biomechanics may also influence radiographic evaluation, as variations in skeletal conformation among breeds can affect the interpretation of the parameters used in FCI scoring systems (Pinna *et al.*, 2023). Given the substantial additive genetic variation estimated in the present study, future studies should estimate annual phenotypic and genetic trends to determine whether this variation is being translated into a selection response.

These results support the development of EBV-based evaluations as a potentially more informative complement to individual radiographic scores, as EBVs incorporate information from relatives and pedigree relationships when estimating genetic merit (Lewis *et al.*, 2013). The potential benefit of EBVs in the KCP registry should be evaluated by estimating EBV accuracy, monitoring genetic trends, and assessing breeder uptake (Pinto Coelho Ribeiro Jardim *et al.*, 2025). Indeed, breeding programs that have incorporated EBVs into selection decisions since the 1980s have demonstrated substantial reductions in the prevalence of hip dysplasia across several dog populations (Thorsrud *et al.*, 2025). Recent studies indicate that genomic information can further improve selection accuracy in the future, but its effective implementation requires larger datasets, validated markers, and integration with pedigree-based evaluations (Ardicli *et al.*, 2025).

Several limitations should be acknowledged. First, the number of animals included in the dataset may have limited the precision of the h^2 estimates, particularly because analyses were conducted separately for each breed. Previous genetic studies have shown that the number of available records can strongly influence the accuracy of heritability estimates, with larger datasets generally providing more reliable genetic parameter estimations (Wang *et al.*, 2020). The pedigree-based Bayesian animal model allowed relationship information to be incorporated but did not eliminate uncertainty due to the limited sample size. Second, radiographic evaluations obtained from routine screening programs may introduce phenotypic variability due to differences in positioning, radiographic quality,

or evaluator interpretation. Diagnostic techniques based on imaging and metric extraction may also be affected by human error during probe positioning or interpretation, which can increase the risk of misclassification and introduce additional variability into phenotypic assessments (Akula *et al.*, 2022). The importance of correct positioning during radiographic examination has been highlighted in previous studies, as small deviations in positioning may influence the subjective interpretation of hip joint conformation and introduce additional variability into phenotypic assessments (Franco-Gonçalo *et al.*, 2023). Unmodeled measurement error in radiographic scoring may have contributed to residual variation and potential misclassification of the binary or ordinal phenotype. Third, a potential limitation of this study is that the binary phenotype definition used a strict threshold (grade A vs. grades B-E), which may not be directly comparable with studies using alternative thresholds such as grades A-B vs. C-E. Although we acknowledge that different threshold definitions could yield different heritability estimates, the A vs B-E threshold was chosen because it aligns with the practical breeding goal of eliminating all degrees of dysplasia, including mild forms. A formal sensitivity analysis comparing multiple binary thresholds was not conducted due to the limited sample size, which could have reduced statistical power for alternative classifications with fewer observations per category. Future studies with larger datasets should evaluate the robustness of heritability estimates across different phenotype definitions. Finally, a potential selection bias may exist because dogs submitted for evaluation may not fully represent the entire breeding population. Because participation in screening may be non-random, both prevalence and variance-component estimates may be affected by ascertainment bias. The results should therefore be interpreted as baseline estimates for the evaluated KCP population rather than definitive parameters for all breeding dogs in Peru.

The results suggest moderate-to-high heritability for CHD status in the evaluated Rottweiler and Golden Retriever populations, although the posterior uncertainty was substantial. These estimates indicate the presence of additive genetic variation that could be relevant for future breeding programs. However, the present study did not estimate genetic trends, EBV accuracy, or expected response to selection; therefore, whether the observed lack of phenotypic progress reflects limitations of current selection practices, phenotyping methods, breeder compliance, or other factors remains hypothetical. Moreover, radiographic screening alone may not fully capture additive genetic risk, emphasizing the potential value of more informative genetic evaluation tools in breeding programs (Kieler *et al.*, 2024). Future research should evaluate whether EBV-based selection could improve the identification of animals with favorable additive genetic merit, and its impact on CHD prevalence should be evaluated through formal estimation of EBV accuracy, selection re-

sponse, and long-term monitoring. While these findings suggest the potential for EBV-based genetic evaluation as a next step toward improving long-term hip health in registered Rottweiler and Golden Retriever populations, this hypothesis requires validation through dedicated genetic trend and selection response studies.

Competing Interests Statement

The authors declare no conflicts of interest.

Ethics Statement

The study was approved by the Institutional Committee of Ethics in Research with Animals and Biodiversity of the Universidad Científica del Sur (Cod. PRE-16-2025-00064, July 22, 2025)

Data Availability Statement

The data are not publicly available due to privacy restrictions from the Kennel Club del Perú (KCP) registry.

Author contributions

Conceptualization, M.J.P.T. and J.L.V.V.; methodology, J.L.V.V.; formal analysis, J.L.V.V.; investigation, M.J.P.T.; data curation, M.J.P.T.; writing—original draft preparation, M.J.P.T.; writing—review and editing, J.L.V.V. All authors have read and agreed to the published version of the manuscript.

