Scientific Papers

REFEREED

The Relationship of Selected Two-dimensional Echocardiographic Measurements to the Racing Performance of 5431 Yearlings and 2003 Two-year-old Thoroughbred Racehorses

Jeffrey A. Seder, AB, MBA, JD, Charles E. Vickery III, BS, and Patrice M. Miller

Two-dimensional echocardiograms (2DE) were recorded for 5431 yearlings and 2003 2-year-old Thoroughbred racehorses offered for sale as potential racehorses at US and European auctions and breeding farms. All horses were unraced and appeared to be in good health. Cardiac measurements included *inter alia*, heart rate, cross-sectional area of the left ventricle in systole and diastole, and interventricular septal wall structural thickness in diastole.

Descriptive statistics (ie, means, medians, standard deviations, correlation coefficients, etc) of the 2DE measurements were provided for the overall population and for subsets defined by sex, age, body size, racing performance, and auction. Racing performance data for each horse included distances raced, earnings, and levels of competition, allowing the comparison of 2DE measurements among horses grouped according to varying levels of racing ability. Statistical descriptions of the 2DE measurements were also provided for individual auctions, and compared in terms of sale prices and subsequent earnings.

Between- and within-subject measurement variability was calculated, and cardiac measurements were shown to be reproducible. Statistically significant relationships were identified between the cardiac measurements and subsequent racing performance (in terms of earnings and distances raced) when horses were grouped by age, sex, and body size. Statistically significant differences were noted between graded stakes winners and cheap claiming horses (ie, horses with large differences in racing ability) and also between groups of allowance and graded stakes horses (ie, horses with small differences in racing ability). Cardiac measurements, grouped by auctions, revealed statistically significant differences among auctions.

From EQB, Inc, West Grove, Pennsylvania.
Reprint requests: EQB, Inc, 501 Hicks Rd, West Grove, PA 19390. 1-800-223-7014; E-mail: cvickery@eqb.com; Web site: *www.eqb.com.*

0737-0806/03/2301-0003 $30.00/0
doi: 1053/jevs.2003.46

SUMMARY

Two-dimensional echocardiographic (2DE) measurements of 5431 yearlings and 2003 two-year-old Thoroughbred racehorses, grouped by age, sex, body size, auction, subsequent earnings, and racing distances, were analyzed. All horses appeared to be in good health and were unraced at the time of cardiac examination. The 2DE measurements included the cross-sectional area of the left ventricle in diastole (LVD) and systole (LVS), and interventricular septal wall structural thickness in diastole (SW). In addition, estimated body size (weight and height) were recorded for each horse.

Statistically significant differences in the means of the 2DE measurements were found between groups of horses of different ages, sexes, body sizes, auctions, race earnings, and racing distances. Multivariate discriminant analysis of the grouped data significantly improved the odds of identifying future success of the racehorses and was predictive of earnings and distances raced. HTWT (the product of height × weight) and SW were the most significant predictive cardiac variables for differentiating between high and low earners. HTWT, LVD, and LVS were the most significant predictive cardiac variables for differentiating between successful sprinters and routers.

Horses in age groups of 20 to 21 months, measured during October through December between the select yearling and two-year-old sales, did not have the same data patterns as were found for other age groups. Most of these horses were measured at private farms and were not preselected on the basis of pedigree or conformation. In addition, many had recently entered training, which may have affected heart rate and cardiac dimensions.[1]

INTRODUCTION

Thoroughbred racing performance relies heavily on the aerobic process, and numerous studies have attempted to link physical heart characteristics with athletic performance. Research studies have shown that 2DE from Thoroughbreds

Table 1. Number of races through three-year-old year among horses categorized as North American

Number of races through 3-year-old year	No. of horses	Percent of total	Cumulative percentage
Unraced	1073	22.58	22.58
1-5	1274	26.80	49.38
6-10	1215	25.56	74.94
11-15	784	16.50	91.44
16-20	297	6.25	97.69
21-25	95	2.00	99.69
26-30	13	0.27	99.96
31-35	1	0.02	99.98
36-40	1	0.02	100.00
Total	**4753**	**100.00**	**100.00**

Table 2. Averages for the breed/Worldwide performances of named Thoroughbred foals born in North America between 1985-1994[10]

Subset of population	Foals of 1985-1994	Foals by top 1% of sires
% Starters/foals	68.9	84.8
% Stakes winners/foals	3.2	9.1
% Graded stakes winners/foals	0.7	3.6
% Grade 1 stakes winners/foals	0.2	1.2
% 2-year-old starters/foals	33.5	46.2
% 3-year-old starters/foals	59.0	76.6
% 4-year-old starters/foals	44.0	57.1
% 5-year-old and up starters/ foals	26.5	36.9
Average career starts/foal	14.5	18.7
Average career starts/starter	21.1	22.0
Average win distance in furlongs	6.82	7.24
Average earnings/starter	$29,102	$71,349
Average earnings/start	$1378	$3242

Top 1% of sires determined by total progeny earnings for 1985 through 1994.

are repeatable[2] and can be used to document serial changes in cardiac dimensions within individual animals.[1]

In this study, 2DE measurements were made from 7434 unraced Thoroughbred racehorses to provide descriptive echocardiographic data for normal horses grouped by age, sex, and body size, and to investigate the relationship of this data to subsequent racing performance.

During this study (1995-2000), 2DE measurements were taken as part of a prepurchase evaluation. EQB acted as a commercial consultant to Thoroughbred buyers and on its own account to collect and analyze 2DE and related pedigree, conformation, and racing performance results. EQB's research took place in the field, inside commercial operations, and at public horse auctions, and was subject to the conditions imposed by those venues.

MATERIALS AND METHODS

Overview. Selected 2DE measurements were recorded for 5431 yearlings and 2003 two-year-old Thoroughbred racehorses between the ages of 12 and 27 months. These were unique, unraced horses. Cardiac measurements were recorded primarily at select public yearling and two-year-old auctions between 1995 and 2000.

All descriptive statistics used only the most current 2DE measurements from each horse to prevent multiple measurements of the same horse from overly influencing statistics within small groups of horses. Use of the most recent measurement of the same horse also maximized the number of 2-year-old Thoroughbreds available for the study. Among the 7434 unique horses, there were 2940 fillies (40%), 4494 colts (60%), 5431 yearlings (73%), and 2003 2-year-olds (27%).

In addition, 5909 horses (79%) were at least 3 years of age by January 1, 2000. Among these horses, 1156 (20%) had raced outside of North America (foreign) and 4753 (80%) stayed in North America by the end of their 3-year-old year. Among the North American horses, 1073 (23%) never raced and 3680 (77%) started at least once (Table 1).

Data from horses with resting heart rates above 40 beats per minute were excluded from this study. The same technician, ultrasound equipment, and measurement protocol, as described in this article, was used for all horses studied. Comments regarding physical appearance, body condition, and conformation were recorded during each examination.

Tables 1 through 4 provide statistical summaries of racing performance among the Thoroughbred breed, in general, and among horses in this study.

Performance records. All horses used to predict performance had race records through their 3-year-old year. Race records included race date, racetrack, race number, distance raced, level of race, claiming price, finish position, and earnings. Horses that raced outside of North America were identified as "foreign," and their race records were not used, because they were often incomplete or difficult to compare with North American records on the basis of dollar value or race level.

Sample bias. There were pedigree and conformation biases, because the horses examined at "select" public auctions were preselected by auction companies on the basis of above-average commercial assessment of pedigree and

Table 3. Sale to racetrack performance of 1990-1999 graduates of major yearling sales[11]

Select yearling auction name and location	No. horses sold	Median sale price ($)	Avg earnings ($)	Starters (%)	Starts per starter	Starts per foal	% winners	% Stakes winners	% Graded stakes winners	Avg winning distance
Fasig-Tipton Kentucky—July	1792	35,000	61,132	1577 (88.0)	19.2	16.9	68.5	6.9	2.7	6.99
Keeneland Kentucky—July	1945	235,000	112,752	1672 (86.0)	14.7	12.7	62.1	11.3	6.3	7.91
Keeneland Kentucky—September	28,176	22,000	48,768	24,130 (85.6)	20.0	17.1	64.9	6.4	2.0	7.01
Fasig-Tipton Saratoga—August	1535	105,000	78,696	1338 (87.2)	16.8	14.7	65.8	10.0	4.5	7.53

conformation. Not all horses at each auction were measured, nor were subjects randomly selected. EQB and its clients, whose management teams included veterinarians and trainers, further preselected horses for cardiac measurement on the basis of their own criteria.

Preselection biases were reflected in the percentage of stakes winners among horses measured. For example, approximately 3.2% of foals become stakes winners (Table 2), whereas 6.25% of horses measured for this study, and which were not known to have raced outside of North America, won a stakes race before they were 4 years old (Table 4).

These biases made predicting poor racing performance difficult because horses measured had no obvious excuses for failing.

Table 4. Averages among horses in this study/statistics through the three-year-old year of study horses

Horses not known to have raced outside of North America	
Subset of Study Population	All horses in this study (%)
% Stakes winners	6.25
% Graded stakes winners	2.90
% Grade 1 stakes winners	1.09
% At least stakes placed (including winners)	12.57
% At least graded stakes placed (including winners)	5.14
% At least grade 1 stakes placed (including winners)	1.48

Includes unraced horses. Race dollar amounts earned can be compared between horses without currency or country distortions. Compare percentages in this table to those of the top 1% of sires' progeny, shown in Table 2.

MEASUREMENT EQUIPMENT

Ultrasound equipment. A Pie Medical digital cineloop scanner 200 from Classic Medical, (Tequesta, Fla), with a 3.5-MHZ annular array, multiring crystal transducer with a 30-cm field of view at 22 frames per second was used for all measurements. The depth of display varied from 15 to 25 centimeters depending on the size of the horse. The ultrasound recorder was equipped with electronic calipers that were used to measure the stored images at the time of the examination.

