Archiv Tierzucht 54 (2011) 5, 439-455, ISSN 0003-9438

Genetic evaluations of traits recorded in British young horse tests

Isobel D. Stewart[1], John A. Woolliams[2] and Susan Brotherstone[1]

[1]Institute of Evolutionary Biology, University of Edinburgh, Edinburgh, UK, [2]The Roslin Institute, Royal (Dick) School of Veterinary Studies, University of Edinburgh, Easter Bush, Midlothian, Scotland, UK

Abstract

Tests for young sport horses were recently introduced in Great Britain. This study characterises the young horse data, examines their suitability for genetic analysis and estimates the genetic parameters needed for breeding value prediction.

Evaluation data from 2006-2009 were used. This included 1 887 evaluations of 1 323 horses, which were evaluated for competing in dressage, show jumping, eventing, endurance or as sport ponies. Traits assessed were conformation, correctness of paces, type and temperament, athleticism and veterinary. The distributions of traits were examined and correlations of traits between disciplines, for the effect of the horse, were estimated. These indicated that traits could be assumed to be genetically identical across disciplines. Variance components were estimated for each of the 5 traits, using an animal model, where random effects were the genetic effect of the horse and the permanent environment of the horse. Bivariate analyses were performed between pairs of traits.

Mean scores for each trait in each discipline were between 8.02 and 8.24, and standard deviations were between 0.54 and 0.83. Heritabilities ranged between 20.3 % for athleticism and 42.2 % for type and temperament. The variance due to the horse's permanent environment ranged from approximately 25 % for correctness of paces and athleticism to 51.6 % for veterinary. The genetic correlations between traits were generally high.

The young horse tests (»Futurity«) recently introduced in the UK are a valuable data source for genetic evaluations. The most appropriate measure will be to combine young horse data with adult competition data to routinely estimate breeding values.

Keywords: genetic evaluation, horse, young horse tests, breeding value

Zusammenfassung

Genetische Beurteilung von Merkmalen anhand Britischer Jungpferdeprüfungen

In Großbritannien wurden kürzlich Prüfungen für junge Sportpferde eingeführt. Die vorliegende Studie beschreibt die Daten der Jungpferde, prüft deren Eignung für genetische Analysen und schätzt genetische Parameter für die Zuchtwertvorhersage. Es wurden die Daten von 1 323 Pferden, die in 1 887 Prüfungen zwischen 2006 und 2009 gewonnen wurden, verwendet.

Die erfaßten Merkmale waren: Exterieur, korrekte Gangarten, Typ und Temperament, Sportlichkeit und veterinärmedizinische Merkmale. Die Verteilung der Merkmale wurden untersucht und Korrelationen der Merkmale zwischen den Disziplinen für die Pferdeeffekte geschätzt. Diese zeigten, dass Merkmale disziplinübergreifend als genetisch identisch angenommen werden könnten. Varianzkomponenten wurden für jedes der 5 Merkmale anhand eines Tiermodells mit den zufälligen Effekten »genetischer Effekt des Pferdes« und »Umwelt des Pferdes« geschätzt. Bivariate Analysen wurden jeweils zwischen zwei Merkmalen durchgeführt.

Die Mittelwerte für jedes Merkmal in jeder Disziplin lagen zwischen 8,02 und 8,24, deren Standardabweichungen zwischen 0,54 und 0,83. Die Heritabilitäten lagen zwischen 20,3% für Sportlichkeit und 42,2 % für Typ und Temperament. Die Varianz, die auf die Umwelt des Pferdes zurückzuführen war, reichte von ca. 25 % für die korrekten Gangarten und Sportlichkeit bis 51,6 % für veterinärmedizinische Merkmale. Die genetischen Korrelationen zwischen Merkmalen waren in der Regel hoch.

Die in Großbritannien kürzlich eingeführten Reitpferdeprüfungen (»Futurity«) sind eine wertvolle Datenquelle für genetische Evaluierungen. Es scheint am Besten zu sein, Jungpferddaten mit den Wettbewerbsdaten erwachsener Tiere zu verknüpfen, um so routinemäßig Zuchtwerte zu schätzen.

Schlüsselwörter: Zuchtwertschätzung, Pferd, Reitpferdeprüfungen, Zuchtwert

Introduction

Data from young horse tests are frequently used in genetic evaluations of sport horses internationally, either alone, or in combination with adult competition data. For instance, young horse data are used in the genetic evaluations in Sweden, The Netherlands, Belgium, Germany and France, amongst others, where »young« may mean ages between 3 and 7 years of age (Thorén Hellsten *et al.* 2006). The genetic evaluations are made possible by the routine conduct of trials organised by the national breeding organisations, over many years and resulting in large datasets. Data are generally collected from field tests, station tests or young horse competitions, which may last between 1 and 100 days, and can involve up to 45 % of eligible horses (Thorén Hellsten *et al.* 2006). The primary aims are both genetic, to select stallions and mares for breeding, and phenotypic, to identify horses with a high potential to succeed in competition. Commonly, breeding objectives focus only on show jumping and/or dressage, and tests examine the conformation, gaits, performance, behaviour and health of the horse.

Analogous young horse tests (Young Horse Evaluations) were introduced in the UK relatively recently, in 2002, and were developed with particular reference to the form and objectives of two Swedish tests – Young Horse Tests (YHT) and Riding Horse Quality Tests (RHQT), aimed at 3 and 4 years and 4 years respectively. The UK tests were initially aimed at 4 to 6 year olds, but recruitment proved difficult due to the introduction of young horse classes for 4 year olds by competition bodies (J. Rogers personal communication), and the accumulated data proved too small for the estimation of genetic parameters (Kearsley 2007). Therefore a set of new young horse tests (»Futurity«) was developed, targeted primarily at 3 year old horses, but including foals to 3 year olds (4 years olds in 2007). These evolved from the Young Horse Evaluations and were established by a national breeding initiative (British Equestrian

Federation's British Breeding) in 2005. There are dual aims of the Futurity tests: to identify elite horses for competition in the different disciplines of dressage, eventing and show jumping, as well as to inform breeding in the UK by providing data for genetic evaluations. Participation has grown since this scheme was introduced and now attracts approximately 900 horses per year with owners motivated by assessing their horse's future potential in competition, and obtaining gradings for marketing purposes.

