THESES OF DOCTORAL (Ph.D.) DISSERTATION

UNIVERSITY OF KAPOSVÁR
FACULTY OF ANIMAL SCIENCE
Production and Breeding of Ruminants and Horse

Head of the doctoral school
DR. PÉTER HORN
Ordinary Member of the Hungarian Academy of Sciences

Supervisor:
DR. WALTER HECKER
Candidate of Agricultural sciences

Co-Supervisor:
DR. JÓZSEF STEFLER
Candidate of Agricultural sciences

POSSIBILITY OF THE SELECTION OF THOROUGHBREDS ON RACE PERFORMANCE BASED ON STEEPLECHASE RACE RESULTS

Written by:
ÁRPÁD BOKOR

KAPOSVÁR
2006

1. BACKGROUND OF THE RESEARCH, OBJECTIVES

After the end of the 2nd World War horses were changed to machines in the field work and transports, and a new breeding goal was determined as sport horse. The direction of the sport horse selection was continuously changed over the years and was substantially affected by Thoroughbreds. The selection of Thoroughbreds is based on their racing performance usually in flat races. Jumping ability, riding ability and temperament have never been applied as selection criteria in Thoroughbred racehorses.

The percentage of Thoroughbred stallions has been decreasing during the last decade because of their special selection, and these horses do not satisfy all requirements that could be important in the sport horse breeding. However the stallions of different kind of sport horse breeds have a high percentage of Thoroughbred blood.

The objective of the sport horse breeding is to breed a horse with good conformation and movement, capable of high performance in the equestrian competitions. The number of Thoroughbreds taking part in the equestrian competitions is low. Neither in the show jumping nor in the thee-day-event competitions are the number of the Thoroughbreds sufficient for the effective selection.

Considering the above mentioned statements and the literature, the only opportunity to select Thoroughbreds for jumping performance is to use the breeding value estimation in the steeplechase race results. It is also a possibility for the sport horse breeders the chose a stallion or mare from a family took part successfully in these kind of races. Breeders prefer to breed Thoroughbreds for flat races rather than for steeplechase races because the horses can take part in the former races from the age of two years, so they can have earnings at this age. On the other hand it also operates as kinds of contra-selection because those horses perform well in flat races usually never compete in steeplechase races.

In Europe, only France, the United Kingdom and Ireland have numerous steeplechase races with numerous participants. During my research I processed the race results of these three countries and examined the possibilities to measure racing performance in steeplechase races.

During my research the following questions were considered:

- Should be the earnings and ranks used as a measure of performance for horses taking part in steeplechase races?
- Develop a model which is suitable for genetic parameter estimation for racehorses taking part in steeplechase races.
- To examine the usefulness of the developed models for breeding value estimation.
- Examination the possibility to adapt the developed models for other populations.

2. MATERIALS AND METHODS

Race records from the United Kingdom and Ireland have been collected from the Race form Interactive (commercially available dataset) (n=51140) and from "France-Galop" for the French races (n=54889). From the United Kingdom and Ireland only the "chase" type of races has been used. In France hurdle races were also excluded from the analysis. The analyzed data was collected between 1998-2003. The records of horses that were distanced from the winning post were excluded from the analysis.

Pedigree information of race participants was also collected from the „RACEFORM INTERACTIVE" and from the „FRANCE-GALOP". The pedigree files were controlled with the „Pedigree Viewer" (KINGHORN ÉS KINGHORN, 2005) software.

After the correction and extension of the available data sets SAS/BASE (SAS 9.1, 2004) module was used to convert the data to the required structure. Preparing the datasets for the (co)variance component estimation was accomplished using the PEST UIUC V3.1 GROENEVELD, 1990) software (under UNIX operation system.

2.1. Measuring performances

During the analysis logarithmic transformations of earnings and ranks were used to measure performance.

2.1.1. Earnings

Divisions of the racing prize were the same in the examined countries. The horse ranked first get 65%; the second get 20%, the third 10% and fourth get 5% of the race prize. "Computed earnings" were calculated for horses without earnings but ranks (V-VI. positions). Unplaced horses (in place VII or later) were arbitrarily assigned the rank and the half of the corresponding earnings for the places 6. Normal distribution of the earnings was obtained by applying a logarithmic transformation.

2.1.2. Ranks

Transformation for ranks was also necessary in order to use a normalized measure of performance. Scores for ordinal data was used (Fisher and Yates, 1957) to obtain a standardized normal scale for the expectation of the k^{th} rank from the n horses competing in the race. Unplaced horses based on the same principle as above received equal value.