Computer equipment and software. SAS release 6.12 (SAS Institute, Cary, NC), for Windows NT (Microsoft, Redmond, Wash) was used for statistical analysis. Universe (IBM) for Windows 2000 (Microsoft) was used to manage the data. The server was a Dell 2300 Poweredge (Dell, Atlanta, Ga) with dual 450-MHZ Intel Pentium (Intel, Santa Clara, Calif) processors, running Windows 2000.

MEASUREMENT TECHNIQUES

The 2DE imaging protocol was carried out on all horses, by the same experienced (≥5 years) technician to reduce measurement variability. Acoustical coupling gel (Aquasonic 100 ultrasound transmission gel, Parker, Fairfield, NJ) was applied liberally over the girth area in the fourth and fifth intercostal spaces, starting just below the level of the point of the shoulder down to the level of the olecranon. Three to five cardiac cycles were measured for each variable. Measurements were not made if the heart rate exceeded 40 beats per minute, if the heart rhythm was irregular, or if the images were unclear.

During 2DE examination, the ultrasound transducer was held in the right hand with the cursor facing caudally. The left forelimb was advanced slightly and the transducer was placed in the fourth or fifth left intercostal space, at a level just dorsal to the point of the olecranon. From this position, a left parasternal short axis view could be obtained by directing the transducer beam perpendicular (horizontal) to the longitudinal cardiac axis. The image provided a nearly circular appearance to the left ventricular lumen. Moving (angling) the transducer beam from the apex to the base of the heart, the moderator band(s), papillary muscle, chordae tendinae, and septal leaf of the mitral valve were identified and then used as intracardiac reference points to obtain reproducible cardiac images in the same tomographic plane.

Except where noted, the short axis images were projected according to international terminology on the basis of the recommendations of the American Society of Echocardiography.[3,4] Short axis images recorded from the left side of the chest were projected as though the tomographic planes were viewed from the base to the apex of the heart.

The 2DE measurements recorded for all 7434 horses were measured using electronic calipers. For all dimensions, the "inner edge" method was used [5,6,7] ie, linear parameters were measured from the inner edge of endocardial surfaces, and areas were traced along the inner borders of the endocardial echoes.

Fig 1. LVS. A left parasternal short-axis echocardiogram of the left ventricle at peak systole from a 2-year-old Thoroughbred filly with a resting heart rate below 40 bpm. This echocardiogram was obtained from the left cardiac window with a 3.5-MHZ probe (see Methods section for more details). The *dotted line* along the endocardium traces the circumference of the left ventricular chamber at peak systole.

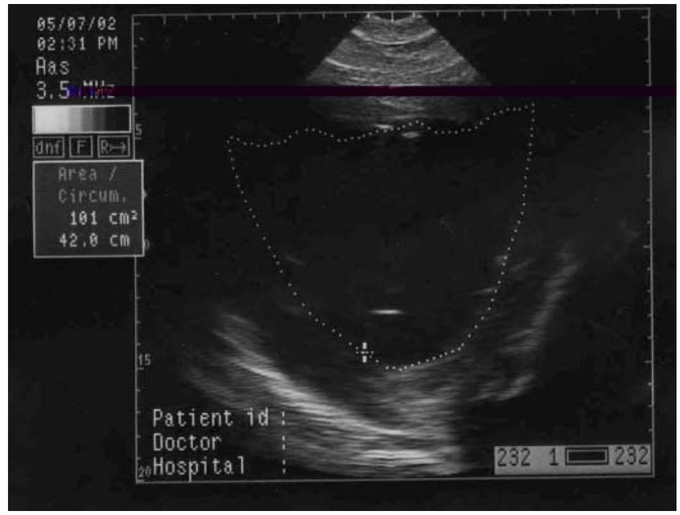


Fig 2. LVD. A left parasternal short-axis echocardiogram of the left ventricle obtained at end diastole from a 2-year-old Thoroughbred filly with a resting heart rate below 40 bpm. This echocardiogram was obtained from the left cardiac window with a 3.5-MHZ probe (see Methods section for more details). The *dotted line* along the endocardium traces the circumference of the left ventricular chamber at end diastole.

Fig 3. SW. A left parasternal short-axis echocardiogram of the left ventricle obtained at end diastole from a 2-year-old filly with a resting heart rate below 40 bpm. This echocardiogram was obtained from the left cardiac window with a 3.5-MHZ probe (see Methods section for more details). The *dotted line* measures from the attachment of the moderator band through the interventricular septum into the right ventricle to the endocardial edge of the right ventricular free wall where it attaches to the intraventricular septum.

The following variables, as shown and described in Figs 1 through 3, were measured from the stored images: left ventricular cross sectional area in diastole (LVD); left ventricular cross sectional area in systole (LVS); and interventricular septal wall structural thickness in diastole (SW).

Percent stroke volume (PS) was computed using the formula:

$$PS = \frac{LVD - LVS}{LVD} \times 100$$

The ultrasound technician estimated height and weight. The variable HTWT, which was the product of height × weight, was used in this research as an estimate of overall body size. See Appendix D for a blind-test statistical comparison of the technician's height and weight estimates versus electronic scale, weight tape, and height stick methods of measurement. (The ultrasound technician, a life-long horseperson, trained horses before this research. While she was a trainer, she had an on-site horse scale in a 40-stall training facility and took daily weight measurements of horses, and compared scale results to weight tape measurements. Thus, these were well educated estimates. Nonetheless, her technique could be difficult for others to duplicate. The discriminant results that follow used the height and weight estimates as reported by the ultrasound technician. Appendix E shows that very similar results were achieved using 5 rating categories [*whole numbers from 1, well below average, to 5, well above average*] to describe height and weight. Using 5 rating categories may be easier for others to duplicate.)

Most cardiac measurements varied depending on age, sex, and weight, making it extremely difficult to compare horses on the basis of cardiac measurements without simultaneously adjusting for the effects of these parameters. Two statistical techniques, percentiles and standardized scores, eliminated the effects of age, sex, and weight. These statistical techniques were only possible because of the large number of horses studied.

Percentiles and standardized scores for LVD, LVS, SW, and PS were calculated by comparing the subject horse to others that were:

Table 5. Summary of measurement variability for combined sexes (LVD and LVS units = mm^2, SW units = mm)

1 Variable	2 Population mean (n = 1464)	3 Between-subject variation	4 Within-subject variation	5 Total variation	6 % of variation due to within-subject variation
LVD	13,282	1490	424	1549	7.50
LVS	4329	496	215	540	15.81
SW	55.5	4.54	1.98	4.96	15.97

- The same sex as the subject horse
- Measured within 30 days of chronologic age of the subject horse
- Measured within one year of when the subject horse was measured
- Within 25 lbs of weight of the subject horse

Percentiles and standardized scores for weight, height, and HTWT were calculated as explained previously, except without weight restrictions on the comparison group.

Subject comparisons were limited to within ± 1 year of the measurement date to minimize the possible effects of gradual small changes in calibration, methodology, and external variables acting on the subjects. Examples of external variables that may have changed over time and affected measurements include sales preparation techniques of horses at auctions, steroid use, growth hormones, wear and tear on equipment, etc.

Percentiles. Percentiles were class rankings of each variable relative to those of similar horses (same sex, similar age, and weight), on a scale from 0 to 100.

Percentiles were calculated as:

$$\frac{R - 0.5}{N}$$

Where R = rank and N = group size, including subject horse. *Example*: If there were 100 colts of similar age and weight (n = 100), and one colt had the 93rd largest LVD (rank = 93), he would have an LVD percentile of:

$$\text{LVD Percentile} = \frac{93.0 - 0.5}{100} = 92.5\%$$

Percentiles were easily interpreted because the numbers ranged from 0 to 100—a familiar scale.

Standardized scores. Technically, percentiles fail to maintain initial distances between variables. Because most data in natural, biological phenomena is located near the middle of the Gaussian-shaped distribution, measurements in the 50th and 52nd percentiles are closer in absolute value than those in the 95th and 97th percentiles. Standardized scores described below maintain the natural spacing between variables, producing a scale-free statistic with a mean of 0, and a standard deviation of 1.

$$\text{Standardized score} = \frac{\text{Observation} - \text{Mean}}{\text{Standard Deviation}}$$

Standardized scores could be difficult to interpret because, whereas they generally ranged from –3 to +3, they tended to congregate around zero. It seems easier to understand that a horse is in the 70th percentile compared with his peers than with knowledge that his standardized score is 0.55.

Statistical analyses and tables in this text are based on percentiles. The same analyses and tables in terms of standardized scores produced virtually the same results. Therefore, they are not shown here, but are available on request.

RESULTS

Reproducibility and Sources of Measurement Variability

Variation (or differences) between cardiac measurements is caused by a combination of within- and between-subject variation. *Within-subject variation*, sometimes called *measurement error*, indicates how accurately or reproducibly the technician and equipment measures a given variable (hearts and horses are moving targets). *Between-subject variation* is the range of expected differences among a particular variable in the general population that is not because of error. Between-subject variation accounted for 84% to 92% of variation in cardiac measurements in this study, and within-subject variation accounted for 8% to 16% of variation.

Measurement variability was calculated for LVD, LVS, and SW among 1464 horses measured in 1999. These cardiac measurements were repeated at least 3 times within a period of several minutes. (1571 horses were measured in 1999. Those excluded from this variability study lacked at least 3 measurements for LVD, LVS, or SW because of auction conditions, during which the technician may have lacked time to repeat measurements, could not sustain a resting heart rate [or behavioral cooperation], or reported only the average.)

Table 5 summarizes between-subject variation (s_B) and within-subject variation (s_W). Column 1 lists the variables studied. Column 2 lists the mean value of each variable among all 1464 horses in this part of the study. Column 3 lists between-subject variation, which is the standard deviation associated with the mean reported in Column 2. Column 4 lists within-subject variation. Column 5 lists total variation. Column 6 lists the percentage of total variation due to within-subject variation (or measurement error).