Given that young horse data have been collected in the UK since 2005, a dataset of adequate size has only recently become available for use in genetic evaluations. The aim of this work therefore was to 1) characterise the young horse data, 2) investigate its suitability for genetic analysis, 3) estimate the genetic parameters needed for breeding value prediction and 4) predict breeding values for all horses in the pedigree.

Material and methods

Test data for 1 323 horses of all sexes were obtained from the Futurity database for the years 2005-2009. Data from 2005 were excluded, as this was a pilot year, with only 72 records. Horses eligible for the Futurity tests must be British bred (foaled in the UK, or to a mare that is usually resident in UK) and can be aged from a few weeks to 3 years old (4 years in 2007). Horses are entered by owners/breeders, and it is acceptable for them to be evaluated in multiple years. Evaluations are conducted in a single day at events held over the summer months at various locations across the UK. The data collected represent a total of 35 evaluation days, across 15 locations, with a range of 7 to 79 horses attending per day.

At an evaluation, horses are judged by a team of up to four judges and a veterinary practitioner. Horses are judged within an arena, without rider, where they are led through various tests. They are examined for their potential future performance in one of the specific disciplines – show jumping, dressage, eventing, endurance or sports ponies (the latter defined as mature height less than 148 cm and at least one parent under 148 cm). The same traits are examined for each discipline: conformation, correctness of paces, type and temperament, athleticism, and veterinary scores. The precise definition of traits differs between disciplines, for example, if the same horse is evaluated for both show jumping and dressage, it might well receive different grades for paces, as the required paces will differ between the disciplines. Subjective scores from 1 (very poor) to 10 (excellent) are awarded for each trait, in increments of 0.25, and an overall grade summarising all the traits is also awarded. Horses are evaluated for the one discipline for which they are entered by the owner (i.e. they are not, as standard, examined for both dressage and show jumping, as is performed in most European evaluations). In the first years of the futurity (2006-2009) horses could be entered for more than one discipline; as of 2010 the rules changed so that horses may only be entered for one discipline. Judges may advise that a horse is evaluated for an alternative discipline if thought appropriate, but whilst the occurrence of this was frequent in the early years, the frequency has decreased over time.

The traits have varied over the years. In 2006, no veterinary trait was recorded and in 2008 »type« became the more general »type and temperament«. Both conformation and veterinary were split into two better defined traits in 2008, as was correctness of paces in 2009; in the analysis the two recordings were averaged to produce a single score. The trait definitions have become more precise over time. For instance, the 2009 definition for type and temperament in eventing was »Suitability of type and temperament to include attention, confidence, expression, efficiency of movement and harmony specifically in relation to intended discipline and relative to age. Is intelligent, willing and honest in his/her attitude to jumping, evaluates the jump and if he/she makes a mistake – seeks to tackle it differently next time. Should give the impression of being trainable«. The veterinary mark (for the same year and discipline) was awarded for the conformation of limbs, hooves, musculo-skeletal and other biomechanical factors which could influence future performance, and assessed the horse's potential to stay sound and free from injury. Good conformation required a rectangular proportional build with horizontal back and proportional legs, a supple poll and head/neck connection with clean throat latch, a long arched neck with muscling to top line, and a strongly built and muscled back and loin. Correctness of paces was judged individually for trot and walk. Walk was an active gait with impulsion and was supple, free in the shoulder and elbow and showed a noticeable over track. Trot was an active gait with impulsion and was supple with balance and self carriage, with a »rounder« action more than a »toe flick« and was rhythmical. Athleticism assessed the canter and gallop, and for three-year olds the jump was assessed.

Also included in the extracted data were the date and location of evaluation, date of birth, age, gender, colour of horse, owner, breeder and handler. Pedigree data were available with the sire, dam and dam's sire identified. Four year old horses were evaluated in 2007 only, and accounted for only 6 records.

Statistical methods

Phenotypic analysis

The distributions of scores from the 5 traits: 1) athleticism 2) conformation 3) correctness of paces 4) type and temperament and 5) veterinary were examined. All scores were greater than or equal to 4, with the exception of one value of 2.5. All other scores for the individual record with the extreme value ranged from 6.5 to 8.25, suggesting that there may be an error in recording, and to be cautious, the value was excluded from the analysis. Although the literature is sparse, there is some evidence examining whether traits measured in foals are the same trait as that measured in a three-year old horse. In the Trakehner, traits measured in foals and mares were found to be very similar, based on sire breeding values estimated from the two data sources (Preisinger *et al.* 1991). In the Holsteiner, high genetic correlations were found between mare studbook registration data and foal performance tests, although correlations between mare station tests and foal tests were low or moderate (Bösch *et al.* 2000).

As the dataset is relatively small at present, there were insufficient data to analyse each trait within each discipline. Therefore it was important to assess whether data could be pooled, assuming that the traits as measured in Futurity were genetically identical across disciplines. However, with so few horses evaluated in multiple disciplines, there were insufficient data to estimate genetic correlations across disciplines with sufficiently low standard errors. Correlations between horse effects for the three main disciplines were therefore estimated. The horse effect included all genetic and non-genetic effects, e.g. permanent environment, after having accounted for age, gender and temporary environment. This analysis used only horses with observations in multiple disciplines. High correlations indicate that the genetic correlations are also very likely to be high unless there was very small additive genetic variance for both traits, which is unlikely. This would provide a justification for the assumption that the traits may be genetically identical across disciplines. Bivariate models including all fixed effects and the random effect of the horse were used to assess correlations. Results (presented below) indicated that the assumption of genetic similarity across disciplines was tenable. The data were therefore pooled over disciplines with any mean differences being accounted for in the models by including discipline as a fixed effect.

Genetic analysis

Univariate analyses for each of the five traits were performed using an individual animal model in ASReml (Gilmour *et al.* 2006):

$$y = mean + dis + loc + sex + loc.date + \beta age + spl(age) + a + v + e \qquad (1)$$

The fixed effects were discipline (*dis*), location (*loc*), gender (*sex*) and location – date interaction (*loc.date*). β is the slope of the regression of the trait on age (*age*) at the time of the evaluation, however the effect of age was fitted as a spline with the smoothing parameter estimated as part of the random model as described in White *et al.* (1999). Random effects were the genetic effect of the horse (*a*), the permanent environment of the horse (*v*) and the residual error (*e*). The random effects *a, v* and *e* were assumed normally distributed with (co) variance matrices $\mathbf{\Sigma_a \otimes A}$, $\mathbf{\Sigma_v \otimes I}$, $\mathbf{\Sigma_e \otimes I}$.

