Analysis of the population of competing British sport horses as measured by lifetime performance: structural distributions at different levels of competition

TC Whitaker1,* and J Hill2
1Faculty of Applied Science and Technology, Writtle College, Chelmsford CM1 3RR, UK
2Institute of Land and Food Resources, University of Melbourne, Victoria 3010, Australia
*Corresponding author: tim.whitaker@writtle.ac.uk

Submitted 20 July 2004: Accepted 18 October 2004
Research Paper

Abstract
A preliminary investigation into the structure of the competing sport horse population within the UK for the year 2000 was undertaken, comprising 30,974 horses and 39,679 individual performances. All performance data were expressed as total lifetime winnings. The level of non-winning horses, expressed as a percentage, was high in all disciplines: dressage 22%, eventing 50% and show jumping 23%. The percentage of horses competing in disciplines at the highest recognized level was small: dressage 2.5%, eventing 10% and show jumping 6.5%. Highest earners for all disciplines all earned considerably more than the threshold for highest performance level in their respective discipline (£2.2 dressage, £25 eventing, £365 show jumping). Data for all three disciplines were highly skewed. Structural analysis of the distributions by level showed a highly skewed hypogeometric distribution at the lowest level of competition for all disciplines, uniform distribution at intermediate levels (bimodal tendencies were observed within dressage) and normal distribution at advanced levels. Product–moment correlations showed a significant correlation between show jumping and eventing (0.228, P < 0.01). Problems with data availability and presentation may mean that up to 12.9% of the population may have had a confounding effect on the results of this study.

Keywords: sports horse; population; Britain; competing

Introduction
Genetic evaluation and breed improvement systems have been applied in many European countries to the production of sports horse populations1. However, the UK equine industry has not applied or developed the methods of selective breeding and breed improvement schemes implemented by its European neighbours. This has resulted in the UK consistently failing to breed sport horses of international class. This is demonstrated by the sire rankings of the World Breeding Federation for Sport Horses. Between 1990 and 1999, not one UK-bred show jumping or dressage sire featured in the top 20 sires2. The situation is slightly better in eventing, with three sires featuring in the top 20 over this period, although the UK is still outperformed by both New Zealand and Ireland2. It is not clear how this situation has occurred. However, diversity in the type of horses used in the various sporting disciplines (dressage, eventing and show jumping) as a result of differing perceptions of competitors, breeders and breeding societies may be a causal factor. Lack of specificity in breeding objective has already been identified by a number of authors3,4 as a major contributory factor in sub-optimal breed improvement. Small populations with non-random matings, long generation intervals, performance recognition at advanced age and high number of colts gelded at a young age are problems affecting the successful production of sport horses5. It is also important to note that there is a strong likelihood that much of
the genetic potential of horses evaluated using performance data is moderated by other ‘environmental’ factors, for example the ability of the rider and the motivation and aspiration of the competitor. It is therefore suggested that, before any formal genetic analysis can be conducted, preliminary analysis of performance data be undertaken with the aim of identifying the structural attributes of the data collated (including the occurrence of missing data).

From the UK perspective, the inclusion of the three disciplines (dressage, show jumping and eventing) in the strategy for sport horse breeding is important. The three disciplines are run unilaterally to national level by three self-governing bodies, the British Show Jumping Association, British Eventing and British Dressage. Each governing body is responsible for recording performance data for all animals competing at national level in its discipline. Until 2000, all data were then collated by the British Horse Database (BHD)\(^5\). This is not an ideal situation but may assist in the development of breed improvement schemes. The present paper reports an evaluation of the UK sport horse population using data from the BHD and highlights the structure of designated sub-populations defined by performance that may be important in the development of breed improvement schemes. A preliminary investigation into sire registration details is also undertaken. Additionally, comment is made on the historical change in number of horses registered for the different disciplines.

**Materials and method**

Approximately 75000 horses had a BHD registration number by the year 2000\(^5\). The database used in this preliminary analysis of UK sport horses was from horses competing in 2000. Data handling problems and data availability resulted in only those horses competing in 2000 being analysed within this study. This population comprised 30974 horses that had a total of 39679 lifetime performances\(^9\). Analysis of registration by discipline (dressage, eventing or show jumping) showed 32766 horse registrations in 2000\(^5\). Further analysis of how the data were presented showed a confounding effect. It was possible for a horse to be registered to compete in a specific discipline in 2000 (e.g. dressage) and to have its lifetime performance in another discipline (e.g. eventing) also recorded, even through they were not registered for that discipline (eventing) in 2000. It was not possible to remove these horses from the data analysis. Neither was it possible or within the scope of this paper to analyse or speculate on the reasons for approximately 44000 horses not competing. Such records were classified as missing at random (MAR).