Confidence intervals based on within-subject variation. The within-subject variations listed in Table 5 were used to compute confidence intervals as reported in Table 6, and to answer the following questions:

How accurately did a single cardiac measurement (ie, not an average of measurements repeated over a period of a few minutes) describe the true value? A statistical solu-

Table 6. 95% confidence intervals (CI) associated with within-subject variations reported in Table 5 (LVD and LVS units = mm2, SW units = mm)

1	2	3	4
Variable	95% CI for a single measurement	95% CI for a measurement repeated 3 times over the course of a few minutes (n = 3)	95% CI for a horse measured on 2 separate dates or for measurements of 2 different horses (n = 3)
LVD	831	480	679
LVS	421	243	344
SW	3.9	2.3	3.2

tion is to use the "95% confidence interval for a single measurement," as shown in Column 2 of Table 6. In this example, the value was 831 mm^2 for LVD. This means that there is a 95% probability that the true LVD lies within 831 mm^2 of a single LVD measurement.

How accurately did the mean of 3 repeated cardiac measurements over the course of a few minutes describe the true value? A statistical solution is to use the "95% confidence interval for repeated measurements," as shown in Column 3 of Table 6. For example, this value was 480 mm^2 for LVD. This means that there is a 95% probability that the true LVD is within 480 mm^2 of the mean of 3 repeated LVD measurements.

How much of a difference between cardiac measurements over some period of time would rule out measurement error as the sole source of the difference? A statistical solution is to use the "95% confidence interval for repeated measurements from two separate dates," as shown in column 4 of Table 6. For example, this value was 679 mm^2 for LVD. This means that if the difference between LVD measurements on different dates exceeded 679, there is a 95% probability that measurement error was not the sole source of that difference.

How much of a difference between repeated cardiac measurements of 2 different horses would rule out measurement error as the sole source of the difference? A conservative statistical solution is to use the 95% confidence interval just mentioned, as listed in column 4 of Table 6. For example, this value was 679 mm^2 for LVD. This means that if the difference between horses' LVDs exceeded 679 mm^2, then there is a 95% probability that measurement error was not the sole source of that difference.

Multiple measurements of the same horse within the same month. The average percent change in cardiac measurements for horses measured twice within the same month of age was calculated for horses 14 to 17 months of age. These were the only individual months of age with at least 5 different horses represented.

The change in the cardiac measurements of these horses fell within the range of expected measurement error described in Table 5. Most change was positive, indicating that growth may have occurred in addition to measurement variation. Measurement variation among horses measured twice within the same month was also influenced by other factors, eg, some horses were re-measured because the ultrasound technician was not satisfied with the initial measurement, likely because of the horse's behavior (ie, suspected illness, medications, or elevated heart rate after start of examination). See Appendix A for detailed data pertaining to horses measured twice within the same month.

Multiple measurements of the same horse measured at different months of age. Appendix A provides detailed data pertaining to horses measured at different months of age. As the time between measurements increased, there was a general increase in the size of cardiac measurements. As horses reached 2 years of age, changes began to level off. Thus, for example, the differences between horses measured at both 13 and 18 months of age were greater than those of horses measured at both 20 and 25 months of age. These results were also evident in the growth curves in Appendix B.

STATISTICAL OVERVIEW OF CARDIAC DATA

Descriptive Statistics

Colts, fillies, and combined sexes.

Means and Standard Deviations

Means and standard deviations of cardiac raw data for combined sexes are presented in Table 7 and Table 8 by months of age. Appendix B includes graphs of each variable versus months of age, including growth curves and related equations.

Growth curves. Cardiac parameter means and growth curves were graphed for colts and fillies in Appendix B. The graph in Figure 4 compares LVD for colts versus fillies, and is typical of sex-related differences. Most growth curves were described well (R^2 ≥0.90) by second-degree polynomial equations, as shown on the graphs. The growth curves should be limited to application over the period from 12 through 27 months of age for which they were calculated (ie, not used to estimate average LVD at 32 months of age).

Anomalies appeared in the data patterns of cardiac measurements versus age at 20 and 21 months of age. These horses were primarily measured during October through December, between the timing of select yearling and select 2-year-old auctions (Appendix C). Horses often enter training during those interim months. Training regimens, and

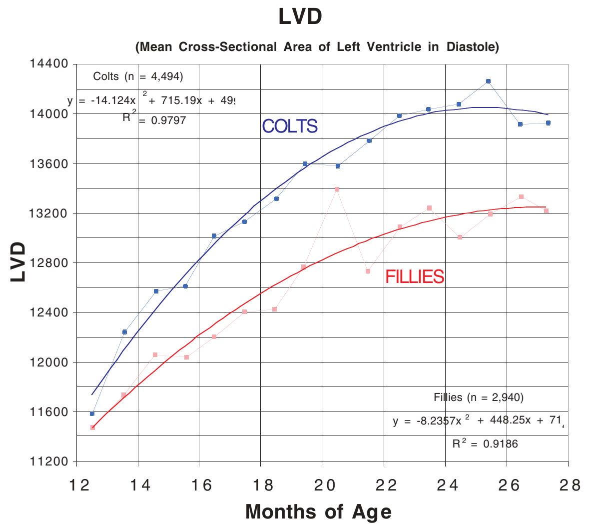


Figure 4

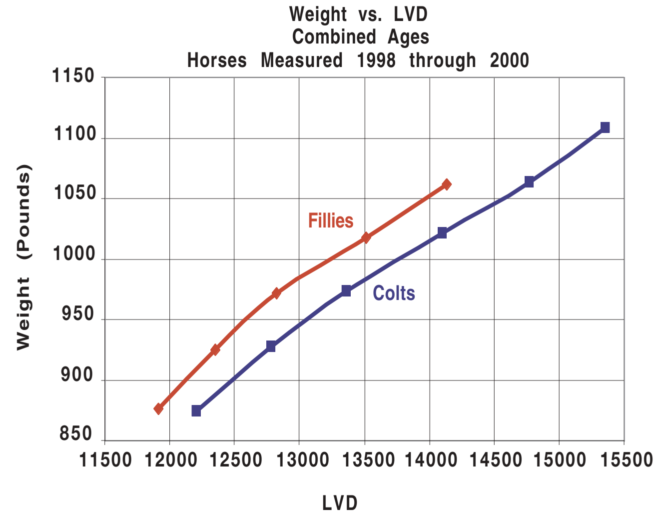


Figure 5

Table 7. Means of cardiac measurements for combined sexes by months of age (total n = 7434)

		Cardiac measurement means						
Age (mos)	No. of horses	Age (mos)	LVD (mm^2)	LVS (mm^2)	SW (mm)	PS (%)	Weight (lbs)	Height (hands)
12	81	12.5	11,534	3823	49.4	66.82	801	14.52
13	155	13.5	12,025	3982	50.7	66.86	875	14.87
14	399	14.6	12,362	4038	50.9	67.32	944	15.12
15	758	15.6	12,395	4024	51.1	67.52	970	15.26
16	1279	16.5	12,689	4133	51.9	67.41	986	15.37
17	1196	17.5	12,843	4182	52.4	67.41	996	15.42
18	856	18.5	12,948	4203	52.4	67.52	1001	15.47
19	551	19.4	13,285	4330	53.5	67.36	1005	15.51
20	248	20.5	13,504	4431	53.9	67.16	1013	15.59
21	337	21.5	13,428	4344	54.0	67.64	1016	15.60
22	440	22.5	13,633	4411	54.7	67.64	1026	15.64
23	485	23.5	13,706	4384	54.7	68.04	1032	15.69
24	333	24.5	13,646	4413	54.6	67.65	1036	15.65
25	184	25.4	13,803	4409	55.2	68.08	1046	15.77
26	95	26.5	13,657	4420	54.6	67.66	1045	15.77
27	37	27.3	13,638	4415	55.1	67.65	1036	15.75

thus each heart's response to training, likely varied greatly during this time.[1] Puberty may play a role among fillies at this age. Most horses were measured during this period at private farms, without any pre-selection on the basis of conformation or pedigree. The ratio of colts to fillies (60%:40%) in this study closely matches those at auctions. This ratio may favor colts because breeding farms keep some of the best-bred, best-conformed fillies for their breeding programs. Therefore, relative to auctions, the fillies seen at private farms may be of higher quality, overall, because they may include the best-bred, best-conformed fillies that never make it to auctions.

Relationship between weight and cardiac measurements. Generally, among the total population measured, the bigger the horse, the bigger its heart, all else being equal. The graphs in Figures 5 through 8 provide a visual overview of the relationship between cardiac measurements and the weight of the horse.

Subsequent high earners and low earners. The graphs in Figures 9 and 10 compare LVD and Weight percentiles for high earners versus low earners, and are typical of performance-related differences (except for PS). Not only were future high earners heavier than low earners (Fig. 9), but, even when normalized by sex, age, and weight, high earners still had higher cardiac measurements (Fig. 10). Sample sizes of high earners in these graphs were small at 19 and 20 months of age.

t tests. Table 9 and Table 10 summarize *t* test comparisons of high versus low earners, and routers versus sprinters. *t* tests assumed equal variances (verified with F statistics).