Although the data included information on the owner and breeder, the recording of this data was not complete (53.8 % and 53.5 % recorded respectively). These effects were included in preliminary analyses as random effects however they reduced the size of the dataset substantially, and reduced the accuracy of the estimates, and in general, in these preliminary analyses, there was no evidence of effect of breeder/ owner. For this reason, the final model excluded the owner and breeder effects. Multivariate analyses were also performed between all pairs of traits, and the models included the same fixed and random effects as in the univariate analyses.

Functions of variance components were calculated as follows:

$$\begin{aligned}&\textit{phenotypic variance: } \sigma_p^2 = \sigma_a^2 + \sigma_v^2 + \sigma_e^2\\&\textit{heritability: } h^2 = \sigma_a^2 / \sigma_p^2\\&\textit{repeatability: } r^2 = (\sigma_a^2 + \sigma_v^2) / \sigma_p^2 \qquad (2)\end{aligned}$$

where σ_a^2 is the additive genetic variance, σ_v^2 is the horse permanent environmental variance (i.e. non-genetic) and σ_e^2 is the residual error variance for the trait.

Breeding values

The distribution of breeding values was examined for each trait. Breeding values were scaled by $20/\sqrt{\sigma_a^2}$, with a mean breeding value of 100. The range and reliabilities of breeding values were also examined, where the reliability was calculated as $1 - [se^2/\sigma_a^2]$, and *se* is the standard error of the breeding values presented in ASReml.

Results

Description of dataset

In total there were 1 887 evaluation records on 1 323 horses (Table 1). The number of evaluations increased throughout the years; 218 were recorded in 2006, 278 in 2007, 529 in 2008 and 862 in 2009. There were approximately equal numbers of evaluations for dressage and eventing (35.4 % and 34.9 % respectively), less for show jumping, sports ponies and endurance horses (19.6 %, 8.9 % and 1.3 % respectively). Due to the small numbers of records for ponies and endurance, results from this point onward are presented only for dressage, eventing and show jumping. The age range for horses at evaluation was 14 days-1 591 days. For horses attending an evaluation date (where it was potentially evaluated for multiple disciplines), 41 % were foals, 24 % yearlings, 21 % 2-year olds, 13 % 3-year olds and <1 % 4-year olds. Horses often had multiple evaluations in one discipline: for dressage, 10.2 % and 1.5 % of horses had 2 and 3 evaluations respectively. Figures for show jumping were 9.6 % and 2.2 %, and eventing were 15.5 % and 1.8 %. Horses were also evaluated in multiple disciplines. Of the horses evaluated for dressage (n=587) 14.7 % and 18.1 % were evaluated for show jumping and eventing respectively, and of the show jumping horses (n=323) 26.6 % and 31.0 % were evaluated for dressage and eventing.

Table 1
Summary of dataset by discipline for all years

Discipline	% of records	% of horses	No. locations	Age (days) min	Age (days) max
Dressage	35.4	44.3	14	14	1 591
Show jumping	19.6	24.3	15	14	1 572
Eventing	34.9	41.3	15	26	1 572
Sport ponies	8.9	10.6	15	34	1 527
Endurance	1.3	1.9	4	47	1 248
Total number	1 887	1 323			

The total number of sires and dams was 607, and 1 096 respectively, which were available for 99.7 % and 99.9 % of records, and 663 sires of dams were available representing 77.7 % of data. The mean progeny size of the half-sib groups was 2.17 (sires) and 1.21 (dams). The maximum sizes of descendant groups were 38, 4 and 11 for sires, dams and dams' sires respectively. Few of the evaluated horses were also sires (n=6) or dams (n=2) of others that were evaluated, although this is mainly due to the short time period that the data covered. A total of 71 sires were also dams' sires. The genetic connections in the dataset are expected to increase rapidly as data accumulate further.

Phenotypic parameters

Summary statistics for the scores for each of the three main disciplines are presented in Table 2. The mean score for each trait in each discipline was between 8.02 and 8.24. The standard deviation ranged between 0.54 and 0.83, and was highest for veterinary. The overall minimum in the data analysed was 4 and maximum 10, indicating that the lower range of the 10 point scale is not used, probably due to pre-selection of horses by owners and the guidelines which attempt to define scoring in absolute terms. Distributions were approximately normal with some slight negative skewness, particularly for veterinary. Histograms of the distributions of scores are presented in Figure 1.

Table 2
Summary statistics for variable scores for each trait for the three main disciplines

Trait	Discipline	N	Mean	SD	Min	Max
Athleticism	D	665	8.02	0.67	6.00	10.00
	SJ	368	8.12	0.78	5.00	10.00
	EV	657	8.05	0.73	4.00	9.75
Conformation	D	668	8.09	0.56	5.50	9.75
	SJ	369	8.14	0.61	6.00	9.62
	EV	658	8.11	0.56	6.00	9.50
Correctness of paces	D	668	8.08	0.57	4.50	10.00
	SJ	367	8.02	0.76	4.00	9.75
	EV	658	8.05	0.54	5.00	9.75
Type & temperament	D	667	8.14	0.60	6.00	10.00
	SJ	367	8.22	0.68	6.00	10.00
	EV	656	8.18	0.63	4.00	10.00
Veterinary	D	601	8.14	0.82	4.00	9.75
	SJ	336	8.24	0.81	4.50	9.75
	EV	560	8.18	0.83	4.00	10.00