2.2. Model development to estimate breeding values

Estimating (co)variance components was accomplished applying the VCE-5 (KOVAC ET AL., 2003) software. To estimate breeding values for earnings and ranks the applied animal and repeatability models followed the method developed by Tavernier (1989, 1990) using the PEST UIUC V3.1 (Groeneveld, 1990) software.

Races were divided into three different classes by distance (under 3500 meter, between 3500 and 4700 meter, over 4700). Considering all race distances together the effect of the race distance was also used in the model.

The repeatability model was the following:

y = Xb+Za+Wpe+e, were

y = vector of observations (log of earnings or transformed ranks)

b = vector of fixed effects, as year, age, sex, racecourse, trainer, jockey, race distance, carried weight, the level of the race

a = vector of additive genetic effects

pe = of permanent non-genetic effects

e = vector of residuals while **X**, **Y**, **W** were the known incidence matrices.

For earnings, the different race levels are not considered because the prize money evaluates the level of the race. For the rankings the level of race had to be considered.

The analysis of variance was conducted by the SAS software package using the GLM (Generalized Linear Model) procedure on race year, age, sex, trainer, jockey, racecourse and race level influencing the race.

Eight different models were used to estimate genetic parameters like environmental and additive genetic effects and repeatability on the log of earnings and transformed ranks using the VCE-5 software (KOVAC ET AL., 2003).

3. RESULTS AND DISCUSSION

Had more information been used the estimated genetic parameters would have been more accurate. However, only those factors were considered that were available in both datasets. All of these factors had significant effect on the examined traits ($p < 0.001$). Analysis of variance was carried out also on the three different race distances and significant effects were found on the measured traits.

Highest parameters were estimated when all race distances were considered in one model and the effect of race distance was considered. Ignoring the effect of the jockey can be explained by their high number (812 and 617), and their dependence on the trainers.

The highest estimated parameters were observed when the effect of the trainer was excluded from the model ($h^2_{FR} = 0.178 \pm 0.013$ and $h^2_{UK, IRE} = 0.062 \pm 0.007$). Several authors noted that ignoring the trainer from the model increased the estimated genetic parameters in flat races based on the earning, ranks or handicap weights (SCHULZE-SCHLEPPINGHOFF ET AL., 1985, 1987; PREISINGER ET AL., 1990). This phenomenon might be caused by numerous environmental effects like racecourse or trainer, or from their interaction with the trainer and resulted in overestimated parameters.

Considering this last finding the following effects were used in the model estimating genetic parameters for the log of earnings: race year, age, sex, trainer and carried weight. Estimated genetic parameters showed very low values ($h^2_{FR} = 0.062 \pm 0.009$, $r_{FR} = 0.208 \pm 0.018$; $h^2_{UK, IRE} = 0.037 \pm 0.009$, $r_{UK, IRE} = 0.225 \pm 0.018$).

Estimated genetic parameters were lower when the transformed rank was evaluated. When the trainer effect was excluded from the model the genetic parameters were also overestimated. In this model the race year, age, sex trainer and the racecourse were considered. Results were lower in France and in the United Kingdom and Ireland also (h^2_{FR} = 0.056±0.007, r_{FR} = 0.129±0.014; $h^2_{UK, IRE}$ = 0.027±0.004, $r_{UK, IRE}$ =0.114±0.009).

Generally the environmental effects have a greater influence on the transformed ranks than on the log of earnings.

Earlier results from OKI ET AL. (1995) pointed out that the racing performances measured by time in different distances are different traits. Thus race results were grouped in two distance intervals under and over 4700 meters. Genetic parameters, like heritability and genetic correlation were estimated by the annual earning per start with the VCE-5 (KOVAC ET AL., 2003) software.

Results ($r_{g_{FR}}$ = -0.329±0.039, $r_{g_{UK, IRE}}$ = 0.087±0.020) showed that the races under and over 47000 meters demand different qualities from the horses. Genetic correlations were low and negative in France, and negligible in the United Kingdom and Ireland.

Genetic trends based on the average breeding values of the successive birth years were different in the two examined groups.

54889 race results of 9041 horses were considered in France. These results showed a negative trend in breeding values (-0.05 % / year) when the estimations were calculated for the log of earnings. It seems that the selection is not based on the estimated breeding values by the BLUP. This may be the consequence of the non-thoroughbred participants in France, where the creation of a mixed population selected by steeplechase races increased the genetic variability.

The accuracy of the estimated breeding values increased with the increasing number of the runs because the repeatability model was used.

In the United Kingdom and Ireland horses achieved a low genetic progress (0.11 % / year). The accuracy of the estimated breeding values was low due to the low heritability, the low number of progeny per sire and the restricted pedigree information.