Funding

This research received no external funding.

Acknowledgments

The authors would like to thank the Kennel Club del Perú (KCP) for providing the database of radiographic evaluations and pedigree records used in this study. We also thank the breeders and veterinarians who contributed to the official screening programs during the 2014–2024 period.

REFERENCES

- Akula, S.P., Akula, P., & Kamati, N. (2022). Detection and classification of canine hip dysplasia according to FCI grading system using 3D CNN's. In *Proceedings of the 2022 First International Conference on Artificial Intelligence Trends and Pattern Recognition (ICAITPR)*; IEEE: Piscataway, NJ, USA, pp. 1–6. <https://doi.org/10.1109/ICAITPR51569.2022.9844209>
- Alsada, I. H. A., (2023). Hip dysplasia in large breed of dogs. In Aguilera-Marcelino L, Younus M, Khan A, Saeed NM and Abbas RZ (eds.). *One Health Triad*. Unique Scientific Publishers, Faisalabad, Pakistan, Vol. 3, pp: 202–207. <https://doi.org/10.47278/book.oht/2023.96>
- Ardicli, S., Yigitgor, P., Ozen, D., Babayev, H., Bozkurt, B., Senturk, N., & Intas, D.S. (2025). The effectiveness of genetic markers and the role of environmental factors in hip dysplasia and osteochondritis dissecans of the shoulder in German Shepherd, Labrador Retriever, and German Wirehaired Pointer (Deutsch Drahthaar) dogs. *Gene Reports*, 38, 102153. <https://doi.org/10.1016/j.genrep.2025.102153>
- Babá, A.Y., de Oliveira, C.A.L., Yoshida, G.M., Costa, M.T.C., Ribeiro, L.B., & Vasconcellos, R.S. (2019). Heritability of hip dysplasia: Preliminary results for German Shepherd dogs in Brazil. *Preventive Veterinary Medicine*, 171, 104745. <https://doi.org/10.1016/j.prevetmed.2019.104745>
- Bell, J. (2018). *What Practitioners Should Know About the Genetics of Hip Dysplasia*. World Small Animal Veterinary Association Congress Proceedings, 2018. Veterinary Information Network. <https://www.vin.com/doc/?id=8896662>
- Binversie, E.E., Momen, M., Rosa, G.J.M., Davis, B.W., & Muir, P. (2022). Across-breed genetic investigation of canine hip dysplasia, elbow dysplasia, and anterior cruciate ligament rupture using whole-genome sequencing. *Frontiers in Genetics*, 13, 913354. <https://doi.org/10.3389/fgene.2022.913354>
- Dreher, C., Wellmann, R., Stratz, P., Schmid, M., Preuß, S., Hamann, H., & Bennewitz, J. (2019). Genomic analysis of perinatal sucking reflex in German Brown Swiss calves. *Journal of Dairy Science*, 102(7), 6296–6305. <https://doi.org/10.3168/jds.2019-16487>
- Fédération Cynologique Internationale (2022). *FCI hip dysplasia classification*. FCI Scientific Commission. <https://www.fci.be/medias/SCI-DYS-HAN-CLA-en-15522.pdf>
- Franco-Gonçalo, P., Alves-Pimenta, S., Gonçalves, L., Colaço, B., Leite, P., Ribeiro, A., Ferreira, M., McEvoy, F., & Ginja, M. (2023). Femoral parallelism: Evaluation and impact of variation on canine hip dysplasia assessment. *Frontiers in Veterinary Science*, 10, 1160200. <https://doi.org/10.3389/fvets.2023.1160200>
- Hadfield, J. D. (2010). MCMC methods for multi-response generalized linear mixed models: The MCMCglmm R package. *Journal of Statistical Software*, 33(2), 1–22. <https://doi.org/10.18637/jss.v033.i02>
- Jacobsen, K.K., Laborie, L.B., Kristiansen, H., Schäfer, A., Gundersen, T., Zayats, T., & Rosendahl, K. (2024). Genetics of hip dysplasia—a systematic literature review. *BMC Musculoskeletal Disorders*, 25, 762. <https://doi.org/10.1186/s12891-024-07795-2>
- Jardim, L.P.C.R. (2023). *Análise de status de quadril de Golden Retrievers no Brasil: Um estudo de saúde e melhoramento genético*. [Master's Thesis, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil]. <https://lume.ufrgs.br/bitstream/handle/10183/262618/001173416.pdf?sequence=1&isAllowed=y>
- Kieler, I.N., Persson, S.M., Hagman, R., Marinescu, V.D., Hedhammar, Å., Strandberg, E., Lindblad-Toh, K., & Arendt, M.L. (2024). Genome wide association study in Swedish Labrador retrievers identifies genetic loci associated with hip dysplasia and body weight. *Scientific Reports*, 14, 6090. <https://doi.org/10.1038/s41598-024-56060-y>
- Lewis, T.W., Blott, S.C., & Woolliams, J.A. (2013). Comparative analyses of genetic trends and prospects for selection against hip and elbow dysplasia in 15 UK dog breeds. *BMC Genetics* 2013, 14, 16. <https://doi.org/10.1186/1471-2156-14-16>
- Lucaci, F.D., Lăcătuș, R., Purdoi, R.C., & Pusta, D.L. (2021). Study on canine hip dysplasia and its diagnosis by imaging and molecular genetics techniques. *Bulletin of University of Agricultural Sciences and Veterinary Medicine Cluj-Napoca*, 78(1). <https://doi.org/10.15835/bu-asmcn-vm:2020.0031>
- Mäki, K., Groen, A.F., Liinamo, A.E., & Ojala, M. (2002). Genetic variances, trends and mode of inheritance for hip and elbow dysplasia in Finnish dog populations. *Animal Science*, 75(2), 197–207. <https://doi.org/10.1017/S1357729800052966>
- Mikkola, L., Kyöstilä, K., Donner, J., Lappalainen, A.K., Hytönen, M.K., Lohi, H., & Iivanainen, A. (2021). An across-breed validation study of 46 genetic markers in canine hip dysplasia. *BMC Genomics*, 22, 68. <https://doi.org/10.1186/s12864-021-07375-x>
- Mostafa, A.A., Nahla, M.A., Ali, K.M., & Berry, C.R. (2022). Modified FCI (Fédération Cynologique Internationale) scoring of the coxofemoral joint in Labrador Retrievers without and with hip dysplasia. *Frontiers in Veterinary Science*, 9, 800237. <https://doi.org/10.3389/fvets.2022.800237>
- Pinna, S., Vezzoni, A., Di Benedetto, M., Lambertini, C., & Tassani, C. (2023). Characterization of FCI (Fédération Cynologique Internationale) grades for hip dysplasia in five dog breeds. *Animals*, 13(13), 2212. <https://doi.org/10.3390/ani13132212>
- Pinto Coelho Ribeiro Jardim, L., Michelsen de Andrade, F., Tyska, D.U., & Cobuci, J.A. (2025). Analysis of the hip status of Golden Retrievers in Brazil—A study of health and genetic improvement. *Veterinary Sciences*, 12, 746. <https://doi.org/10.3390/vetsci12080746>
- Santana, A., Alves-Pimenta, S., Martins, J., Colaço, B., & Ginja, M. (2021). Imaging diagnosis of canine hip dysplasia with and without human exposure to ionizing radiation. *The Veterinary Journal*, 276, 105745. <https://doi.org/10.1016/j.tvjl.2021.105745>
- Soo, M., & Worth, A. (2015). Canine hip dysplasia: Phenotypic scoring and the role of estimated breeding value analysis. *New Zealand Veterinary Journal*, 63(2), 69–78. <https://doi.org/10.1080/00480169.2014.949893>
- Sorensen, D., Andersen, S., Gianola, D. & Fragas, I. (1995). Bayesian inference in threshold models using Gibbs sampling. *Genetics Selection Evolution*, 27, 229. <https://doi.org/10.1186/1297-9686-27-3-229>
- Spiegelhalter, D.J., Best, N.G., Carlin, B.P., & Van Der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 64(4), 583–639. <https://doi.org/10.1111/1467-9868.00353>
- Stock, K.F., Klein, S., Tellhelm, B., & Distl, O. (2011). Genetic analyses of elbow and hip dysplasia in the German shepherd dog. *Journal of Animal*