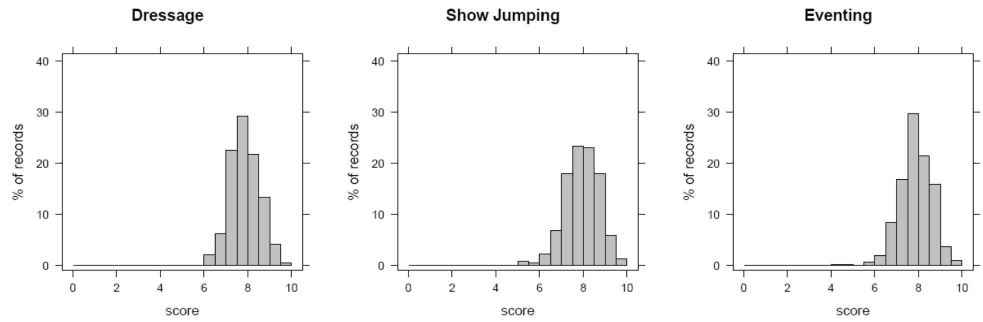
D: dressage, SJ: show jumping, EV: eventing

In the analysis conducted to examine the feasibility of pooling test data across disciplines, the correlations between the three main disciplines for the random effect of the horse were not significantly different from 1 for any pair of traits (Table 3). In the absence of being able to assess genetic correlations between traits directly (due to lack of data), and the magnitude of heritabilities found (reported below), the genetic correlations were also assumed to be high, and it was considered appropriate to pool the data across disciplines.

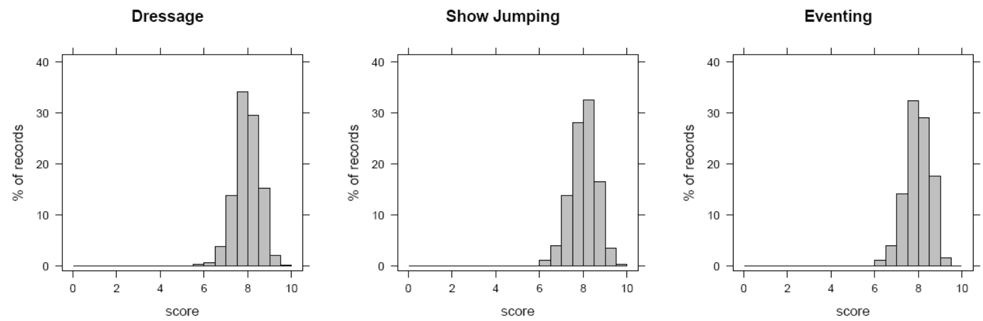
Table 3
Correlations between disciplines for the overall horse effect (additive genetics & permanent environment)

Trait		Discipline	
		SJ	EV
Athleticism	D	0.960	0.996
	SJ		0.990
Conformation	D	0.990	0.990
	SJ		0.986
Correctness of paces	D	0.956	0.975
	SJ		0.960
Type & temperament	D	0.971	0.971
	SJ		0.995
Veterinary	D	0.996	0.990
	SJ		0.999

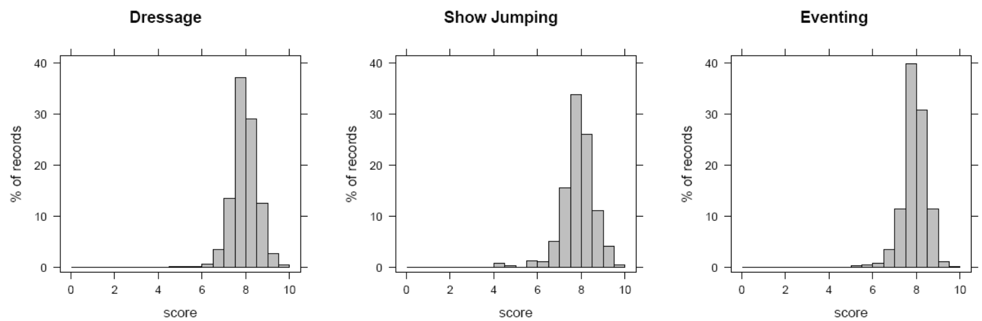
D: dressage, SJ: show jumping, EV: eventing, Analyses were based on 86, 104 and 99/98 horses for D-SJ, D-EV and SJ-EV respectively, except for veterinary, where numbers were 82, 98 and 95 respectively.



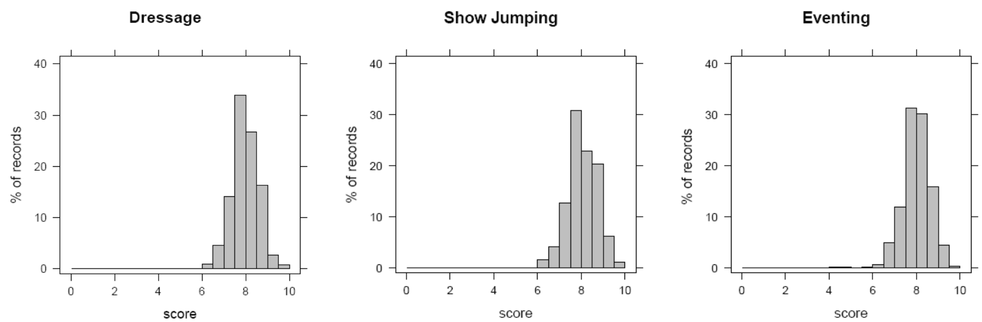
Athleticism



Conformation



Correctness of paces



Type & temperament

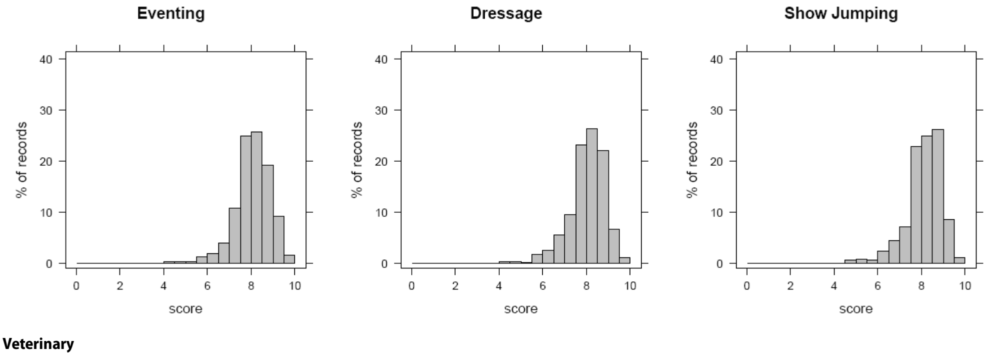


Figure 1
Distributions of scores for each trait, for dressage, show jumping and eventing disciplines

