Animal Science Papers and Reports vol. 22 (2004) no. 4, 637-643
Institute of Genetics and Animal Breeding, Jastrzębiec, Poland

Results of races of different status as a source of information for breeding value prediction of racehorses

Magdalena Sobczyńska, Marek Łukaszewicz

Polish Academy of Sciences Institute of Genetics and Animal Breeding,
Jastrzębiec, 05-552 Wólka Kosowska, Poland

(Received October 12, 2004; accepted December 13, 2004)

The purpose of the study was finding out whether abilities of horses to win in races of different status constitute genetically different traits. Data included information on 12143 starts of 1414 Thoroughbreds, aged 2-9 years and 8754 starts of 928 Arab horses, aged 3-10 years. Thoroughbreds ran in 1693 races over the distances ranging from 1000 to 3200 m, while Arab horses participated in 1157 races on distances from 1400 to 3000 m. One group of high-status (HS) and four groups of low-status (LS) races were distinguished. Ranking (square root) in HS races and rankings in LS races were treated as different traits and analysed with a two-trait animal model. Genetic correlations between traits "ranking in the HS race" and "ranking in one of theLS races", except for that between ranking in prestigious races and ranking in group IV races in the Arab horses ranged from 0.82 to 0.99. It means that we discuss very much the same trait measured only in races of different prestige statuses. LS results should be used as a source of information upon prediction of horses' breeding value for the winning ability.

KEY WORDS: breeding value / race status / racehorses / source of information

Seeking opportunity of increasing the reliability of breeding value estimation breeders try to use all the information available on an evaluated animal. In racing horses in Poland, we face situation when the system of advancing the horses cause that they predominantly race in only a certain quality of races, and hence, they race predominantly against a certain quality of competitors. A problem arises if the results obtained in different presige races (high-status *vs* low-status) regard the same genetic background reflecting a horse's ability to win. In other words the problem is whether ability to

win in a prestigious race is the same trait as ability to win in a race of a lower status. This is not an academic issue. It is a common attitude within the circle of racehorse people that results obtained in low-status are not comparable at all to the results of the high-status races. From a practical point of view it is a question of whether, when estimating breeding value of racing horses, we can combine the results from all the races in a single run.

The breeding stallions are chosen out of the winners of the high-status races. Presumably, they should be the best possible, but since the result of a race is affected by numerous factors, relying just on the results of a limited number of races is risky. One or two outstanding performances can dramatically change a stallion's career regardless the number of poor performances, and *vice versa*. A better solution is to account for all the available performances, including those which can only be attributed to a horse through the pedigree.

Results of the prestigious races (often called "selection races") are intuitively preferred to those obtained in other races. Such an attitude is justified by the plan employed at the racing tracks [Polski Klub Wyścigów Konnych 2001] allowing the race winning horses, or the horses earning much money, to move to a higher status races and compete against, hopefully, more and more intensely selected rivals. On the other hand such a plan does not give the horses a chance to run against every possible rival as not all the horses are given chance to run in prestigious races.

Still, the main body of information comes from the lower-status races. Since the racing success is very much dependent on many factors linked to a particular race (state of track, weather, distance, number and quality of competitors, rider, etc.) ignoring knowledge coming from lower-status races upon estimating the breeding value, would only be justified if ability to win in a high-status race had different genetic background than the ability to win in a low-status race.

Finding out whether the abilities to win in races of different status are genetically different traits was the purpose of the present study.

Material and methods

Placings were analysed of Thoroughbred and Arab horses racing in the years of 1998-2001 at three racing tracks in Poland. The data included information on 12143 starts of 1414 Thoroughbreds aged 2-9 years, and on 8754 starts of 928 Arab horses aged 3-10 years. The Thoroughbred horses ran in 1693 races over the distances of 1000-3200 m, while the Arabs participated in 1157 races and ran distances of 1400-3000 m. Since the genetic and phenotypic correlations between ranking at finish and endowments are high [Fedorski 1988, Chico 1994, Sobczyńska and Kownacki 1997, Sobczyńska and Łukaszewicz 2002] and the money prizes are awarded just to the first five horses, it was decided to measure the ability to win with the ranking at finish. The lower-status (LS) races consisted of four groups (IV, III, II, I – from lowest to highest status) while the high-status (HS) races (A and B categories) were grouped together

(group 0). Placing at finish in LS races and placing at finish in HS races were treated as separate traits. Horses which had never run in a prestigious race had missing data for that type of races. Square root transformation was applied to normalize the distribution of placings.