High earners versus low earners. *t* tests compared high versus low earners of combined sexes and ages, using data normalized for sex, age, and size. Significant differences (P values ≤ .0001) existed between high and low earners

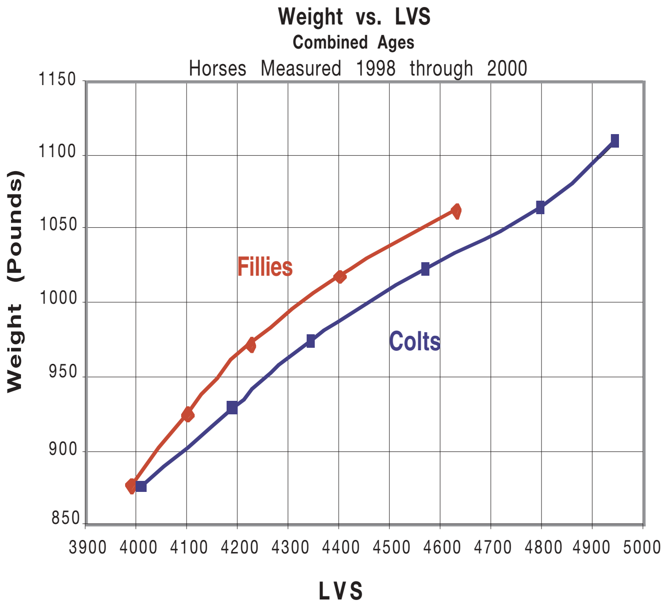


Figure 6

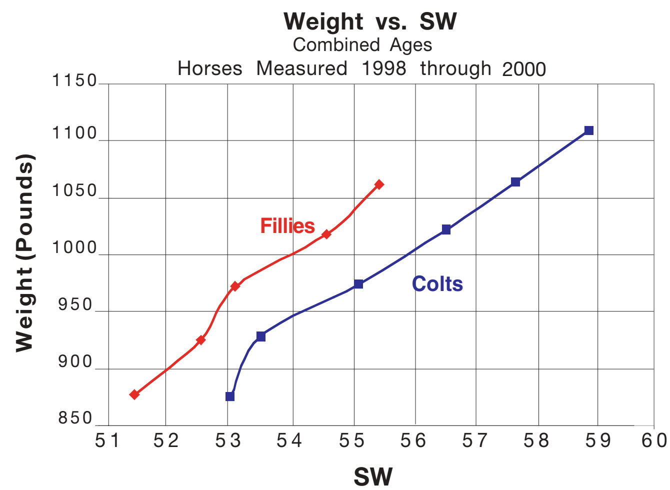


Figure 7

Table 8. Standard deviations of cardiac measurements for combined sexes by months of age corresponding to means shown in Table 7 (total n = 7434)

		Cardiac measurement standard deviations						
Age (mos)	No. of horses	Age (mos)	LVD (mm^2)	LVS (mm^2)	SW (mm)	PS (%)	Weight (lbs)	Height (hands)
12	81	0.2791	1232	424	4.09	1.989	116.2	0.592
13	155	0.2924	1392	505	4.46	2.206	104.6	0.493
14	399	0.2897	1408	519	5.12	2.329	73.9	0.394
15	758	0.2840	1548	553	5.16	2.223	59.6	0.357
16	1279	0.2841	1567	560	5.50	2.297	50.9	0.355
17	1196	0.2902	1541	551	5.29	2.328	49.4	0.369
18	856	0.2783	1595	581	5.45	2.470	50.1	0.368
19	551	0.2800	1504	526	5.24	2.278	49.4	0.367
20	248	0.2898	1347	494	4.34	2.204	45.8	0.358
21	337	0.2908	1459	545	4.84	2.239	46.1	0.367
22	440	0.2760	1404	547	4.59	2.283	41.7	0.354
23	485	0.2889	1366	554	4.56	2.202	44.1	0.359
24	333	0.2965	1493	587	4.63	2.819	48.2	0.366
25	184	0.2886	1519	606	4.53	2.657	37.1	0.362
26	95	0.2887	1410	557	4.74	2.076	36.7	0.328
27	37	0.2498	1590	606	4.74	2.081	40.5	0.375

Table 9. *t* tests—Percentiles (data adjusted for age, sex, and weight)

	Low earners			High earners			
Variables	n	Mean	SD	n	Mean	SD	*P* value
LVD	1061	45.93	28.61	418	53.12	28.32	.0000
LVS	1061	46.45	28.61	418	52.72	28.43	.0001
SW	1061	46.22	27.60	418	53.29	27.17	.0000
PS	1061	50.09	29.56	418	49.89	29.03	.9050
Weight	1091	47.25	29.40	424	60.11	27.38	.0000
Height	1091	53.47	28.58	424	65.81	26.07	.0000
HTWT	1091	45.32	29.02	424	58.53	27.24	.0000

High earners (earnings per start ≥$10,000) versus low earners (earnings per start ≤$2000).

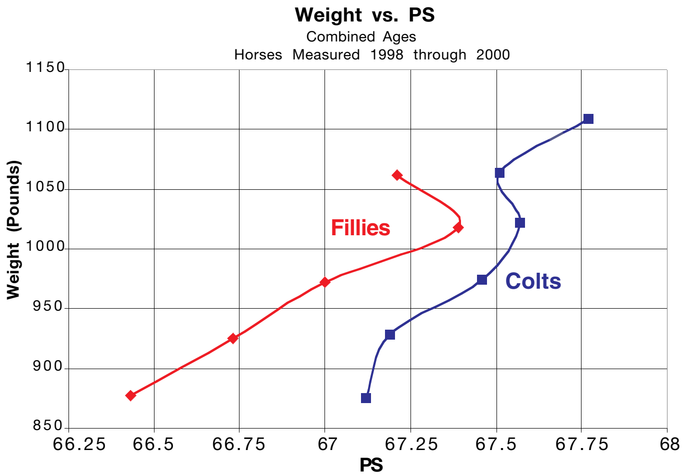


Figure 8

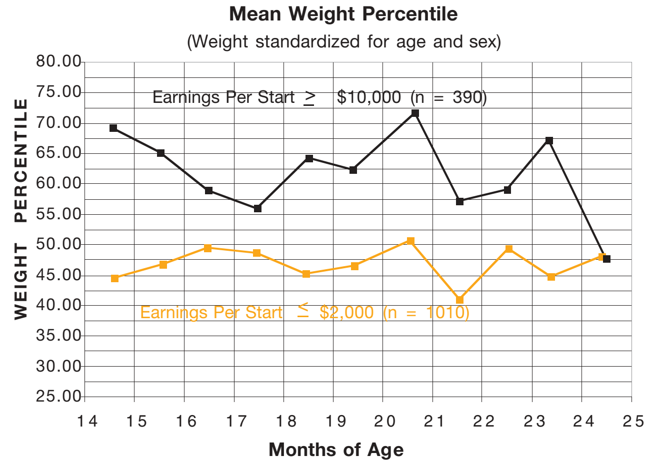


Figure 9

for all of the cardiac parameters listed in Table 9, except for PS. Stepwise analysis, as discussed in this paper, identified SW, LVS, and HTWT as the most significant discriminant variables when differentiating between high and low earners. High earners were defined as horses that raced at least 3 times, with earnings per start of at least $10,000.

High earner routers versus high earner sprinters. *t* tests compared high earner routers versus high earner sprinters of combined sexes and ages, using data standardized for horses of the same age, sex, and size. Significant differences (P values ≤ .05) existed between high earner routers and sprinters for the cardiac variables of LVD, LVS, weight, height, and HTWT, as shown in Table 10. Stepwise analysis, as discussed in this paper, identified LVD, LVS, HTWT, and PS as the most significant discriminant variables when differentiating between high earner routers and sprinters. High earner routers raced at least 3 times at distances of at least 8.5 furlongs, with earnings per start at those route distances of at least $10,000. High earner sprinters raced at least 3 times at distances less than 7.0 furlongs, with earnings per start of at least $10,000 at those sprint distances, and earned < $2000 per start at distances ≥ 8.5 furlongs.

STANDARDIZING CARDIAC MEASUREMENTS ELIMINATED THE EFFECTS OF AGE, SEX, WEIGHT, AND HEIGHT

The high correlation between age and size versus most cardiac measurements (see Table 11) was not present among percentiles (see Table 12). Therefore, when looking at horses of different ages, sexes, and sizes, it is possible to compare their cardiac measurements by standardizing their data (ie, using percentiles, as described previously).

To measure the correlation between age and size versus most cardiac measurements, Pearson correlation coefficients (r) were computed between cardiac measurements

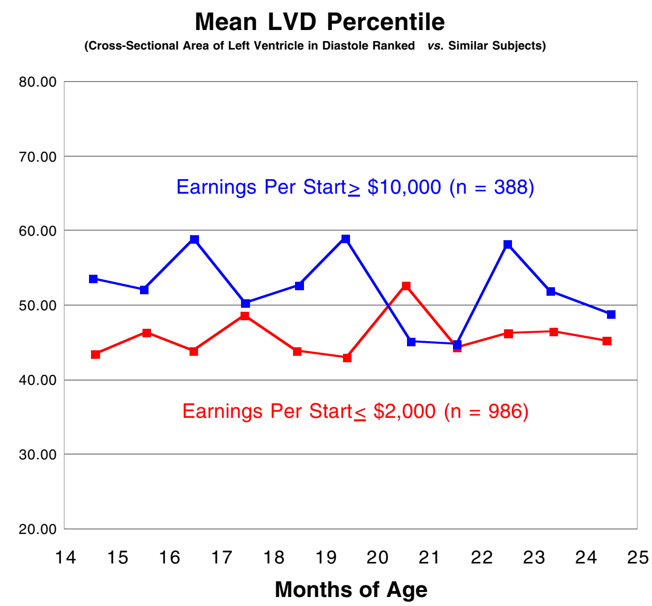


Figure 10

for raw data and percentiles (standardized for sex, age, and weight). Tests for significance of correlation coefficients produced P values ≤ .0001. Correlation coefficients were squared and multiplied × 100 to compute coefficients of determination (R^2), as shown in Table 11 and Table 12.

Essentially, because of the standardization of the data, where there was a high degree of correlation throughout the "Months" column and bottom 3 rows (weight, height, and HTWT) of Table 11, there was little correlation shown in the same column and rows of Table 12 *(bolded areas)*.

COVARIANCE ANALYSIS OF MEANS TO ASSESS THE EFFECTS OF AGE AND SEX ON CARDIAC MEASUREMENTS

Sex-Related Differences. Analysis of covariance showed that age- and weight-adjusted means for cardiac

Table 10. *t* tests—Percentiles (data adjusted for age, sex, and weight)

	Sprinters			Routers			
Variables	n	Mean	SD	n	Mean	SD	P value
LVD	180	48.68	28.69	134	56.95	27.83	.0110
LVS	180	47.68	28.31	134	58.17	28.11	.0012
SW	180	52.08	27.51	134	56.60	27.27	.1495
PS	180	53.04	29.31	134	47.79	27.55	.1085
Weight	180	55.10	26.99	134	64.12	24.11	.0024
Height	180	59.99	26.83	134	69.51	24.60	.0014
HTWT	180	52.85	27.15	134	63.06	24.50	.0007

High earner routers (raced ≥8.5 furlongs) versus high earner sprinters (raced <7 furlongs).