Genetic parameters for single traits

Heritabilities, repeatabilities and the proportion of variance due to the horse's permanent environment for each trait are shown in Table 4. All were significant ($P<0.05$) with the exception of permanent environment for type and temperament. Heritabilities ranged from 0.20 to 0.42, and were highest for type and temperament, and lowest for athleticism. The ratio of the variance due to the horse's permanent environment to the total phenotypic variance ranged between approximately 0.03 (not significant $P>0.05$) for type and temperament and 0.52 for veterinary. The repeatability, representing the total variance due to the horse, including both additive genetic and permanent environmental variance, was lowest for type and temperament and athleticism and highest for veterinary. The residual error – unexplained variance in scoring – may include the effects of the handler, owner and breeder, although some of the longer term influences of the latter may also be in the permanent environment.

Table 4
Heritability, repeatability and horse's permanent environment as a percentage of total phenotypic variance, for each trait, with standard errors in parentheses

Trait	Additive genetics	Horse's permanent environment	Residual error	Repeatability	Total phenotypic variance
Athleticism	20.34 (8.28)*	25.63 (8.58)**	54.03 (3.39)***	45.97 (3.39)***	0.46
Conformation	29.05 (9.36)**	35.09 (9.38)***	35.86 (2.46)***	64.14 (2.46)***	0.30
Correctness of paces	30.42 (9.28)**	24.63 (9.32)**	44.95 (2.95)***	55.05 (2.95)***	0.33
Type & temperament	42.16 (9.70)***	2.80 (9.49)	55.04 (3.47)***	44.96 (3.47)***	0.37
Veterinary	24.95 (11.01)*	51.64 (10.99)***	23.41 (1.82)***	76.59 (1.82)***	0.48

*$P<0.05$, **$P<0.01$, ***$P<0.001$

Fixed effects of gender, discipline and age

Gender effects are given in Table 5. In general geldings scored lower than stallions, although differences were only significant for veterinary and athleticism ($P<0.01$). These differences are partly due to the fact that the best males are kept as stallions. The mean score awarded also differed between disciplines (Table 6). In general, the traits were scored more harshly for dressage: for instance, in type and temperament, show jumping and eventing scored higher than dressage ($P<0.01$). Show jumping was scored lower than eventing for correctness of paces ($P<0.01$), but there were no other significant differences between the two disciplines.

Table 5
Effect of gender on scores awarded

Trait	Gender	Stallion (0 comparison)	Gelding (0 comparison)
Athleticism	stallion		
	gelding	−0.13**	
	mare	−0.02	0.11**
Conformation	stallion		
	gelding	0.01	
	mare	0.04	0.03
Correctness of paces	stallion		
	gelding	−0.08	
	mare	0.02	0.10**
Type & temperament	stallion		
	gelding	−0.08	
	mare	−0.02	0.06
Veterinary	stallion		
	gelding	−0.18**	
	mare	−0.11**	0.06

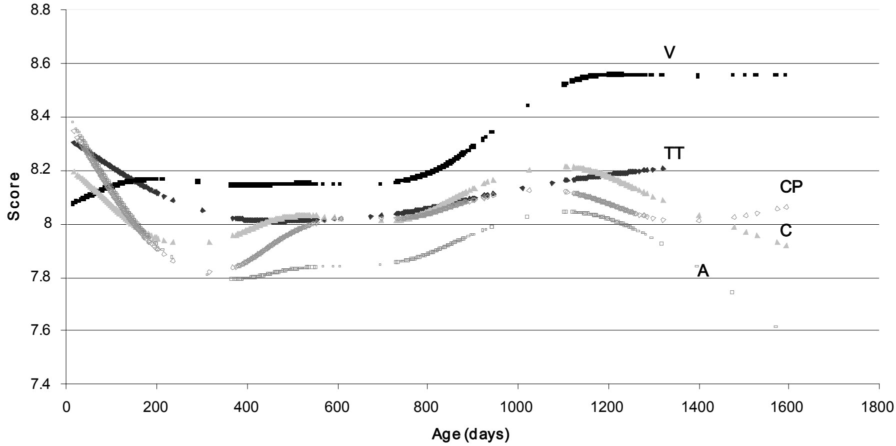
*$P<0.05$, **$P<0.01$, ***$P<0.001$

Table 6
Effect of discipline on scores awarded

Trait		Discipline	
		D (0 comparison)	SJ (0 comparison)
Athleticism	SJ	0.10*	
	EV	0.13***	0.03
Conformation	SJ	0.04	
	EV	0.07*	0.03
Correctness of paces	SJ	−0.04	
	EV	0.07*	0.10**
Type & temperament	SJ	0.11**	
	EV	0.14***	0.02
Veterinary	SJ	0.04	
	EV	0.01	−0.03

D: dressage, SJ: show jumping, EV: eventing, *$P<0.05$, **$P<0.01$, ***$P<0.001$

Results for the effect of age on the scores for each of the traits are shown in Figure 2. In general, with the exception of veterinary scoring, the youngest horses performed relatively well. Scores then declined with age until the horse was a yearling, and then rose again as the horse matured. This implies that scores tend to increase over age, as the horse matures and so temperament, paces etc. improve. However very young horses are either difficult to assess, or tend to be given higher scores through subjective bias towards foals. For veterinary, the foals were scored lowest, and overall scores increased with age. For all traits, high scores were seen in three year olds. Such a trend may reflect improvement of traits, but there are other explanations including selection in submitting for an evaluation. Judges may also be better at discriminating traits in more mature horses. As horses may have been evaluated on several occasions, any improvement in scoring due to experience would give a false indication of an age related change. This is particularly relevant to type and temperament where experience is very likely to improve scoring.



A: athleticism (open medium grey squares), C: conformation (solid light grey triangles), CP: correctness of paces (open medium-light grey diamonds), TT: type and temperament (solid dark grey diamonds), V: veterinary (solid black squares)

Figure 2
Influence of age on scoring for each trait

Genetic correlations between traits

The genetic correlations between traits were generally high and all were significantly different from zero, where standard errors were determined (*P*<0.05) (standard errors were not determined for athleticism – type and temperament, athleticism – correctness of paces and type and temperament – correctness of paces). Apart from correctness of paces and veterinary, and correctness of paces and conformation, they were not significantly different from 1.0 (Table 7). Correlations between traits other than veterinary were particularly high. The magnitude of these correlations indicates that the same genes are largely responsible for all traits.