The variance components were estimated by the REML method with the software of Misztal [1998] using the following animal model:

$$y_{ijklmn} = \alpha + A_i + G_j + R_k + \beta_1 \times weight_{ijklmn} + \beta_2 \times distance_{ijklm} + r_l + pe_m + a_m + e_{ijklmn}$$

where:

y_{ijklmn} – rankings at finish;
α – adequate intercepts;
A_i – fixed effects of age;
G_j – fixed effects of sex;
R_k – fixed effects of race;
$\beta_1 \times weight_{ijklmn}$ – linear regressions on weight carried;
$\beta_2 \times distance_{ijklmn}$ – linear regressions on distance;
r_l – random effects of rider;
pe_m – random effects of permanent environment;
a_m – random genetic additive effect of an animal;
e_{ijklmn} – random errors.

The above model was employed for both pair-wise (two-trait) analyses of the HS races results with each of the LS group results, and for each type of the races separately (single-trait).

As the number of geldings was small (18 Thoroughbreds and one Arab) they were classified together with mares. The Thoroughbred horses older than 3 years were grouped in one age class. Similarly, the Arab horses older than four years were put into one age group. The Thoroughbreds were ridden by 138, and the Arabs by 97 riders. It was also decided not to fit the breeder/owner effect in the model as it has been found to explain too much of the additive genetic variance [Sobczyńska 2003, Sobczyńska and Łukaszewicz 2002]. The pedigrees were three generations deep and included 3018 Thoroughbred and 2025 Arab horses. The computations of variance components were run separately for each breed.

Results and discussion

The numbers of horses and numbers of starts within each race group are presented in Table 1.

Table 1. Distribution of Thoroughbreds (xx) and Arabs (oo) across races of different status

Race status	No. of starts				No. of horses				Mean no. of starts	
	xx		oo		xx		oo		xx	oo
Group IV	2951	(24%)	796	(9%)	645	(46%)	195	(21%)	4.6	4.1
Group III	3287	(27%)	1210	(14%)	845	(60%)	313	(34%)	3.9	3.8
Group II	3851	(32%)	4284	(49%)	1185	(84%)	851	(92%)	3.2	3.2
Group I	982	(8%)	1747	(20%)	504	(36%)	486	(52%)	1.9	5.0
High status	1072	(9%)	717	(8%)	337	(24%)	206	(22%)	3.2	3.5
Total	12143		8754		1414		928		8.6	9.4

The highest proportions of horses and starts were observed in the race group II in which young horses, *i.e.* two years old Thoroughbreds and three years old Arabs, usually begin to run [Polski Klub Wyścigów Konnych 1998-2001]. In both breeds, after one or two starts in the group II races, the horses are moved to neighbouring, higher or lower race groups.

Some inter-breed differences regarding the distribution of the horses between the race types can be seen. While the proportions of horses racing in the prestigious races are similar, the Arabs are more likely to appear in group I races (52% of the horses) and less likely in the races of the lowest status (group IV). A reflection of that fact can be found when analysing the heritability coefficients (Tab. 2). The highest genetic variance, relative to the phenotypic one, was naturally found in group II, in which almost all the horses begin their racing career. It is of the same magnitude as in the entire populations regardless the race status (Tab. 2). But in race groups III and IV the heritability was smaller in the Arab horses (fewer of them are sampled in those groups), while in group I this ratio was more than two times smaller in the Thoroughbreds. When selection for the ability to win begins in the LS race group of a middle status (group II) and is performed diversely towards the extremes, one expects that the ge-

Table 2. Heritabilities and repeatabilities across race types for high (0) and low (I, II, III, IV) status races for Thoroughbreds (xx) and Arabs (oo)

Breed	Statistic	Race status					
		0	I	II	III	IV	combined (0+I+II+III+IV)
xx	heritability	0.14	0.06	0.18	0.10	0.03	0.18*
	repeatability	0.28	0.25	0.27	0.28	0.34	0.34*
oo	heritability	0.09	0.15	0.18	0.03	0.02	0.18**
	repeatability	0.33	0.28	0.46	0.26	0.28	0.45**

*Sobczyńska and Łukaszewicz [2004].
**Sobczyńska and Łukaszewicz [2002].

netic variance would drop as a reaction to the successive steps of selection. Indeed, this phenomenon can be seen in the Arab horses – beginning with h^2=0.18 in group II we arrive at h^2=0.02 in group IV and at h^2=0.09 in the highest group 0 (Tab. 2). The difference between group II and III seems, however, intuitively too big. In the Thoroughbred horses, there was a marked decline of heritability estimated in group I (from 0.18 in group II down to 0.06), which was regained in group 0 (0.14) – still, considerably high genetic variance for purposes of selection. Lower genetic variance is expected in the extreme race groups (0 and IV) also due to the fact that young horses, *i.e.* two years old Thoroughbreds and three years old Arabs are not likely to race in those groups.