List-wise deletion of records identified as MAR was done according to the approach of Little and Rubin\(^6\).

The dataset was subdivided by performance into dressage, eventing and show jumping. The classification of prize fund distribution is fairly detailed and dependent on the level of competition performed at, and it is beyond the scope of the current paper to review this in full. Briefly, in all disciplines points or prize money will be amassed over a horse’s lifetime to give a lifetime performance for the horse. The data for a horse’s performance in the dressage discipline were points awarded for the final percentage score accredited to the horse within competition, and were not adjusted for level of competition\(^7\). The data collated from eventing were points achieved in competition with points being awarded in relation to level of competition and final finishing position within that competition\(^8\). Performance during show jumping was evaluated as financial return (£ won). Financial return in show jumping takes into account the level of competition and position attained in competition\(^9\).

All performance data reported for competing horses were recorded as total lifetime winnings for the animal. For each discipline, a single score (either points won or monetary winnings) was recorded for that animal’s total lifetime performance. Each population of horses competing (as defined by discipline) was subdivided by level of ability as follows:

- **Dressage**: preliminary (0–37 points), novice (38–74 points), elementary (75–149 points), medium (150–249 points), advanced-medium (250–324 points) and advanced (in excess of 324 points);
- **Eventing**: novice (0–20 points), intermediate (21–60 points) and advanced (in excess of 60 points);
- **Show jumping**: grades C (£0–799), B (£800–1799) and A (in excess of £1799).

These individual sub-populations, as well as the complete populations of horses performing in dressage, eventing and show jumping, were described using central tendency theory (mean, median, standard deviation (SD) and coefficient of variation). Measures of normality using skewness and kurtosis tests\(^10\) for each population of horses performing in dressage, eventing and show jumping were also determined. Bivariate product-moment correlations\(^10\) were calculated between points achieved at dressage, eventing and financial returns from show jumping of individual animals performing. The three disciplines were assessed after transformation of the data (log score + 1) according to the method of Silvestrelli *et al.*\(^11\).

Registry of horses by discipline was examined for the year 2000. Further retrospective analysis was
conducted with registry in preceding years 1997–1999\textsuperscript{12–14} and 1988\textsuperscript{15}. The years 1997–2000 represented the life span of the BHD Sports Horse Annual and offered complete and accurate records of registry by discipline. Comparative analysis of registry of horses in 1988 was conducted on the basis of the publication of the Peat Marwick McLintock Report\textsuperscript{15}, the last comprehensive investigation into the equine industry. Analysis of sire details was investigated within the limitations imposed by the manner of data presentation and availability. Sire registration was examined. Additionally, the number of sires offering artificial insemination was determined. A sample of 2\% (n = 618) horses competing was selected by random sampling (random number generation) to determine the incidence of registered sires. This potentially may allow an evaluation of records of sires for completeness and the possible use of records for further analysis for the implementation of breeding schemes. All statistical processing was conducted using SPSS version 10\textsuperscript{16}.

### Results

The population of horses selected was 30 974, comprising 39 679 lifetime performances at affiliated competitions (dressage, eventing, show jumping). The studied population did not include any animals that were not registered to compete in 2000. Descriptive analysis of horses by discipline (show jumping, eventing and dressage) is presented in Table 1. Of the 8705 performances additional to total horses (n = 30 974), it was possible to classify 1792 as being from dual- or triple-registered horses (n = 32 766). The population did, however, include horses that might not have competed in a discipline for which they had been registered in previous years but not in 2000; their inclusion was due to their registration to compete in another discipline during 2000. This amounted to 5121 performances, 12.9\% of the population.

The total number of horses registered for all disciplines increased by 24.5\% over the period 1988–2000. From 1997 to 2000, the annual increase in registrations for all disciplines was 3.6\%. The number of horses registered for show jumping during the period 1988–2000 increased by 9.3\%, whereas the numbers of registrations in eventing and dressage increased more rapidly (54.0\% and 57.9\%, respectively) during the same time. However, the proportion of horses registered for each discipline (period 1997–2000) remained relatively stable (1997 \emph{vs.} 2000, respectively: dressage, 18.1\% \emph{vs.} 20.5\%; eventing, 25.7\% \emph{vs.} 24.7\%; show jumping, 56.2\% \emph{vs.} 54.8\%). The total number of horses registered for show jumping declined by 10\% during the period 1997–2000, whereas total numbers of animals registered for eventing and dressage increased by 5.2\% and 5.9\%, respectively.