Table 7
Genetic correlations (standard error) between traits

Trait	Conformation	Correctness of paces	Type & temperament	Veterinary
Athleticism	0.94 (0.09)	1.00[a]	1.00[a]	0.57 (0.23)
Conformation		0.77 (0.11)**	0.97 (0.05)	0.83 (0.15)
Correctness of paces			1.00[a]	0.44 (0.22)*
Type & temperament				0.80 (0.17)

Significant difference from 1: *P<0.05, **P<0.01, [a]standard errors could not be determined in the analysis

Permanent environment correlations between traits

Correlations between traits for the effect of the permanent environment were significant (P<0.05) and moderate between correctness of paces and all other traits, and the correlation between type and temperament and athleticism was moderate and significant (Table 8). No significant correlations were detected between the other pairs of traits. Thus, compared to the genetic correlations they were smaller and not all significantly different from zero. This indicates that the same aspects of the horses' permanent environment, such as early life influences, stabling, and long-term training and long-term nutritional effects, jointly influence some of the pairs of traits. Also included in the permanent environment effect in this analysis will be maternal effects and non-additive genetic effects.

Table 8
Permanent environment correlations (standard error) between traits

Trait	Conformation	Correctness of paces	Type & temperament	Veterinary
Athleticism	0.37 (0.19)	0.69 (0.05)***	0.58 (0.12)***	0.24 (0.18)
Conformation		0.65 (0.13)***	0.60 (0.34)	0.24 (0.16)
Correctness of paces			0.46 (0.13)***	0.43 (0.18)*
Type & temperament				−0.02 (0.59)

Significant difference from 1: *P<0.05, **P<0.01, ***P<0.01

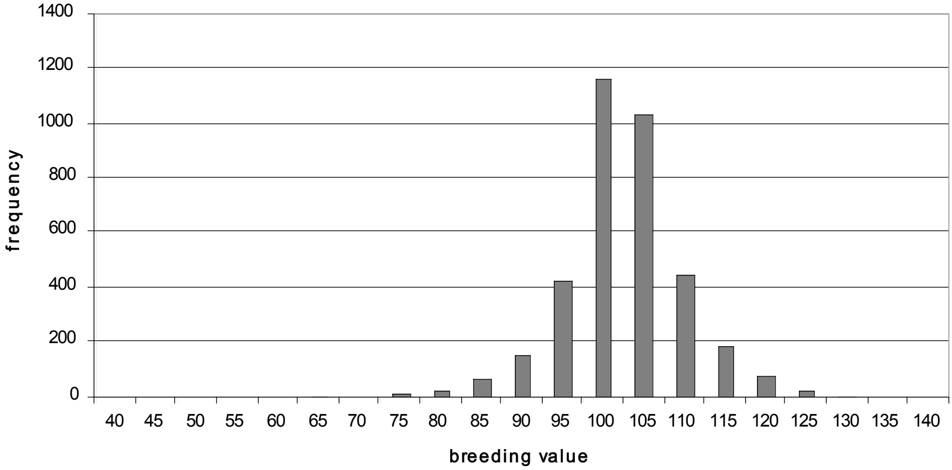
Breeding values

The distribution of breeding values, for all animals in the pedigree, for the examples of athleticism and conformation are given in Figure 3. The highest reliabilities for breeding values were 0.81 (type and temperament), 0.74 (correctness of paces), 0.72 (conformation), and 0.66 (veterinary and athleticism). For publication, industry has decided that only breeding values with a minimum reliability of 30 % will be produced. In these data 39.3 % of horses would quality for type and temperament, 31.9 % for correctness of paces, 37.5 % for conformation, 10.7 % for veterinary and 8.4 % for athleticism.

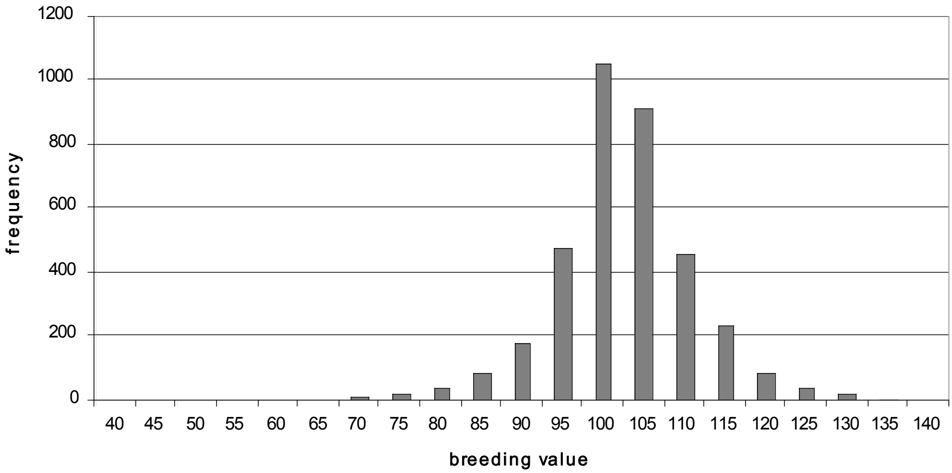
Discussion

The young horse tests (»Futurity«) recently introduced in the UK are a valuable data source for genetic evaluations. Participation has grown substantially since the introduction of the scheme, and it is envisaged that the data will accumulate rapidly. The most appropriate strategy will then be to combine young horse data with adult competition data to routinely estimate breeding values. This study characterised the Futurity dataset to date and performed the first genetic evaluations on traits recorded on young horses in the UK.

To date, genetic evaluations in the UK have been conducted as a research exercise for dressage and eventing using competition data (Kearsley *et al.* 2008, Stewart *et al.* 2010), as large competition datasets exist. In general, both young horse tests and adult competition data have advantages and caveats, and now most international evaluations use a combination of both data sources. Competition data may be available on a greater proportion of the



Athleticism



Conformation

Figure 3
Distribution of breeding values for athleticism and conformation

population however it becomes available late in life and is generally biased in that the horse may compete in only one discipline. Young horse data are obtained much earlier in life, which enables a shorter generation interval thus providing opportunity for greater genetic gain. To be useful, young horse data generally require moderate to high heritabilities for the traits, and high genetic correlations with later competition performance.