Table 11. Coefficients of determination (R^2) (n = 7434) among raw data (unadjusted for sex, age, and weight)

	Coefficients of determination for cardiac measurements—Raw data						
Variables	Mos	LVD	LVS	SW	PS	Weight	Height
LVD	9						
LVS	6	74					
SW	6	70	51				
PS	1	1	19	1			
Weight	22	21	14	10	1		
Height	19	23	16	13	1	73	
HTWT	22	24	15	12	1	97	86

Table 12. Coefficients of determination (R^2) (n ranged between 7288—7434) among percentiles (standardized for sex, age, and weight)

	Coefficients of determination for cardiac measurements—Percentiles						
Variables	Mos	LVD	LVS	SW	PS	Weight	Height
LVD	0						
LVS	0	62					
SW	0	53	30				
PS	0	1	20	2			
Weight	0	0	0	0	0		
Height	0	1	0	0	0	67	
HTWT	0	0	0	0	0	96	79

Table 13. Means adjusted for age and weight (n = 7434)

	LS means		
Variables	Colts	Fillies	*P* values
LVD	13,315	12,832	.0001
LVS	4318	4179	.0001
SW	53.86	52.12	.0001
PS	67.55	67.41	.0100

measurements were significantly different (P values ≤ .01) between colts and fillies 12 through 27 months old, as shown in Table 13.

Age-Related Differences. Analysis of covariance showed that sex- and weight-adjusted means for cardiac measurements were sometimes significantly different between horses of different months of age. The significance of differences varied depending on the variables studied and the number of months apart. In most cases, significant differences ($P \le .05$) were rare or weak when comparing yearlings to yearlings, or 2-year-olds to 2-year-olds, whereas differences were significant when comparing yearlings to 2-year-olds.

USING CARDIAC PARAMETERS TO PREDICT RACING PERFORMANCE

Stepwise Discriminant Analysis of the Relationship of Cardiac Measurements to Performance

Defining performance—3 career starts minimum. It is impossible to know the level of ability of most horses measured that subsequently never raced, or raced only a few times. For this reason, when we formed groups of high versus low earners or routers versus sprinters, the horses we

Table 14. Discriminant model results—High earners versus low earners
Nonblind tests—combined sexes: Names starting with letters A-Z

	Pre-model probability		Post-model probability		
Category	Ratio	%	Ratio	%	*P* value
High earners	418/1479	28.26	256/686	37.32	.0000
Low earners	1061/1479	71.74	631/793	79.57	.0000

Table 15. Discriminant model results—High earners versus low earners
Nonblind tests—Combined sexes: Names starting with letters A-M

	Pre-model probability		Post-model probability		
Category	Ratio	%	Ratio	%	*P* value
High earners	245/883	27.75	154/409	37.65	.0000
Low earners	638/883	72.25	383/474	80.80	.0000

used had to have raced at least 3 times. Raising the minimum number of starts (as many as 6) did not improve or weaken discriminant analyses.

Stepwise analysis was conducted for colts, fillies, and combined sexes, using percentiles for the variables: LVD, LVS, SW, PS, and HTWT.

High earners versus low earners (see also Table 9). Stepwise analysis was used to identify statistically significant variables that could differentiate between groups of horses categorized as high and low earners, defined as: *high earners* raced at least 3 times, with earnings per start of at least $10,000; *low earners* raced at least 3 times, with earnings per start of $2,000 or less.

Throughout the course of EQB's research into the subsequent performance of more than 30,000 Thoroughbreds studied over the past 25 years, EQB has looked closely at the definitions of high and low earners. The levels of earnings per start chosen as "high" and "low" for this study, although not documented here in statistically rigorous fashion, have historically been closely correlated to generally accepted handicapping assessments of individual race performances, eg, Ragozin "sheets" numbers, Beyer's Speed Figures and Experimental Free Handicap assessments. Among the horses in this study that raced at least 3 times in North America, 34% earned $2000 or less per start (categorized as "low" earners) and 13% earned at least $10,000 per start (categorized as "high" earners). Table 2 provides average worldwide performance statistics for the Thoroughbred breed.

For high versus low earners, stepwise analysis identified the following significant variables (listed in order of statistical significance): *combined sexes*—HTWT, SW, LVS; *colts*—HTWT, SW; and *fillies*—HTWT, SW.

High earner sprinters versus high earner routers (see also Table 10). Stepwise analysis was used to identify statistically significant variables that could differentiate between groups of horses categorized as high earner sprinters and high earner routers, defined as: *high earner sprinters* raced at least 3 times at distances <7.0 furlongs, earned at least $10,000 per start at distances <7.0 furlongs, and earned less than $2000 per start at distances ≥8.5 furlongs; and *high earner routers* raced at least 3 times at distances ≥8.5 furlongs, and earned at least $10,000 per start at distances ≥8.5 furlongs.

For high earner sprinters versus high earner routers, stepwise analysis identified the following significant variables: *combined sexes*—HTWT, LVS; *colts*—LVD, HTWT; and *fillies*—PS, HTWT.

MULTIVARIATE DISCRIMINANT ANALYSIS OF THE RELATIONSHIP OF CARDIAC MEASUREMENTS TO PERFORMANCE

Discriminant analysis was used to classify high earners versus low earners, and high earner routers versus high earner sprinters, as defined in the stepwise analysis section.

Classification threshold. Discriminant results were based on a classification threshold of 50%. A classification threshold is the minimum acceptable probability (as defined by the model user) required to classify a horse into a particular group. Thus, no horse was classified into a group unless the models assigned it at least a 50% probability of belonging to that group. In general, the higher the threshold, the better the models performed (ie, a horse with a 70% high earner probability was more likely to be a high earner than a horse with a lower probability, see Appendix G). As the threshold increases for a particular group, the models generally misclassify more members of that group. At public auctions, a high "high earner" threshold would minimize the chances of buying poor performers (Type II errors), while increasing the chances of rejecting good performers (Type I errors).

Z statistics were computed to determine the reliability of discriminant results using the formula below (shown for high earners):

$$Z_H = \frac{P_{H_{post}} - P_{H_{pre}}}{\sqrt{\dfrac{P_{H_{pre}} \bullet (1 - P_{H_{pre}})}{N_{CH_{post}}}}}$$

Where: N_{Hpre} = Number of high earners in model
N_{Tpre} = Total number of horses in model
$N_{HCCpost}$ = Number of high earners correctly classified by model
N_{CHpost} = Number of horses classified as high earners by model
P_{Hpre} = Pre-model probability (N_{Hpre}/N_{Tpre})

Table 16. Discriminant model results—High earners versus low earners
Blind test—Combined sexes: Names starting with letters N-Z

	Pre-model probability		Post-model probability		
Category	Ratio	%	Ratio	%	*P* value
High earners	173/596	29.03	105/278	37.77	.0013
Low earners	423/596	70.97	250/318	78.62	.0027

Table 17. Discriminant model results—High earners versus low earners
Nonblind tests—Colts: Names starting with letters A-Z

	Pre-model probability		Post-model probability		
Category	Ratio	%	Ratio	%	*P* value
High earners	235/880	26.70	143/409	34.96	.0002
Low earners	645/880	73.30	379/471	80.47	.0004

Table 18. Discriminant model results—High earners versus low earners
Nonblind tests—Colts: Names starting with letters A-M

	Pre-model probability		Post-model probability		
Category	Ratio	%	Ratio	%	*P* value
High earners	140/529	26.47	82/246	33.33	.0147
Low earners	389/529	73.53	225/283	79.51	.0226

Table 19. Discriminant model results—High earners versus low earners
Blind test—Colts: Names starting with letters N-Z

	Pre-model probability		Post-model probability		
Category	Ratio	%	Ratio	%	*P* value
High earners	95/351	27.07	63/164	38.41	.0011
Low earners	256/351	72.93	155/187	82.89	.0022

$$P_{Hpost} = \text{Post-model probability } (N_{HCCpost}/N_{CHpost})$$

P values associated with the *Z* statistics were reported in place of *Z* statistics (Table 14 through Table 25).

The model parameters were:

- Horses had to be born by 1997 (so would have racing data through 3-year-old year)
- Horses had to have at least 3 starts (ie, sound enough to race multiple times)

Two types of discriminant analyses, called blind and nonblind tests, were conducted for each model.

- Nonblind test: A nonblind test is one in which the horses classified by a model were used to create the model. Thus, the models "saw" those horses before. A nonblind test is the best-case scenario of how well a model performs.
- Blind test: A blind test is one in which the horses classified by a model were not used to create the model. Thus, the models did not "see" those horses before. A blind test is tougher and more realistic of how well a model performs than a nonblind test. The blind test results, as listed below, are the most important results.

Note that in this study the nonblind and blind tests gave essentially the same predictive results.

To further ensure that this study's results were reproducible, 3 tables were used to summarize each discriminant analysis in terms of blind and nonblind tests. The first table presents nonblind test results based on all horses available for the study. The second table presents nonblind test results based on horses with names beginning with the letters A through M. The third, and most important, table presents blind test results, for which the A through M model was used to classify horses with names beginning with the letters N through Z, which the models had not previously seen.

Each table presents summary statistics as the following describes.

Premodel probability. Discriminating between 2 groups (A and B), the premodel probability is the ratio of all Group A or Group B horses to the total number of horses in the model. This is the probability, using a random selection technique without statistically created models, of correctly classifying a Group A or Group B horse. This probability is shown as a *ratio* and a *percent*. For example, if there are 7 Group A horses and 93 Group B horses, there is a 7% probability of randomly selecting a Group A horse. For Group A horses, this would be shown as a ratio of 7 of 100 and as a percent of 7.00.

Postmodel probability. Discriminating between 2 groups (A and B) the postmodel probability is the ratio of Group A or Group B horses correctly classified by the models to the total number of horses classified by the statistically created models as Group A or Group B horses. This is the probability with discriminant models of correctly classifying Group A or Group B horses. Using the example above, a discriminant model classifying the same 100 horses might classify 25 horses into Group A, of which 5 horses actually belonged to Group A. In this case, the *ratio* for Group A horses would be 5 of 25, or 20%. Thus, in this example, the discriminant models improved the odds of correctly identifying Group A horses from 7% without models to 20% with models. Likewise, they improved the odds of correctly classifying Group B horses from 93% without models to 73 of 75, or 97.3% with models.