Population distribution structures for the three individual performing populations based on the level attained for lifetime winnings (points for dressage and eventing, monetary for show jumping) are given in Table 2. The sub-populations of animals performing in dressage at different levels of competition show considerable variations in their individual structure. At dressage preliminary level (n = 5758 horses), the geometric mean score was 0.850 (SD 0.69). The overall structure (Fig. 1) of the sub-population was hypergeometric (median score at preliminary level was 0), with 31.4\% of the horses competing failing to score any points. At novice level (Fig. 2), the geometric mean was 2.016 (SD 0.09; n = 1284) and modelled a uniform distribution with boundaries of scores ranging from 75.8 to 149.6. At medium level (Fig. 3) of competition, a comparable low level of variation was observed with geometric mean of 2.279 (SD 0.07; n = 638). However, there was a distinct bimodal structure to the sub-population with scores clustered around 154 (geometric mean of 2.188) and 237 (geometric mean of 2.375) points. The number of horses competing at advanced–medium level dressage events was 231. No distinct distribution structure (Fig. 4) could be modelled in the data; however, a low level of variation was observed around the geometric mean of 2.458 (SD 0.04). At the highest level of competition (advanced; n = 175), a normal distribution of scores (Fig. 5) was observed with a geometric mean of 2.610 (SD 0.06).

In the case of eventing, horses competing at novice level (Fig. 6) performed with a mean score of 1.9 (geometric mean of 0.290). There was, however,
considerable variation (SD 2.8 points) and 65.9% of the horses competing at this level failed to attain any points. At novice level, the performance of horses competing was heavily skewed and modelled a hypergeometric distribution with a median score of 0. Performance of horses at intermediate level \((n = 1232)\) (Fig. 7) was consistent (uniform distribution) over a range of scores from 22.4 to 63.1 (geometric mean 1.530, SD 0.13). However, at advanced level (Fig. 8),
the majority of horses (57.5%) competing \((n = 956)\) achieved less than 100 points. The geometric mean score of the negative exponential distribution was 2.110 (SD 0.29).

Approximately a quarter (26.5%) of horses competing \((5189\) of \(19,563)\) at grade C show jumping failed to earn any money at any competition. The geometric mean was 1.350 (SD 0.94) and the structure of the sub-population was modelled to a hypergeometric distribution (Fig. 8). A uniform distribution (Fig. 9) bounded by earnings of £795–1778 with a mean earnings of £1140 (geometric mean 3.057, SD 0.10, \(n = 1266\)) was observed at grade B of competition (Fig. 10). A negative skewed distribution was observed at grade A show jumping (Fig. 11). The average earnings were £4571 (geometric mean 3.660, SD 0.40, \(n = 1449\)); however, 58.7% of the population of horses did not earn more than £5500. The maximum earnings at advanced level was £658,309.
A significant ($P < 0.01$) positive correlation accounting for 29% of variation in performance was observed between show jumping and eventing (Table 3). However, no relationships between performance at show jumping and dressage, or at eventing and dressage, could be identified.

The sample of 618 horses (2% of the whole population) selected at random suggested that 67% were sired by registered stallions and 33% had unrecorded parentage. A total of 797 stallions were registered on the British Horse Database; of these, 67.6% ($n = 539$) have at least one competing sib. The four largest registries of stallions were Weatherbys Non-Thoroughbred Register (17%), Thoroughbred General Stud Book (16%), Anglo European Stud Book (16%) and Sport Horse Breeders of Great Britain (16%). Fourteen per cent of stallions were registered with breeding societies/associations that can be regarded as non-sport horse in nature. Only 0.8% ($n = 6$) appeared to have no form of formal registration. A total of 118 stallions are currently declared as offering artificial insemination, 14.8% of the population of registered stallions.