Selection in the horses evaluated at young horse tests is present in most evaluations internationally. The effect of this will be to bias heritability estimates downwards. In the UK, an estimated 10 % of the 2009 sport horse foal crop was evaluated in the Futurity scheme of that year (Graham Suggett, personal communication). The actual proportion of young horses evaluated at any age is therefore at least 10 % and possibly at much as 30 %. This is in line with the percentages of young horses tested internationally, which range from 10 % to 45 % of the (registered) foals (Thorén Hellsten *et al.* 2006). The Futurity is therefore capturing a comparable proportion of young horses to other international young horse tests, while only in its third year and participation is still growing. Selection in the horses presented is illustrated in the lack of scores in the lower range of the scale. There is anecdotal evidence that the degree of selection has increased over the years (J. Rogers personal communication), as owners become more aware of the characteristics required. A higher proportion of youngstock from stallions with high stud fees tend to be entered (J. Rogers personal communication), and from these stud farms, it is likely that only a limited number of foals are evaluated, suggesting that selection is of the gene pool identified as being more elite. There may be other biases in the horses that are evaluated, for instance related to the fact that participation tends to come from the same owners, and there are limitations in presenting horses for evaluation including the cost of evaluation, and practical resources such as travel costs involved in transporting a horse to the venue.

A team of judges evaluate the traits. In the current analysis, the variance due to the judging teams will be included with the temporary environmental variance (location–date interaction). Training prior to evaluations ensures consistency and teams vary as individual judges rotate between different teams (based on practical issues rather than random assignment or specific groupings). In the years covering the data collection, scoring was done by discussion within the team, which may have led to problems, for instance a more dominant judge may exert an undue influence on the scoring. As of 2011, scoring will be performed by judges individually, and then averaged, which will allow the repeatability of scoring to be evaluated.

The UK tests are distinct in a number of ways from any others internationally. One difference is the ages being evaluated – UK tests cover foals up to 3 (4) year olds, whereas international tests range from 3 to 7 year olds (Thorén Hellsten *et al.* 2006). UK tests also evaluate for only one discipline, although disciplines include eventing, endurance and ponies, whereas international tests often evaluate one horse for both show jumping and dressage. Internationally, eventing evaluations are uncommon. The duration of testing is also relatively short compared to other national tests. UK tests are most comparable to field tests, rather than the station tests or young horse competitions that take place in some other countries. Station tests generally have higher heritabilities compared to field tests, at least in part due to the control of environmental factors (Ricard *et al.* 2000). The longer length of station tests may also produce higher heritability estimates (Dietl *et al.* 2005). Due to the age constraints there is no testing under saddle in the UK whereas tests elsewhere are either ridden or at hand; tests performed without the rider have higher heritabilities than those with the rider, due to the lack of rider variance (Ricard *et al.* 2000). At present, all genders are eligible to participate, without restrictions. However, as of 2012, the eligibility of stallions may change, so that only graded stallions (with studbooks) are eligible, similar to practices on the continent. This change to the rules will introduce bias to the genetic evaluation system, producing downwardly biased heritability estimates due to the restricted genetic pool and thus lower genetic variance. In many foreign evaluations, the breeding goals include subjective terms (Koenen & Aldridge 2002). This is also true of traits assessed in the Futurity, illustrated for example, by the descriptions of »confidence«, »expression«, »harmony« and »honest« within the description of traits. Such subjectivity may result in an increase in phenotypic variance and a reduction in both repeatability and heritability estimates.

This analysis was performed on a relatively small dataset, due to which we had to address strategies of analysing small datasets. Phenotypic correlations between disciplines were estimated. This indicated that genetic correlations were consistent with unity, and justified to an extent the pooling of traits across disciplines in the analysis. The assumption of equal traits across disciplines should be further confirmed when more data are available (however, as horses are now evaluated for only one discipline at any one event, suitable data may take time to accumulate). In contrast to these high estimates, current foreign estimates for genetic correlations between dressage and show jumping related traits from young horse data are mixed (low and either positive or negative), although correlations between canter and jumping tend to be higher than correlations between the other gaits and jumping (Thorén Hellsten *et al.* 2006). If the genetic correlations are indeed higher in the UK then this may well be due to the lower age of testing, when differentiation between disciplines has not yet occurred. There may be various advantages and disadvantages to testing at such a young age. The influence of selection due to training and competing in young horse competitions is reduced, and the earlier the data become available for genetic evaluation, the greater the potential for genetic progress. However, it is more difficult to assess future ability in very young horses, which are still developing rapidly.

Perhaps the most comparable test internationally is that of the Swedish Young Horse Tests (YHT) for 3 and 4 year olds. In these tests, the heritability of total conformation was 0.58, which is twice the estimate of 0.29 from the current study. The heritability of type in the Swedish tests was 0.30, which compares with a higher estimate of 0.42 in the current study. Heritabilities for walk and trot at hand and for free canter were 0.37, 0.45 and 0.37 respectively, i.e. slightly higher than the 0.30 for correctness of paces in this study (Viklund *et al.* 2008). In other countries overall, the average estimates from field performance tests were 0.30, 0.35 and 0.28 for walk, trot and canter, i.e. very close to the 0.30 estimate for correctness of paces in the current study (Thorén Hellsten *et al.* 2006). Thus although there are some differences between estimates in the literature and the current study, other estimates are very close and are of a similar magnitude overall.

The heritability was highest for type and temperament. Type, reflecting physical characteristics, might be expected to be very influenced by genetics. Temperament might be expected to be influenced by both temporary and permanent environment, including early life influences and training. However, no significant variance due to the horse's permanent environment was detected for type and temperament. Heritability estimates for temperament are scarce in the literature. Oki *et al.* (2007) estimated the heritability of temperament in the thoroughbred as 0.23. For young horse tests in Dutch warmbloods, heritability of character was estimated at 0.52 for station tests and 0.06 for field tests (Huizinga *et al.* 1990, Huizinga *et al.* 1991). In the RHQT (1988-2003) of the Swedish warmblood, heritabilities of temperament for jumping and gaits were 0.17 and 0.41 respectively, and for the YHT, temperament for jumping was 0.23 (Viklund *et al.* 2008).