Inter-breed differences can also be observed when the repeatabilities of the performances are compared (Tab. 2). The repeatabilities are fairly stable across the race status groups in the Thoroughbreds indicating that, with heritabilities dropping towards the extremes, specific environment gains importance in moulding the performance. In particular, the covariance between repeated performances has environmental nature in race group IV, which is true for both breeds. The main difference however, regards the repeatability of ranking in group II in which the horses start racing. Young horses entering racing track differ between themselves because of both genetic abilities and permanent environment specific to each horse. While in the Thoroughbred horses the permanent environment component is only half of the additive genetic, it is more than 1.5 times higher in the Arabs (0.09 and 0.28, respectively). This means, that in the Arab horses the environment of rearing animals affects their early performance much more than does their additive genetic background, and more than it is found in the Thoroughbreds. Nevertheless, that difference between breeds vanishes rapidly with ageing of the horses.

The differences between the horses regarding their ability to win in the HS races tend to be due to specific environment rather, than to the differences between their breeding values. The same is true for the LS races.

All the genetic correlations between traits "ranking in the HS race" and "ranking in the LS race", except for correlation between ranking in prestigious races and ranking in group IV races in the Arab horses, appeared high and ranged from 0.82 to 0.99 (Tab. 3). It means that we discuss very much the same trait measured only in races of different status. The lowest genetic correlation of 0.32, estimated between the ability to win in the prestigious races and the ability to win in the group IV races, in the Arabs, is believed to result from very small number of horses racing in both types of races, thus giving a poor estimate of covariance.

Generally, the genetic correlations between rankings found for the Thoroughbreds were higher than those estimated for Arab horses. Considering longer distances, at which the Arab horses race, probably more factors affect the result of the race.

The environmental correlations between ranking at finish in HS races and ranking in group I and II of LS races were all positive and only that between rankings in group 0 and I in Thoroughbreds was low and amounted to 0.18 (Tab. 3). Such a low estimate of the environmental correlation may result from a low number of horses competing in

Table 3. Numbers of Thoroughbred (xx) and Arab (oo) horses participating in a given pair of high status race (0) with low status race (I, II, III or IV), numbers of starts, genetic (r_G) and environmental (r_{PE}) correlations between placings in high and low status races

Breed	Statistic	Pair of race groups				
		0-I	0-II	0-III	0-IV	0-(I+II+III+IV)
xx	no. of horses	241	276	196	90	323
	no. of starts	1405	1883	1412	568	3816
oo	no. of horses	175	167	79	27	187
	no. of starts	1342	1315	465	129	2423
xx	r_G	0.94	0.99	0.91	0.95	0.99
	r_{PE}	0.18	0.48	-0.56	-0.87	0.52
oo	r_G	0.82	0.91	0.92	0.32	0.95
	r_{PE}	0.85	0.64	-0.57	0.22	0.93

group I races. On the other hand, environmental correlations between ranking in group 0 and rankings in group III and IV races tended to be high and negative. Again, in Arab horses, correlation of 0.22 between ranking in group 0 and ranking in group IV, which does not follow the pattern may, again, result from a low number of horses participating in both groups of races.

Ranking at finish describes a horse's ability to compete in a group of rivals. The competition in all the groups of races is similar – it is only the level of competitors which increases all the way from group IV to group 0. A similar problem was met by Fedorski [1977]. He found high genetic correlation between placing in the first race in the season and placings in the remaining races in that season for two-, three- and four-years-old Thoroughbred horses. He suggested that the result of the first race in the season is a good prognosis of the horse's career in the season. Likewise, the results obtained in the present study by horses in group II races predict their quality as competitors in other race groups.

It can be concluded that underestimating of the results of the LS races is unjustified. The LS races are a valuable source of information on horses' ability to win and should be used upon breeding value prediction. Very high genetic correlations between the horse rankings in races of different prestige indicate that ability to win is genetically the same trait across the races of different status. It is only the permanent environment which tends to differentiate the abilities to win in races of different prestige.