**Discussion**

All three disciplines operate a system of thresholds, whereby once a horse has attained a certain level it is no longer allowed to compete below that level. This means that an analysis at each grade or performance level is possible as no animal is represented in two different grades or classes within the data collated. However, there are considerable problems associated with using lifetime performance as an indicator for breed improvement schemes. The problems include inequalities within the distribution of prize and point funds within and between competitions, especially in performance assessment between animals. These problems have been reviewed elsewhere. Additionally, when lifetime points earnings were awarded subjectivity (within the mechanisms to award points), various difficulties became evident in interpretation of the data.

One of the most important problems encountered when assessing the performance of horses for breeding improvement schemes is the effect of the proportion of the population that does not register to compete. The reasons for non-registration for competition are not possible to identify from the data presented in the database. Further analysis of recorded individuals over an extended time period would be required to correctly classify the ‘missing data’. It is, however, reasonable to assume that an animal that did not register may be classified as MAR rather than completely missing at random. Therefore a process of list-wise deletion has been applied to the data.

Lifetime performances of UK sport horses competing in 2000 have been reported in this study for all disciplines (dressage, eventing and show jumping). However, there is a problem associated with the data series. If a horse was registered for discipline and then re-registered for another discipline, its rating for the original discipline was included in the database. This means that an animal may not have competed in a discipline in 2000 but its lifetime performance for that discipline would be recorded if it had performed in either of the other two disciplines in 2000. This accounted for approximately 12.9% of the data available for analysis. Any conclusions that are drawn have to consider this confounding effect on the dataset (Table 1).

The maximum recorded winnings of any individual horses in each discipline are well above the threshold for the highest graded level in that discipline nationally. For dressage, the maximum recorded winnings were £717 (2.2 times greater than the 325 advance points threshold), for eventing £1461 (24 times greater than the 61 points advanced threshold) and for show jumping £658.309 (365 times greater than the grade A threshold of £1800). This may lead to misrepresentation in measures associated with performance when comparing horses that are in excess of the advanced grade for their individual discipline.

**Table 3** Correlations of total lifetime earnings between disciplines for competing horses in 2000

<table>
<thead>
<tr>
<th>Discipline</th>
<th>Dressage</th>
<th>Eventing</th>
<th>Show jumping</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eventing</td>
<td></td>
<td>0.022 ($n = 2493$)</td>
<td></td>
</tr>
<tr>
<td>Show jumping</td>
<td></td>
<td>0.037 ($n = 2448$)</td>
<td>0.228* ($n = 5586$)</td>
</tr>
</tbody>
</table>

*Correlation is significant at the level $P < 0.01$ (two-tailed).
The use and measure of mean winnings for all three populations is unlikely to provide a true reflection of the population structure and therefore has limited value for animal breeding. The mean winnings for dressage for the whole UK sport horse population is £804; however, 88% of the registered population won less than £800. The skewed nature of the distributions is not unexpected and the overall structure of the populations is similar to those found in Irish sport horses\textsuperscript{18}. These factors lead to considerable difficulty in use of this approach to identify potential elite animals for a selective breeding programme.

A degree of commonality in distribution patterns at varying levels of competition is observed. At the lowest levels of competition for each discipline (preliminary dressage, novice eventing and grade C show jumping), a highly negatively skewed, hypergeometric population distribution is observed (Figs 1, 6 and 9). At intermediate levels of competition (novice, medium and advanced–medium dressage, intermediate eventing and grade B show jumping) a more rectangular distribution with varying degrees of bimodality is observed (Figs 2–4, 7 and 10). At the highest levels of competition (advanced dressage, advanced eventing and grade A show jumping), greater normality in distribution patterns is observed (Figs 5, 8 and 11). Further investigation is required into how these variations in distribution pattern may affect the application of any breed improvement scheme to the population.

The use of correlating performances between disciplines as a tool in aiding breeding decisions within a population has been demonstrated in other equine populations\textsuperscript{2,24,25}. Correlations between the various components of the three disciplines in this study were weak or not present. A relatively weak (but significant) relationship between performance at eventing and show jumping was observed (Table 3). This observation is weaker than that reported by Ricard and Chanu\textsuperscript{21}. No apparent relationship between show jumping and dressage (Table 3) was identified in our study. Other authors have reported only weak genetic correlations between these two disciplines\textsuperscript{25,24,25}. Philipsson et al.\textsuperscript{4} reported pre-selection of horses for disciplines as a major problem when assessing the suitability of data for estimating genetic parameters and genetic correlations. Additional to the problems caused by pre-selection for discipline, the effect of environmental variation has been well documented\textsuperscript{20,25}. The preliminary analysis of UK data for all performance data by discipline suggested that 59% of horses competing at eventing have been registered for show jumping at some point in their career, while only 25% of horses competing in show jumping have been registered for eventing. Any comparison between the data presented in this paper and other studies is difficult as the UK data were collected under different environmental conditions and lacked the structured approach to measurement of performance adopted in Europe.