P value. The *P* value was listed corresponding to the *Z* statistic computed for each model.

Table 20. Discriminant model results—High earners versus low earners
Nonblind tests—Fillies: Names starting with letters A-Z

	Pre-model probability		Post-model probability		
Category	Ratio	%	Ratio	%	*P* value
High earners	183/599	30.55	114/270	42.22	.0000
Low earners	416/599	69.45	260/329	79.03	.0002

Table 21. Discriminant model results—High earners versus low earners
Nonblind tests—Fillies: Names starting with letters A-M

	Pre-model probability		Post-model probability		
Category	Ratio	%	Ratio	%	*P* value
High earners	105/354	29.66	68/154	44.16	.0001
Low earners	249/354	70.34	163/200	81.50	.0005

Table 22. Discriminant model results—High earners versus low earners
Blind test—Fillies: Names starting with letters N-Z

	Pre-model probability		Post-model probability		
Category	Ratio	%	Ratio	%	*P* value
High earners	78/245	31.84	47/119	39.50	.0735
Low earners	167/245	68.16	95/126	75.40	.0819

Table 23. Discriminant model results—High earner routers versus high earner sprinters
Nonblind tests—Combined sexes: Names starting with letters A-Z

	Pre-model probability		Post-model probability		
Category	Ratio	%	Ratio	%	*P* value
Routers	134/314	42.68	82/149	55.03	.0023
Sprinters	180/314	57.32	113/165	68.48	.0037

Table 24. Discriminant model results—High earner routers versus high earner sprinters
Nonblind tests—Combined sexes: Names starting with letters A-M

	Pre-model probability		Post-model probability		
Category	Ratio	%	Ratio	%	*P* value
Routers	68/180	37.78	42/81	51.85	.0091
Sprinters	112/180	62.22	73/99	73.74	.0183

Table 25. Discriminant model results—High earner routers versus high earner sprinters
Blind test—Combined sexes: Names starting with letters N-Z

	Pre-model probability		Post-model probability		
Category	Ratio	%	Ratio	%	*P* value
Routers	66/134	49.25	39/64	60.94	.0588
Sprinters	68/134	50.75	43/70	61.43	.0735

High earners versus low earners. The variables HTWT, SW, and LVS, as identified by stepwise analysis, were used in the high earner versus low earner discriminant models.

The predictive results of the blind and nonblind tests were similar.

Results showed that as long as data was first standardized (using percentiles) for each subject's sex, age, and size, each subject's data could be compared with data from subjects of different sexes, ages, and sizes. This made the combined sexes discriminant models just as powerful as separate colt and filly models.

Further comparisons of earnings groups, including \$10,000+ earnings per start versus less than \$7500 earnings per start produced similar results. Thus, the models, using the same independent variables, successfully differentiated between stakes- and allowance-caliber horses, as well as between stakes- and claiming-caliber horses.

Appendix G shows performance summaries of groups of horses receiving various model probabilities (assigned by the combined-sex, nonblind A-Z model) of being high earners ranging from 25% to 75%. In general, horses earned more and raced less frequently as the probability of being high earners, as assigned by the discriminant model, increased.

COMBINED SEXES

Tables 14 through 16 summarize discriminant results for nonblind and blind tests of high earners and low earners, comprised of colts and fillies combined, that had raced at least 3 times (ie, had 3 "starts"). High earners earned at least \$10,000 per start and low earners earned \$2000 or less per start. The improvement associated with discriminant modeling was statistically significant for both high and low earners for all groups studied ($P \leq .0027$).

Non-blind A-Z. Table 14 shows that among 1479 total horses, nonblind discriminant models improved the odds of correctly classifying high earners from 28.26% without models to 37.32% with models. They improved the odds of correctly classifying low earners from 71.74% without models to 79.57% with models. The improvement associated with discriminant modeling was statistically significant for both high and low earners ($P < .0001$).

Nonblind A-M. Table 15 shows that among horses with names beginning with the letters A through M, nonblind discriminant models improved the odds of correctly classifying high earners from 27.75% without models to 37.65% with models. They improved the odds of correctly classifying low earners from 72.25% without models to 80.80% with models. The improvement associated with discriminant modeling was statistically significant for both high and low earners ($P < .0001$).

Blind N-Z. Table 16 shows that among horses with names beginning with the letters N through Z, blind discriminant models based on the A through M horses improved the odds of correctly classifying high earners from 29.03% without models to 37.77% with models. They improved the odds of correctly classifying low earners from 70.97% without models to 78.62% with models. The improvement associated with discriminant modeling was statistically significant for both high and low earners ($P \leq .0027$).

Colts

Tables 17 through 19 summarize discriminant results for high versus low earners among colts.

Nonblind A-Z. Table 17 shows that among 880 colts, nonblind discriminant models improved the odds of correctly classifying high earners from 26.70% without models to 34.96% with models. They improved the odds of correctly classifying low earners from 73.30% without models to 80.47% with models. The improvement associated with discriminant modeling was statistically significant for both high and low earners ($P \leq .0004$).

Nonblind A-M. Table 18 shows that among colts with names beginning with the letters A through M, nonblind discriminant models improved the odds of correctly classifying high earners from 26.47% without models to 33.33% with models. They improved the odds of correctly classifying low earners from 73.53% without models to 79.51% with models. The improvement associated with discriminant modeling was statistically significant for both high and low earners ($P \leq .0226$).

Blind N-Z. Table 19 shows that among colts with names beginning with the letters N through Z, blind discriminant models based on the A through M horses improved the odds of correctly classifying high earners from 27.07% without models to 38.41% with models. They improved the odds of correctly classifying low earners from 72.93% without models to 82.89% with models. The improvement associated with discriminant modeling was statistically significant for both high and low earners ($P \leq .0022$).

Fillies

Tables 20 through 22 summarize discriminant results for high versus low earners among fillies.

Nonblind A-Z. Table 20 shows that among 599 fillies, nonblind discriminant models improved the odds of correctly classifying high earners from 30.55% without models to 42.22% with models. They improved the odds of correctly classifying low earners from 69.45% without models to 79.03% with models. The improvement associated with discriminant modeling was statistically significant for both high and low earners ($P \leq .0002$).

Nonblind A-M. Table 21 shows that among fillies with names beginning with the letters A through M, nonblind discriminant models improved the odds of correctly classifying high earners from 29.66% without models to 44.16% with models. They improved the odds of correctly classifying low earners from 70.34% without models to 81.50% with models. The improvement associated with discriminant modeling was statistically significant for both high and low earners ($P \leq .0005$).

Blind N-Z. Table 22 shows that among fillies with names beginning with the letters N through Z, blind discriminant models based on the A through M horses improved the odds of correctly classifying high earners from 31.84% without models to 39.50% with models. They improved the odds of correctly classifying low earners from 68.16% without models to 75.40% with models. The improvement associated with discriminant modeling was not statistically significant for high or low earners ($P \leq .0819$).

Table 26. Percentage of horses that earned at least $10,000 per start (based on percentiles for individual variables)

	Percentiles			
	0-25%	25-50%	50-75%	75-100%
HTWT	7.6	12.8	14.5	17.8
LVD	11.6	11.1	13.4	17.5
LVS	11.4	11.8	13.9	16.3
SW	10.8	13.1	13.1	16.3
PS	14.3	11.4	14.0	13.3
Average*	10.4	12.2	13.7	17.0

Average was calculated excluding PS, which was not usually predictive.

HIGH EARNER ROUTERS VERSUS HIGH EARNER SPRINTERS–COMBINED SEXES

The variables HTWT and LVS, as identified by stepwise analysis, were used in the high earner routers versus sprinters discriminant models for combined sexes.

Tables 23 through 25 summarize discriminant results for high earner routers versus sprinters.

Nonblind A-Z. Table 23 shows that among 314 high earner horses, nonblind discriminant models improved the odds of correctly classifying routers from 42.68% without models to 55.03% with models. They improved the odds of correctly classifying sprinters from 57.32% without models to 68.48% with models. The improvement associated with discriminant modeling was statistically significant for both routers and sprinters ($P \leq .0037$).

Non-blind A-M. Table 24 shows that among high earner horses with names beginning with the letters A through M, nonblind discriminant models improved the odds of correctly classifying routers from 37.78% without models to 51.85% with models. They improved the odds of correctly classifying sprinters from 62.22% without models to 73.74% with models. The improvement associated with discriminant modeling was statistically significant for both routers and sprinters ($P \leq .0183$).

Blind N-Z. Table 25 shows that among high earner horses with names beginning with the letters N through Z, blind discriminant models based on the A through M horses improved the odds of correctly classifying routers from 49.25% without models to 60.94% with models. They improved the odds of correctly classifying sprinters from 50.75% without models to 61.43% with models. The improvement associated with discriminant modeling was not statistically significant for routers or sprinters ($P \leq .0735$).

CHI-SQUARE ANALYSIS OF PERFORMANCE VERSUS HEART SIZE AND PHYSICAL SIZE

The statistical methods described to this point, and which have shown the predictive nature of cardiac measurements, are perhaps less intuitive than the following examples taken from Appendix J. Once the key variables of HTWT (used as a measure of physical size), LVD, LVS, PS, and SW were standardized for age, sex, and weight, on a scale from 0 *(small)* to 100 *(large)*, we could create groups of horses on the basis of these variables. For example, we could create groups of horses with above or below average LVD, or horses could be grouped into quartiles (ie, from the bottom 25% to the top 25%) based on specific heart measurements or physical size. We could then answer questions such as: *"Was there as high a percentage of high earners among horses with below average LVD as among horses with above average LVD?"* See Appendix J for a more thorough analysis along these lines.

Table 26 shows the percentage of horses that earned at least $10,000 per racing start among horses grouped by physical size and heart size. Overall, 13.3% of the horses in this study's sample earned at least $10,000 per start.