REFERENCES

1. CHICO M. D., 1994 – Genetic analysis of Thoroughbred racing performance in Spain. ***Annales de Zootechnie*** 43, 393-397.

2. FEDORSKI J., 1977 – Odziedziczalność dzielności wyścigowej koni pełnej krwi angielskiej w Polsce. (Heritability of racing ability of Thoroughbreds in Poland). ***Prace i Materiały Zootechniczne*** 14, 121-129.
3. FEDORSKI J., 1988 – Porównanie różnych kryteriów oceny klasy wyścigowej koni pełnej krwi angielskiej (Comparison of different criteria of racing class in Thoroughbreds). In Polish, summary in English. ***Zeszyty Problemowe Postępów Nauk Rolniczych*** 345, 45-50.
4. MISZTAL I., 1998 – BLUPF90 – a flexible mixed model program in Fortran 90. Animal and Dairy Science, University of Georgia, 19 pp.
5. Polski Klub Wyścigów Konnych, Biuletyny 1998-2001- Załącznik nr 1 do Rozporządzenia Ministra Rolnictwa i rozwoju Wsi z dnia 23 sierpnia 2001 w sprawie ustalenia regulaminu wyścigów konnych (Dziennik Ustaw R.P. Nr 90. poz. 1006). Warszawa. (Polish Jockey Club, Bulletins 1998-2001- Ammendment to the Regulation by the Minister of Agriculture and Rural Development of August 23, 2001, on the rules of horse races (Polish Government Gazette 2001, No. 90, Item 1006.
6. SOBCZYŃSKA M., 2003 – Wpływ niektórych czynników na parametry genetyczne cech wyścigowych koni pełnej krwi angielskiej i czystej krwi arabskiej (Influence of some effects on genetic parameter estimates for racing performance in Thoroughbreds and Arabian horses). In Polish, summary in English. ***Roczniki Naukowe Zootechniki***. 30 (2), 223-231.
7. SOBCZYŃSKA M., KOWNACKI M., 1997 – Genetic aspects of racing performance in Polish pure bred Arab horses. I. Genetic parameters. ***Journal of Applied Genetics*** 38 (2), 179-186.
8. SOBCZYŃSKA, M., ŁUKASZEWICZ, M., 2002 – Estimation of genetic parameters for racing performance in Polish Arab Horses. Proceedings of the 53 rd. Annual Meeting of the EAAP. Cairo, Egypt, Book of Abstracts 8, p. 254.
9. SOBCZYŃSKA, M., ŁUKASZEWICZ, M., 2004 – Genetic parameters of racing merit of Thoroughbred horses in Poland. ***Journal of Animal Breeding and Genetics***, 121, 302-306.

Magdalena Sobczyńska, Marek Łukaszewicz

Wyniki gonitw o różnej randze jako źródło informacji o wartości hodowlanej koni wyścigowych

S t r e s z c z e n i e

Celem pracy było zbadanie, czy miejsce zajmowane przez konia w gonitwach o różnym statusie (prestiżu) jest genetycznie tą samą cechą. Analizowano wyniki 12143 startów 1414 koni pełnej krwi w wieku 2 do 9 lat i 8754 starty 928 koni arabskich w wieku 3 do 10 lat. Konie pełnej krwi biegały w 1693 gonitwach, na dystansach od 100 do 3200m, a konie arabskie uczestniczyły w 1157 gonitwach, na dystansach od 1400 do 3000m. Wyróżniono 4 grupy gonitw niższych klas (LS) i jedną, pozagrupową, o wysokim statusie (HS), stosownie do podziału stosowanego na torach. Lokaty (pierwiastek kwadratowy) w gonitwie pozagrupowej i w gonitwie grupowej (I, II, III lub IV) traktowano jako różne cechy i analizowano za pomocą dwucechowego modelu zwierzęcia. Korelacje genetyczne między cechami "lokata w gonitwie o wysokim statusie" a "lokata w gonitwie o niskim statusie", za wyjątkiem korelacji między wynikami w gonitwach pozagrupowych a wynikami w gonitwach IV grupy u arabów, były dodatnie i wysokie – od 0,82 do 0,99. Znaczy to, że lokata konia mierzy jego dzielność wyścigową niezależnie od statusu (prestiżu) gonitwy. Wynika stąd nadto, że wyniki gonitw grupowych mogą (i powinny) być wykorzystywane do szacowania wartości hodowlanej koni wyścigowych.