Estimated breeding values based on the transformed ranks showed a negative genetic trend in France (-0.04 % / year). Accuracy of the estimated breeding vales showed the same tendency as before. Genetic trend was low for the ranks in the United Kingdom and Ireland (0.04 % / year). Accuracy of these results was lower than previously.

4. CONCLUSIONS

Contrary to flat races the traits expressing the performance in steeplechase races (earnings and ranks) are heavily affected by environmental factors therefore the heritability estimates are low. The substantial influence of the environmental factors is primarily the consequence of their large number (number of obstacles, different race levels) and of the different race distances.

Age and sex (of the horses), trainer, racecourse, race distance and carried weight all influence the performance of the horses measured at the race. Breeding value estimation for logarithmic earnings and transformed ranks requires the application of different models. Excluding the trainer effect from the models increased the estimated genetic parameters.

Using the estimated breeding values according to the ranks the genetic trends can be determined. The applied procedure (i.e. breeding value estimation according to the ranks) is suitable to compare the genetic trend of different populations. Contrary to the earnings the ranks provide identical evaluating system in different countries and the difficulty of the races can be expressed with their total prize money. Breeding value estimation based on the earnings would only be suitable to compare populations of different countries if the total prize money is the same.

According to the results of the present study the genetic trend of the steeplechase race horses in the United Kingdom and Ireland exceeds than that of French individuals. One of the reasons for this finding may be that in France the proportion of Thoroughbreds in steeplechase races is about 60%

while in the United Kingdom and Ireland only the Thoroughbred horses compete in these races.

The available information sources had high variability therefore the reliability of the estimations was low. As a consequence the selection should be based on certain families or lines rather then on the whole Thoroughbred population.

My results justify the assumption that presently the selection is only accomplished in flat races and the majority of horses competing in steeplechase races are those that are less successful in the flat races.

The various information sources connected to steeplechase races can be utilized for the selection improving the jumping performance. Applying selection based on breeding value estimation selection response can be achieved for the parameters expressing the race performance. Those horses participating in steeplechase races have to have speed and stamina, besides they possibly possess jumping ability. To evaluate the genetic correlation between the performance showed in steeplechase and show jumping (and in free jumping) further analyses would be necessary. However, presently no adequate information is available as only few horses can be found competing in show jumping or in thee-day-event competitions that had previously been competed in steeplechase races.

5. NEW SCIENTIFIC RESULTS

1. Considering the breeding goal and the racing system of the Thoroughbreds the only opportunity to select Thoroughbreds for jumping performance (beside speed and stamina) is to use steeplechase race results. Thus the breed can be selected indirectly for jumping performance.
2. In steeplechase races heritability estimates for race performance measured by ranks were low (h^2_{FR} = 0.056±0.007, r_{FR} = 0.129±0.014; $h^2_{UK, IRE}$ = 0.027±0.004, $r_{UK, IRE}$ =0.114±0.009) and also less than that of the flat races. Yet applying consistent selection may produce selection response in the long term.
3. Excluding the trainer effect from the model increased the estimated genetic parameters (h^2_{FR} = 0.178 ± 0.013 and $h^2_{UK, IRE}$ =0.062 ± 0.007).
4. Analysing the steeplechase race results of the Thoroughbreds in France and in United Kingdom and Ireland it can be concluded that rank based selection is suitable for the improvement of Thoroughbreds. Breeding value estimation for the ranks is also appropriate for multiple-country comparison.
5. A BLUP model was developed to estimate breeding value for the rank that takes into account the breed characteristics and the specific environment of the steeplechase competitions, derive information from data collecting systems and satisfies reliability requirements. The following effects were used in the model evaluating the ranks: year, age, sex, trainer, racecourse, race distance.

6. SCIENTIFIC PAPERS ON THE SUBJECT OF THE DISSERTATION

Papers published in foreign-language peer-reviewed journals:

Bokor, Á., Blouin, C., Langlois, B., 2006. Possibility of the selection of racehorses on jumping ability based on their steeplechase race results in France, in the United Kingdom and Ireland. J. Anim. Breed. Genet. (*in press*)

Bokor, Á., Stefler, J., Hecker, W., Nagy, I., 2006. Genetic parameters of racing performance of Thoroughbred horses in Hungary. Acta Agraria Kaposváriensis (*in press*)

Bokor, Á., Blouin, C., Langlois, B., Stefler, J. 2005. Genetic parameters of racing merit of Thoroughbred horses in steeplechase races. Ital. J. Anim. Sci., 4, Suppl. 3. 43-45.

Full conference paper in proceedings:

Bokor, Á., Stefler, J., Hecker, W., 2004. Examination of inherited jumping ability of English Thoroughbred mare families from their ranks (Hungary), KRMIVA 46, Zagreb, 3, 141-144.