Table 26 shows that as physical size and heart size measurements increased, except for PS, so did the percentage of high earners. This table shows that 17.8% of horses with HTWT in the 75% to 100% percentile range earned at least $10,000 per start. The percentage of horses that earned at least $10,000 per start was below average (13.3% was average for all horses studied) for groups with cardiac variables below the 50th percentile. Horses with cardiac variables in the 75th and higher percentiles were more likely to earn at least $10,000 per start.

Next, horses were first grouped by physical size, and then by heart measurement size. Table 27 shows that all groups of horses with HTWT percentiles of 75% to 100% (right-hand column) produced higher than average percentages of horses with earnings per start (EPS) ≥$10,000. All groups of horses with HTWT percentiles of 0 to 25% (left-hand column) produced fewer than average percentages of horses with EPS ≥$10,000, regardless of heart measurement size.

Table 27. Percentage of horses that earned at least $10,000 per start (based on percentiles for individual cardiac variables combined with HTWT)

		HTWT			
		0-25%	25-50%	50-75%	75-100%
0-25%	LVD	6.7	10.4	**13.7**	**16.4**
	LVS	6.9	10.5	12.3	**16.2**
	SW	6.7	12.3	11.3	**13.5**
	PS	7.8	13.1	**15.1**	**21.4**
25-50%	LVD	9.4	9.6	11.9	**13.4**
	LVS	8.0	10.9	**14.5**	**14.0**
	SW	7.0	11.4	**14.7**	**18.3**
	PS	5.9	10.4	13.2	**15.4**
50-75%	LVD	4.7	**15.2**	**14.5**	**18.8**
	LVS	7.3	**14.6**	**15.9**	**16.9**
	SW	9.1	12.3	**14.4**	**16.4**
	PS	8.3	**13.4**	**14.6**	**19.2**
75-100%	LVD	11.0	**16.8**	**17.9**	**22.2**
	LVS	8.4	**16.0**	**15.1**	**24.0**
	SW	7.5	**15.8**	**17.6**	**22.7**
	PS	8.2	**14.0**	**15.0**	**15.3**

Above average performance categories are bold.

Bolded areas in Table 27 show groups with higher than average percentages of horses with EPS ≥$10,000. Horses with HTWT percentiles in the 25% to 50% range generally performed as well as average as long as cardiac variables were above average.

The highest percentages of high earners occurred when percentiles for both HTWT and heart size were at least 75%. In cases where HTWT and heart size percentiles were at least 75%, the average percentage of horses with EPS ≥$10,000 was 23.0% (excluding PS)—a 73% improvement over random odds of selecting high earners (13.3% vs 23.0%).

Table 28 shows that high earner routers (earned at least $10,000 per start at distances of ≥8.5 furlongs, with at least 3 starts at those distances) were 4 times as likely to have above-average HTWT and LVD than to have below-average HTWT and LVD.

Table 29 shows that extremely high earners (earned at least $250,000 by the end of their 3-year-old year) were about 3 times more likely to have above average HTWT and LVD than to have below average HTWT and LVD.

See Appendix J for more review of these physical and heart size relationships and associated χ^2 statistics.

Table 28. High earner routers: Originated from these categories in these percentages (earned at least $10,000 per start in at least 3 starts at distances of at least 8.5 furlongs)

	HTWT	
	Below average	Above average
LVD		
Below average	12%	25%
Above average	15%	48%

Table 29. Extremely high earners (at any distance): Originated from these categories in these percentages (earned at least $250,000 by the end of their three-year-old year)

	HTWT	
	Below average	Above average
LVD		
Below average	17%	19%
Above average	15%	50%

DISCUSSION

Any Predictive System Based On Physiology Alone Will Be Far From Perfect

Among horses measured for this study and believed to be in North America, 23% never raced by the end of their arbitrarily defined 3-year-old racing year (ie, using the racing industry's January 1 "birthday" to define years of age). Among those that raced, many had poor performance unrelated to physical ability. For example, mismanagement, illness, injury, and attitude are critical factors that can limit performance. A talented horse can step into a hole or kick its stall and never race again, despite the best care.

PRESELECTION OF THIS STUDY'S SAMPLE (OTHER DATA NOT PRESENTED HERE)

Because of a preselection process of subjects by racing industry experts paid to find the best racing prospects, most horses measured for this study had no obvious reason to fail on the racetrack. (Regardless, some horses soon thereafter failed prepurchase examinations by veterinarians that included the use of radiographs and endoscopic examinations of larynxes.) A comparison of Tables 2 and 3 shows that this study's entire sample had race results similar to the progeny of the top 1% of sires overall (ie, sires of the most successful progeny). *Therefore, this study's sample was of such high racing quality that it should have been much more difficult to separate out the best racehorses. Still, cardiac measurements were able to reproducibly and statistically significantly do so.*

Nota Bene. EQB performed biomechanical gait analysis of hundreds of 2-year-olds in this study at the time of cardiac measurements, filming at 250 frames per second.[8,9] Horses with hearts that were statistically characteristic of high earners didn't necessarily have gaits that were statistically characteristic of high earners, and vice versa. EQB also recorded pedigree and other types of information not presented here for all the horses used as subjects in this study.

EFFECT OF SIZE (HEIGHT AND WEIGHT) ON CARDIAC MEASUREMENTS

Graphs and *t* tests showed that horses that earned at least $10,000 per start, compared with horses that earned no more than $2000 per start, were taller and heavier when measured before ever racing. Therefore, physical size by itself was a predictor of subsequent performance. However, even when comparing horses of the same sex, size, and age when measured before racing, the cardiac measurements of subsequent high earners were significantly larger than those of their similar peers (see graphs in Appendices B and J).

Sources of Measurement Error

LVD and LVS measurement variation. LVD and LVS were calculated by tracing the perimeter of a two-dimensional image on the ultrasound machine, which calculated the area inside the tracing based on a pixel count (512×512 for total screen). Within the areas measured, some variation likely resulted from limitations of pixel and cursor size.

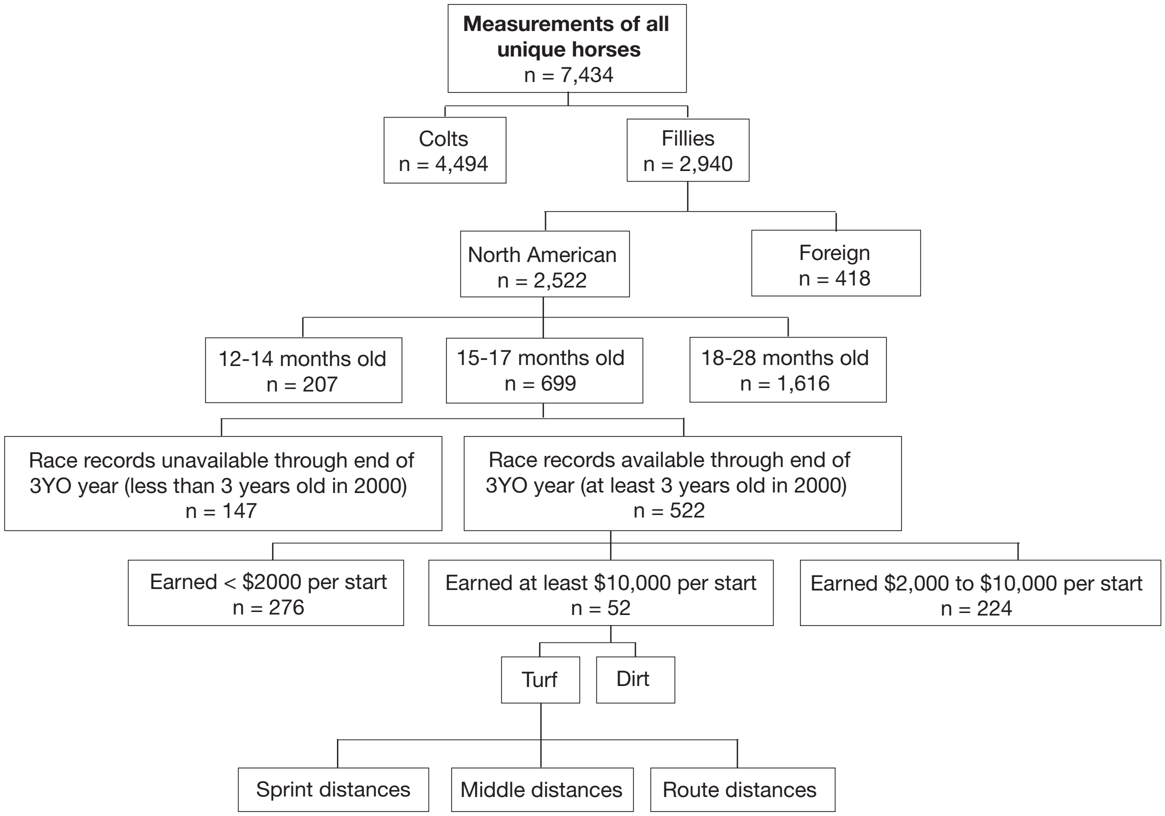
LVD is measured when the left ventricle expands to its largest size in diastole. LVS is measured when the left ventricle contracts to its smallest size in systole. The period of time during which the left ventricle is in diastole lasts longer than when it is in systole. Therefore, when the ultrasound technician manually freezes the diastolic and systolic images before measurement, she is more likely to capture LVD at the right moment than LVS. This may partly explain why there is more variation in LVS than in LVD.

SW variation. SW measurements were taken from reference points that were more difficult to reproduce than LVD and LVS. This difficulty likely accounted for much of the measurement error.

DISCRIMINANT ANALYSIS—PREDICTING RACING PERFORMANCE

Discriminant results showed that a horse's weight and height were important predictive indices of subsequent performance, in terms of earnings and successful distances raced. In addition, SW, as defined in Fig 3, was the most important predictive variable when differentiating between high and low earners. In addition to physical size, the left ventricle in diastole and systole (LVD and LVS) were the most important predictive variables when differentiating between successful sprinters and routers.

Several of the variables studied were highly correlated (ie, similar). Discriminant models typically had very similar results when one or two variables were replaced with



Diagram

other variables with which they were highly correlated (eg, LVS and LVD, or WT and HTWT).

In most cases, combined-sex discriminant models correctly identified the same horses that were correctly identified by the same-sex models.