For use in genetic evaluations, young horse trait data must show moderate/ high heritabilities and high genetic correlations with later competition performance. The next requirement in these studies is to investigate the genetic correlations between the young horse data and adult competition data. Internationally, estimates for genetic correlations between young horse tests and adult competition are generally positive and moderate/high indicating that the young horse tests are effective in selecting for adult competition performance (Thorén Hellsten *et al.* 2006). For example Viklund *et al.* (2010) estimated correlations between competition data and traits recorded in the Swedish RHQT, and found the correlation between show jumping competition success and the jumping traits was 0.87 to 0.89 (depending on the competition performance trait) and between dressage competition success and the average for gaits was 0.68 to 0.75.

Pedigree data/recording are often a major limitation in the genetic evaluation of horses. The pedigree available in this analysis was shallow (including only sire, dam and dam's sire). A greater depth of pedigree, which should accumulate over time, would provide more accurate estimates of the genetic variance. The young horse data were recorded comparatively well, however horses are identified by name only. The use of a unique identifier, such as the Universal Equine Life Number or passport number would greatly add to the integrity of the data. The recording of the pedigree data is of particular importance for the purposes of genetic evaluations and this is an issue that will escalate over time unless resolved. This analysis was performed on the first four years of data and was therefore based on a relatively small dataset. As the data accumulate over the years, the accuracy of estimates and thus extent of interpretations that can be drawn will increase.

In conclusion, the young horse tests introduced in GB have been successful and participation continues to grow. Equal numbers of evaluations are for dressage and cross-country, with less in show jumping. The data are a potentially valuable data source for genetic evaluations. The magnitudes of heritabilities indicate that, if genetic correlations between scored traits and later competition performance are high, the data will be useful in selecting for later competition performance. At present, the reliability of breeding values is low to moderate but reliabilities will increase rapidly with additional data. The young horse data could be used to predict breeding values for the traits recorded as part of the scheme although best use of the data would be made by combining it with adult competition data for use in the prediction of breeding values for competition.

Acknowledgements

The authors are grateful for financial support from BBSRC, British Equestrian Federation and Biosciences KTN, and for data and collaboration from British Equestrian Federation.

References

Bösch M, Reinecke S, Röhe R, Kalm E (2000) Genetic analysis of traits from Holsteiner riding horses - Variance components for performance traits of foals on field test, mares at studbook registration and mares on field and station test. Züchtungsk 72, 161-171 [in German]

Dietl G, Hoffmann S, Reinsch N (2005) Impact of trainer and judges in the mare performance test of Warmblood Horses. Arch Tierz 48, 113-120

Gilmour AR, Gogel BJ, Cullis BR, Thompson R (2006) ASReml User Guide Release 2.0 VSN International Ltd, Hemel Hempstead, HP1 1ES, UK

Huizinga HA, Boukamp M, Smolders G (1990) Estimated parameters of field performance testing of mares from the Dutch Warmblood riding horse population. Livest Prod Sci 26, 291-299

Huizinga HA, van der Werf JHJ, Korver S, van der Meij GJW (1991) Stationary performance testing of stallions from the Dutch Warmblood riding horse population: 1. Estimated genetic parameters of scored traits and the genetic relation with dressage and jumping competition from offspring of breeding stallions. Livest Prod Sci 27, 231-244

Kearsley C (2007) Genetic Evaluation of Sport Horses in Britain. PhD thesis, University of Edinburgh, UK

Kearsley CGS, Woolliams JA, Coffey MP, Brotherstone S (2008) Use of competition data for genetic evaluations of eventing horses in Britain: Analysis of the dressage, showjumping and cross country phases of eventing competition. Livest Sci 118, 72-81

Koenen EPC, Aldridge LI (2002) Testing and genetic evaluation of sport horses in an international perspective. 7th World Congress on Genetics Applied to Livestock Production, Montpellier, 19th-23rd August, France

Oki H, Kusunose R, Nakaoka H, Nishiura A, Miyake T, Sasaki Y (2007) Estimation of heritability and genetic correlation for behavioural responses by Gibbs sampling in the Thoroughbred racehorse. J Anim Breed Genet 124, 185-191

Preisinger R, Wilkens J, Kalm E (1991) Estimation of genetic parameters and breeding values for conformation traits for foals and mares in the Trakehner population and their practical implication. Livest Prod Sci 29, 77-86

Ricard A, Bruns E, Cunningham EP (2000) Genetics of performance traits. In: Bowling AT, Ruvinsky A (eds.) The genetics of the horse. CABI Publishing, Wallingford, UK, 411-438

Stewart ID, Woolliams JA, Brotherstone S (2010) Genetic evaluation of horses for performance in dressage competitions in Great Britain. Livest Sci 128, 36-45

Thorén Hellsten E, Viklund Å, Koenen EPC, Ricard A, Bruns E, Philipsson J (2006) Review of genetic parameters estimated at stallion and young horse performance tests and their correlations with later results in dressage and show-jumping competition. Livest Sci 103, 1-12

Viklund Å, Thorén Hellsten E, Näsholm A, Strandberg E, Philipsson J (2008) Genetic parameters for traits evaluated at field tests of 3- and 4-year-old Swedish Warmblood horses. Animal 2, 1832-1841

Viklund Å, Braam Å, Näsholm A, Strandberg E, Philipsson J (2010) Genetic variation in competition traits at different ages and time periods and correlations with traits at field tests of 4-year-old Swedish Warmblood horses. Animal 4, 682- 691

White IMS, Thompson R, Brotherstone S (1999) Genetic and environmental smoothing of lactation curves with cubic splines. J Dairy Sci 82, 632-638

Received 15 March 2011, accepted 22 September 2011.

Corresponding author:

Isobel D. Stewart
email: i.d.stewart@sms.ed.ac.uk

Institute of Evolutionary Biology, University of Edinburgh, Kings Buildings, West Mains Road, Edinburgh, EH9 3JF, UK