- Breeding and Genetics*, 128(3), 219–229. <https://doi.org/10.1111/j.1439-0388.2010.00901.x>
- Strandberg, E., Andersson, L., Emanuelson, U., Bjørnvad, C.R., Ringmark, S., Hedhammar, Å., & Höglund, K. (2023). Heritability and genetic trend of body weight in dogs of different breeds in Sweden. *Journal of Animal Science*, 101, skad173. <https://doi.org/10.1093/jas/skad173>
- Thorsrud, J.A., Evans, K.M., Quigley, C.K., Srikanth, K., Reverter, A., Porto-Neto, L.R., & Huson, H.J. (2025). Breeding values and index creation for health and behavior traits in Labrador Retriever guide dogs. *Frontiers in Veterinary Science*, 12, 1628161. <https://doi.org/10.3389/fvets.2025.1628161>
- Tomé, I., Alves-Pimenta, S., Colaço, B., & Ginja, M. (2026). Ultrasonographic evaluation of canine hip dysplasia: Comparison with FCI radiographic scoring system. *Veterinary Sciences*, 13, 20. <https://doi.org/10.3390/vetsci13010020>
- Wang, S., Friedrich, J., Strandberg, E., Arvelius, P., & Wiener, P. (2020). Methods to improve joint genetic evaluation of canine hip dysplasia across BVA/KC and FCI screening schemes. *Frontiers in Veterinary Sciences*, 7, 386. <https://doi.org/10.3389/fvets.2020.00386>
- Woody, N.R. (2025). Diagnosis, Treatment, Genetic Basis, and Prevention of Canine Hip Dysplasia. [Honors Thesis, Liberty University], 2. <https://digitalcommons.liberty.edu/honors/1472/>