Blind tests showed that cardiac parameters predicted subsequent racing performance with far greater accuracy than possible selecting horses from these groups at random. Models successfully differentiated not only between stakes- and claiming-caliber horses, but also between stakes- and allowance-caliber horses.

On average, blind test discriminant models improved random odds of identifying high earners (or routers) by 35% (ie, going from a 30% probability of correctly identifying high earners without models to a 40% probability with models).

Stepwise and discriminant analyses beyond those presented here sometimes produced exceptional results for 1 group in the comparison, but unexceptional results for the other group. For example, a high versus low earners model may accurately predict high earners, whereas just meeting random expectations among low earners. Multiple models differentiated by level of earnings may be needed in such instances. Model limitations have to be assessed relative to potential applications. *Z* tests were helpful in determining the statistical strength of discriminant results for each individual group represented in the models.

THE NEED FOR A LARGE DATABASE

Better discriminant results can sometimes be achieved by more narrowly defining the groups to be differentiated. For example, a model can be designed to predict the earnings potential of large fillies 15 to 17 months old (ie, good-looking, well developed fillies at the September Keeneland auctions). A large database was essential when narrowly defining models in these ways.

Selecting subsets of the data by criteria including: sex, age, racing continent (North American or foreign), surface (dirt or turf), distance preference, earnings (and other performance criteria, including graded stakes, stakes, or other winners), number of races (ie, starts), etc, can reduce a database of 7500 horses to just a few.

The above diagram illustrates how various selection criteria affects group sizes. Note that the diagram traces only one of countless possible branches on the data tree.

GROWTH SPURT?

Age- and weight-adjusted means showed that a period of rapid growth occurred between 18 to 21 months of age. This phenomenon was seen repeatedly in groupings by months-of-age, and in graphs of cardiac variables by months of age. The growth that occurred during this period was not always along a smooth curve relative to age, such that you could not be sure whether the rapid growth would occur at 18 or 21 months of age, or any point in-between, for a particular horse. However, if you knew the measurement at 17 months of age (ie, before the growth spurt), you could reasonably estimate the measurement size past 21 months of age (ie, after the growth spurt). One or more of the following effects may have caused the growth spurt:

- Training effect: horses measured as yearlings were usually measured before 20 months of age (ie, before October). Horses measured as 2-year-olds were usually measured after 20 months of age (see Appendix C). Horses at 2-year-old sales had been in training with a rider, whereas yearlings did not have that level of training.
- Age-related growth, independent of training: some growth during this period occurs along the expected growth curve as a standard part of maturity with age.
- Onset of puberty: the onset of puberty may have an especially large effect on fillies, as seen in many of the graphs at 20 and 21 months of age.
- Market forces bias among fillies: some of the largest hearts relative to sex and age were measured among extremely well bred and well conformed fillies at prominent private farms (ie, large, successful, high-end breeding operations) during October through December, when there were no major high-priced auctions. Many of these types of fillies are kept and/or raced by their breeders so they are retained within the operation eventually for breeding purposes, and don't go through auctions, or they may go through different types of auctions than those covered in this study (ie, breeding stock auctions). Thus, within this study, average cardiac measurements for fillies measured during October through December may exceed those of fillies measured at auctions. Consider that, among the typical auctions included in this study, 40% of auction horses were fillies and 60% were colts.

RECOMMENDED FURTHER RESEARCH

Aside from refinements to research presented here, such as more narrowly defining groups of horses for this research, we would like to see further research to assess the following areas of study.

Auctions. Initial research (see Appendix H) has shown statistically significant differences among the means (by age, sex, and size) of cardiac measurements taken from different auctions.

Pedigree. We recorded dam, sire, and damsire for each horse. It would be worthwhile to look for relationships between pedigree and cardiac parameters. This information could be combined with other disciplines to identify genetic markers for cardiac parameters. Appendix I provides some initial tables summarizing cardiac measurements for dams, sires, and damsires with large numbers of offspring in this study.

Four-year-old data. EQB historically limited race record assessment to the 2- and 3-year-old years because many high earners (by the end of their 3-year-old year) subsequently got injured and worked their way down the claiming ranks. Thus, a stakes-caliber horse as a 3-year-old might later become a cheap claiming horse. However, many horses racing today (including those that are not a part of this study) have their best performances after 3 years of age. We have enough data now to include racing records beyond the 3-year-old year and to manage anomalies, including high earners that subsequently race at the lowest levels.

Visual ratings of ultrasound cardiac images. Late in this study, the ultrasound technician began recording additional information for each horse (25% of horses have data for the additional variables). These were subjective ratings, ranging from 1 *(poor)* to 5 *(excellent)*, that the ultrasound technician used to describe the images on the ultrasound machine—visual impressions of ecogenicity and of the general shape of the 2DE images. These ratings, discussed in Appendix F (which included discriminant analysis), may help to further predict performance. Their use was minimal in this study because too few horses for which we had these ratings had racing data through their three-year-old year.

Combining cardiac assessments with other scientific assessments. Some horses measured for this study failed veterinary examinations during the same auction. However, veterinary assessments of health were not used in this study. Many 2-year-olds that were measured were filmed in slow motion during the same auction at 250 pictures per second and digitized for biomechanical gait analysis. Some horses with statistically exceptional cardiac measurements (ie, as described in the discriminant analysis results section of this article) did not have statistically exceptional gaits, and vice versa.[8,9] The added value of using veterinary, gait, and other scientific assessments along with the cardiac measurements described in this study should be further investigated.

We thank J. Richard Trout, PhD (biostatistics).

GLOSSARY OF TERMS

Chronological age: Age based on actual calendar days, months, or years of age, as opposed to being based on the racing industry's arbitrary January 1st "birthday."
EPS: Earnings per start through three-year-old year
Height: Measured in hands
LVD: Left ventricle area at diastole (see Fig 2)
LVS: Left ventricle area at systole (see Fig 1)
Money: Total earnings through three-year-old year
Months: Months of age when measured, calculated as (days of age/30.4167)
PS: Percent of blood pumped per stroke ([LVD – LVS]/LVD)
Sex: "C" (used for colts, ridglings, and geldings) or "F"
Starts: Total starts through three-year-old year
SW: Interventricular septal wall structural thickness (see Fig 3)
Weight: Measured in pounds

REFERENCES

1. Young LE. Cardiac responses to training in 2-year-old thoroughbreds: an echocardiographic study. Equine Vet J 1999;Suppl 30:195-8.
2. Young LE, Scott GR. Measurement of cardiac function by transthoracic echocardiography: day-to-day variability and repeatability in normal thoroughbred horses. Equine Vet J 1998;30:117-22.
3. Henry WL, DeMaria A, Gramiak R, King DL, Kisslo JA, Popp RL, et al. Report of the American Society of Echocardiography Committee on nomenclature and standards in two-dimensional echocardiography. Circulation 1980;62:212-7.
4. Feigenbaum H. Echocardiography. 4th ed. Philadelphia (PA): Lea & Febiger; 1986. p.1-33.
5. Wyatt HL, Heng MK, Meerbaum S, Hestenes JD, Cobo JM, Davidson RM, et al. Cross-sectional echocardiography. I. Analysis of mathematical models for quantifying mass of the left ventricle in dogs. Circulation 1979;60:1104-13
6. Weyman AE. Cross-sectional echocardiography. Philadelphia (PA): Lea & Febiger; 1982. p. 497-504.
7. O'Grady MR, Bonagura JD, Powers JD, Herring DS. Quantitative cross-sectional echocardiography in the normal dog. Vet Radiol 1986;27:34-49.
8. Seder JA, Vickery CE. Temporal and kinematic gait parameters of thoroughbred racehorses at or near racing speeds. Presented at 1993 AESM Proceedings, Fallbrook, CA, 14 March, 1993.
9. Seder JA, Vickery CE. Double and triple fully airborne phases in the gaits of racing speed thoroughbreds. From proceedings of the Twelfth Meeting of the Association of Equine Sports Medicine, Fallbrook, CA, 13-16 March, 1992.

FOR ADDITIONAL INFORMATION ALSO SEE:

1. Voros K, Holmes JR, Gibbs C. Anatomical validation of two-dimensional echocardiography in the horse. Equine Vet J 1990;22:392-7.
2. Voros K, Holmes JR, Gibbs C. Left ventricular volume determination in the horse by two-dimensional echocardiography: an in vitro study. Equine Vet J 1990;:398-402.
3. Pascoe JR, Hiraga A, Hobo S, Birks EK, Yarbrough TB, Takahashi T, et al. Cardiac output measurements using sonomicrometer crystals on the left ventricle at rest and exercise. Equine Vet J 1990;Suppl 30:148-52.
4. Hiraga A, Hobo S, Birks EK, Takahashi T, Hada T, Smith BL, et al. Changes in left ventricular dynamics during graded exercise. Equine Vet J 1990;Suppl 30:122-5.
5. Sampson SN, Tucker RL, Bayly WM. Relationship between VO_{2max}, heart, score and echocardiographic measurements obtained at rest and immediately following maximal exercise in thoroughbred horses. Equine Vet J 1990;Suppl 30:190-4.
6. O'Callaghan MW. Comparison of echocardiographic and autopsy measurements of cardiac dimensions in the horse. Equine Vet J 1985;17:361-8.
7. Kriz NG, Hodgson DR, Rose RJ. Changes in cardiac dimensions and indices of cardiac function during deconditioning in horses. AM J Vet Res 2000;61:1553-60.
8. Slater JD, Herrtage ME. Echocardiographic measurements of cardiac dimensions in normal ponies and horses. Equine Vet J 1995;Suppl 19:28-32.
9. Marr CM, Bright JM, Marlin DJ, Harris PA, Roberts CA. Pre- and post-exercise echocardiography in horses performing treadmill exercise in cool and hot/humid conditions. Equine Vet J 1999;Suppl 30:131-6.
10. Averages for the breed. Thoroughbred Times 8 June 2002:31.
11. Sparkman JP. A report card for competitors. Thoroughbred Times 6 July 2002:20.

Appendices will appear online at: *www.j-evs.com*