Problems associated with recorded parentage are apparent in the UK sports horse population. A substantial proportion (33%) of the competing population has no recorded parentage. Of the remaining population that have recorded parentage, further investigation is required to ascertain the numbers that are verifiable through genetic analysis. Furthermore, 32% of stallions have no registered competing progeny, making any assessment of the stallion population difficult. Of the registered stallions, 16% are registered with the General Stud Book. These horses will not have been bred as sport horses but will have been ‘imported’ to the sport horse population from the racing industry. When the registration in the General Stud Book is combined with the 14% of sires registered with other non-sport horse societies, approximately 30% of stallions recorded in the \textit{Sports Horse Annual} are not bred initially for the sports horse industry. This proportion compares with estimates of 60–70% in the New Zealand sport horse population\textsuperscript{26}.

The number of horses competing in the UK has increased between 1988 and 2000 by over 24%. Substantial increases are recorded in both dressage and eventing. The apparent increase in horses registered reflects the expansion in these two disciplines, in the light of improved management (i.e. the formation of British Dressage and British Eventing in 1999 and 1997, respectively) and accessibility of competition in the UK. The number of mares in the sport horse industry covered by stallions in the UK is difficult to ascertain. Sports Horse Breeders of Great Britain reported in 2000 that the average stallion registered (total number of stallions registered 107) covered 16 mares\textsuperscript{27}. The top-ranked show jumping (by lifetime progeny winnings) stallion in 2000 was reported to have covered 15 mares whereas the top-ranked dressage stallion covered 20 mares\textsuperscript{5}. Age profiling of the UK population would enable a more accurate picture of the number of brood mares to be established. Unfortunately, the current data recorded do not allow the analysis. If the age profile of the UK population is similar to that in the Irish sport horse population\textsuperscript{26}, the largest single age group accounts for approximately 13% of the population or about 4200 horses in the UK. If a fertility rate of 65%\textsuperscript{27} is applied to the mare population, an estimate of a maximum of just over 6700 brood mares is produced. If this is correct, the average stallion is covering approximately eight mares.
Conclusions

If the UK is to compete with its European neighbours and become a leader in the breeding of sport horses, it needs to embark upon a comprehensive and systematic review of its current sport horse industry. A number of conclusions and recommendations regarding the UK sport horse population, with particular reference to the feasibility of introducing a breed improvement scheme, may be drawn from the present study.

The data used herein were in raw form and therefore subjected to environmental as well as genetic effects. Analysis of genotype and environmental factors has to be performed; however, the current recording method lacks a degree of quantification. Ideally, all performance for all horses - whether currently registered or not - should be recorded. There is also a lack of data regarding age and sex, making it impossible to predict the effect of these factors on competition performance.

The structure and manner in which performance levels are recorded needs to be reviewed, particularly at the higher and lower levels of competition. It is clear that there is a degree of commonality in data distribution patterns between disciplines at the same level. This can be concluded as hypergeometric at novice level, uniform at intermediate levels and Poisson distribution at advanced levels. The causative factors of these patterns need further investigation. Many horses are competing well above the current highest threshold in the UK and therefore may be compared unfavourably with horses that have just reached this threshold. Additionally, many horses within disciplines are recorded as having no points or monetary winnings. Few conclusions can be drawn from correlations between disciplines. Where significant correlations do exist, between show jumping and eventing, pre-selection of horses will have played a major role. It is also apparent that the vast majority of competitors are competing at a relatively low level within their sport. The implications of this for elite riders and horses need to be considered within any breed improvement model as a major potential source of environmental variation.

Lack of recorded parentage is a large problem when consideration is given to implementation of any breeding model. Currently two-thirds of the population have recorded sires, although the number of these that are verifiable requires further investigation. Studies into any change in pattern of parentage recording should be conducted with regard to the level at which the horse is competing. The number of sports horse stallions in the UK is high compared with the competing population as a whole. The number of mares covered on average is too small for the elite stallion population to have any real effect on the overall improvement of horse quality in the UK. The implications of this and the previous comments need careful consideration before the implementation of a breed improvement model for sports horses within the UK.

References


